bioRxiv preprint first posted online Aug. 30, 2018; doi: http://dx.doi.org/10.1101/395160. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. All rights reserved. No reuse allowed without permission.

Total lung capacity without plethysmography

08/18/18

Ori Adam^{1,7}, Inon Cohen⁷, Wai-Ki Yip⁶, Robert J. Shiner^{2,7}, Peter Calverley^{3,7}, Zachi Peles⁷, Adam LaPrad⁷, Yoni Dagan⁷, Robert Brown^{4,7*}, Julian Solway^{5,7*}, and Jeffrey J. Fredberg^{6,7*}

¹Department of Geophysical, Atmospheric and Planetary Sciences, Tel Aviv University, Tel Aviv, Israel

²Department of Respiratory Medicine, NHLI, Imperial College, London, United Kingdom

³School of Medicine, University of Liverpool, Liverpool, United Kingdom

⁴Pulmonary and Critical Care Medicine Unit, Department of Medicine, Massachusetts General Hospital,

Harvard Medical School, Boston, MA, United States

⁵Department of Medicine, University of Chicago, Chicago, IL, United States

⁶Harvard School of Public Health, Harvard University, Boston, MA, United States

⁷PulmOne Advanced Medical Devices, Ltd, Ra'anana, Israel

*senior co-authors

Corresponding author: Jeffrey J. Fredberg, Ph.D.; Program in Molecular and Integrative Physiological Sciences; Department of Environmental Health; Harvard School of Public Health Room 1-313; 665 Huntington Avenue; Boston, MA 02115; 617-432-0198 phone; 617-432-3468 fax; jfredber@hsph.harvard.edu

Running Head: TLC without plethysmography

Key Words: Pulmonary Function Testing, Physiology, Lung Volume, Lung Mechanics

Author Contributions: O.A., R.J.S., P.C., R.B., J.S., and J.J.F. designed the study. R.J.S., P.C., R.B., J.S., and J.J.F. oversaw the study. O.A. and Z.P. performed the study. I.C., W.Y., and Y.D. analyzed the data. A.L., J.S., and J.J.F. contributed to the writing of the manuscript. All authors have seen and approved the final version of the manuscript.

ABSTRACT

Background

Among the most basic measures of respiratory function is the total lung capacity (TLC). TLC is the pulmonary gas volume at maximal lung inflation, which is the sum of the volume of gas that can be exhaled –the vital capacity (VC)– and the volume of gas that cannot –the residual volume (RV). Determination of VC requires only spirometry whereas determination of RV or TLC requires body plethysmography, gas dilution or washout, or thoracic imaging, each of which is more complex than spirometry, and none of which is suited to routine office practice, population screening, or community medicine. To fill this gap, we describe here a new approach to determine TLC without plethysmography.

Methods

In a heterogeneous population of 434 volunteers (265 male, 169 female; 201 healthy, 170 with airflow obstruction, and 63 with ventilatory restriction), we determined TLC in the standard fashion using conventional body plethysmography (TLC_{pleth}). In the same individuals, we also determined TLC in a novel fashion using the MiniBoxTM (TLC_{MB}). To obtain TLC_{MB}, population-based data from traditional spirometry together with flow-interruption transients were subjected to data mining and machine-learning to create for each individual subject an unbiased statistical determination of TLC.

Results

For the combined heterogeneous population, we found $TLC_{pleth} = 1.02TLC_{MB}$ -0.091 L, adjusted r^2 =0.824. For the heterogeneous population as a whole, and for each subpopulation, TLC_{MB} closely tracked TLC_{pleth} . For 26 healthy subjects measured on different days, the coefficient of variation for repeated measurements in was 3.3% for TLC_{pleth} versus 1.6% for TLC_{MB} . *Conclusions*

These results establish the validity and potential utility of a new method for rapid, accurate, and repeatable determination of TLC in a heterogeneous patient population, but without the need of a plethysmograph.

INTRODUCTION

The most common test of pulmonary function is spirometry, in which the volume of air flowing into and out of the respiratory system is measured directly. Spirometry can quantify volume differentials such as tidal volume (V_T), forced vital capacity (FVC), or expiratory reserve volume (ERV), but cannot measure absolute volumes such as residual volume (RV), functional residual capacity (FRC), or total lung capacity (TLC). Absolute thoracic gas volumes (TGVs), such as RV, FRC, and TLC are useful in the diagnosis and management of respiratory system diseases, but their measurement requires technologies that are more complex and labor intensive than spirometry. Accordingly, RV, FRC, and TLC are often not available in adult or pediatric office practices.

