ALIGNING LUNG FUNCTION EQUIPMENT AND REFERENCE VALUES IN ADULTS.

Sylvia Verbanck 1, Daniel Schuermans 1, Bruce Thompson 2,3 and Eef Vanderhelst 1.

1 Respiratory Division, University Hospital UZ Brussel, Vrije Universiteit Brussel, 1090 Brussels, Belgium.

2 Allergy, Immunology, and Respiratory Medicine, Alfred Hospital, Melbourne, Australia.

3 Central clinical school, Monash University, Melbourne Australia.

Running head: lung function abnormality following new reference values or new equipment.

Corresponding author: Sylvia Verbanck

Respiratory Division, University Hospital UZ Brussel

Laarbeeklaan 101

1090 Brussels, Belgium

Tel: +32 2 477 6352

Email: sylvia.verbanck@uzbrussel.be

Keywords: lung function indices, reference values, limits of normal.
ABSTRACT

Background: When introducing new equipment or reference equations in the lung function laboratory, systematic z-score deviations could arise due to local population or equipment differences.

Objectives: To propose a workable method to align reference equations with lung function equipment.

Method: Using two cases of equipment transition in our laboratory as a test case, we first performed lung function after the transition on a control group of 40 normal young adults (20M/20F; 20-30yrs). For those indices with an average z-score in excess of ±0.5, adapted reference values were obtained by an offset or scaling factor on the M-coefficient in the so-called LMS method recommended by GLI, and z-scores were computed again.

Results: Following a transition involving instrumental dead space reduction, lung clearance index was predictably reduced resulting in a mean (+SD) z-score of -1.9 (+1.1) in the control group; by adapting reference values with an offset on M, z-score became -0.1 (+1.1). Applying the same method to a transition of standard lung function equipment, z-scores became centered around zero in the control group, but also became properly aligned in a test group of 81 other subjects spanning a wider age range (20-80yrs).

Conclusions: We proposed and verified a method to align local equipment with reference values obtained elsewhere, or following a local change in equipment. The key is to measure a relatively small young adult group, identifying those lung function indices that need an adaptation based on z-score, to then obtain laboratory-specific reference values that can be applied over the entire age range.
INTRODUCTION

Following reference equations for spirometry [1] and carbon monoxide transfer factor [2], the global lung function initiative (GLI) now proceeds with static lung volumes to replace previous normative values [3]. When locally incorporating reference equations obtained elsewhere, lung function indices may show a population- or equipment-specific bias of unknown origin, partly dependent on hardware, software and how exactly guidelines are implemented (e.g., how lung divisions are derived [4]). Implementing the GLI equations for spirometry has led to some researchers being concerned with the fact that these are not representative of their local population [5] although local equipment effects cannot be excluded. More complex lung function tests such as multiple breath washout (MBW) [6] may present an equipment-specific bias of known origin (e.g., effect of instrumental dead space). In either case, a direct and unverified implementation into existing equipment of reference equations obtained elsewhere or on other equipment, could lead to a local laboratory not obtaining zero average z-scores on selected indices, if a normal control group were to be measured. An equipment- or population-dependent average z-score deviation in excess of $+0.5$ is considered physiologically significant [2].

Recently, we obtained reference equations for all lung function parameters including lung volumes and MBW indices [7] on a group of 252 normal subjects aged 20-80 years. This was done using LMS statistical handling characteristic of GLI and measuring at least 20 male and 20 female subjects per decade as per indication of sample sizes required for establishing reference values locally [1,8]. Subsequently, two cases of equipment transition in our laboratory have led to a bias (z-score deviation in excess of $+0.5$), one of known and one of unknown origin, requiring an adaptation of our reference values. The aim of the present study is to propose a simple method to verify, and if needed, adapt reference values to an adult lung function laboratory, without having to measure 250 subjects between the age of 20 and 80 years each time a new set of (global) reference equations is being implemented, or new equipment is being installed, or software is being upgraded. We first
explored this method for a case of equipment bias of known origin, i.e., following a reduction in instrumental dead space of MBW apparatus, which has a predictable effect on MBW indices [9]. We then also applied the method to a bias of unknown origin, after an in-depth hardware and software review of a systematic change in static lung volume measurements following a commercial equipment transition.

The proposed method consists of measuring a control group of healthy young adults (20M/20F; 20-30 years) to verify and adapt reference values after the equipment transition. We hypothesized that in the adult range, a population- or equipment-dependent bias will leave the overall dependence of the reference equations on age largely unaltered. If true, the reference value adaptations based on the young control group, can be readily applied to a wider adult age range. We could explicitly test the latter hypothesis in subjects having visited our laboratory before and after the standard lung function equipment transition, identifying a subgroup of subjects who were not expected to show marked differences in their lung function post- versus pre-equipment transition.

