Machine Learning-based predictions of prognosis in Cystic Fibrosis

UCLA

Ahmed M. Alaa[1], Thomas Daniels[2], R. Andres Floto[3], and Mihaela van der Schaar[1,4]

[1]UCLA, [2] University Hospital Southampton, [3] Department of Medicine, University of Cambridge & Royal Papworth Hospital, Cambridge, [4] Alan Turing Institute

University of Oxford

This work was supported by the UK Cystic Fibrosis Trust.

Objectives

Using an **automated machine learning** framework (**AutoPrognosis**) to build a model for predicting 3-year mortality in CF patients using data from the UK CF registry.

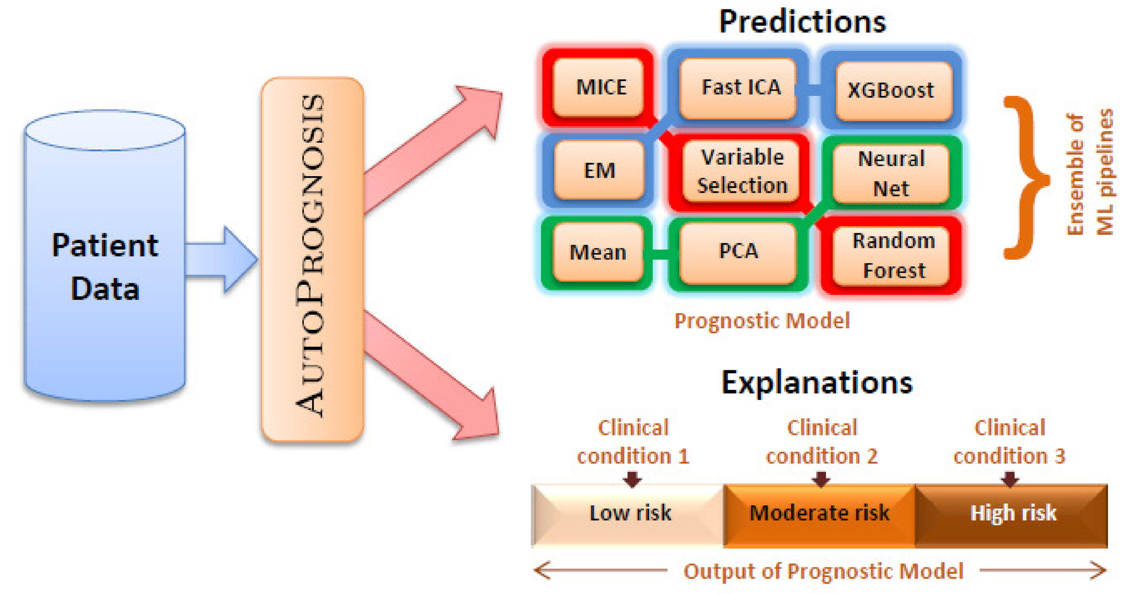
Crucial for many clinical decisions:

- Establishing the optimal timing for referring patients for **lung transplantation**.
- Administering different types of **treatments**.

Our method: AutoPrognosis

- **Automatically** constructs ensembles of prognostic modeling **pipelines**, provides **"clinical explanations"** for the learned models, and can easily update its learned models as more data is collected over time.

- **A prognostic modeling pipeline:** data imputation, feature processing and classification algorithms.



