

Artificial Intelligence and its Use in the Clinic

An analysis of adherence to best practices by AI models in the last 10 years

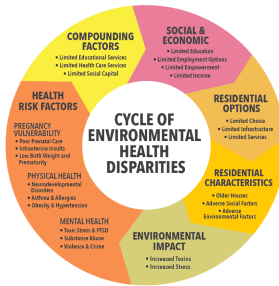
John Snell



Improved Outcomes



Health Disparities



Improved Value



ARTIFICIAL
INTELLIGENCE

Broad Utility



Improved Transparency



But...

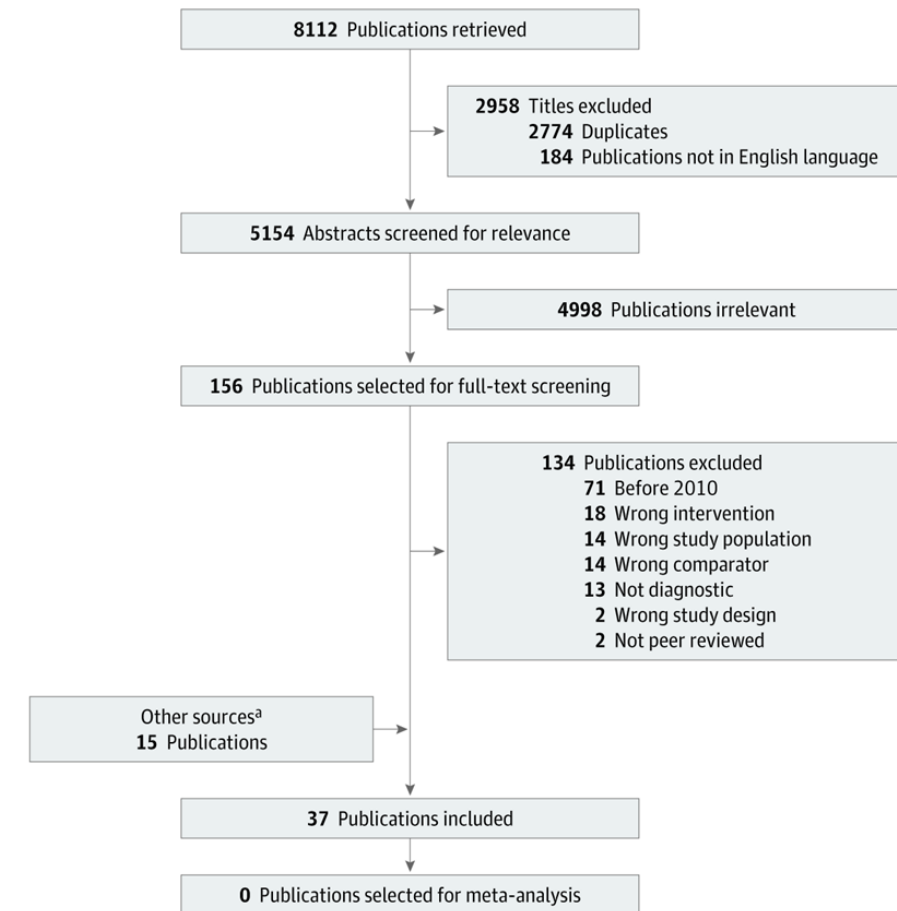
Is this true?

Core Principles

1. AI tools should aim to alleviate existing health disparities
2. Outcomes of AI tools should be clinically meaningful
3. AI tools should aim to reduce overdiagnosis and overtreatment
4. AI tools should aspire to have high healthcare value and avoid diverting resources from higher priority areas
5. AI tools should consider the biographical drivers of health
6. AI tools should be designed to be easily tailored to the local population
7. AI tools should promote a learning healthcare system
8. AI tools should facilitate shared decision-making

Data Collection

- 42 Papers were analyzed from the years 2010 through 2022
- Initial data collection based on Vasey et al. 2021 (37 papers)
 - Their metric of interest was association between clinician diagnostic performance and use of machine-learning based clinical diagnostic systems
 - They found no robust evidence of using algorithms to support rather than replace clinicians
- 5 additional papers of interest were added to diversify results

