Validation of Risk Models for AD Using Independent Progression Cohort Studies

by Colin Birkenbihl

4th April 2019

AETIONOMY Fraunhofer SCAI ucb imi innovative medicines initiative efpia

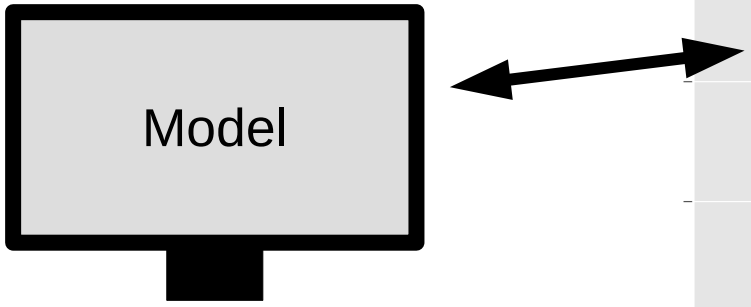
Motivation

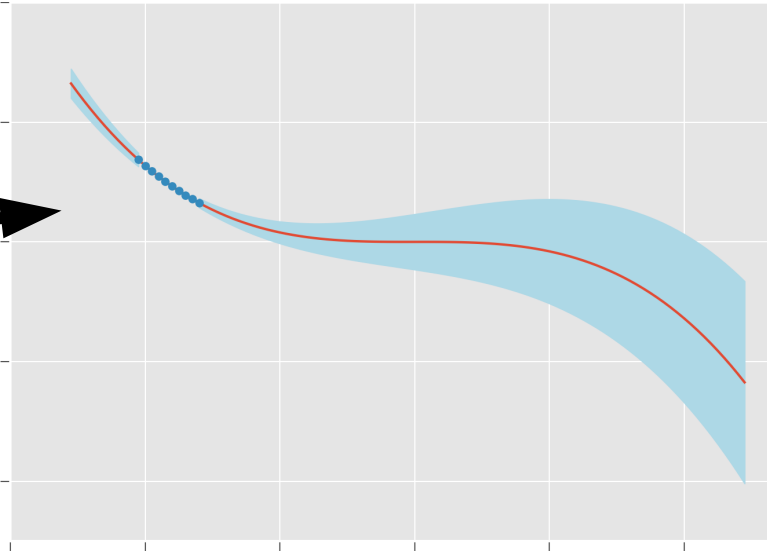
- Pre-symptomatic diagnosis of AD is vital
- Machine Learning allows to assess an individual's disease risk years before diagnosis
- In 2018 Khanna *et. al* published an AD risk model
  - Cross validation C-index ~ 0.86
  - Not validated on external cohort data