To measure absolute lung volumes, the ATS/ERS Consensus Statement identifies five methods: whole body plethysmography, multi-breath helium dilution, nitrogen wash-out, computed tomography, and chest radiography.¹ Among these, body plethysmography is used most commonly and is widely regarded as the gold standard.¹⁻⁸ Since its inception by Dubois in 1956⁹, body plethysmography has remained simple in principle but inherently complex, capital intensive, and physically imposing in practice. The plethysmograph can be uncomfortable or intimidating for the patient enclosed within it and, moreover, is dependent upon a skilled technician for calibration, operation and maintenance. Gas dilution and gas washout are well-established alternatives to body plethysmography, but each presents its own technical challenges.

For these reasons, investigators have explored alternative avenues to determine absolute lung volumes by other means, but with no success. Respiratory system impedance, even when extended to a wide range of forcing frequencies, has been shown to be inadequate to infer absolute lung volumes in the individual subject.¹⁰⁻¹⁷ Similarly, forced expiratory maneuvers have been shown to be inadequate.¹⁸ These failures may be attributable in part to the fact that the dynamics of gas distribution within the human lung are complex, and especially so in obstructive lung disease. Moreover, data interpretation in these approaches often rests upon fitting data to idealized mathematical models wherein there exists a wide range of TGV values that might fit the data equally well. When this happens, no useful determination of TGV can be inferred and the problem of mathematical inference is said to be non-unique or ill-posed.

To determine TLC in the individual subject without using a plethysmograph, here we take a different approach. Across a heterogeneous population of volunteers we measured

traditional spirometry together with flow-interruption transients using a MiniBoxTM, described below. In the same individuals we also measured TLC using traditional body plethysmography (TLC_{pleth}). To obtain an unbiased statistical determination of TLC from corresponding MiniBoxTM data (TLC_{MB}) in the individual subject, we then used data mining and machinelearning. Because TLC_{MB} for the individual subject is based on population-based data mining and machine-learning rather than a direct physiological measurement, its main advantage is that it rests upon no idealizing assumptions concerning respiratory system structure or function. Nevertheless, in any given subject this approach was able to determine TLC in a fashion that is accurate and repeatable when compared to TLC_{pleth} but without the need of a plethysmograph.

METHODS

The study comprised three parts. First, in a heterogeneous population of 300 qualified volunteers, as described below, we measured TLC in the conventional manner using body plethysmography. In these same volunteers, we measured conventional spirometry and flow-interruption transients using a desktop device called the MiniBox (PulmOne Advanced Medical Devices, Ltd., Ra'anana, Israel). Based upon these data, we used a statistical algorithm –the LASSO^{19,20}– to find the strongest set of statistical predictors of TLC_{pleth}. We arrived at a final statistical model with which to calculate TLC from the statistical predictors, called TLC_{MB}. Second, we validated this statistical prediction using N-fold cross-validation.^{21,22} Third, to evaluate this statistical model still further, we compared TLC_{MB} against TLC_{pleth} in a prospective heterogeneous cohort of 134 qualified volunteers.

Subject population. We recruited volunteers at 6 institutions (Soroka University Medical Center, Beer Sheva, Israel; Rambam University Medical Center, Haifa, Israel; Maccabi HaShalom, Tel-Aviv, Israel; Maccabi HaSharon, Kefar-Saba, Israel; Assaf HaRofeh Hospital, Tzrifin, Israel; Tel Aviv Medical Center, Tel-Aviv, Israel) under a research protocol approved by the Ethical Review Board of each. The prospective clinical study was registered with ClinicalTrials.gov (NCT 01952431).

The population comprised three groups (Table 1): 1) healthy subjects; 2) subjects with airflow obstruction, such as chronic obstructive lung disease (COPD) or asthma and with varying severity level (mild, moderate, and severe); and 3) subjects with restrictive ventilatory disorders.

