ONLINE SUPPLEMENTARY MATERIAL

Reference equations for pulmonary diffusing capacity using segmented regression show similar predictive accuracy as GAMLSS models

[1]Gerald S. Zavorsky Ph.D., [2]Jiguo Cao Ph.D.

[1]Pulmonary Services, University of California, Davis, Medical Center, Sacramento, California, United States

[2]Department of Statistics and Actuarial Science, Simon Fraser University, Burnaby, British Columbia, Canada

This supplement has been peer-reviewed.

Gerald Zavorsky: ORCID ID: https://orcid.org/0000-0002-4473-1601

How simulated raw data was created from the paper by Munkholm and colleagues

After repeated unsuccessful attempts to obtain the raw data from Munkholm and colleagues[1], it was decided to simulate their data so that the reference equations for pulmonary diffusing capacity, originally published in 2017 in the European Respiratory Journal[2], could be updated. Only adults were used in their study[1], so this newly created or "simulated" raw data had the same number of adult women and adult men as their paper (142 females, 140 males). The age data (range = 18 to 97 years of age) was generated from the normal distribution with their mean and standard deviation (53 ± 23 years old for women, 54 ± 22 years old for men) using a statistical technique called truncation[3]. A truncated sample can be thought of as being equivalent to an underlying sample with all values outside the bounds entirely omitted. The "*rtruncnorm*" library (version1.0-8) in R language (https://www.r-project.org/) was used for random generation of truncated normally distributed values. There is an approximate correlation of 0.6 between height and weight when both sexes are combined in previous studies that generated reference equations for DLNO in adult subjects[4-6](pooled data), the mean and SD for height (165 ± 7 cm for women, 179 ± 8 cm for men) and weight (64.6 ± 9.0 kg for women, 78.5 ± 11.0 cm for men) were generated from a bivariate normal distribution. The mean and standard deviation for males and females from Munkholm *et al.*[1] and the correlation coefficient ~0.6 was truncated by the range of height (range = 149-184 cm in women, 156-198 cm in men) and weight (45-97 kg in women, and 52 to 109 kg in men). The data for DLCO, DLNO, and VA are generated from the linear model from their paper[1], and the residuals are generated from the normal distribution with mean 0 and the standard deviation provided for its residual standard errors[1].

Only 25% of the total pooled data (272 out of 1076 subjects) included the simulated Munkholm data. Of all the subjects that had their lung function tested using the Jaeger MasterScreen Pro or MasterLab Pro lung function device, 40% of these subjects are from the simulated Munkholm data (see **Figure 2** in their article). The results of the accuracy of the simulated data compared to their prediction equations are presented in the following four pages (**Tables S1, S2A, S2B, S3**).

Table S1. A comparison between anthropometric characteristics generated from simulated data created from Munkholm and colleagues[1], and the actual data.

Female (N=142)	**Munkholm *et al.* (2018)**	**Simulated raw data created from Munkholm's population characteristics**
Age (yrs)	53 (23) [18-97]	55 (18) [20-95]
Height (cm)	165 (7) [149-184]	166 (7) [152-182]
Weight (kg)	64.6 (9.0) [45.1-97.0]	65.0 (9.0) [45.1-92.2]
BMI (kg/m^2)	23.6 (2.7) [18.1-29.8]	23.5 (3.0) [15.9 to 32.1]
Male (N=140)		
Age (yrs)	54 (22) [18-97]	55 (19) [19-95]
Height (cm)	179 (8) [156-198]	179 (8) [160-197]
Weight (kg)	78.5 (11.0) [52.0-108.8]	78.4 (11.0) [54.7-100.7]
BMI (kg/m^2)	24.4 (2.6) [18.0-30.0]	24.4 (2.8) [18.0-32.5]

Mean and SD. The SD are in parentheses; the range is in brackets.

Table S2A. The simulated raw data created from Munkholm and colleagues[1] female study population characteristics is nearly equivalent to their actual raw data.

Female	Age	Age^2	height	Constant	R^2	SEE
DLNO Munkholm *et al.* (2018) (mL/min/mmHg)		-0.00753	0.766	-2.36	0.80	11.4
DLNO based on simulated raw data from Munkholm's (2018) population characteristics (mL/min/mmHg)		-0.00769	0.799	-7.74	0.66	12.0
DLCO Munkholm *et al*. (2018) (mL/min/mmHg)		-0.00166	0.192	-3.58	0.77	2.8
DLCO based on simulated raw data from Munkholm's (2018) population characteristics (mL/min/mmHg)		-0.00170	0.251	-12.84	0.67	2.8
VA Munkholm *et al.* (2018) (L)	0.039	-0.000426	0.047	-3.55	0.53	0.49
VA based on simulated raw data from Munkholm's (2018) population characteristics (L)		-0.000126	0.035	-0.48	0.30	0.54

DLNO = pulmonary diffusing capacity for nitric oxide; DLCO = pulmonary diffusing capacity for carbon monoxide; VA = alveolar volume.

Table S2B. The simulated raw data created from Munkholm and colleagues[1] male study population characteristics is nearly equivalent to their actual raw data.

Male	Age	Age^2	height	Constant	R^2	SEE
DLNO Munkholm *et al.* (2018) (mL/min/mmHg)		-0.0125	0.97	5.72	0.82	16.6
DLNO based on simulated raw data from Munkholm's (2018) population characteristics (mL/min/mmHg)		-0.0119	1.151	-7.74	0.75	15.3
DLCO Munkholm *et al*. (2018) (mL/min/mmHg)		-0.00258	0.252	-5.01	0.81	3.7
DLCO based on simulated raw data from Munkholm's (2018) population characteristics (mL/min/mmHg)		-0.00248	0.22	0.14	0.68	3.7
VA Munkholm *et al.* (2018) (L)	0.0387	-0.000442	0.077	-7.90	0.59	0.69
VA based on simulated raw data from Munkholm's (2018) population characteristics (L)		-0.000090	0.073	-6.16	0.48	0.64

DLNO = pulmonary diffusing capacity for nitric oxide; DLCO = pulmonary diffusing capacity for carbon monoxide; VA = alveolar volume.

Table S3. Reference equations from Munkholm *et al.* [1] are compared to estimated equations based on simulated raw data created from Munkholm's study population characteristics.

	DLNO	DLCO	VA
Correlation coefficient	0.99	0.99	0.99
Coefficient of Variation (%)	1.4%	1.2%	2.1%
Bland-Altman Differences 95% CI (mL/min/mmHg)	-4 to +5	-0.9 to +0.9	-0.34 to +0.34
Paired *t*-test differences (mL/min/mmHg)	+0.5 (2.4)	0 (0.5)	0 (0.2)

For the paired *t*-test differences, the numbers in parentheses indicate the SD. DLNO = pulmonary diffusing capacity for nitric oxide; DLCO = pulmonary diffusing capacity for carbon monoxide; VA = alveolar volume.

Table S4. Reference equations using segmented regression with sex and equipment devices included in the equation.

DLCO (n = 1076) (mL/min/mmHg) Breakpoint = 24.3 yrs old [95% CI = 23.1 to 25.5 yrs old] AIC = 5801; BIC = 5841	**Estimate**	**SE**	**95% CI**	**Adjusted R^2**	**RSE**
$Intercept_1$ (for 5.0 to 24.2 yrs old)	-19.49	1.46	-12.4, -16.6	0.81	2.86_1
$Intercept_2$ (for 24.3 to 95.0 yrs old)	-6.74				3.90_2
Age^2_1 (for 5.0 to 24.2 yrs old)	0.0194	0.0017	0.016, 0.023		
Age^2_2 (for 24.3 to 95.0 yrs old)	-0.0021	0.00007	-0.002, -0.002		
Height (cm)	0.223	0.011	0.20, 0.24		
Altitude (m)	0.005	0.00106	0.003, 0.007		
Sex	4.86	0.25	4.38, 5.34		
DLNO (n = 1076) (mL/min/mmHg) Breakpoint = 23.5 yrs old [95% CI = 21.9 to 24.9 yrs old] AIC = 9044; BIC = 9084	**Estimate**	**SE**	**95% CI**	**Adjusted R^2**	**RSE**
$Intercept_1$ (for 5.0 to 23.4 yrs old)	-108.1	6.80	-121.4, -94.8	0.84	13.31_1
$Intercept_2$ (for 23.5 to 95.0 yrs old)	-61.0				17.36_2
Age^2_1 (for 5.0 to 23.4 yrs)	0.0755	0.0092	0.058, 0.094		
Age^2_2 (for 23.5 to 95.0 yrs old)	-0.010	0.003	-0.016, 0.004		
Height (cm)	1.17	0.052	1.07, 1.27		
PFT equipment	18.0	1.21	15.6, 20.4		
Sex	23.9	1.12	21.7, 26.1		
VA (n = 1076) (L) Breakpoint = 29.5 yrs old [95% CI = 27.7 to 31.2 yrs old] AIC = 2018; BIC = 2058	**Estimate**	**SE**	**95% CI**	**Adjusted R^2**	**RSE**
$Intercept_1$ (for 5.0 to 29.4 yrs old)	-5.93	0.243	-6.41, -5.45	0.86	0.52_1
$Intercept_2$ (for 29.5 to 95.0 yrs old)	-4.39				0.67_2
Age^2_1 (for 5.0 to 29.4 yrs old)	0.0016	0.0002	0.001, 0.002		
Age^2_2 (for 29.5 to 95.0 yrs old)	-0.00015	0.00001	0.000, 0.000		
Height (cm)	0.060	0.002	0.056, 0.064		
PFT equipment	0.280	0.045	0.192, 0.368		
Sex	0.76	0.042	0.68, 0.85		

SE = standard error, RSE = residual standard error. AIC = Akaike information criterion; BIC = Bayesian information criterion; For the PFT equipment, 1 = Hyp'Air Compact, and 0 = Jaeger MasterScreen Pro. Sex, 1 = male, 0 = female. Only one age^2 and one intercept is used, depending on the age of the subject. Age^2_1 and $intercept_1$ is used before the breakpoint. Age^2_2 and $intercept_2$ is used at the breakpoint and beyond. For example, for a male who is 23.0 years old, 180 cm tall, and had his measurements performed on the Hyp'Air device, the predicted DLNO (mL/min/mmHg) = $0.0755\cdot(23^2) + 1.17\cdot(180) + 18\cdot(1) + 23.9\cdot(1) - 108.1 = 184.3$ mL/min/mmHg with an LLN of $184.3 - (13.31\cdot1.645) = 162.4$ mL/min/mmHg. For a male 24 years old, with the same height and equipment used, the predicted DLNO (mL/min/mmHg) = $-0.01\cdot(24^2) + 1.17\cdot(180) + 18\cdot(1) + 23.9\cdot(1) - 61.0 = 185.7$ mL/min/mmHg with an LLN of $185.7 - (17.36\cdot1.645) = 157.1$ mL/min/mmHg.

