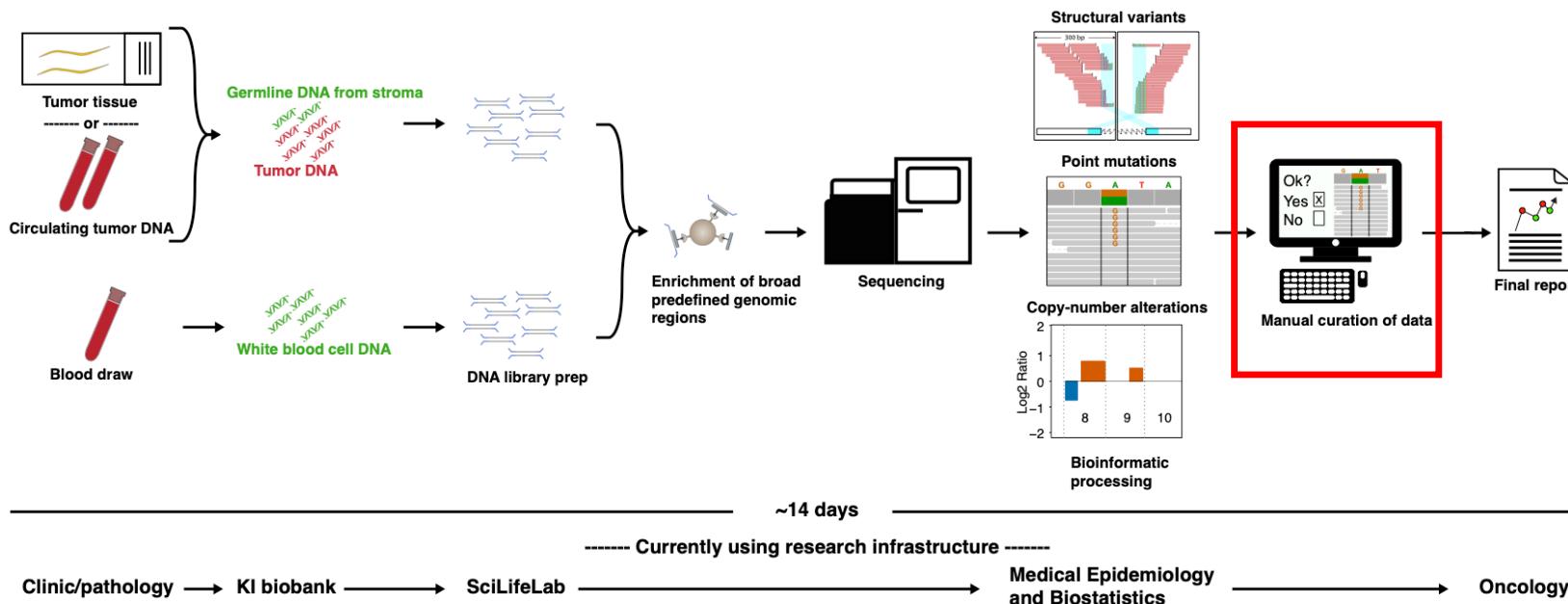


How to curate somatic- and germline variation for clinical (trial) use

Standardizing NGS-interpretation

- Manual review is still needed to avoid false positive variants
- To ensure performance over time, curation need to take place using standard operating procedures with “trained” staff.



A great paper for an introduction

- Purpose
 - Manual review is needed to obtain list of high-quality variants
 - Present a suggestion of standardized manual review

- Methods
 - SOP containing
 - 4 different calls
 - Somatic/Germline/Ambiguous/Fail
 - 19 tags
 - Data features, common artifacts, providing support for call
 - 4 reviewers classified the variants
 - prior to SOP
 - after reading SOP
 - Accuracy assessed with orthogonal sequencing

© American College of Medical Genetics and Genomics

ARTICLE | Genetics
inMedicine

Open

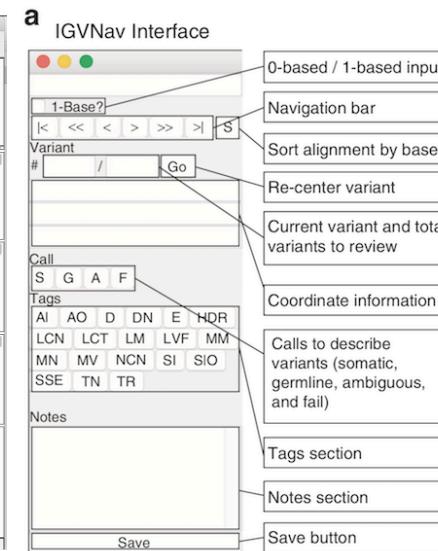
Standard operating procedure for somatic variant refinement of sequencing data with paired tumor and normal samples

