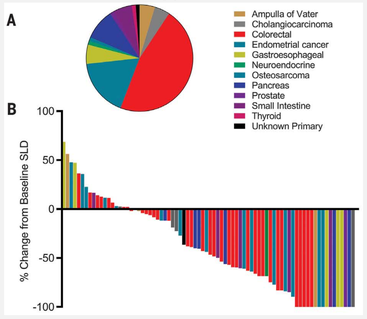
A Multifactorial Model of T Cell Expansion and Durable Clinical Benefit in Response to a PD-L1 Inhibitor

*Mark DM Leiserson[1,2], Vasilis Syrgkanis[1], Amy Gilson[1], Miroslav Dudik[1], Sharon Gillett[1], Jennifer Chayes[1], Christian Borgs[1], Dean F Bajorin[3,4], Jonathan Rosenberg[3], Samuel Funt[3,4], Alexandra Snyder[3,5], **Lester Mackey**[1,*]*

*[1]Microsoft Research; [2]Department of Computer Science, University of Maryland, College Park; [3]Department of Medicine, Memorial Sloan Kettering Cancer Center; [4]Department of Medicine, Weill Cornell Medical College; [5]Adaptive Biotechnologies; *Email:* lmackey@microsoft.com*

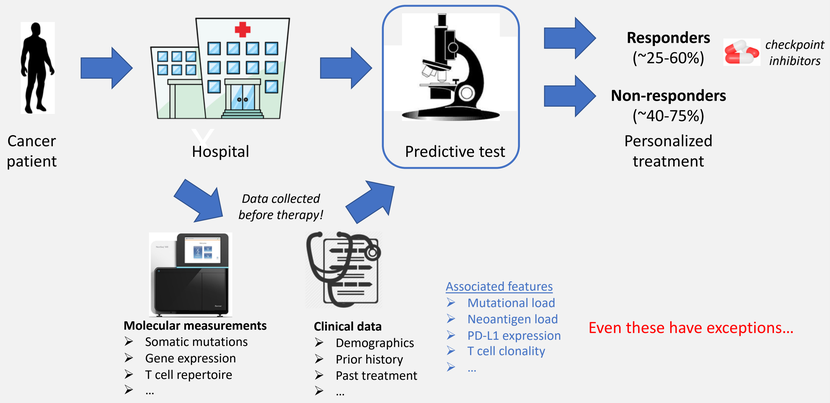
Predicting response to checkpoint inhibitor immunotherapy

Promise of checkpoint inhibitors
- Success in previously untreatable late-stage cancers
- Treatment scales across patients and cancer types



Mismatch-repair deficiency predicts response of solid tumors to PD-1 blockade
Le *et al.* (Science, 2017)

Biomarkers for response are a priority
- Efficacy
- Adverse effects
- Side effects



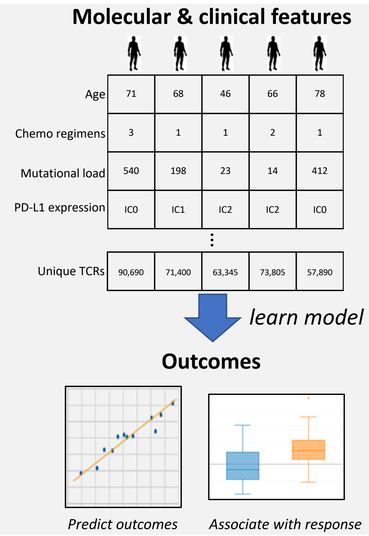
Our approach: predicting *molecular* response

Several key challenges
1. **Integrate** molecular and clinical features
2. **Select** appropriate outcome to predict
   A. Clinical responses are noisy, impacted by confounding factors, and poor proxies for molecular response
3. **Assess** value for new patients from small datasets
   A. Predictive accuracy
   B. Feature importance

How to predict the response of immune system post-therapy?

Our approach: Predict T cell clones in the tumor (TILs) that expand in the blood

Advantages over predicting standard clinical outcomes (e.g., overall survival or durable clinical benefit)
- More immediate outcome
- Finer-grained measurement



Features of 21 bladder cancers

N=21 patients treated with PD-L1 inhibitor (*atezoliuzmab*)

Clinical	Tumor	Circulating
Prior intravesical Bacillus Calmette–Guérin (BCG)	Missense SNV count (mutation load)	Productive unique TCR count
Age	Expressed missense SNV count	Clonality (TCR)
Albumin < 4	Neoantigen count	Diversity (TCR)
Baseline neutrophil to lymphocyte ratio	Expressed neoantigen count	T cell fraction
Time since last chemotherapy (days)	Clonality (TCR)	Top clone frequency (%)
5-factor score	Diversity (TCR)	
Number of chemotherapy regimens	T cell fraction	

Public dataset from Snyder, et al. (*PLoS Medicine*, 2017)

A multifactorial model of immune response

Approach: Predict number of **expanded TIL clones** in blood three weeks after treatment as a function of 19 clinical, tumor, and circulating features

