Deep Learning Utilizing Discarded Spirometry Data to Improve Lung Function and Mortality Prediction

Davin Hill[1], Max Torop[1], Aria Masoomi[1], Peter J. Castaldi[2], Michael Cho[2,3], Jennifer Dy[1], Brian D. Hobbs[2,3]

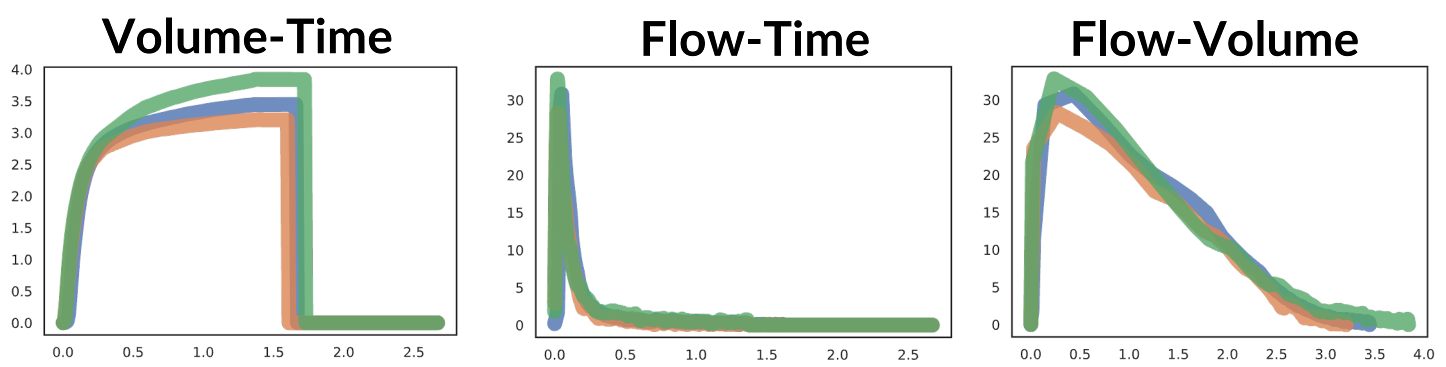
[1] Northeastern University, Department of Electrical and Computer Engineering, Boston, MA
[2] Channing Division of Network Medicine, Brigham and Women's Hospital, Boston, MA [3] Division of Pulmonary and Critical Care Medicine, Brigham and Women's Hospital, Boston, MA

