Prognostic Models in Oncology: From Feature Selection to Web Apps

James P. Long

University of Texas MD Anderson Cancer Center

UT Health San Antonio – December 5, 2023

Contact: jplong@mdanderson.org

Collaboration

James Long
MDACC

Yu Shen
MDACC

Amy Heimberger
Northwestern

THE UNIVERSITY OF TEXAS
MD Anderson Cancer Center
Making Cancer History®

Northwestern Medicine®
Feinberg School of Medicine

Long, J.P. and Shen, Y., 2023. Detection method has independent prognostic significance in the PLCO lung screening trial. *Scientific reports*, *13*(1), p.13382.

Sherise D Ferguson, Tiffany R Hodges, Nazanin K Majd, Kristin Alfaro-Munoz, Wajd N Al-Holou, Dima Suki, John F de Groot, Gregory N Fuller, Lee Xue, Miao Li, Carmen Jacobs, Ganesh Rao, Rivka R Colen, Joanne Xiu, Roel Verhaak, David Spetzler, Mustafa Khasraw, Raymond Sawaya, **James P Long**, Amy B Heimberger. A validated integrated clinical and molecular glioblastoma long-term survival-predictive nomogram. *Neuro-oncology advances*, *3*(1), p.vdaa146.

Outline

Prognostic Models

Feature: Cancer Detection Method

Models and Algorithms: Random Survival Forests

Web Apps for Sharing Models

Outline

What are Prognostic Models?

Variables / Features / Predictors

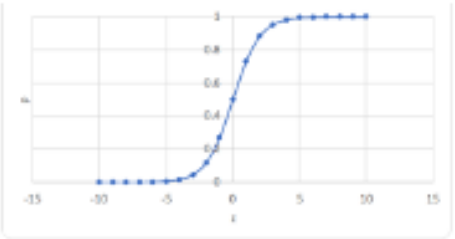
Clinical
- Age
- Disease Subtype
- Prior Treatments
- .
- .
- .

Genomic
- PTEN Mutation
- RB1 Mutation
- KRAS Expression
- .
- .
- .

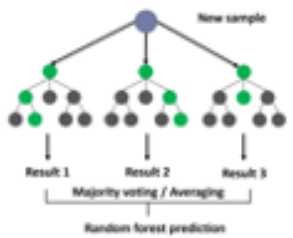
Imaging
- T1 Flair
- Necrosis Volume
- .
- .
- .

Model / Algorithm

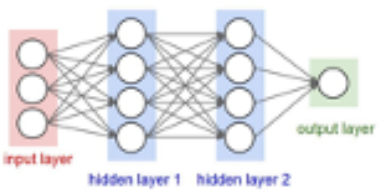
Logistic Regression



Random Forests



Deep Neural Network

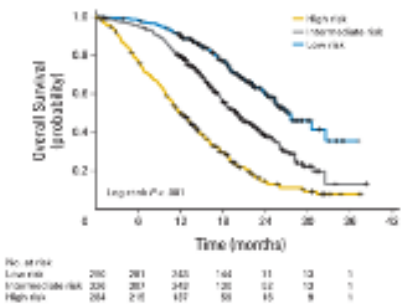


Predictions

Responder

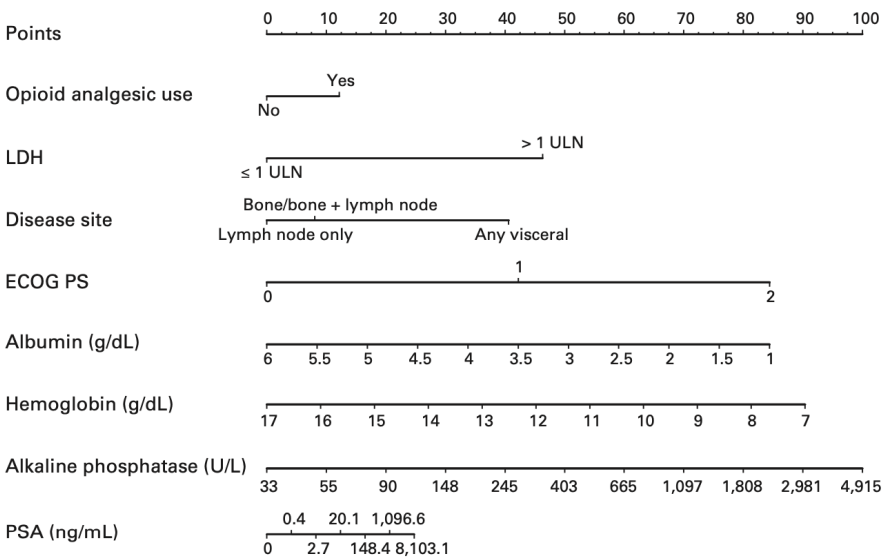
vs.

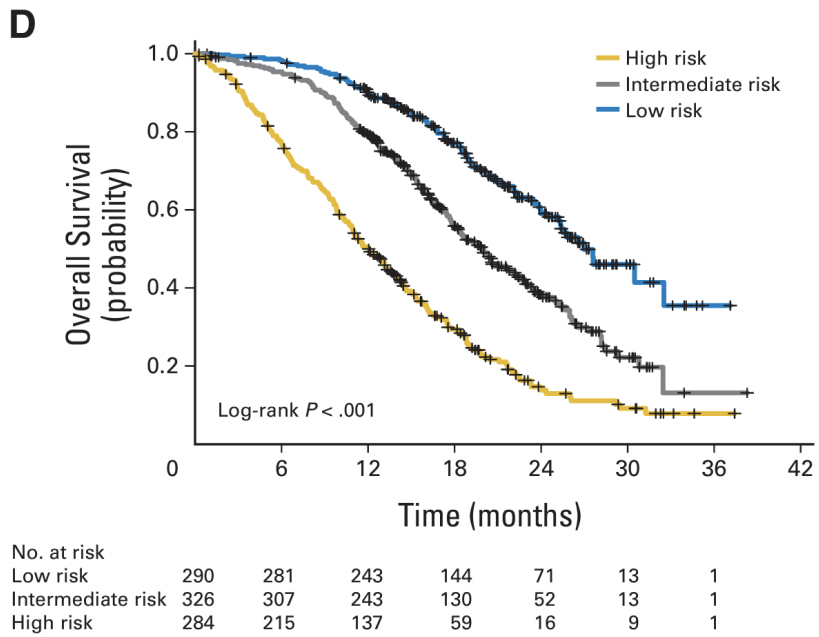
Non-Responder



Example 1: mCRPC

- Halabi et al. [2014] developed prognostic model for predicting time from metastasis to death in metastatic castrate resistant prostate cancer (mCRPC)
- Model score stratifies patients into risk groups [Halabi et al., 2023]





