Variation between observers in the estimation of airway resistance and thoracic gas volume

PHILIP W. LORD, ALAN G. F. BROOKS, AND JANET M. EDWARDS

From the MRC Environmental Hazards Unit, St Bartholomew's Hospital Medical College, Charterhouse Square, London EC1M 6BQ

Lord, P. W., Brooks, A. G. F., and Edwards, Janet M. (1977). Thorax, 32, 67–70. Variation between observers in the estimation of airway resistance and thoracic gas volume. Variations in the estimation of airway resistance (Rw) and thoracic gas volume (Vtg) made by different observers using body plethysmography have been investigated. Five observers determined Rw and Vtg in normal subjects using prerecorded signals so that intrasubject variations were eliminated. The slopes of the oscillograph tracings required to make the determinations were assessed using a cursor and a scale fitted to the oscilloscope. Significant and consistent differences in Rw determined by different observers were found, with a mean range of 37%. Great care should therefore be taken when comparing results obtained by different observers. The variation between observers measuring Vtg was significant and consistent but of much smaller magnitude than was the case for Rw, having a mean range of 12%.

In the course of experiments (Lawther et al., in preparation) we have seen that there are significant and substantial inter-observer variations in the measurement of airway resistance (Rw) and thoracic gas volume (Vtg) by whole body plethysmography using the methods described by Comroe et al. (1959). Significant differences in Rw as measured by two observers of the order of 0.005 kPa l⁻¹ s have been noted by Guvatt et al. (1967). Experiments have also shown that there is a tendency to estimate progressively lower values of Rw in successive measurements (Lawther et al., 1973; Lawther et al., in preparation). In a comparison of measurements of Rw obtained simultaneously by an automatic method and by an experienced observer (AB) it was shown that this tendency was more likely to be due to a perceptual error made by the observer than to genuine changes in Rw (Lord and Brooks, 1977). The experiment reported here was performed in order to investigate more fully the magnitude of observer differences and to see if the decline between successive readings could be demonstrated with other observers. Since Rw varies considerably within individuals both during the day and between different days (McDermott, 1966; Zedda and Sartorelli, 1971; Hruby and Butler, 1975), the comparisons reported here were made by recording the signals output from the plethysmograph and replaying them to each of the observers under test so that variations within subjects would be eliminated and so that each observer assessed Rw and Vtg from identical signals.

Method

As part of another experiment (Lord and Brooks, 1977), Rw and Vtg were assessed by constant volume plethysmography (Dubois et al., 1956a, b). The electrical signals representing plethysmograph pressure, airflow, and mouth pressure, which are normally displayed in pairs on the axes of an oscilloscope, were recorded using two channels of a four-channel Hewlett Packard 3960 analogue tape recorder. The signals were recorded at 2.38 cm sec⁻¹. The plethysmograph pressure signal was passed through a high-pass filter (cutoff at 0.05 Hz, 12 dB/octave) before recording to suppress slow drift due to thermal changes in the plethysmograph. Signals were recorded from 12 subjects on five occasions and each time 10 successive measurements of Rw and Vtg were made. Seven panting respirations were allowed for determining Rw and seven for Vtg. The subjects were instructed to pant with a frequency of 2 Hz and with a tidal volume of 200 ml around a lung
volume close to their functional residual capacity. These prerecorded signals were replayed on to the two axes of an oscilloscope and were all passed through two low-pass filters to remove tape noise (cut-off at 14 Hz, 12 dB/octave). Care was taken with the calibration of the tape recorder to ensure that the replayed signal levels were as close as possible to the prerecorded levels. One of the five occasions for each of the 12 subjects was selected at random for replaying. The 10 successive values of $R_{aw}$ and $V_{aw}$ were assessed by aligning a rotatable cursor parallel with the traces displayed on the oscilloscope screen and reading the slope of the trace from a circular scale. $R_{aw}$ was assessed between flows of 0 to $0.4$ litres sec$^{-1}$ in the inspiratory direction. From these slopes $R_{aw}$, $SR_{aw}$ (specific resistance), and $V_{aw}$ were calculated. All the estimates for a given subject were made successively by all the observers, whose order in assessing the readings was randomised so as to minimise the effects of any drift which might have occurred in the replayed signal levels (no drift was noticed).

The five observers were all members of the staff of the Environmental Hazards Unit. Observer AB was considered very experienced in this technique, having made thousands of measurements over 12 years. Observers PL and CB were very familiar with the technique, while observers JE and BB were less experienced and were given periods of practice before the experiment.

Results

The values of $V_{aw}$, $R_{aw}$, and $SR_{aw}$ were examined for observer differences by analysis of variance. In all these analyses $SR_{aw}$ behaved in the same manner as $R_{aw}$ and so only the results for $V_{aw}$ and $R_{aw}$ are discussed.

Both $V_{aw}$ and $R_{aw}$ showed highly significant observer and subject differences ($p<0.001$). The observer variations were systematic in that the rankings of the values obtained by the observers varied little between subjects (Fig. 1a, b); the means of the values of $R_{aw}$ and $V_{aw}$ obtained by BB, AB, and PL were similar.

The mean values of $R_{aw}$ obtained by the different observers for each subject varied substantially, and the ranges of these were proportional to the overall subject means. The range, divided by the subject’s overall mean, gave a measure of the observer variability independent of the level of $R_{aw}$; its average value over all subjects was 37%. This variation was larger in the female subjects (43%) than in the males (34%).

There was a significant decrease in $R_{aw}$ between the successive 10 measurements on all subjects and over all observers. This trend (Fig. 2a) was caused by consistently high first readings in each set of 10, and when these were excluded from the analyses the trend was no longer significant. The high first readings in sets appeared to be true high values in the subjects rather than an observer error, as the same trend was found when the same sets from the tape recordings were assessed by an automatic method (Lord and Brooks, 1977) and the results analysed for reading number variations.

The measurements of $V_{aw}$ showed less variation than $R_{aw}$ both between observers and with reading number. The variation between the different observers’ measurements was also proportional to the mean $V_{aw}$ of a subject, and the ratio of observer range to subject mean had an average value of 12% (Fig. 1b). No difference between these ratio values in male and female subjects was noted. There was a significant increase in $V_{aw}$ with reading number, which over all subjects amounted to 0-1 litre between the first and last readings. This trend was barely significant when the first reading was omitted from the analyses. The effect of reading number was independent of the observer, and the trend was also seen in the values of $V_{aw}$ made by the automatic method, indicating that it was a genuine increase in volume (Fig. 2b).

Discussion

This study supports the findings in our survey work (Lawther et al., in preparation) in showing significant and substantial differences in $R_{aw}$ and $V_{aw}$ as determined by different observers using the method of Comroe et al. (1959). The values of $V_{aw}$ were more consistent than those of $R_{aw}$, and our results show that great caution should be exercised when comparing measurements of $R_{aw}$ obtained by different observers from normal subjects during survey and experimental work. We have no results for patients and subjects with high airway resistance, but since the variability in $R_{aw}$ between observers in our series was proportional to $R_{aw}$ we see no reason to doubt that similar variation would be observed in subjects with much higher resistances than those studied here. The consistency of the differences between the observers as shown by their nearly uniform rankings for $V_{aw}$ and $R_{aw}$ with respect to subject and reading number indicates that they are perf...
Variation in the estimation of airway resistance and thoracic gas volume

Fig. 1 Ranges of operator means for each subject: (a) for $R_{aw}$ and (b) for $V_{tg}$. (See Fig. 2 for operator symbol codes.) Subjects JP, AK, AM, and MC females. Standard errors of the means are 0.009 kPa l$^{-1}$ s ($R_{aw}$) and 0.071 ($V_{tg}$).

Receiving and assessing the slopes in different ways. Such perceptual differences may be important in similar tasks such as compliance determination, and even in other fields such as radar tracking. A significant interaction term in our analysis of variance between subjects and operators may indicate that the characteristic loop shapes produced by different subjects affect the judgement of slopes by an operator.

The increased variability seen with increased $V_{tg}$ may have been due to the increased absolute error in slope for given angular error in the position of the circular cursor at large angles. For $V_{tg}$ the angles were in the range $34^\circ$ to $71^\circ$, and a $1^\circ$ error in alignment would give rise to errors of 0.08 1 and 0.56 1 in $V_{tg}$ respectively. A similar argument does not apply to $R_{aw}$, for the range in angle was $12^\circ$ to $33^\circ$, giving, at a fixed $V_{tg}$ of 3 1, errors of 0.001 kPa l$^{-1}$ s in $R_{aw}$ at both extremes.

We cannot explain the higher consistency of the $R_{aw}$ determinations in the male subjects.

In this study all the observers' means for each reading number followed the same pattern as that obtained by an automatic method (Fig. 2). There was little evidence of a linear trend between successive readings of $R_{aw}$ once the first in the sets of 10 had been excluded from the analyses, contrary to what was found in our other series (Lawther et al., 1973 and in preparation; Lord and Brooks, 1977). In this present series fewer sets of signals were examined and the lack of a significant trend may be due to this.

In all our studies the first reading of $R_{aw}$ within a set has been found to be high, and therefore it should be used with caution since it is not clear what is the significance of this anomaly.

We should like to thank our colleagues within the Air Pollution Unit for their cooperation and help with the work.
Fig. 2 Means for each operator for each of 10 successive readings: (a) for $R_{aw}$ and (b) for $V_{tg}$. Standard errors are 0.008 kPa 1⁻¹ s ($R_{aw}$) and 0.061 ($V_{tg}$).

References


Requests for reprints to: P. W. Lord, Environmental Hazards Unit, MRC Air Pollution Unit, Bartholomew's Hospital Medical College, Charterhouse Square, London EC1M 6BQ.