Exercise peripheral oxygen saturation (SpO2) accurately reflects arterial oxygen saturation (SaO2) and predicts mortality in systemic sclerosis

J J Swigris,1 X Zhou,1 F S Wamboldt,2 R du Bois,1 R Keith,1 A Fischer,1 G P Cosgrove,1 S K Frankel,1 D Curran-Everett,3 K K Brown1

ABSTRACT
Background: Measures of oxygenation have not been assessed for prognostic significance in systemic sclerosis-related interstitial lung disease (SSc-ILD).

Methods: 83 subjects with SSc-ILD performed a maximal cardiopulmonary exercise test with an arterial line. The agreement between peripheral oxygen saturation (SpO2) and arterial oxygen saturation (SaO2) was examined and survival differences between subgroups of subjects stratified on SpO2 were analysed. Cox proportional hazards analyses were used to examine the prognostic capabilities of SpO2.

Results: At maximal exercise the mean (SD) difference between SpO2 and SaO2 was 2.98 (2.98) and only 15 subjects had a difference of >4 points. The survival of subjects with SSc-ILD whose maximum exercise SpO2 (SpO2max) fell below 89% or whose SpO2max fell >4 points from baseline was worse than subjects in comparator groups (log rank p = 0.01 and 0.01, respectively). The hazard of death during the median 7.1 years of follow-up was 2.4 times greater for subjects whose SpO2max fell below 89% (hazard ratio 2.4, 95% CI 1.1 to 4.9, p = 0.02) or whose SpO2max fell >4 points from baseline (hazard ratio 2.4, 95% CI 1.1 to 5.0, p = 0.02).

Conclusion: In patients with SSc-ILD, SpO2 is an adequate reflection of SaO2 and radial arterial lines need not be inserted during cardiopulmonary exercise tests in these patients. Given the ease of measurement and its prognostic value, SpO2 should be considered as a meaningful clinical and research outcome in patients with SSc-ILD.

Measurement of static pulmonary physiology and, more recently, assessments of high-resolution CT (HRCT) scans have been shown to provide important prognostic information about patients with fibrosing interstitial lung disease related to systemic sclerosis (SSc-ILD). For patients with idiopathic fibrosing ILD, assessments of blood oxygenation—particularly those done while patients are exerting—have also been shown to predict outcome.

Maximal cardiopulmonary exercise test
All subjects underwent maximal cardiopulmonary exercise testing at our institution according to a standardised protocol. A radial arterial line and a peripheral pulse oximeter were placed before commencing exercise. The peripheral pulse oximeter was placed on the index finger of the hand opposite the arterial line. If an adequate pulse oximeter signal was not obtained (inability to obtain correct pulse), an earlobe probe was used. In more than 95% of subjects, finger probes were used. Baseline measurements were collected after

1 Interstitial Lung Disease Program and Autoimmune Lung Center, National Jewish Health, Denver, Colorado, USA; 2 Division of Psychosocial Medicine, National Jewish Health, Denver, Colorado, USA; 3 Division of Biostatistics and Bioinformatics, National Jewish Health, Denver, Colorado, USA

Correspondence to: Dr J J Swigris, Interstitial Lung Disease Program and Autoimmune Lung Center, National Jewish Medical and Research Center, 1400 Jackson Street, Denver, CO 80206, USA; swigrisj@njc.org

Received 20 November 2008
Accepted 26 March 2009
Published Online First 7 April 2009
models were used to examine the prognostic capabilities of SpO2. We confirmed the proportionality assumption was met for the dichotomised SpO2 variables by examining log(-log) plots. All statistical analyses were performed using SAS Version 9.1 (SAS Institute, Cary, North Carolina, USA). A p value of <0.05 was considered to be statistically significant.