Erica K. Barnell, BS¹, Peter Ronning, BS¹, Katie M. Campbell, BS¹, Kilannin Krysiak, PhD^{1,2}, Benjamin J. Ainscough, PhD^{1,3}, Lana M. Sheta¹, Shahil P. Pema¹, Alina D. Schmidt, BS¹, Megan Richters, BS¹, Kelsy C. Cotto, BS¹, Arpad M. Danos, PhD¹, Cody Ramirez, BS¹, Zachary L. Skidmore, MEng¹, Nicholas C. Spies, BS¹, Jasreet Hundal, MS¹, Malik S. Sediqzad¹, Jason Kunisaki, BS¹, Felicia Gomez, PhD¹, Lee Trani, BS¹, Matthew Matlock, BS¹, Alex H. Wagner, PhD¹, S. Joshua Swamidas, MD/PhD^{1,5}, Malachi Griffith, PhD^{1,2,3,6} and Obi L. Griffith, PhD^{1,2,3,6}

Standard operating procedure for somatic variant refinement of sequencing data with paired tumor and normal samples, Gen Med, 2018.

Clinical Cancer Genomics – vt 2022

IGVnav – addition software to IGV



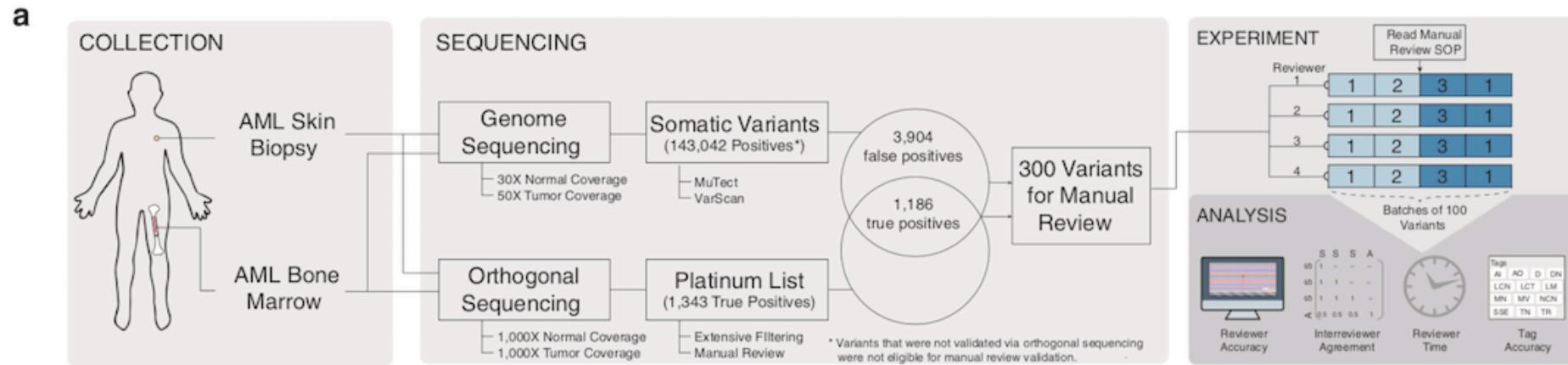
b IGVNav Input File

chr	start	stop	ref	var	call	tags	notes
10	26174201	26174201	C	A	S		
14	44505849	44505849	A	G	F	MM	
3	67004225	67004225	T	C	A	SI	
10	26174114	26174114	A	T	A	SI	
10	70753879	70753879	C	A	F	LVF	
10	94227337	94227337	C	T	F	HDR	
11	5390168	5390168	T	G	F	SI	
12	180263686	180263686	T	C			
12	122190233	122190233	C	A			
13	189125155	189125155	A	G			
13	23250679	23250679	G	A			
13	38691806	38691806	G	A			