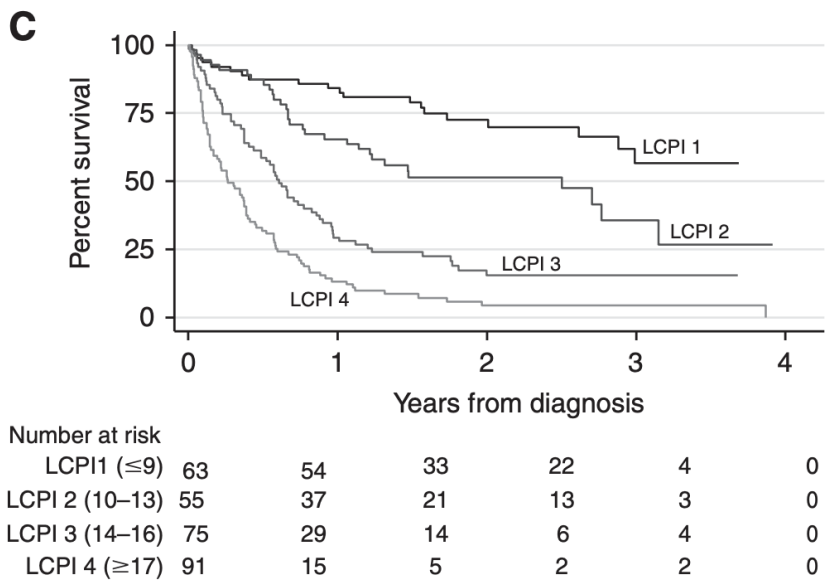
Example 2: Lung Cancer

- Alexander et al. [2017] developed Lung Cancer Prognostic Index (LCPI) for predicting time from diagnosis to death in Lung Cancer

Table 3. Weighted scores for predictive model for overall survival in newly diagnosed non-small-cell lung cancer

	LCPI points	m-LCPI
Stage group		
I	0	0
II	2	2
IIIA	5	4
IIIB	7	6
IV	9	8
NSCLC NOS	3	3
No proven actionable mutation[a]	3	3
ECOG performance status ⩾2	3	Excluded
Ever smoker	2	2
Respiratory comorbidity	2	2
Weight loss >10%	2	Excluded
Male sex	1	1
Age group 50 or less		
⩽50 years	0	0
51–70	1	1
71–90	2	2
⩾91 years	3	3
LCPI Group	LCPI score	m-LCPI score
LCPI 1	⩽9	⩽8
LCPI 2	10–13	9–11
LCPI 3	14–16	12–14
LCPI 4	⩾17	⩾15

Abbreviations: ALK = anaplastic lymphoma kinase; ECOG = Eastern Cooperative Oncology Group; EGFR = epidermal growth factor receptor; LCPI = Lung Cancer Prognostic Index; m-LCPI = modified lung cancer prognostic index; NOS = not otherwise specified; NSCLC = non-small-cell lung cancer.
[a]Includes EGFR/ALK negative, KRAS positive, mutation not tested.



Today's Talk

Variables / Features / Predictors

Clinical
- Age
- Disease Subtype
- Prior Treatments
- .
- .
- .

Genomic
- PTEN Mutation
- RB1 Mutation
- KRAS Expression
- .
- .
- .

Imaging
- T1 Flair
- Necrosis Volume
- .
- .
- .

1) Detection Method as Prognostic Feature

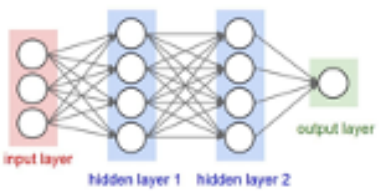
Model / Algorithm

Logistic Regression

2) Random Survival Forests as Alternative to Cox PH Model

Random Forests

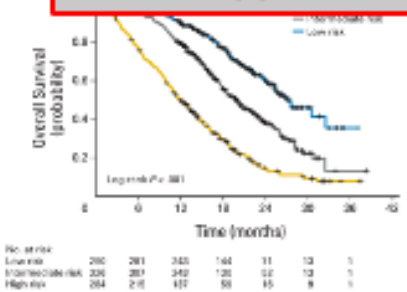
Deep Neural Network



Predictions

Responder
vs.
Non-Responder

3) Sharing Models with Shiny Web Apps

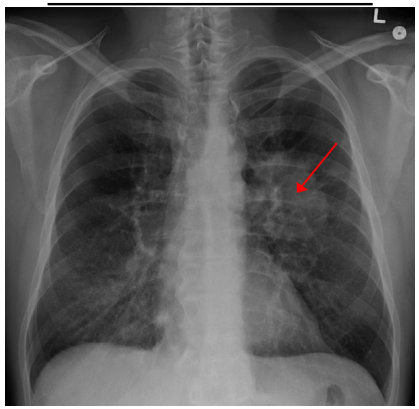


Outline

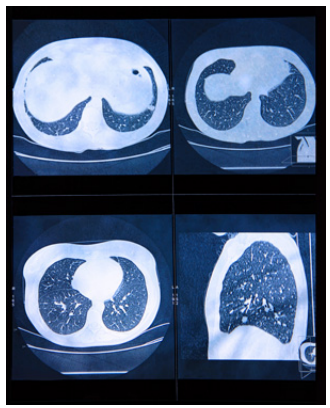
Cancer Screening

- Cancer screening has the potential to detect tumors early, prior to clinical symptoms.
- Cancers diagnosed at an early stage (I or II) have better prognosis than those diagnosed at a late stage (III or IV)

