Generalizable Deep Learning Model for Early Alzheimer's Disease Detection from Structural MRIs

Narges Razavian

Assistant Professor
New York University Langone Health

The Power of NACC Data, May 2nd 2023

Connect

Narges Razavian – Assistant Professor @ NYU Langone

Narges.Razavian@nyulangone.org

NACC

UNIVERSITY of WASHINGTON

Dementia Disease Disparities

- 66% of AD patients are women
- Risk factors correlate with Race & Socio-economic status
- Black and Hispanic patients 30% and 40% less likely than White patients to be seen by neurologists
  - Lower education, low income, and being uninsured → lower neurologist visits
- Dementia screening instruments (MOCA, MMSE) & tools build on majority white healthy research cohorts
  - ADNI Alzheimer's Disease Neuroimaging Initiative (92% White)
  - ADNI excludes participants with comorbidities, even though the main risk factors are comorbidities
- NACC National Alzheimer's Coordination Center (83% White)

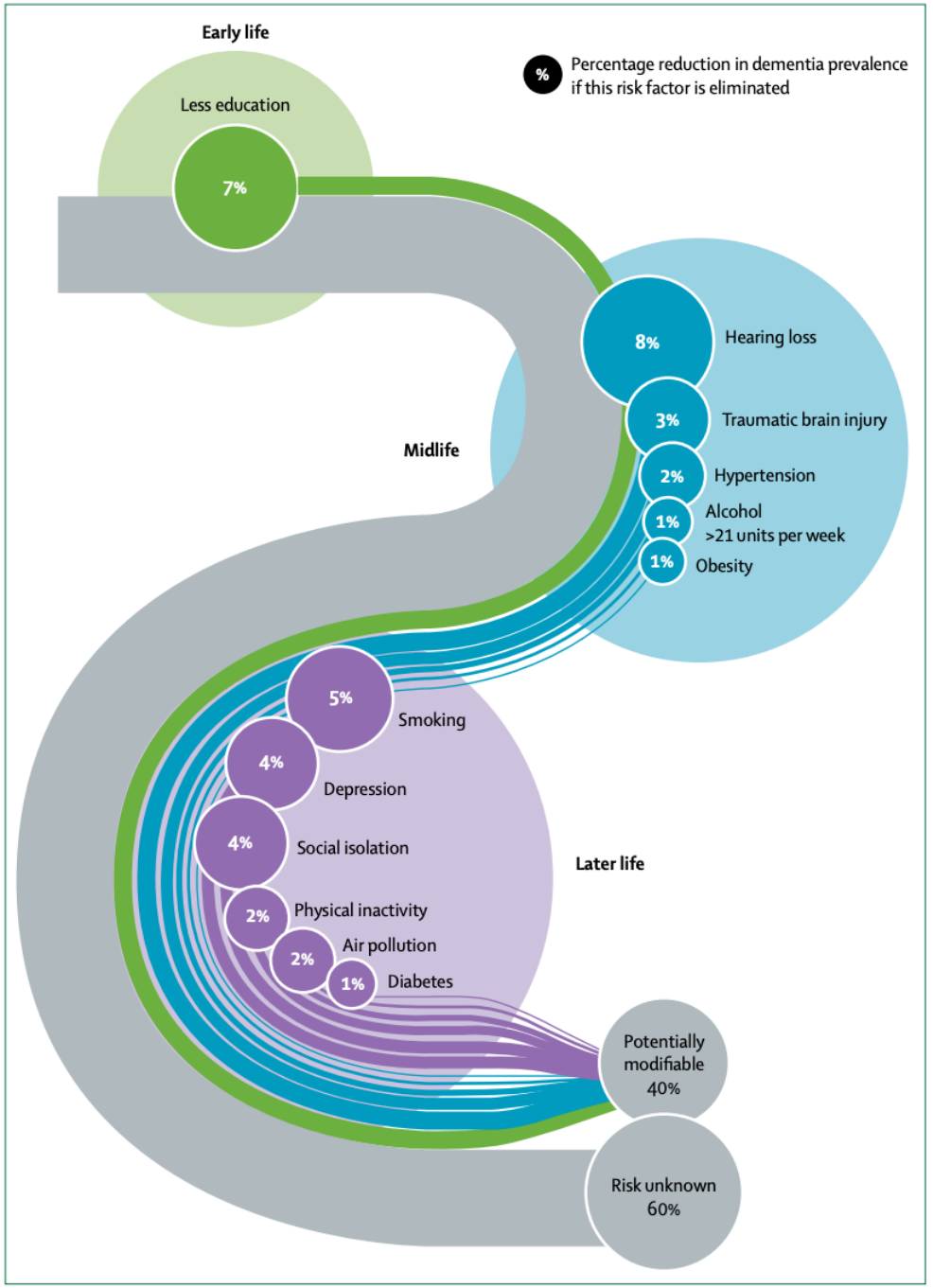


Figure 7: **Population attributable fraction of potentially modifiable risk factors for dementia**

Early Detection (in the real world) Matters

- All new clinical trials address “Mild to moderate AD”
- Early detection for preventative care
  - SPRINT MIND large scale randomized trial: Intensive hypertension control helps prevent conversion to MCI/AD
  - PREVENTABLE trial underway to study statins & cholesterol control
- Improved caregiver support & financial planning
- Better enrollment for clinical trials
- Real-world early detection may improve disparities in clinical trial enrolments too.
  - Aducanumab: Only 6 Black individuals were in the randomization arm with FDA approved dose. Only 3% identified as Hispanic, 0.03% (1 person) as American Indian, and of the 9% identified as Asian, 94% were recruited in Asia.