c IGVNav Output File

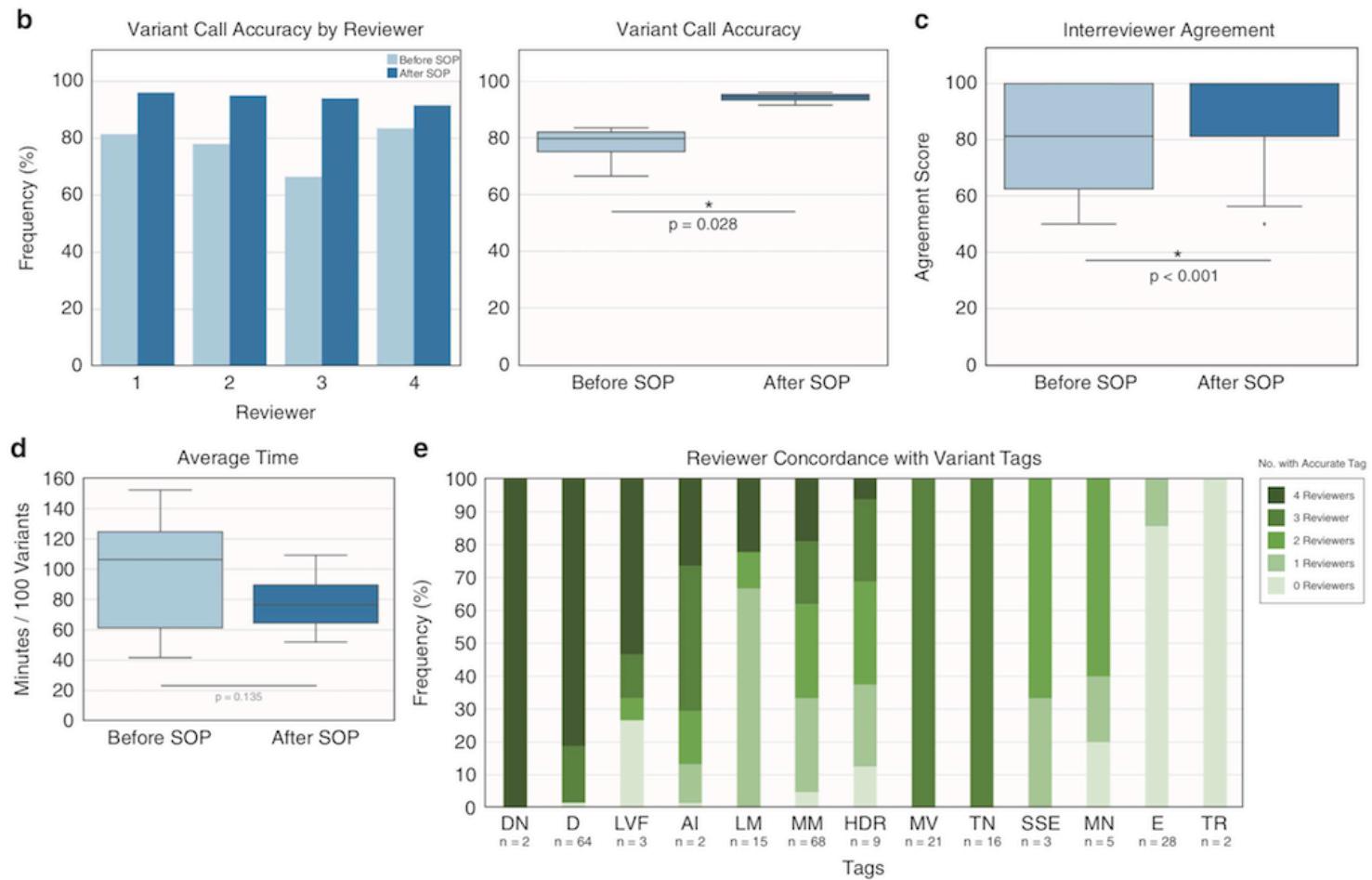
chr	start	stop	ref	var	call	tags	notes
10	26174201	26174201	C	A	S		
14	44505849	44505849	A	G	F	MM	
3	67004225	67004225	T	C	A	SI	
10	26174114	26174114	A	T	A	SI	
10	70753879	70753879	C	A	F	LVF	
10	94227337	94227337	C	T	F	HDR	
11	5390168	5390168	T	G	F	SI	
12	180263686	180263686	T	C	F	SI	
12	122190233	122190233	C	A	S	'dinucleotide'	
13	189125155	189125155	A	G	A	SIO	
13	23250679	23250679	G	A	A	SI	
13	38691806	38691806	G	A	F	MM	

Results



- Blinded novice reviewers manually reviewed 200 variants in two batches
- Important to use multiple tags to motivate the “Call”
- To fail a variant, the reviewer must confidently determine that the variant was called because of a sequencing or analysis artifact.