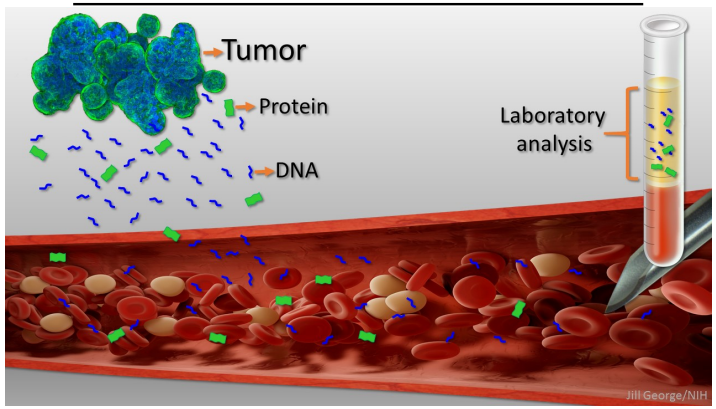
Chest X–Ray



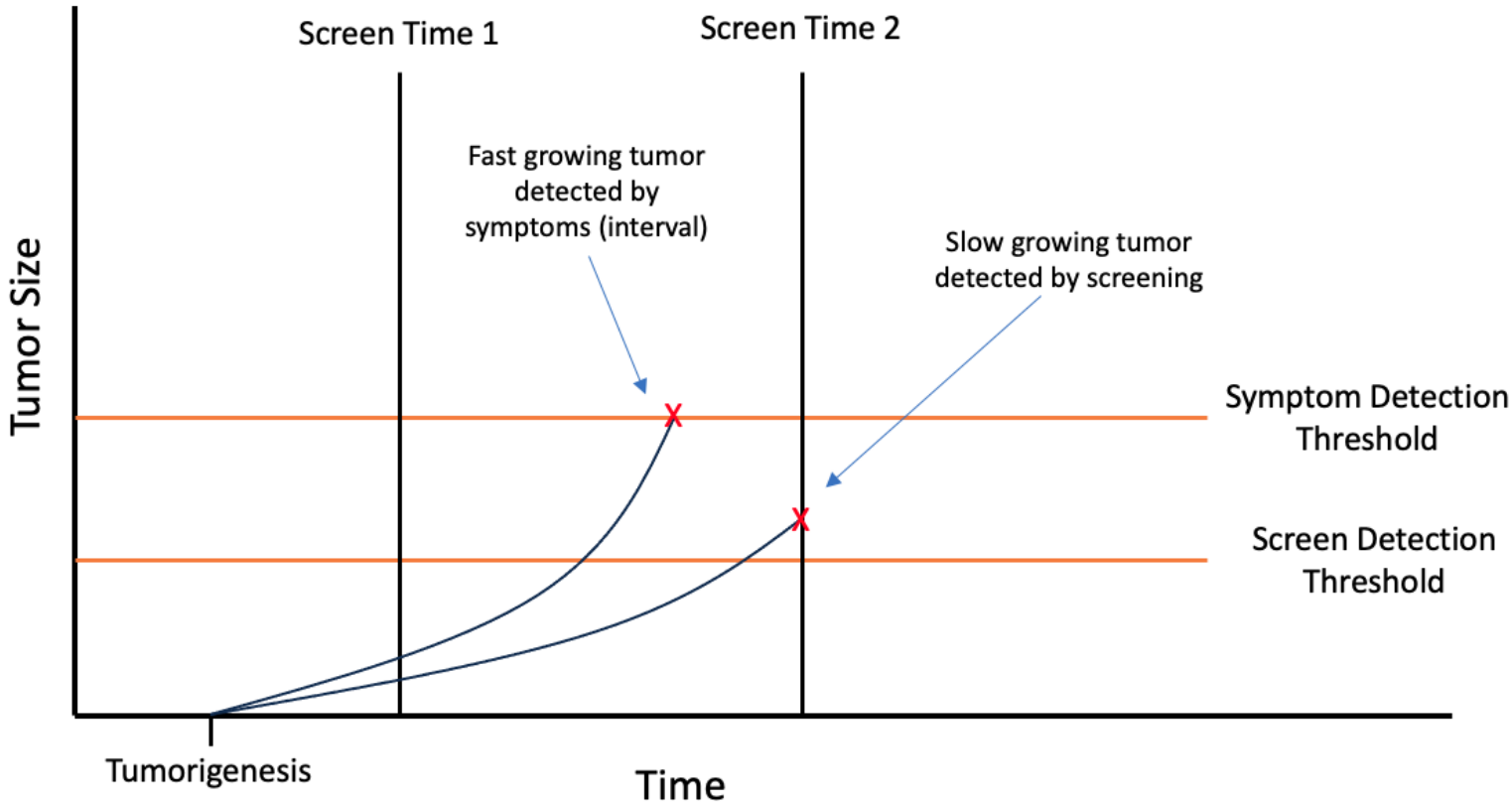
Low–Dose CT



Circulating Biomarkers



Detection Method and Tumor Growth Rate



Screen detected tumors are (on average) slower growing than interval detected tumors. [Albert et al., 1978, Morrison, 1992, Kramer et al., 2021]

Detection Method and Prognostic Significance

- Detection Method = Screen or Interval
- Expect screen detected tumors to have better prognosis:

  Screen detection $\rightarrow$ earlier stage at diagnosis $\rightarrow$ better prognosis
- Screen detected tumors tend to be slower growing than interval detected tumors.
  - Tumor growth rate is rarely known, not used in prognostic models.
  - Detection method may be a proxy for tumor growth rate.

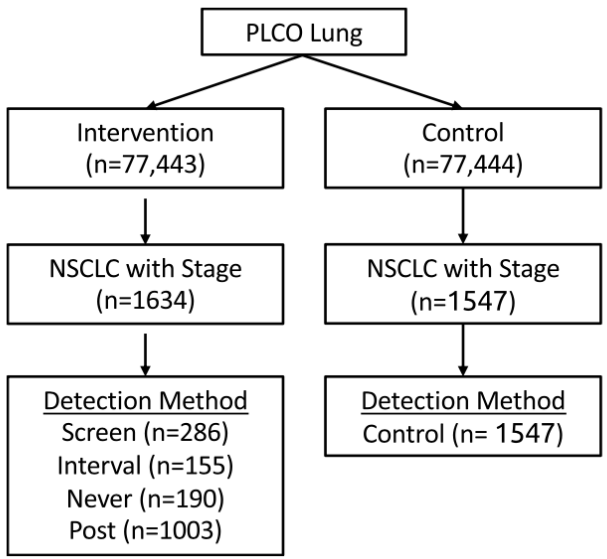
Existing Literature and Hypothesis

- Joensuu et al. [2004] found lower risk of distant recurrence in Screen detected breast cancers relative to interval detected.
- Shen et al. [2005] found better survival in screen–detected breast cancer
- Mook et al. [2011] found screen detected breast tumors have better survival
- Most existing work studying prognostic value of detection method focused on breast cancer and mammograms

Hypothesis: Detection Method has independent prognostic significance in predicting survival time in patients diagnosed with lung cancer.

PLCO Screening Trial

- We obtained data from the PLCO Cancer Screening Trial [Andriole et al., 2005]
- PLCO assessed efficacy of screening in 4 cancer types
- The Lung cohort of PLCO compared 4 annual chest X-rays to standard of care



Detection Method:

- Screen = Detected at annual X-ray screening
- Interval = Detected between annual screenings
- Never = Detected in patient who did not attend screenings
- Post = Detected after screening period ended
- Control = Detected in control group.

