Is it time to include oxygen needs as an endpoint in clinical trials in patients with fibrosing interstitial lung disease? If so, how?

Kerri Aronson,1 Susan S Jacobs,2 Dawn Repola,3 Jeffrey J Swigris

ABSTRACT

Many patients with fibrosing interstitial lung disease (fILD) will need to use supplemental oxygen (O₂) to maintain normoxia at some point in their illness. If it is not needed at the time of diagnosis, then if fILD progresses—or if a comorbid condition like pulmonary hypertension develops—O₂ will become necessary, often, initially, during exertion and all-too-often, eventually, at rest as well. But presumably, if all else remains stable, if fILD progression is halted or slowed, O₂ needs follow in parallel. Despite perceived or unnoticed benefits of O₂, and prescribers’ good intentions to improve patients’ sense of well-being, patients with fILD generally view O₂ with frustration and fear, as it threatens their already-impaired quality of life. Because of how meaningful and impactful O₂ is to the lives of patients with fILD, ‘O₂ need’ is a critically important—and perhaps the most—patient-centred metric that should be considered for incorporation as an endpoint in therapeutic trials. It is unclear how this should be done, but in this paper, we offer some possible approaches that merit consideration.

Practitioners prescribe oxygen (O₂) to patients with fibrosing interstitial lung disease (fILD) in hopes of the following: (1) that it will limit desaturation events and combat breathlessness; (2) that it will allow patients to be more active physically and socially; (3) that it will stave off putative complications of hypoxaemia (eg, cognitive dysfunction, pulmonary hypertension) and (4) that it will improve health-related quality of life (HRQL).

However, despite the rationale for O₂, and prescribers’ good intentions, patients with fILD generally view O₂ with frustration and fear—it threatens their HRQL which is already impaired by having a condition that imposes itself on every aspect of their lives; nasal cannulas call unwanted attention to patients when they are out in public; O₂ users feel stigmatised and are, in our patient–author’s experience, viewed as ‘smokers who get what they deserve’, even if they never smoked a day in their lives; O₂ delivery equipment is typically heavy, unwieldy and intimidating. Also, O₂ disrupts the home environment, and, sadly, O₂ is a constant reminder to patients they are living with a condition that could shorten their lives.

To no surprise, and as results from several observational studies reveal, the need for O₂ is associated with disease progression and/or shortened survival.5–7 But, as shown in the Effect of Ambulatory Oxygen on Quality of Life for Patients with Fibrotic Lung Disease (AmbOx) trial (which included patients with fILD who desaturated on a timed walk test but were normoxic at rest), compared with no use, the unblinded use of O₂ over a 2-week period was associated with better HRQL.8 A follow-up analysis revealed the cost-effectiveness of ambulatory oxygen for improving HRQL in fILD.9 However, in larger, cross-sectional studies, compared with patients with fILD who do not need O₂, those who need it report worse quality in multiple life domains, including emotional well-being, social participation and independence.10–12 Thus, for many, O₂ is a double-edged sword. Studies are ongoing to add to the evidence on the efficacy and cost effectiveness of O₂ at rest and/or with ambulation in patients with fILD.13 14 The clinician–authors who have cared for patients with fILD know firsthand patients’ resilience and adaptability to worsening disease; however, the need for O₂ (needing it at all or needing increasingly higher flows) remains an inescapable concern for most of them.

Given its tremendous meaningfulness to patients and its potential effect on HRQL, ‘O₂ need’ is a critically important and inarguably patient-centred endpoint that we and others believe should be considered for incorporation into therapeutic trials.15 Below, we discuss the challenges and possibilities for including ‘O₂ needs’ as an endpoint. Recognising the subtle differences between ‘O₂ needs’, ‘O₂
prescription’ and ‘O₂ use’, we cover certain aspects of each and argue that ‘O₂ needs’ is the best option of the three for an O₂-centric endpoint. Our hope is that we as a field can extend the conversation around O₂-centric endpoints and devote research efforts to figure out how best to incorporate O₂ needs into therapeutic trials.

THE PROBLEMS WITH ‘O₂ PRESCRIPTION’
Significant barriers to operationalising an O₂-centric endpoint in therapeutic trials include the variability in defining use/need and the lack of a robust foundation endpoints and devote research efforts to figure out how successful it is to incorporate O₂ needs into therapeutic trials.

