Single-Breath Diffusing Capacity for Carbon Monoxide Instrument Accuracy Across 3 Health Systems

Matthew J Hegewald MD, Boaz A Markewitz MD, Emily L Wilson MSc,
Heather M Gallo, and Robert L Jensen PhD

BACKGROUND: Measuring diffusing capacity of the lung for carbon monoxide (D_{LCO}) is complex and associated with wide intra- and inter-laboratory variability. Increased D_{LCO} variability may have important clinical consequences. The objective of the study was to assess instrument performance across hospital pulmonary function testing laboratories using a D_{LCO} simulator that produces precise and repeatable D_{LCO} values. METHODS: D_{LCO} instruments were tested with CO gas concentrations representing medium and high range D_{LCO} values. The absolute difference between observed and target D_{LCO} value was used to determine measurement accuracy; accuracy was defined as an average deviation from the target value of < 2.0 mL/min/mm Hg. Accuracy of inspired volume measurement and gas sensors were also determined. RESULTS: Twenty-three instruments were tested across 3 healthcare systems. The mean absolute deviation from the target value was 1.80 mL/min/mm Hg (range 0.24–4.23) with 10 of 23 instruments (43%) being inaccurate. High volume laboratories performed better than low volume laboratories, although the difference was not significant. There was no significant difference among the instruments by manufacturers. Inspired volume was not accurate in 48% of devices; mean absolute deviation from target value was 3.7%. Instrument gas analyzers performed adequately in all instruments. CONCLUSIONS: D_{LCO} instrument accuracy was unacceptable in 43% of devices. Instrument inaccuracy can be primarily attributed to errors in inspired volume measurement and not gas analyzer performance. D_{LCO} instrument performance may be improved by regular testing with a simulator. Caution should be used when comparing D_{LCO} results reported from different laboratories. *Key words: accuracy; quality control; simulator; single-breath diffusing capacity for carbon monoxide.* [Respir Care 2015;60(3):430–436. © 2015 Daedalus Enterprises]

Introduction

The single-breath diffusing capacity of the lung for carbon monoxide (D_{LCO}) is a common and clinically useful test providing a quantitative measure of gas transfer in the lungs. It is important for diagnosing and managing patients with respiratory disease and monitoring those exposed to drugs and toxins.[1-5] Because clinical decisions are made based on absolute values and changes in D_{LCO}, it is important to minimize intra-laboratory and inter-laboratory variability in test measurement. D_{LCO} test variability can be attributed to both technical (instrument, technician, test gases) and biological factors.[6,7] Significant intra-laboratory variability has been primarily attributed to inaccuracy of gas or volume measurement with variations in

temperature, barometric pressure, and breath-hold time contributing less to measurement inaccuracy.[6,8-11] Significant inter-laboratory variability in D$_{LCO}$ measurement has been reported and attributed to differences in devices, software, test gases, and testing procedure.[6,7,9,12,13] This study was designed to assess D$_{LCO}$ instrument performance and quantify inter-laboratory variability in D$_{LCO}$ measurement in Utah and southern Idaho. Our hypothesis was that inter-laboratory variability is large and that poorly performing instruments could be identified using a D$_{LCO}$ simulator (Hans Rudolph, Inc, Kansas City, Missouri) that are not identified with use of standard laboratory quality control measures. Some of these results have been published in abstract form.[14]

Methods

We tested 23 D$_{LCO}$ devices in 15 hospital-based pulmonary function testing (PFT) laboratories in Utah (22 instruments) and southern Idaho (1 instrument) using the same simulator over an 18-month period from 2009 to 2011; all measurements were made by 2 experienced technicians. The PFT laboratories were operated by 3 separate health systems (Intermountain Healthcare, University of Utah Health Systems, and Veterans Administration). The instrument manufacturers were: SensorMedics/Viasys/Carefusion Healthcare (Yorba Linda, California) for 13 devices, Collins/nSpire (Louisville, Kentucky) for 8 devices, and Medgraphics (Saint Paul, Minnesota) for 2 devices.

All instruments were maintained according to manufacturer instructions and calibrated according to manufacturer specifications on test days. Quality control measures used by the laboratories varied, but all included regular bio-control testing. Bio-control testing procedures were not standardized in most laboratories.