MATERIALS AND METHODS

Since our local reference equations produce very similar z-scores compared to GLI (at least for those indices that are also available with GLI), we used our own reference equations for all indices under study here [7]. These had been obtained with our bag-in-box system and N₂ analyzer (with 50 ml instrumental dead space) and with state-of-the-art commercial lung function equipment (VmaxEncore20c, 22, 22d; Cardinal Health, USA) [7]. Our reference values had been based on measurements in approximately 20 males and 20 females per decade in the age range 20 to 80, guided by prior recommendations about sample size [8] and using the LMS method to derive the predicted value and z-score [1].
In the present work we adhered to the same sample size, but now limited to the first decade of the adult age range (20-30 years), to obtain a “control group” of 40 normal subjects (Local ethics committee BUN143201836071); measurements were performed between June and October 2018. We measured MBW using our bag-in-box system and N2 analyzer but with instrumental dead space reduced to 15ml. We also measured spirometry [10], gas transfer factor [11] and static lung volumes [4] using a different commercial lung function system (MasterscreenPFT, Sentriesuite2.19; Vyaire, USA), and computed z-scores for this group based on our reference equations. Firstly, the indices were identified for which average z-score was in excess of 0.5 in absolute value. Secondly, all related indices were selected, e.g., if FEV1 was identified, FVC and FEV1/FVC were selected as well. Thirdly, we determined for each index whether the difference between measured and predicted value was correlated with the magnitude of the predicted value. If this correlation reached significance (P<0.05 in Pearson correlation), the adaptation of the reference value for that index consisted of a scaling factor on M. If not, an offset on M was considered instead. In either case (offset or scaling on the M coefficient in the LMS method), L and S coefficients remained unaffected, and corresponding upper and lower limits of normal (LLN, ULN) were computed as M.(1 + 1.645.S.L)1/2.

Finally, we selected a “test group” of subjects having visited our lung function lab within the year before and after equipment transition, and who had pre-transition z-scores for lung function within 1.645 limits, to verify the impact of adapting the reference value post-transition.

RESULTS

Table 1 shows MBW and lung function values obtained in the control group after equipment transition, using our original reference equations which were available for all lung function indices under study. For FVC, FEV1, FEV1/FVC, TLco and Kco, z-scores in Table 1 were very similar to those
obtained by applying available GLI equations for Caucasian subjects, respectively 0.5±0.8(SD), 0.3±0.9(SD), 0.3±0.8(SD), 0.1±1.0(SD). For internal consistency between all indices under study, and for the sake of illustration, we use our original reference equations for all indices from this point forward.

For MBW indices, lung clearance index (LCI) predictably showed the most marked deviation with mean(±SD) z-score= -1.9 (±1.1). The absolute changes for all three MBW indices (±0.5 for LCI, +0.005 L⁻¹ for Scond, -0.01 L⁻¹ for Sacin) were very close to those previously obtained by switching instrumental dead space between 15 and 50ml in a given subject [9]. Figure 1A is a graphical representation of LCI predicted and LLN and ULN corresponding to the original (thin lines) reference values [7] and the adapted ones (thick lines) obtained by applying an offset of -0.5 units on the M coefficient for LCI (Table 1); superimposed are the individual data points from the control group on which the offset is based (circles).

Of the other lung function indices obtained in the control group, only TLco and Kco did not require adaptation of reference values (i.e., z-scores within 0.5 limits). The deviation of measured versus predicted value was not correlated with predicted value for : LCI, Scond (and related index Sacin), RV (and related indices TLC and RV/TLC) (P>0.1 for all) and therefore an offset was applied for these indices. In fact, the absolute RV increase was not significantly different from the absolute TLC increase (P>0.1), averaging +0.33 L; corresponding offset on RV/TLC was +0.04. By contrast, FEV₁ and FVC deviations in the control group correlated with their respective predicted value (FEV₁: r=0.41, P=0.009 and FVC: r=0.45, P=0.004; Pearson correlation) indicating that a scaling factor was more appropriate. When adapting the reference values using the offset or scaling in Table 1 (italics), and applying these to the control group, resulting z-scores were centered on zero (by design) and standard deviations remained unaltered. In the Online Supplement, the impact is shown of considering instead of a scaling factor on the M coefficient for FEV₁ of 1.080 (Figure1B), an offset of 0.34 L. This shows that the effect of not taking into account the contribution from the magnitude of the FEV₁ value on the adaptation of M, by using an offset instead of a scaling, is in fact very small.
To then test the impact of the adapted reference values on a group of all aged subjects, we selected the 518 subjects having undergone a full lung function within the year before as well as within the year after equipment transition (29% COPD, 26% asthma, 16% cystic fibrosis, 12% interstitial lung disease, 6% thorax pathology, 6% other pathologies and 5% normal). When identifying those subjects with pre-transition z-scores within 1.645 limits, 81 subjects could be retained as a “test group” (30% normal, 26% asthma, 19% cystic fibrosis, 16% interstitial lung disease, 10% other pathologies) with a median time lapse between pre- and post-equipment transition measurement of 10 months. For the test group, Table 2 shows how the unadapted reference values would increase average RV from 101% pred (z-score=0.0) to 124% pred (z-score=1.0), merely due to equipment differences. Applying the adapted reference values (based on the control group aged 20-30 years; 50% male) to the test group ranging 20-80 years (49% male), brings z-scores for RV, TLC and RV/TLC close to pre-transition values.