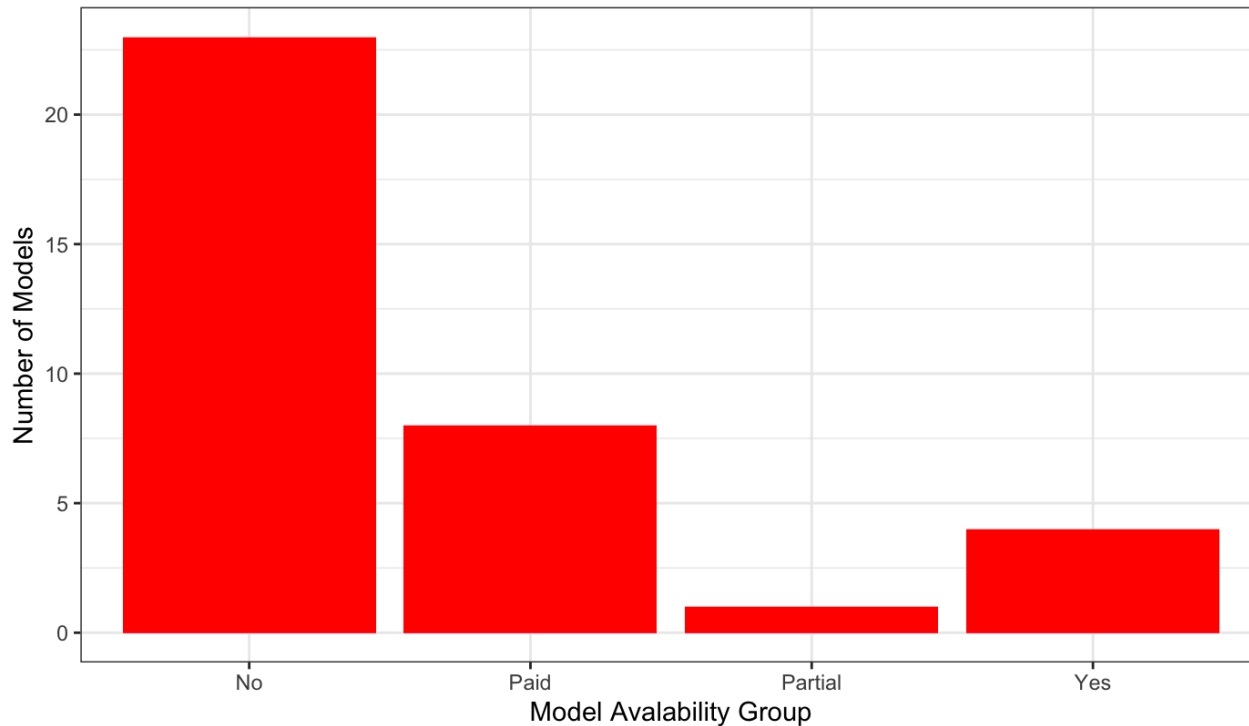


Vasey et al. 2021 Figure 1

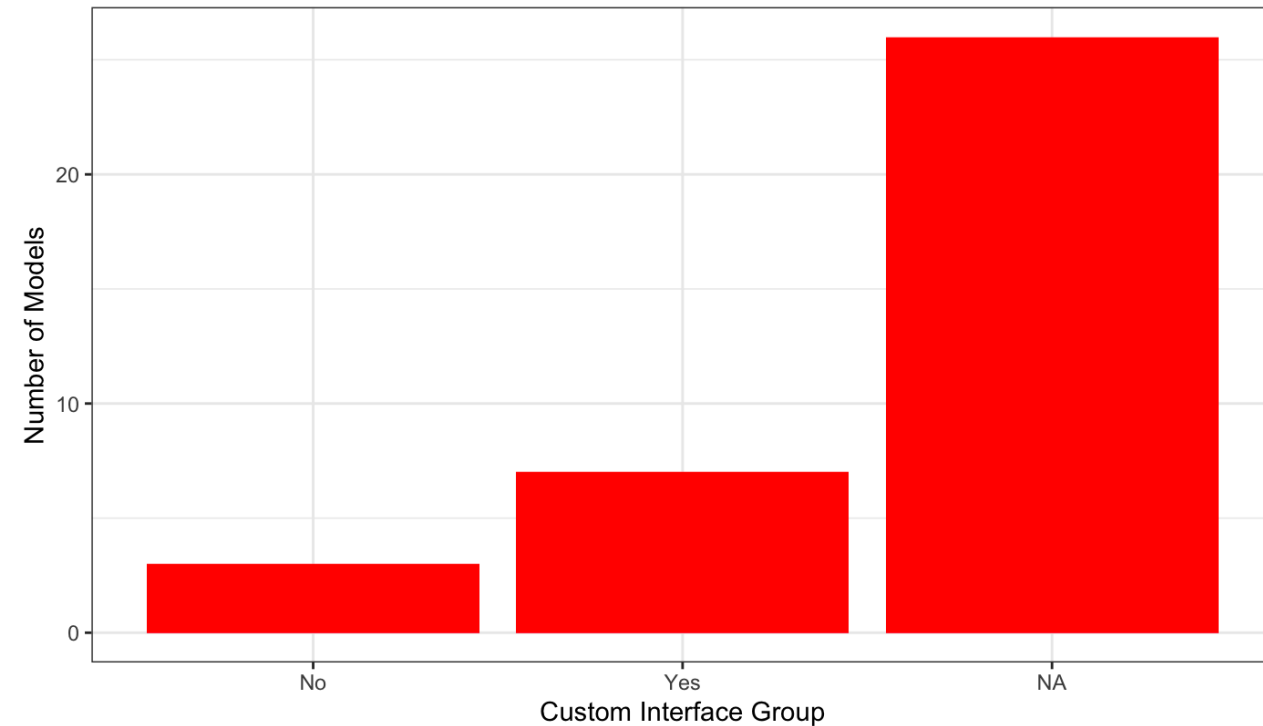
Principle 1: AI tools should aim to alleviate existing health disparities