NOTE: If the DLCO is measured at an altitude that is more than 300 meters, we recommend converting the measured DLCO to sea level first, based on the data by Gray and colleagues[7], and then omitting the altitude covariate from the equation (as the converted DLCO will be adjusted to an altitude of 0 meters). Adjusting the DLCO to sea level (mL/min/mmHg) = measured DLCO at altitude · (0.505 + 0.00065·barometric pressure in mmHg at altitude). The formula to estimate barometric pressure at altitude in mm Hg is: 760·exp[– 0.284·altitude in meters / (8.314·Temperature in Kelvin)], where Kelvin = ºC + 273.15. (see: https://planetcalc.com/938/)

Table S5. Mean z-scores, skewness, and mean percent predicted values fitted to the reference equations.

Fitted z-scores [mean (SD)]	**GAMLSS**	**Segmented linear regression**	**GLI**
$DLCO_{MasterScreen}$	-0.03 (0.95) [-0.02]	-0.03 (0.90) [-0.03]	0.28 (0.90) [0.05]
$DLCO_{Hyp'Air}$	0.11 (1.08) [-0.03]	0.04 (1.38) [0.07]	0.99 (1.11) [-0.27]
$DLNO_{MasterScreen}$	-0.02 (1.02) [0.03]	-0.01 (0.88) [-0.12]	N/A
$DLNO_{Hyp'Air}$	-0.01 (1.00) [-0.06]	0.01 (1.17) [0.03]	N/A
$VA_{MasterScreen}$	-0.01 (0.98) [-0.02]	0.04 (0.93) [0.23]	0.02 (0.91) [0.13]
$VA_{Hyp'Air}$	-0.03 (1.04) [0.01]	0.05 (1.09) [0.18]	0.63 (1.09) [0.12]
Fitted % predicted [Mean (SD)]	**GAMLSS**	**Segmented linear regression**	**GLI**
$DLCO_{MasterScreen}$	100 (12)	100 (12)	105 (14)
$DLCO_{Hyp'Air}$	102 (15)	100 (15)	116 (18)
$DLNO_{MasterScreen}$	100 (11)	100 (12)	N/A
$DLNO_{Hyp'Air}$	100 (14)	100 (14)	N/A
$VA_{MasterScreen}$	100 (10)	101 (14)	101 (10)
$VA_{Hyp'Air}$	100 (11)	100 (11)	107 (12)

n = 691 for the Jaeger MasterScreen Pro; n = 385 for the Hyp'Air Compact. DLNO = pulmonary diffusing capacity for nitric oxide; DLCO = pulmonary diffusing capacity for carbon monoxide; VA = alveolar volume; N/A = not applicable. Brackets represent the skewness, and parentheses represent the SD.

Table S6. Pearson's r Correlational matrix of fitted z-scores between models.

	DLCO z-scores GAMLSS	**DLCO z-scores segmented linear regression**	**DLCO z-scores GLI**	**DLNO z-scores GAMLSS**	**DLNO z-scores segmented linear regression**	**VA z-scores GAMLSS**	**VA z-scores segmented linear regression**	**VA z-scores GLI**
DLCO z-scores GAMLSS		0.94	0.82	0.62	0.65	0.44	0.43	0.41
DLCO z-scores segmented linear regression	0.94		0.80	0.59	0.69	0.42	0.45	0.40
DLCO z-scores GLI	0.82	0.80		0.51	0.54	0.35	0.33	0.48
DLNO z-scores GAMLSS	0.62	0.59	0.52		0.93	0.46	0.43	0.42
DLNO z-scores segmented linear regression	0.65	0.69	0.54	0.93		0.47	0.49	0.44
VA z-scores GAMLSS	0.44	0.42	0.35	0.46	0.47		0.95	0.91
VA z-scores Segmented linear regression	0.43	0.45	0.33	0.43	0.49	0.95		0.88
VA z-scores GLI	0.41	0.40	0.48	0.42	0.44	0.91	0.88	

$p < 0.001$ for all. The correlations between the Global Lung Function Initiative (GLI) fitted z-scores (http://gli-calculator.ersnet.org/index.html), and all others include 1065 subjects as GLI fitted z-scores do not go above 85 years of age. Otherwise, 1076 subjects. DLNO = pulmonary diffusing capacity for nitric oxide; DLCO = pulmonary diffusing capacity for carbon monoxide; VA = alveolar volume. Note: these correlations are based on non-diseased subjects, so the range of z-scores are narrow (about -3.0 to +3.0). If diseased subjects were included then the range of z-scores would be wider and the correlations would likely be higher than what is shown, here.

Table S7. Pearson's r correlations between measured variables and the fitted predicted values from various models.

	Measured DLCO	Measured DLNO	Measured VA
Predicted DLCO segmented linear regression	0.91	0.92	0.88
Predicted DLCO GAMLSS	0.91	0.91	0.88
Predicted DLNO segmented linear regression	0.91	0.93	0.86
Predicted DLNO GAMLSS	0.90	0.93	0.87
Predicted VA segmented linear regression	0.86	0.85	0.93
Predicted VA GAMLSS	0.86	0.85	0.94
Predicted DLCO GLI	0.88	0.88	0.88
Predicted VA GLI	0.84	0.83	0.93

$p < 0.001$ for all. Both equipment devices were combined. DLNO = pulmonary diffusing capacity for nitric oxide; DLCO = pulmonary diffusing capacity for carbon monoxide; VA = alveolar volume. n = 1076 for GAMLSS and segmented linear regression; n = 1065 for GLI. The correlations between the predicted values calculated by the Global Lung Function Initiative (GLI) (http://gli-calculator.ersnet.org/index.html) and the current GAMLSS and segmented regression models were based on 1052 subjects as GLI predictions do not go above 85 years of age.

Table S8. Receiver-Operating Characteristic (ROC) analysis for evaluating the performance of DLNO compared to DLCO when the estimated prevalence of an abnormal result in a population is five percent (i.e., when 5% of the population is below the LLN based on DLCO).

	Segmented regression	GAMLSS
Area under the ROC Curve (AUC)	0.68 [0.65, 0.70]	0.69 [0.66, 0.72]
Youden's J statistic	0.35 [0.24, 0.47]	0.38 [0.25, 0.50]
Sensitivity	0.38 [0.26, 0.50]	0.42 [0.29, 0.55]
Specificity	0.97 [0.96, 0.98]	0.97 [0.95, 0.98]
Positive predictive value	40% [30%, 52%]	38% [29%, 49%]
Negative predictive value	97% [96%, 97%]	97% [96%, 98%]
Positive likelihood ratio	12.7 [8.0, 20.2]	11.8 [7.6, 18.2]
Negative likelihood ratio	0.64 [0.53, 0.77]	0.60 [0.49, 0.75]

AUC = The percent chance that when using DLCO to detect an abnormal result, the DLNO can also distinguish an abnormal result in the same patient. Brackets represent the 95% bootstrapped confidence interval.

Youden's J statistic (sensitivity + specificity – 1): Measures the effectiveness of using DLNO as a diagnostic test compared to DLCO.

Sensitivity (true positive rate): Probability that DLNO is abnormal (< LLN) when DLCO is also abnormal (< LLN).

Specificity (true negative rate): Probability that DLNO is normal (≥ LLN) when DLCO is also normal (≥ LLN).

Positive predictive value (precision): Probability DLCO is abnormal (< LLN) when DLNO is also abnormal (< LLN).

Negative predictive value: Probability that DLCO is normal (≥ LLN) when DLNO is also normal (≥ LLN).

Positive likelihood ratio (True positive rate ÷ false positive rate): The ratio between the probability that DLNO is abnormal (< LLN) given that DLCO is abnormal (< LLN) and the probability that DLNO is abnormal (< LLN) given that DLCO is normal (≥ LLN).

Negative likelihood ratio (false-negative rate ÷ true negative rate): The ratio between the probability that DLNO is normal (≥ LLN) given that the DLCO is abnormal (< LLN) and the probability that the DLNO is normal (≥ LLN) given that the DLCO is normal (DLCO is ≥ LLN).

Table S9. The approximate percent predicted at the lower limit of normal and its interquartile range varies across age groups between the two different models.

Segmented Regression	**5-7 years (n=51)**	**8-10 years (n=93)**	**11-50 years (n=594)**	**51-70 years (234)**	**71-95 years (n=104)**
DLNO	68 [61-73]	78 [74-80]	82 [81-84]	78 [76-79]	71 [67-74]
DLCO	66 [60-70]	75 [72-77]	82 [80-86]	77 [76-79]	71 [67-79]
VA	64 [55-70]	76 [73-79]	83 [82-86]	81 [80-82]	79 [77-81]
GAMLSS	**5-7 years (n=51)**	**8-10 years (n=93)**	**11-50 years (n=594)**	**51-70 years (234)**	**71-95 years (n=104)**
DLNO	83 [83-83]	83 [83-83]	83 [77-83]	82 [77-83]	82 [77-83]
DLCO	84 [80-84]	83 [81-84]	82 [80-83]	79 [77-80]	74 [72-75]
VA	87 [85-87]	86 [85-87]	85 [84-85]	85 [81-85]	81 [80-85]

The 25th to 75th percentile of the percent predicted at the LLN is a weighted average in brackets with its median outside. These values correspond to a z-score of –1.645. Notice that in the segmented regression models, the youngest and oldest age groups have a much lower median percent predicted compared to the middle age groups for DLNO and DLCO. For VA (segmented regression), the youngest age group is much different compared to all other age groups. With the GAMLSS models, the percent predicted at the LLN is very consistent for all age groups except for DLCO.

Table S10. The US Department of Health median values for height at each age for white subjects[8].

Age (yrs)	Height female (cm)		Age (yrs)	Height male (cm)
5	112.1		5	112.4
6	119.3		6	118.0
7	123.7		7	126.1
8	129.8		8	131.8
9	136.5		9	136.4
10	142.3		10	141.1
11	150.8		11	148.3
12	154.3		12	153.9
13	157.7		13	163.6
14	161.2		14	170.0
15	160.0		15	172.7
16	161.7		16	172.6
17	162.4		17	174.9
18	162.3		18	175.5
19	161.2		19-39	178.0
20-39	164.4		40-59	177.5
40-59	163.1		60-95	174.3
60-95	159.8			

Table S11. A worked example in traditional units for calculating the predicted, percent predicted, and DLCO z-score in a hypothetical individual using the GAMLSS model.