Dementia (ADRD) Patients at NYU Langone

Features	All Patients with Age≥65 NYU Langone (N=844,852) (%)	All Patients Age≥65, With diagnosis of AD/ADRD or MCI (N=40,035) (%)	All Patients with Age≥65 & AD/ADRD/ MCI & Referred to NYU Barlow (N=3,847)(%)
Age (mean, SD)	75.76 (8.4)	80.38 (8.1)	78.26 (7.4)
Female (N, %)	489441(57.9)	22360(55.8)	2045(53.1)
Asian	33163(3.9)	2165(5.4)	120(3.1)
Black	65859(7.8)	2724(6.8)	221(5.7)
White	486903(57.6)	27033(67.5)	2695(70.0)
Hispanic	41650(4.9)	2627(6.5)	164(4.2)
Hypertension	295761(35.0)	23744(59.3)	1768(45.9)
Diabetes	122496(14.5)	10558(26.3)	855(22.2)
Diabetes, complex	26900(3.1)	2952(7.3)	218(5.6)
Hyperlipidemia	224252(26.5)	17601(43.9)	1521(39.5)
On Statins	285692(33.8)	19087(47.6)	1932(50.2)
On Aspirin	216695(25.6)	16118(40.2)	1822(47.3)
On BP Meds	327697(38.7)	22768(56.8)	2003(52.0)
On any vascular med*	419073(49.6)	28487(71.1)	2859(74.3)
AD/ADRD Screening	14042(1.6)	7611(19.0)	2193(57.0)
Any Vascular Risk*	396079(46.8)	**29361(73.3)**	**2450(63.6)**
Among Patients with Vascular Risk:			
≥140 SBP or ≥90 DBP	297021(74.9)	23821(81.1)	1782(72.7)
On any vascular med	305155(77.0)	23881(81.3)	2068(84.4)

Need For Accessible And Accurate Methods For Early Detection

For Real World Clinics
With More Diversity and More Realistic Comorbidity Burden

Imaging Biomarkers for Real World Clinical Setting

- PET imaging with β-amyloid & Tau tracers → Not covered by insurance, expensive, different non-standardized tracers (tau), accessibility disparities

- Structural MRIs
  → Show atrophies
  → Historically using hippocampal volume (not accurate at MCI stage).

→ Can we use deep learning on 3D volumes to better identify?

→ Can we integrate these models into clinical settings and measure their impact?

→ Do the model eventually change patient outcome (i.e. rate of early detection)

Data - Publicly available large cohorts

Alzheimer's Disease Neuro-Imaging Initiative (ADNI)

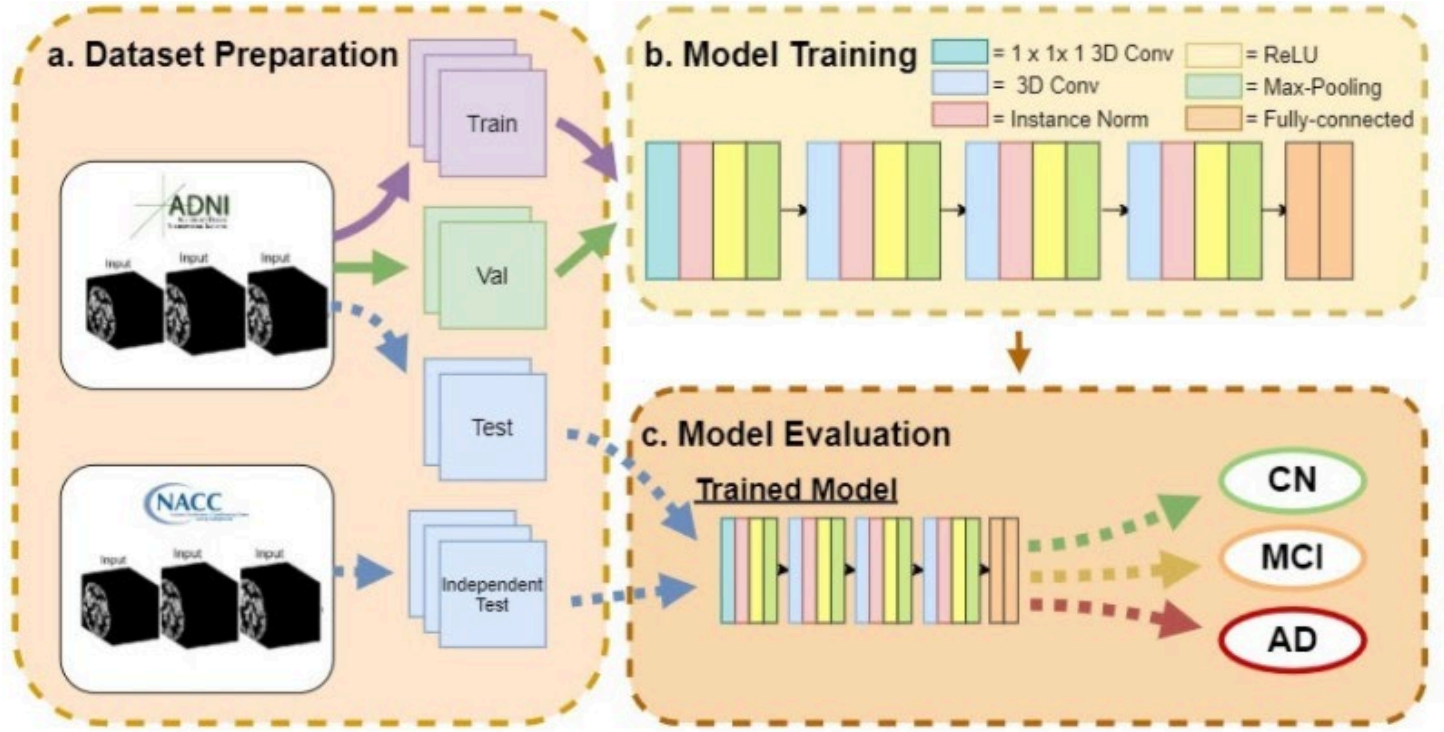
- 652 individuals with T1 MRIs
  - 2619 MRI scans
- Somewhat AD specific (later years)
- Healthy AD phenotype
- 92% White

National Alzheimer's Coordinating Center (NACC)

- 1522 individuals with T1 MRIs
  - 2045 MRI scans
- Allows comorbidity and other disease burdens (more realistic)
- 83% white (as of 2019)
- Mixed ADRD subtypes
  - Wei et.al *On gaps of clinical diagnosis of dementia subtypes: A study of Alzheimer's disease and Lewy body disease*, Front. Aging Neurosci., 2023
    - *Among participants diagnosed as AD in the clinic, over 32% had concurrent LBD neuropathology at autopsy. Among participants diagnosed as LBD, 32% to 54% revealed concurrent autopsy-confirmed AD pathology.*