Can disadvantaged groups equally access and benefit from the AI model?
Can less specialized physicians use the model in smaller settings?

Model Availability



Custom Interface Index



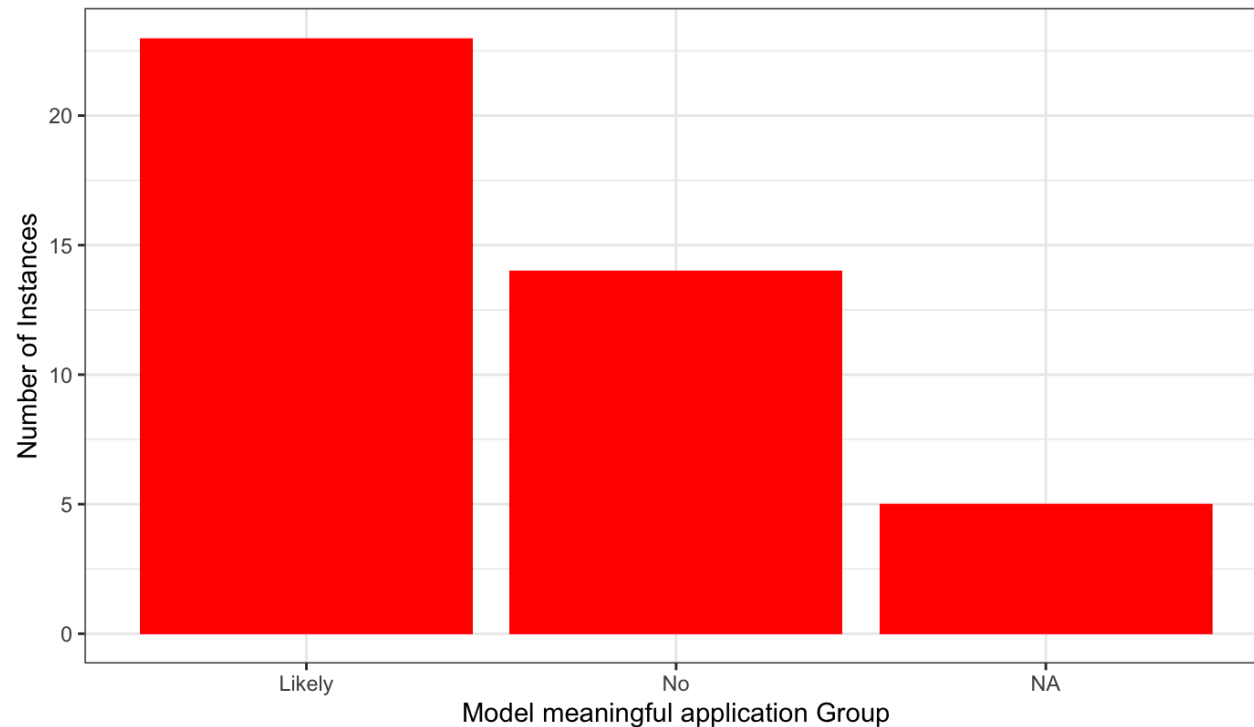
Principle 1: AI tools should aim to alleviate existing health disparities

Public Availability	Custom Interface			
	No	Yes	NA	
	No	0	3	20
	Paid	0	3	5
	Partial	0	0	1
Yes	3	1	0	

