

# Area under the expiratory flow-volume curve: predicted values by regression and deep learning methods and recommendations for clinical practice

Octavian C Ioachimescu ,<sup>1,2</sup> José A Ramos,<sup>3</sup> Michael Hoffman,<sup>3</sup> James K Stoller<sup>4</sup>

**To cite:** Ioachimescu OC, Ramos JA, Hoffman M, *et al.* Area under the expiratory flow-volume curve: predicted values by regression and deep learning methods and recommendations for clinical practice. *BMJ Open Resp Res* 2021;**8**:e000925. doi:10.1136/bmjresp-2021-000925

► Additional online supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/bmjresp-2021-000925>).

Received 11 March 2021  
Revised 19 April 2021  
Accepted 20 April 2021



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

<sup>1</sup>School of Medicine, Department of Medicine, Division of Pulmonary, Allergy, Critical Care and Sleep Medicine, Emory University, Atlanta, Georgia, USA

<sup>2</sup>Department of Medicine, Sleep Medicine Section, Atlanta Veteran Affairs Healthcare System, Decatur, Georgia, USA

<sup>3</sup>Cleveland Clinic, Respiratory Institute, Cleveland, Ohio, USA

<sup>4</sup>Cleveland Clinic, Education Institute, Cleveland, Ohio, USA

## Correspondence to

Dr Octavian C Ioachimescu;  
oioac@yahoo.com

## ABSTRACT

**Background** In spirometry, the area under expiratory flow-volume curve (AEX-FV) was found to perform well in diagnosing and stratifying physiologic impairments, potentially lessening the need for complex lung volume testing. Expanding on prior work, this study assesses the accuracy and the utility of several models of estimating AEX-FV based on forced vital capacity (FVC) and several instantaneous flows. These models could be incorporated in regular spirometry reports, especially when actual AEX-FV measurements are not available.

**Methods** We analysed 4845 normal spirometry tests, performed on 3634 non-smoking subjects without known respiratory disease or complaints. Estimated AEX-FV was computed based on FVC and several flows: peak expiratory flow, isovolumic forced expiratory flow at 25%, 50% and 75% of FVC (FEF<sub>25</sub>, FEF<sub>50</sub> and FEF<sub>75</sub>, respectively). The estimations were based on simple regression with and without interactions, by optimised regression models and by a deep learning algorithm that predicted the response surface of AEX-FV without interference from any predictor collinearities or normality assumption violations.

**Results** Median/IQR of actual square root of AEX-FV was 3.8/3.1–4.5 L<sup>2</sup>/s. The per cent of variance (R<sup>2</sup>) explained by the models selected was very high (>0.990), the effect of collinearities was negligible and the use of deep learning algorithms likely unnecessary for regular or routine pulmonary function testing laboratory usage.

**Conclusions** In the absence of actual AEX-FV, a simple regression model without interactions between predictors or use of optimisation techniques can provide a reasonable estimation for clinical practice, thus making AEX-FV an easily available additional tool for interpreting spirometry.

## INTRODUCTION

Interpretation of spirometry relies on visual pattern recognition and on quantitative comparisons between measured flows and volumes versus predicted reference values derived in healthy subjects from similar populations.<sup>1–3</sup> Previous studies described the diagnostic utility of a global spirometric parameter called area under expiratory flow-volume curve (AEX or AEX-FV)<sup>4–9</sup> for characterising

## Key messages

### What is the key question?

► Area under expiratory flow-volume loop (AEX or AEX-FV) is a spirometric measurement that could differentiate between functional impairments and patterns. How can we derive AEX-FV based on available volumes and flows and make it widely available for clinical practice?

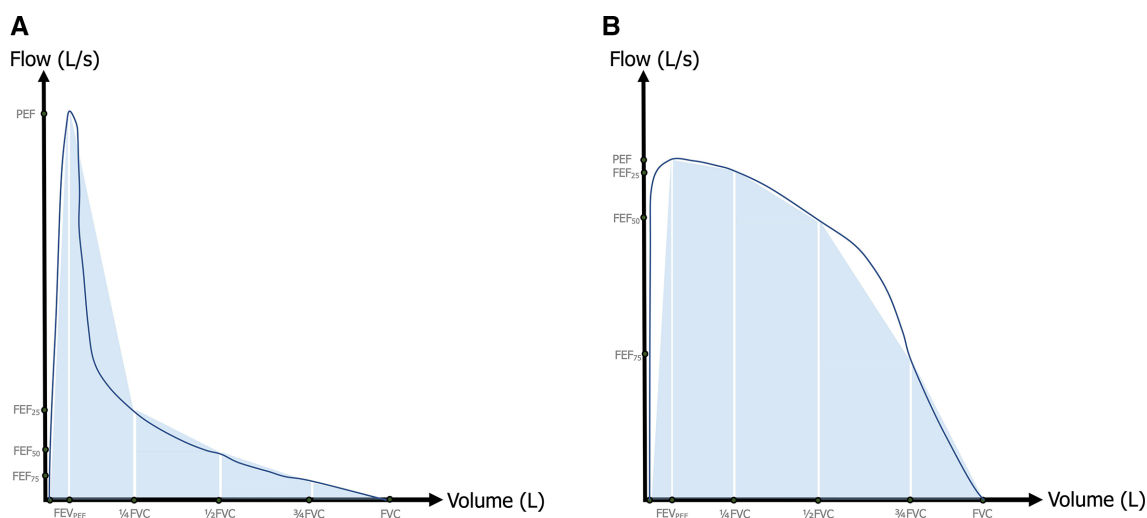
### Why read on?