Patients were recruited from the Lung Function Laboratory at each institution. In each case, disease severity was defined by the criteria in ATS/ERS guidelines.²³

Subjects were considered eligible if they: a) provided informed consent; b) were at least 18 years of age; and c) were cooperative and capable of following instructions. Healthy subjects were eligible if they: a) never smoked; b) had no known history of respiratory, cardiovascular, hepatic, renal or metabolic disease; c) had a BMI < 35 kg/m^2 ; d) had no persistent (lasting greater than 3 days) respiratory symptoms during the last 12 months (e.g., dyspnea, chronic cough, wheezing or phlegm); and e) had no history suggesting upper respiratory infection during the three weeks prior to testing. Non-healthy subjects were eligible if they had a documented obstructive or restrictive respiratory disorder.

Subjects were excluded from the study if they: a) were pregnant at the time of the study; b) had performed any significant physical activity that has resulted in breathlessness during 1 hour prior to the study; c) had a tracheostomy; d) were unable to satisfactorily perform routine, full lung function testing including body plethysmography (e.g., due to claustrophobia or inability to perform the panting maneuver required for plethysmography); e) were unable or unwilling to give informed consent; or f) were unable to complete the protocol.

For each subject, all measurements were made in the same laboratory, by the same technician, and were competed within two hours. The technician also recorded the subject's gender, date of birth, height, weight, and summary medical history.

Body plethysmography. Depending on the study site, different commercial body plethysmographs were used: a) ZAN 500 (nSpire Health, Inc) - Soroka, Rambam, Maccabi HaSharon, Assaf HaRofeh; b) Platinum Elite-Series (MedGraphics) – Maccabi HaShalom; c) MasterScreen Lab (Erich Jaegar, CareFusion) – Tel Aviv, Assaf HaRofeh. Associated transducers were calibrated in accordance with the manufacturers' user manuals. Device calibration and device agreement between institutions were verified using manually operated isothermal containers (3 L or 5 L) filled with copper wool, as well as by measuring a healthy control subject with a known TLC.

Body plethysmography measurements were performed in accordance with manufacturer recommendations and ATS/ERS guidelines.¹ Subjects panted at FRC at 0.5 to 1 Hz against a closed valve and then inhaled to TLC followed by slow exhalation to RV. The thoracic gas

volume (TGV) at FRC was calculated as the mean of the first 3 individual FRCs that were within 5% of each other in which the 2 highest inspiratory capacity (IC) measurements were within 10% (or 0.15 L) of each other. TLC_{pleth} was calculated by adding the largest of the three ICs to the mean TGV.

MiniBox. The MiniBoxTM (PulmOne Advanced Medical Devices, Ltd., Ra'anana, Israel) is a table-top unit that includes a spirometer and a flow-interruption device. The flow-interruption device (Figure 1A) consists of a rigid 16.3 L container, a rapidly closing valve (<10 msec), and a hotwire anemometer-type flowmeter (working range \pm 5 L/s) (Figure 1B). Calibration of the MiniBox flowmeters was performed daily using a standard 3 L syringe.

Subjects first performed spirometry using either the spirometer associated with the body plethysmograph or the hand-held spirometer associated with the MiniBox. For each participating laboratory, spirometry measurements were performed in accordance with the manufacturers' user manuals and in a manner typical for that laboratory, as assessed by the local laboratory director. By design, spirometry measurements corresponded to real-world circumstances so as to include real-world methodological variability. That being the case, these measurements did not necessarily conform to ATS/ERS guidelines.²⁴ For example, in retrospective analysis we found that not all spirometry efforts continued for at least 6 seconds. At least 3 spirometry efforts were recorded. For SVC and forced vital capacity (FVC), additional maneuvers were performed until the 2 highest values of the 3 measurements comprising the selected group were within 5% (or 0.15 L) of each other. Similarly for IC, the 2 highest values of the 3 measurements comprising the selected group were required to be within 10% (or 0.15 L) of each other.

Subjects were next measured with the MiniBox flow-interruption device. With cheeks supported and a nose clip in place, each subject sat upright in a chair and breathed through a disposable bacterial filter attached to the MiniBox flow-interruption device (Figure 1A). The subject was asked to breathe normally for a short time until comfortable with the device. Then, brief flow interruptions (~70 msec) were automatically triggered in the vicinity of mid-inspiration of each tidal breath (Figure 1C). After a minimum of 25 such interruptions or a maximum of 150 seconds of tidal breathing, the subject performed a maximal inspiration twice to reach total lung capacity (TLC). The subject then exhaled slowly to residual volume (RV).