Simulator Testing

All D$_{LCO}$ simulations were performed using a single-breath D$_{LCO}$ simulator (Hans Rudolph) as previously described.[7] Previous reports indicate the simulator produces a precise and highly repeatable D$_{LCO}$ value in any instrument. The D$_{LCO}$ simulator utilizes 2 precision syringes containing precision gases with 2 concentrations of CO and tracer gas (methane, helium, or neon). For each simulation, a syringe is used to manually withdraw (inhale) a precise volume (V_I = 4.5 L) of gas from the D$_{LCO}$ instrument. After approximately 10 s, the second syringe is used to manually inject (exhale) the precision test gas mixture into the PFT instrument. The target D$_{LCO}$ value is then calculated using the EasyLab QC software provided with the simulator, while accounting for the breath-hold time,

QUICK LOOK

Current knowledge

The single-breath diffusing capacity for carbon monoxide (D$_{LCO}$) is a common useful test providing a quantitative measure of gas transfer in the lungs. It is important for diagnosing and managing patients with respiratory disease and monitoring those exposed to drugs and toxins. Because clinical decisions are made based on absolute values, it is important to minimize intra-laboratory and inter-laboratory variability in test measurement.

What this paper contributes to our knowledge

D$_{LCO}$ instrument accuracy was unacceptable in 43% of tested devices. Instrument inaccuracy was primarily due to errors in inspired volume measurement, not gas analyzer performance. D$_{LCO}$ instrument performance may be improved by regular testing with a simulator. Caution should be used when comparing D$_{LCO}$ results reported from different laboratories.

ambient temperature, and barometric pressure, using the following formula[7]:

$$D_{LCO} = CorrV_I/PB\text{-}47 \times F_{ITR}/F_{ETR} \times 60/t \times$$

$$(STPD\ factor) \times \ln(F_{ICO} \times F_{ETR}/F_{ECO} \times F_{ITR})\ [1]$$

in which CorrV$_I$ (corrected inhaled volume) is the measured V$_I$ minus instrument dead space, PB-47 is the barometric pressure minus water vapor pressure (saturated at 37°C), F$_{ITR}$ is the inspired tracer concentration, and F$_{ETR}$ is the expired tracer concentration, t is the breath-hold time (recorded for each measurement), 60/t adjusts the breath-hold time from seconds to minutes, STPD is the factor converting D$_{LCO}$ to standard temperature and pressure dry (STPD) conditions, F$_{ICO}$ is the inspired CO concentration, and F$_{ECO}$ is the expired CO concentration.

The CO and tracer gas concentrations used corresponded to a D$_{LCO}$ target of approximately 34 mL/min/mm Hg and 45 mL/min/mm Hg. Four simulations were performed at each gas concentration (8 total for each instrument); the first simulation from each gas concentration for an instrument was discarded to assure that each instrument system had been completely flushed.