► Different AEX approximations can be derived using the trapezoid-and-triangle segmentation method, and based on forced vital capacity (FVC), peak expiratory flow (PEF) and several instantaneous isovolumic flows. One estimation is based on FVC, PEF, forced expiratory flow at 25% (FEF<sub>25</sub>), FEF<sub>50</sub>, FEF<sub>75</sub> and forced expiratory volume at PEF (FEV<sub>PEF</sub>) (the latter not often reported or used in spirometry). We developed here several predictive models for AEX-FV based only on FVC, PEF, FEF<sub>25</sub>, FEF<sub>50</sub> and FEF<sub>75</sub>.

### What is the bottom line?

► Whenever actual AEX-FV is not available, estimated AEX-FV could be computed easily and included in the standard spirometry reports, providing additional value for diagnosis and severity stratification of functional impairments. A simple linear regression model for estimating AEX-FV that does not include FEV<sub>PEF</sub> was found to be highly performing and adequate for clinical practice. Due to their higher precision, deep learning models could be used in more advanced epidemiological or research studies.

pulmonary function test (PFT) impairments, potentially lessening the need to perform lung volume testing by body plethysmography or other complex methods. The measurement was found to be particularly useful in separating traditional patterns of normal function, obstruction, restriction, mixed ventilatory defects and small airway disease, in further refining the response to bronchodilators, as well as in stratifying the severity of



**Figure 1** Two examples of how AEX-FV<sub>4</sub> reconstruction or approximation by the triangle-and-trapezoid method using PEF, FEF<sub>25</sub>, FEF<sub>50</sub> and FEF<sub>75</sub> could overestimate (A) or underestimate (B) the actual AEX-FV. AEX-FV, area under expiratory flow-volume curve; FVC, forced vital capacity; FEF<sub>25</sub>, FEF<sub>50</sub> and FEF<sub>75</sub>, forced expiratory flows at 25%, 50% and 75%; PEF, peak expiratory flow.

these impairments, including the possibility to use it as a surrogate marker and severity grading system for airway hyperinflation.<sup>4-9</sup>

While AEX-FV represents the integral function of airflow by expired volume (an easily available measurement in digital spirometry), it is currently made available in the interpretative reports by only one PFT equipment vendor. To overcome this limitation, in previous work we approximated the AEX-FV from forced vital capacity (FVC), peak expiratory flow (PEF), forced expiratory volume at PEF (FEV<sub>PEF</sub>) and several instantaneous isovolumic flows, that is, forced expiratory flows at 25%, 50% and 75% of FVC (FEF<sub>25</sub>, FEF<sub>50</sub> and FEF<sub>75</sub>, respectively).<sup>10</sup> However, the triangle-and-trapezoid segmentation method for approximating AEX-FV could lead to either overestimation (one example in figure 1A) or underestimation (eg, figure 1B) of the actual measurement, which may affect its diagnostic and prognostic value. Further, the FEV<sub>PEF</sub> is rarely included in the standard spirometry reports, hence not easily available in simple equations such as the ones used for AEX<sub>1-4</sub> (the AEX-FV approximations based on 1–4 flows, that is, PEF, FEF<sub>25</sub>, FEF<sub>50</sub> and FEF<sub>75</sub>).<sup>11</sup>

In the current study, we aim to compute the AEX-FV by using various predictive models that consider only widely available parameters such as FVC, PEF, FEF<sub>25</sub>, FEF<sub>50</sub> and FEF<sub>75</sub>, effectively searching for an optimised combination of coefficients for the predictive variables, so that AEX can become a standard reported parameter in spirometry.