Results

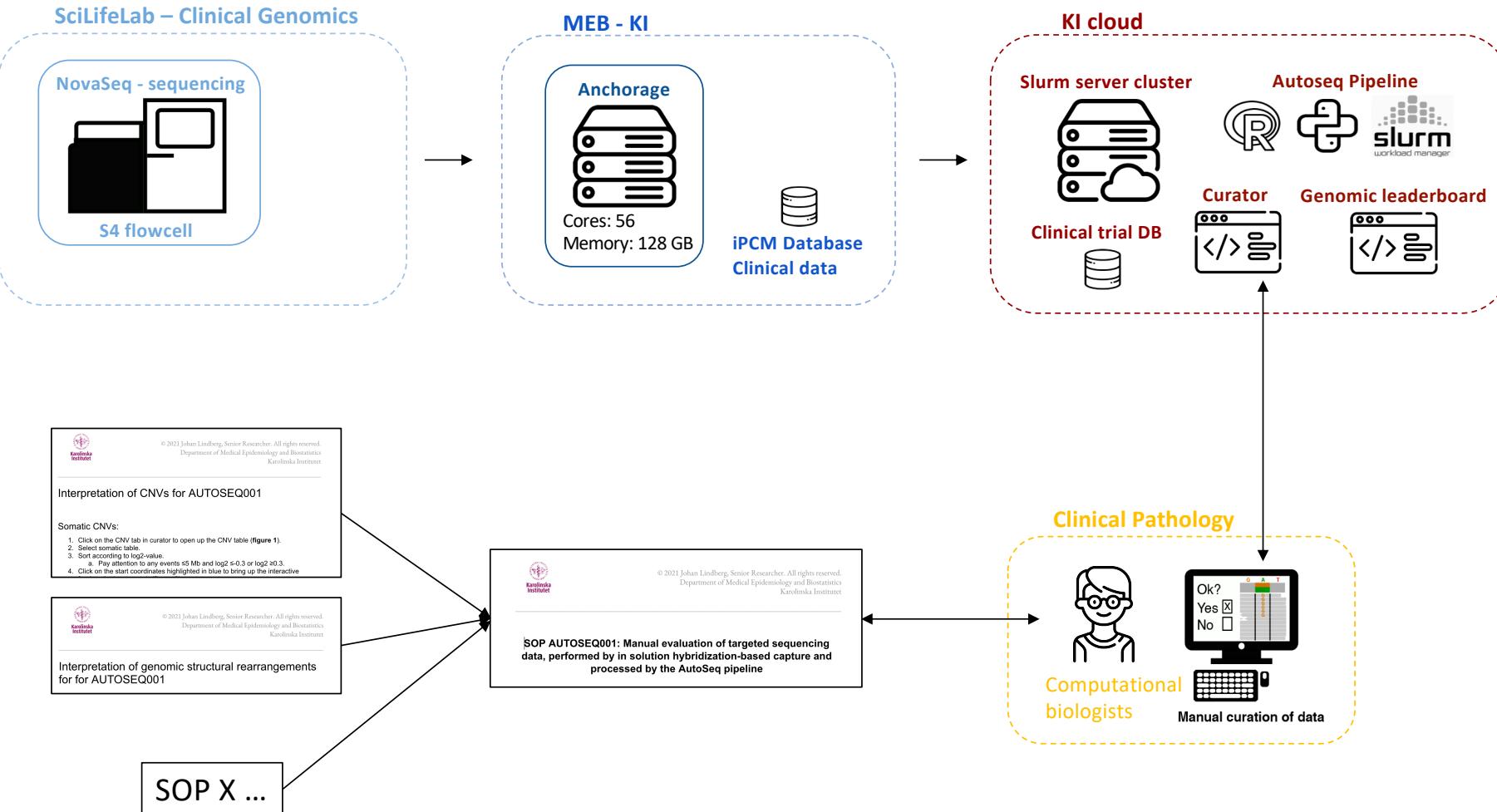


Warning ...