Figure 2A is a graphical representation of RV/TLC predicted and LLN and ULN corresponding to the unadapted reference values [7] superimposed with data from the test group subjects (Table 2) obtained pre- (open triangles) and post- (closed triangles) equipment transition. For clarity, only females are represented here (n=43 out of 81) and corresponding male data points can be found in the Online Supplement. Figure 2B illustrates the effect of applying an offset of +0.04 units on the M coefficient for RV/TLC (Table 1) on the predicted and ULN and LLN lines in the age range 20-80 years, superimposed with post-equipment transition data from the control group (open circles) and test group (closed triangles). It can be seen that with the adapted reference values, the test group data nicely follow the predicted, LLN and ULN throughout the age range, whereas they would systematically fall between predicted and ULN if the adaptation had not been done. To better appreciate the difference between unadapted and adapted reference values in both panels of Figure 2, we also show here the predicted and ULN for RV/TLC for females according to the 1995 ATS/ERS reference values [3].
In the present study, we proposed a simple method to adapt reference values in a typical situation that can occur in a clinical lung function laboratory, when an equipment change results in a bias of known or unknown origin. However, the method described here can also be used when equipment hardware and software is unaltered, but new reference equations such as those from GLI are to be incorporated, also taking into account local study populations. The method consists of measuring a young adult control group to determine which are the lung function indices that could benefit from an adaptation of their reference values. The predicted value (M in the LMS method) is then simply scaled or offset to obtain a laboratory-specific reference value for that index. The adaptation in terms of an offset or scaling on M does carry the implicit assumption that this can be applied to an age range above 30 years. For the standard lung function indices, the stability of z-scores across the equipment transition in our test group, confirms that the extrapolation from the young control group to a wider age range is likely to be valid (Table 2). The data in Table 2 and in Figure 2 also show the impact of not aligning reference equations with equipment at all: patients without air trapping in terms of RV/TLC pre-transition would appear to have considerable air trapping only 10 months later.

The proposed method is not excessively time-consuming in terms of subject recruitment and measurement, and the criterion to decide whether an index needs adaptation is flexible. The scaling and offset on M can be readily altered or repeated when new equipment or new reference equations are to be implemented. For instance, in the case of FEV\textsubscript{1} and FVC (Table 1), one could argue that the z-score deviations being in the vicinity of 0.5, an adaptation may not be strictly necessary, particularly because z-score deviation for FEV\textsubscript{1}/FVC was only 0.3. In a cross sectional setting it may be pragmatic to only adapt reference values when z-score deviations are worse than this. However, in a case where historical data are considered, for instance to establish individual trends in lung
function decline, a more rigorous alignment of reference equations and equipment throughout each portion of the observation period may be crucial. In the case of FEV$_1$, the impact of applying either an offset or a scaling factor was small (Online Supplement). In the case of the three MBW indices of ventilation heterogeneity (Table 1), the effect of instrumental dead space was expected to be independent of the degree of ventilation heterogeneity itself [12]. On the one hand, this implies an offset rather than a scaling factor. On the other hand, this implies that the offset due to dead space would be valid across the entire adult age range, given that the effect of age is an increased degree of ventilation heterogeneity itself [7,9].

With the obvious benefits brought to the lung function community by GLI, which steadily replaces outdated reference values, the more complex measurement such as lung volume measurement also comes with its own challenges. If we accept the premise that the most recent equipment is likely to be the most accurate one, then our data after equipment transition (closed triangles in Figure 2B) suggest that the predicted RV/TLC values from upcoming GLI will actually be in the vicinity of those offered by current 1995 ATS/ERS reference values [3], yet that predicted and ULN will no longer be linear as a function of age. If however, a different instrument is used (open triangles in Figure 2A), RV/TLC may well be offset and an adaptation of reference values may be warranted.