Statistical Analysis

The absolute difference between the observed and calculated target D$_{LCO}$ value was used as the measure of accuracy. The 3 differences for each instrument were
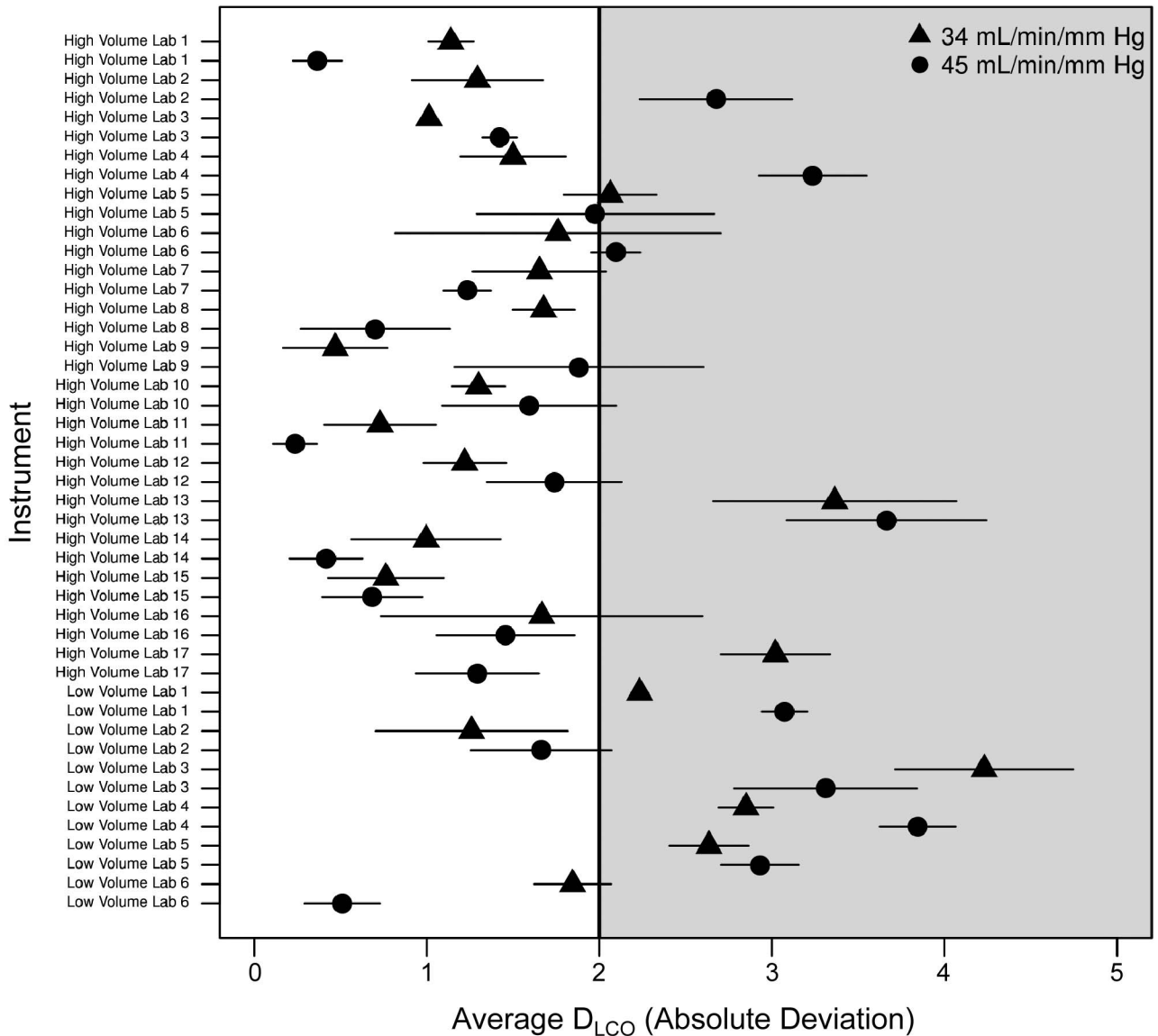
Fig. 1. The mean absolute deviation from the 2 D_{LCO} target values is listed by instrument. Error bars represent standard error of mean. Acceptable accuracy is defined as a deviation < 2.0 mL/min/mm Hg. The gray shaded area shows unacceptable deviation.

averaged, and unacceptable instrument accuracy was defined as an average difference greater than 2.0 mL/min/mm Hg from the target.[7] V_I accuracy was also measured by determining the absolute percent error for each device. Average volume measurement percent errors greater than 3.5% were considered unacceptable.[15] Accuracy of device gas sensors were assessed for carbon monoxide (CO) and inert gas (methane, helium, or neon) by measuring absolute percent error. Gas measurement errors exceeding 10% on average were considered unacceptable. Device reproducibility was assessed by determining the SD of the raw differences between each measured D_{LCO} value and the target value from each individual simulation and the raw percent errors for V_I, CO, and inert gas. The SD was

calculated for all instruments, by hospital PFT laboratory volume and by manufacturer. High volume laboratories were defined as performing more than 20 D_{LCO} tests per week. Chi-square tests for equality of proportions were used to compare the proportions of acceptable and unacceptable instruments by hospital volume and instrument manufacturer.