- Manual reviewers have reported reviewer fatigue, especially when evaluating tumors with a high variant burden ...
- Demo the local setup

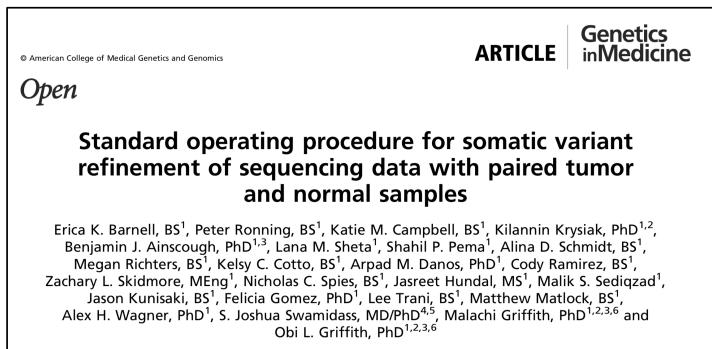


The iPCM study – data flow



- Demo: PB-P-00G02-CFDNA-03932575-KH-C3
 - SOP
 - Stratification based on tumor burden
 - SNV/Indel + tags
 - Homdel
 - GSR
 - IGV

Why manual when can be automated



Generate a lot of data

Machine learning →



Improve speed!

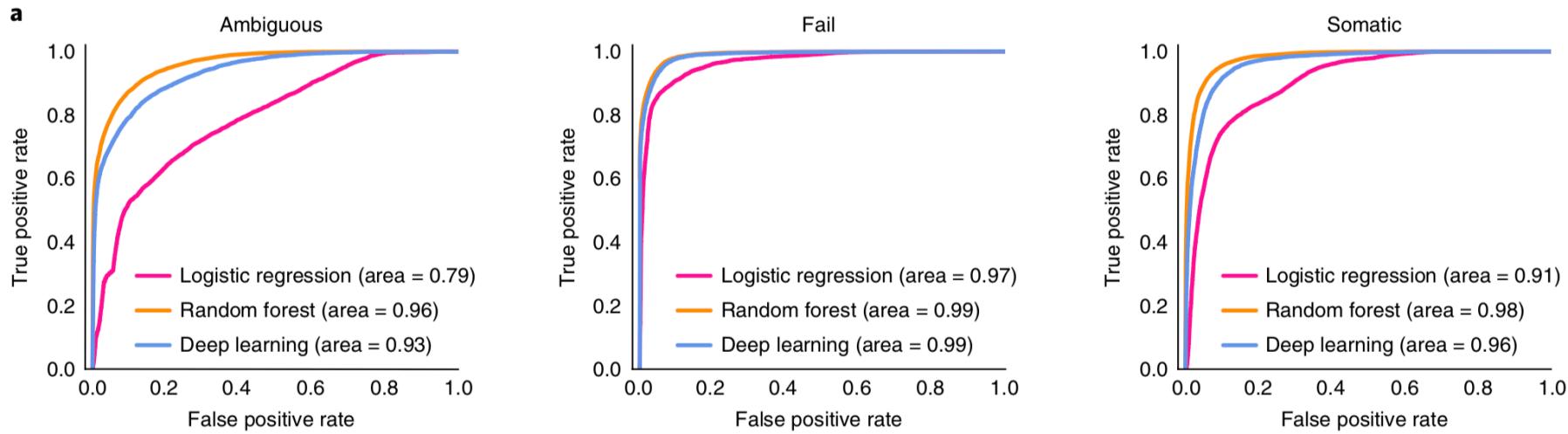
Results

- Potential of using machine learning to replace manual assessment
 - Training data set:
 - 41,000 variants from 21 studies
 - 440 cases derived from nine cancer subtypes.
 - Paired tumor and normal samples
 - Estimated manual effort: 585 hours
 - 71 features was assigned to train the model
 - For example: cancer type, sample type, tumor read depth, normal read depth, tumor VAF, normal VAF, base quality, mapping quality