A white female 65 yrs old, 167 cm tall) has a measured DLCO of 17.0 mL/min/mmHg at 300-m elevation. To predict the DLCO value for her age, sex, height, ethnicity, and elevation, see the reference equation in Table 2B of the main article and see the supplementary excel spreadsheet for the spline tables.

From the excel tables, the Mspline for DLCO, female = – 0.458, and the Sspline for DLCO, female = 0.5711

M (predicted value) = exp[-4.481 + 1.406· ln(height) + 0.194·ln(age) + 0.0002·(altitude) + Mspline]
M = exp[-4.481 + 1.406·ln(167) + 0.194·ln(65) +0.0002·(300) – 0.458] = exp(3.1267)
M (predicted value) = **22.8 mL/min/mmHg**

S (variability around the median) = exp[0.642·ln(age) – 1.018·ln(height) + Sspline]
S = exp[0.642·ln(65) – 1.018·ln(167) + 0.5711] = exp(-1.959)
S = **0.141**

L (index of skewness) = **0.325**

% predicted = (measured/M) ·100
% predicted = (17.0/22.8) ·100 = **74.6%**

Lower limit of Normal (LLN) (5th percentile) = exp[ln(M) + ln(1 – 1.645·L·S)/L)]
Lower limit of Normal (LLN) (5th percentile) = exp[ln(22.8) + (ln(1 – 1.645·0.325·0.141)/0.325)]
Lower limit of Normal (LLN) (5th percentile) = exp[3.1268 + (ln(0.9246)/0.325)]
Lower limit of Normal (LLN) (5th percentile) = exp[3.1268 + (– 0.0784/0.325)]
Lower limit of Normal (LLN) (5th percentile) = exp[3.1268 – 0.2412] = exp(2.8856)
Lower limit of Normal (LLN) (5th percentile) = **17.9 mL/min/mmHg**

Z-score = $[(measured/M)^{L} – 1]/(L·S)$
Z-score = $[(17.0/22.8)^{0.325} – 1]/(0.325·0.141)$
Z-score = $[(0.7456)^{0.325} – 1]/0.0458$
Z-score = [0.909-1]/0.0458 = **–1.987.**
Her z-score is below the LLN for the 5th percentile and 2.5th percentile. Her z-score of –1.987 is more negative than a z-score of –1.645 AND –1.96.

NOTE: If the DLCO is measured at an altitude that is more than 300 meters, we recommend converting the measured DLCO to sea level first based on the data by Gray and colleagues[7], and then omitting the altitude covariate from the equation (as the converted DLCO will be at an altitude of 0 meters). Adjusted DLCO to sea level (mL/min/mmHg) = measured DLCO at altitude·(0.505 + 0.00065·barometric pressure in mmHg at altitude). The formula to estimate barometric pressure at altitude in mm Hg is: 760·exp[– 0.284·altitude in meters / (8.314·Temperature in Kelvin)], where Kelvin = ºC + 273.15. (see: https://planetcalc.com/938/).

Figure S1. The DLNO to DLCO ratio across age groups and lung function devices. The ratio is relatively independent of age. Those that had the DLNO/DLCO ratio measured on the Hyp'Air Compact device (n=385, solid **green** bars) had a higher ratio compared to those that had the ratio using the Jaeger MasterScreen Pro (n=691, solid **blue** bars) (4.87 ± 0.44 vs. 4.58 ± 0.57). The Hyp'Air device provided a ratio that was 0.23 to 0.35 units higher than the Jaeger MasterScreen Pro (95% CI, $p < 0.001$). This is because the Hyp'Air Compact measures DLNO that is about 16 to 20 mL/min/mmHg more compared to the Jaeger MasterScreen Pro, yet both devices have similar DLCO values. Open circles are outliers [3rd quartile +1.5·interquartile range or 1st quartile – 1.5·interquartile range]. Extreme outliers are presented by an asterisk [more than the 3rd quartile +3·interquartile range or 1st quartile – 3·interquartile range].

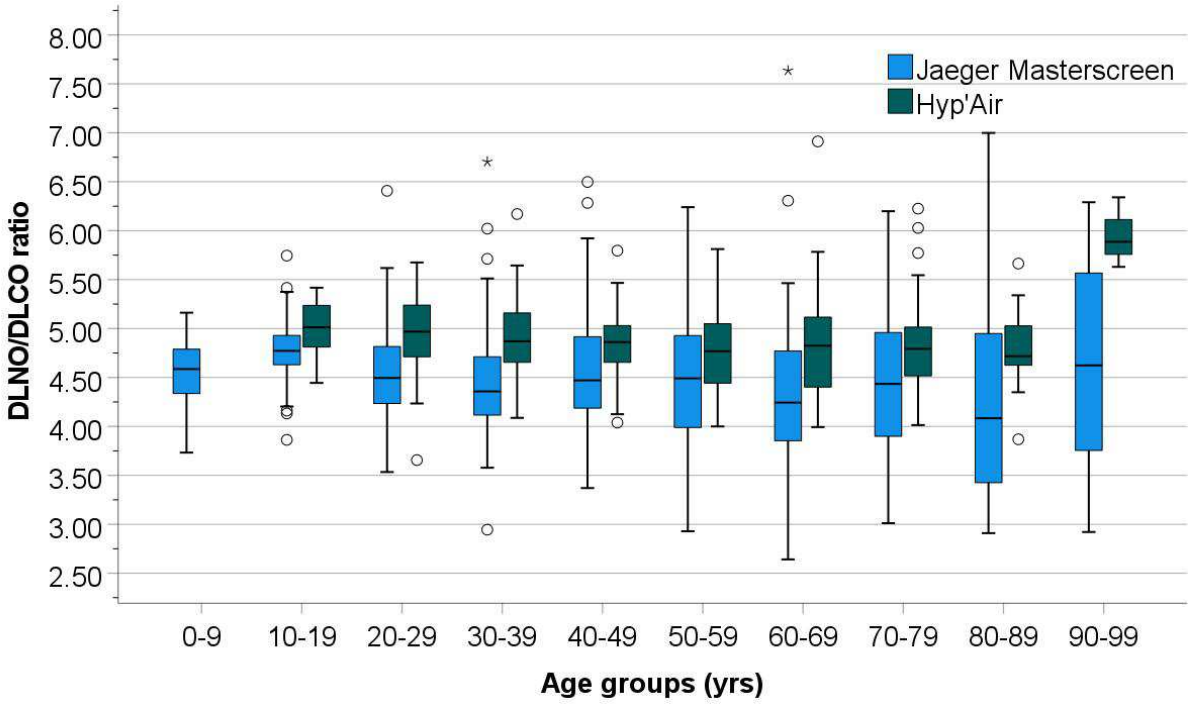


Figure S2. Fitted z-scores and % predicted DLNO, DLCO, and VA using segmented linear regression models. Both devices are combined (n=1076).

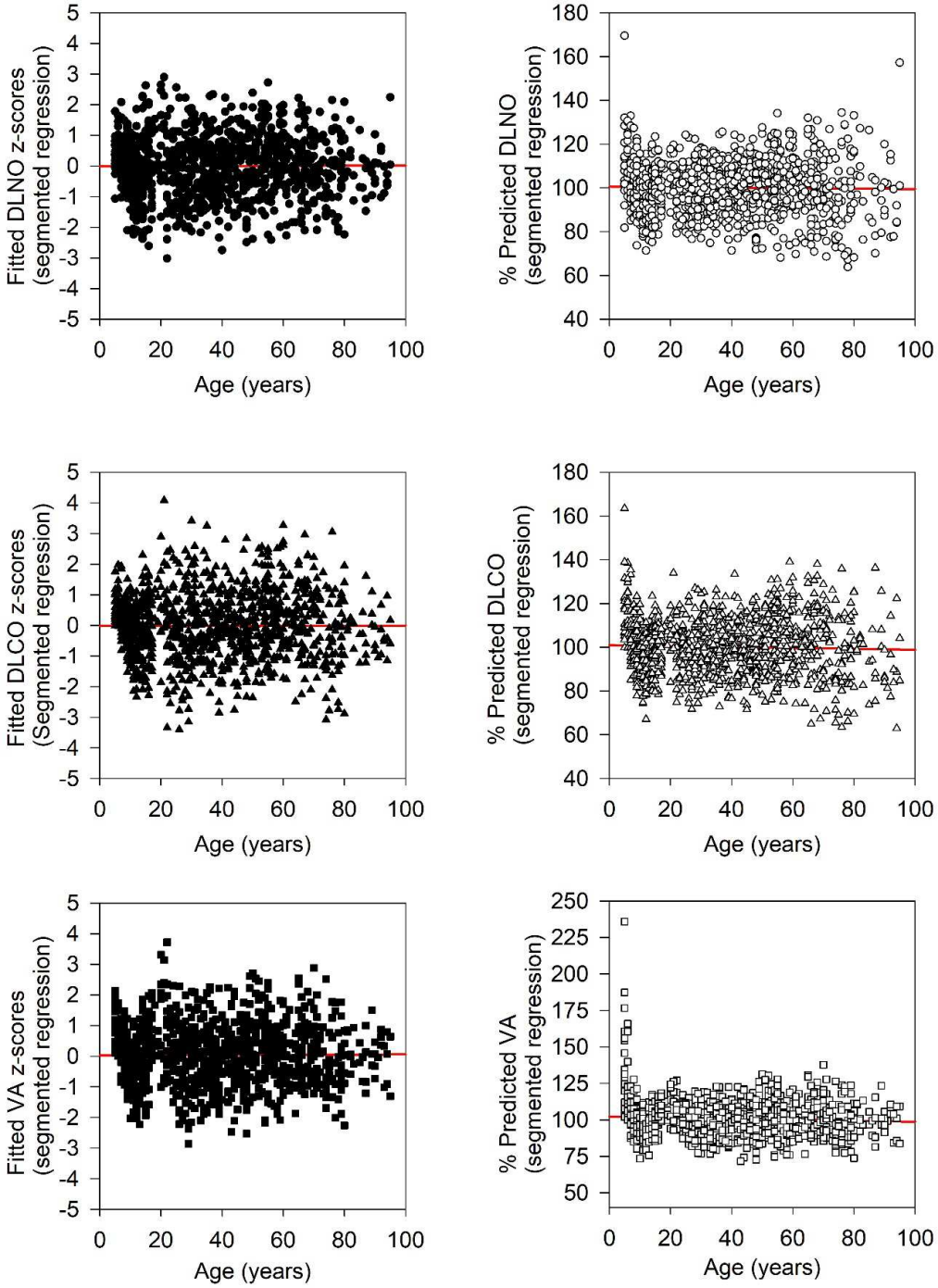


Figure S3. Fitted z-scores and % predicted DLNO, DLCO, and VA using GAMLSS models. Both devices are combined (n = 1076).