Results

Twenty-three D_{LCO} instruments were tested in 15 hospital-based PFT laboratories across 3 separate healthcare systems. The instruments tested encompassed approximately 60% of all those in hospital-based PFT laboratories
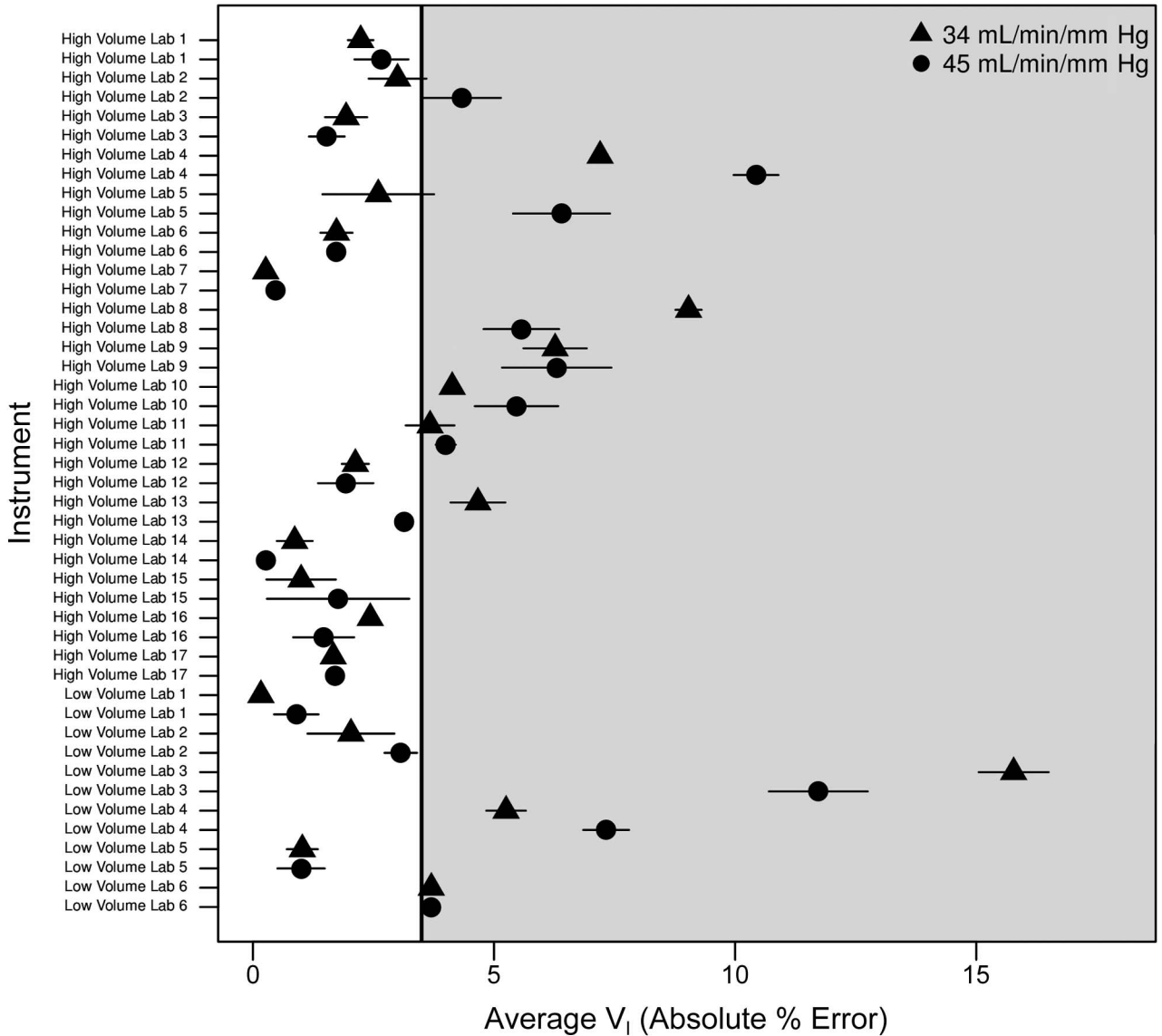
Fig. 2. The mean absolute percent error in inhaled volume (V$_I$) measurement is listed by instrument. Bars represent standard error of mean. Acceptable accuracy is defined an error < 3.5%. The gray shaded area shows unacceptable deviation.

