Reproducibility of pulmonary function tests under laboratory and field conditions

R G LOVE, M D ATTFIELD,* AND K D ISLES†

From the Institute of Occupational Medicine, Edinburgh EH8 9SU, UK

ABSTRACT The reproducibility of pulmonary function tests in the laboratory and in a mobile field survey vehicle has been studied. Groups of laboratory workers were studied at base and a random sample of 38 coalminers was examined in the mobile laboratory. The intra-subject variability of some newer tests of lung function, including closing volume and maximum flow at low lung volumes, has been compared with that of well-established tests, such as lung volumes and forced expiratory volume from two measurements made more than one day apart. Most measurements were slightly less reproducible in the study of coalminers than in the laboratory personnel. Conventional tests, such as forced expiratory volume in one second, lung volumes, single breath CO transfer factor, and exercise ventilation were very reproducible, the coefficients of variation (cov) being generally between 5% and 10%. The closing volume test, maximum expiratory flow at low lung volumes, and the single breath N₂ index were less reproducible: cov between 15% and 39% in the miners. The forced expired time and volume of isoflow, measured only on laboratory workers, however, exhibited greater reproducibility than previously reported (cov = 10% and 15% respectively). It is suggested that, when assessing the repeatability of lung function tests, account should be taken of the circumstances in which the intra-subject variability was measured.

Pulmonary function tests, which are carried out on selected workers at the nine collieries that form the National Coal Board’s Pneumoconiosis Field Research programme, have recently been extended to include tests considered to reflect changes in the function of small airways as a means of clarifying the known relationship between functional changes and occupational dust exposure in coalminers. The present investigation reports the within-subject variability of the main tests of lung function performed in these surveys. The reproducibility of conventional and newer tests of lung function conducted on subjects under typical laboratory conditions and on naive industrial workers under field conditions has been investigated.

Table 1 Number, age, and smoking habits of participants in laboratory and field trials

<table>
<thead>
<tr>
<th>Test</th>
<th>No</th>
<th>Age, years Mean (range)</th>
<th>Smoking habit§</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>S</td>
</tr>
<tr>
<td>Laboratory</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Static lung volumes and TlCoss*</td>
<td>40</td>
<td>37-7 (19-64)</td>
<td>19</td>
</tr>
<tr>
<td>Forced expiration and closing volume</td>
<td>16</td>
<td>32-2 (23-53)</td>
<td>3</td>
</tr>
<tr>
<td>Exercise ventilation and TlCoss*</td>
<td>36</td>
<td>33-8 (21-59)</td>
<td>10</td>
</tr>
<tr>
<td>FET and V isoV†</td>
<td>19</td>
<td>30-2 (19-50)</td>
<td>7</td>
</tr>
<tr>
<td>Field</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All tests except FET and V isoV</td>
<td>382</td>
<td>40-2 (18-61)</td>
<td>22</td>
</tr>
</tbody>
</table>

*Diffusing capacity for carbon monoxide during single breath and steady state.
†FET = Forced expiratory time; V isoV = Volume of isoflow.
*Includes three subjects whose smoking history was unknown.
§S = Smoker; N = Non-smoker; X = Ex-smoker.
**Methods**

Before the introduction of each pulmonary function test in our mobile survey unit, measurements of between-visit and, if necessary, within-visit variability were made in the laboratory over a three-year period. A trial was also held in the mobile laboratory used for field surveys, and all the tests then currently used in the survey were assessed for reproducibility on working coalminers at a local colliery.

Male staff employed at this institute volunteered to act as subjects in the laboratory trials. Table I details the numbers, age, and smoking habits of the men carrying out each series of tests. No one was rejected specifically because of the presence of respiratory symptoms. In all cases the subjects attended the laboratory on two separate occasions at about the same time of day and where possible on the same day of the following week, although in practice the intervals between tests on any subject varied from one day to several weeks, mainly due to the unavailability of the subjects. It appears, however, that within- and between-day variations are not significant for measurements derived from flow-volume and closing volume manoeuvres.2 3

For the field trial 40 working miners were randomly selected from five ten-year age groups, and this was the sole basis for selection. Two of these men failed to attend leaving 38 subjects for the trial (table I). Each man attended the mobile laboratory at the same time of the same day of two successive weeks. Height, weight, and smoking habits were recorded at the time of the trials.