RESULTS

The baseline characteristics of the subjects are shown in table 1. SpO2 overestimated SaO2 both at rest and at maximal exercise. The mean (SD) difference between SpO2 and SaO2 (keeping overestimates as positive values and underestimates as negative values) was similar at maximal exercise and rest (1.5 (4) vs 1.5 (3), p = 1, fig 1). Similar results were obtained when absolute values of the difference between SpO2 and SaO2 were used (2.24 (1.91) vs 2.98 (2.98), p = 0.06). SpO2 misclassified four subjects at maximum exercise (SpO2 observed to be >88% but SaO2 was <88%). For these subjects, median (IQR) values for baseline SpO2, SaO2 and arterial oxygen tension (PaO2) and maximum exercise SpO2, SaO2 and PaO2 were 91.5 (89–96), 88.5 (88–90), 61 (59.5–67), and 89.5 (89–93), 82 (79–88), 47.5 (43.5–62.5).

Over the study period (median follow-up 10.3 years, IQR 4–17) there were 40 deaths. Truncating follow-up at 20 years, we observed 39 deaths over the study period (median follow-up after truncation 7.1 years). Median survival was 9.5 years (IQR 4–16). Subjects whose SpO2 at maximum exercise (SpO2max) fell to <89% had shorter survival than those whose SpO2max remained ≥89% (p = 0.01, log rank test; fig 2). The results were similar when stratifying subjects on whether SpO2max fell >4 points from baseline (p = 0.01, log rank test). In Cox proportional hazards models, SpO2 was a significant predictor of mortality; over a median 7.1 years of follow-up, the risk of death was 2.4 times greater for subjects whose SpO2max fell below 89% (hazards ratio (HR) 2.4, 95% CI 1.2 to 4.9, p = 0.02) than for subjects whose SpO2max remained ≥89%. Similarly, the risk of death was 2.4 times greater for subjects whose SpO2max fell >4 points from baseline (HR 2.4, 95% CI 1.1 to 5.0, p = 0.02) than for subjects whose SpO2max remained within 4 points of baseline values. When analysed as a continuous variable, the difference between baseline SpO2 and SpO2max remained a significant predictor (HR 1.08, 95% CI 1.01 to 1.14, p = 0.002). Controlling for FVC%, the difference between baseline SpO2 and SpO2max (continuous variable) remained a significant predictor (HR 1.07, 95% CI 1.01 to 1.14, p = 0.02). Figure 3 shows the relationship between percentage lung carbon monoxide transfer factor (TiCO2%) and SpO2max.

DISCUSSION

With two goals in mind—to assess the usefulness of SpO2 as a prognostic marker in patients with SSC-ILD and to determine whether an arterial line is needed to accurately assess oxygenation—we conducted a study to first examine agreement between SpO2 and SaO2 at rest and maximal exercise and then to analyse the ability of SpO2 to predict mortality in patients with SSC-ILD. We hypothesised that, in subjects with SSC-ILD, SpO2 would inaccurately reflect SaO2 at rest and the disparity would be even greater at maximal exercise. In contrast, we found that SpO2 was an accurate reflection of SaO2 both at rest and maximal exertion in these subjects. Moreover, we observed that SpO2, a simple non-invasive and inexpensive measure to collect, was a predictor of mortality in patients with SSC-ILD.

Recently there has been a groundswell of attention on the use of non-invasive markers of exertional blood oxygenation (eg,
nadir $\text{SpO}_2$ during a 6-minute walk test (6MWT) or statistical manipulations of $\text{SpO}_2$ over the course of a timed walk test) as outcome measures in therapeutic trials and clinical studies enrolling subjects with ILD.\textsuperscript{71 41 5} This increased attention raises three important distinct but related questions regarding the use of $\text{SpO}_2$ at maximal exercise as an outcome metric:

![Figure 1](image1.png)

**Figure 1** Plots of peripheral oxygen saturation ($\text{SpO}_2$) against arterial oxygen saturation ($\text{SaO}_2$) and Bland-Altman plots for subjects with fibrosing interstitial lung disease related to systemic sclerosis (SSc-ILD) at (A) baseline and (B) maximum exercise. (A) In the upper graph the diagonal line marks the line of unity, and in the lower graph the middle line marks the mean difference between $\text{SpO}_2$ and $\text{SaO}_2$ and the two outer lines mark twice the standard deviation in either direction. (B) In the upper graph the diagonal line marks the line of unity, and in the lower graph the middle line marks the mean difference between $\text{SpO}_2$ and $\text{SaO}_2$ and the two outer lines mark twice the standard deviation in either direction.