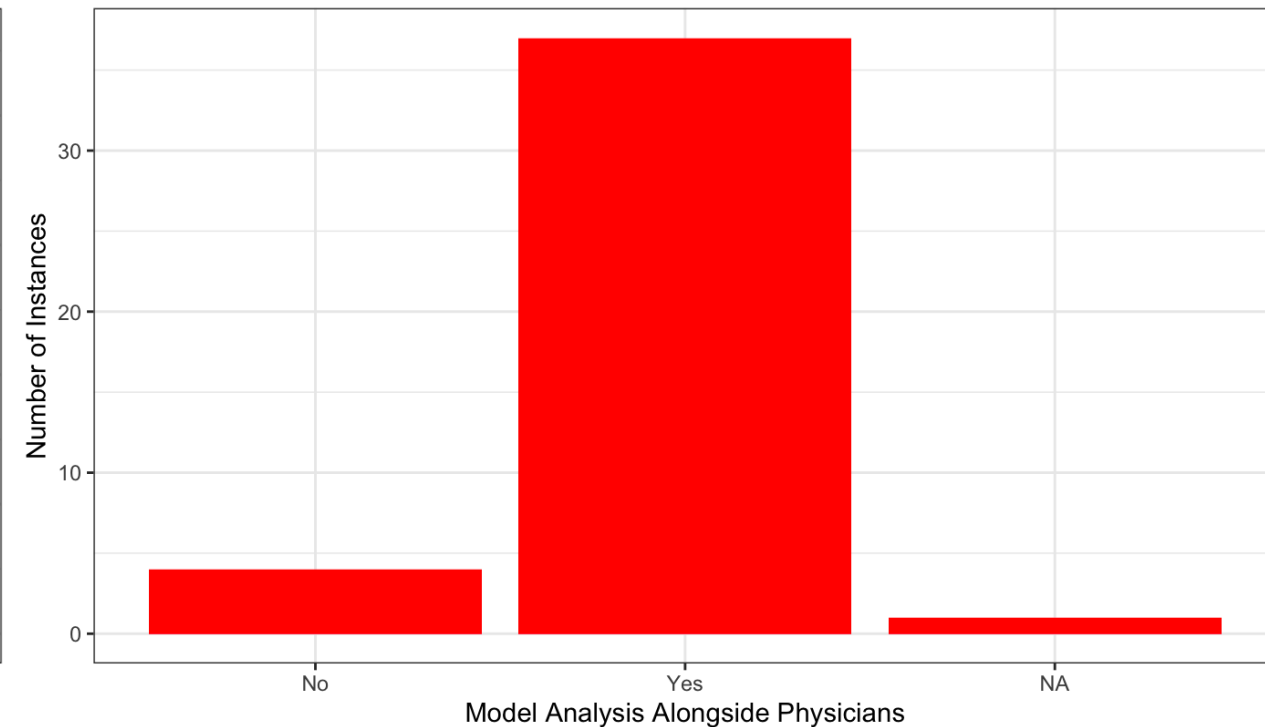
Principle 2: Outcomes of AI tools should be clinically meaningful

Does model perform in a meaningful manner in any capacity clinically? Has it been tested against clinicians?

