Quality control of DL_{CO} instruments in global clinical trials

R. Jensen*, M. Leyk[#], R. Crapo*, D. Muchmore[#] and P.Y. Berclaz[#]

ABSTRACT: High inter- and intra-laboratory variability exists for the single-breath diffusing capacity of the lung for carbon monoxide (DL_{CO}) test. To detect small changes in diffusing capacity in multicentre clinical trials, accurate measurements are essential. The present study assessed whether regular DL_{CO} simulator testing maintained or improved instrument accuracy and reduced variability in multicentre trials.

The 125 pulmonary function testing laboratories that participated in clinical trials for AIR[®] Inhaled Insulin validated and monitored the accuracy of their DL_{CO} measuring devices using a DL_{CO} simulator, which creates known target values for any device. Devices measuring a simulated DL_{CO} different from target by $>3 \text{ mL} \cdot \text{min}^{-1} \cdot \text{mmHg}^{-1}$ failed testing and were serviced. Device accuracy was assessed over time and with respect to differences in several variables.

Initially, 31 (25%) laboratories had a DL_{CO} device that failed simulator testing. After fixing or replacing devices, 124 (99%) laboratories had passing devices. The percentage of failed tests significantly decreased over time. Differences in geographical region, device type, breath-hold time, temperature and pressure were not associated with meaningful differences in DL_{CO} device accuracy.

Regular diffusing capacity of the lung for carbon monoxide simulator testing allows pulmonary function testing laboratories to maintain the accuracy of their diffusing capacity measurements, leading to reduced variability across laboratories in multicentre clinical trials.

KEYWORDS: Accuracy, inhaled insulin, multicentre trials, quality control, simulator, single-breath diffusing capacity of the lung for carbon monoxide

Pulmonary function tests (PFTs), such as spirometry (forced expiratory volume in one second and forced vital capacity), lung volume measurements (total lung capacity) and the single-breath diffusing capacity of the lung for carbon monoxide (DL_{CO}), are essential for evaluating the efficacy of new therapies for asthma, chronic obstructive pulmonary disease and other lung diseases or to assess the safety of treatments administered by the pulmonary route [1, 2]. The DL_{CO} , which is already a valuable tool for disease diagnosis [3–6], is also an important measurement used to identify subtle adverse effects of inhaled therapies on the diffusing capacity in clinical trials.

Although the utility of the DL_{CO} is well established, an issue facing investigators in large-scale, multicentre clinical trials is the high inter- and intra-laboratory variability for this measurement [7–9]. Measuring the DL_{CO} is a complex process [10], largely automated in modern devices, and sources of inter-laboratory variability include differences in devices, test gases, software [11] and testing procedures. Within a

single laboratory, the accuracy of gas or volume measurements may drift over time and daily changes in temperature and barometric pressure need to be accounted for [8, 12]. Increases in variability reduce the sensitivity of DL_{CO} measurements to detect potential harmful or beneficial effects of an inhaled product in a test population.

To control the accuracy and reproducibility of DL_{CO} data in clinical trials, processes for quality control must be established. To this end, clinical trials for AIR[®] Inhaled Insulin (Alkermes, Inc., Cambridge, MA, USA) have used a centralised quality-assurance programme, in which each laboratory performing PFTs was qualified and individually quality controlled throughout the study period [13]. To reduce inter- and intra-laboratory variability due to instrumentation, PFT laboratories were required to first certify and then monitor the accuracy of their DL_{CO} devices using a simulator (Hans Rudolph, Kansas City, MO). The DL_{CO} simulator creates a precisely known, repeatable DL_{CO} value in any device. Comparing measured DL_{CO} values with

AFFILIATIONS

*LDS Hospital, Pulmonary Division, Salt Lake City, UT, and
[#]Lilly Research Laboratories, Eli Lilly and Company, Lilly Corporate Center Indianapolis, IN, USA.

CORRESPONDENCE

P.Y. Berclaz
Eli Lilly and Company
Lilly Corporate Center
Indianapolis
IN 46285
USA
Fax: 1 3172776351
E-mail: berclazpi@lilly.com

Received:

June 16 2008

Accepted after revision:

November 25 2008

STATEMENT OF INTEREST

Statements of interest for all of the authors and the study itself can be found at www.erj.ersjournals.com/misc/statements.dtl

For editorial comments see page 722.