Validation Principles

Models are only applicable on comparable data

- Same features as in training dataset
- Similar value ranges

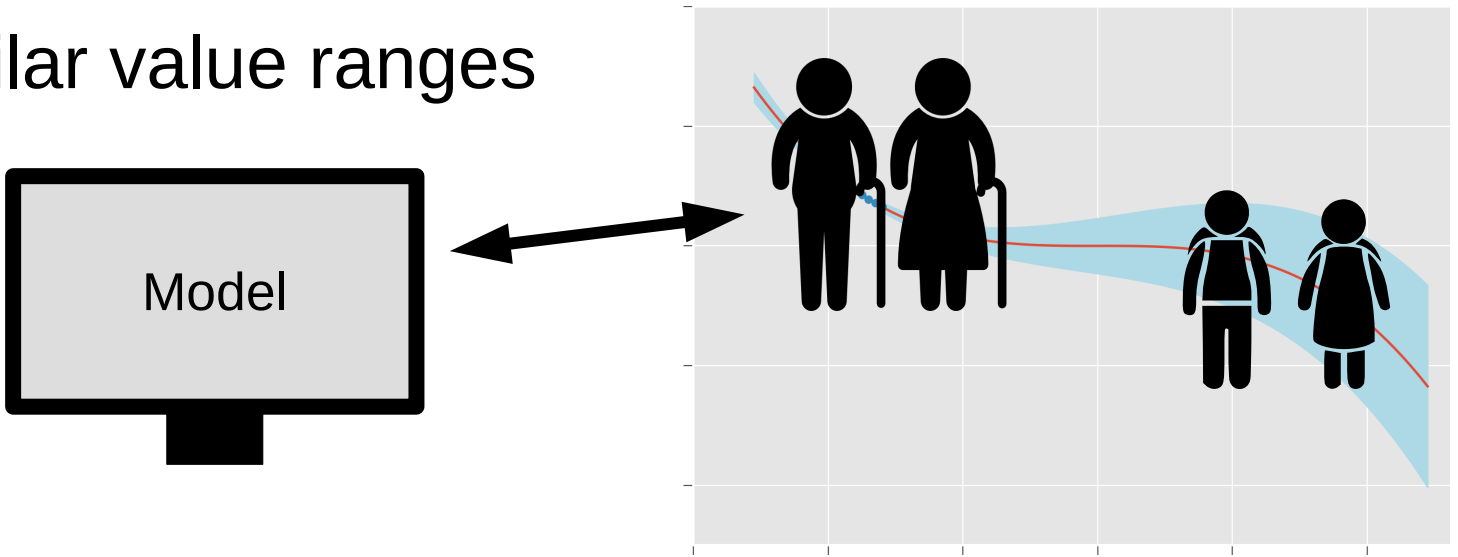




Validation Principles

Models are only applicable on comparable data

- Same features as in training dataset
- Similar value ranges



Motivation

- The AD data landscape is scattered
  - Often external cohort study data not accessible
  - Each study with own assumptions and biases
  - We do not know how comparable the data really are
- It remains unclear if AI models build based on one study generalize

Goals

- Systematic statistical comparison of two major AD studies
  - ADNI
  - AddNeuroMed (incl. ART Cohort & Dementia Case Register)
- Demonstrate that despite evident differences comparable subcohorts can be found
- Validation of the AD riskmodel on a comparable AddNeuroMed subcohort

COMPARISON

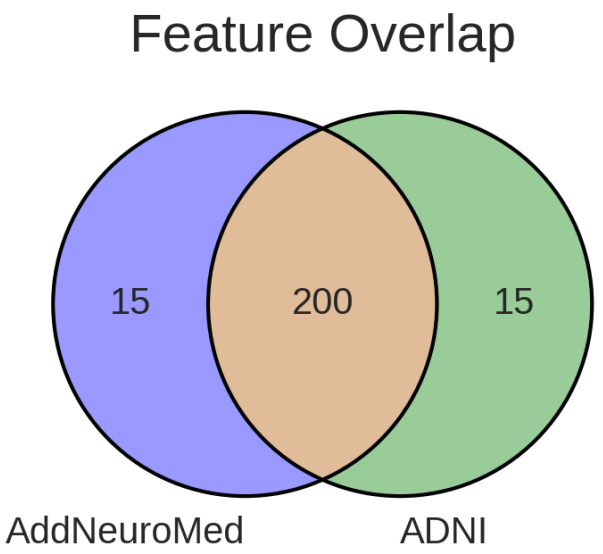
ADNI vs. AddNeuroMed

Comparison Methods

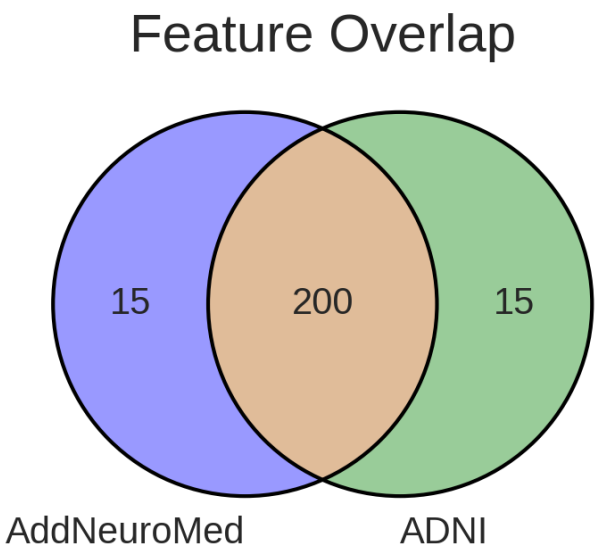
- Compare patients with same diagnosis
- Focus on demographic, clinical and imaging features

- Compare feature overlap
- Nonparametric hypothesis testing + FWE correction
- Assess significantly different features