The following tests were carried out in separate trials in the laboratory but were performed at the same visit during the field trial.

**LUNG VOLUMES AND SINGLE BREATH TRANSFER FACTOR**

Lung volumes and single breath transfer factor (TLCOsb) were measured on an automated spirometer (Resparameter Mk 3, PK Morgan Ltd, Chatham, Kent). Vital capacity (VC), functional residual capacity (FRC) residual volume (RV), and total lung capacity (TLC) were determined once by helium dilution, and duplicate measurements of TLCO were made during breathholding4 and their mean calculated.

**FORCED EXPIRATORY VOLUME AND MAXIMUM EXPIRATORY FLOW-VOLUME CURVES**

Forced expiratory volumes and maximum expiratory flow-volume curves were recorded on an electronic waterless spirometer (Ohio Medical Products, Wisconsin—Model No 800). After one or two practice blows three forced expirations were traced as flow-volume (FV) curves on a fast response X-Y recorder (Hewlett-Packard, Model No 7045A), from which forced expiratory volume in one second (FEV1), forced vital capacity (FVC), and maximum expiratory flow at 50% and 25% of forced vital capacity (Vmax50 and Vmax25) were measured. The mean values of these variables were calculated, if the FVCs were within 200 ml of each other, by measuring 25% and 50% FVC from RV. If an otherwise technically satisfactory FVC was more than 200 ml but less than 10% below the highest FVC, Vmax50 and Vmax25 were determined using the FVC of that particular FV curve.

A one-second timer, calibrated from 50 cycles/second mains frequency, was activated by expulsion of the first 100 ml of air and traced. via an event marker, a one-second interval on the volume axis of the X-Y recorder, from which FEV1 was read.

Forced expired time (FET), defined as the time taken to expel a vital capacity breath, was measured by means of a digital timer, which was activated by expulsion of the first 100 ml of air and switched off when flow had fallen below about 50 ml/s.

The subject was subsequently connected to a gas mixture containing 80% He/20% O2, from which he took three near VC inspirations. At the end of the third inhalation he was instructed to exhale forcibly into the spirometer. Two, or more if necessary, flow-volume curves were recorded and the curve with the highest peak flow and FVC was used to compare expiratory flows for determination of the volume of isoflow (V isoV) by the method of Hutcheon et al.5 If FVC on He/O2 differed from FVC on air the relevant He/O2 curve was superimposed with the best air curve for measurement of V isoV. FET and V isoV were not determined on the miners.

It was only possible to measure FEV1 on a modified Gaensler water-filled spirometer during the field study, so that each subject, allocated alternately, had to blow three times into one spirometer followed by three blows into the second spirometer to record Vmax50 and Vmax25.

Although the same timer was used in both studies, a small systematic difference between the spirometers led to the mean FEV1 values from the Gaensler spirometer being slightly higher in the group of miners than if the Ohio spirometer had been used. This should not have affected the between visit variation of FEV1, however, since such systematic effects have been allowed for in the subsequent analysis.

**CLOSING VOLUME AND THE SINGLE BREATH NITROGEN INDEX**

Closing volume and the single breath nitrogen index
were estimated by the single breath nitrogen method. The outputs from the N₂ analyser (Ohio Medical Products, Wisconsin—Model No 700) and volume output from the spirometer (Ohio—Model No 800) were recorded on the X-Y recorder. It was possible to measure from this tracing the single breath N₂ index between 750 ml and 1250 ml and closing volume (CV). Closing capacity (CC) was calculated using the previously measured RV and TLC.

Inspiratory and expiratory flows were controlled by in-line resistances, designed to keep airflow below about 0.5 l/min. At these low flow rates differences in response-times of the N₂ meter and spirometer, if present, would not cause any significant distortion of the tracings. Any expirations that were clearly too rapid were rejected by the technician conducting the test.