European Respiratory Journal
Print ISSN 0903-1936
Online ISSN 1399-3003

the target DL_{CO} value allows the devices' accuracy and precision to be assessed. Using the DL_{CO} simulator on a regular basis facilitates early detection of technician or device issues that could lead to test failures.

The aim of the present study was to demonstrate that a quality-assurance programme that includes DL_{CO} simulator testing can establish and maintain the accuracy of DL_{CO} devices in global multi-site clinical trials, thereby reducing inter- and intra-laboratory variability. The results of simulator tests performed throughout clinical trials for human insulin inhalation powder are reported. Absolute differences between measured and target values, as well as trends over time in these differences, were analysed. In addition, the accuracy of DL_{CO} measurements was assessed with respect to variations in room temperature, barometric pressure, breath-hold time (BHT), region, machine type and target simulator values.

METHODS

Study design

In total, 125 PFT laboratories from North America (Canada and the USA), the European Union (EU; Belgium, Bulgaria, Croatia, Germany, Hungary, Italy, Poland and Portugal), Central/South America (Argentina, Brazil, Chile, Colombia, Mexico and Puerto Rico) and Asia (India, Philippines, Singapore and Taiwan) were recruited to clinical trials for AIR[®] Inhaled Insulin. DL_{CO} measurements were performed throughout the trials in study subjects as a safety evaluation.

During the trials, each laboratory used a DL_{CO} simulator to demonstrate that their device's measurements were accurate to within 3 mL·min⁻¹·mmHg⁻¹. Before patient testing, all laboratories were required to pass an initial certification test meeting this standard. After passing the initial certification, laboratories performed simulator testing weekly for 3 months, and then

bi-weekly for the duration of the trials (fig. 1). Devices not meeting this standard at any time were fixed or replaced and re-tested before being used.

Simulator testing

Simulator testing procedures have been described previously [11]. The DL_{CO} simulator delivered known inspired volumes (V_I) and exhaled concentrations of CO and tracer gas (helium, methane or neon). Two CO concentrations (low 0.1% and high 0.13%) and two V_I (3.5 and 4.5 L) were used to generate four tank gas/V_I combinations. Each simulation used one tank gas/V_I combination. Target DL_{CO} values for this combination were calculated while accounting for the BHT, temperature and barometric pressure. Target values for alveolar volume (V_A) were also calculated. The absolute differences between all measured and target values were used as measures of accuracy. Changes in accuracy were assessed over time. The effects of different regions, device type, BHT, temperature and pressure on the accuracy of DL_{CO} devices were also analysed.

Statistical analyses

Data from all sites were combined. Outliers were removed from the data set. These data points represented nonphysiological values or values that would be impossible to obtain in a simulator test and probably were the result of a data entry error or a leak in the connections between the simulator and the DL_{CO} measuring device. Outliers were defined as follows: measured and target DL_{CO} values <5 mL·min⁻¹·mmHg⁻¹ or >60 mL·min⁻¹·mmHg⁻¹; V_A values <1 L or >7 L; V_I values less than the simulated V_I or >7 L; simulated expired CO concentration >100%; measured expired CO concentration <5% or >120% of the simulated CO concentration; simulated expired tracer gas concentrations <0.10%; measured expired tracer gas concentrations <50% or >150% of the simulated value.

The Cochran–Armitage trend test was used to detect trends over time in the percentage of failed simulation tests. To assess changes in accuracy over time or the effects of regions, device type, BHT, temperature and pressure on DL_{CO} accuracy, a repeated measures model was fitted using sites as the random effect and the following covariates: years from initial certification test and tank gas/V_I combination; region and the region and tank gas/V_I combination interaction (to assess effect of regions); device type and the device type and tank gas/V_I combination interaction (to assess effect of devices); BHT and the BHT and tank gas/V_I combination interaction (to assess effect of BHT); temperature, pressure and the temperature and tank gas/V_I combination interaction (to assess the effects of temperature and pressure). Compound symmetry was used as the covariance structure. Interaction terms that were not significant at the 0.05 level were removed from the model. A similar model was used to assess changes in accuracy for V_A, V_I and CO and tracer gas concentrations with the following covariates: years from initial certification test and tank gas/V_I combination. Statistical significance was concluded when the p-value for two-sided test was <0.05.