Comparison Results



Comparison Results



	Unmatched
CTL	53
MCI	152
AD	148

Majority of compared features differed significantly

Demographic Comparison

- Significant differences across all demographic features

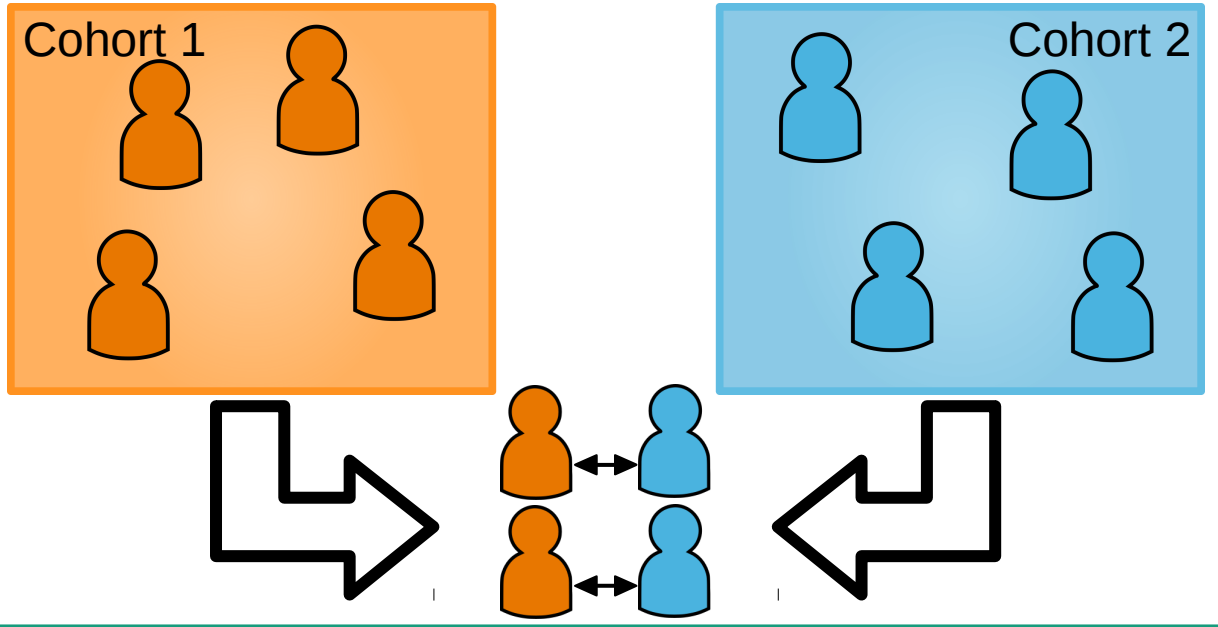
	CTL p-value	MCI p-value	AD p-value
Gender	0.0	0.0	0.0
Age	0.94	7.6e-9	7.0e-9
Education	0.0	0.0	0.0
APOE4	2.1e-5	0.0	1.1e-14

Cohorts are not comparable as are….