The point of inflection was estimated according to the criteria suggested by the US National Heart and Lung Institute. Two observers read the traces independently. There were no consistent differences in the means and coefficients of variation determined by the observers but the values of only one observer (KS) are included in this analysis, since his readings were marginally more consistent.

VENTILATORY RESPONSE TO EXERCISE
The ventilatory response to exercise was measured by having the subjects pedal an electrically braked constant-load bicycle ergometer (Elema-Schonander, Stockholm) at successive five-minute loads of 50, 75, 100, and 125 watts at 60 revolutions a minute. The subject inspired a gas mixture of 0.05% carbon monoxide in air and the inspiratory minute volume \( \dot{V}_1 \) was measured on a Parkinson-Cowan dry gas meter. Expired air was collected during the fourth and fifth minutes of each exercise period for analysis of \( \text{FE}_{O_2} \) and \( \text{FE}_{CO} \), from which \( \text{no}_2 \) (O₂ consumption in mmol/min) and steady state transfer factor, TLCSS, were respectively derived. The latter was calculated by method 1 (assumed dead space) of Bates et al. Anatomical dead space (ml) was estimated from the sum of the subject's age in years and weight in pounds and added to the instrumental dead space (70 ml). Back-pressure of CO in the blood was determined before and after each workload by the subject rebreathing from a bag for six breaths.

The reproducibility of the tests was assessed by considering the variability within subjects not attributable to possible systematic differences between visits. It is assumed that the within-subject variability is the same for all subjects, and it was estimated by calculating the standard deviation of the between-visit differences recorded for each subject. Expressed as a coefficient of variation, this standard deviation provides a dimensionless (inverse) measure of reproducibility, after the removal of systematic between-visit effects.

Results

Reproducibility of each lung function test between visits is shown in tables 2-4. Coefficients of variation for static lung volumes and TLCSS (table 2) are generally less than 10%, except for RV and FRC in the field study.

Measurements of forced expiratory volumes and maximum expiratory airflow (Vmax) differed in their reproducibility (table 3). FEV₁ and FVC were more reproducible than Vmax 50 and Vmax 25 in both groups and Vmax was considerably less reproducible in the miners. FET and V isov measurements were not available from the field.

Table 2  Lung volumes and single breath transfer factor

<table>
<thead>
<tr>
<th>Variable</th>
<th>No of subjects</th>
<th>Visit 1</th>
<th>Visit 2</th>
<th>Between visits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>-------------------</td>
<td>------</td>
<td>-------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td><strong>Laboratory</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FRC (1)</td>
<td>40</td>
<td>3.82</td>
<td>3.71</td>
<td>0.77</td>
</tr>
<tr>
<td>VC (1)</td>
<td>40</td>
<td>5.27</td>
<td>5.20</td>
<td>0.82</td>
</tr>
<tr>
<td>RV (1)</td>
<td>40</td>
<td>2.12</td>
<td>2.06</td>
<td>0.59</td>
</tr>
<tr>
<td>TLC (1)</td>
<td>40</td>
<td>7.38</td>
<td>7.26</td>
<td>0.97</td>
</tr>
<tr>
<td>RV/TLC (%)</td>
<td>40</td>
<td>28.8</td>
<td>28.3</td>
<td>6.93</td>
</tr>
<tr>
<td>TLCSS (mmol/min/kPa)</td>
<td>40</td>
<td>10.65</td>
<td>10.69</td>
<td>2.06</td>
</tr>
<tr>
<td><strong>Field</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FRC (1)</td>
<td>38</td>
<td>3.76</td>
<td>3.78</td>
<td>1.07</td>
</tr>
<tr>
<td>VC (1)</td>
<td>38</td>
<td>5.16</td>
<td>5.11</td>
<td>0.97</td>
</tr>
<tr>
<td>RV (1)</td>
<td>38</td>
<td>2.06</td>
<td>2.17</td>
<td>1.02</td>
</tr>
<tr>
<td>TLC (1)</td>
<td>38</td>
<td>7.22</td>
<td>7.28</td>
<td>1.13</td>
</tr>
<tr>
<td>RV/TLC (%)</td>
<td>38</td>
<td>28.1</td>
<td>29.2</td>
<td>11.37</td>
</tr>
<tr>
<td>TLCSS (mmol/min/kPa)</td>
<td>38</td>
<td>10.45</td>
<td>10.69</td>
<td>2.14</td>
</tr>
</tbody>
</table>