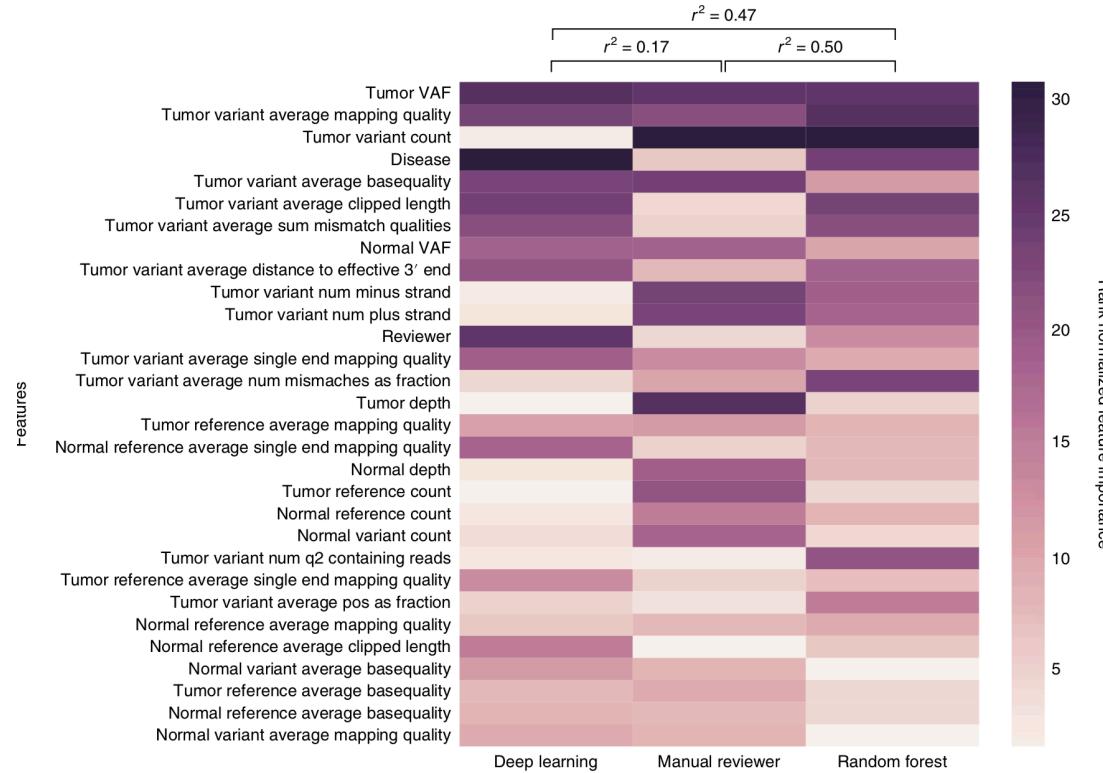
Results

- Manually reviewed variant calls
 - 18,381 were confirmed as somatic
 - 10,643 were assessed as ambiguous
 - 8,854 as failed
 - 3,122 as germline
- 2/3 used to train the model.
 - 10-fold cross-validation to check performance
 - 1/3 held-out dataset behaved as expected

Good enough?