In conclusion, we proposed a relatively simple method to align reference equations with specific equipment and a local population using on a flexible z-score based criterion. It can be repeated when guidelines change or when hardware or software modifications are carried out. The alternative is to measure patients as is, with equipment and reference equations provided by the manufacturer, without any prior testing of normal subjects at all. In that case, the resulting z-scores may not properly reflect true abnormality and lead to misdiagnosis.
STATEMENT OF ETHICS
Subjects have given their written informed consent. The study protocol has been approved by the research institute’s committee on human research.

DISCLOSURE STATEMENT
SV has no conflicts of interest; DS has no conflicts of interest; BT has no conflicts of interest; EV has no conflicts of interest.

FUNDING SOURCES
N/A.

AUTHOR CONTRIBUTIONS
SV designed and conducted the study; DS performed the lung function measurements; BT and EV reviewed the manuscript.
REFERENCES


Figure 1: 
Panel A: Predicted values (solid line) and lower and upper limits of normal (dashed lines) for the lung clearance index (LCI), using original (thin lines) and adapted (thick lines) reference values. The 40 data points (open circles: male; grey circles: female) are from the control group on which the adaptation (-0.5 unit offset in Table 1) was based.

Panel B: Predicted values (solid line) and lower and upper limits of normal (dashed lines) for FEV₁ in male subjects using original (thin lines) and adapted (thick lines) reference values. The 20 data points (open circles: male) are from the control group on which the 1.08 scaling in Table 1 was based. The predicted values and limits of normal by applying an offset of +0.34 instead of a 1.08 scaling can be found in the Online Supplement.

Figure 2: Test group data for residual volume-to-total lung capacity ratio (RV/TLC). Female data only; corresponding male data are in the Online Supplement. Identical in both panels are the RV/TLC predicted (thin solid line) and ULN (thin dotted line) values according to ATS/ERS1995 (ref3).

Panel A: Test group data, pre- (grey triangles) and post- (black triangles) equipment transition and predicted values (thick solid line) and LLN,ULN (thick dashed lines) using original reference values.

Panel B: Test group data post-equipment transition (black triangles) and predicted values (thick solid line) and LLN,ULN (thick dashed lines) using adapted reference values. Also shown are the 20 data points (grey circles: female) from the control group on which the adaptation (+0.04 offset in Table 1) was based.
Figure 2

A

B

\( \frac{R}{V_{T/LC}} \)

age (years)

\( \frac{R}{V_{T/LC}} \)

age (years)
TABLE 1: Lung function and MBW indices measured after equipment transition in the control group (n=40).

<table>
<thead>
<tr>
<th>MBW indices</th>
<th>Original Reference Values</th>
<th>offset or scale (*)</th>
<th>adapted Reference Values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>%pred mean (SD)</td>
<td>z-score mean (SD)</td>
<td>%pred mean (SD)</td>
</tr>
<tr>
<td>LCI</td>
<td>92 (4)</td>
<td>-1.9 (1.1)</td>
<td>offset: -0.5</td>
</tr>
<tr>
<td>S_{cond}</td>
<td>119 (39)</td>
<td>0.8 (1.7)</td>
<td>offset: +0.005 L^{-1}</td>
</tr>
<tr>
<td>S_{acin}</td>
<td>86 (22)</td>
<td>-0.6 (0.8)</td>
<td>offset: -0.010 L^{-1}</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Standard Lung Function Indices</th>
<th>%pred mean (SD)</th>
<th>z-score mean (SD)</th>
<th>%pred mean (SD)</th>
<th>z-score mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FEV1</td>
<td>108 (11)</td>
<td>0.6 (0.9)</td>
<td>scale: 1.080</td>
<td>102 (11)</td>
</tr>
<tr>
<td>FVC</td>
<td>106 (11)</td>
<td>0.4 (0.9)</td>
<td>scale: 1.056</td>
<td>98 (11)</td>
</tr>
<tr>
<td>FEV1/FVC</td>
<td>0.3 (1.2)</td>
<td>scale: 1.023</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TL_{CO}</td>
<td>99 (10)</td>
<td>-0.1 (0.9)</td>
<td>0</td>
<td>99 (10)</td>
</tr>
<tr>
<td>K_{CO}</td>
<td>98 (12)</td>
<td>-0.2 (1.1)</td>
<td>0</td>
<td>98 (12)</td>
</tr>
<tr>
<td>RV</td>
<td>125 (21)</td>
<td>0.9 (0.8)</td>
<td>offset: +0.33 L</td>
<td>99 (21)</td>
</tr>
<tr>
<td>TLC</td>
<td>105 (8)</td>
<td>0.4 (0.7)</td>
<td>offset: +0.33 L</td>
<td>100 (8)</td>
</tr>
<tr>
<td>RV/TLC</td>
<td>121 (22)</td>
<td>0.7 (0.8)</td>
<td>offset: +0.04</td>
<td>101 (22)</td>
</tr>
</tbody>
</table>