Model Meaningfulness



Physician Benchmarking



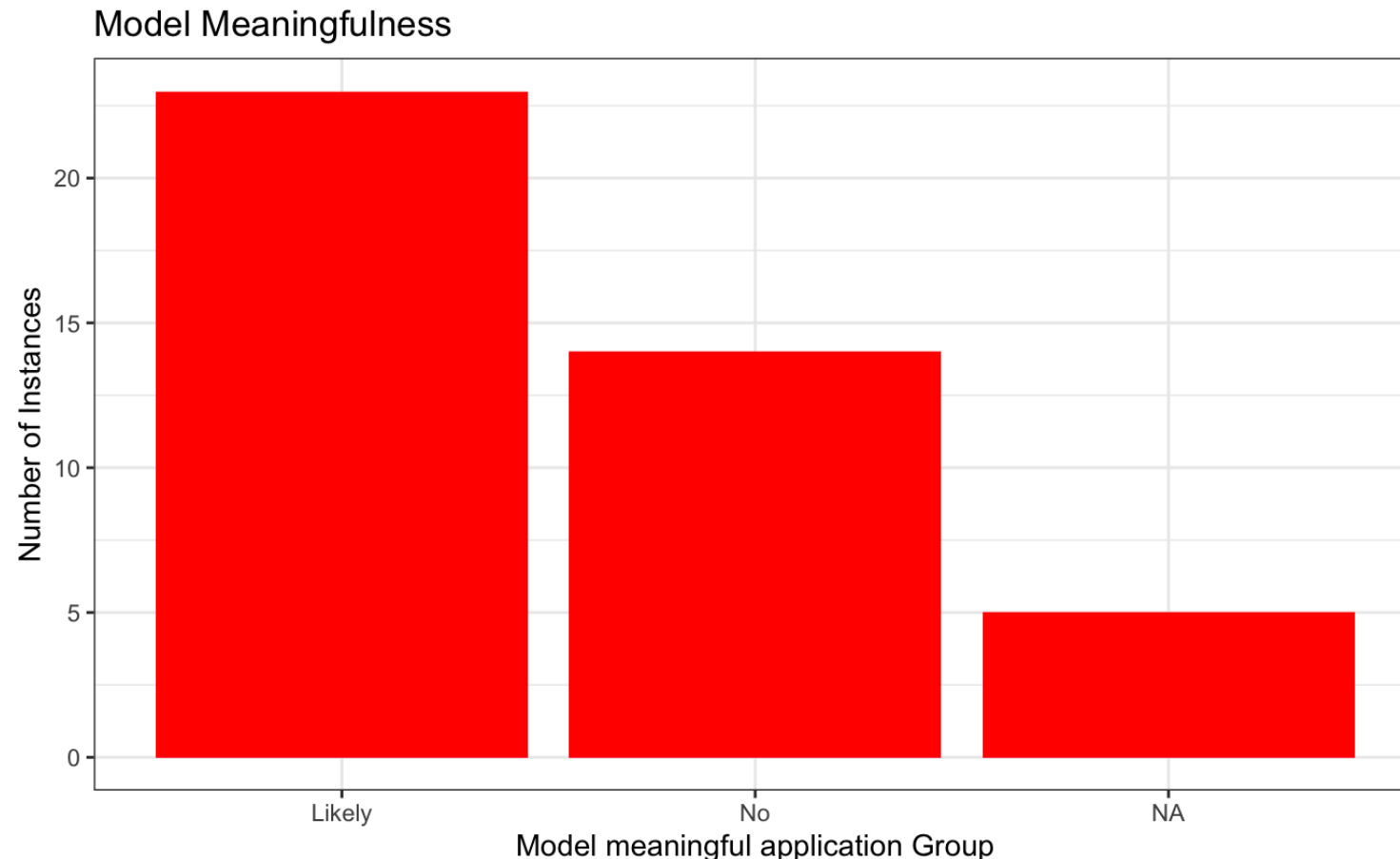
Principle 2: Outcomes of AI tools should be clinically meaningful

		Benchmarked Against Physicians		
Model Is Likely to be Meaningful		No	Yes	NA
	Likely	1	22	0
	No	0	14	0
	NA	3	1	1

Principle 3: AI tools should aim to reduce overdiagnosis and overtreatment

Does the model limit incorrect diagnoses in patients, limiting improper treatment?

This metric is adjacent to this goal. Tested models were measured for efficacy, not generally clinically applied.



Principle 4: AI tools should aspire to have high healthcare value and avoid diverting resources from higher priority areas

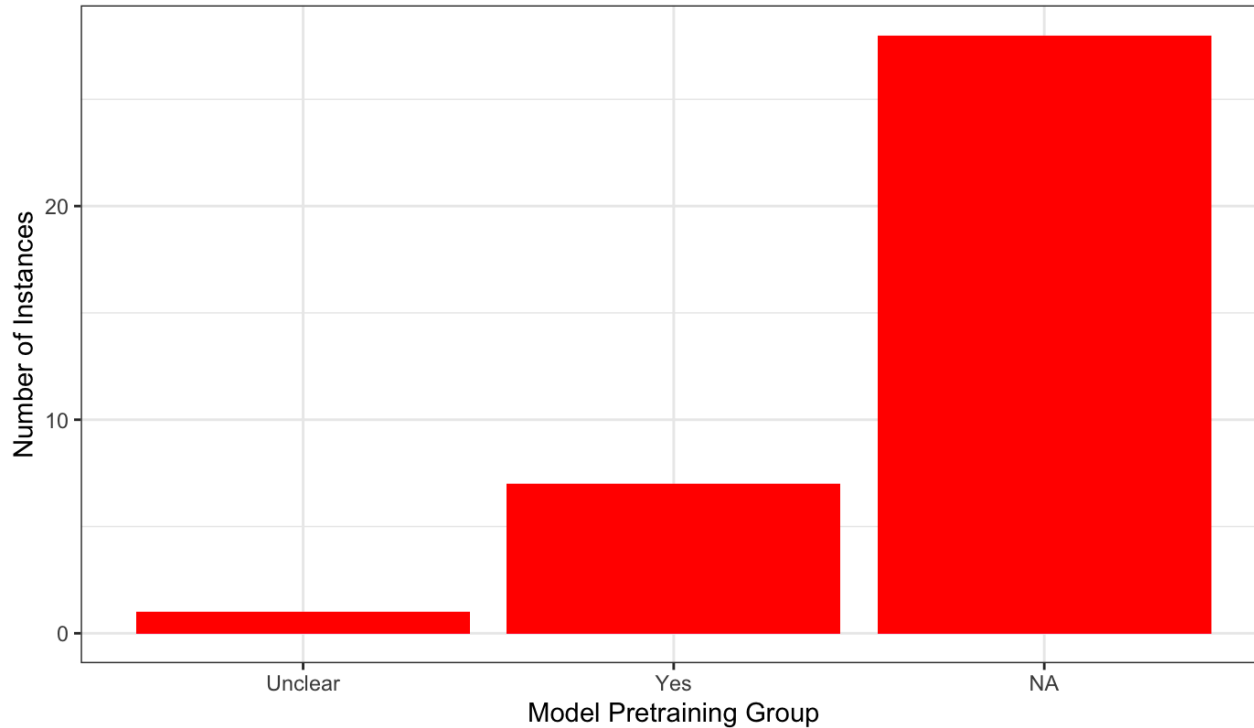
Does the model have a high value per dollar ratio?