Model Architecture

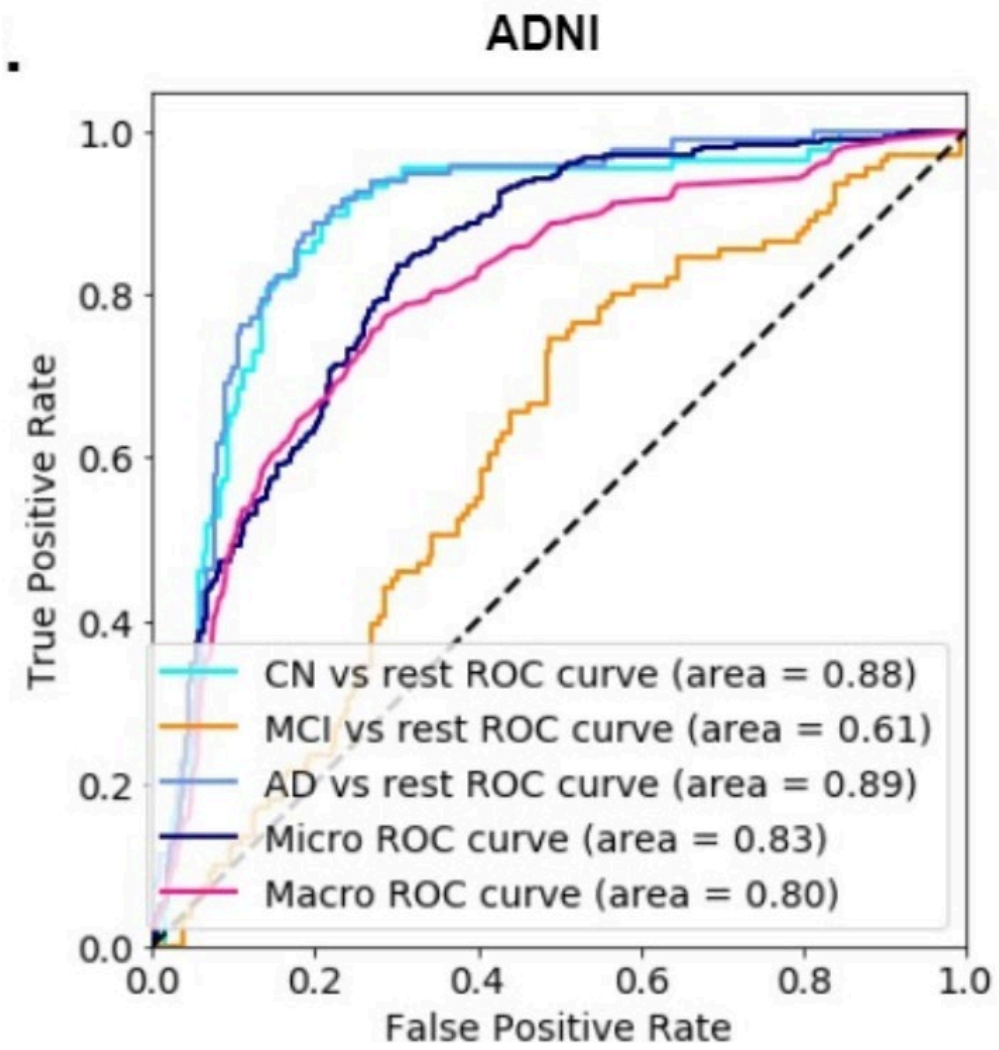
- Improved architecture via
  - Instance normalization outperforms Batch normalization
  - Less early spatial down-sampling
  - Widening the layers brings consistent gains while increasing the depth does not