The above flow-interruption measurement was repeated up to 3 times. The entire measurement was deemed acceptable if the SVC measured with the MiniBox flow-interruption device was within 10% (or 0.15 L) of the SVC measured with the spirometer. MiniBox flow interruption data were pre-processed and filtered based on pre-defined criteria.

Statistical modelling. To construct an unbiased statistical model for TLC, and thus calculate TLC_{MB} for each individual volunteer, we identified 137 plausible predictors of TLC from the spirometry and flow interruption data in 300 qualified volunteers. The metrics included conventional spirometry indices, transient flows and pressures measured at different time points during the flow-interruption, and their time derivatives. We then used a statistical algorithm – the LASSO^{19,25} – to find the smallest possible set of predictors that produced a statistically significant determination of TLC. The LASSO is an extension of multiple linear regression that finds a combination of parameters while forcing all but a few coefficients to be precisely zero, thereby providing a minimal statistical model. Here, the LASSO was accomplished using a tunable parameter that constrains the coefficients with cross-validation using random sampling with replacement (bootstrapping)²⁶⁻²⁹ repeated 300 times for each value from a range of possible values. Each sampling was constrained according to the same ratio of male/female and healthy/non-healthy as the entire group of subjects. Using the LASSO applied to the dataset of 300 subjects, we arrived at a final statistical model to calculate TLC_{MB}.

N-fold cross-validation. To assess the predictive ability of the model, 10-fold cross-validation was used on the dataset of 300 subjects.^{21,22} The dataset was randomly divided using 10-fold for 50 times and the samples for each fold were selected randomly for each time. A 5-fold and leave-one-out cross validation (LOOCV) was also performed as comparison.

Prospective validation. Last, we performed an independent prospective study to further validate the TLC_{MB} equation. In a prospective heterogeneous cohort of 134 additional volunteers not previously studied (Table 1), we repeated the protocol of MiniBox and body plethysmography measurements. We then used the new MiniBox data and the TLC_{MB} equation derived from the initial cohort of 300 subjects to calculate TLC_{MB} and compared it to TLC_{pleth} in the prospective cohort of 134 subjects.

RESULTS

Subject characteristics. A total of 564 subjects were enrolled, of whom 4 were unable to complete the protocol and 126 were disqualified based on quality assurance criteria. There were not any adverse events. The final qualified dataset comprised 300 subjects in the first cohort and 134 subjects in the prospective cohort (Table 1). Both cohorts included healthy individuals and patients with a range of diseases and a range of disease severities.

TLC_{MB} versus TLC_{pleth}. Across the entire mixed population of 300 qualified subjects, TLC_{MB} tracked TLC_{pleth} closely (Figure 2A; TLC_{pleth} = $1.02TLC_{MB} - 0.091$ L, adjusted r^2 =0.824). In the subset of 150 healthy individuals, the variability was smallest (Figure 2B; TLC_{pleth} = $0.991TLC_{MB} + 0.0414$ L, adjusted r^2 =0.852) while in the subset of 114 obstructed subjects (Figure 2C; TLC_{pleth} = $1.02TLC_{MB} - 0.004$ L, adjusted r^2 =0.739) and in the subset of 36 restricted subjects (Figure 2D; TLC_{pleth} = $0.844TLC_{MB} - 0.474$ L, adjusted r^2 =0.653), the variability was somewhat greater. Nonetheless, in each of these subpopulations, TLC_{MB} closely tracked TLC_{pleth}.

To examine differences between results from both methods and their dependence on lung size, we performed Bland-Altman analyses.³⁰ In the population as a whole (Figure 2E), and in each of the subpopulations (healthy - Figures 2F; obstructed - Figure 2G, and restrictive - Figure 2H), the coefficients of variations were 9.91%, 7.93%, 11.30%, and 13.70% respectively; the mean biases were small (0.01 L, -0.01 L, 0.11 L, and 0.20 L, respectively); also, there was no systematic trend of variability or bias with lung size.