*Mean of two values per subject at each visit.
†Between visits variation has been calculated after removal of any consistent visits effect.
‡Coefficient of variation: the denominator of this ratio is the grand mean.
Table 3  Measurements of ventilatory capacity and components of the maximum expiratory flow-volume curve and closing volume

<table>
<thead>
<tr>
<th>Variable</th>
<th>No of subjects</th>
<th>Visit 1</th>
<th>Visit 2</th>
<th>Between visits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>FEV₁ (l)</td>
<td>14</td>
<td>4.36</td>
<td>0.61</td>
<td>4.37</td>
</tr>
<tr>
<td>FVC (l)</td>
<td>16</td>
<td>5.55</td>
<td>0.68</td>
<td>5.51</td>
</tr>
<tr>
<td>FEV₁/FVC (%)</td>
<td>14</td>
<td>77.7</td>
<td>6.09</td>
<td>78.3</td>
</tr>
<tr>
<td>Vmax₁ (l/sec)</td>
<td>16</td>
<td>4.62</td>
<td>0.98</td>
<td>4.54</td>
</tr>
<tr>
<td>Vmax₂ (l/sec)</td>
<td>16</td>
<td>1.86</td>
<td>0.59</td>
<td>1.83</td>
</tr>
<tr>
<td>FET (sec)</td>
<td>19</td>
<td>5.97</td>
<td>3.42</td>
<td>5.88</td>
</tr>
<tr>
<td>V isoV (l)</td>
<td>16</td>
<td>0.98</td>
<td>0.61</td>
<td>1.04</td>
</tr>
<tr>
<td>V isoV (%)</td>
<td>16</td>
<td>17.3</td>
<td>12.63</td>
<td>18.3</td>
</tr>
<tr>
<td>CV/VC (%)</td>
<td>16</td>
<td>12.3</td>
<td>6.2</td>
<td>11.5</td>
</tr>
<tr>
<td>CC/TLC (%)</td>
<td>15</td>
<td>30.4</td>
<td>5.17</td>
<td>30.1</td>
</tr>
<tr>
<td>N₁ 750-1250 (%)</td>
<td>16</td>
<td>0.44</td>
<td>0.16</td>
<td>0.49</td>
</tr>
<tr>
<td>FEV₁ (l)</td>
<td>38</td>
<td>3.69</td>
<td>0.97</td>
<td>3.65</td>
</tr>
<tr>
<td>FVC (l)</td>
<td>35</td>
<td>4.68</td>
<td>0.82</td>
<td>4.69</td>
</tr>
<tr>
<td>FEV₁/FVC (%)</td>
<td>35</td>
<td>75.4</td>
<td>10.15</td>
<td>74.8</td>
</tr>
<tr>
<td>Vmax₁ (l/sec)</td>
<td>38</td>
<td>3.80</td>
<td>1.68</td>
<td>3.78</td>
</tr>
<tr>
<td>Vmax₂ (l/sec)</td>
<td>38</td>
<td>3.19</td>
<td>0.84</td>
<td>1.32</td>
</tr>
<tr>
<td>CV/VC (%)</td>
<td>35</td>
<td>15.2</td>
<td>7.52</td>
<td>15.9</td>
</tr>
<tr>
<td>CC/TLC (%)</td>
<td>35</td>
<td>36.5</td>
<td>12.99</td>
<td>38.2</td>
</tr>
<tr>
<td>N₁ 750-1250 (%)</td>
<td>37</td>
<td>0.88</td>
<td>0.64</td>
<td>0.78</td>
</tr>
</tbody>
</table>

†See footnotes in table 2.