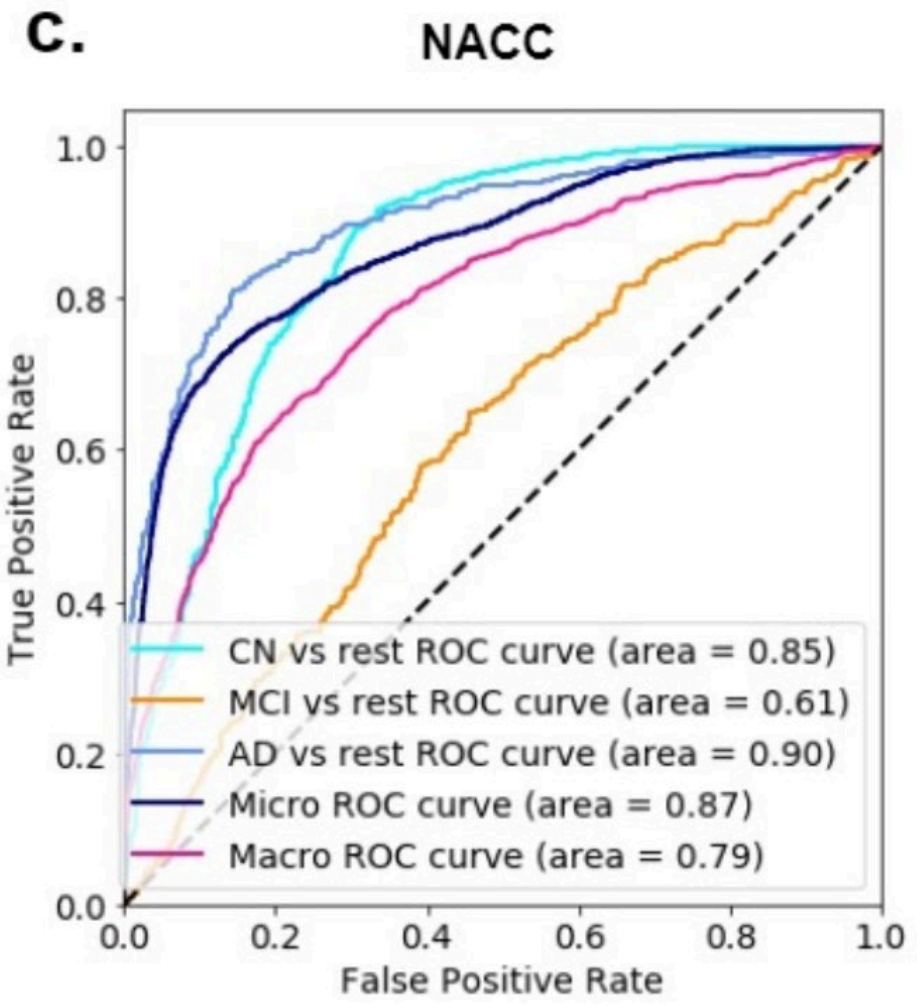


Characteristics Table

	ADNI (n=2619)			NACC (n=2025)		
Patient Characteristics	**Cognitively Normal (n =782)**	**Mild Cognitive Impairment (n=1089)**	**Alzheimer's Disease (n=748)**	**Cognitively Normal (n=1281)**	**Mild Cognitive Impairment (n = 322)**	**Alzheimer's Disease (n = 422)**
Age, mean (sd)	77.3 (5.6)	76.5 (7.3)	76.5 (7.3)	69.1 (9.4)* (p-val<0.01)	74.4 (8.5)* (p-val<0.01)	73.9 (8.8)* (p-val:<0.01)
Sex, n (%)						
Male	394 (50.4%)	659 (60.5%)	406 (54.3%)	489 (38.2%)* (p-val<0.01)	128 (39.8%)* (p-val<0.01)	219 (49.5%) (p-val:0.433)
Female	388 (49.6%)	430 (39.5%)	342 (45.7%)	792 (61.8%)* (p-val<0.01)	194 (60.2%)* (p-val<0.01)	223 (50.5%)* (p-val:0.02)
Education, avg years (sd)	17.2 (3.1)	16.7 (3.2)	16.1 (3.5)	16.3 (2.6)* (p-val<0.01)	15.7 (2.8)* (p-val<0.01)	15.1 (3.3)* (p-val<0.01)
APOE4, n (%)	224 (28.6%)	567 (52.1%)	496 (66.3%)	479 (37.4%)* (p-val<0.01)	146 (45.3%)* (p-val:0.03)	202 (45.7%)* (p-val<0.01)

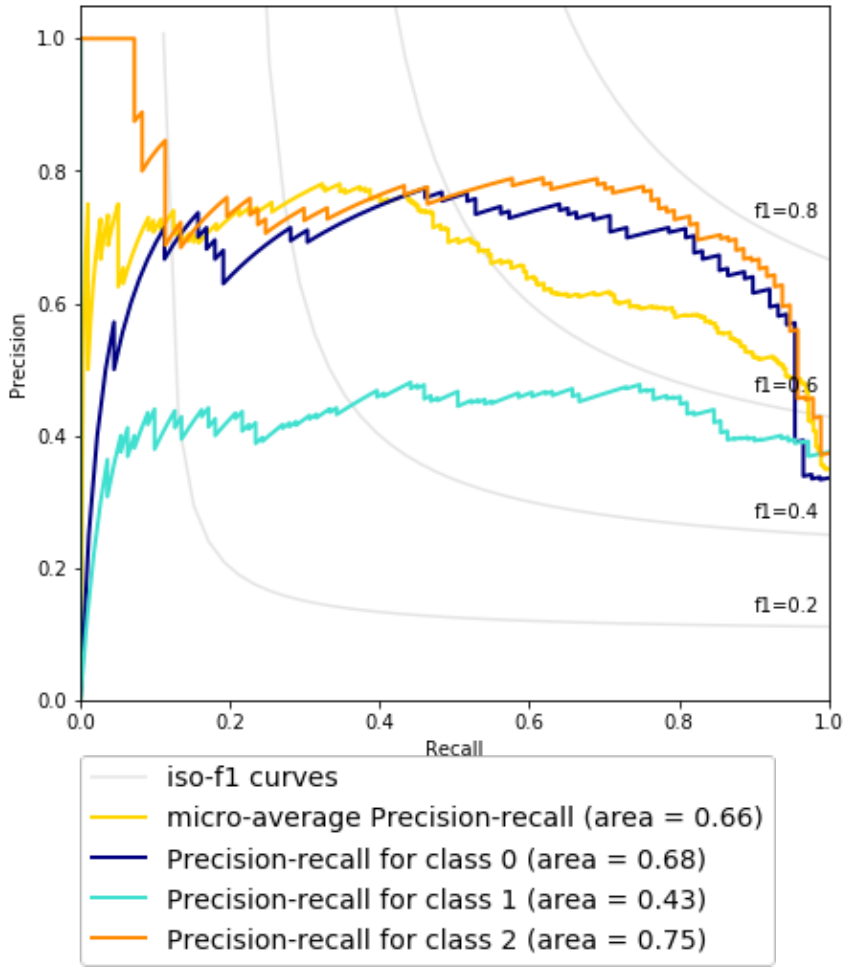
Results


ADNI
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
CN vs rest ROC curve (area = 0.88)
MCI vs rest ROC curve (area = 0.61)
AD vs rest ROC curve (area = 0.89)
Micro ROC curve (area = 0.83)
Macro ROC curve (area = 0.80)