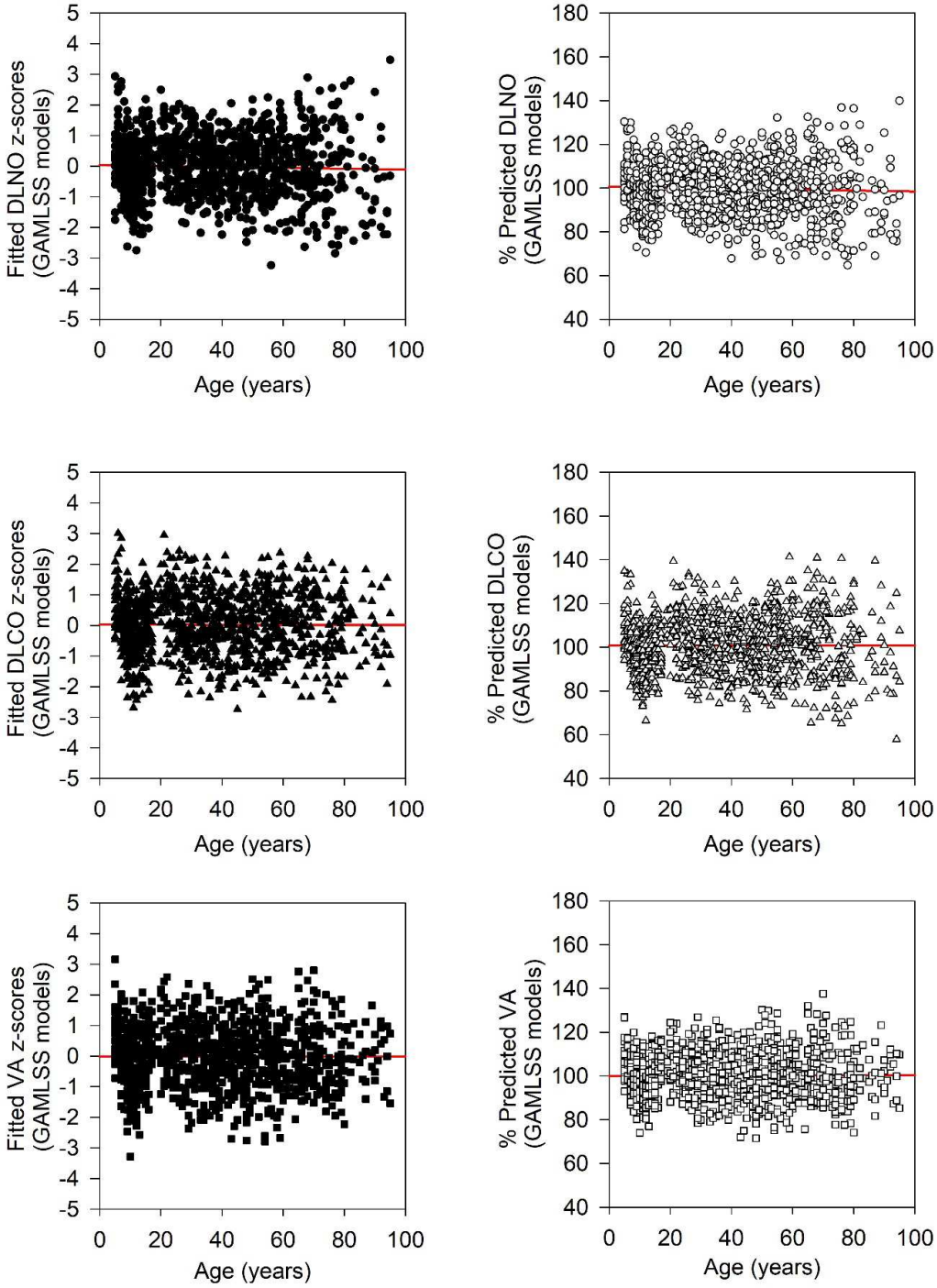


Figure S4. Quantile-Quantile (Q-Q) plots for DLNO, DLCO, and VA as generated by segmented regression models. These graphs depict the observed quantiles of a variable's distribution (red circles) against the quantiles that we would expect to see if the data is approximately normally distributed (solid black line). A normal distribution can approximate the z-scores data since most data points fall on the straight line. As for the percent predicted data, DLNO, DLCO, and VA shows two, one, and 12 outliers, respectively.

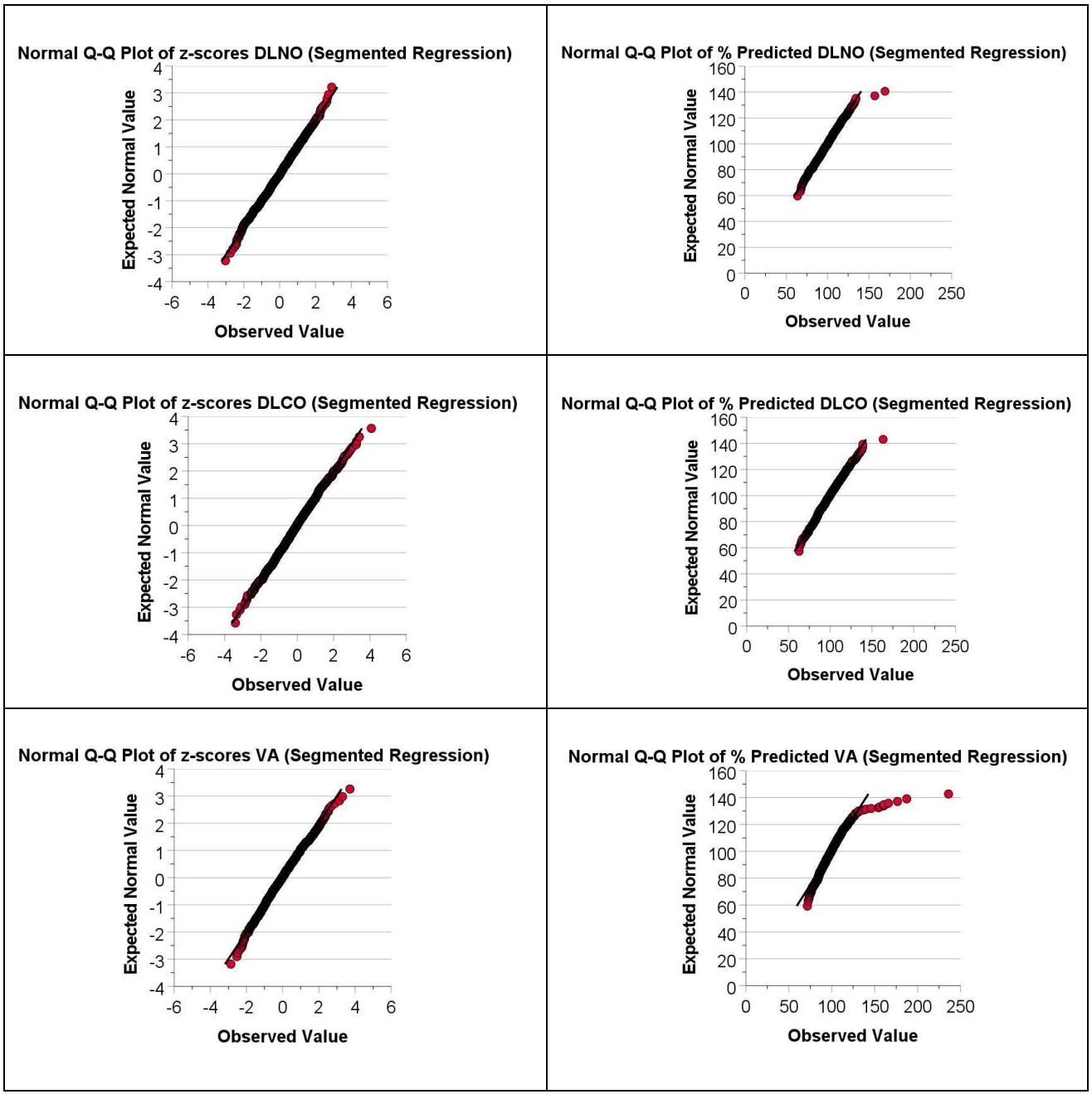


Figure S5. Quantile-Quantile (Q-Q) plots for DLNO, DLCO, and VA as generated by GAMLSS models. These graphs depict the observed quantiles of a variable's distribution (red circles) against the quantiles that we would expect to see if the data is approximately normally distributed (solid black line). The z-scores and % predicted values for all variables could be approximated by a normal distribution since almost all data points are straight.