## METHODS

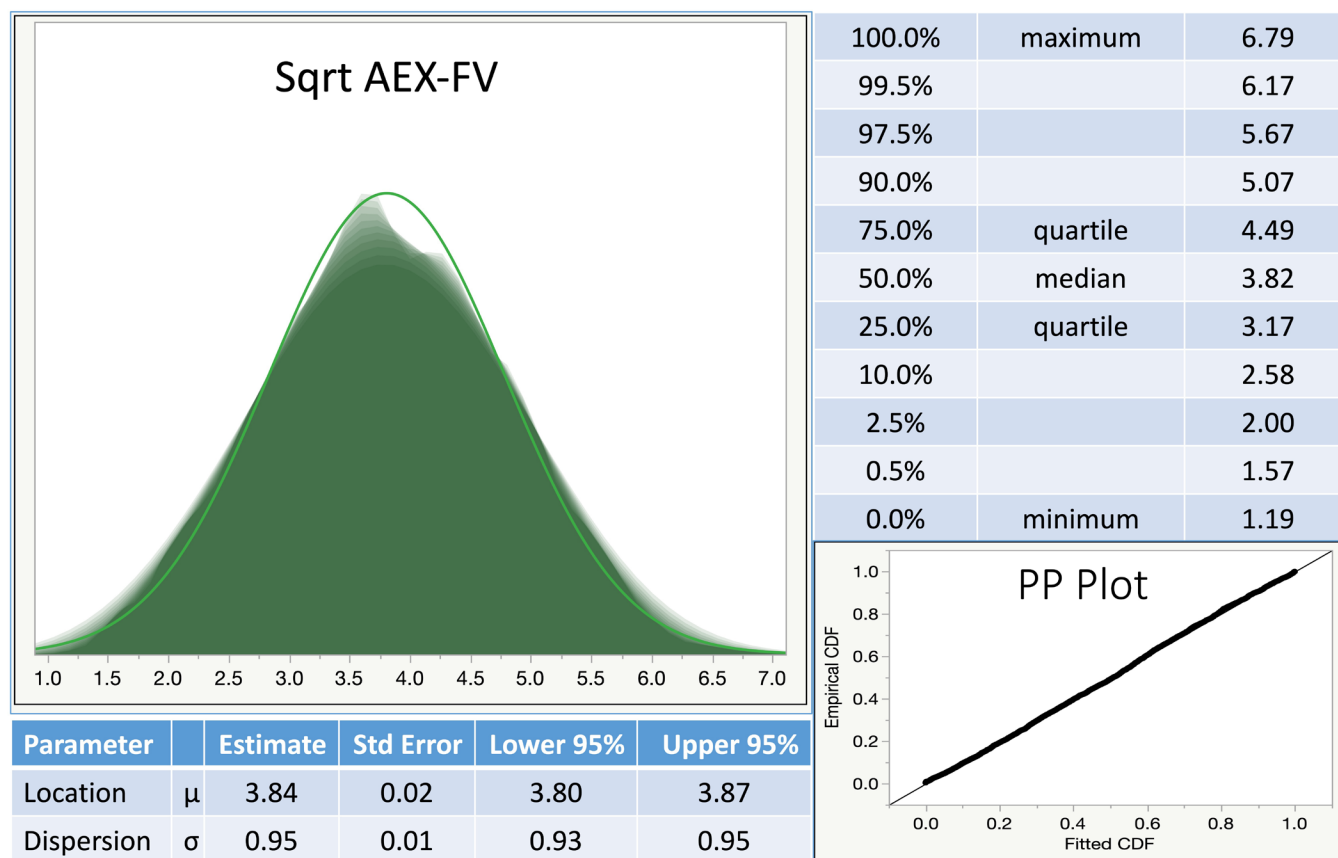
Analyses were performed on a dataset of 4980 normal spirometry tests including AEX-FV values, done in the Cleveland Clinic PFT Laboratories between 2 April 2019 and 17 August 2020. The spirometry tests were performed on 3634 distinct, consecutive adult non-smoking subjects

without any known lung disease or reported respiratory symptoms. In 135 of these ‘normal’ tests, the best FVC, FEV in 1 s (FEV<sub>1</sub>), FEV<sub>1</sub>/FVC or ‘distal’ flow FEF<sub>75</sub> were below the lower limits of normal, as defined by Global Lung Initiative (GLI) equations,<sup>2 12</sup> so they were excluded from further analyses even if they represented normal, isolated or ‘non-specific’ reductions in one of these functional parameters<sup>13</sup>; this exclusion led to a total number of normal spirometry tests analysed of 4845.

Spirometry was performed and interpreted according to the current American Thoracic Society/European Respiratory Society standards and recommendations.<sup>1 14-16</sup> The tests were performed using Jaeger Master Lab Pro systems (Wurzburg, Germany). The most recent, comprehensive and applicable reference values, as published by the GLI were used for spirometry interpretation.<sup>2 12</sup>

Statistical analyses were done in JMP Pro V.15 (SAS Institute). Descriptive statistical analysis of available variables was performed. Categorical variables were presented as frequencies or group percentages. Continuous variables were characterised as mean±SD (for normally distributed variables) or as median and 25–75<sup>th</sup> IQR (for non-normal distributions, as determined by the Shapiro-Wilk or Anderson-Darling test). Student’s t test and analysis of variance were used to compare mean values, while categorical variables were compared using variations of the  $\chi^2$  test. The Tukey-Kramer honest significant differences method was used to compare means among pairs when the variances were similar, while the Wilcoxon or Kruskal-Wallis rank sum tests were performed as non-parametric methods when variances were unequal, as appropriate.

In a prior publication,<sup>10</sup> we defined AEX-FV<sub>4</sub> (also called AEX<sub>4</sub>) by the triangle-and-trapezoid method of segmentation of AEX-FV, and we used for computations the following formula:



**Figure 2** Distribution of the variable square root of AEX-FV (Sqrt AEX-FV) illustrated as a shadowgram and main quantiles; estimates of the distribution and dispersion, and the PP plot for a normal distribution fit. Which shows the relationship between the empirical cumulative distribution function (CDF) and the fitted CDF obtained using the estimated parameters. AEX-FV, area under expiratory flow-volume curve; PP plot, percentile-percentile plot.

$$\text{AEX-FV}_4 = 0.5 \cdot \text{FEV}_{\text{PEF}} \cdot \text{PEF} + 0.5 \cdot (\text{PEF} + \text{FEF}_{25}) \cdot \left(\frac{1}{4} \cdot \text{FVC} - \text{FEV}_{\text{PEF}}\right) + 0.5 \cdot (\text{FEF}_{25} + \text{FEF}_{50}) \cdot \left(\frac{1}{4} \cdot \text{FVC} + 0.5 \cdot (\text{FEF}_{50} + \text{FEF}_{75}) \cdot \frac{1}{4} \cdot \text{FVC} + 0.5 \cdot \text{FEF}_{75} \cdot \frac{1}{4} \cdot \text{FVC}\right)$$

In the current investigation, instead of using the above coefficients (1 or 0.5), we aimed: (1) to identify a combination of coefficients which ‘smooth’ (reduce or augment) the contributions of various flows, and (2) not to rely on  $\text{FEV}_{\text{PEF}}$ , a spirometric variable not often used or included in the standard PFT reports.