This metric is not able to well characterized through the method of analysis performed here.

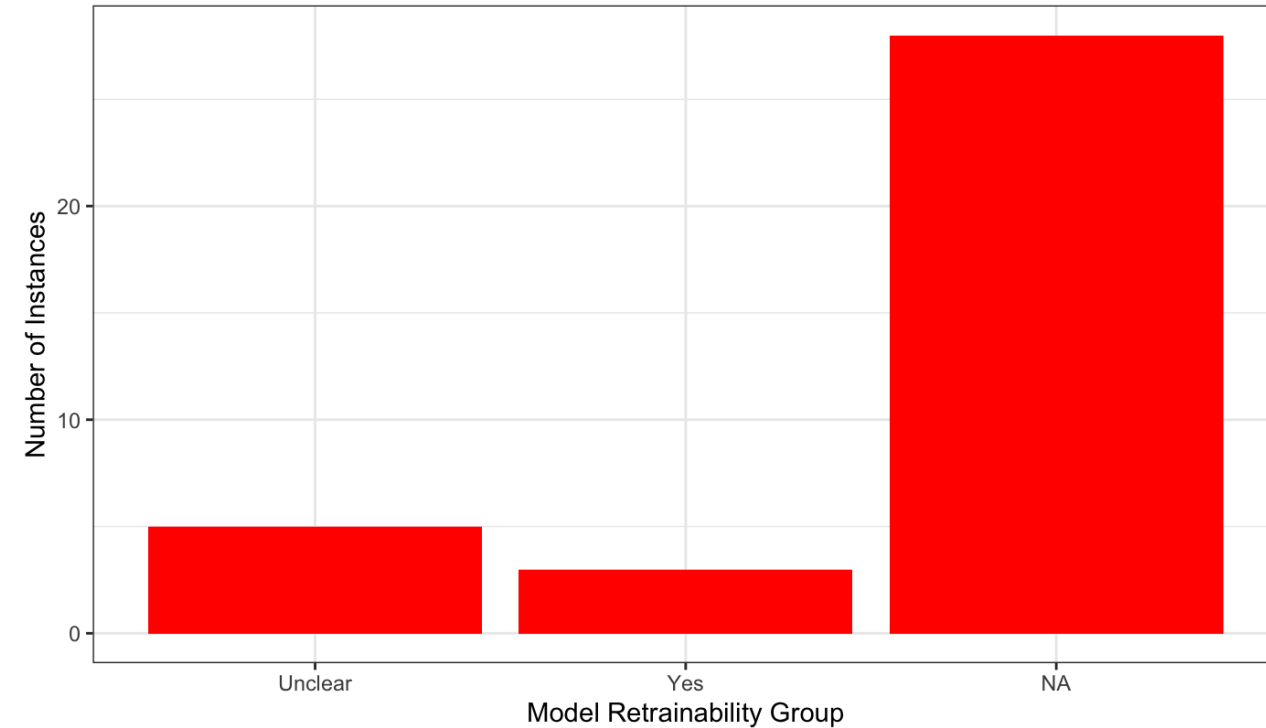
Principle 5: AI tools should consider the biographical drivers of health

Are models either pretrained or retrainable to include social/genetic determinants of health?