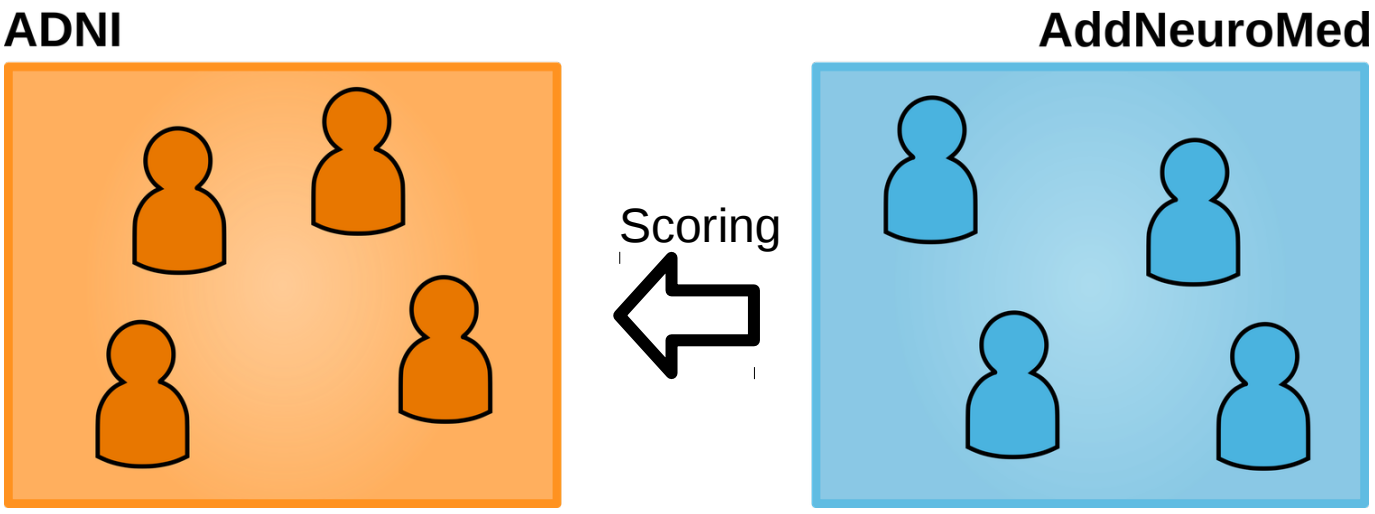
Propensity Score Matching

Propensity Score Matching: Goal

Goal:
Match patients from C1 with similar counterpart from C2

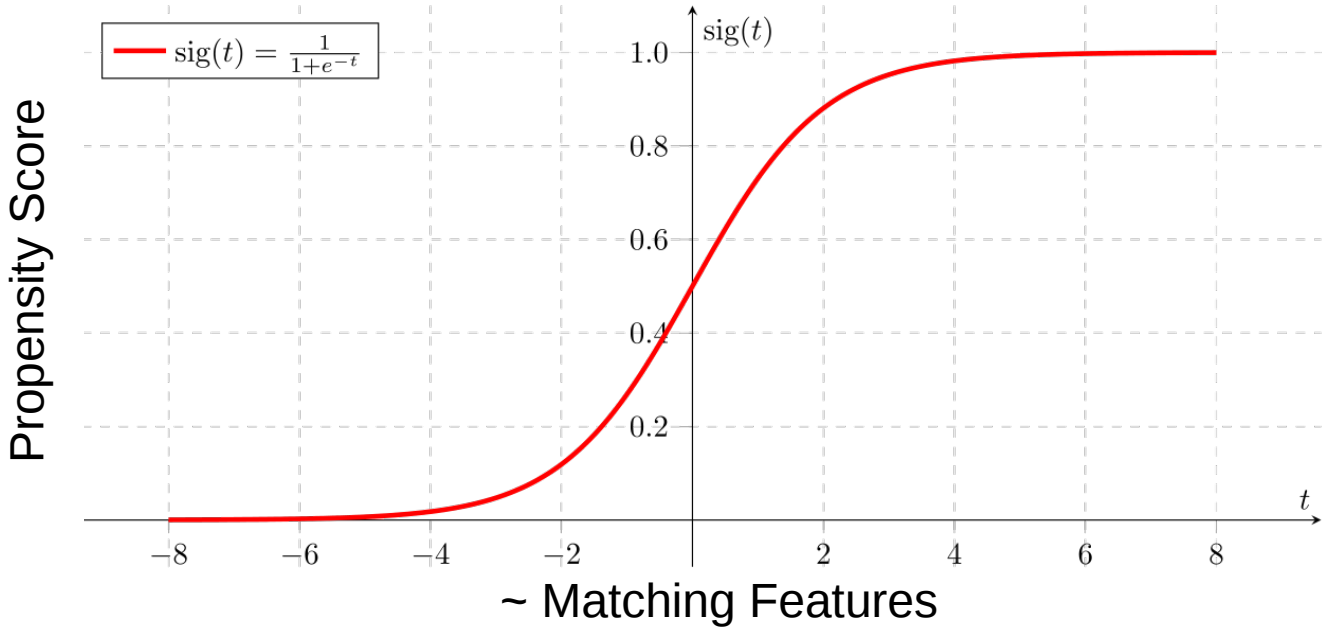