PLCO Screening Trial

- Performed two analyses:
  - 4-year cohort: Patients diagnosed lung cancer within 4 years after randomization
  - Extended cohort: All patients diagnosed with NSCLC with Stage information
- Grouped Never and Post detected tumors into category Other
  - Detection Method $= \{$Screen,Interval,Other,Control$\}$

Cohort Summary

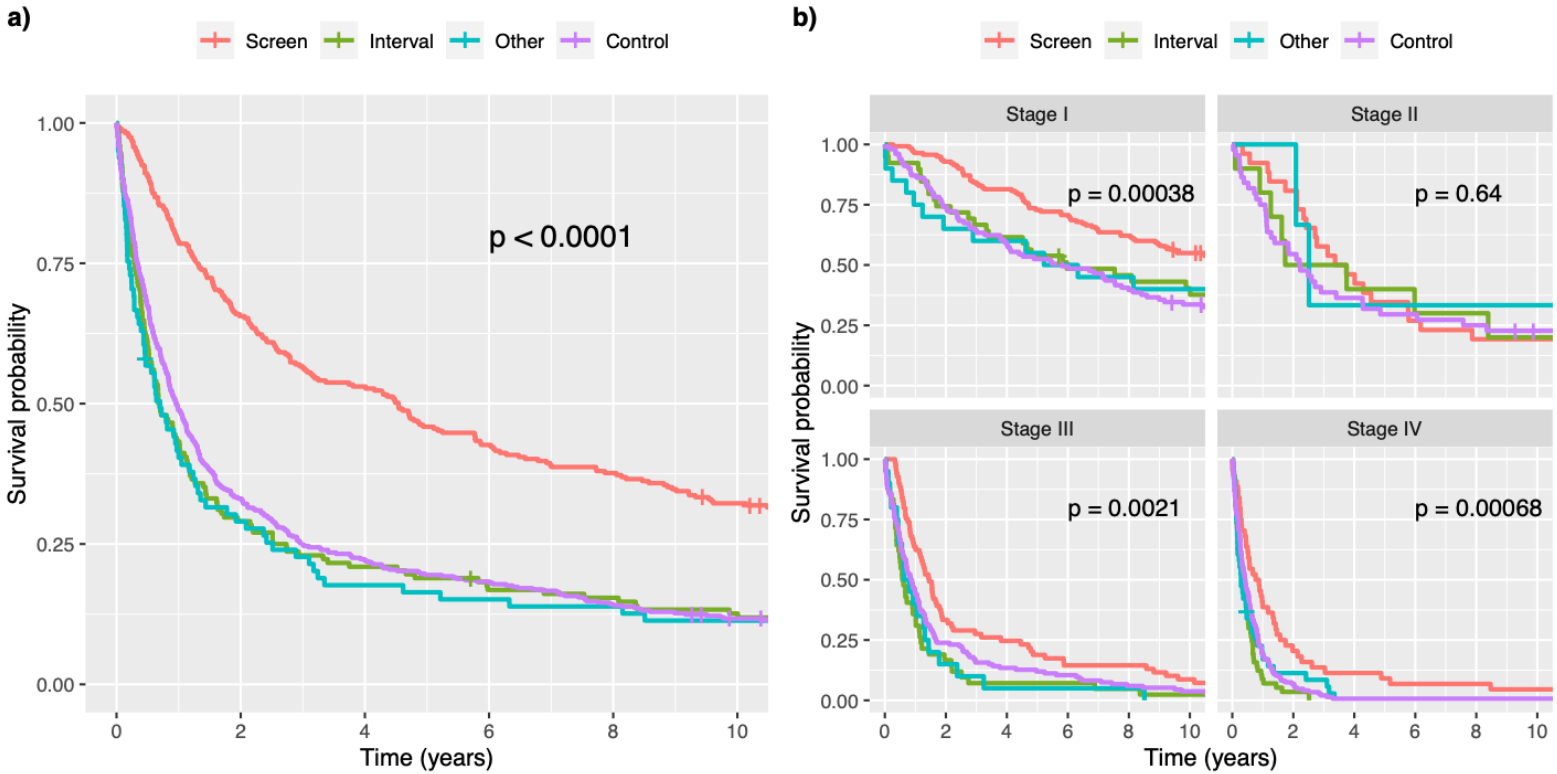
Table 1. Characteristics of lung cancers detected in PLCO.

Characteristic	**Screen**, N = 279	**Interval**, N = 148	**Other**, N = 81	**Control**, N = 426	**p-value**
Stage					<0.001
Stage I	140 (50%)	39 (26%)	20 (25%)	101 (24%)	
Stage II	26 (9.3%)	10 (6.8%)	3 (3.7%)	44 (10%)	
Stage III	69 (25%)	42 (28%)	20 (25%)	134 (31%)	
Stage IV	44 (16%)	57 (39%)	38 (47%)	147 (35%)	
Age					0.5
<=59	58 (21%)	37 (25%)	11 (14%)	83 (19%)	
60-64	67 (24%)	35 (24%)	24 (30%)	125 (29%)	
65-69	98 (35%)	46 (31%)	28 (35%)	126 (30%)	
>=70	56 (20%)	30 (20%)	18 (22%)	92 (22%)	
Sex					0.5
Female	113 (41%)	59 (40%)	36 (44%)	157 (37%)	
Male	166 (59%)	89 (60%)	45 (56%)	269 (63%)	
Smoked					<0.001
No	23 (8.2%)	8 (5.4%)	5 (6.2%)	29 (6.8%)	
Yes	256 (92%)	139 (94%)	63 (78%)	377 (88%)	
Unknown	0 (0%)	1 (0.7%)	13 (16%)	20 (4.7%)	
Histology					<0.001
Adenocarcinoma	134 (48%)	58 (39%)	29 (36%)	189 (44%)	
Bronchiolo-alveolar	33 (12%)	10 (6.8%)	2 (2.5%)	20 (4.7%)	
Squamous cell	59 (21%)	33 (22%)	23 (28%)	109 (26%)	
Large cell	21 (7.5%)	10 (6.8%)	3 (3.7%)	26 (6.1%)	
Other NSC	6 (2.2%)	4 (2.7%)	2 (2.5%)	4 (0.9%)	
Carcinoma, NOS	21 (7.5%)	28 (19%)	21 (26%)	74 (17%)	
Other/Unknown	5 (1.8%)	5 (3.4%)	1 (1.2%)	4 (0.9%)	

n (%)

Fisher's Exact Test for Count Data with simulated p-value(based on 2000 replicates)

Survival by Detection Method



- Left: Screen detected tumors how much longer survival times than other types.
- Right: Difference remains significant after controlling for stage.
  - This suggests Detection Method has independent prognostic significance (perhaps because proxy for tumor growth rate).

LCPI [Alexander et al., 2017]

1. Used a Cox Proportional Hazards model (Cox PH) to identify important predictors of survival for NSCLC in Australian cohort.
2. Points assigned to values of variable (Stage II $= 2$ points, Stage IV $= 9$ points) based on hazard ratios in Cox model.
3. Cutoffs chosen based on total points to classify patients into LCPI I through IV.
4. Kaplan Meier survival curves compared for LCPI I - IV on validation sets.