RESULTS

Initial DL_{CO} device performance in pulmonary function laboratories

A total of 125 PFT laboratories were recruited to the present study. Initially, 31 (25%) sites had DL_{CO} devices that measured

FIGURE 1. Each pulmonary function laboratory was required to certify the accuracy of their diffusing capacity of the lung for carbon monoxide (DL_{CO}) measuring device using a simulator. Each test used one of four tank gas/inspired volume (V_I) combinations: 4.5 or 3.5 L V_I and high or low concentration of CO. If a laboratory found its device to measure a DL_{CO} different from the target by >3 mL·min⁻¹·mmHg⁻¹, it used the results of the simulator test and additional testing to isolate the source of the malfunction. After certifying a device, patient testing could begin. Regular simulator testing continued for the duration of the trials.

a DL_{CO} value $>3 \text{ mL}\cdot\text{min}^{-1}\cdot\text{mmHg}^{-1}$ different from the target and were considered to have failed. Of the sites that failed the initial certification, 24 (19.8%) passed after replacing parts and/or repairing the equipment. Devices at six (4.8%) other sites passed after new equipment was purchased. One (0.8%) site was unable to certify its DL_{CO} device and was disqualified from the study. After correcting device issues or purchasing new equipment, 124 sites (99%) had devices that were measuring the DL_{CO} within $3 \text{ mL}\cdot\text{min}^{-1}\cdot\text{mmHg}^{-1}$ of the simulated value.

DL_{CO} device performance over time

After the initial certification test, these 124 PFT laboratories tested their equipment using the simulator weekly for 3 months and then bi-weekly for the duration of the trials. When devices were found to be malfunctioning, they were serviced. Data was collected from 9,083 DL_{CO} simulation tests and for up to 3 yrs after certification. After outliers were removed, data for 9,025 DL_{CO} measurements, 7,271 VA measurements, 8,810 V_I measurements, 7,770 expired CO concentration measurements and 8,300 expired tracer gas concentration measurements were analysed.

In figure 2, the percentage of all simulator tests in which a device failed is displayed in 2-month intervals from the date of the initial certification test. After a transient increase from 5.0% in the first interval (0–2 months) to 7.1% in the second interval (2–4 months), the proportion of test failures significantly decreased to 3.1% for all tests conducted between 12 months and up to 36 months from the initial certification test ($p<0.001$).

To more closely examine how well sites maintained the accuracy of their device, the results of all simulator tests were analysed over time. A regression line relating the absolute difference from the target DL_{CO} value for each test and the time since the initial certification test for each site was fitted to data from all four tank gas/V_I combinations (fig. 3). At first, the mean $\pm$ SE absolute difference from target for all tests was $1.54 \pm 0.05 \text{ mL}\cdot\text{min}^{-1}\cdot\text{mmHg}^{-1}$ and decreased significantly by

FIGURE 2. The percentage of simulated measurements that were different from the target by $>3 \text{ mL}\cdot\text{min}^{-1}\cdot\text{mmHg}^{-1}$ was plotted in 2-month time intervals after the initial certification test. The percentage of failed tests decreased significantly over time. Cochran–Armitage trend test, two-sided probability: $p<0.001$.

$-0.13 \text{ mL}\cdot\text{min}^{-1}\cdot\text{mmHg}^{-1}\cdot\text{yr}^{-1}$ ($p=0.0002$). A similar analysis was used to examine trends in the accuracy of the measured values for V_I, V_A, expired CO and tracer gas concentrations. A decrease in the mean $\pm$ SE absolute difference between measured and target V_A and V_I of $-13.03 \pm 4.61 \text{ mL}\cdot\text{yr}^{-1}$ ($p=0.0047$) and $-16.14 \pm 2.07 \text{ mL}\cdot\text{yr}^{-1}$ ($p<0.0001$), respectively, was observed (table 1). There was no significant change in the accuracy of measured expired CO concentration or tracer gas concentration.