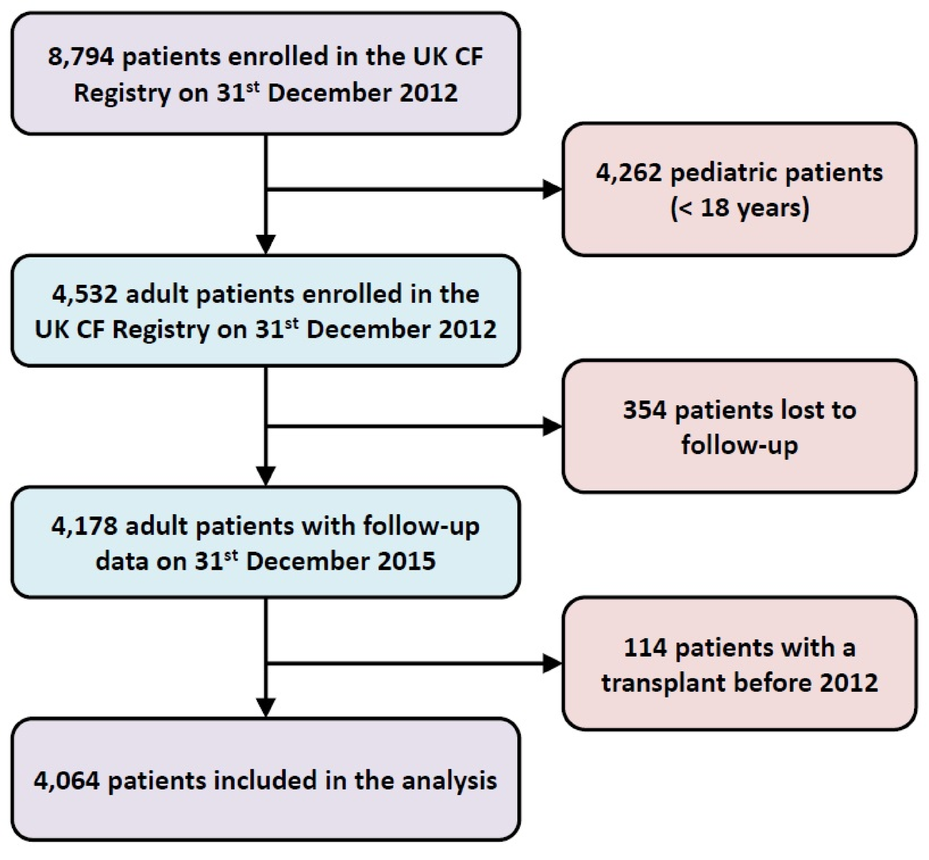
Patient Cohort

Inclusion criteria:

- Enrolled in the UK CF registry with annual follow-up data available on the 31st of **December, 2012**.
- Adult patients who are over **18 years** of age.
- Follow-up data on the 31st of **December, 2015**.

Cohort statistics:

- A total of **4,064 patients** included.
- A total of **115 variables** for each patient.
- Mortality rate was **9.4%**.



Diagnostic Accuracy

How should diagnostic accuracy be evaluated?

- **Most previous works focused on AUC-ROC,** but AUC-ROC can be deceptively large because true negatives can be trivially predicted + AUC-ROC does not account for imbalanced outcomes.
- **Alternative:** area under the precision-recall curve (average precision) focuses only on positive cases.

Prognostic model	AUC-ROC	Youden's J statistic	AUC-PR	Average Precision	F_1 score
AutoPrognosis	0.89 ± 0.01	0.67 ± 0.02	0.58 ± 0.04	0.59 ± 0.04	0.60 ± 0.03
Nkam *et al.*	0.86 ± 0.01	0.58 ± 0.03	0.50 ± 0.03	0.48 ± 0.03	0.52 ± 0.02
Buzzetti *et al.*	0.83 ± 0.01	0.54 ± 0.03	0.42 ± 0.02	0.44 ± 0.03	0.49 ± 0.02
CF-ABLE-UK	0.77 ± 0.01	0.48 ± 0.05	0.28 ± 0.04	0.20 ± 0.02	0.34 ± 0.02
FEV_1% predicted criterion	0.70 ± 0.01	0.41 ± 0.02	0.50 ± 0.02	0.27 ± 0.02	0.47 ± 0.01