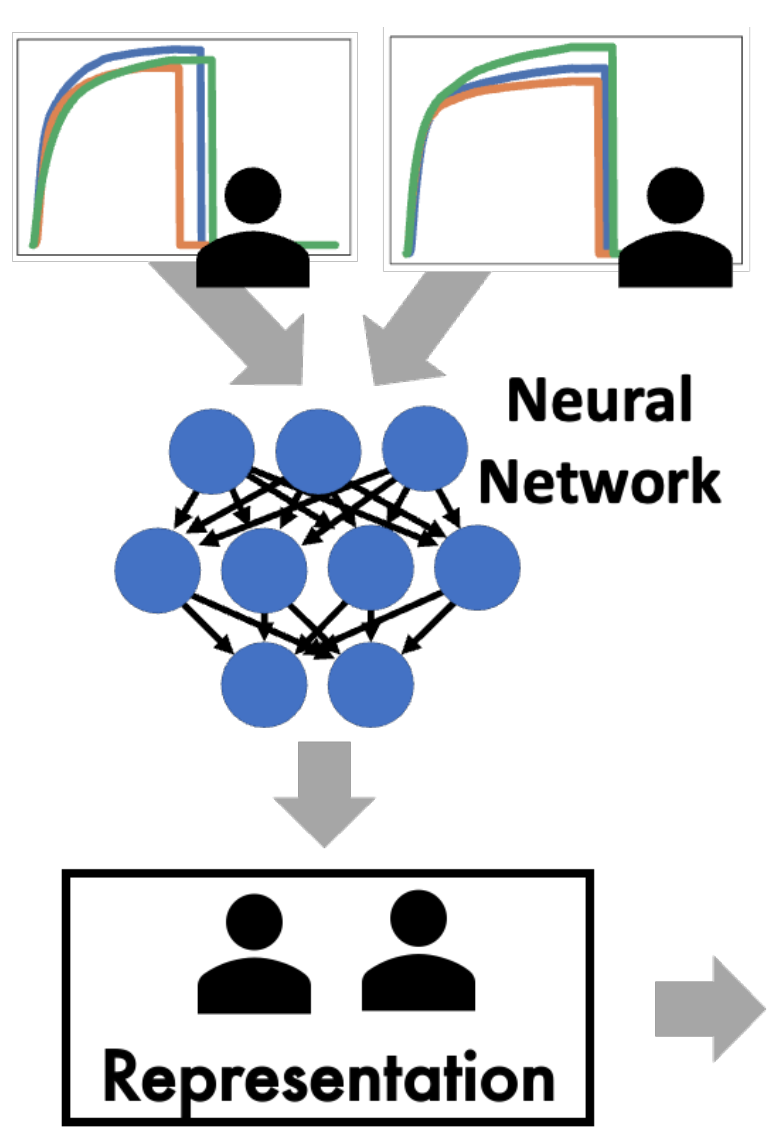
Introduction

- Spirometry measurements are often noisy; two-three blows are often required to estimate spirometry metrics such as FEV_1 and FVC.
- Traditionally only the best FEV_1 and FVC are reported, with data from submaximal and quality control (QC) failing blows being discarded
- We hypothesize that discarded spirometry blow data meaningfully contributes to the prediction of lung health and all-cause mortality.

Method

- Examined 268,473 UK Biobank participants with 2-3 spirometry efforts and at least 1 QC-passing effort
- Spirometry data were volume-time data at 10 ms intervals
- All efforts for each individual were used to train a Deep Learning model based on SimCLR [1]
- Randomly applied transformations to each effort:





- The model is trained to predict whether any two randomly selected, transformed blows are produced by the same individual.
- The learned SimCLR representation (100-dimensional vector) is used in downstream tasks as a proxy for lung function