DL_{CO} device accuracy associated with each region, tank gas/V_I combination and device type

The results of the simulator testing were also used to examine the effects of variations in several different conditions on the accuracy of DL_{CO} devices. For geographical regions, the average absolute differences between measured and target simulator values ranged from 1.1 to $2.0 \text{ mL}\cdot\text{min}^{-1}\cdot\text{mmHg}^{-1}$ (North America 1.2–1.7, Central/South America 1.2–1.8, EU 1.2–1.8 and Asia 1.1–2.0). For each tank gas/V_I combination that was used, the average absolute difference between the measured and target ranged from 1.2 to $1.7 \text{ mL}\cdot\text{min}^{-1}\cdot\text{mmHg}^{-1}$ (table 2).

Each site used a DL_{CO} measuring device of their choice for the duration of the trial, giving a total of 12 device types used across all of the sites. For most of the DL_{CO} device types, the average absolute difference between the measured and target DL_{CO} value ranged from 0.6 to $2.1 \text{ mL}\cdot\text{min}^{-1}\cdot\text{mmHg}^{-1}$ (fig. 4). Only one of the devices made measurements that were on average $3.7 \text{ mL}\cdot\text{min}^{-1}\cdot\text{mmHg}^{-1}$ different from the target at one of the four tank gas/V_I combinations (V_I 4.5, high concentration of CO).

DL_{CO} device accuracy over ranges of pressure, BHT and temperature

Variations in barometric pressure (range ~ 550 – 800 mmHg) and BHT (range ~ 1 – 15 s) did not significantly affect the accuracy of DL_{CO} devices ($p=0.18$ and $p=0.24$, respectively). Although there was a statistically significant effect of variations in temperature (range ~ 10 – 37°C) on DL_{CO} accuracy, this was only true for two of the simulated target values (V_I 3.5, high concentration of CO, and V_I 4.5, high concentration of CO; table 3), and a mean $\pm$ SE change of -0.03 ± 0.01 or $-0.009 \pm 0.01 \text{ mL}\cdot\text{min}^{-1}\cdot\text{mmHg}^{-1}$ per increase of 1°C might not be considered a meaningful effect on the accuracy of DL_{CO} devices.

DISCUSSION

A reported high inter- and intra-laboratory variability [7–9] has limited the effective use of the DL_{CO} measurement in clinical trials. The analysis reported in the present study suggests that, in order to collect quality DL_{CO} data in large multicentre trials, PFT laboratories need to be monitored, and that regular DL_{CO} simulator testing is an effective method for maintaining device accuracy and precision for the duration of a clinical trial. By maintaining accurate devices, variability of DL_{CO} data is reduced, facilitating detection of small changes in diffusing capacity in a test population.

To the current authors' knowledge, the present study is the first and largest study of DL_{CO} simulator testing data collected during global, multicentre clinical studies. Data was collected from $>9,000$ simulator tests performed in 124 PFT laboratories,

FIGURE 3. The absolute difference between the target diffusing capacity of the lung for carbon monoxide value and the measured value decreased, on average, over time for all four tank gas/inspired volume (Vi) combinations. a) Vi 3.5 L, high concentration of CO; b) Vi 3.5 L, low concentration of CO; c) Vi 4.5 L, high concentration of CO; d) Vi 4.5 L, low concentration of CO. Regression lines are shown for change in absolute difference over time (years since initial certification test) on each graph. The slope of the line is $-0.13 \text{ mL CO}\cdot\text{min}^{-1}\cdot\text{mmHg}^{-1}\cdot\text{yr}^{-1}$.

in several different countries and using several different machines over several years.

Of note, 25% of PFT laboratories included in the present study were not measuring the DL_{CO} within a targeted degree of accuracy prior to corrective intervention. This has significant implications for any multi-site clinical trial that requires DL_{CO} measurements, because it suggests that a large percentage of

existing PFT laboratories have DL_{CO} devices that are either malfunctioning or are being used improperly. To reduce variability to an acceptable standard, simulator testing is necessary, in order to identify these laboratories and mitigate their deficiencies before inclusion into clinical trials.