N-Fold cross validation: The mean prediction error using 10-fold cross-validation was 0.437 L with the mean prediction SE of 0.00171 L. The 5-fold and LOOCV also produced similar mean prediction errors. Thus, if the initial 300 volunteer cohort is representative of the population in which TLC_{MB} would be measured in practice, the statistical model has good predictive power.

Independent validation in the prospective cohort. To further validate TLC_{MB}, we then used the statistical model equation derived from the original cohort of 300 to calculate TLC_{MB} for each

member of an independent prospective cohort of 134 (Figure 3). Although slopes and the adjusted r^2 were slightly lower in that prospective cohort, TLC_{MB} closely tracked TLC_{pleth} and followed similar regression lines and confidence intervals.

Post hoc statistical analysis determined that the predictive contribution of both spirometry and flow interruption transients were statistically significant (p<0.01) and that spirometry contributed a majority of the predictive power.

Day-to-day repeatability of TLC_{MB}. From the initial 300 subject pool, we selected 26 healthy subjects at random to assess day-to-day repeatability with a minimum of 12 days between the measurements. Day-to-day repeatability was expressed as a coefficient of variation (CV; Figure 4). For TLC_{MB}, the CV was 1.6% whereas for TLC_{pleth} the CV was 3.3%.

DISCUSSION

Based upon spirometry and flow interruption data taken together, here we apply data mining and machine learning to population-based observations in order to determine TLC in the individual subject. Within a highly heterogeneous population of subjects, we show that this approach yields accurate and reproducible determinations of TLC in the individual subject. Unlike the TLCpleth, the TLC_{MB} is not calculated based upon a physical principle or a mechanistic respiratory system model; physiological mechanisms were not a consideration. Instead, here we used inductive statistics and nonlinear systems identification, reminiscent of other applications of big data, to infer a relationship from which we could then make an accurate determination of TLC in the individual subject.

Summary of clinical results. Across the entire population studied, across specific patient subgroups, and across a prospective heterogeneous population, our results show that TLC_{MB} is accurate compared to TLC_{pleth}. Among our prospective cohort of 134 subjects, who were healthy or had varying severities of obstructive and restrictive diseases, TLC_{MB} correlated well with TLC_{pleth} (adjusted $r^2 = 0.795$) with a slope close to unity (slope = 0.977) (Figure 3). Furthermore, in a subset of healthy subjects, TLC_{MB} was appreciably more repeatable from day-to-day than was TLC_{pleth} (Figure 4), suggesting that TLC_{MB} might be particularly useful in longitudinal clinical management.

Comparison to helium dilution and CT imaging. How does TLC_{MB} compare to other alternative technologies, such as gas dilution or computed tomography (CT), to measure absolute

lung volumes? In a cohort of healthy, obstructive, and restrictive subjects, O'Donnell et al.⁶ performed Bland-Altman analyses to compare TLC measured using both helium dilution (TLC_{He}) and CT imaging (TLC_{CT}) to TLC measured using plethysmography (TLC_{pleth}).⁶ For TLC_{He} and TLC_{CT}, the analysis showed coefficients of variation of 18.9% and 15.6%, respectively, together with systematic biases and trends for increasing error in subjects with larger TLCs (Figure 5B and 5C). Although we studied a different cohort, and results may therefore not be strictly comparable, Bland-Altman analysis of TLC_{MB} showed a coefficient of variation of 12.3% in our prospective cohort (N = 134 subjects), no systematic bias, and no trend of increasing error with increasing TLC (Figure 5A). While each of these technologies is based on a different mechanism-of-action, and thus is not expected to mimic plethysmographic TLC faithfully in all subjects, TLC_{MB} values had the smallest deviations from those of TLC_{pleth}.

Limitations: The TLC_{MB} approach is limited in at least three important ways. First, TLC_{MB} is a population-based statistical approach. To the extent that specific populations might differ, different data training sets might be required. For example, pediatric populations, geriatric populations, or specific racial populations might require different training sets. Second, the flow interruption parameters improved the accuracy of TLCMB determinations by an amount that was small but nonetheless was highly significant statistically (p<0.01). Third, TLC_{MB} is designed to recapitulate as closely as possible TLC_{pleth}. But TLC_{pleth} is itself subject to artifacts and is known to be an imperfect measure of TLC.³¹⁻³⁴ As such, any biases or errors inherent to TLC_{pleth} are necessarily inherent in TLC_{MB}.