in Utah. Overall, 10 of 23 devices (43%) failed to meet the standard for accuracy established for this study (average absolute deviation from one or both target values greater than 2.0 mL/min/mm Hg) (Fig. 1). For the 23 devices tested, the mean absolute deviation from the medium target value (approximately 34 mL/min/mm Hg) was 1.77 mL/min/mm Hg (range 0.47–4.23; SD 0.93) and the mean absolute deviation from the high target value (approximately 45 mL/min/mm Hg) was 1.83 mL/min/mm Hg (range 0.24–3.85; SD 1.11). Five devices failed to accurately measure both the medium and high target values, 3 devices failed in measurement of the high target value only, and 2 devices inaccurately measured only the medium target value. Most of the inaccurate instruments

underestimated the target value (8 of 10 inaccurate devices). High volume laboratories performed better than low volume laboratories with 11 of 17 (65%) devices from high volume laboratories meeting the accuracy standards compared with 2 of 6 (33%) devices from low volume laboratories. The difference in proportions was not statistically significant ($P = .39$), although this may be attributable to the small sample size. The proportion of devices achieving accurate results was not significantly different by manufacturer ($P = .33$). However, the small numbers of instruments from some manufacturers precluded a conclusive determination of whether there was a significant difference in accuracy between manufacturers.
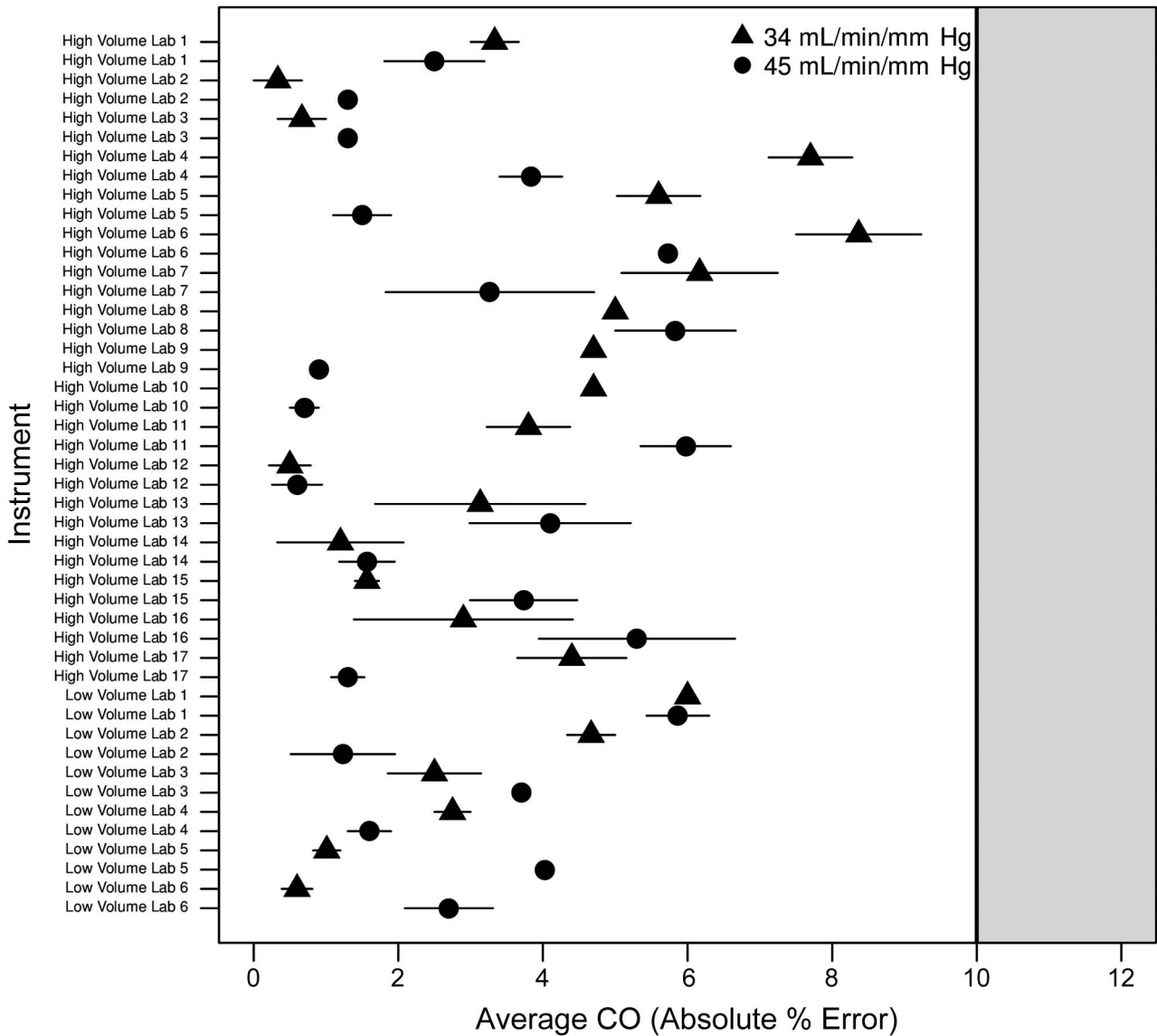
Fig. 3. The mean absolute percent error in CO measurement is listed by instrument. Bars represent standard error of mean. Acceptable accuracy is defined as an error < 10%. The gray shaded area shows unacceptable deviation.