Step 1

Table 2. Predictors of survival in the derivation cohort by multivariate Cox pro

	LCPI			
	HR	Log HR	95% CI	*P*-value[a]
Stage group				
I	1.0	0.00		
II	1.57	0.45	0.96–2.58	0.09
IIIA	3.06	1.12	2.05–4.57	<0.01
IIIB	4.45	1.49	2.91–6.84	<0.01
IV	7.98	2.08	5.55–11.52	<0.01
NSCLC NOS	1.91	0.65	1.48–2.65	<0.01
No proven actionable mutation[b]	1.91	0.65	1.31–2.53	<0.01
ECOG performance status ⩾2	1.78	0.57	1.40–2.23	<0.01
Ever Smoker	1.65	0.50	1.19–2.35	<0.01
Respiratory comorbidity	1.46	0.38	1.11–1.68	<0.01
Weight loss >10%	1.42	0.35	1.13–1.79	<0.01
Male sex	1.36	0.31	1.11–1.68	0.01
Age (per 20 years aged >50)	1.25	0.22	1.05–1.50	<0.01

Abbreviations: ALK = anaplastic lymphoma kinase; CI = confidence interval; ECOG, Eastern Cooperative Oncology
LCPI = lung cancer prognostic index; m-LCPI = modified lung cancer prognostic index; NOS = not otherwise specifie
[a] *P*-value for the comparison of overall survival for patients with and without the specified factor, or for stage group
[b] Includes EGFR/ALK negative, KRAS positive, mutation not tested.

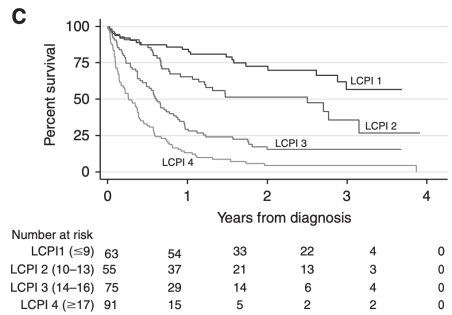
Step 2 + 3

Table 3. Weighted scores for predictive model for overall survival in newly diagnosed non-small-cell lung cancer

	LCPI points	m-LCPI
Stage group		
I	0	0
II	2	2
IIIA	5	4
IIIB	7	6
IV	9	8
NSCLC NOS	3	3
No proven actionable mutation[a]	3	3
ECOG performance status ⩾2	3	Excluded
Ever smoker	2	2
Respiratory comorbidity	2	2
Weight loss >10%	2	Excluded
Male sex	1	1
Age group 50 or less		
⩽50 years	0	0
51–70	1	1
71–90	2	2
⩾91 years	3	3
LCPI Group	LCPI score	m-LCPI score
LCPI 1	⩽9	⩽8
LCPI 2	10–13	9–11
LCPI 3	14–16	12–14
LCPI 4	⩾17	⩾15

Abbreviations: ALK = anaplastic lymphoma kinase; ECOG = Eastern Cooperative Oncology Group; EGFR = epidermal growth factor receptor; LCPI = Lung Cancer Prognostic Index; m-LCPI = modified lung cancer prognostic index; NOS = not otherwise specified; NSCLC = non-small-cell lung cancer.
[a] Includes EGFR/ALK negative, KRAS positive, mutation not tested.

Step 4



Testing Hypothesis

1. Identified all significant prognostic variables in LCPI which were also available in PLCO data set.
2. Added Detection Method variable to this set.
3. Fit a Cox Proportional Hazards model on data (following Alexander et al. [2017]).

Cox PH Model Summary

Characteristic	log(HR)	95% CI	p-value
Detection			
Screen	—	—	
Interval	0.61	0.39, 0.83	<0.001
Other	0.46	0.19, 0.74	0.001
Control	0.43	0.26, 0.60	<0.001
Stage			
Stage I	—	—	
Stage II	0.50	0.24, 0.77	<0.001
Stage III	1.3	1.1, 1.5	<0.001
Stage IV	2.0	1.8, 2.3	<0.001
Age			
<=59	—	—	
60-64	0.28	0.07, 0.49	0.008
65-69	0.36	0.16, 0.57	<0.001
>=70	0.71	0.49, 0.93	<0.001
Sex			
Female	—	—	
Male	0.17	0.02, 0.31	0.024
Smoked			
No	—	—	
Yes	0.35	0.06, 0.65	0.020
Unknown	0.37	-0.10, 0.83	0.12
Histology			
Adenocarcinoma	—	—	
Bronchiolo-alveolar	-0.54	-0.85, -0.23	<0.001
Squamous cell	0.13	-0.05, 0.31	0.2
Large cell	-0.22	-0.51, 0.07	0.13
Other NSC	-0.22	-0.76, 0.32	0.4
Carcinoma, NOS	0.15	-0.05, 0.35	0.14
Other/Unknown	0.17	-0.35, 0.70	0.5

HR = Hazard Ratio, CI = Confidence Interval

- p-values for Detection Method highly significant
- Hazard ratios for Detection Method larger than for several other prognostic variables such as Sex and Histology
- C–Indices:

	C-Index
LCPI	0.74
PLCO Lung	0.76

Outline

Proportional Hazards (PH) Assumption

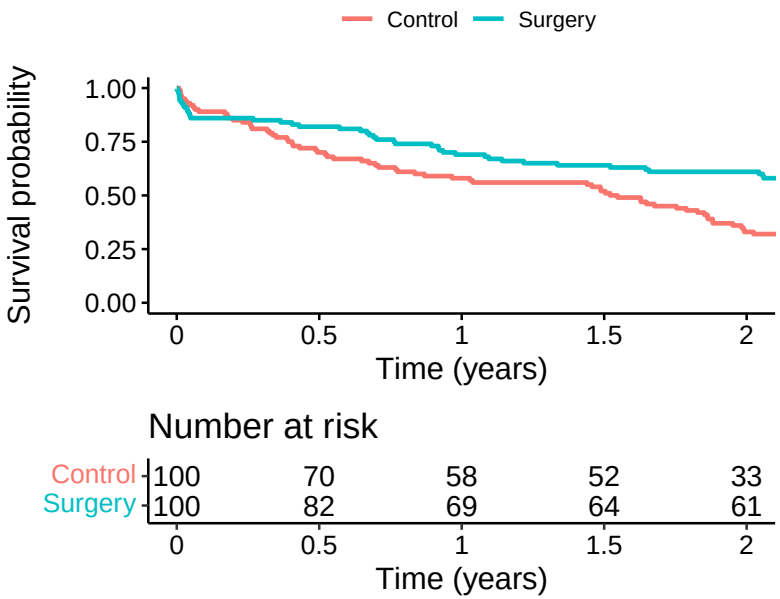
- Hazard at time t = death rate at time t
- Cox Proportional Hazards (Cox PH) model assumes that hazard ratios do not vary with time