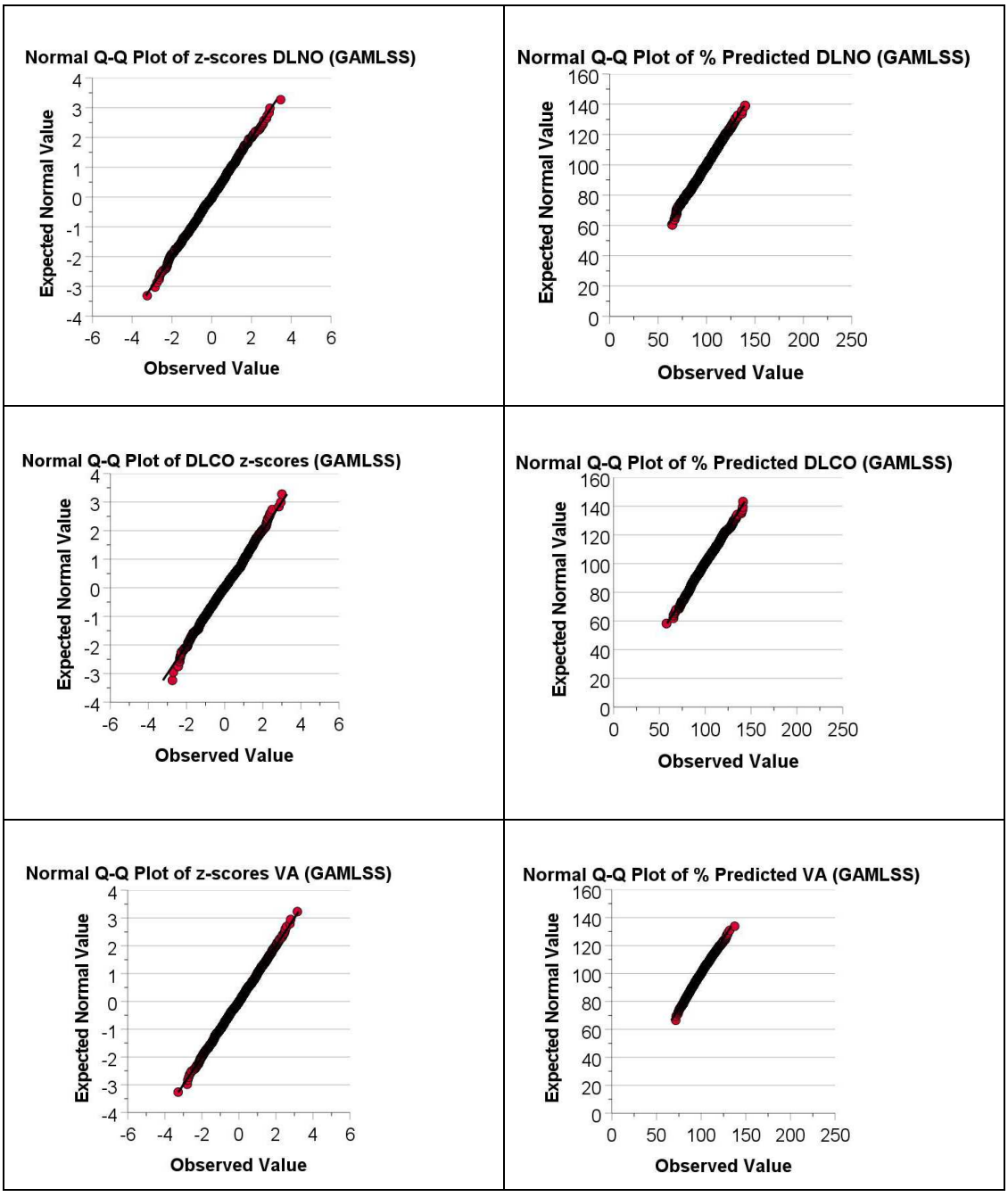


Figure S6. The mean and SD of the fitted scores for DLCO per model are displayed against the pulmonary function device used. There is a statistical difference between the fitted z-scores per model ($p < 0.001$) and pulmonary function device used ($p < 0.001$), and the interaction between model and device ($p < 0.001$). When the data were fitted to the Global Lung Function (GLI) Initiative GAMLSS reference equations, the DLCO z-scores were about 0.3 SD units higher than the z-scores obtained from the current study's segmented regression and GAMLSS two models when the data was obtained from the MasterScreen Pro device. However, when the data originated from the Hyp'Air Compact device, the z-scores fitted to the GLI reference equations were about 1.0 SD units higher than the z-scores obtained from the other two models ($p < 0.001$). The GLI GAMLSS reference equations are based on a 10 s breath-hold time, allowing for a better homogenous gas penetration in the lung. Indeed, a 5-6 s breath-hold time will result in poorer gas penetration in the lung. The Hyp'Air Compact device is represented by open white circles, and the Jaeger MasterScreen Pro is represented by solid, black-filled circles.

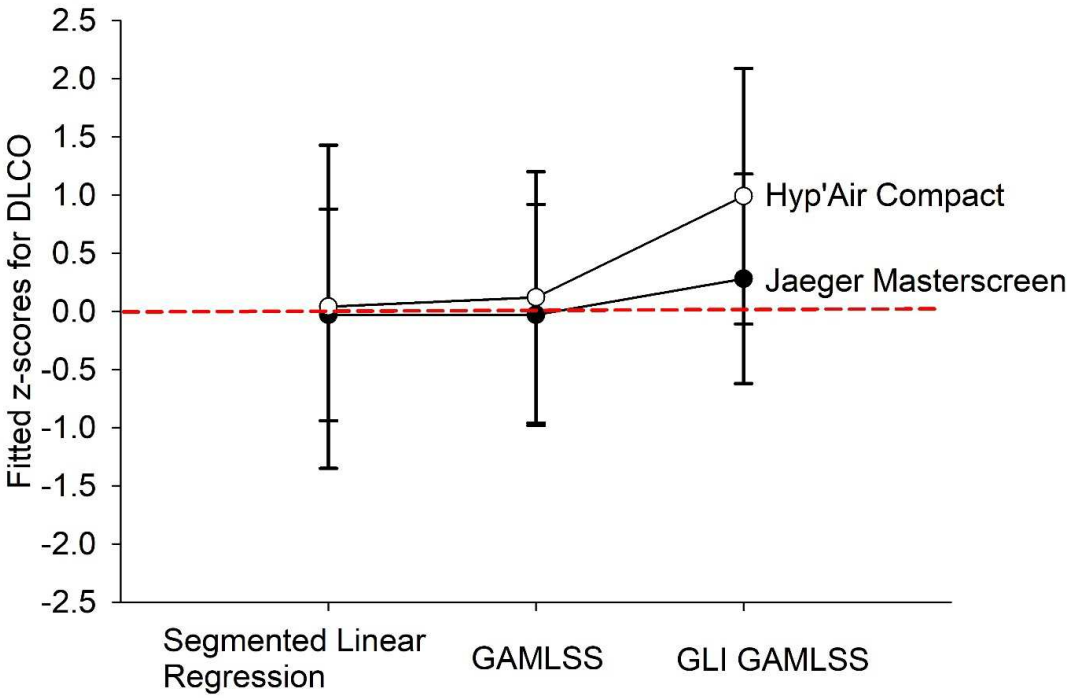


Figure S7. The mean and SD of the fitted z-scores for DLNO per model are displayed against the pulmonary function device used. There is no statistical difference between the fitted z-scores per model (p = 0.11), pulmonary function device used (p = 0.86), or the interaction between model and device (p = 0.67). Open white circles represent the Hyp'Air Compact device, and the Jaeger MasterScreen Pro is represented by solid, black-filled circles.

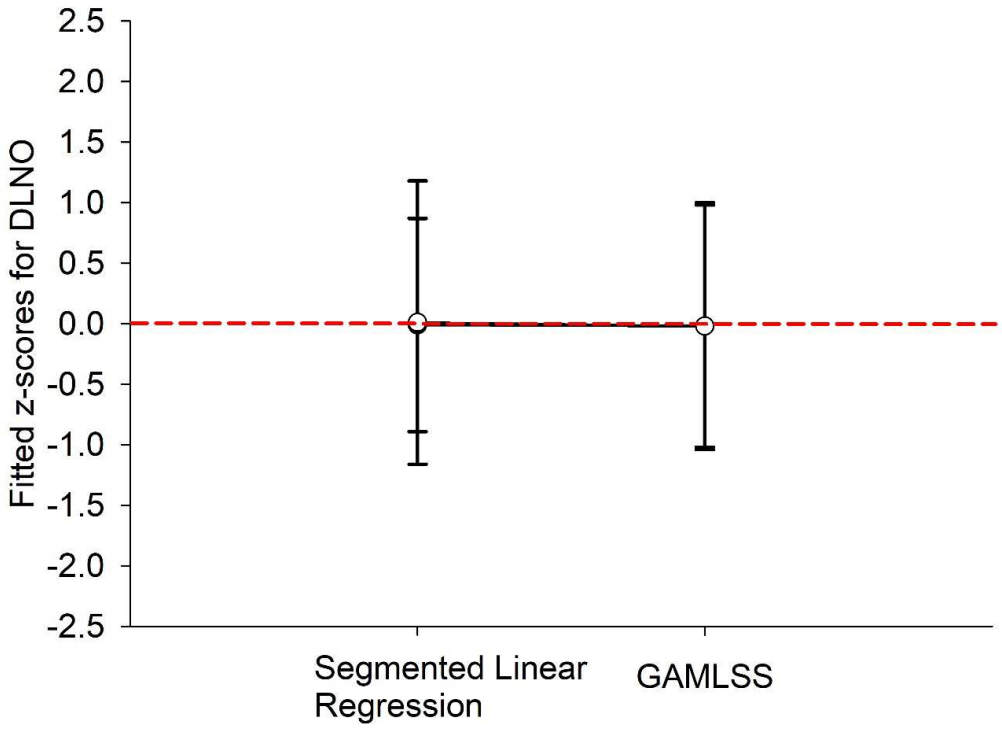


Figure S8. The mean and SD of the fitted z-scores for alveolar volume (VA) per model are displayed against the pulmonary function device used. There is a statistical difference between the fitted z-scores per model ($p < 0.001$) and pulmonary function device used ($p = 0.001$), as well as the interaction between model and device ($p < 0.001$). When the data originated from the Hyp'Air Compact device, the z-scores fitted to the GLI reference equations were 0.6 SD units higher than the z-scores obtained from the other two models ($p < 0.001$). The GLI GAMLSS reference equations are based on a 10 s breath-hold time, allowing for a better homogenous gas penetration in the lung. Indeed, a 5-6 s breath-hold time will result in poorer gas penetration in the lung. The Hyp'Air Compact device is represented by open white circles, and the Jaeger MasterScreen Pro is represented by solid, black-filled circles.