Table 4  Ventilatory response to exercise and steady state transfer factor for carbon monoxide (TLCO35)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Workload (W)</th>
<th>Visit 1</th>
<th>Visit 2</th>
<th>Between visits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No</td>
<td>Mean</td>
<td>SD</td>
<td>No</td>
</tr>
<tr>
<td>V₁ (1/min)</td>
<td>50</td>
<td>36</td>
<td>24.1</td>
<td>3.8</td>
</tr>
<tr>
<td>NO₂ (mmol/min)</td>
<td>50</td>
<td>39.7</td>
<td>5.9</td>
<td>38.8</td>
</tr>
<tr>
<td>TLCO35 (mmol min/kPa)</td>
<td>50</td>
<td>12.4</td>
<td>2.17</td>
<td>12.63</td>
</tr>
<tr>
<td>V₁ 67% (l/min)</td>
<td>40.7</td>
<td>5.6</td>
<td>42.0</td>
<td>6.7</td>
</tr>
<tr>
<td>TLCO35 (mmol/min/kPa)</td>
<td>15.48</td>
<td>2.87</td>
<td>15.81</td>
<td>3.02</td>
</tr>
<tr>
<td>V₁ (1/min)</td>
<td>100</td>
<td>36</td>
<td>24.4</td>
<td>3.6</td>
</tr>
<tr>
<td>NO₂ (mmol/min)</td>
<td>100</td>
<td>39.7</td>
<td>5.4</td>
<td>39.3</td>
</tr>
<tr>
<td>TLCO35 (mmol/min/kPa)</td>
<td>100</td>
<td>12.4</td>
<td>2.17</td>
<td>12.63</td>
</tr>
<tr>
<td>V₁ 67% (l/min)</td>
<td>37</td>
<td>41.8</td>
<td>5.7</td>
<td>42.3</td>
</tr>
<tr>
<td>TLCO35 (mmol/min/kPa)</td>
<td>36</td>
<td>16.75</td>
<td>4.09</td>
<td>16.41</td>
</tr>
</tbody>
</table>

*V₁ 67 = V₁ at NO₂ of 67 mmol/min; TLCO35 = TLCO at V₁ of 35 l/min.
†See footnotes in table 2.

study but their coefficients of variation in the laboratory group were of the same order as those of Vmax.

Closing volume, as CV/VC, and the N₂ index exhibited a low level of reproducibility in both groups (table 3), particularly in the miners, but expressing the results as CC/TLC considerably reduced the coefficient of variation of this measurement.

The reproducibility of measurements made during steady state exercise, at 50 W and 100 W only, is shown in table 4 to represent one mild and one moderate workload. Results from the field trial showed good reproducibility between visits for V₁ and NO₂ but a very high systematic difference between visits for TLCO, which was higher on the first visit at each workload, especially the lower ones. While the coefficient of variation of TLCO at 100 W was acceptably low, however, that of TLCO at 50 W was much higher. Reproducibility was much better in the laboratory subjects, although there was a larger systematic difference in NO₂ for this group.
Discussion

The aim of the present investigation was to establish levels of reproducibility for various tests of lung function, particularly those measured in a mobile field laboratory with inexperienced subjects. Clearly in general the index of reproducibility, the between visits coefficient of variation, was higher for measurements carried out in the field. Nevertheless, the reproducibility of many of these field measurements is acceptable when compared with those made under laboratory conditions or with values previously published.

Coefficients of variation of 1.5% for VC,11 2% for TLC, 5.5% for RV,12 and 3.3% for FRC13 are some of the lower values available from earlier studies, confirming the generally repeatable nature of tests of static lung volume.

TLC by Stancescu,14 measured on an automated spirometer system, requires several different measurements for its calculation, yet its reproducibility in our hands is good, comparing favourably with published values of about 5%.15

The reproducibility of Vmax50 and Vmax25 in our laboratory subjects confirms the findings of other workers—for example a cov of 5.4% for Vmax50 and 9% for Vmax25 reported by Stancescu.15 He reported slightly higher values for patients with airways obstruction as other workers have done for normal subjects.16 17 There are several reasons why Vmax at low lung volumes might be less reproducible than FEV1 and FEV1/FVC. The flow-volume curve is subject to irregular and unpredictable fluctuations; there may be different volumes expired on each occasion; the subject's effort may vary and our use of the mean of three, rather than the maximum, values may increase the apparent within individual variability. These factors may well be of greater significance in the group studied in the field.