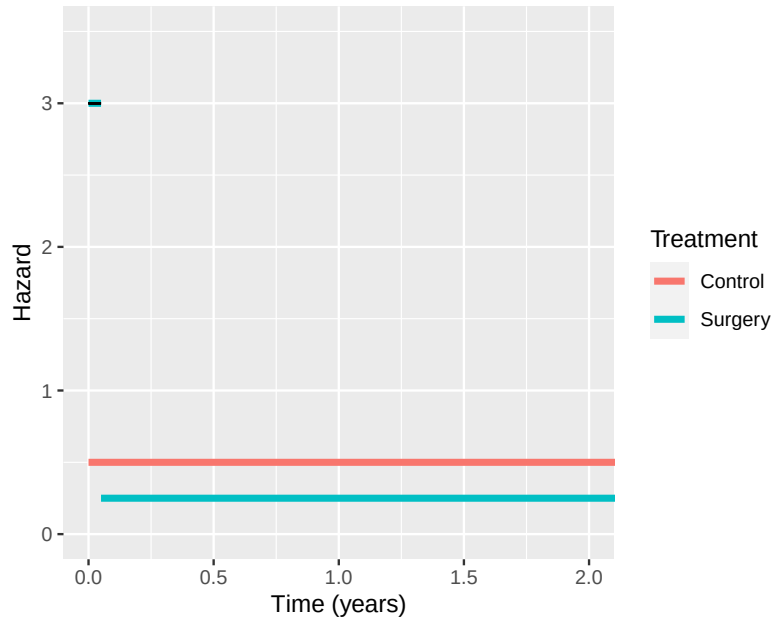
$$\widehat{\text{HR}} = \frac{\text{Interval Hazard}}{\text{Screen Hazard}} = 1.84$$

- Reality can be much more complicated
  - The hazard ratio may change with time.
  - The hazard ratio may be different for different values of other covariates, e.g. males and females.

Example of Violation of PH Assumption

- Patients randomized to risky surgery or control
- Surgery hazard: Initially high because of surgical complications but lower at later times due to tumor removal
- Control hazard: Constant





$$HR = \frac{\text{Surgery Hazard}}{\text{Control Hazard}} = \begin{cases} 6 & 0 < t < 0.1 \\ 0.5 & t > 0.1 \end{cases}$$

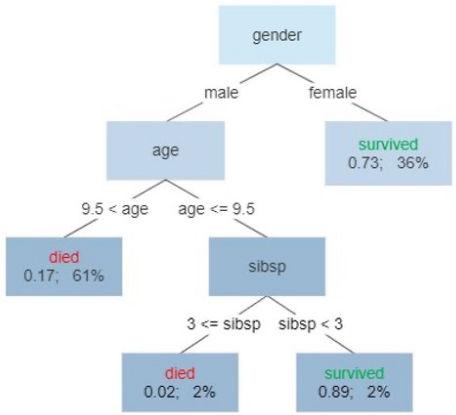
$$\text{Cox PH Estimated Hazard} = 0.57$$

Consequences of Violations in Model Assumptions

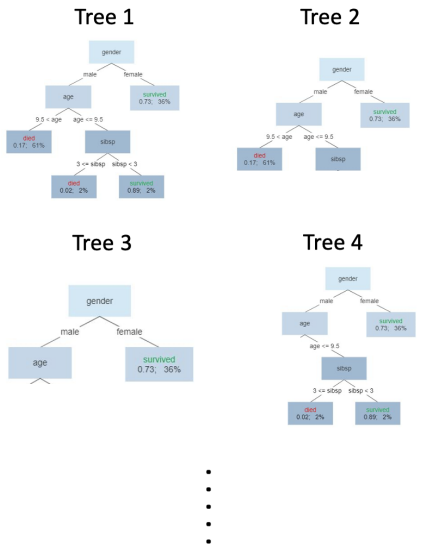
- "All models are wrong, some models are useful."
  - George E.P. Box
- Severity in violation of PH assumption is important
- Concerns with PLCO Lung Application:
  1. Detection Method may no longer be a significant predictor if a different model is used
  2. Other models may obtain better prediction performance (e.g. C–index)

Random Survival Forests (RSF)

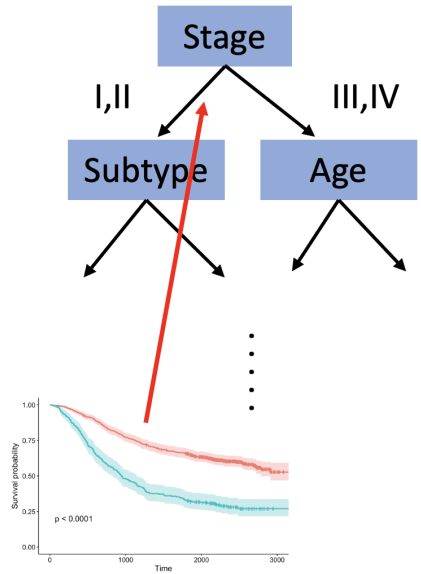
CART (1984)



Random Forest (2001)



RSF (2008)



	Interpretation	Prediction	Assumptions	Data Size
Cox PH	✓			Smaller
RSF		✓	✓	Larger

Breiman et al. [1984], Breiman [2001], Ishwaran et al. [2008]

RSF Variable Importance

- Project Goal: Assess importance of Detection Method in Prognostic models
  - CoxPH variable importance usually measured by size of HR and p–values
  - RSF does not directly compute HR or p–values
- RSF Variable Importance
  1. Values of a variable X are permuted among the patients.
  2. Compute C-index using permuted X variable.
  3. VIMP $=$ C-index with true X - C-index with permuted X.

Results of Random Survival Forests

- Variable Importance (VIMP) Scores:

	4-Year Cohort	Extended Cohort
Stage	0.17631	0.16060
Detection	0.01440	0.00425
Age	0.00671	0.00601
Histology	0.00382	0.00713
Sex	0.00085	0.00311
Smoked	0.00064	-0.00012

Detection Method is second most important predictor in 4-year cohort.

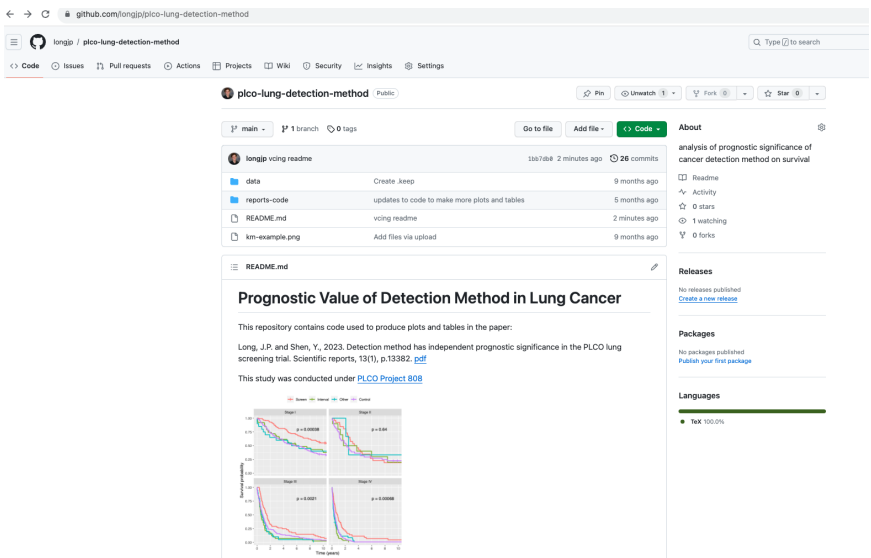
- C–Indices:

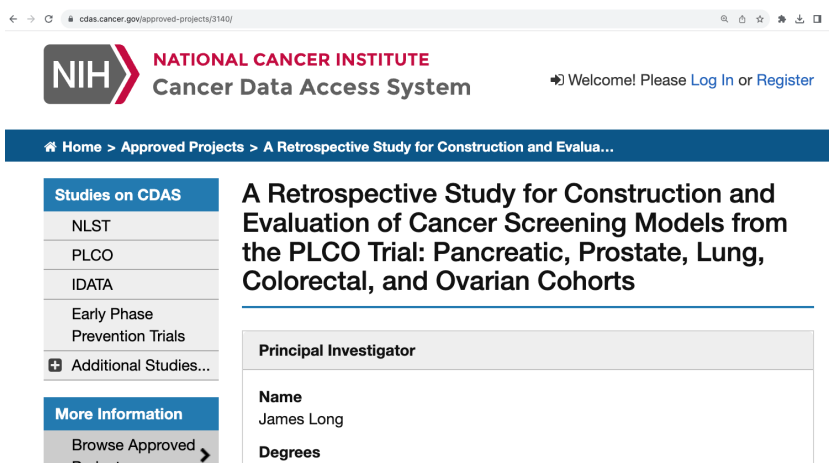
	C-Index
Cox PH	0.76
RSF	0.75

Similar between two models.

Reproducible Results

- Project analysis done in RMarkdown
- All code publicly available
- PLCO data may be requested from CDAS





`https://github.com/longjp/plco-lung-detection-method`

Outline

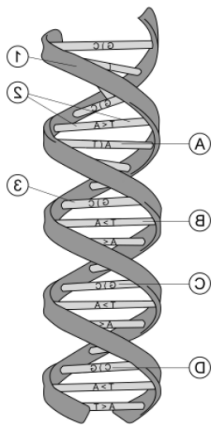
Glioblastoma Multiforme (GBM)

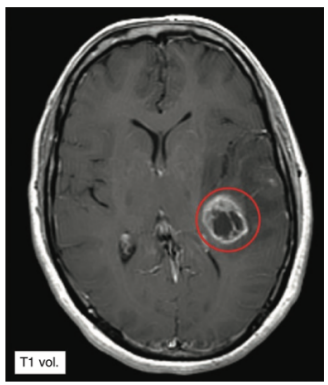
- GBM is the most deadly form of brain cancer
- 5-year survival rate $< 10\%$
- Question: What variables (clinical, imaging, genomic) are predictive of survival? Can we identify long-term survivors?

Project Outline

MD Anderson GBM Patient Cohort

- STS: < 6 months (n=37)
- MTS: 6 months - 5 years (n=22)
- LTS: >5 years (n=21)





Proportions Odds Ordinal Regression Model with 5 features

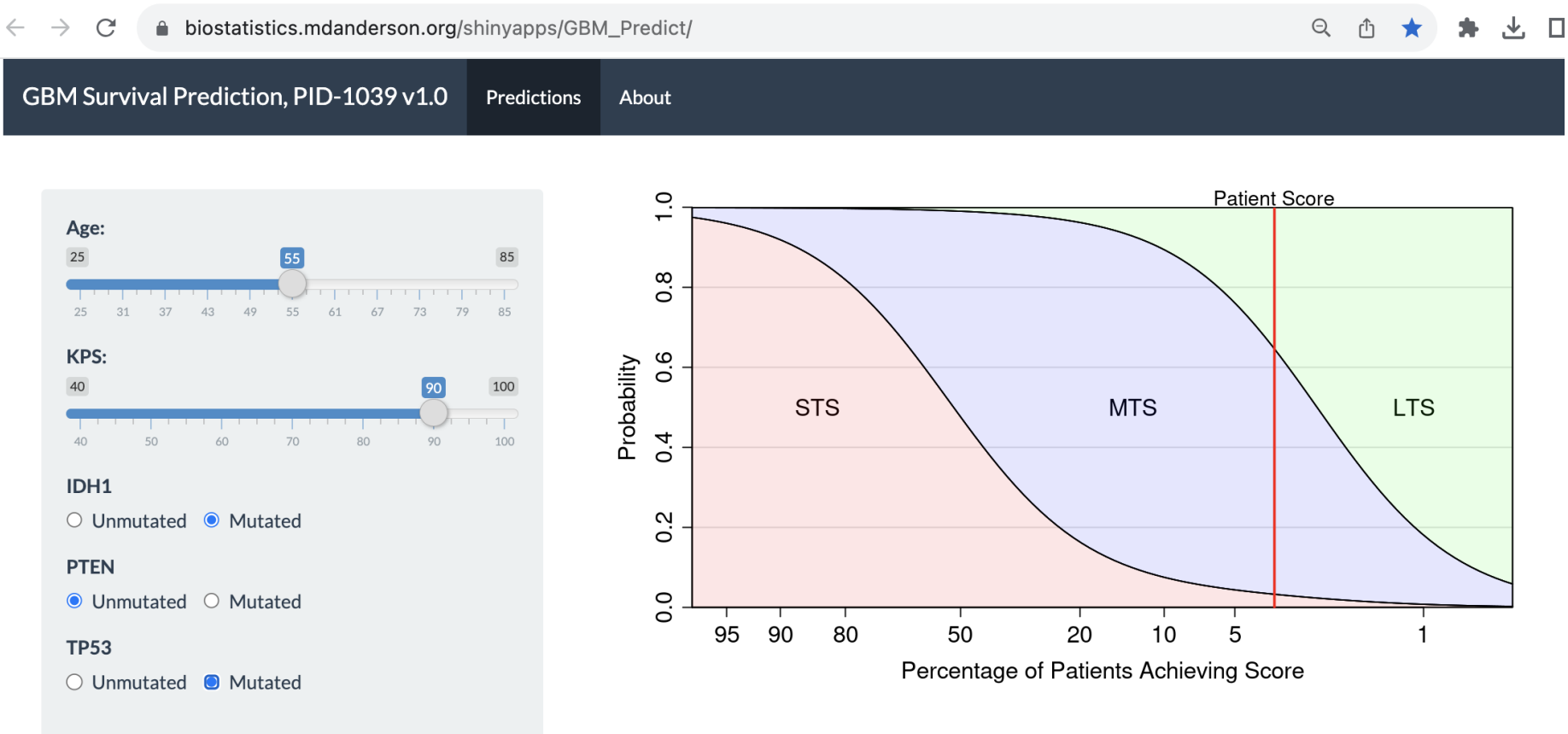
THE CANCER GENOME ATLAS

Validation Cohort

?