Downstream Prediction

Results

A) $FEV_1/FVC<0.7$ Prediction

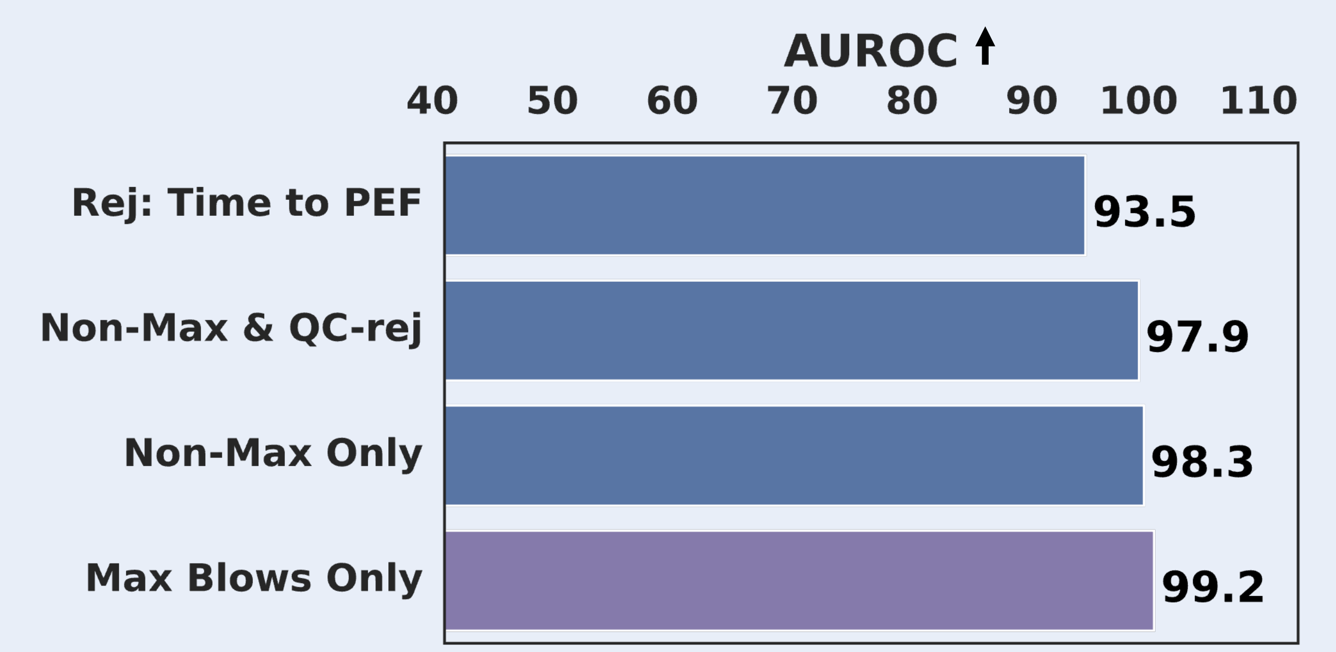


Figure 1: We use SimCLR representations, filtered on only QC-rejected and non-maximal efforts, to predict if the maximal FEV_1/FVC ratio is less than 0.7. AUROC results show that using only QC-rejected and non-maximal efforts can achieve similar performance as using the maximal effort only.