Elastic Net: High-dimensional regression procedure designed to reduce overfitting and be robust to irrelevant features (*Hastie & Zhou, JRSS B, 2005*)

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}} ||y - X\beta||_2^2 + \lambda_1 ||\beta||_2^2 + \lambda_2 ||\beta||_1$$

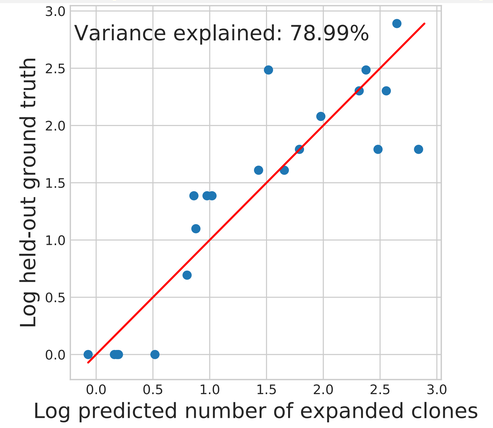
Cross-validation (CV): Assess prediction accuracy on unseen patients by repeating predictive analysis with each patient withheld

Permutation test: Test whether cross-validation error significantly smaller than error under permuted labeling ("Permutation Tests for Studying Classifier Performance," *Ojala & Garriga, JMLR, 2010*)

Our code: https://github.com/lrgr/multifactorial-immune-response

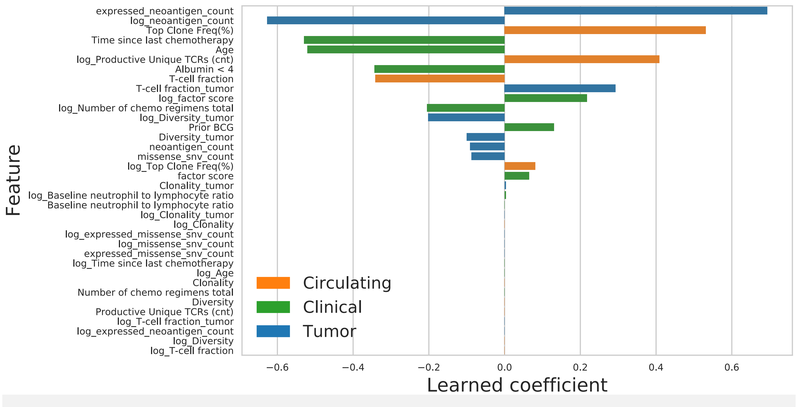
Accurate prediction of TIL clone expansion

Learned model **explains 79% of variance in held-out patients**

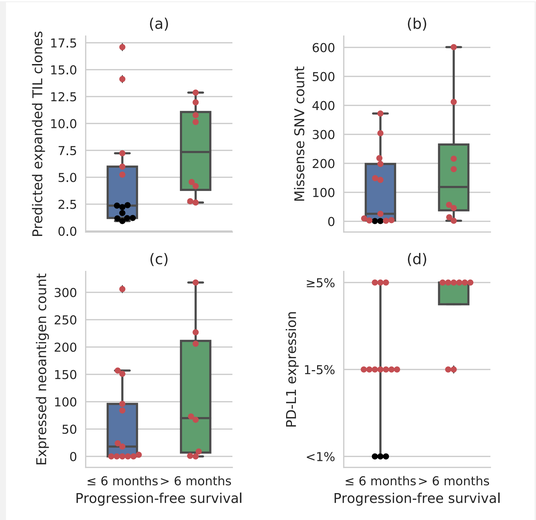


Clinical, tumor, and circulating feature importance

Clinical, tumor, and circulating features all contribute significantly
- Only 26% of variance explained if all tumor features excluded
- Only 13% of variance explained if all circulating features excluded
- 0% of variance explained if all clinical features excluded

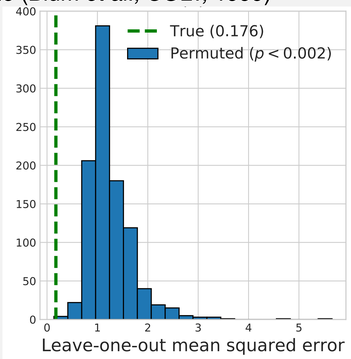


Implications for durable clinical benefit



Sanity-check test of CV error

Cross-validation error significantly worse when labels randomly permuted (1000 permutations)
- **Intuition:** CV error uninformative if comparable error achieved with wrong (permuted) labels
- Rules out perverse case where CV known to be unreliable (*Blum et al., COLT, 1999*)



Open questions

1. Can more predictive signal be extracted from previously completed studies of immunotherapy benefit?
2. What other cancer immunotherapy prediction problems could benefit from multifactorial modeling?
3. What value can patients or trial designers derive from accurate TIL expansion prediction?
4. What other "non-standard" phenotypes would be of value to predict?