- Considered publishing model as nomogram in paper
  - These can be difficult to use / inaccessible to patients / caregivers

Shiny Web App



https://biostatistics.mdanderson.org/shinyapps/GBM_Predict/

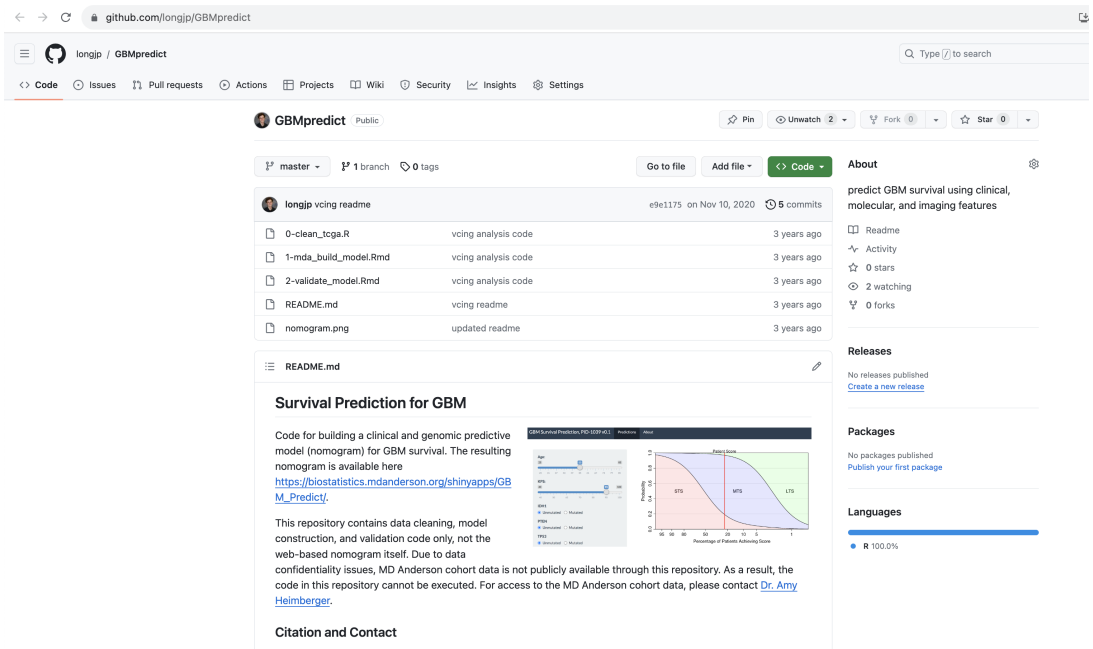
Reproducible Research

Neuro-Oncology Advances

3(1), 1–10, 2021 | doi:10.1093/noajnl/vdaa146 | Advance Access date 31 October 2020

A validated integrated clinical and molecular glioblastoma long-term survival-predictive nomogram

Sherise D. Ferguson†, Tiffany R. Hodges†, Nazanin K. Majd, Kristin Alfaro-Munoz, Wajd N. Al-Holou, Dima Suki, John F. de Groot, Gregory N. Fuller, Lee Xue, Miao Li, Carmen Jacobs, Ganesh Rao, Rivka R. Colen, Joanne Xiu, Roel Verhaak, David Spetzler, Mustafa Khasraw, Raymond Sawaya, James P. Long†, and Amy B. Heimberger†



https://github.com/longjp/GBMpredict

Discussion

- Detection Method should be recorded in clinical trial data bases / cancer registries and considered as variable when constructing prognostic models.
- Modern machine learning / AI tools should be considered when constructing prognostic models. But they are not necessarily superior to existing methods on given data set. Comparison of new tool with existing methods is critical.
- Prognostic models can be deployed as web applications to facilitate / accelerate use by clinicians, patients, and caregivers

Thank you. Questions?

Bibliography I

A. Albert, P. M. Gertman, T. A. Louis, and S.-i. Liu. Screening for the early detection of cancer—ii. the impact of screening on the natural history of the disease. *Mathematical Biosciences*, 40(1-2):61–109, 1978.

M. Alexander, R. Wolfe, D. Ball, M. Conron, R. G. Stirling, B. Solomon, M. MacManus, A. Officer, S. Karnam, K. Burbury, et al. Lung cancer prognostic index: a risk score to predict overall survival after the diagnosis of non-small-cell lung cancer. *British journal of cancer*, 117(5):744–751, 2017.

G. L. Andriole, D. L. Levin, E. D. Crawford, E. P. Gelmann, P. F. Pinsky, D. Chia, B. S. Kramer, D. Reding, T. R. Church, R. L. Grubb, et al. Prostate cancer screening in the prostate, lung, colorectal and ovarian (plco) cancer screening trial: findings from the initial screening round of a randomized trial. *Journal of the National Cancer Institute*, 97(6):433–438, 2005.

L. Breiman. Random forests. *Machine learning*, 45:5–32, 2001.

L. Breiman, J. Friedman, R. Olshen, and C. Stone. Cart. *Classification and regression trees*, 1984.

S. Halabi, C.-Y. Lin, W. K. Kelly, K. S. Fizazi, J. W. Moul, E. B. Kaplan, M. J. Morris, and E. J. Small. Updated prognostic model for predicting overall survival in first-line chemotherapy for patients with metastatic castration-resistant prostate cancer. *Journal of Clinical Oncology*, 32(7):671, 2014.

S. Halabi, Q. Yang, A. Roy, B. Luo, J. C. Araujo, C. Logothetis, C. N. Sternberg, A. J. Armstrong, M. A. Carducci, K. N. Chi, et al. External validation of a prognostic model of overall survival in men with chemotherapy-naïve metastatic castration-resistant prostate cancer. *Journal of Clinical Oncology*, 41(15):2736–2746, 2023.

H. Ishwaran, U. B. Kogalur, E. H. Blackstone, and M. S. Lauer. Random survival forests. *The Annals of Applied Statistics*, 2 (3):841 – 860, 2008. doi: 10.1214/08-AOAS169. URL https://doi.org/10.1214/08-AOAS169.

H. Joensuu, T. Lehtimäki, K. Holli, L. Elomaa, T. Turpeenniemi-Hujanen, V. Kataja, A. Anttila, M. Lundin, J. Isola, and J. Lundin. Risk for distant recurrence of breast cancer detected by mammography screening or other methods. *Jama*, 292 (9):1064–1073, 2004.

B. S. Kramer, J. K. Gohagan, and P. C. Prorok. *Cancer screening: theory and practice*. CRC Press, 2021.

S. Mook, L. J. Van't Veer, E. J. Rutgers, P. M. Ravdin, A. O. van de Velde, F. E. van Leeuwen, O. Visser, and M. K. Schmidt. Independent prognostic value of screen detection in invasive breast cancer. *Journal of the National Cancer Institute*, 103 (7):585–597, 2011.

A. S. Morrison. Screening in chronic disease. *(No Title)*, 1992.

Y. Shen, Y. Yang, L. Y. Inoue, M. F. Munsell, A. B. Miller, and D. A. Berry. Role of detection method in predicting breast cancer survival: analysis of randomized screening trials. *Journal of the National Cancer Institute*, 97(16):1195–1203, 2005.