The simple linear and generalised (optimised) regression models used the following pool of variables: FEV, PEF, forced expiratory time (FET),  $\text{FEF}_{25}$ ,  $\text{FEF}_{50}$  and  $\text{FEF}_{75}$  to predict square root of AEX-FV (Sqrt AEX-FV, a transformation that lead to a near-normal distribution, figure 2).<sup>17 18</sup> Generalised regression was performed using optimisation techniques such as ridge regression, lasso, elastic net and double lasso, with and without adaptive features. The variables that contributed by <10% to a model (eg, FET) were pruned out from the models; for this, we used k nearest neighbours’ method of dependent variables resampling in JMP. Similarly, interactions that contributed by <5% of the model’s variance were not included in the linear regression models, even if statistically significant (ie, a false discovery rate or FDR <0.001). A random 70/30 partition was generated for

cross-validation of all the models. Statistical significance was defined as  $p < 0.001$ .

Patients or the public were not involved in the design, conduct, reporting or dissemination plans of our research.

Patient and public involvement statement: patients and public were not involved in the design, conduct or analyses of the study results.

## RESULTS

From the total of 4845 normal, best-value spirometry sets performed on 3634 distinct participants (83% men), 2475 (51%) were coded as ‘single’ tests, 1250 (26%) as pre-bronchodilator and 1120 (23%) as postbronchodilator tests. In order to avoid weighing due to multiple tests performed in the same subjects, and to assess the potential impact of bronchodilator administration, we used for models’ development and validation one random spirometry set from each of the 3634 tested subjects. Approximately 71% of the tested individuals were self-identified as white, 17% as black, 0.6% North-East Asians, 0.4% South-East Asians and 11% as other or multiracial. The median age (IQR) was 60 (47–70) years. Median (IQR) height, weight and body mass index were

**Table 1** Simple linear regression model without interactions for predicting square root AEX-FV

Term	Estimate	SE	T ratio	Prob > T	VIF	Main effect	R <sup>2</sup>	RASE
Intercept	-0.070	0.007	-10.027	<0.0001	-	-	T: 0.993	
FVC	0.501	0.002	218.464	<0.0001	2.48	0.24		T: 0.078
PEF	0.062	0.002	37.790	<0.0001	5.24	0.21		
FEF <sub>25</sub>	0.097	0.002	51.527	<0.0001	5.88	0.20	V: 0.993	V: 0.077
FEF <sub>50</sub>	0.118	0.002	47.515	<0.0001	4.98	0.19		
FEF <sub>75</sub>	0.168	0.005	35.138	<0.0001	3.19	0.16		

In both training (T, n=2577) and validation (V, n=1057) sets, model's R<sup>2</sup> was 0.993 and RASE was 0.078–0.077.

AEX-FV, area under expiratory flow-volume curve; FEF<sub>25</sub>, forced expiratory flow at 25%; FVC, forced vital capacity; PEF, peak expiratory flow; RASE, root average square error; VIF, variance inflation factor.

1.74 (1.68–1.80) m, 88 (77–102) kg and 29 (26–34) m/kg<sup>2</sup>, respectively. All tests had available, actual AEX-FV measurements, as provided by the spirometry testing equipment vendor. Mean ( $\mu$ ) Sqrt AEX-FV was 3.81 (95% CI or CI 3.78 to 3.84), and the standard deviation ( $\sigma$ ) was 0.97 (95% CI 0.95 to 0.97) L<sup>2</sup>/s. Median/IQR of actual Sqrt AEX-FV was 3.79/3.12–4.49 L<sup>2</sup>/s. Mean ( $\mu$ , 95% CI)/SD ( $\sigma$ , 95% CI) of Sqrt AEX-FV was 3.89 (3.85–3.92)/0.90 (0.88–0.90) in 'single' tests, 3.64 (3.57–3.72)/1.01 (0.96–1.02) and 3.84 (3.74–3.93)/1.06 (1.00–1.07) L<sup>2</sup>/s, in prebronchodilator and postbronchodilator tests, respectively. By gender, mean ( $\mu$ , 95% CI)/SD ( $\sigma$ , 95% CI) Sqrt AEX-FV was 2.76 (2.72–2.80)/0.65 (0.61–0.65) in women and 4.09 (4.06–4.11)/0.85 (0.83–0.85) L<sup>2</sup>/s in men.

Table 1 and online supplemental figure S1 show the linear regression model that predicts Sqrt AEX-FV using FVC, PEF, FEF<sub>25</sub>, FEF<sub>50</sub> and FEF<sub>75</sub> and assuming

no interactions between the X factors. All five variables performed well in univariate and multivariate models with significance levels <0.0001, and contributed between 16% and 24% in predicting the Y variable Sqrt AEX. While the variance inflation factor was generally lower than 10, one of the assumptions of the regression model was violated, that is, that of normal distribution of the predictors, and their distribution did not 'normalise' with any of the usual transformations. In both testing and validation sets, the percentage of variance explained by the model (R<sup>2</sup>) was very high, 0.993. While some of the interactions between these factors were statistically significant (table 2), their main effects were very small and the overall R<sup>2</sup> remained <0.995, pointing towards the fact that more complex models that include interactions are not necessary (despite collinearities), as the