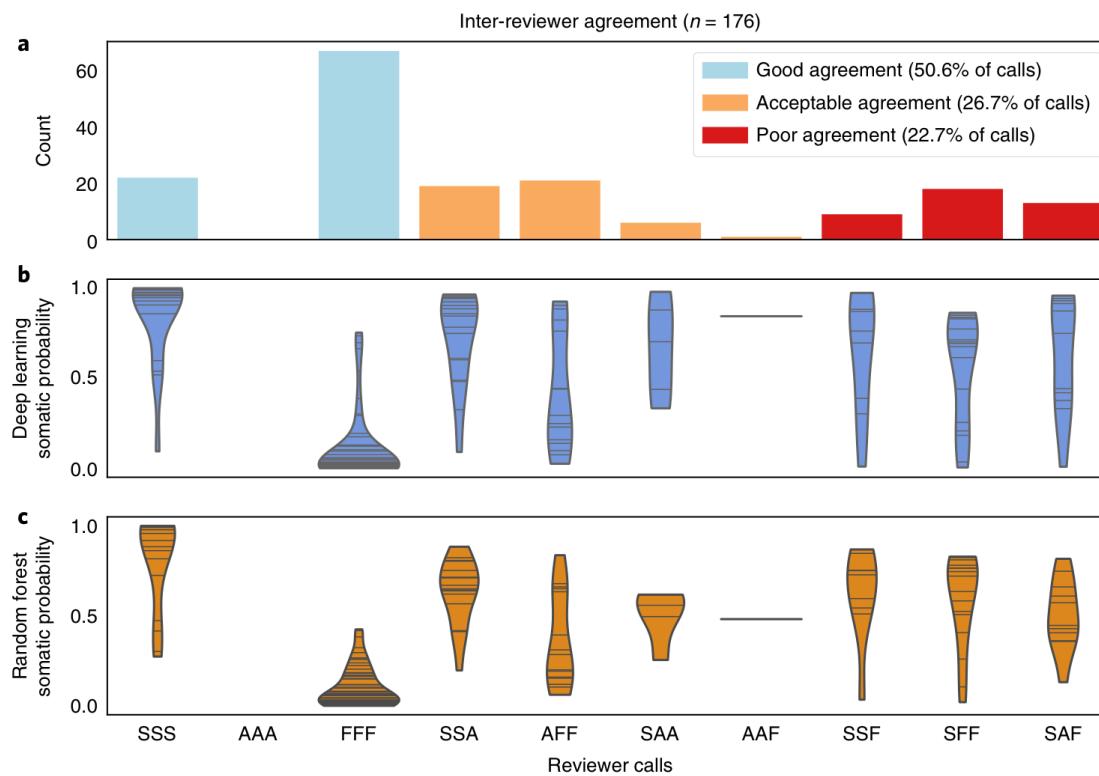


Important features



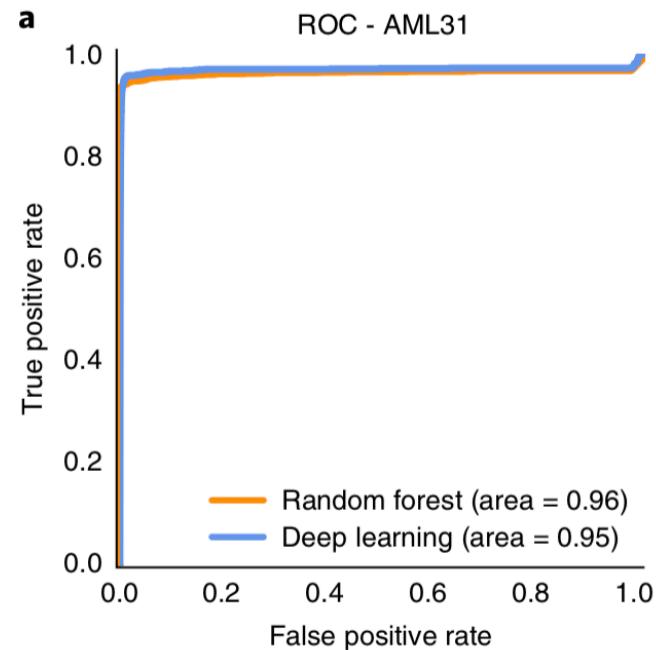
- Place for semi automated curation?

When reviewers agree - performance is good



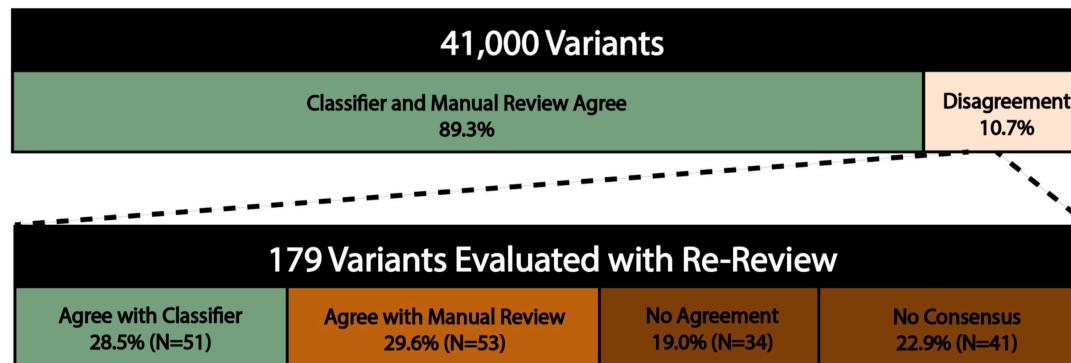
Validation

- When the training data + predicton data is the same = good performance
 - Otherwise re-traning is needed
- Independent validation of an AML case.
 - 192,241 putative somatic variants
 - All variants validated using targeted sequencing



Discordant variants

- 89.3% (35,622/41,000) calls are concordant between manual & classifier calls.
 - Unbiased manual re-review was performed on 179 discordant variants.
 - Seven individuals proficient in manual review re-reviewed IGV snapshots of the 179 variants and a consensus call was used.
- Authors conclude that there exist variants that are hard to classify even among experts ...