For patients with fILD, it seems there is general agreement around prescribing O₂ to patients with resting hypoxaemia; however, that is not necessarily the case for prescribing O₂ for isolated exertional hypoxaemia. Several things equal or greater than blood oxygen levels could influence whether a patient is prescribed O₂, some of which have to do with the practitioner, including their individualised habits around whether, when and how to test patients. Some practitioners may not prescribe O₂ regardless of the degree of exertional desaturation. Practitioners who prescribe, typically do so on a trial basis, and (ideally) after a thorough discussion of theoretical (and possibly real) benefits—and the likely burdens O₂ imposes on the patient—while incorporating patients’ preferences, goals and values into a shared decision.

### Table 1: Potential trial endpoints focused on supplemental oxygen

<table>
<thead>
<tr>
<th>Potential endpoint</th>
<th>Definition</th>
<th>Potential scenarios for implementation/use in a therapeutic trial</th>
<th>Results if therapeutic agent is effective in the trial</th>
<th>Drawbacks/limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>O₂ prescription</td>
<td>New prescription for O₂ in an O₂-naive individual.</td>
<td>▶ Compare proportions of O₂-naive subjects in each arm who are prescribed O₂ after Visit 1 of the trial.</td>
<td>▶ Proportion of O₂-naive subjects in treatment arm prescribed O₂ will be lower than in the placebo arm.</td>
<td>▶ Variability and potential bias introduced if O₂ prescription left to subject’s treating physician.</td>
</tr>
<tr>
<td></td>
<td>- Made up on randomisation (eg, O₂ concentrator or flow).</td>
<td>▶ Perform time-to-O₂ prescription analysis among O₂-naive subjects in each arm.</td>
<td>▶ Among O₂-naive subjects, time to O₂ prescription will be longer in the treatment arm than in the placebo arm.</td>
<td>▶ Enough O₂-naive subjects will need to be recruited (and equitably randomised to each arm) to allow for the detection of between-group differences.</td>
</tr>
<tr>
<td>O₂ use</td>
<td>O₂ use measured in time, volume or rate.</td>
<td>▶ Compare mean, per 24-hour O₂ use between arms.</td>
<td>▶ Normalised for activity (eg, per MET expended) relative to baseline use (eg, litres/week), use of O₂ among subjects in treatment arm will be less than use of O₂ among subjects in the placebo arm.</td>
<td>▶ Results will be markedly influenced by subjects’ activity level and whether they adjust flows throughout the day.</td>
</tr>
<tr>
<td></td>
<td>- Made up on randomisation (eg, via wearable or activity questionnaire).</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O₂ need</td>
<td>Among O₂-naive subjects, new need for O₂ and among O₂ users, flow or volume of O₂ needed to maintain blood oxygen levels (at rest or during standardised activity).</td>
<td>▶ Compare proportions of O₂-naive subjects in each arm who need O₂ after Visit 1 of the trial according to ‘O₂ needs test’.</td>
<td>▶ Difference in O₂ need to maintain SpO₂&gt;89% on a constant speed treadmill walk test will be lower for subjects in treatment arm than subjects in the placebo arm.</td>
<td>▶ Walk efficiency issues with treadmill.</td>
</tr>
<tr>
<td></td>
<td>- Made up on randomisation (eg, via wearable or activity questionnaire).</td>
<td>▶ Compare change in O₂ needs from baseline needed to maintain SpO₂ on O₂ needs test.</td>
<td>▶ Time to increased O₂ needs (ie, flow needed to maintain SpO₂&gt;89%) on a constant speed treadmill walk test will be longer in the placebo group than in the treatment group.</td>
<td>▶ Are there learning effects?</td>
</tr>
<tr>
<td></td>
<td>- Made up on randomisation (eg, via wearable or activity questionnaire).</td>
<td>▶ Analyse time-from-baseline-to-O₂ need among O₂-naive subjects in each arm.</td>
<td>▶ After Visit 1, proportion of O₂-naive subjects who need O₂ to maintain SpO₂&gt;89% on a constant speed treadmill walk test will be lower in the treatment arm than in the placebo arm.</td>
<td>▶ Uncertainties around pace and duration.</td>
</tr>
<tr>
<td></td>
<td>- Made up on randomisation (eg, via wearable or activity questionnaire).</td>
<td>▶ Perform time-from-baseline-to-increased-O₂ needs (litres as continuous variable vs categorical; eg, new need for high-flow) analysis among O₂-naive subjects in each arm.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

MET, metabolic equivalent; O₂, oxygen; SPO₂, oxygen saturation.
Country-specific or region-specific standards of care vary starkly, with some not offering any portable O₂ option, and others doing so only after standardised (in some cases blinded) testing to document objective benefit. Geographic and regional discrepancies in O₂ availability/supply and patients’ social determinants of health will also influence access to O₂. Given the variability in prescribing patterns and other potential influencers, without a strict study protocol for prescription, we believe the proportion of patients prescribed O₂ since baseline visit would not be a good endpoint for clinical trials.