Propensity Score Matching



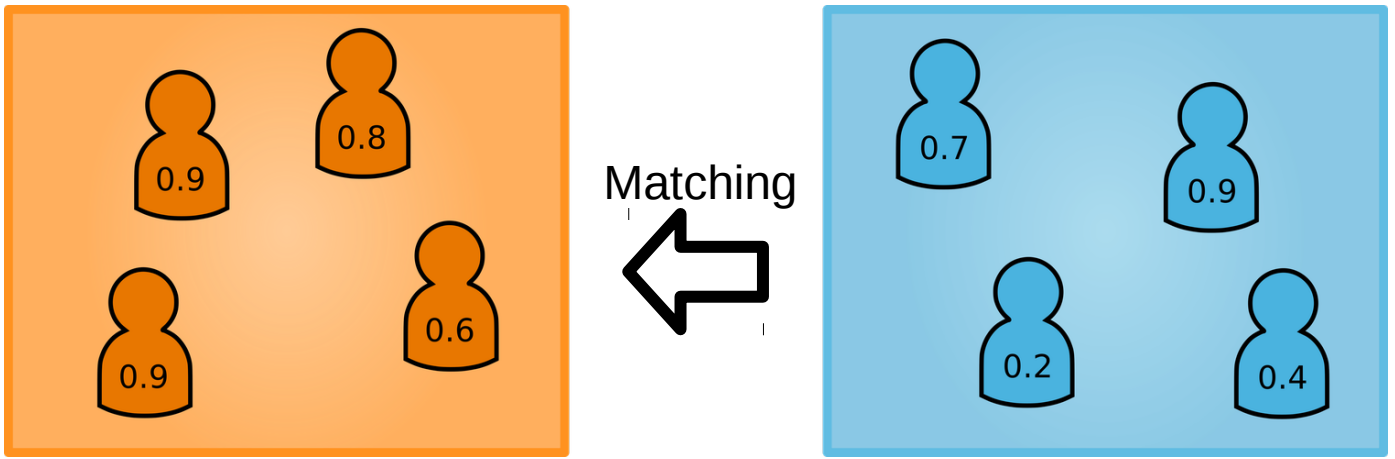
- Propensity scoring based on AddNeuroMed

Propensity Scoring



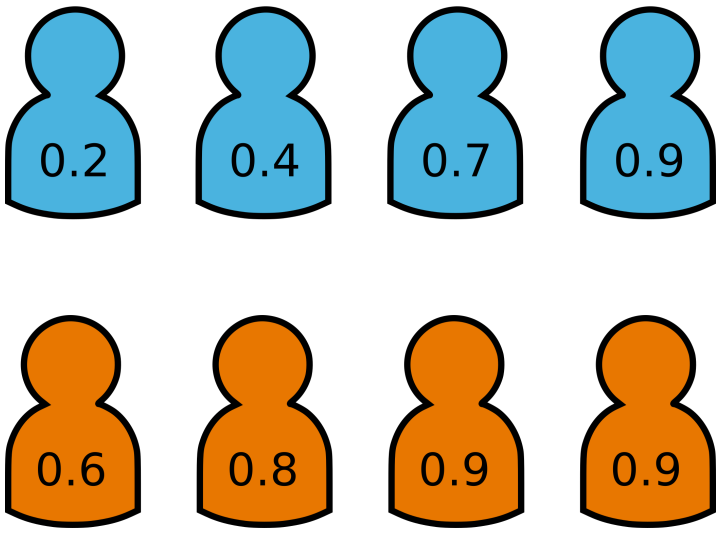
Matching Features: Sex, Age, Education, APOE4, MMSE