Impact on lung transplant referral decisions

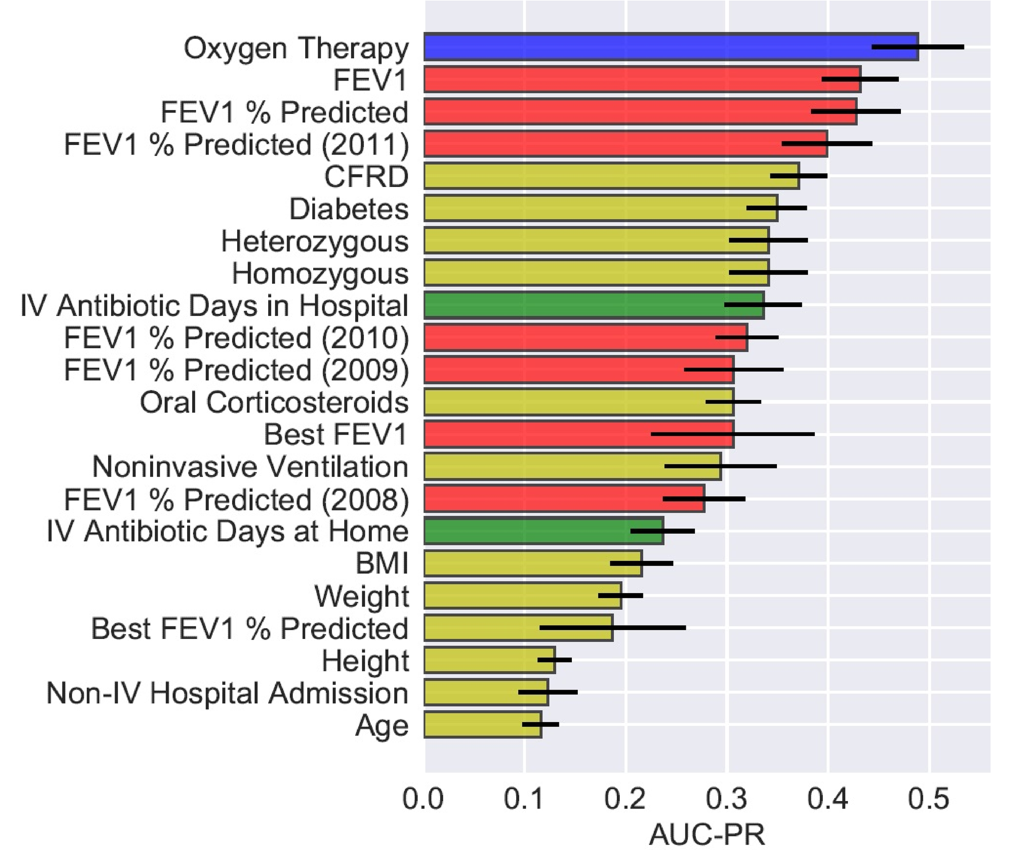
- **Operating point:** fix the negative predictive value (NPV) at the one achieved by the "$FEV_1 < 30\%$" criterion.
- **Improvement in the positive predictive value (PPV) from 48% to 65%.**

