

RESEARCH

Open Access



Direct and indirect evidence of efficacy and safety of rapid exercise tests for exertional desaturation in Covid-19: a rapid systematic review

Asli Kalin¹, Babak Javid², Matthew Knight^{3*} , Matt Inada-Kim⁴ and Trisha Greenhalgh¹

Abstract

Background: Even when resting pulse oximetry is normal in the patient with acute Covid-19, hypoxia can manifest on exertion. We summarise the literature on the performance of different rapid tests for exertional desaturation and draw on this evidence base to provide guidance in the context of acute Covid-19.

Main research questions:

1. What exercise tests have been used to assess exertional hypoxia at home or in an ambulatory setting in the context of Covid-19 and to what extent have they been validated?
2. What exercise tests have been used to assess exertional hypoxia in other lung conditions, to what extent have they been validated and what is the applicability of these studies to acute Covid-19?

Method: AMED, CINAHL, EMBASE MEDLINE, Cochrane and PubMed using LitCovid, Scholar and Google databases were searched to September 2020. Studies where participants had Covid-19 or another lung disease and underwent any form of exercise test which was compared to a reference standard were eligible. Risk of bias was assessed using QUADAS 2. A protocol for the review was published on the Medrxiv database.

Results: Of 47 relevant papers, 15 were empirical studies, of which 11 described an attempt to validate one or more exercise desaturation tests in lung diseases other than Covid-19. In all but one of these, methodological quality was poor or impossible to fully assess. None had been designed as a formal validation study (most used simple tests of correlation). Only one validation study (comparing a 1-min sit-to-stand test [1MSTST] with reference to the 6-min walk test [6MWT] in 107 patients with interstitial lung disease) contained sufficient raw data for us to calculate the sensitivity (88%), specificity (81%) and positive and negative predictive value (79% and 89% respectively) of the 1MSTST. The other 4 empirical studies included two predictive studies on patients with Covid-19, and two on HIV-positive patients with suspected pneumocystis pneumonia. We found no studies on the 40-step walk test (a less demanding test that is widely used in clinical practice to assess Covid-19 patients). Heterogeneity of study design precluded meta-analysis.

(Continued on next page)

* Correspondence: matthewknight@nhs.net

³West Hertfordshire Hospitals NHS Trust, Vicarage Rd, Watford, Hertfordshire WD18 0HB, UK

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

(Continued from previous page)

Discussion: Exertional desaturation tests have not yet been validated in patients with (or suspected of having) Covid-19. A stronger evidence base exists for the diagnostic accuracy of the 1MSTST in chronic long-term pulmonary disease; the relative intensity of this test may raise safety concerns in remote consultations or unstable patients. The less strenuous 40-step walk test should be urgently evaluated.

Keywords: Covid-19, Desaturation, 1-min sit-to-stand test, 6-min walk test, 40-step test, Normoxia, Silent hypoxia, Hidden hypoxia

Background

A substantial proportion of patients with acute coronavirus 19 (Covid-19) develop a potentially critical form of the illness requiring intensive care unit admission [1]. The degree of lung involvement in acute Covid-19 is variable, producing a spectrum of illness from mild upper respiratory tract symptoms to acute respiratory distress syndrome [2]. Patients with mild initial symptoms can rapidly deteriorate to severe or critical cases. In hospitalised patients, hypoxaemia and the need for oxygen are independent predictors of severe outcomes [3, 4]. The usual time from symptom onset to the development of severe hypoxemia is between 7 and 12 days [5, 6]. Recent prognostic tools such as the 4C score have emphasised the importance of identifying hypoxia early [3, 7] and there are physiological reasons for managing this hypoxia actively [8, 9].

The poor correlation between both subjective feeling of shortness of breath (dyspnoea) and objective measures of breathlessness and hypoxia in patients with Covid-19 has resulted in UK guidelines recommending that the assessment of the breathless, unwell or high-risk patient should include oximetry [5]. For example, a retrospective cohort study of 64 Covid-19 patients considered eligible for home oximetry monitoring showed that the presence of dyspnoea had a positive predictive value of only 42% for hypoxemia and absence of dyspnoea had a negative predictive value of 86% for excluding it [10].

Indeed, the mismatch between relatively mild subjective respiratory distress and objective evidence of peripheral hypoxia is now recognised as a feature of Covid-19 and has been termed ‘silent’ or ‘happy’ hypoxemia [11, 12]. Similarities with PCP have been drawn by others, and whilst a detailed discussion of pulmonary physiology is outside the scope of this literature review, this feature has been attributed to moderate to severe ventilation-perfusion mismatch [11], intra-pulmonary shunting, loss of lung perfusion regulation, intravascular microthrombi or reduced lung compliance [12, 13].