The 124 sites that passed the certification improved their ability to maintain the accuracy of their devices within the pre-defined

TABLE 1		
Rate of change of absolute differences between measured and simulated values for volume and gas concentrations		
	Decrease in absolute difference per yr	p-value [#]
VA mL	-13.03 ± 4.61	0.005
Vi mL	-16.14 ± 2.07	<0.001
Expired CO ppm	6.58 ± 9.28	0.48
Expired tracer ppm	-0.50 ± 88.96	0.99

Data are presented as least squared mean $\pm$ SE, unless otherwise stated. VA: alveolar volume; Vi: inspired volume. [#]: calculated for difference from zero.

TABLE 2		
Average absolute differences for each tank gas/inspired volume (Vi) combination		
Tank gas/Vi combination	Average absolute difference [#] mL·min ⁻¹ ·mmHg ⁻¹	p-value [†]
Vi 3.5 L, high CO	1.4 ± 0.05	<0.001
Vi 3.5 L, low CO	1.2 ± 0.05	<0.001
Vi 4.5 L, high CO	1.7 ± 0.05	<0.001
Vi 4.5 L, low CO	1.5 ± 0.05	<0.001

Data are presented as least squared mean $\pm$ SE, unless otherwise stated. [#]: between measured and target value of diffusing capacity of the lung for carbon monoxide; [†]: calculated for difference from zero.

FIGURE 4. Accuracy was assessed for all 12 device types for measuring diffusing capacity of the lung for carbon monoxide (DL_{CO}) used by the sites in the present study, by calculating the mean of the absolute values of the difference between the measured and target DL_{CO} values. Mean absolute differences plus 95% confidence intervals are presented for each machine at each of the tank gas/inspired volume (Vi) combinations used. a) Vi 4.5 L, low concentration of CO; b) Vi 4.5 L, high concentration of CO; c) Vi 3.5 L, low concentration of CO; d) Vi 3.5 L, high concentration of CO. Manufacturers of the devices were as follows. Zan300: nSpire Health, Longmont, CO, USA. Spirotech: VIASYS Healthcare, Yorba Linda, CA, USA. SensorMedic: Sensor Medics, Yorba Linda. SandMKeystone: S&M Instruments Co., Doylestown, PA, USA. Piston: Piston, Budapest, Hungary. MediSoft: Medisoft, Dinant, Belgium. MedGraph: Medical Graphics Corp., St Paul, MN, USA. Lungtest 1000: MES, Krakow, Poland. Jaeger: Jaeger, Würzburg, Germany. GansHorn: Ganshorn Medizin Electronic, Niederlauer, Germany. CosMed: Cosmed, Rome, Italy. Collins: Ferraris Respiratory, Louisville, KY, USA.

guidelines ($<3 \text{ mL}\cdot\text{min}^{-1}\cdot\text{mmHg}^{-1}$ difference from target). Despite a slight initial increase, the percentage of failed simulation tests significantly decreased over time. In addition, sites improved the overall accuracy of their DL_{CO} devices over time. The observed initial increase in simulation test failures may reflect level of technician inexperience with respect to

device maintenance in the first few months after certification. It is likely that the positive and negative feedback that the simulator testing provided helped the technicians become more aware and capable with respect to maintaining and calibrating their DL_{CO} devices, resulting in a decrease in simulation test failures over time.

The simulator-based quality control programme was able to limit the effects of environmental and operational conditions that vary between laboratories and within a laboratory over time and, thus, data collected in different regions and different laboratories were highly comparable. On average, DL_{CO} measurements made in different regions and by different devices were <2 and $<2.1 \text{ mL}\cdot\text{min}^{-1}\cdot\text{mmHg}^{-1}$, respectively, and DL_{CO} device accuracy was stable over a range of barometric pressure, temperature and BHT.