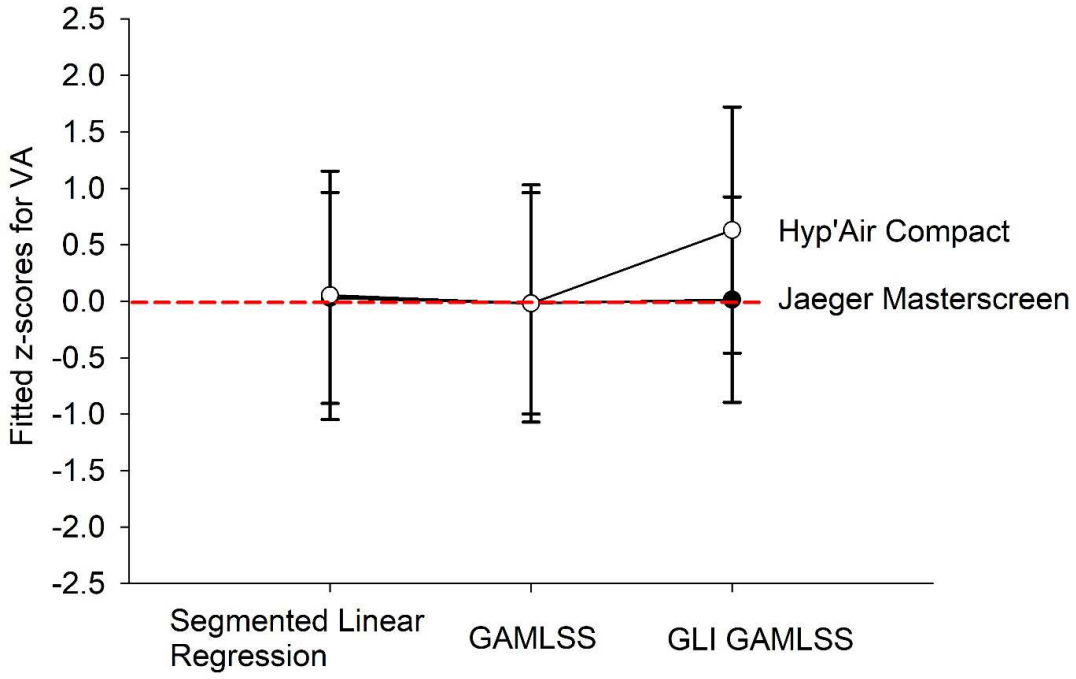


Figure S9A. Comparison between LLN for DLCO obtained from segmented linear regression models and the LLN for DLCO obtained from this study's GAMLSS equations. Both devices are combined (n = 1076). Units are mL/min/mmHg. The dotted black line is the line of identity; the solid red line is the regression line. The mean absolute difference between the LLN from both models was 1.1 (SD 1.1) mL/min/mmHg [range = 0 to 7.4 mL/min/mmHg]. Systematic differences (y-intercept) = 1.98 [95% CI = 1.77 to 2.20]; Proportional differences (slope) = 0.91 [0.90 to 0.92], Random differences (residual standard deviation) = 0.96 [95% CI = -1.88 to 1.88]; Linear model validity (Cusum test for linearity) = $p < 0.01$; Spearman rank correlation coefficient = 0.98.

LLN for DLCO GAMLSS = 0.91·(LLN for DLCO Segmented regression) + 1.98

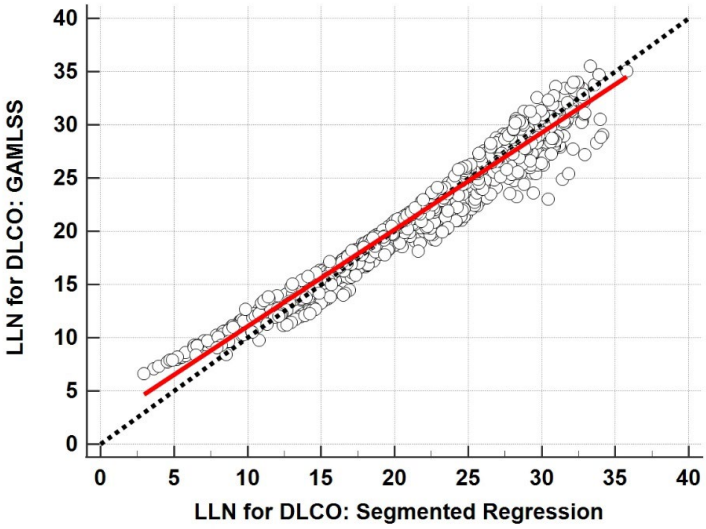


Figure S9B. Residuals plot by rank number. The residuals are the differences between the predicted values and the observed values. Red boxes indicate outliers (4 SD).

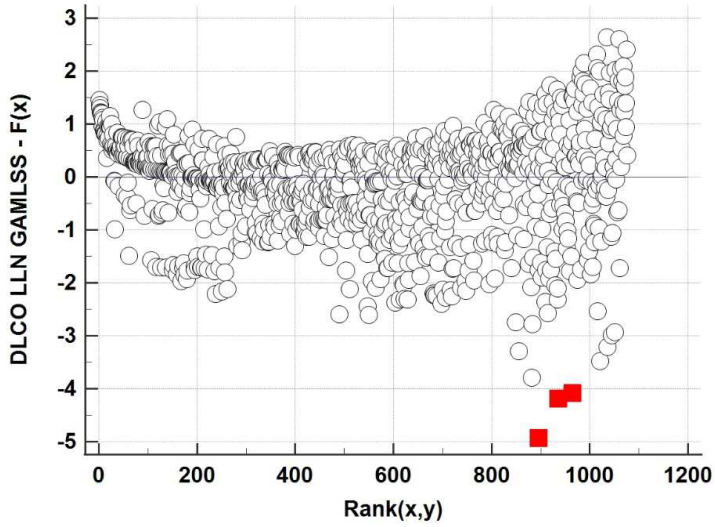
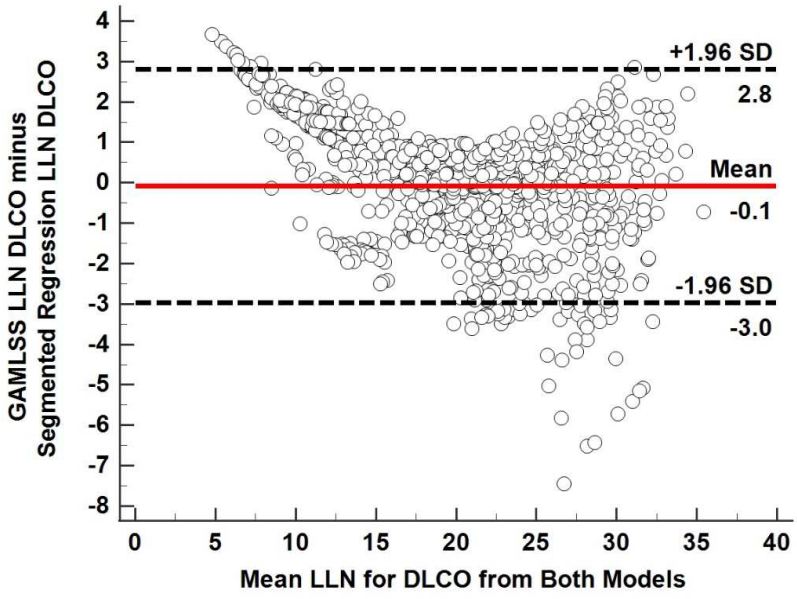


Figure S10A (Top panel). This plot demonstrates that when the LLN for DLCO tends to be on the low end (i.e., < 12.5 mL/min/mmHg), the GAMLSS models usually have a higher LLN. When the LLN for DLCO > 25 mL/min/mmHg, the segmented linear regression models can sometimes have a much higher LLN compared to GAMLSS models. Fifty-one subjects (4.7%) fall outside of ± 3.0 mL/min/mmHg difference. Both lung function devices are combined (n = 1076). Units are mL/min/mmHg. **Figure S10B (Lower panel).** The absolute difference in the LLN between both models versus the mean LLN from both models (mL/min/mmHg). Kendall's Tau correlation coefficient = – 0.05 [95% CI = – 0.10 to – 0.001] (p = 0.01).



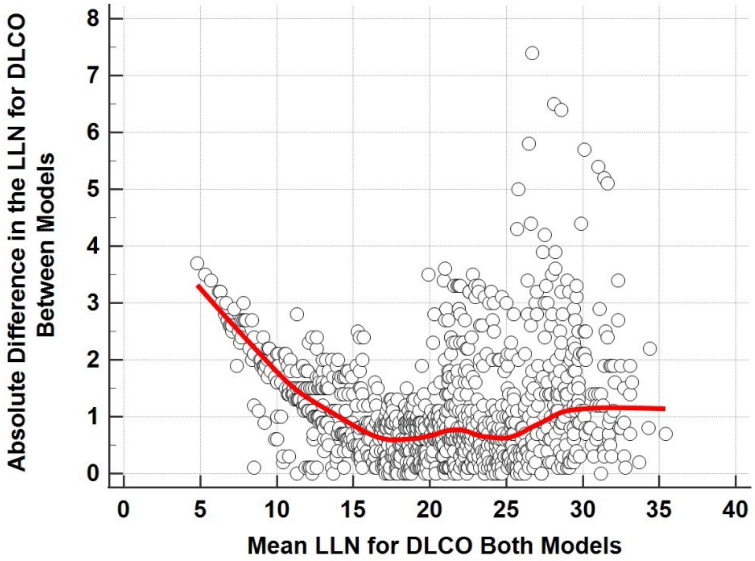


Figure S11A. Comparison between LLN for DLNO obtained from segmented linear regression models and the LLN for DLNO obtained from this study's GAMLSS equations. Both devices are combined (n = 1076). Units are mL/min/mmHg. The dotted black line is the line of identity; the solid red line is the regression line. The mean absolute difference between the LLN from both models was 6.2 (SD 5.5) mL/min/mmHg [range = 0 to 37 mL/min/mmHg]. Systematic differences (y-intercept) = 14.0 [95% CI = 12.9 to 15.0]; Proportional differences (slope) = 0.86 [0.85 to 0.87], Random differences (residual standard deviation) = 4.7 [95% CI = -9.2 to 9.2]; Linear model validity (Cusum test for linearity) = $p < 0.01$; Spearman's rank correlation coefficient = 0.98.

LLN for DLNO GAMLSS = 0.86·(LLN for DLNO Segmented Regression) + 14.0

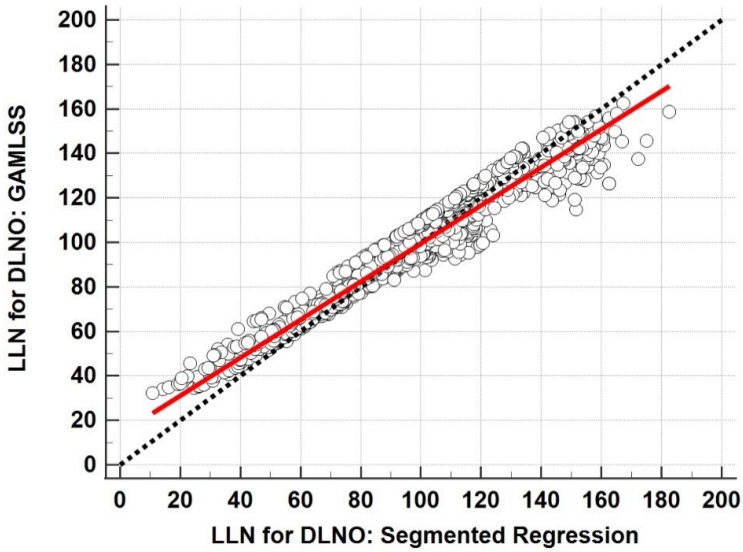


Figure S11B. Residuals plot by rank number. The residuals are the differences between the predicted values and the observed values. Red boxes indicate outliers (4 SD).