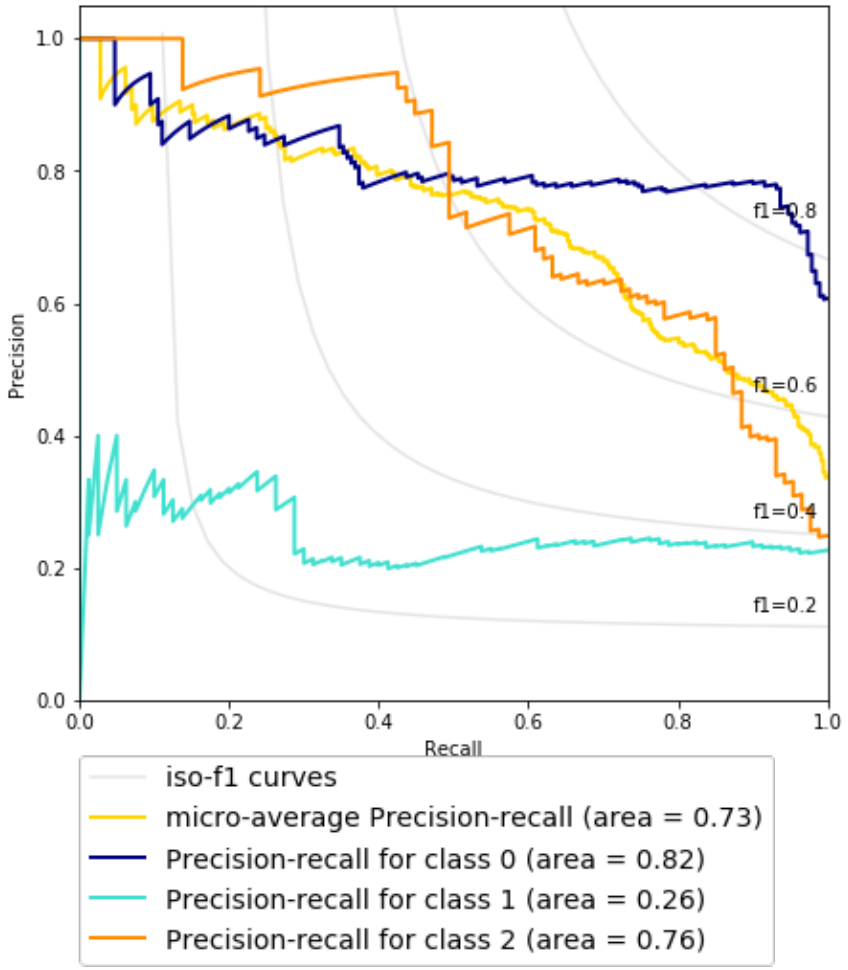

C.
NACC
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
CN vs rest ROC curve (area = 0.85)
MCI vs rest ROC curve (area = 0.61)
AD vs rest ROC curve (area = 0.90)
Micro ROC curve (area = 0.87)
Macro ROC curve (area = 0.79)

Precision/Recall Curves - Clinically More Actionable

ADNI Heldout Test Set



NACC (External Validation Data)



How does deep learning compare to Freesurfer based model?

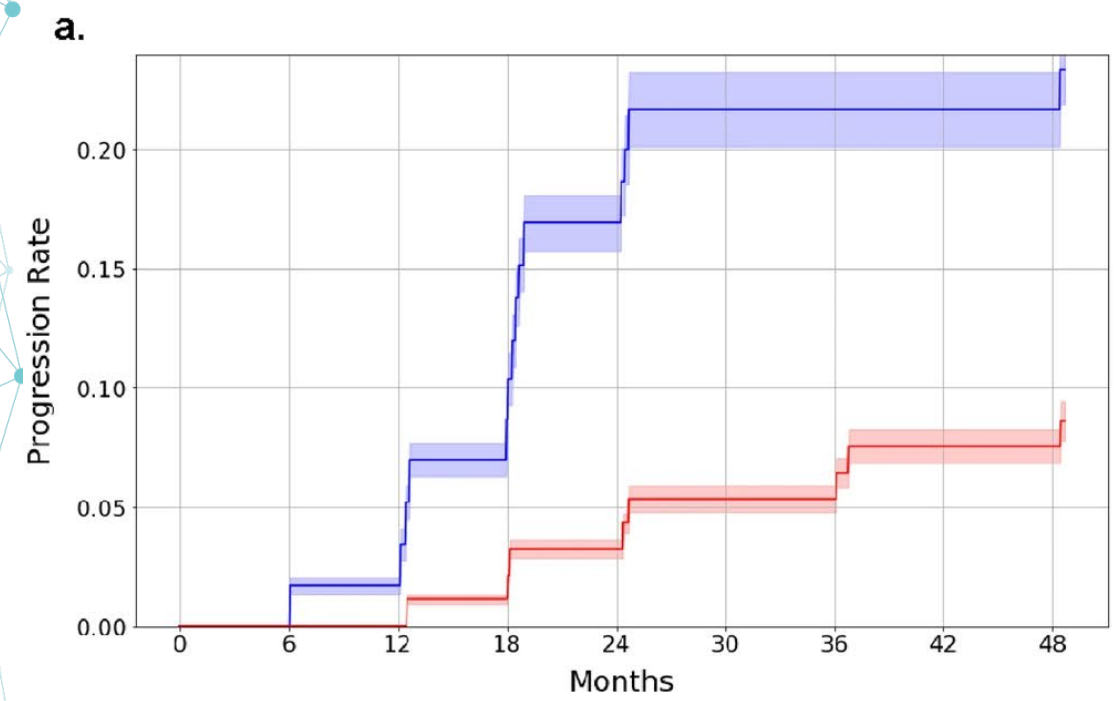
	ADNI Heldout (n=90 individuals, 297 scans)		NACC external validation (n=1522 individuals, 2025 scans)	
	Deep learning model Area under ROC curve	Freesurfer-based model Area under ROC curve	**Deep learning model Area under ROC curve**	Freesurfer-based model Area under ROC curve
Cognitively Normal	**87.59 (95% CI: 87.13 - 88.05)**	84.45 (95% CI: 84.19 - 84.71)	**85.12 (95% CI: 85.26 - 84.98)**	80.77 (95% CI: 80.55 - 80.99)
Mild Cognitive Impairment	**62.59 (95% CI: 62.01 - 63.17)**	56.95 (95% CI: 56.27 - 57.63)	**62.45 (95% CI: 62.82 - 62.08)**	57.88 (95% CI: 57.53 - 58.23)
Alzheimer's Disease Dementia	**89.21 (95% CI: 88.88 - 89.54)**	85.57 (95% CI: 85.16 - 85.98)	**89.21 (95% CI: 88.99 - 89.43)**	81.03 (95% CI: 80.84 - 81.21)

Freesurfer also takes **11 hours** per MRI vs. Deep learning model that takes **7.8mins** (7 min of pre-processing, 0.07s of the model running)