Model Pretraining Status



Model Retraining Status



Principle 5: AI tools should consider the biographical drivers of health

Are models either pretrained or retrainable to include social/genetic determinants of health?

Model Pretraining Status	Model Retrainingability		
	Unclear	Yes	NA
Unclear	1	0	0
Yes	4	3	0
NA	0	0	28

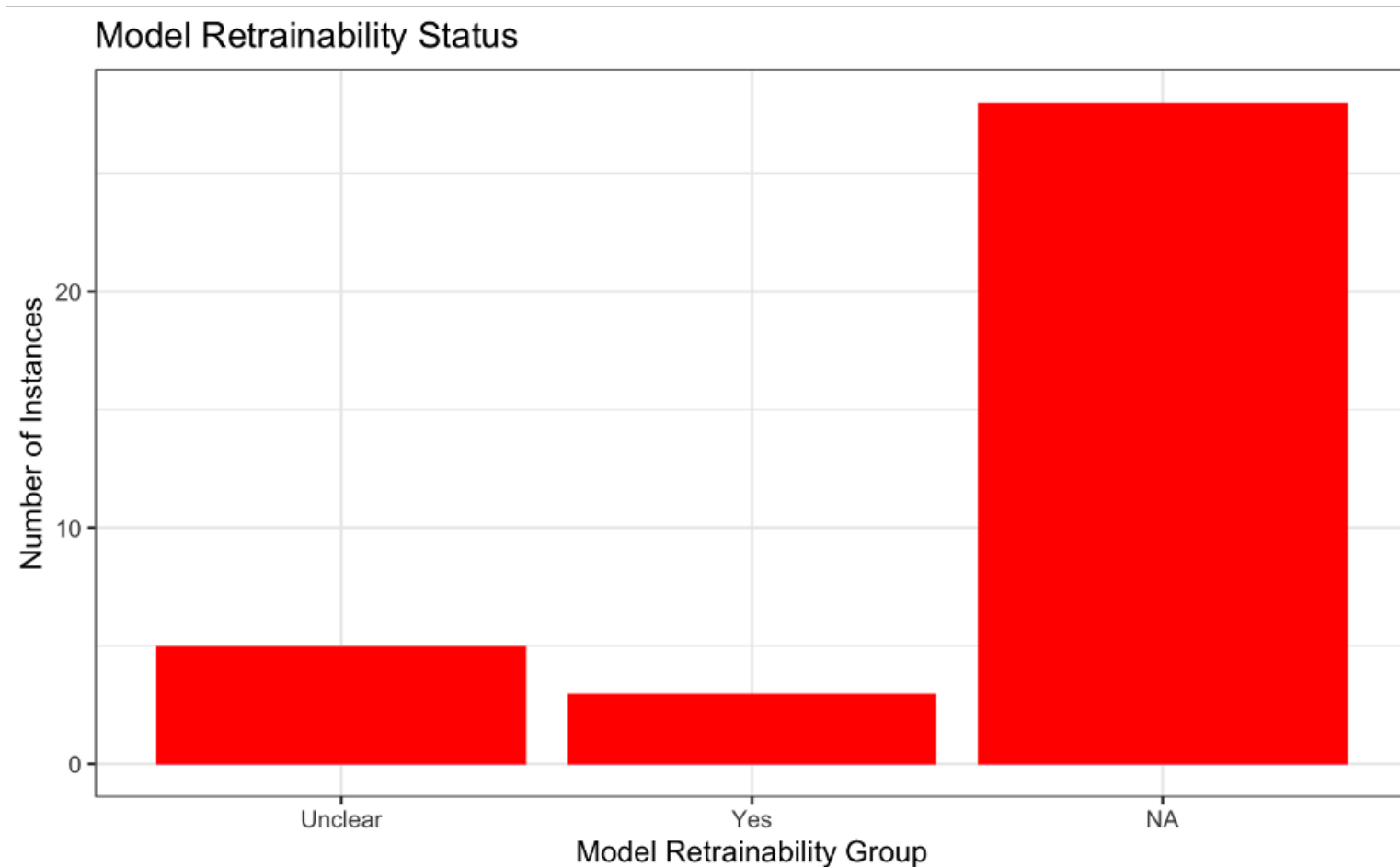
Principle 5: AI tools should consider the biographical drivers of health

Are models either pretrained or retrainable to include social/genetic determinants of health?