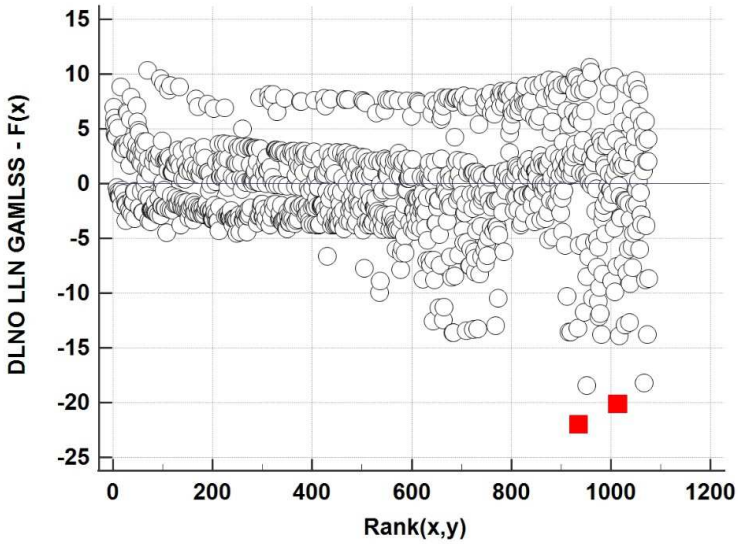
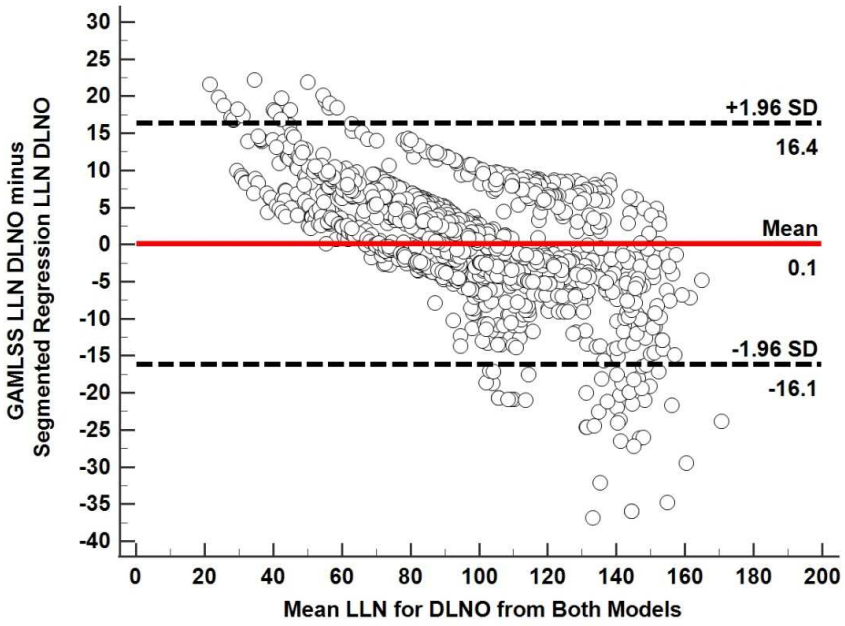


Figure S12A (Top panel). This plot demonstrates that when the DLNO is on the low end (< 60 mL/min/mmHg), the GAMLSS models usually have a higher LLN, and when the LLN for DLNO > 130 mL/min/mmHg, the segmented linear regression models can sometimes have a much higher LLN compared to GAMLSS models. Sixty-nine subjects (6.4%) fall outside of ±16.0 mL/min/mmHg. Both lung function devices are combined (n = 1076). Units are mL/min/mmHg.
Figure S12B (Lower panel). The absolute difference in the LLN between both models versus the mean LLN from both models (mL/min/mmHg). Kendall's Tau correlation coefficient = 0.06 [95% CI = 0.01 to 0.10] (p = 0.006).



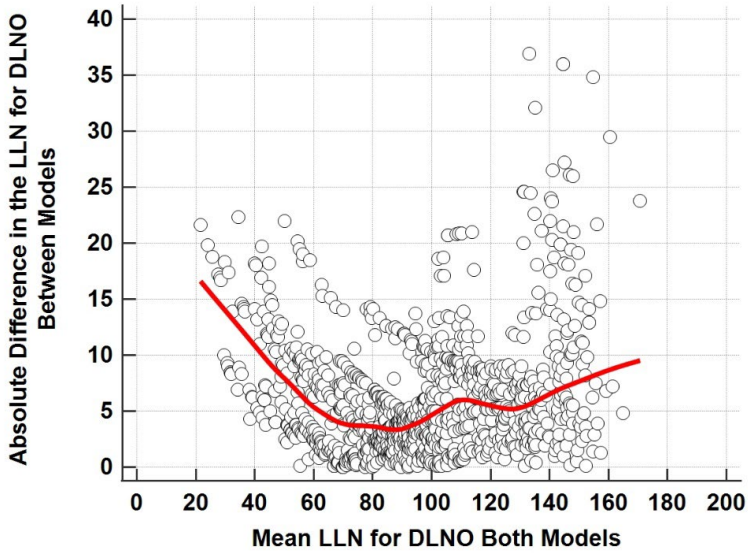


Figure S13A. Comparison between LLN for alveolar volume (VA) obtained from segmented linear regression models and the LLN for VA obtained from this study's GAMLSS equations. Both devices are combined (n = 1076). Units are in L. Dotted black line is the line of identity; the solid red line is the regression line. The mean absolute difference between the LLN from both models was 0.18 (SD 0.16) L [range = 0.0 to 1.38 L]. Systematic differences (y-intercept) = 0.22 [0.18 to 0.26]; Proportional differences (slope) = 0.97 [0.96 to 0.98], Random differences (residual standard deviation) = 0.14 [–0.28 to 0.28]; Linear model validity (Cusum test for linearity) = $p <$ 0.01; Spearman rank correlation coefficient = 0.99.

LLN VA GAMLSS = 0.968·(LLN for VA Segmented Regression) + 0.22

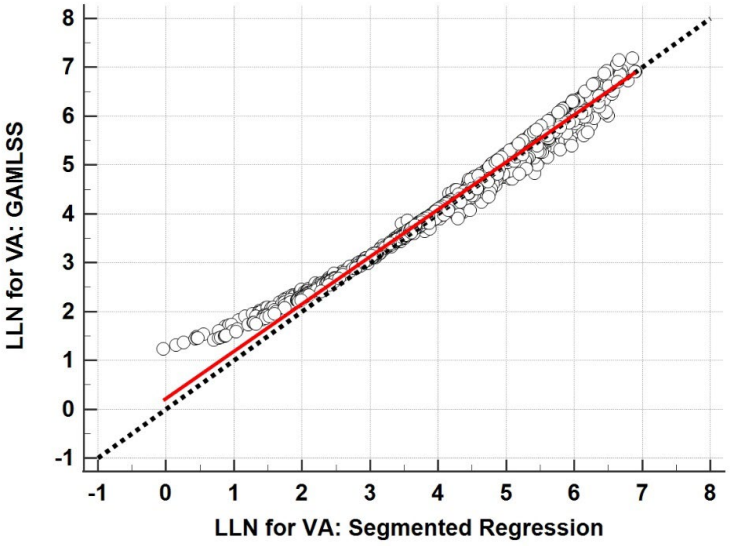


Figure S13B. Residuals plot by rank number. The residuals are the differences between the predicted values and the observed values. Red boxes indicate outliers (4 SD).

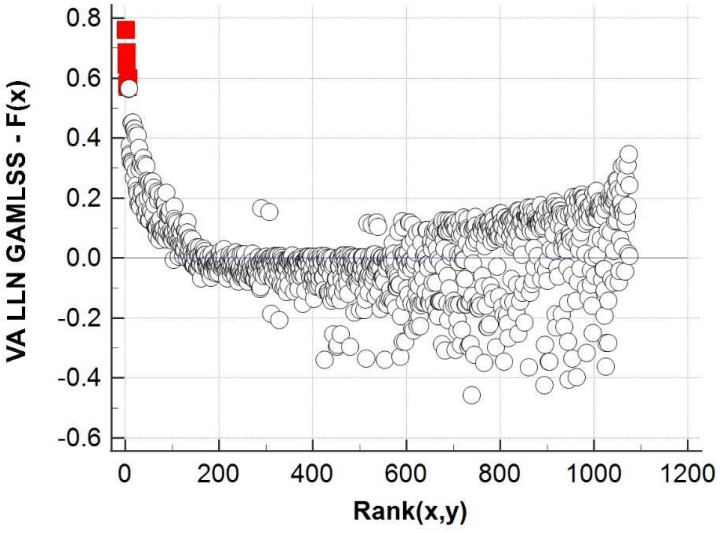


Figure S14A (Top panel). This plot demonstrates that when the VA is on the low end (< 2.5 L), the GAMLSS models developed in this study tend to have a higher LLN. When VA is > 4 L, the segmented linear regression models can sometimes have a much higher LLN compared to GAMLSS models Seventy-nine subjects (7.3%) fall outside of ± 0.40 L. Both lung function devices are combined (n = 1076). Units are L. **Figure S14B (Lower panel).** The absolute difference in the LLN between both models versus the mean LLN from both models (L). Kendall's Tau correlation coefficient = 0.07 [95% CI = 0.02 to 0.13] (p = 0.0004).