Propensity Score Matching



- Match AddNeuroMed Patients to comparable ADNI patients

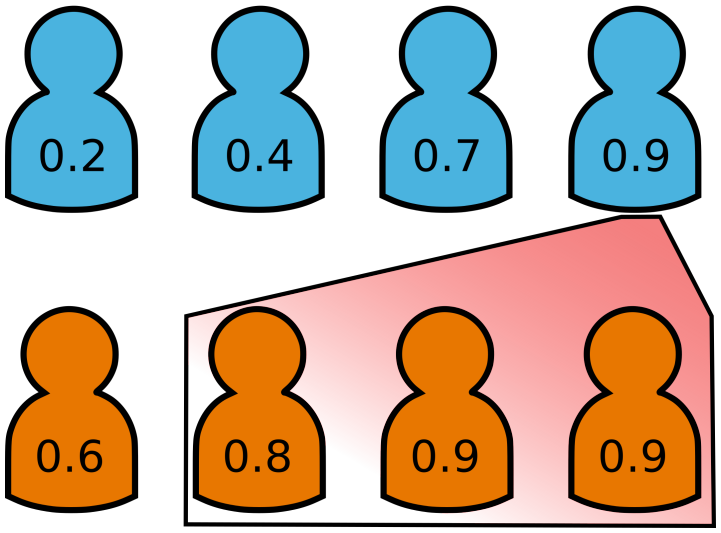
Caliper Based Matching



Caliper = 0.2

- Assign patients a counterpart within caliper range

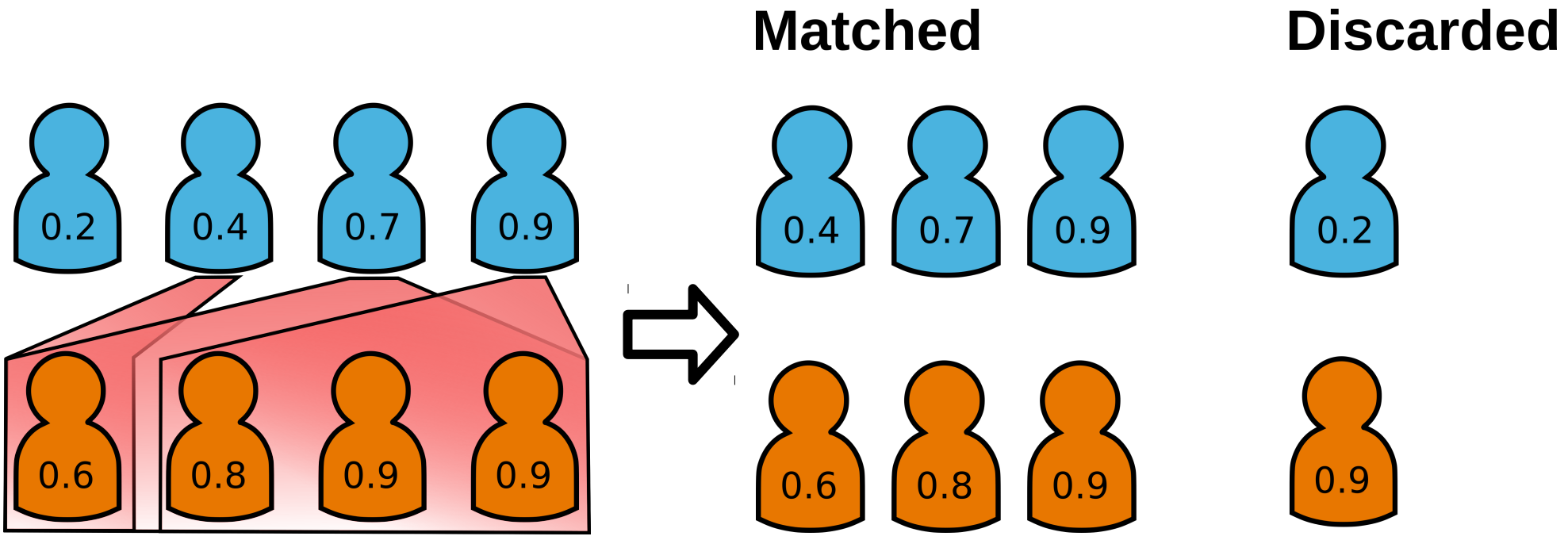
Caliper Based Matching



Caliper = 0.2

- Assign patients a counterpart within caliper range

Caliper Based Matching



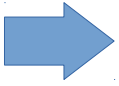
Comparison after Matching

Data Availability

- PSM only works for complete cases
- Inherent data loss when applied to clinical study data

	CTL			MCI			AD		
	n	CC	match	n	CC	match	n	CC	match
ADNI	417	415	199	872	866	147	342	338	111
AddNeuroMed	793	266	199	397	238	147	512	262	111