The recent reintroduction of FET as a screening test of lung function prompted us to examine its repeatability. Its usefulness has been questioned because of a high degree of variability within subjects: coefficients of 13%18 and 21-26%18 have been reported. Our results were more reproducible than these, possibly because of different methods of measurement. Previous workers have measured FET from a volume/time spirometric trace or by using the stop-watch/stethoscope method. Our method, which relies on a timing system de-activated by a threshold flow device, tends by its nature and in practice to give rise to rather longer expiratory times than published values even in normal subjects (typically 4-6 seconds). These were repeatable in a given subject however, although some of the longer times found in older men may have been related to lack of motivation as much as to abnormalities in airways.

On the other hand, the volume of isoflow (V isoV) is in principle likely to be more difficult to replicate successfully on different occasions in the same subject. Estimation of the isoflow point on the flow-volume curve may be influenced by test to test differences in (1) inspiratory VC; (2) expiratory VC; (3) level of maximum expiratory flow on air or He/O2 curves; (4) subjective assessment by the observer of the point of isoflow, which may in turn depend on the technique used for recording and the physical size of the recording scale; and (5) distribution of helium during the preceding three VC breaths and subsequent washout between tests, particularly in those with expiratory airflow limitation. Possibly because of some of the above factors the reproducibility of V isoV has been found to be poor. Sudlow et al.19 observed wide variation of V isoV/VC (cov = 60.8%) in five repetitions of this measurement carried out over a period of six months in a group of healthy non-smokers experienced in performing lung function tests, which was considerably greater than the variation reported here. Its usefulness as a sensitive discriminatory test of lung function in apparently healthy smokers, however, has previously been reported.20

The large variability of duplicate determinations of CV/VC in both laboratory and field populations confirms the findings of other workers.2 23 The reduction in the coefficient of variation but not the SD, when the results were expressed as CC/TLC (= (CV + RV)/(VC + RV)), is due to the relatively greater increase in the numerator causing an increased absolute mean value with no increase in the SD of the between-visit measurement, again confirming the findings of Becklake et al.2 There can be genuine differences, however, between determinations of the point at which inflection occurs. CV may be influenced by incomplete filling or emptying of the lungs and by daily variations in VC, which may not be related to the former, although it appears not to depend on diurnal rhythms nor to be amenable to training.2 3 We may have underestimated the variation of CC/TLC, since single estimations of RV (by helium dilution) were used.

The coefficient of variation of the N2 index was unusually large. One reason for this was the difficulty of reading accurately on the recording paper alterations of about 1 mm (representing 0·5% N2). When related to the average normal value of 0·5-1·0% (that is 1-2 mm on the record) small errors in reading these alterations would produce large c ovs. In abnormal states, however, the value of this index may increase by two- to ten-fold and hence
tends to eliminate this difficulty. The larger co-
gefficient of variation in our field study is unsatis-
factory and compares unfavourably with the labora-
tory results and other published values. For example, McCarthy et al21 found that the cov for
each of 20 subjects studied weekly for 10 weeks was
about 20%.

Ventilatory response to exercise was found to be a
very reproducible test, although less so at the lowest
workload in the miners. The small systematic error in
\(\text{FIO}_2\) of the laboratory subjects may have been due
to some undetected instrumental fault, because
omission of five subjects on one day and reanalysis
(not shown in table 4) eliminated the systematic
difference at all but the highest workloads.

Reproducibility of TlCO3 between visits has not
been reported before, but Holmgren,22 who used
arterial blood analyses to calculate physiological
dead space, has reported a cov of 7-2\% in duplicate
determinations made on the same day. Holmgren’s
analysis, which differed slightly from our own,
might have yielded a similar cov if conducted on
measurements made on separate days, but one would
expect his within-day cov to be less than our
between-day cov. The repeatability of this test in both
groups was generally good at the higher workloads
but there was a large systematic difference in the
miners. This difference is difficult to explain com-
pletely, since it appears to be unrelated to hyper-
ventilation on either occasion, or to differences in
ventilation, inspired gas concentrations, technicians,
or environmental factors.