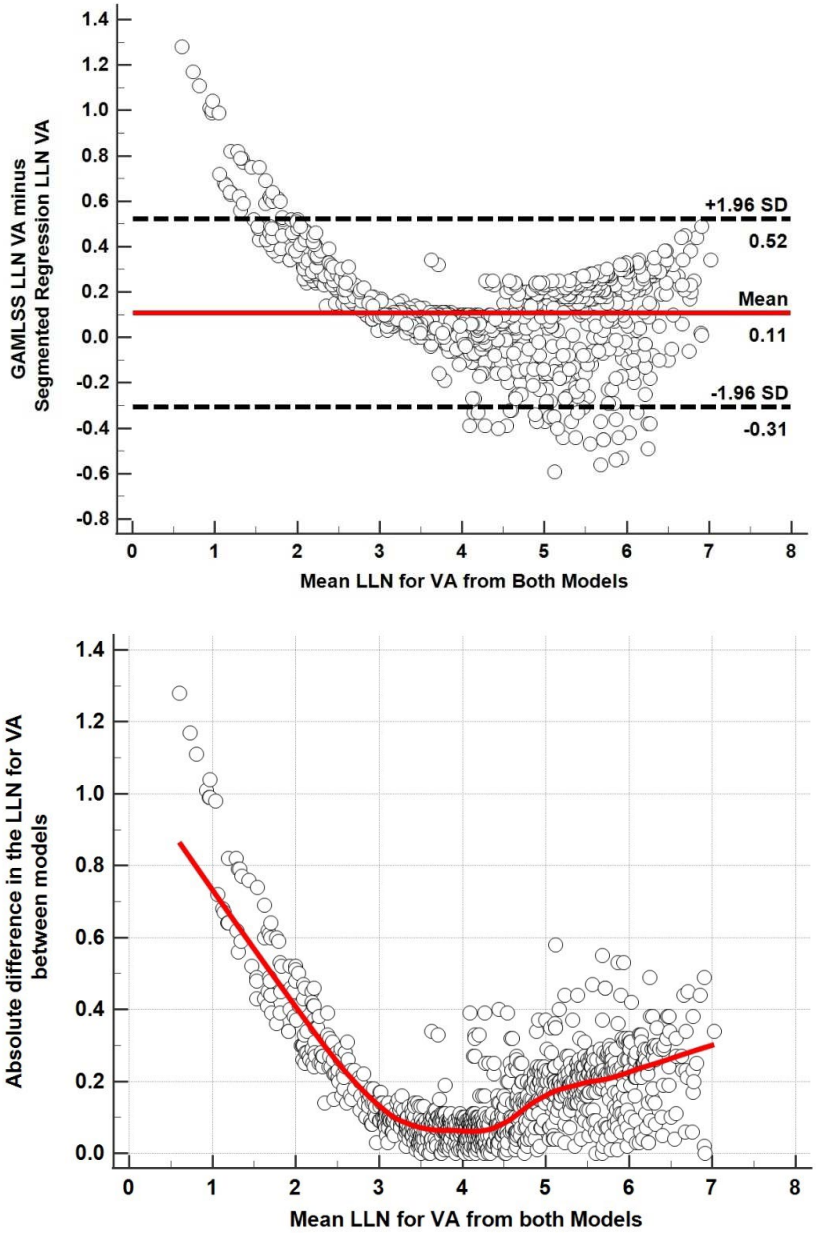
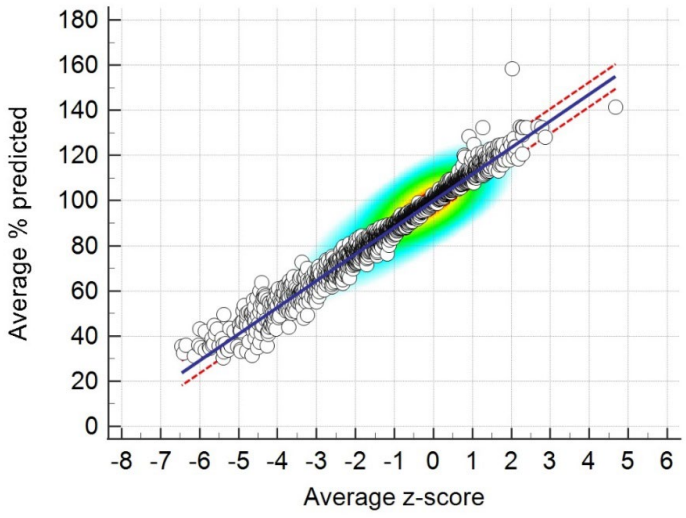
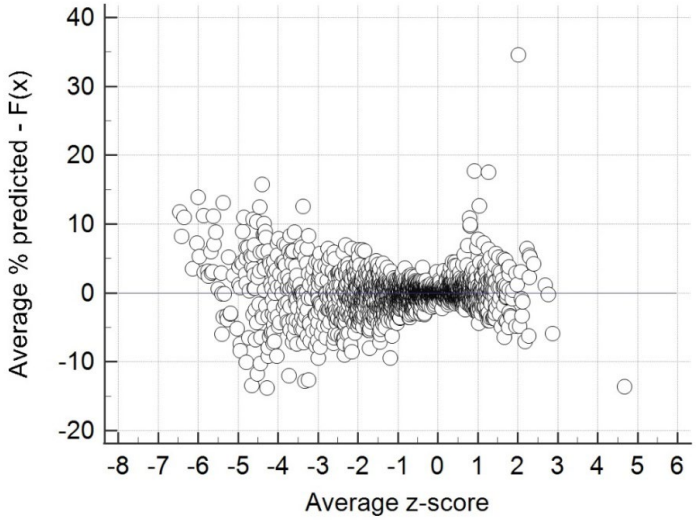


Figure S15A. Regression analysis demonstrates the linear association between percent predicted values and z-scores in both healthy and diseased subjects combined (n=2206). Each circle represents an average z-score combining DLNO, DLCO, and VA from both GAMLSS and segmented regression models. The data consisted of pooled data on healthy subjects [1 4-6 9] [n = 1076 subjects 5 to 95 years old, 546 males, 530 females] as well as diseased patients [n = 683 patients, 8 to 87 years of age, 289 males, 394 females][10-20]. Also, 447 control subjects from studies with cardiopulmonary disease were used[13 14 17 18]. Three-hundred twenty-nine out of 2206 subjects (15%) include unpublished data from Cochin Hospital in Paris, France (Courtesy of Dr. Anh-Tuan Dinh Xuan, with permission). The mean breath-hold times were 6.1 s (SD = 1.9 s) for all data.

% predicted = 11.80·(z-score) + 99.93 R^2 = 0.98, F(1, 2204) = 94775, p < 0.001 Residual standard deviation = 2.8% 95% CI of slope = 11.7 to 11.9% 95% CI of intercept = 99.8 to 100.1% Shapiro-Wilk test for Normal Distribution, W = 0.85, reject normality (p < 0.001). Red dashed lines represent the 95% prediction interval. The heat map represents where most of the data lie.	
Figure S15B. Residuals plot by rank number. The residuals are the differences between the predicted values and the observed values.	

References

1. Munkholm M, Marott JL, Bjerre-Kristensen L, et al. Reference equations for pulmonary diffusing capacity of carbon monoxide and nitric oxide in adult Caucasians. *Eur Respir J* 2018;52(1):1500677. doi: 10.1183/13993003.00677-2015

2. Zavorsky GS, Hsia CC, Hughes JM, et al. Standardisation and application of the single-breath determination of nitric oxide uptake in the lung. *Eur Respir J* 2017;49(2):1600962. doi: 10.1183/13993003.00962-2016

3. Robert CP. Simulation of Truncated Normal Variables. *Stat Comput* 1995;5(2):121-25. doi: Doi 10.1007/Bf00143942

4. Aguilaniu B, Maitre J, Glenet S, et al. European reference equations for CO and NO lung transfer. *Eur Respir J* 2008;31(5):1091-7. doi: 10.1183/09031936.00063207

5. van der Lee I, Zanen P, Stigter N, et al. Diffusing capacity for nitric oxide: reference values and dependence on alveolar volume. *Respir Med* 2007;101(7):1579-84. doi: 10.1016/j.rmed.2006.12.001

6. Zavorsky GS, Cao J, Murias JM. Reference values of pulmonary diffusing capacity for nitric oxide in an adult population. *Nitric Oxide* 2008;18(1):70-9. doi: 10.1016/j.niox.2007.10.002

7. Gray G, Zamel N, Crapo RO. Effect of a simulated 3,048 meter altitude on the single-breath transfer factor. *Bull Eur Physiopathol Respir* 1986;22(5):429-31. [published Online First: 1986/09/01]

8. Fryar CD, Carroll MD, Gu Q, et al. Anthopometric Reference Data for Children and Adults: United States, 2015-2018. Analytical and Epidemiological Studies. National Center for Helath Statistics. Washington, DC: Vital Health Statistics, 2021:36 pages.

9. Thomas A, Hanel B, Marott JL, et al. The single-breath diffusing capacity of CO and NO in healthy children of European descent. *PLoS One* 2014;9(12):e113177. doi: 10.1371/journal.pone.0113177

10. Radtke T, Benden C, Maggi-Beba M, et al. Intra-session and inter-session variability of nitric oxide pulmonary diffusing capacity in adults with cystic fibrosis. *Respir Physiol Neurobiol* 2017;246:33-38. doi: 10.1016/j.resp.2017.08.002

11. van der Lee I, Zanen P, Biesma DH, et al. The effect of red cell transfusion on nitric oxide diffusing capacity. *Respiration* 2005;72(5):512-6. doi: 10.1159/000087676

12. Moinard J, Guénard H. Determination of lung capillary blood volume and membrane diffusing capacity in patients with COLD using the NO-CO method. *Eur Respir J* 1990;3(3):318-22.

13. van der Lee I, Gietema HA, Zanen P, et al. Nitric oxide diffusing capacity versus spirometry in the early diagnosis of emphysema in smokers. *Respir Med* 2009;103(12):1892-7. doi: 10.1016/j.rmed.2009.06.005

14. Magini A, Apostolo A, Salvioni E, et al. Alveolar-capillary membrane diffusion measurement by nitric oxide inhalation in heart failure. *Eur J Prev Cardiol* 2015;22(2):206-12. doi: 10.1177/2047487313510397

15. Apostolo A, Paolillo S, Contini M, et al. Comprehensive effects of left ventricular assist device speed changes on alveolar gas exchange, sleep ventilatory pattern, and exercise performance. *J Heart Lung Transplant* 2018;37(11):1361-71. doi: 10.1016/j.healun.2018.07.005

16. Magini A, Contini M, Spadafora E, et al. Variability in pulmonary diffusing capacity in heart failure. *Respir Physiol Neurobiol* 2020;280:103473. doi: 10.1016/j.resp.2020.103473

17. Barisione G, Brusasco C, Garlaschi A, et al. Lung diffusing capacity for nitric oxide as a marker of fibrotic changes in idiopathic interstitial pneumonias. *J Appl Physiol* 2016;120(9):1029-38. doi: 10.1152/japplphysiol.00964.2015

18. Sorensen JK, Buchvald F, Berg AK, et al. Ventilation inhomogeneity and NO and CO diffusing capacity in ex-premature school children. *Respir Med* 2018;140:94-100. doi: 10.1016/j.rmed.2018.06.006

19. Dressel H, Filser L, Fischer R, et al. Lung diffusing capacity for nitric oxide and carbon monoxide: dependence on breath-hold time. *Chest* 2008;133(5):1149-54. doi: 10.1378/chest.07-2388

20. Degano B, Soumagne T, Delaye T, et al. Combined measurement of carbon monoxide and nitric oxide lung transfer does not improve the identification of pulmonary hypertension in systemic sclerosis. *Eur Respir J* 2017;50(4) doi: 10.1183/13993003.01008-2017