Causal Modeling of CALGB 80405 (Alliance) Identifies Network drivers of Metastatic Colorectal Cancer

Rahul K. Das,[1] Leon Furchtgott,[1] Daniel Cunha,[1] Fang-Shu Ou,[2] Federico Innocenti,[3] Heinz-Josef Lenz,[4] Jeffrey Meyerhardt,[5] Kelly Rich,[1] Jeanne Latourelle,[1] Donna Niedzwiecki,[6] Andrew Nixon,[7] Eileen M. O'Reilly,[8] Diane Wuest,[1] Boris Hayete,[1] Iya Khalil,[1] Alan Venook,[9]

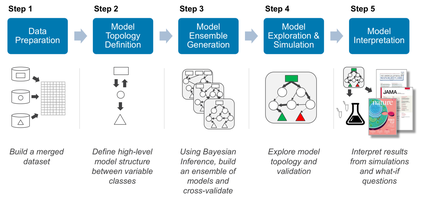
[1]GNS Healthcare, Cambridge, MA; [2]Alliance Statistics and Data Center, Mayo Clinic, Rochester, MN; [3]University of North Carolina, Chapel Hill, NC; [4]University of Southern California, Los Angeles, CA; [5]Dana-Farber Cancer Institute/Partners Cancer Care, Boston, MA; [6]Alliance Statistics and Data Center, Duke University, Durham NC; Duke Cancer Institute, [7]Duke University Medical Center, Durham, NC; [8]Memorial Sloan Kettering Cancer Center, New York, NY; [9]University of California San Francisco, San Francisco, CA

BACKGROUND

- CALGB 80405 is a recently-completed phase III clinical trial of FOLFOX and FOLFIRI with randomly assigned cetuximab (cet) or bevacizumab (bev) in metastatic CRC (mCRC) patients.
- Hypothesis-free machine learning approaches to this study dataset can provide valuable insights into mCRC prognosis and management of mCRC progression.
- Causal modeling identifies the set of conditional dependencies between variables leading to outcomes.
- We built multivariate causal models of mCRC and examined the network drivers of mCRC survival.

METHODS

Fig. 1: Schematic of REFS™ Reverse Engineering & Forward Simulation Workflow



- Using our Bayesian causal machine learning platform REFS™, an ensemble of 128 network models were built for overall survival (OS) of mCRC.
- The ensemble enables estimation of model uncertainty and identification of key drivers by model consensus.
- Simulations were performed on the ensemble to identify causal drivers of OS after accounting for confounders. Causal effect was quantified by median hazard ratio (HR). For continuous variables, 3rd & 1st quartile values were used to compute HR.
- Analysis of NanoString data:
  - Consensus molecular subtypes (CMS) were computed using published code (Guinney et al., *Nat. Med.* 2015) on GitHub.
  - Molecular clusters were computed using consensus clustering.
- Patients with both KRAS wild-type and mutant tumors were included and those who received both cet and bev treatments were excluded. Molecular data from primary tumors were included.
- Two independent cohorts (N=117 for mutations, N=206 for nanostring data) were withheld and used for causal drivers validation.

CAUSAL MODELS

- **Model1:** Clinical variables only (N=1463, 68 variables)
- **Model2:** Clinical+molecular variables without raw nanostring data (N=430, 84 vars)
- **Model3:** Clinical+all molecular variables (N=430, 900 vars)

RESULTS

Model 1: Clinical Causal Drivers of OS

- 1º side, ECOG performance score, concentrations of aspartate aminotransferase (AST), hemoglobin (HGB), absolute neutrophil counts (ANC), lactate dehydrogenase (LDH) and metastases at intra-abdominal, lung, and liver were the strongest causal drivers of OS.

Fig. 2: Reverse Engineering: Consensus Subnetwork to OS

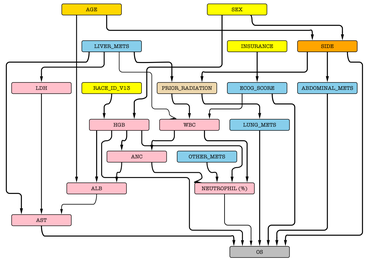
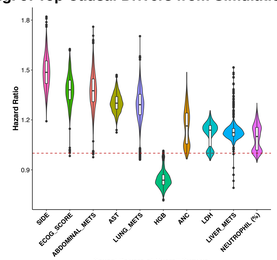


Fig. 3: Top Causal Drivers from Simulations



Model 2: Molecular Causal Drivers of OS

- Clustering of NanoString data revealed three molecular clusters with upregulation of different signatures: (1) WNT-signaling, (2) Angiogenesis & ECM remodeling, (3) Immune infiltration.
- BRAF mutation, RAS mutation, CMS4, and angiogenesis signature were the top molecular drivers of OS.
- Causal effects of 1º side on OS was found to be driven by a molecular pathway.