Overall, inspired volume (V$_I$) measurement did not meet accuracy criteria: 11 of 23 devices (48%) had an average absolute percent error greater than 3.5%. The mean absolute percent error for V$_I$ measurement for all machines was 3.69% (range 0.17–15.78; SD 3.24) (Fig. 2).

The instrument gas analyzers performed adequately. All instruments had average CO measurements within 10% of the target value, with a mean absolute percent of 3.26% (range 0.33–8.37; SD 2.10) (Fig. 3). Average inert gas measurements were also acceptable with a mean absolute percent error of 1.52% (range 0.17–5.50; SD 1.28).

Device reproducibility, determined using actual differences between measured and target values for each mea-

surement and expressed in SD, is listed in Table 1. The instruments were less variable at the high target values for all test parameters. Variability among the high and low volume laboratories and manufacturers was not consistent across the measured parameters.

Discussion

This is the first study, to our knowledge, to test D$_{LCO}$ instrument accuracy in clinical pulmonary function laboratories across a geographic region. Consistency of test measurement across a geographic area is clinically relevant, as patients are often evaluated at different hospitals

Table 1. Standard Deviation for D$_{LCO}$ Test Parameters

Category	SD for D$_{LCO}$ Test Parameters			
	D$_{LCO}$ (mL/min/mm Hg)	V$_I$ (%)	CO (%)	Inert Gas (%)
Overall	1.97	4.03	3.59	1.89
Target D$_{LCO}$				
34 mL/min/mm Hg	2.01	4.29	3.95	1.91
45 mL/min/mm Hg	1.91	3.79	3.21	1.88
Hospital volume				
Low (< 20 tests/wk)	1.58	5.18	3.45	2.03
High (> 20 tests/wk)	1.81	3.31	3.54	1.57
Manufacturer				
Collins/nSpire	1.83	1.93	3.54	1.59
SensorMedics/Viasys	2.12	4.42	3.49	1.83
Medgraphics	0.96	2.94	1.37	0.66

Data are presented as SD of the raw differences between each measured D$_{LCO}$ value and the target value from each individual simulation for D$_{LCO}$ and the raw percent errors for V$_I$, CO, and inert gas.
D$_{LCO}$ = diffusing capacity of the lung for carbon monoxide
V$_I$ = inspired volume

in the same region. The important findings of this hospital-based PFT laboratory quality control study of D$_{LCO}$ instruments were as follows: (1) 43% of devices did not meet our accuracy criteria for D$_{LCO}$ measurement (within 2.0 mL/min/kg of target value), with the largest absolute errors exceeding 4.0 mL/min/mm Hg; (2) high volume laboratories generally performed better than low volume laboratories, although the difference was not statistically significant; (3) gas analyzers for CO and inert gas measurements performed adequately; and (4) device pneumotachographs generally performed poorly; 48% did not accurately measure V$_I$. All instruments tested were maintained and calibrated according to the manufacturer's specifications, and the laboratories routinely performed bio-control testing. However, bio-control testing was not well standardized in most laboratories, and there was often a lack of consistent personnel to serve as bio-controls. This suggests that routine quality control measures are not adequate for many pulmonary function laboratories, and routine testing with a D$_{LCO}$ simulator will identify poorly performing instruments and likely improve laboratory performance. Our findings reinforce the caution that care should be taken in comparing D$_{LCO}$ results obtained in different laboratories.