**Table 2** Linear regression model with interactions for predicting square root AEX-FV

Term	Estimate	Std error	T ratio	Prob > T	VIF	Main effect	R <sup>2</sup>	RASE
Intercept	-0.084	0.007	-11.598	<0.0001	-	-	T: 0.994	T: 0.074
FVC	0.497	0.002	220.145	<0.0001	2.63	0.24		
PEF	0.065	0.002	37.790	<0.0001	6.33	0.20		
FEF <sub>25</sub>	0.096	0.002	47.751	<0.0001	7.41	0.20		
FEF <sub>50</sub>	0.118	0.003	45.661	<0.0001	5.90	0.19		
FEF <sub>75</sub>	0.182	0.005	33.196	<0.0001	4.57	0.16		
(PEF-8.75238)·(FEF <sub>25</sub> -7.10942)	-0.003	0.001	-5.385	<0.0001	5.40			
(PEF-8.75238)·(FVC-4.11503)	-0.002	0.002	-1.448	0.1476	6.74		V: 0.994	V: 0.075
(PEF-8.75238)·(FEF <sub>50</sub> -3.71047)	-0.006	0.002	-2.660	0.0079	22.28			
(PEF-8.75238)·(FEF <sub>75</sub> -1.0233)	0.014	0.005	2.870	0.0041	17.35			
(FEF <sub>25</sub> -7.10942)·(FVC-4.11503)	0.012	0.002	5.441	<0.0001	12.33			
(FEF <sub>25</sub> -7.10942)·(FEF <sub>50</sub> -3.71047)	0.002	0.002	1.148	0.2510	16.86			
(FEF <sub>25</sub> -7.10942)·(FEF <sub>75</sub> -1.0233)	-0.004	0.005	-0.754	0.4506	20.73			
(FVC-4.11503)·(FEF <sub>50</sub> -3.71047)	0.014	0.003	4.513	<0.0001	11.07			
(FVC-4.11503)·(FEF <sub>75</sub> -1.0233)	-0.021	0.005	-3.875	0.0001	6.57			
(FEF <sub>50</sub> -3.71047)·(FEF <sub>75</sub> -1.0233)	-0.019	0.003	-6.096	<0.0001	5.39			

In both training (T, n=2577) and validation (V, n=1057) sets, model's R<sup>2</sup> was ~0.994 and RASE was 0.074–0.075.

AEX-FV, area under expiratory flow-volume curve; FEF<sub>25</sub>, forced expiratory flow at 25%; FVC, forced vital capacity; PEF, peak expiratory flow; RASE, root average square error; VIF, variance inflation factor.

**Table 3** Model Comparisons between linear regression (standard least squares) and various techniques of optimisation (generalised regression) for predicting square root AEX-FV

Response distribution	Estimation method	AICc	BIC	Generalised R <sup>2</sup> (T)	Generalised R <sup>2</sup> (V)	RASE (T)	RASE (V)	Lambda (penalty, T)
Normal	Standard Least Squares	-5843.0	-5802.1	0.993	0.993	0.078	0.077	-
Normal	Ridge	-5843.0	-5802.1	0.993	0.993	0.078	0.077	0.000
Normal	Lasso	-5841.8	-5800.9	0.993	0.993	0.078	0.077	0.071
Normal	Adaptive Lasso	-5843.0	-5802.1	0.993	0.993	0.078	0.077	0.000
Normal	Elastic Net	-5842.2	-5801.3	0.993	0.993	0.078	0.077	0.050
Normal	Adaptive Elastic Net	-5843.0	-5802.1	0.993	0.993	0.078	0.077	0.000
Normal	Double Lasso	-5841.8	-5800.9	0.993	0.993	0.078	0.077	0.071
Normal	Adaptive Double Lasso	-5843.0	-5802.1	0.993	0.993	0.078	0.077	0.000
Gamma	Maximum Likelihood	-1812.5	-1771.5	0.969	0.968	-	-	-
Gamma	Ridge	-1793.6	-1752.7	0.969	0.969	-	-	10.708
Gamma	Lasso	-1808.3	-1767.4	0.969	0.968	-	-	38.550
Gamma	Adaptive Lasso	-1812.5	-1771.5	0.969	0.968	-	-	0.000
Gamma	Elastic Net	-1807.2	-1766.3	0.969	0.968	-	-	42.452
Gamma	Adaptive Elastic Net	-1812.5	-1771.5	0.969	0.968	-	-	0.000
Gamma	Double Lasso	-1808.3	-1767.4	0.969	0.968	-	-	38.550
Gamma	Adaptive Double Lasso	-1812.5	-1771.5	0.969	0.968	-	-	0.000
LogNormal	Maximum Likelihood	-1355.9	-1315.0	0.965	0.966	-	-	-
LogNormal	Ridge	-1339.8	-1298.9	0.965	0.966	-	-	8.279
LogNormal	Lasso	-1353.4	-1312.5	0.965	0.966	-	-	26.082
LogNormal	Adaptive Lasso	-1355.9	-1315.0	0.965	0.966	-	-	0.000
LogNormal	Elastic Net	-1352.4	-1311.5	0.965	0.966	-	-	30.105
LogNormal	Adaptive Elastic Net	-1355.9	-1315.0	0.965	0.966	-	-	0.000
LogNormal	Double Lasso	-1353.4	-1312.5	0.965	0.966	-	-	26.082
LogNormal	Adaptive Double Lasso	-1355.9	-1315.0	0.965	0.966	-	-	0.000

In both training (T, n=2577) and validation (V, n=1057) sets, the models' R<sup>2</sup> was between 0.993 and 0.965 and RASE between 0.078 and 0.077. The lambda penalty coefficient was included, where appropriate.