B) FEV_1pp Prediction

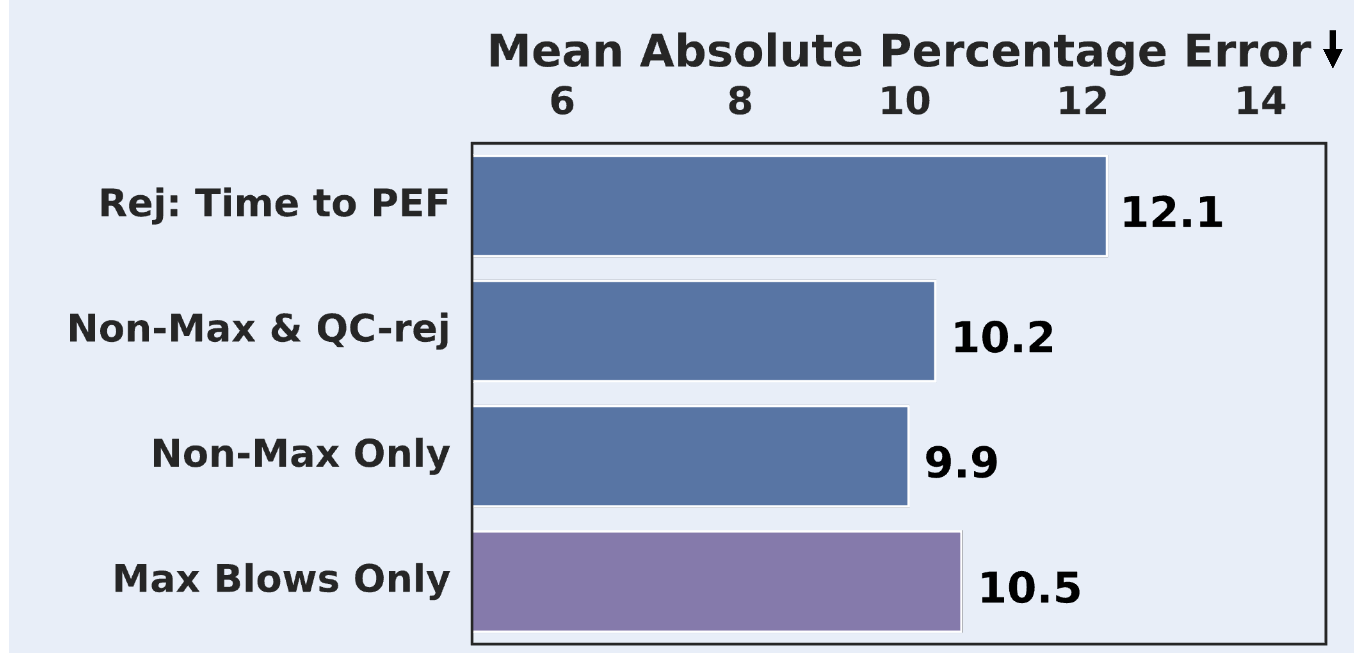


Figure 2: Similar to Fig.1, we use the SimCLR representation filtered on QC-rejected and non-maximal efforts to predict FEV_1 %Predicted (using GLI2012). Using the non-maximal efforts achieves similar mean absolute percentage error as the using only the maximal efforts.

C) Mortality Prediction

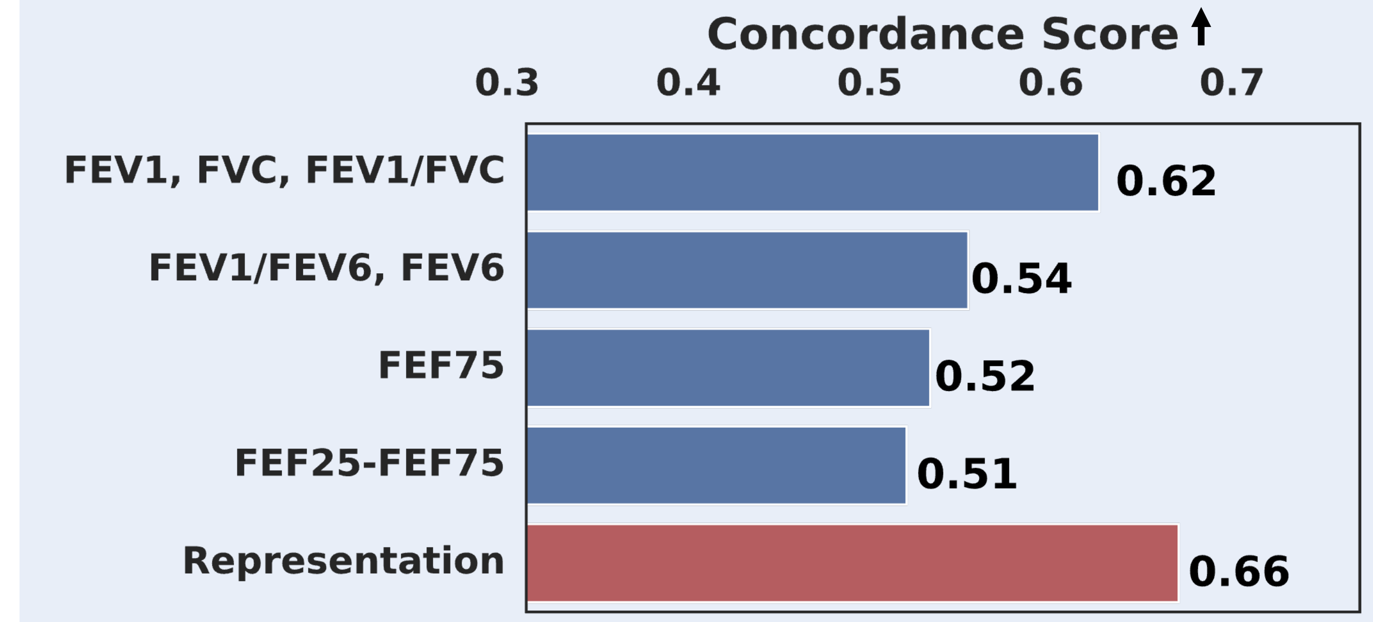


Figure 3: The learned representation is used in time-to-event analysis to evaluate its predictive power compared to traditional spirometry measurements. We see that the concordance score results indicate that our representation results in more accurate predictions.