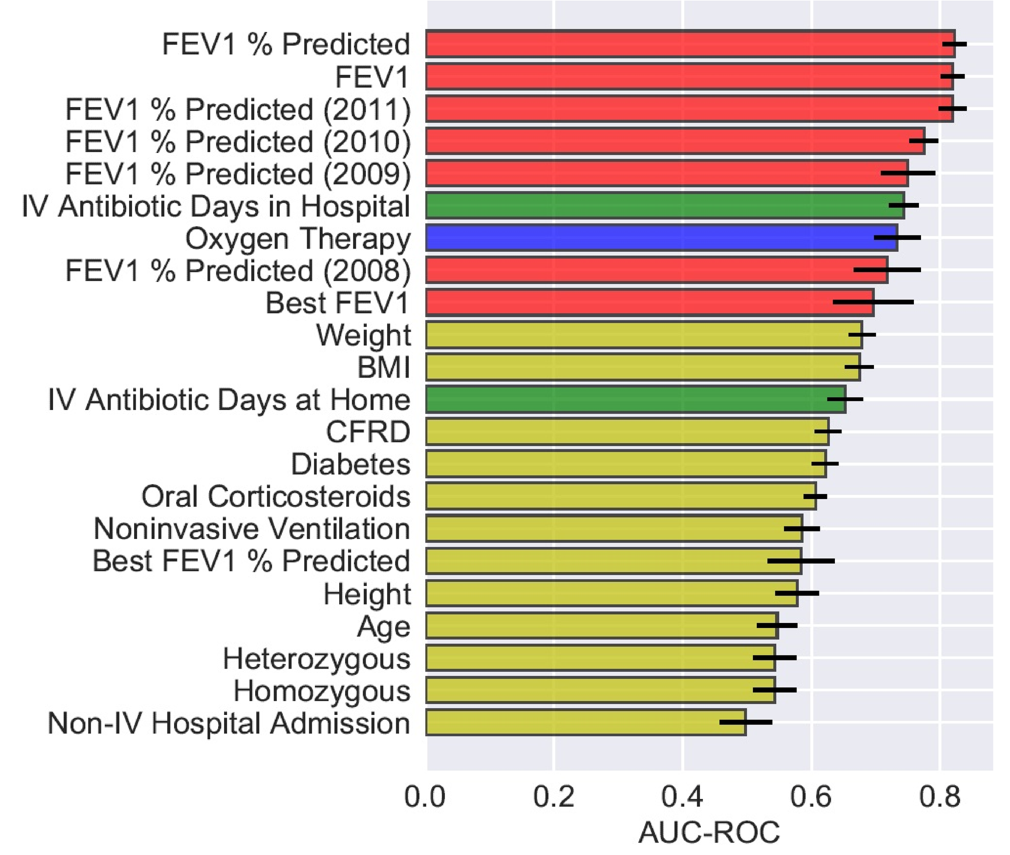
	Cutoff	PPV (95% CI) (%)	NPV (95% CI) (%)	Sens (95% CI) (%)	Spec (95% CI) (%)	Accuracy (%)	F_1 score
FEV_1% predicted	<20	66 (62, 70)	92 (91, 93)	13 (9, 17)	99 (98, 100)	92 (91, 93)	21 (19, 23)
	<30	48 (44, 52)	95 (94, 96)	46 (42, 50)	95 (94, 96)	91 (90, 92)	47 (45, 49)
	<40	29 (27, 31)	96 (95, 97)	62 (60, 64)	86 (84, 88)	84 (83, 85)	40 (38, 42)
	<50	21 (19, 23)	97 (96, 98)	73 (71, 75)	75 (73, 77)	75 (74, 76)	33 (31, 35)
Nkam *et al.*[36]	>6.5	75 (64, 86)	92 (91, 93)	13 (11, 15)	99 (98, 100)	92 (91, 93)	22 (19, 25)
	>4	56 (52, 60)	95 (94, 96)	46 (44, 48)	96 (95, 97)	92 (91, 93)	50 (49, 51)
	>2.5	42 (37, 47)	96 (95, 97)	61 (60, 62)	91 (90, 92)	88 (87, 89)	49 (45, 53)
	>2	31 (27, 35)	97 (96, 98)	73 (72, 74)	83 (79, 87)	82 (78, 86)	43 (39, 47)
	>0.50	88 (79, 97)	92 (91, 93)	13 (12, 14)	99 (98, 100)	92 (91, 93)	23 (22, 24)
AutoPrognosis	>0.33	65 (61, 69)	95 (94, 96)	46 (45, 47)	97 (96, 98)	93 (92, 94)	53 (51, 55)
	>0.15	49 (43, 55)	96 (95, 97)	62 (61, 63)	93 (92, 94)	90 (89, 91)	54 (50, 58)
	>0.10	36 (32, 40)	97 (96, 98)	74 (73, 75)	87 (86, 88)	86 (84, 88)	48 (45, 51)

Risk Factors

Which patient variables best explain accuracy gains?

- **Variable importance rankings depend on the diagnostic accuracy metric used.**
- **Oxygen therapy is the variable used by machine learning to improve precision.**





Conclusions

- **The area under precision-recall curve is a more appropriate metric than AUC-ROC for evaluating prognostic scores. This fact was overlooked by previous works.**

- Our results indicate that competitive machine learning approaches significantly improve prognostic forecasting and will support optimized referral for lung transplantation.

- Incorporating variables related to gas exchange (Oxygenation) into predictive models in addition to spirometric variables can significantly boost the precision of lung transplant referral decisions.