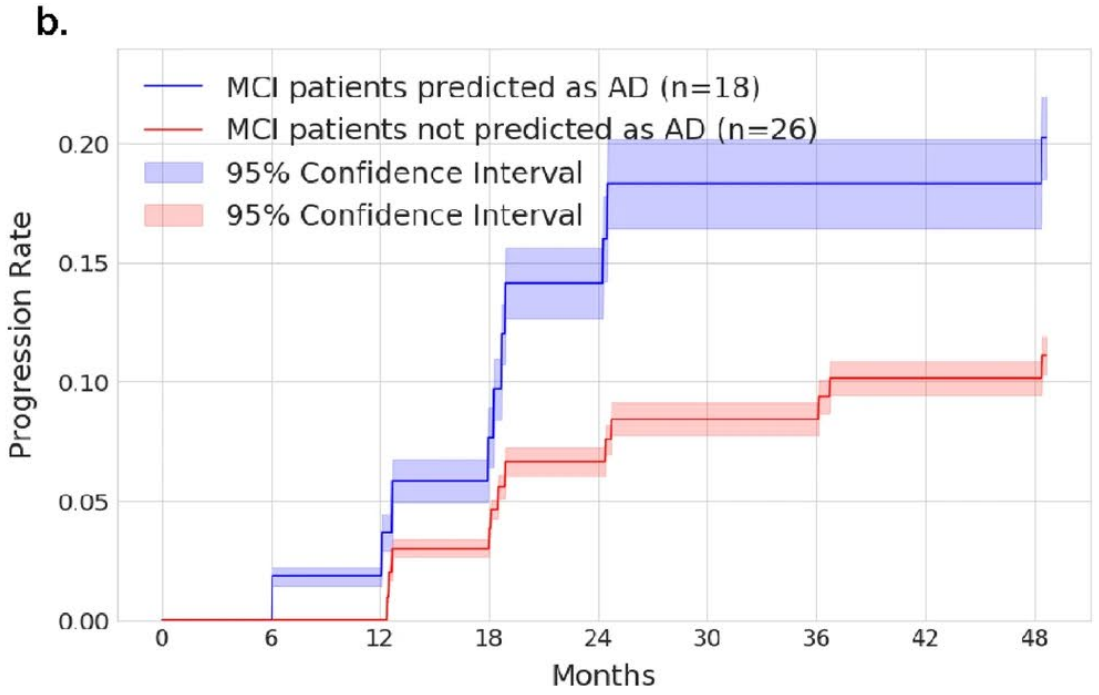
Progression to Dementia

*For MCI patients **predicted as AD** vs **not AD***

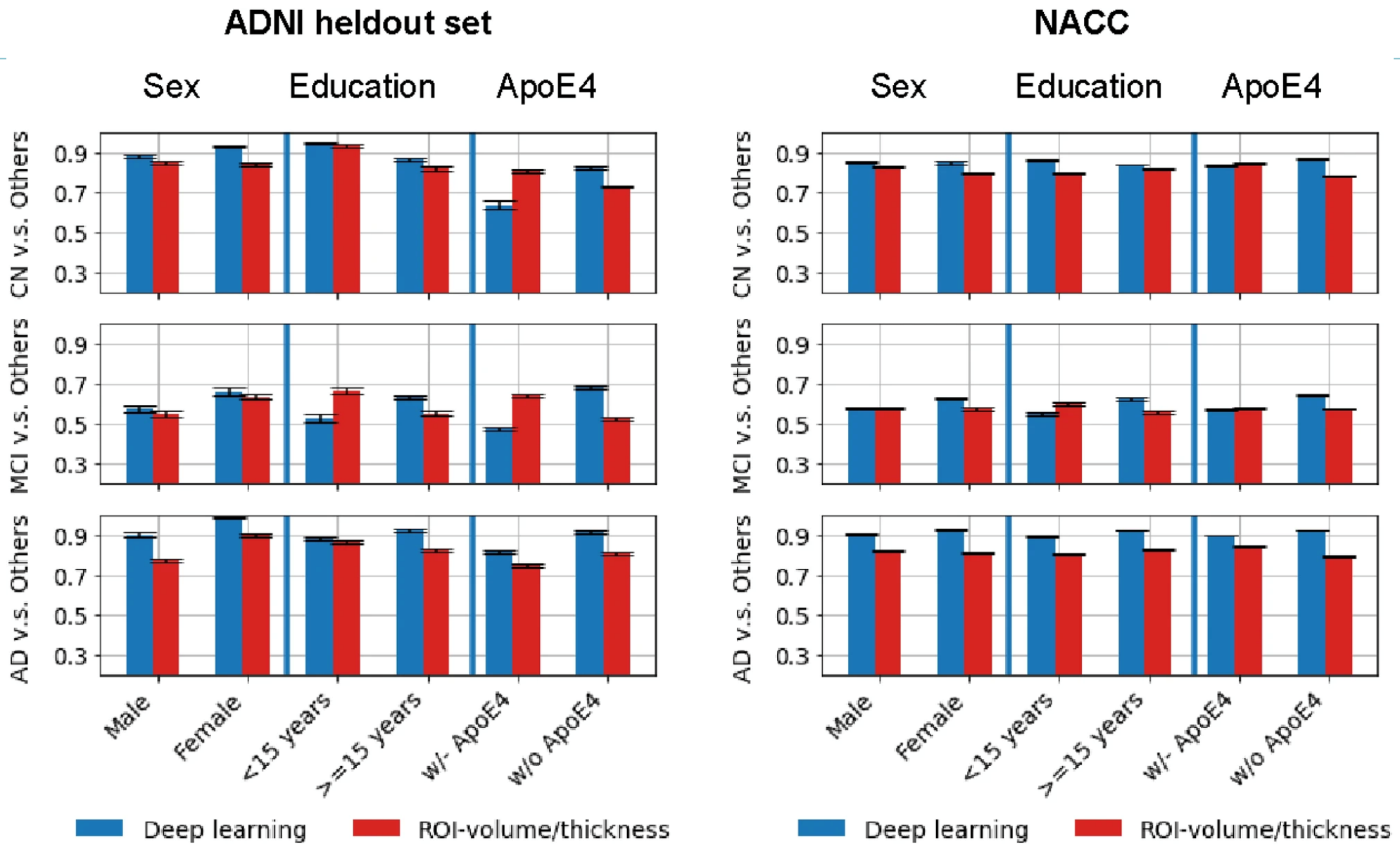
23.02% of subjects in group A (blue line) progress to AD, compared to 8.81% of subjects in group B (red line).