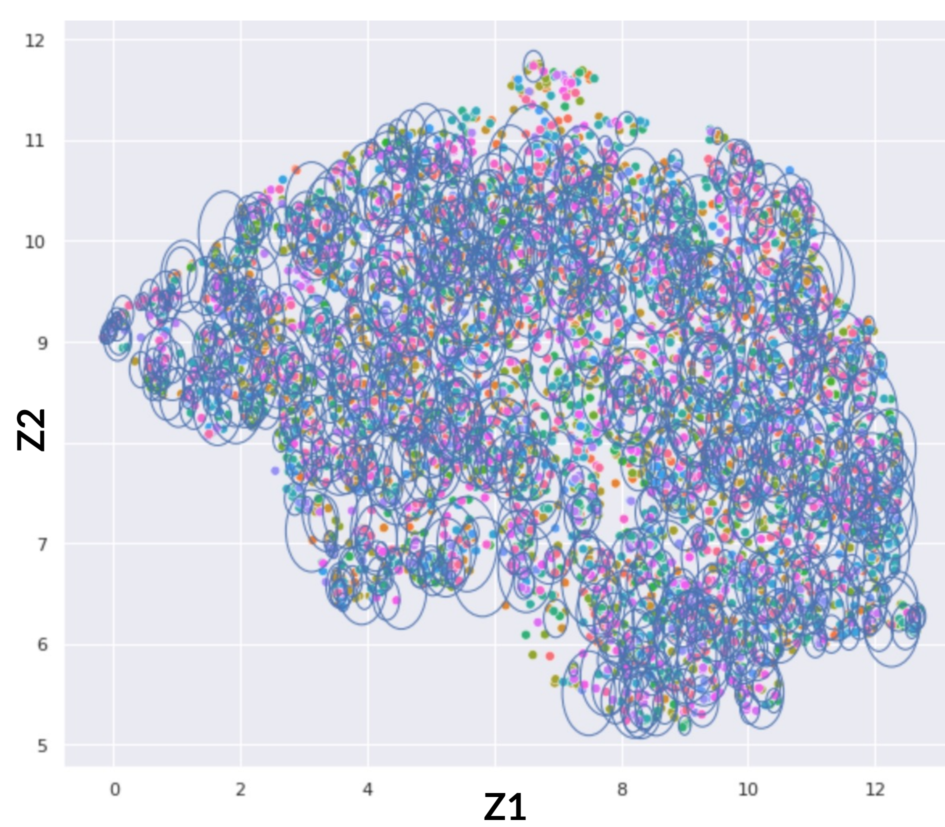
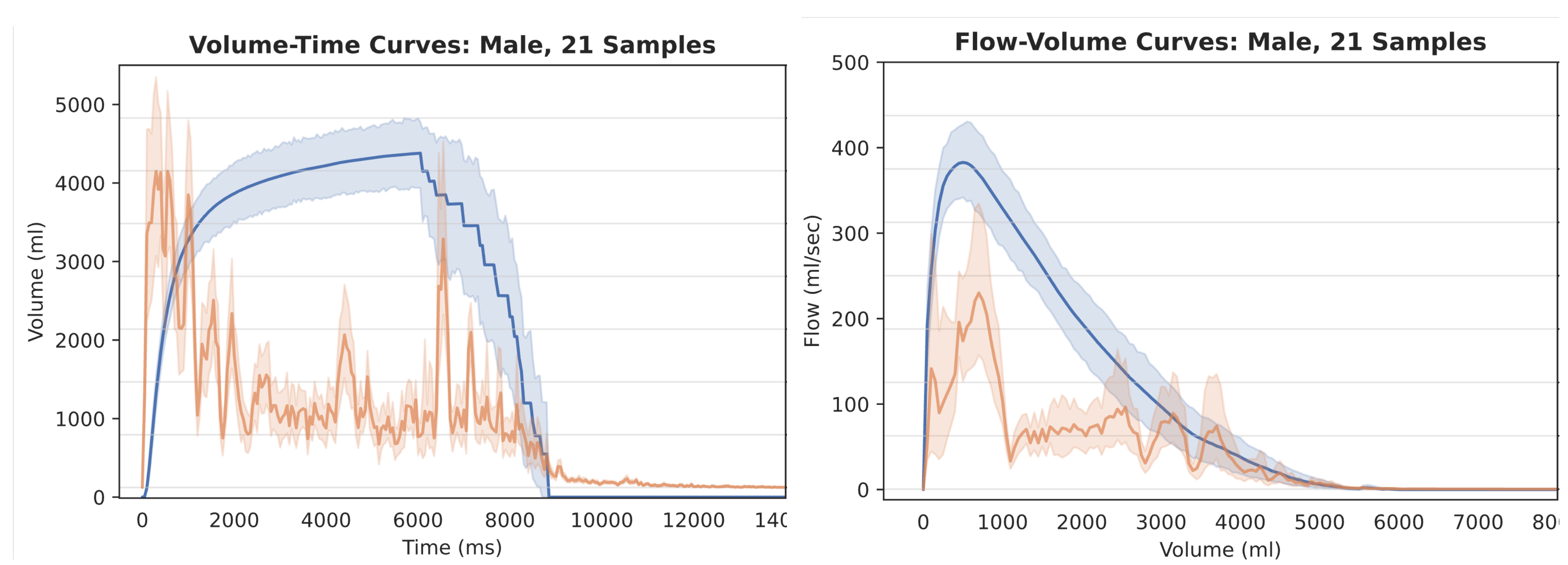