Conclusions. The NHLBI, ATS, and ERS have encouraged innovation in technologies to measure absolute lung volumes so as to attain improved accuracy, ease of use, and rapidity of testing³⁵, and have recommended rigorous testing to ensure no substantial differences in results compared with standard techniques. In all sub-populations tested, the MiniBoxTM performed in a manner that compared favorably with body plethysmography. Also, the day-to-day variability of TLC_{MB} was smaller than that of TLCs derived from helium dilution, CT imaging, or body plethysmography. Accordingly, this study establishes the validity of TLC_{MB} for rapid, accurate, and repeatable determination of TLC in a heterogeneous population of healthy adults and those with respiratory system diseases.

FOOTNOTES

Author Disclosures: All authors except Robert Brown, MD have a financial interest in

PulmOne Advanced Medical Devices, Ltd.

Funding: The study was sponsored by PulmOne Advanced Medical Devices, Ltd.

Previous Publications: Parts of this manuscript were presented previously in abstract form.

REFERENCES

1. Wanger J, Clausen JL, Coates A, et al. Standardisation of the measurement of lung volumes. Eur Respir J 2005;26:511-22.

2. Brown MS, Kim HJ, Abtin F, et al. Reproducibility of lung and lobar volume measurements using computed tomography. Acad Radiol 2010;17:316-22.

3. Coates AL, Peslin R, Rodenstein D, Stocks J. Measurement of lung volumes by plethysmography. Eur Respir J 1997;10:1415-27.

4. Garcia-Rio F, Dorgham A, Pino JM, Villasante C, Garcia-Quero C, Alvarez-Sala R. Lung volume reference values for women and men 65 to 85 years of age. Am J Respir Crit Care Med 2009;180:1083-91.

5. Hankinson JL, Stocks J, Peslin R. Reproducibility of lung volume measurements. Eur Respir J 1998;11:787-90.

6. O'Donnell CR, Bankier AA, Stiebellehner L, Reilly JJ, Brown R, Loring SH. Comparison of plethysmographic and helium dilution lung volumes: which is best for COPD? Chest 2010;137:1108-15.

7. Quanjer PH, Stocks J, Polgar G, Wise M, Karlberg J, Borsboom G. Compilation of reference values for lung function measurements in children. Eur Respir J Suppl 1989;4:184S-261S.

8. Stocks J, Quanjer PH. Reference values for residual volume, functional residual capacity and total lung capacity. ATS Workshop on Lung Volume Measurements. Official Statement of The European Respiratory Society. Eur Respir J 1995;8:492-506.

9. Dubois AB, Botelho SY, Bedell GN, Marshall R, Comroe JH, Jr. A rapid plethysmographic method for measuring thoracic gas volume: a comparison with a nitrogen washout method for measuring functional residual capacity in normal subjects. The Journal of clinical investigation 1956;35:322-6.

10. Farre R, Peslin R, Rotger M, Barbera JA, Navajas D. Forced oscillation total respiratory resistance and spontaneous breathing lung resistance in COPD patients. Eur Respir J 1999;14:172-8.

11. Lutchen KR, Jackson AC. Statistical measures of parameter estimates from models fit to respiratory impedance data: emphasis on joint variabilities. IEEE Trans Biomed Eng 1986;33:1000-9.

12. Jackson AC, Lutchen KR, Dorkin HL. Inverse modeling of dog airway and respiratory system impedances. J Appl Physiol 1987;62:2273-82.

13. Lutchen KR, Jackson AC. Reliability of parameter estimates from models applied to respiratory impedance data. J Appl Physiol 1987;62:403-13.

14. Dorkin HL, Lutchen KR, Jackson AC. Human respiratory input impedance from 4 to 200 Hz: physiological and modeling considerations. J Appl Physiol 1988;64:823-31.

15. Chalker RB, Celli BR, Habib RH, Jackson AC. Respiratory input impedance from 4 to 256 Hz in normals and chronic airflow obstruction: comparisons and correlations with spirometry. Am Rev Respir Dis 1992;146:570-6.