Deep Learning Model



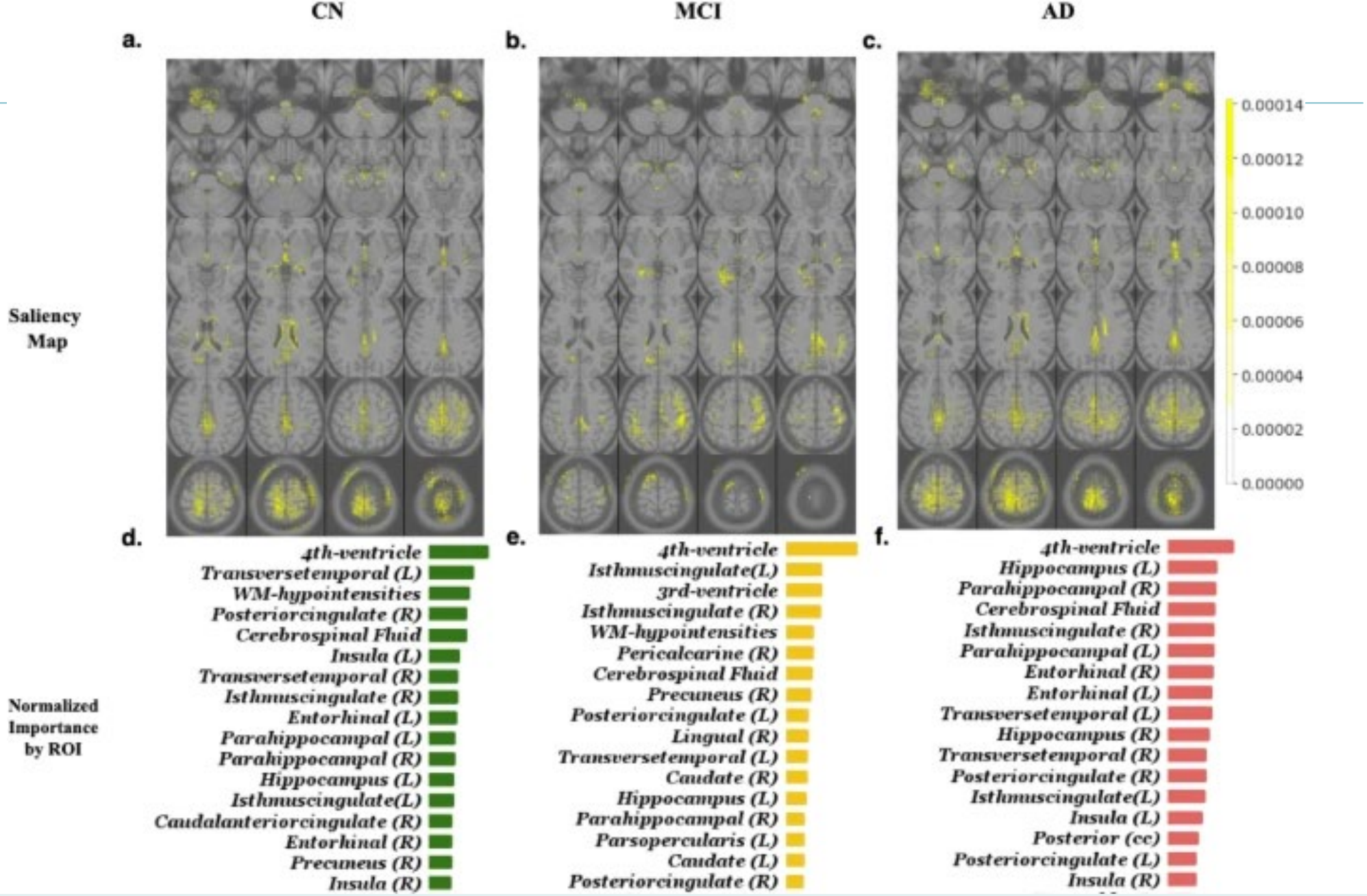
Free-surfer Based Model


ADNI heldout set
NACC
Sex
Education
ApoE4
CN v.s. Others
MCI v.s. Others
AD v.s. Others
0.9
0.7
0.5
0.3
Male
Female
<15 years
>=15 years
w/- ApoE4
w/o ApoE4
Deep learning
ROI-volume/thickness

Validation Under Different Comorbidities

NACC Subcohort Comorbidities (Active or in the past)	Number of NACC participants with the comorbidity	AUC of detecting early stage AD/ADRD (CN vs MCI/AD)	AUC of detecting later stage AD/ADRD (AD vs MCI/CN)
Diabetes	173	80.14%	83.85%
Congestive Heart Failure	18	84.64%	88.97%
Hypertension	702	85.55%	87.78%
Hypercholesterolemia	845	86.93%	90.69%
Atrial fibrillation	110	88.92%	83.94%
Angioplasty/endarterectomy/stent O	78	90.29%	88.27%
Cardiac bypass procedure	33	82.72%	79.14%
Angina	10	60.71%	55.56%
Other cardiovascular disease	251	79.12%	91.70%
Stroke	43	93.38%	93.76%