ORIGINAL ARTICLE

Olaparib for Metastatic Castration-Resistant Prostate Cancer

J. de Bono, J. Mateo, K. Fizazi, F. Saad, N. Shore, S. Sandhu, K.N. Chi, O. Sartor, N. Agarwal, D. Olmos, A. Thiery-Vuillemin, P. Twardowski, N. Mehra, C. Goessl, J. Kang, J. Burgents, W. Wu, A. Kohlmann, C.A. Adelman, and M. Hussain

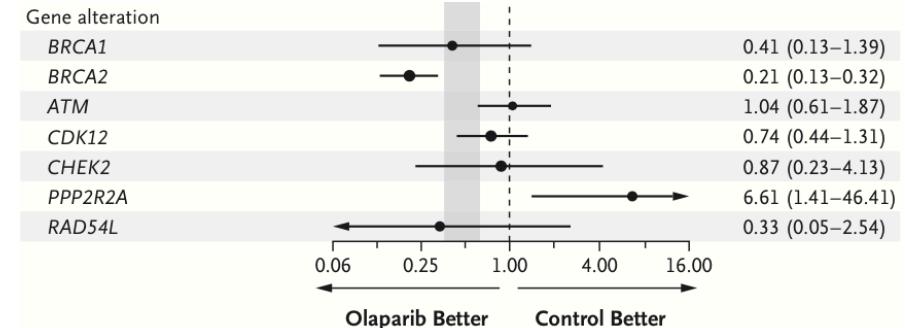
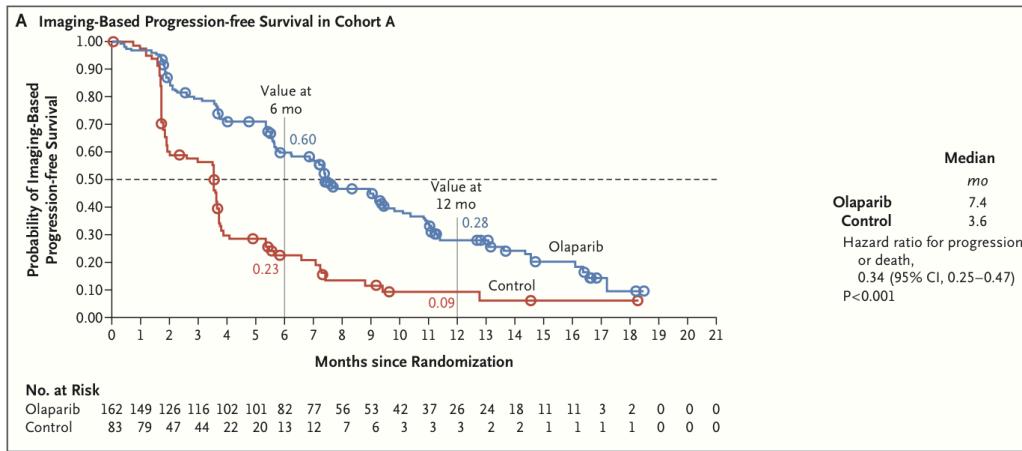
The Profound trial ..

Listen to the podcast of Vinay Prasad
<https://soundcloud.com/plenarysession/ep252>

Profound ...

- Phase 3 trial
- PARP in olaparib in men with metastatic castration-resistant prostate cancer
- Prerequisite: progression on enzalutamide or abiraterone
- Cohort A (245 patients) had at least one alteration in BRCA1, BRCA2, or ATM
- Cohort B (142 patients) had alterations in any of 12 other prespecified genes, prospectively and centrally determined from tumor tissue.
- Patients were randomly assigned (in a 2:1 ratio) to receive olaparib or the physician's choice of enzalutamide or abiraterone (control).
 - Not cabazitaxel which has a proven effect in 2nd line mCRPC!
 - Abi -> enza has ~5% response rate
 - Enza -> abi has no responses

Results



- Led to approval in US for ATM/BRCA1/BRCA2
- 3 other studies only show relevant responses in genes in BRCA2 complex

Joint curation!