THE PROBLEMS WITH ‘O₂ USE’
Likewise, variable use of O₂ by patients with fILD—and unreliable availability and delivery of O₂—exposes the outcome ‘O₂ use’ to similar biases and inaccuracies as O₂ prescription, thus limiting its utility as a valid endpoint for clinical trials. Whether O₂ use is measured in time, volume or rate (eg, litres of O₂ per unit of time), measured use will depend, to a large degree, on a person’s activity: assuming they increase the flow of O₂ when active, patients who are more active will use more O₂. Although physical activity could be considered as a stratification factor for trial enrolment, precisely how stratification would incorporate duration and/or intensity of physical activity makes it complicated at best. When and on what basis subjects change their flow rates throughout the day would also significantly bias results. As an example of how most patients with fILD typically use O₂—they reluctantly accept it and then judge for themselves how and when their lives accommodate it—consider results from an observational study of 50 patients with ILD (the majority with fILD; 16 of whom used O₂): in this study, the investigators found that the 34 patients who did not use O₂ had a nadir oxygen saturation (SpO₂) of 84±7% during continuous monitoring in their homes. On average, these patients spent about 50% of their in-clinic timed walk test (6-minute walk test (6MWT))—and 16% of their monitored time at home—with an SpO₂<90%. Startlingly, the 16 patients who used O₂ also spent 50% of their 6MWT—and over 20% of their monitored time at home—with an SpO₂<90% despite using O₂ at their prescribed flow rates. The results from the study clearly highlight the difference between merely using O₂ and precisely identifying O₂ needs (and using amounts needed to maintain a target SpO₂); the latter is seemingly far better suited for consideration as a formal research endpoint.

HOW SHOULD WE CONSIDER ‘O₂ NEEDS’
Perhaps the first challenge to overcome is developing a sensical, valid and reliable definition for ‘O₂ needs’ in fILD. Resting alveolar-arterial O₂ difference (A-a) is

![Figure 1](http://bmjopenrespres.bmj.com/)

Figure 1 Adapted from the conceptual framework developed by the Food and Drug Administration for assessment and selection of patient-reported outcome measures. ADL, activities of daily living; COA, clinical outcome assessments; ESWT, endurance shuttle walk test; HRQL, health-related quality of life; ILD, interstitial lung disease; ISWT, incremental shuttle walk test; O₂, oxygen; SpO₂, oxygen saturation; 6MWT, 6-minute walk test.
apparently not the answer: it was included in a composite endpoint in a trial of interferon gamma-1b in patients with idiopathic pulmonary fibrosis whose partial pressure of arterial oxygen on room air was greater than 55 mm Hg.19 Variability in room air, resting A-a from screening to enrolment led investigators to deem it unreliable and thus ill-suited for use as a stand-alone endpoint or a component in a composite endpoint.20 Although pulmonary alveolar proteinosis (PAP) is not an fILD, problems with A-a were also seen in a trial of inhaled molgramostim for PAP.21

So, for resting O2 needs, maybe it is simpler: the lowest O2 flow needed to maintain an SpO2 of at least 90% at rest. But, on closer inspection, it may not be so straightforward. For example, how long to monitor SpO2 to determine need? Five seconds? Thirty seconds? One minute? What position should the patient be in? Standing? Seated upright with feet flat on the floor? How are patients supposed/allowed to breathe: is deep or pursed-lip breathing allowed? Perhaps SpO2 taken at 1 min via ear probe (without patients able to see the read out or hear any alarm), while the patient is seated, feet flat on floor and breathing quietly without pursed lips is a reasonable, although not evidence-based, starting point.

Even more subtleties must be considered for defining exertional O2 needs, particularly when referring to a determination employed within the confines of a therapeutic trial. Should we define O2 need as the lowest amount of O2 required to maintain SpO2 90% or greater throughout the duration of a challenge test? Variable effort during a field walk test (eg, 6MWT, incremental or endurance shuttle walk test, Glitter-activities of daily living test) could create error in determining O2 needs.22 Some experts have proposed calculating the distance saturation product (DSP), or the product of the distance walked and lowest SpO2 during the walk test to 'normalize' for effort.23 However, the DSP fails to precisely determine O2 needs.