Fig.4: Molecular Clusters

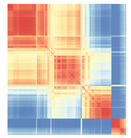


Fig. 5: Over-represented GO Biological Processes in Angiogenesis Cluster

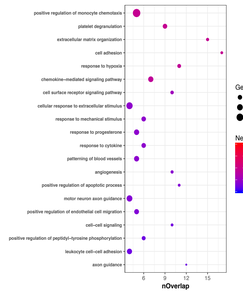


Fig. 6: Consensus Subnetwork to OS

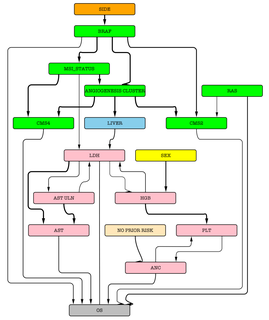
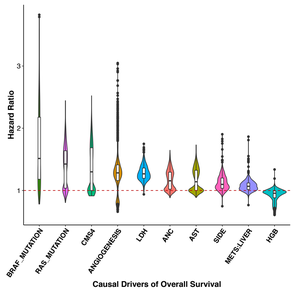


Fig. 7: Top Causal Drivers from Simulations



Model 3: Causal Driver Genes of OS

- ALOX5 and CDX2 were among the top causal driver genes of OS.
- The causal genes in the molecular pathways leading to OS are involved in ECM remodeling and angiogenesis, thereby corroborating the findings from Model 2.

Fig. 8: GO Biological Processes where Causal Driver Genes are Over-represented

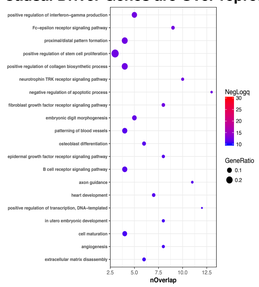


Fig. 9: Consensus Subnetwork to OS

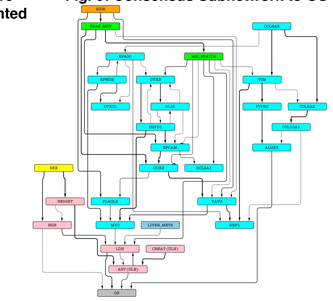
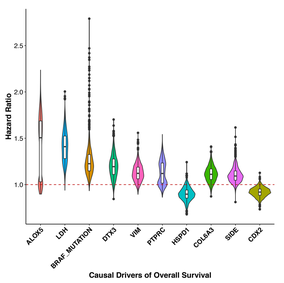


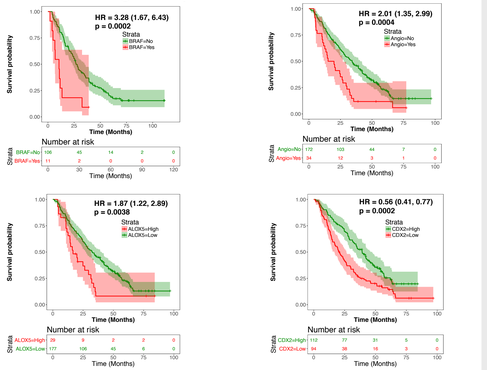
Fig. 10: Top Causal Drivers from Simulations



Validation of Causal Drivers of OS

- Identified causal drivers were validated in independent cohorts using univariate Cox proportional hazard model. HR, 95% CI, and p-value are shown in the plots below.

Fig. 11: Kaplan Meier Survival Curves



CONCLUSIONS

- Bayesian causal modeling identified clinical and molecular causal drivers (prognostic biomarkers) of OS for mCRC. The molecular drivers were validated in independent cohorts.
- 1º side, ECOG score, AST, LDH, HGB, and metastases (intra-abdominal, and liver) were the top clinical drivers of OS.
- BRAF & RAS mutations, CMS4, and angiogenesis/ ECM remodeling signature were top molecular drivers of OS.
- Consistent with previous studies, ALOX5 and CDX2 were identified as causal driver genes of OS.
- A molecular pathway between 1º side and OS was identified. Investigation into the molecular underpinnings of sidedness in driving OS is currently in progress.
- The availability of the measures for the drivers at baseline will allow better risk stratification at initiation of treatment.
- Additional research, including prospective studies, is necessary to confirm these findings.

Support: U10CA180821, U10CA180882, U10CA180888; Eli Lilly and Company, Genentech, Pfizer.

ClinicalTrials.gov Identifier: NCT00265850