- Assessed number of significant features for 100 matchings



Comparable subcohorts present in ADNI and AddNeuroMed

Risk Model Validation

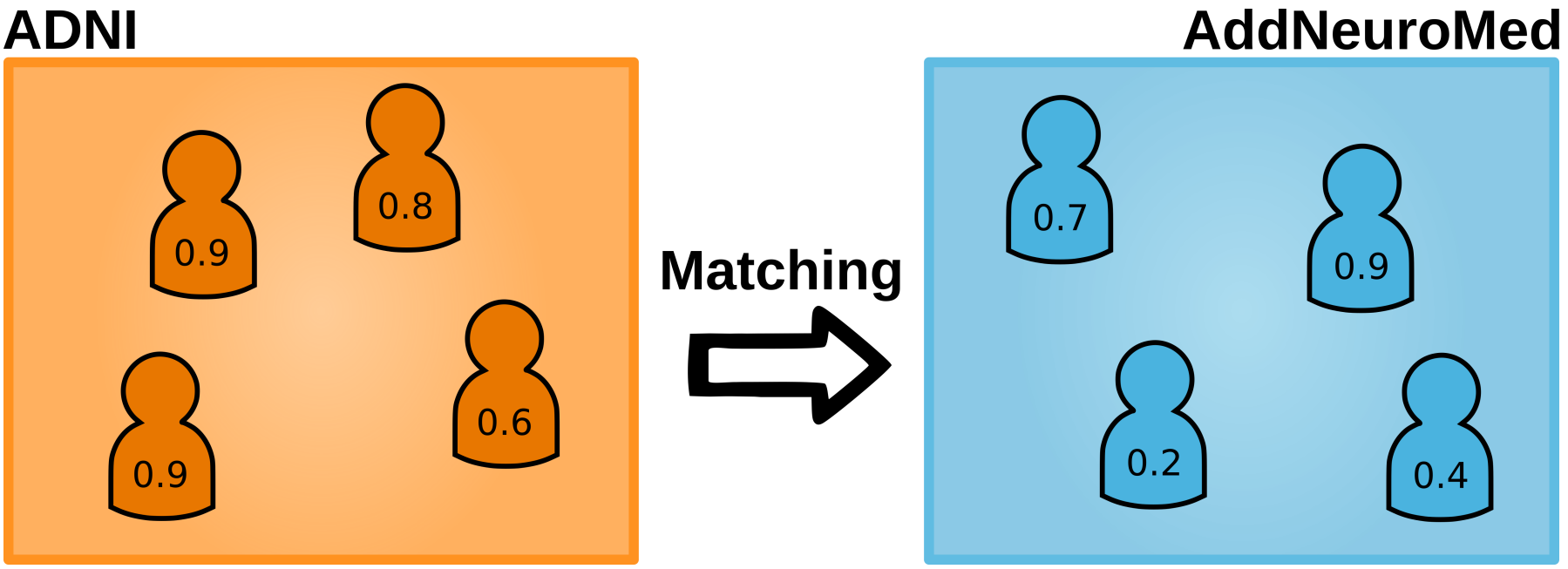

Risk Model

- Gradient Boosting Machine
- Predicts time to event (AD onset)
- Trained on ADNI baseline data of control and MCI patients