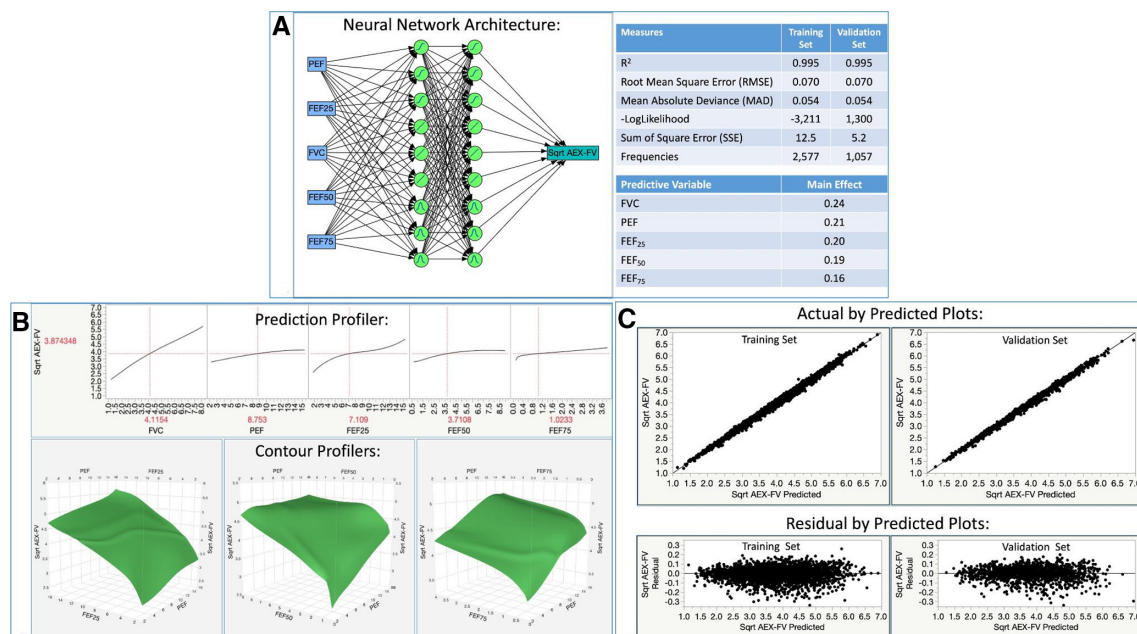
AEX-FV, area under expiratory flow-volume curve; AICc, Akaike Information Criterion; BIC, Bayesian Information Criterion; RASE, root average square error.

model's performance remains high (similar dispersion as expressed by the root average square error between 0.074 and 0.075).

Further attempts to improve the models' performance by using various response variable distributions (eg, gamma or lognormal transformations) and using optimisation (generalised regression) techniques such as ridge regression, elastic net, lasso and double lasso, with or without adaptive features, did not lead to major improvements in generalised R<sup>2</sup> of the validation set (table 3).

To counteract the issue of non-normal distributions and collinearities between variables (assumed by default in regression to be completely independent), we also

developed a neural network model using the same variables (figure 3A–C). We ran several neural network architectures, and the optimised, fastest models were based on a deep learning architecture with one output (Sqrt AEX-FV, maintained as transformed for purpose of comparability with the other models), same five inputs (PEF, FVC, FEF<sub>25</sub>, FEF<sub>50</sub> and FEF<sub>75</sub>) and two hidden layers, each with three sigmoid  $[(e^{2x}-1)/(e^{2x}+1)]$ , 3 gaussian  $(1/e^{x-x})$  and three linear activation function nodes. We used transformed covariates (with either Johnson Su or Johnson Sb distributions), a robust fit method that uses absolute deviations instead of least squares (in order to minimise the effects of outliers), 10 additional tours



**Figure 3** (A) Neural network architecture, model performance and variable importance. The average of the absolute values of the differences between the response and the predicted response. (B) Prediction and contour profilers for the neural network model. (C) Actual and residual Sqrt AEX values versus predicted Sqrt AEX. AEX-FV, area under expiratory flow-volume curve; FVC, forced vital capacity; FEF<sub>25</sub>, FEF<sub>50</sub> and FEF<sub>75</sub>, forced expiratory flows at 25%, 50% and 75%; mean abs dev, mean absolute deviance; PEF, peak expiratory flow; RMSE, root mean square error; SSE, sum of squares error.

for the fitting process, and a squared penalty function to avoid overfitting (the latter being preferred when all X variables are expected to contribute to the predictive ability of the model). Generalised R<sup>2</sup> in this instance simplifies the traditional R<sup>2</sup> for continuous normal responses in the standard least squares setting, and is also known as the Nagelkerke or Craig-Uhler R<sup>2</sup>, which is a normalised version of Cox and Snell's pseudo R<sup>2</sup>.<sup>19</sup> The root mean square error (RMSE) is equivalent to the SD, while Mean Absolute Deviance is represented by the average of the absolute differences between the predicted response and the actual Y variable (figure 3A). The SSE (figure 3A) represents the sum of squares error, while the main effects of the variables is based on the k nearest neighbours' method of resampling dependent variables. Overall, the R<sup>2</sup> was very high (0.995 in both the training and the validation set), while the RMSE was even further reduced to 0.070 (figure 3A). The prediction profiler in figure 3B shows the non-linear relationship between Sqrt AEX-FV and the X variables, while the contour profilers show the response surfaces for various combination of variables versus Sqrt AEX-FV (figure 3B). Another illustration of the neural network model's performance is shown in figure 3C, which showcases the high correlations between Sqrt AEX-FV observed and predicted values, and the residual values plotted against the predicted Sqrt AEX-FV in the training and validation sets, respectively. The figure 3C also shows that the residuals do not increase at higher response variable values (visual inspection confirming that the homoscedasticity condition is met).