As pointed out by West,23 however, the method of
calculation of TlCO used here, which is based on an
assumed dead space/tidal volume ratio and a
measured CO extraction, is apt to cause an over-
estimation of TlCO when the ratio is high—for
example, during low workloads or when CO
extraction is high, as in a fit young man. A com-
bination of these factors will give rise to an artificially
high TlCO especially at low workloads. It was found
that estimation of TlCO at a workload of 100 W
provided the most reproducible results and
would probably be our method of choice if a single esti-
mation were to be made during exercise.

The finding that most measurements were
generally less reproducible in the field study of
coalminers is not unexpected in view of the subjects’
lack of familiarity with the apparatus and the
operators. Caution should therefore be exercised in
applying levels of reproducibility which have been
derived from “trained” subjects in the laboratory to
a group of the general population or to industrial
workers. Although the more recently introduced
tests, FET and V isoV, were not examined for their
repeatability in the mobile laboratory, it is en-
couraging that these tests exhibited better repro-
ducibility than was indicated by previous reports.

Detels et al24 also measured lung function in a
laboratory by means of spirometry, single breath
oxygen test, and body plethysmography. Measure-
ments on their 94 adult subjects were repeated later
in a pulmonary function laboratory. They reached
broadly similar conclusions regarding the repeat-
ability of these tests.

The conventional index of reproducibility dis-
cussed above (cov) is convenient for comparing
different tests because it is dimensionless. It is not
to be regarded, however, as the sole criterion when
considering tests for possible use in epidemiological
surveys. The ability of a test to discriminate between
groups is a function not only of its reproducibility in
individuals but also of the scatter within groups and
of the numbers of individuals in each group. If costs
allow the numbers examined may be increased at
will. However, the specificity of a test to the physi-
ological feature of interest is equally important.
Whether or not a “specific test of small airways
function” with poor reproducibility is to be preferred
to a non-specific but highly repeatable test such as
FEV1 will be determined by their continued use in
long-term follow-up studies of lung function in
groups of closely monitored industrial workers.

We thank Mr J Birkett, Miss E Donaldson, Mr K
Sweetland, and Mr TD Wilson for their technical
help, Mrs MG Garden for typing the manuscript,
and all the miners and other volunteers who took
part in this study.

References

1. Rogan JM, Attfield MD, Jacobsen M, Rae S, Walker DD,
  Walton WH. Role of dust in the working environment
  in development of chronic bronchitis in British coal
  closing volume test in population studies. Sources of
  variation and reproducibility. Am Rev Respir Dis
  1975;111:141-7.
3. McFadden ER Jr, Holmes B, Kiker R. Variability of
  closing volume measurements in normal man. Am Rev
4. Meade F, Saunders MJ, Hyett F, Reynolds JA, Pearl N,
  Cotes JE. Automatic measurement of lung function.
5. Hutcheon M, Griffin P, Levison H, Zamel N. Volume of
  isoflow: a new test in detection of mild abnormalities of
6. Fowler WS. Lung function studies III. Uneven pulmonary
  ventilation in normal subjects and in patients with
7. Anthonisen NR, Dawson J, Robertson PC, Ross WRD.
  Airway closure as a function of age. Respir Physiol
8. Leblanc P, Ruff F, Milic-Emili J. Effects of age and body


19 Sudlow MF, Costello JF, Flanley DC, Millar J. Reproducibility and correlation of tests of "small airways disease." Bull Eur Physiopathol Respir 1975;11:16P-20P.


Reproducibility of pulmonary function tests under laboratory and field conditions.
R G Love, M D Attfield and K D Isles

doi: 10.1136/oem.37.1.63

Updated information and services can be found at:
http://oem.bmj.com/content/37/1/63

These include:

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/