DL_{CO} device accuracy was also stable across the tank gas/Vi combinations that were used. The tank gas/Vi combinations were used to create different DL_{CO} target values, which could be calculated after accounting for the local temperature, barometric pressure and the BHT of the device. That the averages of the absolute differences calculated between the measured and the simulated targets for each of the tank gas/Vi combinations were similar is consistent with what has been seen in clinical testing. PUNJABI *et al.* [14] analysed repeated DL_{CO} measurements made in patients with a range of

TABLE 3 Effect of temperature on device [#] accuracy with respect to four tank gas/inspired volume (Vi) combinations		
Tank gas/Vi combination	Change in absolute difference [‡] mL·min ⁻¹ ·mmHg ⁻¹ per 1°C	p-value ⁺
Vi 3.5 L, high CO	-0.03±0.01	0.03
Vi 3.5 L, low CO	-0.002±0.01	0.87
Vi 4.5 L, high CO	-0.05±0.01	<0.001
Vi 4.5 L, low CO	-0.009±0.01	0.44

Data are presented as least squared mean±SE, unless otherwise stated. [#]: for measuring diffusing capacity of the lung for carbon monoxide (DL_{CO}); [‡]: between measured and target value of DL_{CO} ; ⁺: calculated for difference from zero.

obstructive and restrictive ventilatory impairment. They showed that the absolute difference between repeated DL_{CO} measurements in individual patients remained stable over a wide range of DL_{CO} values.

With respect to patient testing in large, multicentre clinical trials, maintaining accurate DL_{CO} devices and diminishing inter- and intra-laboratory variability can lead to detection of smaller changes in DL_{CO} over time. This improvement in sensitivity is particularly important because the effects of some inhaled therapeutic agents are likely to be small. It has been reported that inhaled insulin products (insulin human (rDNA origin) inhalation powder and human insulin inhalation powder) on average caused reversible reductions in DL_{CO} of 0.5–1.0 mL·min⁻¹·mmHg⁻¹ after 12–24 weeks of treatment [13, 15–17], an amount difficult to reliably assess without optimum performance of DL_{CO} testing procedures across all study sites.

Although simulator testing reduces the variability of DL_{CO} measurements due to instrumentation, it does not address all sources of variability. The DL_{CO} manoeuvre is complex for study subjects and requires substantial effort and cooperation [10]. Each subject is an additional source of variability [8, 12], and it has been shown that the patient can account for between 30 and 60% of intra-subject variation in DL_{CO} measurements [18].

To minimise this source of variability in multicentre clinical trials, additional standardisation is necessary for patient testing procedures. Another study has shown that a centralised review of patient test results, along with standardised instrumentation and a demonstrated competency of technicians, can reduce the variability of DL_{CO} tests in patients in multicentre clinical trials [19]. A root mean-squared coefficient of variation (RMSCV) of 6.0% was found, compared with a previously reported RMSCV >9.0% in trials that did not use these methods [15, 16]. In clinical trials for AIR[®] Inhaled Insulin, only highly experienced PFT laboratories were used to collect the data used in the present study. In addition, patient test data underwent independent reviews throughout the trials. Combining these types of standards for patient testing with DL_{CO} simulator testing may further decrease the variability of DL_{CO} measurements in multicentre clinical trials and in clinical testing.

The results of the present study demonstrate that, in order to obtain high quality diffusing capacity of the lung for carbon monoxide data in multicentre, global clinical trials, appropriate monitoring of diffusing capacity devices is essential. Using a diffusing capacity simulator allows pulmonary function testing laboratories to maintain accuracy, thus reducing inter- and intra-laboratory variability of the measurement of diffusing capacity of the lung for carbon monoxide throughout the study.

ACKNOWLEDGEMENTS

The authors would like to acknowledge D. Shrom (Eli Lilly and Company, Indianapolis, IN, USA) and K.S. Shields (inVentiv Clinical Solutions LLC, Indianapolis) for support in the preparation of this manuscript.