![Figure 2](image2.png)

**Figure 2** Kaplan–Meier survival curves for subjects with fibrosing interstitial lung disease related to systemic sclerosis (SSc-ILD) stratified on peripheral oxygen saturation at maximal exercise ($\text{SpO}_2$ max) < 89% or ≥ 89%. The solid line represents subjects whose $\text{SpO}_2$ max remained ≥ 89% and the dashed line represents subjects whose $\text{SpO}_2$ max fell to < 89%. Tic = censored observation. Median survival for the entire cohort = 9.5 years.

![Figure 3](image3.png)

**Figure 3** Relationship between peripheral oxygen saturation ($\text{SpO}_2$) at maximum exercise and resting percentage lung carbon monoxide transfer factor ($\text{TLCO}$%).

$$r = 0.62$$

$$p < 0.0001$$
Is it valid—does it in fact measure what it is purported to measure (eg, true blood oxygenation or $\text{SaO}_2$)?

Is it reliable—if it is measured at two separate time points in a subject whose clinical status has not changed, will it produce similar results?

Is it responsive to underlying change—if a subject’s blood oxygenation at maximal exercise changes from baseline, will $\text{SpO}_2$ reflect those changes?

The current study shows that $\text{SpO}_2$ at maximal exercise is a valid measure of blood oxygenation at maximal exercise in patients with SSC-ILD, and $\text{SpO}_2$ does in fact accurately track changes in $\text{SaO}_2$.

The Bland-Altman plots reinforce this finding. These plots give a graphical presentation of the agreement between two methods of measurement, they depict an estimate of the bias (or systematic error which is simply the overestimation or underestimation of one measure compared with the other) as the mean difference between the two measures. The precision of that estimate is reflected in its standard deviation. Whereas correlation coefficients express the relationship between two variables, Bland-Altman plots depict agreement between them. When one is trying to determine the accuracy with which one measure (eg, $\text{SpO}_2$) reflects another (eg, $\text{SaO}_2$) or whether one measure might be used in place of another measure, correlation may not tell the true story—there can be extremely high correlation between two measures but, at the same time, poor agreement. This study shows that, for patients with SSC-ILD, $\text{SpO}_2$ is an accurate reflection of $\text{SaO}_2$ at rest or maximal exercise.

Several studies have examined the agreement between $\text{SpO}_2$ and $\text{SaO}_2$ but, to our knowledge, this is the first in a cohort with SSC-ILD. The importance and clinical relevance of this study centres on the peripheral circulation issues in SSC that make most clinicians reluctant to place an arterial line (digits have been lost as a consequence) and wary of $\text{SpO}_2$ accuracy in these patients. There is therefore a need to validate $\text{SpO}_2$ in SSC-ILD. In general, $\text{SpO}_2$ may either overestimate or underestimate $\text{SaO}_2$. In a meta-analysis Jensen and colleagues reported that, among 23 studies for which bias and precision estimates were available, the absolute mean (SD) bias was 1.99 (0.23) (ie, on average, $\text{SpO}_2$ overestimated $\text{SaO}_2$ by 1.99 points). In those studies the mean (SD) difference between $\text{SpO}_2$ and $\text{SaO}_2$ ranged from −13.2 (8.0) to 12.0 (13.3). The authors commented that severe or rapid desaturation; hypotension, hypothermia, or other unstable haemodynamic or low perfusion states; dys-haemoglobinaemia or use of vital dyes; and motion may all confound agreement between $\text{SpO}_2$ and $\text{SaO}_2$. The mean differences between $\text{SpO}_2$ and $\text{SaO}_2$ in the current study fall well within the range mentioned in that analysis.