The models performed well in both prebronchodilator and postbronchodilator studies, their residuals being very low (~0.001 in regression and ~-0.007 in neural networks) and similar in the two groups. Further, for all predictive models tried, the performance was almost identical in prebronchodilator versus postbronchodilator spirometry tests (data not shown). When we assessed the coefficients of variation for AEX-FV, we found it to be very low: 5 (2–6)% in prebronchodilator tests and 6 (4–12)% in postbronchodilator spirometries, making the promise of a reliable, reproducible functional parameter, that is, with low intraindividual variability.

## DISCUSSION

The AEX-FV has emerged as a useful spirometric tool, able to separate traditional patterns of normal function, obstruction, restriction, mixed ventilatory defects and small airway disease, to characterise the response to bronchodilators, to stratify the severity of these impairments, and to assess airway hyperinflation, without employing more advanced, expensive and laborious lung volume testing.<sup>4–9</sup>

In a previous study, we have shown that computing AEX-FV<sub>4</sub> (or AEX<sub>4</sub>) using the trapezoid-and-triangle method of approximation, and based on PEF, FVC, FEF<sub>25</sub>, FEF<sub>50</sub>, FEF<sub>75</sub> and FEV<sub>PEF</sub> represents a reasonable approximation of the AEX-FV.<sup>10</sup> The main finding of this article is that a simple linear regression model using only PEF, FVC, FEF<sub>25</sub>, FEF<sub>50</sub> and FEF<sub>75</sub> as predictors for Sqrt AEX-FV (table 1) is likely sufficient to estimate the AEX-FV when

the actual measurement is not available. As such, the dependence on spirometric variables that are either not reported ( $FEV_{PEF}$ ) or not made available by the digital spirometry platform (AEX-FV or  $FEV_{PEF}$ ) is easily avoided. Furthermore, especially when advanced computational capabilities are not available, more complex regression or neural network models are likely non-essential for the purpose of this endeavour.

Our previous work showed that AEX compared favourably with several traditional spirometric measurements and parameters in diagnosing physiologic derangement and in estimating the severity of respiratory impairments. The actual AEX was able to predict with reasonable accuracy inspiratory capacity, inspiratory capacity/total lung capacity and residual volume/total lung capacity ratios, thus potentially obviating the need for complex lung volume testing.<sup>4,5</sup> The current investigation extends the value of the AEX-FV concept by showing that it can be easily estimated using widely available spirometric volumes and instantaneous flows. This becomes important especially when existing PFT software does not compute or report the actual AEX values. Further, we recently developed methodologies for predicting AEX-FV values, which allows the clinician to derive the percent predicted values and to further characterise physiological impairment.<sup>20</sup>

The current work also extends prior evaluations of AEX, which have tested the use of this spirometric functional measurement either in paediatric populations, or for assessing bronchoconstriction or bronchodilation responses.<sup>21–24</sup> In a more recent article, the authors constructed predicted  $AEX_4$  (called ‘reference flow-volume loop’) and compared it against actual AEX, thus assessing the degree of airway hyperinflation in adult chronic obstructive pulmonary disease patients.<sup>25</sup>

For any new functional measurement to be used in clinical practice, it is important to establish not only its performance, but also its test-to-test, inherent variability. As such, in a separate sub-analysis of valid and acceptable spirometry tests represented by 93 prebronchodilator and 122 post-bronchodilator trials, with a median of 3 and 5 trials per subject, respectively, we evaluated the inter-test variability of the AEX-FV. The average (range) for the coefficient of variation, defined as  $SD \cdot 100 / \text{mean}$  of AEX-FV was overall low, that is, 6 (1%–11%) in prebronchodilator tests and 7 (1%–20%) in postbronchodilator assessments, respectively. Their ranges, defined as mean differences between the best (largest) and the worst (smallest) AEX-FV values, were 3.5 (95% CI 3.2 to 3.7) and 2.9 (95% CI 2.8 to 3.2)  $L^2/s$  in prebronchodilator and postbronchodilator repeat trials, respectively. The predicted AEX-FV by various models were highly correlated with the actual AEX-FV variable, thus the variability of the predicted AEX-FV was, unsurprisingly, very similar. Given that the relative differences (range  $\cdot 100 / \text{mean}$ ) could be considerable (22% and 21% in prebronchodilator and postbronchodilator tests, respectively), perhaps an expression of the inherent variability of the lung volumes and instantaneous flows that determine

AEX-FV, it is likely best to consider for reporting and analysis purposes the best (largest) obtained values, similar to the rules applied to  $FEV_1$ , FVC and their ratio.

One of the strengths of our investigation is represented by the large dataset of normal PFTs extracted from a laboratory-based population of non-smokers without any known respiratory symptoms or pathologies, and for which actual AEX-FV were made available by the digital spirometry platform. In order to assess more advanced modelling capabilities, we employed both regular and optimised regression models and several simplified models of machine or deep learning. One weakness of this investigation is the single-centre collection of PFTs; as such, these findings will need to be validated externally in other populations. Additionally, some PFT laboratory-based ‘normal’ populations may include various uncaptured, relevant pathologies or could hide early functional impairments, while still above the lower limit of normal cut-offs. This may affect the population-level estimations or normality. Further, some demographic groups are under-represented in our cohort; as such, our models require further evaluation using broader populations and the models may require further refinement by including gender, race or ethnicity predictor factors.