16. Jackson AC, Neff KM, Dorkin HL, Lutchen KR. Interpretation of respiratory input impedance in healthy infants. Pediatr Pulmonol 1996;22:364-75.

17. Frey U, Suki B, Kraemer R, Jackson AC. Human respiratory input impedance between 32 and 800 Hz, measured by interrupter technique and forced oscillations. J Appl Physiol 1997;82:1018-23.

18. Cliff IJ, Evans AH, Pantin CF, Baldwin DR. Comparison of two new methods for the measurement of lung volumes with two standard methods. Thorax 1999;54:329-33.

19. Fu WJ. Penalized Regressions: The Bridge versus the Lasso. Journal of Computational and Graphical Statistics 1998;7:397-416.

20. Tibshirani R. Regression shrinkage and selection via the Lasso. J R Statist Soc B 1996;58:267-88.

21. Picard RR, Cook RD. Cross-Validation of Regression Models. Journal of the American Statistical Association 1984;79:575-83.

22. Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. Proceedings of the 14th international joint conference on Artificial intelligence - Volume 2. Montreal, Quebec, Canada: Morgan Kaufmann Publishers Inc.; 1995:1137-43.

23. Pellegrino R, Viegi G, Brusasco V, et al. Interpretative strategies for lung function tests. The European respiratory journal 2005;26:948-68.

24. Miller MR, Hankinson J, Brusasco V, et al. Standardisation of spirometry. The European respiratory journal 2005;26:319-38.

25. Tibshirani R. Regression shrinkage and selection via the lasso: a retrospective. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 2011;73:273-82.

26. Bickel P, Freedman D. Some Asymptotic Theory for the Bootstrap. The Annals of Statistics 1981;9:1196-217.

27. Efron B. Bootstrap Methods: Another Look at the Jackknife. The Annals of Statistics 1979;7:1-26.

28. Singh K. On the Asymptotic Accuracy of Efron's Bootstrap. The Annals of Statistics 1981;9:1187-95.

29. Efron B, Tibshirani R. Improvements on Cross-Validation: The 632+ Bootstrap Method. Journal of the American Statistical Association 1997;92:548-60.

30. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307-10.

31. Brown R, Hoppin FG, Jr., Ingram RH, Jr., Saunders NA, McFadden ER, Jr. Influence of abdominal gas on the Boyle's law determination of thoracic gas volume. J Appl Physiol 1978;44:469-73.

32. Brown R, Ingram RH, Jr., McFadden ER, Jr. Problems in the plethysmographic assessment of changes in total lung capacity in asthma. Am Rev Respir Dis 1978;118:685-92.

33. Brown R, Scharf S, Ingram RH, Jr. Nonhomogeneous alveolar pressure swings in the presence of airway closure. J Appl Physiol 1980;49:398-402.

34. Brown R, Scharf SM, Ingram R, Jr. Nonhomogeneous alveolar pressure swings: effect of different respiratory muscles. J Appl Physiol 1982;52:638-41.

35. Consensus statement on measurement of lung volumes in humans. In: Clausen JL, Wanger JS, eds. NHLBI workshop12 Nov 2003.

FIGURE LEGENDS

Figure 1: Overview of the MiniBoxPFT flow-interruption device and measurement

procedure. a) Photograph of the MiniBox flow-interruption device during operation. The device sits on a table and the subject breathes tidally through a bacterial filter. **b**) Schematic illustration of the components of the MiniBox flow-interruption device, depicting the relative positions of the cylindrical container, valve, and flowmeter. Arrows indicate the direction of flow during inspiration. **c**) Schematic illustration of the breathing pattern required for lung volume measurement using the MiniBox flow-interruption device. Increasing lung volume is shown on the vertical axis and time increases to the right. During tidal breathing, brief interruptions are triggered in the vicinity of mid-inspiration (dots). After a minimum of 25 such interruptions or a maximum of 150 seconds, the subject then inhales maximally to TLC twice (double inspiratory capacity) and then exhales slowly to RV (slow expiratory vital capacity). In the illustration, only 9 interruptions are shown due to space limitations.