The results of the current study not only suggest that $\text{SpO}_2$ is a valid surrogate for $\text{SaO}_2$ in patients with SSC-ILD, but also suggest that desaturation, as measured by $\text{SpO}_2$, is a significant predictor of mortality in this patient group. Our results are in line with the work by Lama and colleagues that suggested desaturation (as measured by $\text{SpO}_2$) during a 6MWT is an important prognosticator in patients with idiopathic interstitial pneumonia. In so far as the 6MWT accurately reflects functional exercise capacity in patients with SSC-ILD—it does so in patients with fibrotic idiopathic interstitial pneumonia—we hypothesise that our results would hold for measures of $\text{SpO}_2$ collected during the 6MWT in this patient population. Not surprisingly, we found TLCO% was the only significant predictor. Like other investigators, we also found TLCO% to be a potent predictor of survival in our cohort (data not shown). Because of the strong relationship between TLCO% and $\text{SpO}_{2\text{max}}$, and because our goal was merely to begin to examine $\text{SpO}_{2\text{max}}$ as a prognostic marker, we performed our survival analysis adjusting for FVC% and not TLCO%.

Although the results are novel and clinically relevant, this study has limitations. This is a retrospective analysis of data collected prospectively over a period of three decades. Different pulse and co-oximeters were used during different time periods; however, each instrument is purported to be accurate within two percentage points for $\text{SaO}_2$ values from 70–100% so we can be confident in the readings. Data for this study were collected at a centre situated 5280 feet above sea level. In Denver, patients probably “live” closer to the steep portion of the oxygen dissociation curve (probably on or very close to the shoulder) than patients at lower altitudes. How this affects the results merits consideration and examination in future studies. Given the lack of systematic examinations for pulmonary hypertension and the changes in available technology to assess for pulmonary hypertension over the study period, we cannot be certain how many subjects truly had the condition. Even more complex is the issue of exercise-induced pulmonary hypertension: how many subjects had it is unknown but, as with other studies of subjects with ILD, the possibility of its presence and its effects on exercise $\text{SpO}_2$ must be considered. Given these limitations, we believe the results should be viewed as hypothesis-generating and will hopefully spark continued investigation in this area. These results will need prospective confirmation at other altitudes. Future studies should examine whether $\text{SpO}_2$ values collected during the 6MWT are as meaningful as those collected during cardiopulmonary exercise tests, and efforts should be made to further delineate the relationship between resting or exercise-induced pulmonary hypertension and $\text{SpO}_2$.

CONCLUSION

In SSC-ILD, at both baseline and maximal exercise, $\text{SpO}_2$ is an accurate reflection of $\text{SaO}_2$. In patients with SSC-ILD, $\text{SpO}_2$ carries prognostic value. Because of the ease with which it is assessed, consideration should be given to measuring exercise $\text{SpO}_2$ as a marker of clinical status or as an outcome in clinical trials enrolling subjects with SSC-ILD. Future research could clarify a number of outstanding and important questions related to $\text{SpO}_2$ in this patient population.

Competing interests: None.

REFERENCES


Paediatric respiratory syncytial virus infections: rates and risk factors for hospitalisation

This population-based prospective study set out to establish the rates of respiratory syncytial virus (RSV)-associated paediatric respiratory tract infections in both inpatient and outpatient settings. Cases were identified through the New Vaccine Surveillance Network, a Centre for Disease Control initiative to survey acute respiratory infections in children aged <5 years in three geographically diverse US counties. The study was conducted during November to April between 2000 and 2004.

Nasal and throat swabs were analysed by RT-PCR and culture for RSV and other viruses including influenza, parainfluenza, rhinovirus and metapneumovirus. Of 6225 children with acute respiratory infections identified during the surveillance period, 5067 (81%) were enrolled of which 919 (18%) were RSV-positive; 6% of those were co-infected, most commonly with influenza virus. Three hundred and fifty-five RSV-infected children (39%) were treated in the community or emergency department. Those hospitalised (61%) were more likely to be under 6 months of age (58% vs 25% of outpatients) and to have been born prematurely (16% vs 8% of outpatients). The only socioeconomic factors found to be significantly associated with the risk of hospitalisation were breast feeding of less than 1 month duration and the presence of preschool children within the household.

This study attempts to identify the burden of RSV-associated acute respiratory infections in the community and hospital settings. While infection trends may be similar in the UK paediatric population, rates of hospitalisation are likely to vary.