Prior studies have addressed D$_{LCO}$ accuracy using the same D$_{LCO}$ simulator employed in the present study. Jensen et al[7] measured mean absolute accuracy of 5 D$_{LCO}$ instruments, from different manufacturers, tested over 90 d, and reported that the mean deviation from the target D$_{LCO}$ ranged from 1.0 to 4.0 mL/min/mm Hg. Jensen et al[9] also tested 125 D$_{LCO}$ instruments involved in a multi-center international inhaled insulin trial using the same D$_{LCO}$

simulator as used in the present study. That study found that 25% of laboratories failed initial testing (mean absolute deviation from target value 1.54 mL/min/mm Hg); instrument failure was defined as a deviation from target value of greater than 3.0 mL/min/mm Hg. We defined instrument failure as a deviation from target value of greater than 2.0 mL/min/mm Hg, based on a large clinical PFT laboratory study that found that 91% of subjects had an absolute difference of < 2.0 mL/min/mm Hg between 2 D$_{LCO}$ measurements.[16] In our study, the overall mean deviation from the 2 target values was 1.80 mL/min/mm Hg.

Variability in D$_{LCO}$ measurement can be due to both technical (instrument, testing procedure, instrument gases) and biological/patient factors.[12] Technical sources of variability need to be minimized, both because they contribute more to overall D$_{LCO}$ testing inaccuracy and because they are more readily controlled than biological/patient factors. Jensen et al[6] assessed D$_{LCO}$ instrument accuracy and variability using normal subjects and a D$_{LCO}$ simulator over 3 months and reported that variability was primarily due to the instrument and that there were significant differences between manufacturers. D$_{LCO}$ instrument inaccuracy can be due to multiple factors. The main sources of error are inaccurate measurements of V$_I$, CO gas concentration and inert gas concentration (see Equation 1 above). Inaccurate CO gas analysis, often attributed to non-linearity of CO analyzers, was thought to be the main source of error in measurement of D$_{LCO}$.[12] In our study, the largest source of error was attributable to inaccurate measurement of V$_I$ (mean absolute error of 3.69%) rather than inaccuracy in CO gas analysis (mean absolute percent error of 3.26%). Inhaled volume accuracy needs to be closely monitored with daily 3-L syringe testing, and poorly performing pneumotachographs need to be replaced. The inert gas analyzers performed well (mean absolute percent error of 1.52%). Another potential source of inaccuracy results from errors in instrument gas concentration (F$_{ICO}$ and F$_{ITR}$ in Equation 1). Instrument gases were either locally provided or obtained from the manufacturer and are required to be accurate to within 2.0% of stated concentration. However, some laboratories used gravimetrically mixed test gases that are accurate to within 1.0%. Although we did not test instrument gases for accuracy, it is not likely that they were a source of significant error.

Simulator testing only assesses technical sources of variability due to the D$_{LCO}$ instrument. It does not assess variability due to the technician or biological/patient factors. Large errors can be observed if the technician does not allow adequate washout of the dead space or if gas leaks in the system are not identified. Biological/patient sources of error can include a failure to inhale to total lung capacity, prolonged inhalation or exhalation time, maintaining airway pressure during the breath-hold that is significantly higher or lower than atmospheric pressure, leaks at the mouth or nose

or cough during testing, and failing to correct for variations in hemoglobin level or elevated carboxyhemoglobin. Small variations in temperature and barometric pressure observed during testing have minimal effects on accuracy.[9]

It is important to minimize inaccuracy and variability in D$_{LCO}$ testing so changes due to disease process or changes due to therapeutic interventions can be detected. The clinical utility of D$_{LCO}$ measurement will be enhanced by maintaining accurate instruments. The magnitude of some of the errors in D$_{LCO}$ measurement observed in this study was clinically important. Inaccurate D$_{LCO}$ measurement may affect clinical management in a variety of clinical settings including idiopathic pulmonary fibrosis,[17] lung cancer,[3] patients being monitored for drug-induced pulmonary toxicity,[2,18] and sarcoidosis.[19,20]

Conclusions

Our results confirm that inter-laboratory variability in D$_{LCO}$ testing is large and that simulator testing can detect instruments performing poorly beyond those that can be identified with standard PFT laboratory quality control measures. More importantly, this study re-emphasizes the need for caution when comparing D$_{LCO}$ measurements made in different laboratories.