A treadmill walk is not recommended as a test of exercise capacity or physical functioning in patients with COPD (and by extension, patients with fILD), due to poor walking efficiency on the treadmill in people who are not used to doing it. However, since our main focus here is on O2 needs and not distance, holding the pace constant—as could be done with the treadmill—would alleviate much of the variability stemming from patient effort. O2 needs could be determined as the lowest flow necessary to maintain SpO2 at least 90% for a timed walk test, performed at a given speed, held constant for an individual for all trial assessments. Since distance is not the outcome, walking efficiency is less of a concern. But, many important unanswered questions remain about the ability to get an accurate assessment of O2 needs using this method. For example, how do we determine the speed at which to set the treadmill (gender and height would have to be considered)? What impact does treadmill walking learning effect have on the outcome of interest? Presumably, the speed would be the same for all subjects—or all subjects of the same gender or height strata. Would the pace be the same for a subject who does not require O2 and one requiring 6 L/min at rest? How long should the walk be? Six minutes? Two minutes? Inaccuracies of SpO2 from skin pigmentation will also have to be considered and reconciled.24

We believe an ‘O2-2-3-4’ test is a reasonable starting point at which to begin the reliability/validity investigation. In this test, the patient walks on a treadmill for 2 min, at 3 miles/hour and an incline of 4%. This is derived from the Dyspnoea Challenge for COPD in which the treadmill is fixed at 3km/hour and an incline of 4%.25 Although the purpose here is to identify O2 needs, the test allows for capture of other potentially important metrics; for example, desaturation ≥4% from start of test.

The pros, cons and unanswered questions related to potential endpoints of supplemental O2 in fILD are outlined in table 1. The US Food and Drug Administration has provided updated guidance surrounding how to collect and interpret information to define endpoints that guide development of therapeutics and regulatory decision-making.26 This includes an approach to selecting, modifying, developing and validating clinical outcome assessments (COA) to measure patient-centred outcomes in clinical trials. In figure 1, we use this framework to display how we might consider an O2-centric endpoint in fILD therapeutic trials. Once there is agreement around the definition of O2 needs and the methods for assessing it, other considerations will need to be addressed. For example, particularly if O2 needs become a high-tier endpoint, the effects of skin pigmentation, altitude and resting SpO2 will need to be carefully considered—and perhaps used as a stratification variable(s) at randomisation. A host of analyses could be considered, including (but certainly not limited to) the following: proportion of O2-naïve patients with new O2 need; proportion of patients with stable or increased O2 needs, categorised by degree of increase using various cutoffs (>1L/min, >4L/min); and various time-to-event analyses. In pragmatic, real-world trials, testing could be performed in centres with pulmonary rehabilitation programmes (or even virtually if patients have access to a treadmill).

Supplemental O2 needs have been identified by patients themselves as the real-life ‘staging system’ of living with fILD.2 A new or increased need for O2 is inarguably one of the most important milestones for patients with fILD, as it signals a new way of living for patients and their families. We need collaborative efforts among clinicians, researchers and patients, as well as support from funding agencies to design and implement effective studies that will address how to best operationalise O2 needs as a COA. A Delphi approach could be useful here. The need for validation work and the challenges should not dissuade consideration of this patient-centred, critically important and clinically meaningful metric as a potential stand-alone or component of a composite outcome measure.

Patients with fILD deserve to be armed with information
around if or when to expect a prescription for $O_2$ (or significant increases in flow rates) and what therapies may delay or prevent this occurrence. Now is our opportunity to bring resources and patients’ input to bear on sorting out how to assess $O_2$ needs in clinical trials as an initial step toward reaching this critically important goal.

**Twitter** Jeffrey J Swigris @SwigOutFishing

**Acknowledgements** We dedicate this manuscript to DR who passed away after the paper was written. DR provided thoughtful contributions to this manuscript and was a strong advocate for patients living with interstitial lung disease. She will be dearly missed.

**Contributors** KA, SSJ, DR and JJS: Substantial contributions to the conception or design of the work. KA, SSJ, DR and JJS: Drafting the work and revising it critically for important intellectual content. KA, SSJ, DR* and JJS: Final approval of the version to be published. KA and SSJ: Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. "Approved prior version. There are no competing interests for any author. This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

**Funding** The authors have declared no specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interests** None declared.

**Patient consent for publication** Not applicable.

**Ethics approval** Not applicable.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

**ORCID iD** Jeffrey J Swigris http://orcid.org/0000-0002-2643-8110

**REFERENCES**


24. Sjoding MW, Ivashyna TJ, Valley TS. Change the framework for pulse Oximeter regulation to ensure clinicians can give patients the oxygen they need. *Am J Respir Crit Care Med* 2023;207:661–4.