		Model Retrainingability		
		Unclear	Yes	NA
Model Availability	No	0	0	23
	Paid	4	0	4
	Partial	0	0	1
	Yes	1	3	0

Principle 6: AI tools should be designed to be easily tailored to the local population

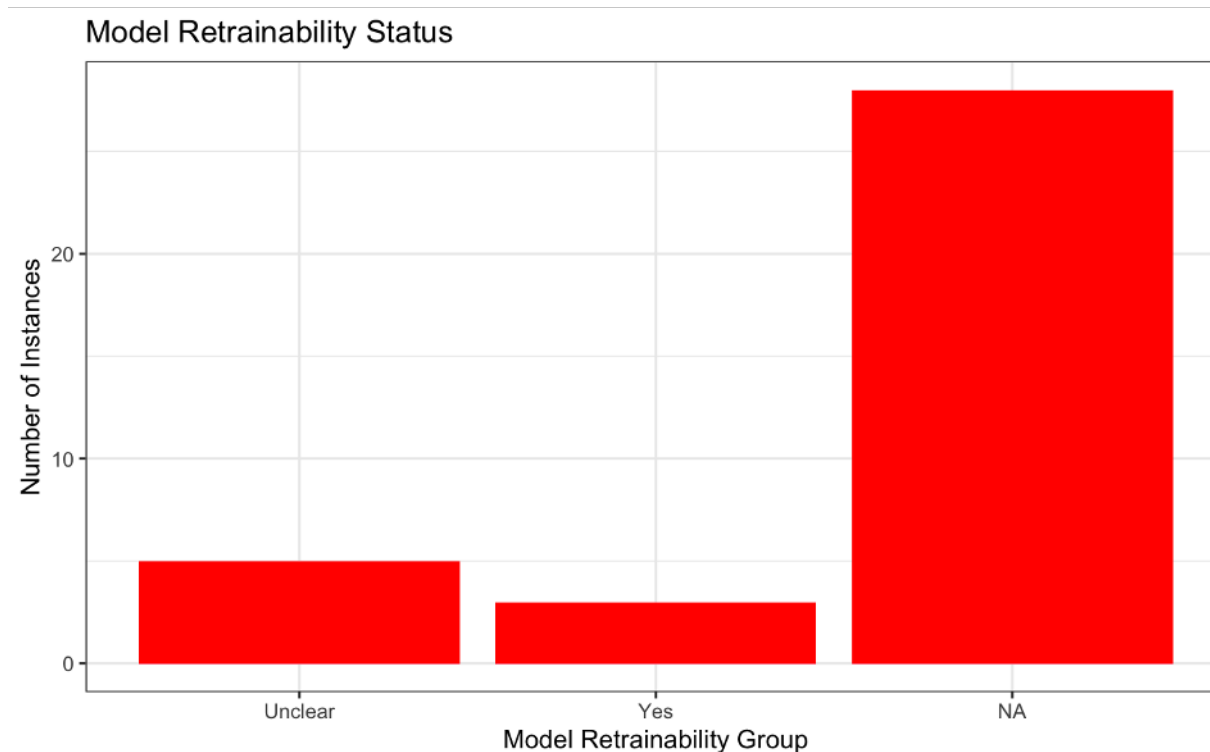
Can the model be tailored to a physician's local populations?



Principle 7: AI tools should promote a learning healthcare system

Is the model set up in such a way that it can be regularly updated as new data comes in so as to improve learning by both the model and physicians?

This metric is not able to well characterized through the method of analysis performed here.



Principle 8: AI tools should facilitate shared decision-making

Is the model set up in such a way that it facilitates shared decision making between physicians and patients?

This metric is not able to well characterized through the method of analysis performed here.

Overall Adherence to Best Practices

Principle	Adherence
Principle 1	19.4%
Principle 2	42.9%
Principle 3	Up to 54.8%
Principle 4	Not Characterized
Principle 5	8.3%
Principle 6	Up to 22.2%
Principle 7	Not Characterized *** (Up to 8.3%)
Principle 8	Not Characterized

Overall Adherence to Best Practices

Principle	Adherence
Principle 1	19.4%
Principle 2	42.9%
Principle 3	Up to 54.8%
Principle 4	Not Characterized
Principle 5	8.3%
Principle 6	Up to 22.2%
Principle 7	Not Characterized *** (Up to 8.3%)
Principle 8	Not Characterized

Steps For Improvement

- Focus on How to Alleviate Health Disparities
- Be presented alongside proposed value
- Reduce 'Black Box' nature of current models
- Improve ability to retrain/augment existing models
- Make non-paid models from research publicly available!!!

Conclusions

Questions?