REFERENCES

- 1 Muchmore DB, Gates JR. Inhaled insulin delivery – where are we now? *Diabetes Obes Metab* 2006; 8: 634–642.
- 2 Rosenstock J, Muchmore D, Swanson D, Schmitke J. AIR Inhaled Insulin System: a novel insulin-delivery system for patients with diabetes. *Expert Rev Med Devices* 2007; 4: 683–692.
- 3 Brunelli A, Refai MA, Salati M, Sabbatini A, Morgan-Hughes NJ, Rocco G. Carbon monoxide lung diffusion capacity improves risk stratification in patients without airflow limitation: evidence for systematic measurement before lung resection. *Eur J Cardiothorac Surg* 2006; 29: 567–570.
- 4 Aduen JF, Zisman DA, Mobin SI, *et al.* Retrospective study of pulmonary function tests in patients presenting with isolated reduction in single-breath diffusing capacity: implications for the diagnosis of combined obstructive and restrictive lung disease. *Mayo Clin Proc* 2007; 82: 48–54.
- 5 Ates F, Hacievliyagil SS, Karıncaoglu M. Clinical significance of pulmonary function tests in patients with acute pancreatitis. *Dig Dis Sci* 2006; 51: 7–10.
- 6 Mohsenifar Z, Lee SM, Diaz P, *et al.* Single-breath diffusing capacity of the lung for carbon monoxide: a predictor of PaO₂, maximum work rate, and walking distance in patients with emphysema. *Chest* 2003; 123: 1394–1400.
- 7 Wanger J, Irvin C. Comparability of pulmonary function results from 13 laboratories in a metropolitan area. *Respir Care* 1991; 36: 1375–1382.
- 8 Crapo RO, Jensen RL. Standards and interpretive issues in lung function testing. *Respir Care* 2003; 48: 764–772.
- 9 Kangalee KM, Abboud RT. Interlaboratory and intra-laboratory variability in pulmonary function testing. A 13-year study using a normal biologic control. *Chest* 1992; 101: 88–92.
- 10 American Thoracic Society. Single-breath carbon monoxide diffusing capacity (transfer factor). Recommendations for a standard technique – 1995 update. *Am J Respir Crit Care Med* 1995; 152: 2185–2198.
- 11 Jensen RL, Teeter JG, England RD, White HJ, Pickering EH, Crapo RO. Instrument accuracy and reproducibility in measurements of pulmonary function. *Chest* 2007; 132: 388–395.
- 12 Becklake MR. Concepts of normality applied to the measurement of lung function. *Am J Med* 1986; 80: 1158–1164.
- 13 Garg S, Rosenstock J, Silverman BL, *et al.* Efficacy and safety of preprandial human insulin inhalation powder versus injectable insulin in patients with type 1 diabetes. *Diabetologia* 2006; 49: 891–899.
- 14 Punjabi NM, Shade D, Patel AM, Wise RA. Measurement variability in single-breath diffusing capacity of the lung. *Chest* 2003; 123: 1082–1089.
- 15 Hollander PA, Blonde L, Rowe R, *et al.* Efficacy and safety of inhaled insulin (exubera) compared with subcutaneous insulin therapy in patients with type 2 diabetes: results of a 6-month, randomized, comparative trial. *Diabetes Care* 2004; 27: 2356–2362.
- 16 Quattrin T, Bélanger A, Bohannon NJ, Schwartz SL, Exubera Phase III Study Group. Efficacy and safety of inhaled insulin (Exubera) compared with subcutaneous

- insulin therapy in patients with type 1 diabetes: results of a 6-month, randomized, comparative trial. *Diabetes Care* 2004; 27: 2622–2627.
- 17** Skyler JS, Weinstock RS, Raskin P, *et al.* Use of inhaled insulin in a basal/bolus insulin regimen in type 1 diabetic subjects: a 6-month, randomized, comparative trial. *Diabetes Care* 2005; 28: 1630–1635.
- 18** Jensen RL, Teeter JG, England RD, *et al.* Sources of long-term variability in measurements of lung function: implications for interpretation and clinical trial design. *Chest* 2007; 132: 396–402.
- 19** Wise RA, Teeter JG, Jensen RL, *et al.* Standardization of the single-breath diffusing capacity in a multicenter clinical trial. *Chest* 2007; 132: 1191–1197.