## CONCLUSION

This study analyses the performance of several models of estimating the AEX-FV, based on FVC, PEF and several isovolumic flows, that is,  $FEF_{25}$ ,  $FEF_{50}$  and  $FEF_{75}$ . We found that AEX-FV estimated by a simple regression model without interactions performs with acceptable accuracy as a surrogate marker of actual AEX, which makes it potentially useful in diagnosing physiologic derangement of pulmonary function and in stratifying the severity of such impairment in clinical practice. In more advanced epidemiological or research settings, the use of a neural networks or deep learning models may be preferred due to their higher precision.

**Collaborators** None.

**Contributors** OCI, JAR, MH and JKS contributed to the writing of this article; JAR and MH contributed with data extraction; OCI contributed with statistical analyses.

**Funding** The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interests** None declared.

**Patient consent for publication** Not required.

**Ethics approval** Institutional research ethical oversight approval was obtained to conduct the study (Cleveland Clinic IRB EX#0504 and Emory IRB# 00049576/Atlanta Veteran Affairs Medical Center R&D loachimescu-002).

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** No data are available. Data sharing: in place in-between investigators’ institutions

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines,

terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

#### ORCID iD

Octavian C Ioachimescu <http://orcid.org/0000-0001-9047-6894>

#### REFERENCES

- Standardization of spirometry, 1994 update. American thoracic Society. *Am J Respir Crit Care Med* 1995;152:1107–36.
- Quanjer PH, Stanojevic S, Cole TJ, et al. Multi-Ethnic reference values for spirometry for the 3–95-yr age range: the global lung function 2012 equations. *Eur Respir J* 2012;40:1324–43.
- Staitieh BS, Ioachimescu OC. Interpretation of pulmonary function tests: beyond the basics. *J Investig Med* 2017;65:301–10.
- Ioachimescu OC, Stoller JK. An alternative spirometric measurement. area under the expiratory flow-volume curve. *Ann Am Thorac Soc* 2020;17:582–8.
- Ioachimescu OC, McCarthy K, Stoller JK. Alternative measurements to aid interpretation of spirometry: the role of AREA-UNDER-THE-EXPIRATORY flow-volume curve. *Chest* 2006;130:119S.
- Vermaak JC, Bunn AE, de Kock MA. A new lung function index: the area under the maximum expiratory flow-volume curve. *Respiration* 1979;37:61–5.
- Bunn AE, Vermaak JC, De Kock MA. A comprehensive on-line computerised lung function screening test. *Respiration* 1979;37:42–51.
- Bunn AE, De Brandt HM, Vermaak JC. Analogue device for measurement of area under the maximum expiratory flow-volume curve. *Med Biol Eng Comput* 1979;17:695–6.
- Bunn AE, Vermaak JC, De Kock MA. The area under the maximum expiratory flow-volume curve. *Lung* 1978;155:72–3.
- Ioachimescu OC, Stoller JK. Area under the expiratory flow-volume curve (AEX): actual versus approximated values. *J Investig Med* 2020;68:403–11.
- National Center for Health Statistics (NCHS). National health and nutrition examination survey (NHANES) data department of health and human services, centers for disease control and prevention (CDC).
- Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S. population. *Am J Respir Crit Care Med* 1999;159:179–87.
- Iyer VN, Schroeder DR, Parker KO, et al. The nonspecific pulmonary function test: longitudinal follow-up and outcomes. *Chest* 2011;139:878–86.
- Lung function testing: selection of reference values and interpretative strategies. American thoracic Society. *Am Rev Respir Dis* 1991;144:1202–18.
- Miller MR, Hankinson J, Brusasco V, et al. Standardisation of spirometry. *Eur Respir J* 2005;26:319–38.
- Pellegrino R, Viegi G, Brusasco V, et al. Interpretative strategies for lung function tests. *Eur Respir J* 2005;26:948–68.
- Steyerberg EW. *Clinical prediction models*. Springer Science + Business Media, LLC, 2009.
- Saltelli A. Making best use of model evaluations to compute sensitivity indices. *Comput Phys Commun* 2002;145:280–97.
- Nagelkerke NJD. A note on a general definition of the coefficient of determination. *Biometrika* 1991;78:691–2.
- Ioachimescu OC, Stoller JK, Garcia-Rio F. Area under the expiratory flow-volume curve: predicted values by artificial neural networks. *Sci Rep* 2020;10:16624.
- Zapletal A, Chalupová J. Forced expiratory parameters in healthy preschool children (3–6 years of age). *Pediatr Pulmonol* 2003;35:200–7.
- Zapletal A, Hladíková M, Chalupová J, et al. Area under the maximum expiratory flow-volume curve—a sensitive parameter in the evaluation of airway patency. *Respiration* 2008;75:40–7.
- Sovijarvi AR. Flow-volume response to inhaled methacholine in asthmatics; comparison of area under the curve (AFV) with conventional parameters. *Eur J Respir Dis Suppl* 1986;143:18–21.
- Stein D, Stein K, Ingris S. [Aex - the area under the expiratory flow-volume loop]. *Pneumologie* 2015;69:199–206.
- Das N, Topalovic M, Aerts J-M, et al. Area under the forced expiratory flow-volume loop in spirometry indicates severe hyperinflation in COPD patients. *Int J Chron Obstruct Pulmon Dis* 2019;14:409–18.