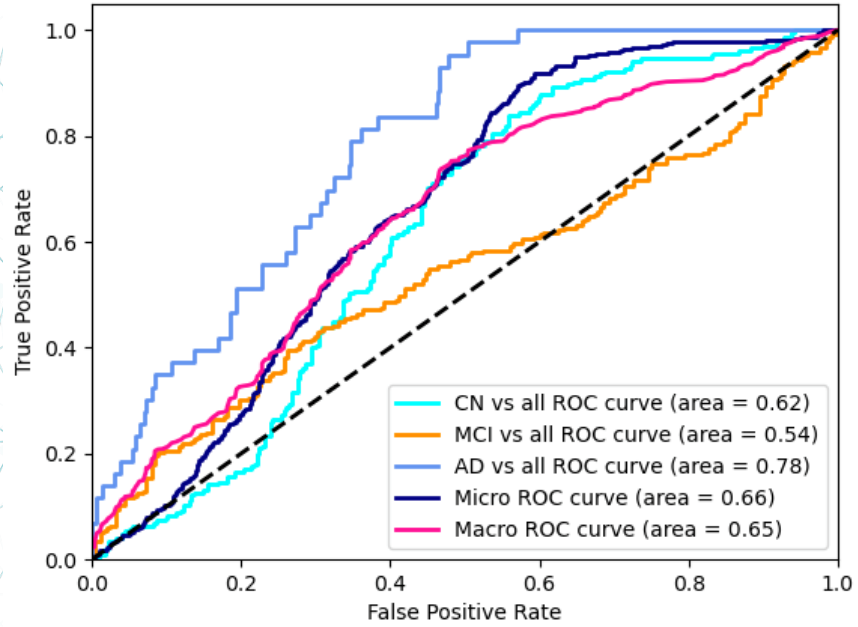
AI Model Explanations



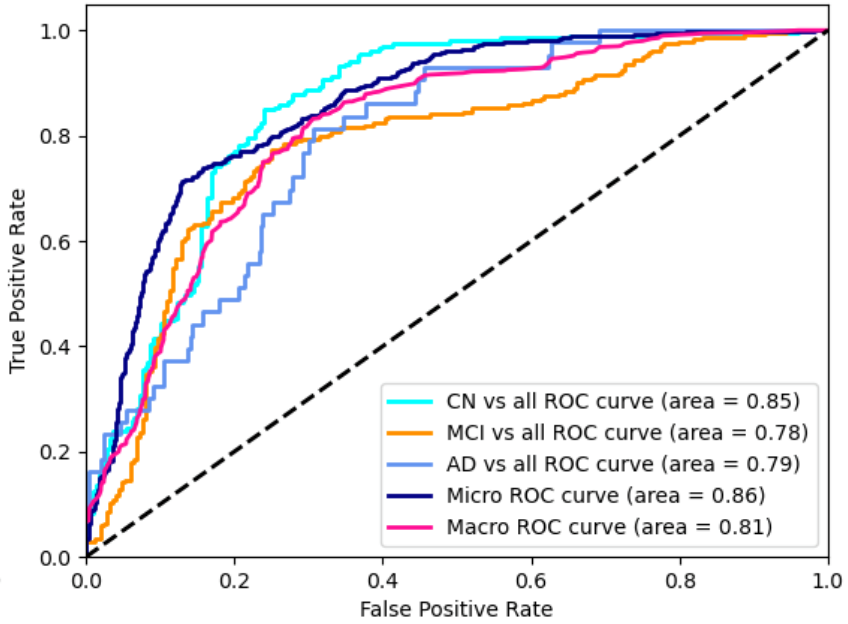
Clinical MRI data from NYU Barlow Memory Center (NIH Designated AD research center ADRC) (Patients who visited 10 neurologists there and had MRIs)

	Full cohort Age>65 (% of n=4945)	**Age>65 with Dementia (all subtypes) (% of n=3187)**
Age mean (standard deviation)	80.19 (7.60)	80.79 (7.43)
Gender: Female	2663(53.85%)	1728(54.22%)
Race: Asian	166(3.36%)	92(2.89%)
Race: Black	311(6.29%)	173(5.43%)
Race: White	3469(70.15%)	2192(68.78%)

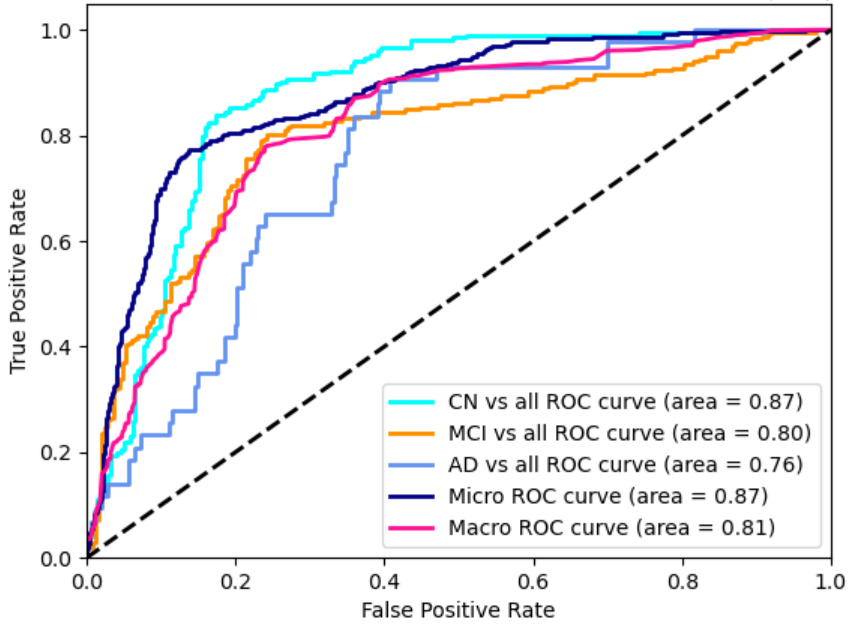
Evaluating our previous model on T1 MRIs



Direct evaluation without re-training



Only fine-tuning the last MLP layers



Re-training the full network

Summary

- NACC enabled closer “real world” validation of our AI tool
- Model is open-source & free to download & use!
  - https://github.com/NYUMedML/CNN_design_for_AD
  - link is also in our paper
- Clinical preliminary results seem promising
- Our next steps
  - Retrospective and prospective clinical evaluation
  - Randomized Trial of the AI tool to assess impact
  - Continue to evaluate on new NACC data

scientific reports

Explore content · About the journal · Publish with us

nature › scientific reports › articles › article

Article | Open Access | Published: 17 October 2022

Generalizable deep learning model for early Alzheimer’s disease detection from structural MRIs

Sheng Liu, Arjun V. Masurkar, Henry Rusinek, Jingyun Chen, Ben Zhang, Weicheng Zhu, Carlos Fernandez-Granda & Narges Razavian

Scientific Reports **12**, Article number: 17106 (2022) | Cite this article

16k Accesses | **3** Citations | **98** Altmetric | Metrics

Abstract

Early diagnosis of Alzheimer’s disease plays a pivotal role in patient care and clinical trials. In this study, we have developed a new approach based on 3D deep convolutional neural networks to accurately differentiate mild Alzheimer’s disease dementia from mild cognitive impairment and cognitively normal individuals using structural MRIs. For comparison, we have built a reference model based on the volumes and thickness of previously reported brain regions that are known to be implicated in disease progression. We validate both models on an internal held-out cohort from The Alzheimer's Disease Neuroimaging Initiative (ADNI) and on an external independent cohort from The National Alzheimer's Coordinating Center (NACC). The deep-learning model is accurate, achieved an area-under-the-curve (AUC) of 85.12 when distinguishing between cognitive normal subjects and subjects with either MCI or mild Alzheimer’s dementia. In the more challenging task of detecting MCI, it achieves an AUC of 62.45. It is also significantly faster than the volume/thickness model in which the volumes and thickness need to be extracted beforehand. The model can also be used to forecast progression: subjects with mild cognitive impairment misclassified as having mild Alzheimer’s disease dementia by the model were faster to progress to dementia over time. An analysis of the features learned by the proposed model shows that it relies on a wide range of regions associated with Alzheimer's disease. These findings suggest that deep neural networks can automatically learn to identify imaging biomarkers that are predictive of Alzheimer's disease, and leverage them to achieve accurate early detection of the disease.

Acknowledgements

NACC and **ADNI**

Sheng Liu
Arjun Masurkar
Henry Rusinek
Jingyun Chen
Ben Zhang
Weicheng Zhu
Carlos Fernandez-Granda

Yin Aphinyanaphongs

Funding Sources

NIH National Institutes of Health
Turning Discovery Into Health

The Leon Lowenstein Foundation

NSF

NYU Langone Health

NYU Center for Data Science

Thank you!