REFERENCES

1. Camus P, Martin WJ 2nd, Rosenow EC 3rd. Amiodarone pulmonary toxicity. Clin Chest Med 2004;25(1):65-75.
2. Dimopoulou I, Galani H, Dafni U, Samakovii A, Roussos C, Dimopoulos MA. A prospective study of pulmonary function in patients treated with paclitaxel and carboplatin. Cancer 2002;94(2):452-458.
3. Donington J, Ferguson M, Mazzone P, Handy J Jr, Schuchert M, Fernando H, et al. American College of Chest Physicians and Society of Thoracic Surgeons consensus statement for evaluation and management for high-risk patients with stage I non-small cell lung cancer. Chest 2012;142(6):1620-1635.
4. Martinez FJ, Flaherty K. Pulmonary function testing in idiopathic interstitial pneumonias. Proc Am Thorac Soc 2006;3(4):315-321.
5. Sun XG, Hansen JE, Oudiz RJ, Wasserman K. Pulmonary function in primary pulmonary hypertension. J Am Coll Cardiol 2003;41(6): 1028-1035.
6. Jensen RL, Teeter JG, England RD, Howell HM, White HJ, Pickering EH, Crapo RO. Sources of long-term variability in measurements of lung function: implications for interpretation and clinical trial design. Chest 2007;132(2):396-402.
7. Jensen RL, Teeter JG, England RD, White HJ, Pickering EH, Crapo RO. Instrument accuracy and reproducibility in measurements of pulmonary function. Chest 2007;132(2):388-395.
8. Crapo RO, Jensen RL. Standards and interpretive issues in lung function testing. Respir Care 2003;48(8):764-772.
9. Jensen R, Leyk M, Crapo R, Muchmore D, Berclaz PY. Quality control of DL,CO instruments in global clinical trials. Eur Respir J 2009;33(4):828-834.
10. Kangalee KM, Abboud RT. Interlaboratory and intralaboratory variability in pulmonary function testing: a 13-year study using a normal biologic control. Chest 1992;101(1):88-92.
11. Wanger J, Irvin C. Comparability of pulmonary function results from 13 laboratories in a metropolitan area. Respir Care 1991;36(12): 1375-1382.
12. Crapo RO, Jensen RL, Wanger JS. Single-breath carbon monoxide diffusing capacity. Clin Chest Med 2001;22(4):637-649.
13. Jensen RL, Crapo RO. Diffusing capacity: how to get it right. Respir Care 2003;48(8):777-782.
14. Hegewald MJ, Morris A, Gallo H, Li Y. DLco instrument accuracy across a Utah health system. Am J Respir Crit Care Med 2011;183: A6291.
15. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, et al. Standardisation of spirometry. Eur Respir J 2005;26(2):319-338.
16. Punjabi NM, Shade D, Patel AM, Wise RA. Measurement variability in single-breath diffusing capacity of the lung. Chest 2003;123(4): 1082-1089.
17. Flaherty KR, Andrei AC, Murray S, Fraley C, Colby TV, Travis WD, et al. Idiopathic pulmonary fibrosis: prognostic value of changes in physiology and six-minute-walk test. Am J Respir Crit Care Med 2006;174(7):803-809.
18. Goldschlager N, Epstein AE, Naccarelli G, Olshansky B, Singh B. Practical guidelines for clinicians who treat patients with amiodarone. Practice Guidelines Subcommittee, North American Society of Pacing and Electrophysiology. Arch Intern Med 2000;160(12):1741-1748.
19. Sietsma K. Sarcoidosis and the diffusing capacity for carbon monoxide. Sarcoidosis 1990;7(1):12-14.
20. Wasfi YS, Rose CS, Murphy JR, Silveira LJ, Grutters JC, Inoue Y, et al. A new tool to assess sarcoidosis severity. Chest 2006;129(5): 1234-1245.