- Features:
  - Clinical, Demographic, MRI, Pathway impact scores, Genetic

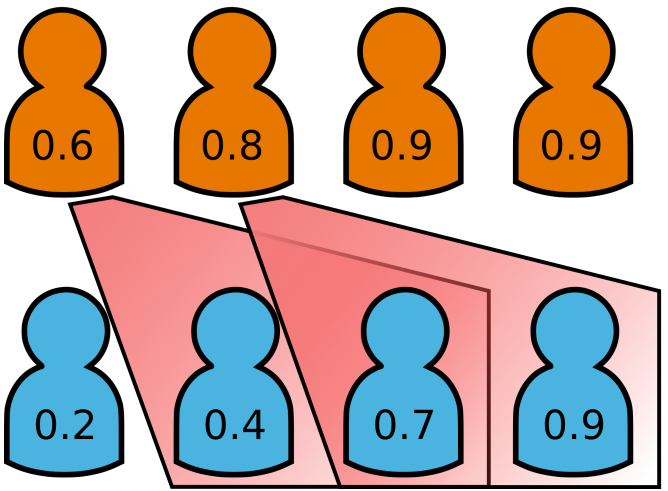
Matching for Validation

- Based on ADNI to find comparable AddNeuroMed patients

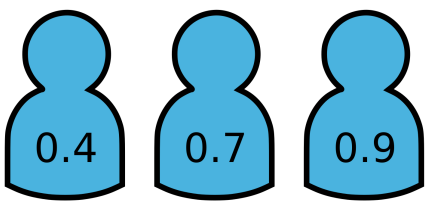


Matching for Validation

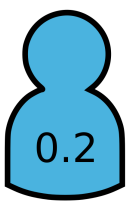
Caliper based patient matching



Used in Validation



Discarded



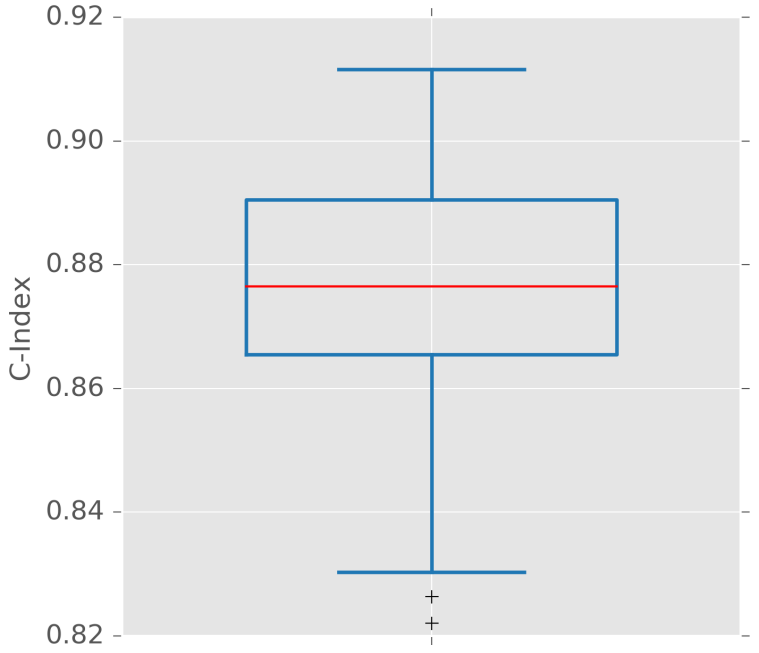
Validation Methods

100 Matching and validation runs

Average matched validation set composed of…

- ~ 160 control / MCI Patients
- 30 Converters (Events)

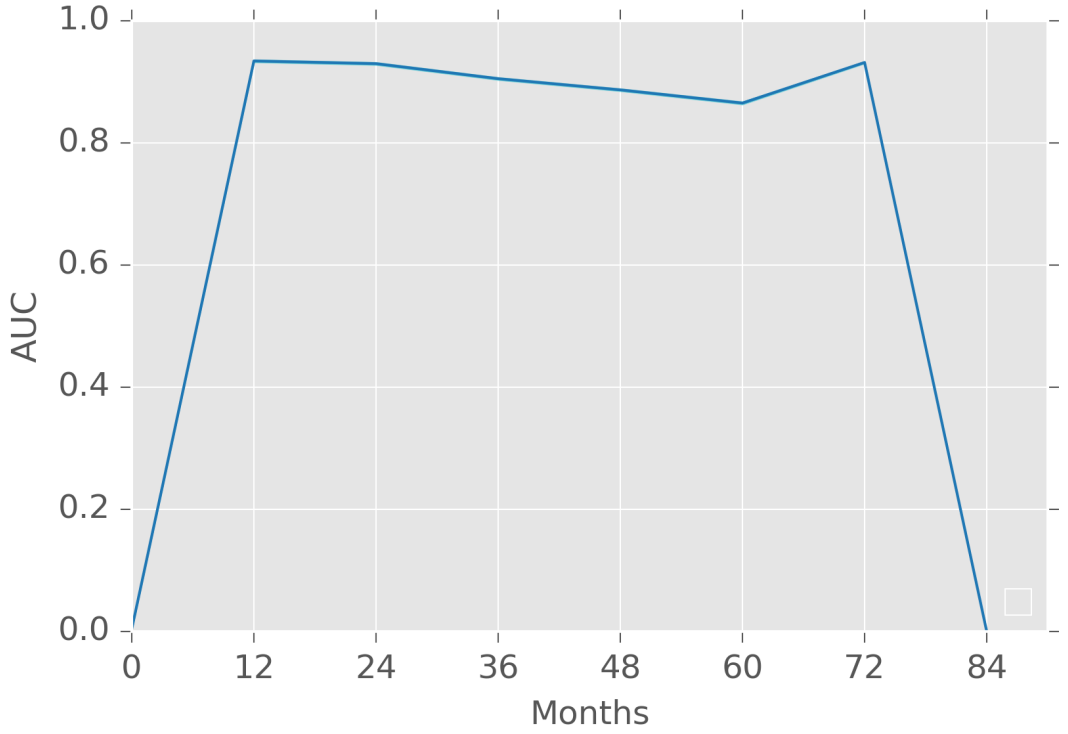
Validation Performance



- C-index ~ 0.88

Performance similar to cross-validation on training data

Validation Performance



- Predicts AD reliably up to 6 years before onset

Conclusion

- Absence of data and features is still very much limiting
- Even in largely different cohorts comparable subcohorts can be found
- Statistical matching allows for finding comparable subcohorts
- The model we build and validated shows to predict AD reliably up to 6 years before onset

Thank You!

And thanks to:

- Holger Fröhlich
- Martin Hofmann-Apitius
- Asif Emon Kahn
- Sarah Westwood
- Henri Vrooman
- Simon Lovestone

The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking under grand agreement n° 115568 resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2017-2013) and EFPIA companies in kind contribution.