Figure 4 (Left): UMAP [2] low-dimensional projection of the SimCLR representation of each effort. The entirety of an individual's efforts are circled in blue. Each individual's SimCLR spirometry representations are close together, which indicates averaging across representations is a reasonable approach to generating a single representation per individual. This averaging approach improved performance.

Figure 5 (Below): Evaluation of different data transformations (lower is better) on representation similarity. Each cell represents the 1 or 2 transformations (row & column) on the data. We calculate the average distance between blow representations for a given individual after applying the transformations. Increasing the number of transformations increases the distance between representations

	Volume-Time	Flow-Time	Flow-Volume
Volume-Time	0.128	0.282	0.215
Flow-Time	0.282	0.152	0.291
Flow-Volume	0.215	0.291	0.176

Figure 6 (Right): We apply [3] which estimates the relative importance (orange) of the blow values (blue) over time. These importance values are calculated per individual, then averaged over the test set (shaded region represents 2 std. dev.). We see that regions corresponding to PEF and FVC generally have the highest importance towards prediction.



Conclusions

- We trained a Deep Learning SimCLR representation of spirometry data from the UK Biobank, including "discarded" QC-rejected and non-maximal efforts
- We observed feature similarity across the entirety of a participant's blows, indicating the discarded blows contain complementary information to the best blows.
- The SimCLR representation incorporating all QC-passing spirometry efforts improved on prediction of FEV_1pp compared to using only maximal efforts
- The SimCLR representation also improves time-to-event mortality prediction compared to traditional spirometry metrics.

References / Funding

[1] Chen, T., Kornblith, S., Norouzi, M. & Hinton, G.. (2020). A Simple Framework for Contrastive Learning of Visual Representations.

[2] McInnes, Leland, John Healy, Nathaniel Saul, and Lukas Grossberger. 2018. "UMAP: Uniform Manifold Approximation and Projection."

[3] Frye, Christopher, Colin Rowat, and Ilya Feige. 2020. "Asymmetric Shapley Values: Incorporating Causal Knowledge into Model-Agnostic Explainability."

The project was funded by various NIH-NHLBI grants.