LCI: lung clearance index; S_{cond}, S_{acin}: indices of conductive and acinar airways ventilation heterogeneity; FEV1: forced expired volume in 1 s; FVC: forced vital capacity; TL_{CO}: carbon monoxide diffusing capacity; K_{CO}: TL_{CO} divided by alveolar volume; RV: residual volume; TLC: total lung capacity.

(*) offset or scaling is applied to coefficient M of the LMS method (Ref7).
TABLE 2: Lung function measured before and after equipment transition in the test group (n=81).

<table>
<thead>
<tr>
<th></th>
<th>PRE-TRANSITION</th>
<th>POST-TRANSITION</th>
<th>adapted Reference Values (*)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>original Reference Values</td>
<td>original Reference Values</td>
<td></td>
</tr>
<tr>
<td></td>
<td>%pred mean (SD)</td>
<td>z-score mean (SD)</td>
<td>%pred mean (SD)</td>
</tr>
<tr>
<td>FEV₁</td>
<td>100 (11)</td>
<td>-0.1 (0.9)</td>
<td>103 (13)</td>
</tr>
<tr>
<td>FVC</td>
<td>98 (10)</td>
<td>-0.2 (0.8)</td>
<td>100 (12)</td>
</tr>
<tr>
<td>FEV₁/FVC</td>
<td>0.1 (1.0)</td>
<td></td>
<td>0.4 (1.1)</td>
</tr>
<tr>
<td>TLCO</td>
<td>104 (13)</td>
<td>0.2 (1.0)</td>
<td>101 (17)</td>
</tr>
<tr>
<td>KCO</td>
<td>104 (13)</td>
<td>0.3 (1.0)</td>
<td>96 (16)</td>
</tr>
<tr>
<td>RV</td>
<td>101 (17)</td>
<td>0.0 (0.8)</td>
<td>124 (19)</td>
</tr>
<tr>
<td>TLC</td>
<td>99 (9)</td>
<td>-0.1 (0.8)</td>
<td>104 (10)</td>
</tr>
<tr>
<td>RV/TLC</td>
<td>103 (13)</td>
<td>0.1 (0.7)</td>
<td>121 (15)</td>
</tr>
</tbody>
</table>

FEV₁: forced expired volume in 1 s; FVC: forced vital capacity; TLco: carbon monoxide diffusing capacity; Kco: TLco divided by alveolar volume, RV: residual volume, TLC: total lung capacity. (*) adapted Ref Eq, based on the offsets and scalings in Table 1.
ALIGNING LUNG FUNCTION EQUIPMENT AND REFERENCE VALUES IN ADULTS.

Sylvia Verbanck, Daniel Schuermans, Bruce Thompson and Eef Vanderhelst.

Online Data Supplement
FIGURE LEGENDS

**Figure OLS 1:**
Predicted values (solid line) and lower and upper limits of normal (dashed lines) for FEV$_1$ in male subjects using original (thin lines) and adapted (thick lines) reference values; adaptation is done by a scaling factor of 1.08 (Table 1). The dotted lines are the adapted predicted values and limits of normal by applying an offset of +0.34 instead of a 1.08 scaling.

**Figure OLS 2:** Test group data for residual volume-to-total lung capacity ratio (RV/TLC). Male data only; corresponding female data are Figure 2). Identical in both panels are the RV/TLC predicted (thin solid line) and ULN (thin dotted line) values according to ATS/ERS1995 (ref3). **Panel A:** Test group data, pre- (open triangles) and post- (black triangles) equipment transition and predicted values (thick solid line) and LLN,ULN (thick dashed lines) using original reference values. **Panel B:** Test group data post-equipment transition (black triangles) and predicted values (thick solid line) and LLN,ULN (thick dashed lines) using adapted reference values. Also shown are the 20 data points (open circles: male) from the control group on which the adaptation (+0.04 offset in Table 1) was based.
Figure OLS2