Figure 2: Comparison of TLC_{MB} and TLC_{pleth} in the initial 300 subject cohort. Top row: Scatter plots of plethysmographic TLC (TLC_{pleth}) vs. MiniBox TLC (TLC_{MB}) for all subjects (**A**), healthy subjects only (**B**), obstructed subjects only (**C**), and restricted subjects only (**D**). Males are represented by closed circles and females are represented by open circles. For subjects that were measured more than once on the device, the TLC is presented as the average value of all measurements. The dashed lines represent the unity line and the dotted lines represent the confidence intervals. The linear regression equation and the adjusted R² are displayed within each graph. **Bottom row:** Associated Bland-Altman plots comparing TLC_{MB} to TLC_{pleth} for all subjects (**E**), healthy subjects only (**F**), obstructed subjects only (**G**), and restricted subjects only (**H**). The dotted lines represent the mean bias while the dashed lines represent the upper and lower limits (±1.96*SD). The coefficient of variation (CV) is displayed within each graph.

Figure 3: Prospective validation of TLC_{MB} in an independent cohort of 134 subjects. Top row: Scatter plots of plethysmographic TLC (TLC_{pleth}) vs. MiniBox TLC (TLC_{MB}) for all subjects (A), healthy subjects only (B), obstructed subjects only (C), and restricted subjects only (D). Males are represented by closed circles and females are represented by open circles. For

subjects that were measured more than once on the device, the TLC is presented as the average value of all measurements. The dashed lines represent the unity line and the dotted lines represent the confidence intervals. The linear regression equation and the adjusted R^2 are displayed within each graph. **Bottom row:** Associated Bland-Altman plots comparing MiniBoxPFT-derived TLC_{stat} to plethysmographic TLC for all subjects (**E**), healthy subjects only (**F**), obstructed subjects only (**G**), and restricted subjects only (**H**). The dotted lines represent the mean bias while the dashed lines represent the upper and lower limits (±1.96*SD). The coefficient of variation (CV) is displayed within each graph.

Figure 4: Repeatability of TLC determined using MiniBoxPFT or body plethysmography on two different days. Top row: TLC measured on day 1 and day 2 with the TLC_{MB} (**A**) and TLC_{pleth} (**B**). **Bottom row** (**C**, **D**): TLC measured on day 2 normalized to the day 1 value. In 26 healthy subjects, the TLC_{MB} day-to-day repeatability was 1.6% compared to 3.3% for body plethysmograph.

Figure 5: Bland-Altman Plots for TLC values determined by MiniBox, CT, or multibreath helium dilution, as compared to TLC by body plethysmography. In comparison to plethysmographic TLC, the absolute errors in MiniBoxPFT-derived TLC_{MB} are smaller than the errors in TLC determined using CT (TLC_{CT}) or multibreath helium dilution (TLC_{He}), as measured in O'Donnell et al.⁶ The dotted lines represent the mean bias while the dashed lines represent the upper and lower limits (±1.96*SD). The coefficient of variation (CV) is displayed within each graph.

TABLE 1

Table 1: Basic Subject Characteristics.			
Characteristic	Entire	Initial Model	Prospective
	Dataset	Development	Validation
	(N = 434)	(N = 300)	(N = 134)
Male	265	185	80
	(61.1%)	(61.7%)	(59.7%)
Female	169	115	54
	(38.9%)	(38.3%)	(40.3%)
Age (years)*	45.8 ± 19.2	45.0 ± 20.1	47.8 ± 17.0
Height (cm)*	168.6 ± 9.9	168.2 ± 9.7	169.6 ± 10.4
Weight (kg)*	75.0 ± 17.3	73.6 ± 16.5	78.1 ± 18.8
Plethysmographic TLC (L)*	5.6 ± 1.4	5.6 ± 1.3	5.6 ± 1.5
Respiratory Condition: †			
Healthy	201	150	51
	(46.3%)	(50.0%)	(38.1%)
Obstructed	170	114	56
	(39.2%)	(38.0%)	(41.8%)
Restrictive	63	36	27
	(14.5%)	(12.0%)	(20.1%)
Obstruction Severity: †			
Mild	74	42	31
	(17.1%)	(14.0%)	(23.2%)
Moderate	240	179	62
	(55.3%)	(59.7%)	(46.4%)
Severe	120	79	41
	(27.6%)	(26.3%)	(30.4%)

* Values are means ± SD.

⁺ Defined according to ATS standards.²²



Figure 1:















