Automated detection of the phase III slope during inert gas washout testing

Christopher R. Stuart-Andrews,1 Vanessa J. Kelly,1,2,3,4 Scott A. Sands,5,6 Angela J. Lewis,1 Matthew J. Ellis,1 and Bruce R. Thompson1,2,3

1Physiology Service, Department of Allergy, Immunology and Respiratory Medicine, The Alfred Hospital, Melbourne; 2Department of Medicine, Nursing and Health Science, Monash University, Clayton, Victoria; 3Co-operative Research Council for Asthma and Airways, Glebe, New South Wales, Australia; 4Department of Medicine, Pulmonary and Critical Care Unit, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts; 5The Ritchie Centre for Baby Health Research, Monash Institute of Medical Research, Monash University, Clayton, Victoria, Australia; and 6Division of Sleep Medicine, Brigham and Women’s Hospital and Harvard Medical School, Boston, Massachusetts

Submitted 28 March 2011; accepted in final form 14 December 2011

In the current study we developed and critically evaluated an innovative methodology for automated detection of the phase III slope during inert gas washout testing. J Appl Physiol 112: 1073–1081, 2012. First published December 15, 2011; doi:10.1152/japplphysiol.00372.2011.—We describe a method to determine the phase III slope for the purpose of calculating indexes of ventilation heterogeneity, $S_{\text{acin}}$ and $S_{\text{cond}}$, from the multiple breath nitrogen washout test (MBNW). Our automated method applies a recursive, segmented linear regression technique to each breath of the MBNW test and determines the best point of transition, or breakpoint, between each phase of the washout. A sample set of 50 MBNW tests (controls, asthma, and COPD) was used to establish the conditions in which the phase III slope obtained from the automated technique best matched that obtained by two manual interpreters. We then applied our technique to a test set of 30 subjects (with an even number of subjects in each of the above groups) and compared these results against the manual analysis of a third independent manual interpreter. Indexes of ventilation heterogeneity were determined using both methods and compared. The phase III slopes determined by the automatic technique best matched the manual interpreter when the phase III slope was calculated from the phase II–III transition point plus the addition of 50% of the phase II volume to the end of the expiration. Calculation of the indexes $S_{\text{acin}}$ and $S_{\text{cond}}$ showed no overall difference between analysis methods in either $S_{\text{acin}} (P = 0.14)$ or $S_{\text{cond}} (P = 0.59)$ when the set threshold was applied to our automated analysis. Our analysis method provides an alternate means for rapid quantification of the MBNW test, removing operator dependence without alteration in either $S_{\text{acin}}$ or $S_{\text{cond}}$.

Inert gas washout; computer algorithm; alveolar slope

SINGLE AND MULTIPLE BREATH inert gas washout tests of peripheral airway function have been shown to provide potentially useful indexes of ventilation heterogeneity across a range of pulmonary diseases (3, 10, 13, 15, 16, 29, 30). Despite the additional information the multiple breath nitrogen washout (MBNW) test provides over tests such as spirometry, it has not become routine in pulmonary function laboratories in part because of the complex and time-consuming analysis required to generate indexes of ventilation heterogeneity. As a result of their infrequent use, interpretation of ventilation heterogeneity indexes has, as yet, not been well established.

One of the methods of MBNW analysis that is gaining acceptance measures the change in slope of the normalized nitrogen ($N_2$) alveolar plateau ($S_{\text{mbnw}}$) following consecutive 100% oxygen ($O_2$) inhalations using 1 liter tidal volumes ($V_T$) until a defined number of lung turnovers (TO) has been reached (8). Analysis of the change in $S_{\text{mbnw}}$ over the series of inhalations is used as the input variable to determine indexes of ventilation heterogeneity within the diffusive-convective and convective regions of the lung. Several methods of calculating these indexes have previously been proposed (9, 11, 26, 30, 31). However, the quality of calculated indexes of ventilation heterogeneity from MBNW testing is highly dependent on the accuracy and precision of the input variable, $S_{\text{mbnw}}$. The end point of the alveolar plateau is relatively easy to define as it coincides with the start of a subsequent inspiration. In contrast, the start point of the alveolar plateau is often difficult to determine as the transition region between phase II ($P_{II}$) and phase III of the washout is broad, particularly in the presence of lung disease (Fig. 1). Simply detecting the point of transition between phase II and phase III leads to difficulties, as the section of trace that is of interest is the alveolar plateau, which is further into expiration than this transition.

Previous studies have overcome the difficulty in identifying the phase III start point by measuring the gradient of the gas concentration between a fixed fraction (11) or fixed volume point (31) during expired volume and end expiration. However, all subjects have differing dead space volumes (12) and, therefore, the volume expired during the phase II portion of the washout varies between individuals. Starting the alveolar plateau calculation late in expiration will mean the slope will be calculated over a smaller volume, and artifacts such as cardio-genic oscillations would, therefore, have greater impact on the measured alveolar plateau. Trained scientists take this into account, and it is common for manual alteration of the alveolar plateau start point to occur to maximize the power of the linear regression to overcome cardio-genic oscillation effects (10, 31).

Unfortunately, manual adjustment adds considerable time to analysis of a MBNW test and can introduce an expert scientific bias into an otherwise quantitative measurement. Because of the increasing number of clinical respiratory laboratories utilizing the MBNW technique worldwide (10, 22, 25, 31), there is a need for an analysis technique capable of determining phase III slopes in an unbiased and reliable manner, leading to the automated calculation of ventilation heterogeneity indexes.

In the current study we developed and critically evaluated an analysis algorithm that can automatically determine the alveolar slope in a manner equivalent to that determined manually by the method of Verbanck et al. (31), which is becoming accepted in both the clinical and research setting (7).
Innovative Methodology

AUTOMATED ANALYSIS FOR THE DETECTION OF PHASE III GRADIENTS

Fig. 1. A typical expirogram recorded during inert gas washout testing. The phase I period reflects the anatomical dead space, phase II consists of a mixture of dead space volume and alveolar gas, phase III and the alveolar plateau consist of expired alveolar gas from which the slope may be used to assess ventilation heterogeneity.

METHODS

Following approval from our institution’s human research ethics committee, a total of 80 subjects were recruited (30 controls, 30 subjects with asthma, and 20 subjects with a clinical diagnosis of COPD). Written informed consent was obtained prior to any testing being conducted. All subjects underwent spirometric testing using standard lung function laboratory equipment (Platinum Elite series, Medgraphics, St. Paul, MN) according to ATS/ERS criteria (19) and a MBNW test. Asthma was defined by the National Asthma Education and Prevention Program guidelines (1), and all asthma subjects had standard lung function laboratory equipment (Platinum Elite series, Medgraphics, St. Paul, MN) according to ATS/ERS criteria (19) and a MBNW test. Asthma was defined by the National Asthma Education and Prevention Program guidelines (1), and all asthma subjects had baseline forced expiratory ratio (FER) below the lower limit of normal (14) with a significant (>12% and 200 ml) increase in either their forced expiratory volume in 1 s (FEV1) or forced vital capacity (FVC) following administration of bronchodilator (300 μg Salbutamol). COPD subjects had a clinical diagnosis of airflow obstruction with confirmation of low diffusing capacity for carbon monoxide and a fixed obstructive pattern on post-bronchodilator spirometry according to the Global Initiative for Chronic Obstructive Lung Disease (GOLD) criteria (21).

For the purpose of determining an automated algorithm for phase III slope detection, subject data were separated into two groups: a training group and a test group. The training group contained 50 subjects (20 asthmatics, 20 controls, and 10 subjects with COPD), whereas the test group contained the remaining 30 subjects (10 asthmatics, 10 controls, and 10 subjects with COPD).

For all subjects, MBNW was performed as described below followed by spirometric testing.

Equipment Setup

The MBNW circuit consisted of a double “bag in box” system, where one bag in the system was prefilled with 100% O2 prior to commencement of the test, whereas the other was used for the collection of expired gases as has been previously described (10, 31).

The MBNW test was performed with the subject in the seated position wearing a nose clip. Following a period of quiet breathing of ambient air, the subject was asked to inspire a series of 1-liter breaths of 100% O2 at their resting respiratory rate. Assistance in inspiring the requested volume was provided by way of a computer display. Pneumatic computer-controlled valves (8600 series, Hans-Rudolph, Shawnee, KS) controlled the direction of gas flow throughout the system. Nitrogen concentration was measured continuously at the subject’s mouth by an analyzer employing the emission photo spectroscopy method (AU9240–4032, Medgraphics) with constant suction through the analyzer being provided by a vacuum pump (UVlac rotary vacuum pump, G-100, Sinku Kiko, Yokohama, Japan). The N2 analyzer had been previously calibrated against a set of 9 O2/N2 National Association of Testing Authorities (NATA) approved calibration gases (ranging between 5.1 and 84.7% N2) and 100% O2 (0% N2). Tidal volumes were measured by volume displacement from within the bag in box system by use of a precalibrated pneumotachograph (Fleisch type, flow range 0–5 l/s). Each MBNW test was concluded on washout of N2 to a level below 2%. MBNW tests were repeated three times with a sufficient time interval permitted between trials (>5 min) to allow for complete re-equilibration of N2 within the lung.

Nitrogen concentration, expired volume, and instantaneous flow data were collected throughout at a sampling rate of 50 Hz via a 16-bit data-acquisition card (PCI-MIO-16XE-50, National Instruments, Austin, TX) with data stored on computer for later analysis. Following data collection, a cross-correlation algorithm was applied to the data to determine, and correct for, the delay in gas sampling and analysis with respect to the flow and volume measurements (typically a 200-ms time correction). MBNW data was then subdivided into individual breaths for further manual and automated analysis.

Manual Analysis

For the 50 subjects in the training group, manual determination of the alveolar plateau from all MBNW traces was performed by two trained respiratory scientists (MAN1 and MAN2) using a previously described technique (31). The MBNW data from the remaining 30 subjects that formed the test group were analyzed by a third independent scientist (MAN3). From the measured slopes of the alveolar plateau for each test, ventilation heterogeneity indexes from both the diffusive-convective (Sacin) and convective regions (Scond) of the lung were calculated for all subjects as previously described (31). Briefly, the alveolar slope for each breath was normalized by dividing the slope by the mean N2 concentration of the whole breath. Normalized slopes (Smil) from the three tests were then plotted against lung TO flow (cumulative expired volume divided by functional residual capacity). Scond was calculated as the gradient of the least squares linear regression plotted between TO 1.5 and TO 6. Smil was determined by subtracting the conductive airways contribution from the Smil of the first breath (31).

Breakpoint Analysis

The automated analysis technique was developed and employed using MATLAB software (The Mathworks, Natick, MA). A method for automated phase III slope detection was developed using segmented linear regression or “breakpoint” analysis similar to that previously described by Beaver et al. (4). The consecutive data for a single breath of the MBNW test was first split into two distinct overlapping sections with the algorithm selecting the split to occur at the N2 midpoint for the breath (end-tidal N2 - inspired N2); a point approximately midway during the phase II portion of the washout. The first data segment therefore contained data that included the entire phase I and approximately half of the phase II, whereas segment two contained the majority of the phase II and all remaining data for that breath including the entire alveolar plateau. Each individual segment was then analyzed by the breakpoint analysis technique separately.

Breakpoint determination of the phase I-II transition point was undertaken initially with the first data segment (of total length n) being further split into two components (n1 and n2). Initially n1 was set to contain the first three points of the breath (n1 = 3) with n2 set to contain all other points in the segment (n2 = n-3). Linear regression...
was performed across both \( n_1 \) and \( n_2 \), and the resulting overall coefficient of determination (\( R^2 \)) was calculated. This regression process was repeated by incrementally increasing \( n_1 \) to 4 and reducing \( n_2 \) to \( n-4 \), where a new \( R^2 \) value was calculated. The process of increasing \( n_1 \) and decreasing \( n_2 \) was continued until \( n_1 = n-3 \) and \( n_2 = 3 \), after which a plot of \( R^2 \) against expired volume was constructed. The breakpoint was determined from this plot (Fig. 2) as the volume at which the \( R^2 \) was a local maxima, and this point used as the phase I-II transition point. This same process was then repeated on the second data segment that contained the phase II-III transition point.

For the purpose of determining the phase III slope, all breaths with expired volume \( >0.9 \) or \( >1.4 \) liters were excluded from analysis, removing the requirement for the automated technique to account for phase IV slope. Because \( N_2 \) data immediately following the phase II-III breakpoint is influenced by a component of late phase II volume (i.e., the breakpoint does not sit precisely on the alveolar plateau portion of the washout curve), the final determination of the alveolar slope was calculated between the phase II and III transition, plus a fixed percentage portion of the \( P_1 \) volume. The alveolar slope was calculated as the gradient of the least squares linear regression between this point (alveolar plateau start volume) and the end of expiration. For comparison with manual analysis, the alveolar slope was calculated at multiple values of added \( P_1 \) volumes (0, 10, 20, 30, 40, 45, 50, 55, 60, 70, 80, 90, 100%) to determine the \( P_1 \) volume that yielded best agreement with manual analysis from the training set.

### Statistical Analysis

Assessment of the accuracy of the automated phase II-III transition points and the calculated \( S_{\text{III}} \) for all recorded breaths at differing values of added \( P_1 \) were analyzed and compared against the mean \( S_{\text{III}} \) results of the two manual interpreters (MAN1 and MAN2) for those 50 subjects in the training group. This analysis allowed for the determination of the proportion of added \( P_1 \) yielding the closest level of agreement between manual and automated analyses. Following the determination of the proportion of added \( P_1 \) yielding the closest level of agreement between manual and automated analyses, the automated analysis method was performed on the additional 30 subjects (test group), and the results were compared with those obtained by a third independent manual interpreter (MAN3). For the test group, manual and automated \( S_{\text{acin}} \) and \( S_{\text{cond}} \) values were calculated separately for each subject, and indexes were compared against each other by the method of Bland and Altman (6). Differences between data sets were assessed using a Wilcoxon signed rank test for paired samples with \( P < 0.05 \) regarded as being significant. Data are presented as median [interquartile range (IQR)] unless otherwise specified.

In our test group we assessed the utility of the manual and automated analysis methods to detect differences in \( S_{\text{acin}} \) and \( S_{\text{cond}} \) between control, asthma, and COPD subject populations for the study.

### Table 1. Demographics of the control, asthma and COPD subject populations for the study

<table>
<thead>
<tr>
<th></th>
<th>Control Group</th>
<th>Asthma Group</th>
<th>COPD Group</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Training Group</strong></td>
<td>( n = 20 )</td>
<td>( n = 20 )</td>
<td>( n = 10 )</td>
</tr>
<tr>
<td>Men/women</td>
<td>11/9</td>
<td>13/7</td>
<td>6/4</td>
</tr>
<tr>
<td>Age, yr</td>
<td>47.3 (15.0)</td>
<td>53 (10.3)*</td>
<td>69 (6.8)*</td>
</tr>
<tr>
<td>Height, cm</td>
<td>168.6 (9.0)</td>
<td>171.3 (10.1)</td>
<td>172.6 (12.1)</td>
</tr>
<tr>
<td>Weight, kg</td>
<td>75.0 (13.7)</td>
<td>82.8 (16.6)</td>
<td>79.1 (24.3)</td>
</tr>
<tr>
<td>FEV1, liters [%predicted]</td>
<td>3.5 (0.9) [104.1%]</td>
<td>2.19 (0.8) [63.6%]</td>
<td>2.3 (0.8) [73.3%]*</td>
</tr>
<tr>
<td>FVC, liters [%predicted]</td>
<td>4.5 (1.2) [105.2%]</td>
<td>3.98 (1.3) [89.5%]</td>
<td>3.9 (1.0) [95.3%]*</td>
</tr>
<tr>
<td>Test Group</td>
<td>( n = 10 )</td>
<td>( n = 10 )</td>
<td>( n = 10 )</td>
</tr>
<tr>
<td>Men/women</td>
<td>5/5</td>
<td>5/5</td>
<td>7/3</td>
</tr>
<tr>
<td>Age, yr</td>
<td>33.4 (4.6)</td>
<td>39.3 (15.6)</td>
<td>66.0 (7.9)*</td>
</tr>
<tr>
<td>Height, cm</td>
<td>171.1 (8.5)</td>
<td>166.7 (8.8)</td>
<td>172.5 (9.0)</td>
</tr>
<tr>
<td>Weight, kg</td>
<td>76.8 (13.4)</td>
<td>76.8 (12.6)</td>
<td>77.4 (16.5)</td>
</tr>
<tr>
<td>FEV1, liters [%predicted]</td>
<td>3.5 (0.6) [93.8%]</td>
<td>2.0 (0.7) [59.1%]*</td>
<td>1.7 (0.6) [54.7%]*</td>
</tr>
<tr>
<td>FVC, liters [%predicted]</td>
<td>4.4 (0.9) [94.9%]</td>
<td>3.0 (0.9) [71.4%]*</td>
<td>3.6 (0.7) [89.2%]*</td>
</tr>
<tr>
<td>FER, %</td>
<td>81.4 (5.9)</td>
<td>58.2 (12.2)*</td>
<td>44.5 (10.8)*</td>
</tr>
</tbody>
</table>

Pulmonary function data are presented as mean (SD) [mean percentage predicted]. *statistically significant differences \( P < 0.05 \) in a demographic when compared with control group.

**J Appl Physiol** • doi:10.1152/japplphysiol.00372.2011 • www.jappl.org
between disease groups using Kruskal-Wallis one-way analysis of variance by ranks with Dunn’s test for post hoc comparisons across subject groups.

RESULTS

Examination of the demographics of our different subject groups showed no significant differences in height or weight. However, we did find a statistically significant difference in age between our control group and our asthma and COPD groups ($P < 0.05$, Table 1).

Training Group

MBNW tests from 50 subjects resulted in 3,397 discrete breaths being obtained (1,340 breaths from the asthma group, 943 from the COPD group, and 1,114 breaths from controls). Two hundred seventy-eight individual breaths (97 asthma, 91 COPD, and 90 control) obtained during MBW testing were excluded as these were found to lie outside of the specified volume inclusion range of 0.9–1.4 liters. Once excluded, neither manual nor automated analysis was performed on those breaths.

Comparison of manual analysis between two interpreters. Between the two separate interpreters’ manual analysis (MAN1 and MAN2), the start of the selected alveolar plateau differed by median [IQR] 1.03 [-5.79–15.93] ml, resulting in a 0.20/liter overall difference in the normalized alveolar slope ($P < 0.01$) as shown in Fig. 3, top. Calculated indexes of ventilation heterogeneity ($S_{ac,in}$ and $S_{cond}$) were significantly different between interpreters, with a 1.73% difference between median $S_{ac,in}$ values ($P < 0.01$) and a 1.03% difference in $S_{cond}$ ($P = 0.04$). The absolute difference between $S_{ac,in}$ values was 0.003 [−0.010–0.001]/liter and for $S_{cond}$ 0.0004 [−0.001–0.002]/liter.

Comparison between manual and automated analysis. Mean $S_{HI}$ data from both manual interpreters (MAN1 and MAN2) were calculated for each breath (MAN) and compared for alveolar slope against our automated (AUTO) method. The $S_{HI}$ resulting from the addition of various proportions of $P_{HI}$
Table 2. Effect of addition of a fixed percentage of phase II volume to the normalized phase III slope as assessed by our automated breakpoint analysis method when compared with manual analysis

<table>
<thead>
<tr>
<th>BP + Fixed % of PII Volume</th>
<th>SnIII Median % error</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%</td>
<td>-21.23</td>
</tr>
<tr>
<td>10%</td>
<td>-16.61</td>
</tr>
<tr>
<td>20%</td>
<td>-11.57</td>
</tr>
<tr>
<td>30%</td>
<td>-7.34</td>
</tr>
<tr>
<td>40%</td>
<td>-3.62</td>
</tr>
<tr>
<td>45%</td>
<td>-2.21</td>
</tr>
<tr>
<td>50%</td>
<td>0.15</td>
</tr>
<tr>
<td>55%</td>
<td>0.54</td>
</tr>
<tr>
<td>60%</td>
<td>1.86</td>
</tr>
<tr>
<td>70%</td>
<td>3.73</td>
</tr>
<tr>
<td>80%</td>
<td>5.54</td>
</tr>
<tr>
<td>90%</td>
<td>8.86</td>
</tr>
<tr>
<td>100%</td>
<td>7.07</td>
</tr>
</tbody>
</table>

PII, phase II; SnIII, phase III slope; BP, breakpoint.

Volume to the breakpoint is detailed in Table 2. When no PII volume was added to the breakpoint, the SnIII was lower when compared with the manually determined slope (median percentage difference 21.23%, P < 0.001). Furthermore, the phase III start point selection was 113.72 [86.98–140.38] ml earlier than the start point selected by MAN analysis. Minimization of the error between MAN and AUTO SnIII occurred with the addition of 50% of the PII volume when there was only a 0.15% difference between the normalized phase III slopes (P < 0.01) as shown in Fig. 3, bottom.

Test Group

MBNW tests in these 30 subjects resulted in 2,192 individual breaths being recorded, of which 80 breaths (13 asthma, 46 COPD, and 21 control) were removed from further analysis as the expired volume of these breaths fell outside of the range of 0.9–1.4 liters and, as with the training group, no further analysis was undertaken on these breaths by either manual or automated analysis.

Comparison between manual analysis and automated breakpoint analysis incorporating 50% of PII. When examining the agreement in our test group between our third manual interpreter (MAN3) and our automated analysis (including the addition of 50% of the PII volume; AUTO) we found that there was a median difference of −33.92 [−93.22–15.16] ml in the selected start point of the alveolar slope, which equated to a −0.01/liter difference in SnIII (P = 0.17). Assessment of the agreement in both Sacin and Scond indexes revealed no significant difference in either median Sacin (0.004 [−0.006–0.014]/liter; P = 0.14) or Scond (0.0002 [−0.003–0.002]/liter, P = 0.59), as shown in Fig. 4, A and B.

The effect of disease status on the determination of ventilation heterogeneity indexes. Overall, there were observable effects of disease status on the determination of alveolar plateau, but these effects did not affect the calculation of Sacin and Scond. Specifically, automated versus manual analysis determined the alveolar plateau start volume 7.96 [−43.60–54.08] ml (P = 0.05) later in expiration in asthma, 76.07 [7.02–146.89] ml earlier in COPD (P < 0.01), and 19.86 [4.87–44.33] ml earlier in controls (P < 0.01). The selection of differing alveolar starting volumes resulted in minimal effects on the alveolar slopes, with a 0.11% decrease in asthma (P = 0.06), a 4.46% increase in COPD (P < 0.01), and a 1.69% increase in controls (P < 0.01). When assessing the ventilation heterogeneity indexes, Sacin and Scond, we found no statistically significant differences between MAN3 and AUTO for any disease group (Fig. 5, A and B).

Our automated method also demonstrated the same utility as the manual method for assessing the effect of disease on ventilation heterogeneity indexes. We found no significant difference between values for Sacin between the control and asthma groups (MAN3: 0.13 [0.08–0.22] vs. 0.12 [0.11–0.16], AUTO: 0.14 [0.08–0.21] vs. 0.13 [0.11–0.15]). We found statistically significant increases in Sacin values in subjects with COPD when compared against the other disease groups (MAN3: 0.63 [0.43–0.85], AUTO: 0.63 [0.40–0.80]; P < 0.05 vs. control and asthma). Both methods also detected significantly greater Scond values in asthma than in control (MAN3: 0.03 [0.01–0.05] vs. 0.01 [−0.00–0.01]; AUTO: 0.03 [0.02–0.05] vs. 0.01 [−0.01–0.01]; P < 0.05) and in COPD (MAN3: 0.03 [0.03–0.06], AUTO: 0.04 [0.02–0.06]; P < 0.05).

DISCUSSION

In the current study we developed a technique to automatically detect the slope of the phase III portion of the washout curve, enabling the determination of the alveolar plateau as outlined by Verbanck et al. (31). We demonstrated no clinically significant differences between our automated algorithm and manual analysis in either the normalized phase III slope (SnIII), Sacin or Scond. Our algorithm will likely improve the inter- and intra-interpreter variability of the analysis of inert gas washout tests, leading to an increase in the precision of the indexes generated from the washout test. Application of this algorithm would greatly reduce the time required to analyze the MBNW tests, which is typically in excess of 15 min. By overcoming difficulties in manual analysis, our computational method allows rapid feedback to the operator, overcoming one of the obstacles to the routine use of inert gas washout techniques in clinical practice. Our choice of ventilation heterogeneity indexes (31) to assess agreement between methods of SnIII calculation was driven by current clinical practice (7, 10); however, other indexes such as those proposed originally by Crawford et al. (8) or by Emery et al. (11) could equally well be used.

We chose to assess the accuracy of our automated analysis technique across a range of subjects: control subjects, who have a relatively flat phase III slope (10, 31), and in subjects with asthma and COPD, where changes in respiratory dead space and the slope of phase III is steeper and alterations in ventilation heterogeneity are expected (10, 27, 29, 31). Within these subject groups we did not find any significant alteration in the ability of our automated analysis technique to assess ventilation heterogeneity indexes that agreed with those generated by manual analysis.

We first examined the differences in SnIII and the resulting indexes of ventilation heterogeneity (Sacin and Scond) determined manually by two separate interpreters in our training group of 50 subjects. We found a small but statistically significant difference in SnIII values between the two interpreters of 0.20/liter (P < 0.001). The resulting Sacin and Scond indexes...
showed a small but statistically significant difference between the two interpreters; however, these differences fall within the reported day to day variability of MBNW indexes for control and asthmatic subjects (5) and those with COPD (20). The observed variability in S\text{III},S_{acin}, and S\text{cond} is one reason that an automated analysis is preferable over manual methods because of its superior repeatability and reproducibility.

One potential concern with the use of an automated analysis method is that the phase III transition from phase II is not instantaneous; that is, the initial portion of the phase III may be contaminated by late phase II volume. Inappropriately including a portion of phase II within the phase III will result in an increase in the measured alveolar slope, resulting in large changes in S\text{acin} (although the measurement of S\text{cond} should be largely unaffected). To overcome this limitation, we added a percentage of P\text{II} volume to the phase II-III transition point (phase III breakpoint) and assessed the effect on the ventilation heterogeneity indexes S\text{acin} and S\text{cond}. The use of 50% of the phase II volume, as opposed to an arbitrary fixed value, was chosen because it effectively corrects ventilation heterogeneity indexes for differences in dead space volume between subjects (11).

Inasmuch as there is no gold standard for determination of the alveolar slope, previous studies (10, 22, 25, 31) have used a variety of different analysis techniques when deriving ventilation heterogeneity indexes. We found that the addition of 50% of phase II volume gave the best overall agreement between the alveolar slopes determined by the automated algorithm and the manual method described by Verbanck et al. (31) (median percentage difference 0.15%). Further examination of these S\text{III} values in our training group allowed calculation of the ventilation heterogeneity indexes S\text{acin} and S\text{cond}. We found that, with addition of 50% of the phase II volume, there was a 2.84% median difference in S\text{acin} ($P = 0.91$) and 4.88% difference in S\text{cond} ($P < 0.01$) between the automated and manual analysis techniques. Although addition of greater percentages of P\text{II} volume may have given better agreement for both indexes, we felt that the most robust design of an automated method for S\text{III} calculation would maximize the volume of phase III used in the calculation of alveolar slope. We were willing to accept a small, clinically insignificant difference in both S\text{acin} and S\text{cond} to match the S\text{III} values between methods. Finally, when validating our automated technique using a test group of 30 subjects, we found no differences in either S\text{acin}
(\(P = 0.14\)) or \(S_{\text{cond}} (P = 0.59)\) and a median percentage difference of \(-0.01/\text{liter}\) in \(S_{\text{III}} (P = 0.17)\).

There have been other methods to detect the alveolar plateau published in the literature (11, 31) that are worthy of comment. Table 3 provides a comparison of the results of our independent manual analysis (\(\text{MAN}_3\)) against our automated technique (AUTO) and three other techniques previously published in the literature: 1) taking an early (0.5–0.75 \(\times V_T\)) or 2) late (0.75–0.95 \(\times V_T\)) fraction of the alveolar plateau as described by Emery (11) and 3) using a fixed volume point to define the start of the alveolar plateau (0.65 liters expired) as described originally by Verbanck et al. (31). Unlike the comparison of our automated algorithm against the manual technique, there were statistical differences between the selected volume points for early and late portions of phase III and 0.65 liter when compared against the manual technique. We further examined the change in this volume across the entire washout and also, using the \(S_{\text{III}}\) values from each technique, calculated \(S_{\text{ac}}\) and \(S_{\text{cond}}\). We found that there was no significant change in the selected alveolar plateau start volume during the washout using any of the analysis techniques (\(P = 0.68\)) and the calculation of the indexes \(S_{\text{ac}} (P = 0.10)\) or \(S_{\text{cond}} (P = 0.38)\) were not affected. However, we did find that using the early part of the alveolar slope appeared to result in overestimation of \(S_{\text{ac}}\), whereas the late component and the fixed 0.65 liter start volume resulted in underestimation of \(S_{\text{ac}}\) when compared against \(\text{MAN}_3\) and AUTO. These results were not statistically different. Importantly, there was a significantly larger interquartile range in \(S_{\text{ac}}\) using the early fraction of alveolar plateau, suggesting that this calculation is more susceptible to noise likely introduced as a result of the small sampling volume used to calculate the alveolar plateau.

Despite small differences in \(S_{\text{ac}}\) and \(S_{\text{cond}}\) determined from automatic and manual \(S_{\text{III}}\) analysis in the training group, the changes observed between control, asthma, and COPD groups in the current study are far greater. We found a 132% greater median \(S_{\text{cond}} (0.006/\text{liter absolute median change})\) in subjects with COPD and a 403% increase in \(S_{\text{ac}} (0.434/\text{liter absolute difference})\) when compared against \(\text{MAN}_3\) and AUTO. These results were not statistically different. Importantly, there was a significantly larger interquartile range in \(S_{\text{ac}}\) using the early fraction of alveolar plateau, suggesting that this calculation is more susceptible to noise likely introduced as a result of the small sampling volume used to calculate the alveolar plateau. In this study the automated analysis technique was tested on data obtained from a control group and two groups of obstructed patients (asthma and COPD), and thus a range of diseases were assessed. The range of \(S_{\text{III}}\) values recorded from our different disease groups examined in this study cover

---

**Table 3. Comparison of the alveolar plateau start volumes, the change in volume over the course of the MBNW test (\(\Delta\)alveolar plateau start volume), \(S_{\text{ac}}\) and \(S_{\text{cond}}\) determined by 4 different analysis techniques using our test group of 30 subjects**

<table>
<thead>
<tr>
<th></th>
<th>(\text{MAN}_3)</th>
<th>AUTO</th>
<th>0.5–0.75 (\times V_T)</th>
<th>0.75–0.95 (\times V_T)</th>
<th>0.65 liters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alveolar plateau start volume, ml</td>
<td>492.55 [330.68–594.80]</td>
<td>502.26 [424.26–590.84]</td>
<td>540.86* [509.95–554.57]</td>
<td>811.29* [764.93–831.85]</td>
<td>650.00* [650.00–650.00]</td>
</tr>
<tr>
<td>(\Delta)alveolar plateau start volume, ml</td>
<td>11.78 3.71</td>
<td>0.173 [−25.47–38.73]</td>
<td>[−17.95–21.70]</td>
<td>[−17.95–21.70]</td>
<td>0.000 0.014</td>
</tr>
<tr>
<td>(S_{\text{ac}})/liter</td>
<td>0.176 [0.108–0.362]</td>
<td>0.241 [0.164–0.784]</td>
<td>0.141 [0.080–0.305]</td>
<td>0.013 [0.009–0.043]</td>
<td>0.021 [0.011–0.049]</td>
</tr>
<tr>
<td>(S_{\text{cond}})/liter</td>
<td>0.017 [0.009–0.043]</td>
<td>0.019 [0.011–0.049]</td>
<td>0.013 [0.000–0.029]</td>
<td>0.026 [0.003–0.056]</td>
<td>0.021 [0.004–0.056]</td>
</tr>
</tbody>
</table>

\(\text{MAN}_3\): traditional manual analysis where the alveolar plateau start volume is selected by a human interpreter; AUTO, results from our automated analysis algorithm; 0.5–0.75 \(\times V_T\) and 0.75–0.95 \(\times V_T\), effect of a previously described technique using early and late fixed fractions of exhaled tidal volume (11); and 0.65 liters, results of using a fixed volume to define the start of phase III (31). *Statistical significance between groups when compared with the \(\text{MAN}_3\) (\(P < 0.05\)).
much of the expected range from other patient populations. Previous studies have shown that the values of $S_{III}$ in asthma and COPD subjects are elevated above the levels of controls with the entire range of expected $S_{III}$ values being shown to fall between 0.1 and 1.7/liter (29–31). We found that the $S_{III}$ values determined by our automated technique covered the range of 0.12–1.57/liter in all subjects from our test group ($n = 30$) with no change in mean bias across the entire range of recorded $S_{III}$ values (Fig. 3). The strong agreement between manual and automated analysis across all disease groups shows that our technique is a robust and reliable method for the analysis of MBNW tests.

There is a potential for a bias to occur with the automated technique given that dead space can alter with disease, an effect that would act to systematically alter the phase II volume. Studying two disease groups (asthma and COPD) allowed us to assess the effect of any alteration in respiratory dead space on our automatic detection of $S_{III}$. However no such bias was seen (Fig. 6B) and, therefore, we conclude that any potential differences attributable to altered phase II volume with disease are similar between the manual and automated methods.

Although it is still not known if the accuracy of the automated analysis will be affected following an induced change in ventilation heterogeneity (such as that seen post bronchial provocation challenge) the agreement between manual and automated analysis across all three subject groups suggests that there will not be any significant changes in agreement. We further studied this by measuring the difference in breakpoint volume between the first and last breath in our test group (Fig. 6A). The overall median [IQR] difference was small (3.36 [−7.09–15.38] ml) irrespective of disease group. This finding was consistent with the constant Fowler dead space (12) calculated throughout the maneuver as seen in Fig. 6B. This has also been shown by others (8). As we add only 50% of the total measured dead space volume to our automated phase III slope formula, the small changes in measured dead space volume observed during the MBNW test will only have a very small effect on the selected start volume of the alveolar plateau. The resulting effect on the $S_{III}$ will be insignificant. Under circumstances where it is imperative that similar portions of specific ventilation are assessed, a fixed volume analysis technique such as that described previously by Emery et al. (11) may indeed be preferable. However, it is likely that the noise in the measurement of $S_{III}$ by taking such an approach will increase because of the reduction in sample volume the phase III slope is calculated across (Table 3). Finally we believe the repeatability and reproducibility of an automated analysis technique is preferable over manual methods, as small shifts in ventilation heterogeneity will not be confounded by variability in analysis.

Although we chose to limit testing of our automated algorithm to the more time-consuming analysis of MBNW, our method could also be applied to the analysis of the single breath washout test where determination of the alveolar slope is less challenging because of the increased volume of alveolar gas expired during the test. We would further expect that the analysis techniques presented in this manuscript are equally applicable to the analysis of MBNW tests where the tidal volume of a subject lies outside of the suggested volume range (8, 31). Furthermore, the method described can equally be applied to determination of any phase IV breakpoints recorded at expired volumes above the subject’s closing capacity as is commonly seen during single breath washout tests. Finally, our method would also be appropriate for the analysis of other inert gas species used for gas washout testing. Gases such as sulphap-hexafluoride ($SF_{6}$) (2, 3, 18, 22–24) and helium (He) (2, 16) have been used successfully by others to perform MBNW tests with the only notable difference in analysis being that the phase IV slope for $SF_{6}$ would be expected to deviate in a negative direction (17, 23, 25, 27).

In conclusion, we developed an alternate analysis method to automatically detect phase transitions for the determination of alveolar slope during the inert gas washout test. This method eliminates any operator selection bias that may occur during manual analysis of the MBNW test without affecting the ability to observe differences in ventilation heterogeneity indexes in control and in subjects with either asthma or COPD. With the removal of human interference from the analysis of inert gas washout testing, reliability in the derivation and calculation of the alveolar slope and resulting indexes of ventilation heterogeneity will be improved.

**GRANTS**

Financial support was provided to V. J. Kelly by research grants from the National Health and Medical Research Council of Australia (Grant 521808), the Cooperative Research Centre for Asthma and Airways (CRCAA), and the American Australian Association. S. A. Sands received support from an American Heart Association Post-Doctoral Fellowship. B. R. Thompson also received financial support through the National Health and Medical Research Council of Australia (Grants 486101 and 491103).

**DISCLOSURES**

No conflicts of interest, financial or otherwise, are declared by the authors.

**AUTHOR CONTRIBUTIONS**


**REFERENCES**


**AUTHOR CONTRIBUTIONS**


**REFERENCES**