Hypoxia is common in acute severe Covid-19. Richardson et al. found that 28% of 5700 patients admitted to hospital with acute Covid-19 in New York City were sufficiently hypoxemic to need supplemental oxygen on admission [14]. Yet dyspnoea was reported in only 18.7% of 1099 hospitalised Covid-19 patients, despite low

PaO₂/FiO₂ ratios and requirement for supplemental oxygen in 41% [15]. In a study of the 19 worst-affected countries worldwide, Goyal et al. reported that rates of mortality were significantly higher in those countries where policies recommended lower oxygen saturations before administering supplemental oxygen than in those that aimed to normalised saturations, though there may have been other explanations for these differences including different degrees of preparedness for a pandemic and different incidence rates [16].

It is widely reported that some patients with acute Covid-19 have normal pulse oximetry at rest but their readings deteriorate on exertion (unpublished data). UK national guidance, for example, recommends the use of exertional desaturation tests to detect early deterioration of patients with COVID-19 in community and emergency settings [17–19]. In consensus exercises, front-line clinicians have identified the late transfer of patients with exertional desaturation (i.e. a fall of 3% or more in pulse oximetry reading on exercise) as a possibly remediable cause of poor outcome [20].

Several publications recommend or suggest the use of exertional tests in the assessment of COVID-19 patients. Mantha et al. discuss the use of a modified 6-min walk test (6MWT), wearing masks and only in those under 70, on day 4 or 5 of clinical illness to risk-stratify patients [6]. This suggestion is echoed by Noormohammadpour and Abolhasani, who propose that deterioration in the 6MWT test can be used to identify those who need referral to hospital [21]. Pandit et al. recommend the 6MWT for those under 60 who are not short of breath at rest and a 3-min walk test (3MWT) for those over 60 who are unable to complete the longer test, but warn that the test is contra-indicated in patients who are hypoxic at rest (SpO₂ < 94%), short of breath at rest, not able to walk unassisted, have Eisenmenger’s syndrome, severe anaemia, unstable angina or valvular heart disease [2]. They suggest that the 6MWT or 3MWT test can be performed at home or in hospital under the supervision of either a family member or a healthcare professional. These authors define exertional hypoxia as an absolute drop in oximeter reading by 3% or more from baseline. The level of 3% reduction from normal levels is drawn from the British Thoracic Society

emergency oxygen guidelines [22] consensus amongst clinicians and a recent empirical study [23].

In sum, consensus guidance and editorials recommend tests for exertional hypoxia in Covid-19, but the evidence base for these tests has not previously been formally reviewed. Indeed, the prognostic utility of exertional desaturation remains unknown. Another unknown is the safety of such tests in patients with suspected Covid-19, especially when used remotely without a clinician physically present. Establishing a validated tool to assess exertional desaturation will help to ensure that future research on this topic can be undertaken in a consistent way.

Review objective and research questions

The overall objective of this rapid review was to examine the published evidence base for the use of rapid exercise tests and assess their application to the assessment of patients with acute Covid-19. We were particularly interested in tests that could be used outside of the hospital setting, since the reality of acute Covid-19 often involves a remote assessment (with the patient at home at a distance from the clinician) or one in a bespoke ambulatory setting such as a ‘hot hub’ (where exercise tests may be performed outside in car parks, for example, for infection control reasons). We were, however, aware that when we began this study, little or no direct research evidence existed on the use of these tests in patients with Covid-19.

The research questions were:

1. What exercise tests have been used to assess exertional hypoxia at home or in an ambulatory setting in the context of Covid-19 and to what extent have they been validated?
2. What exercise tests have been used to assess exertional hypoxia in other lung conditions, to what extent have they been validated and what is the applicability of these studies to acute Covid-19?

Methods

We followed Cochrane Rapid Review guidelines [24] and reported our findings using the updated version of the Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies (PRISMA-DTA) checklist [25]; this is appended as Supplementary table S1. Our team included experienced health librarians as well as systematic reviewers and clinicians (including a general practitioner, general physician, respiratory consultant and emergency medicine consultant with experience of undertaking exercise tests). References were downloaded to Endnote (version 9.0) to maintain and manage citations and facilitate the review

process. A protocol for the review was published on the Medrxiv database [26].

Search strategy and selection criteria

Three independent searches were conducted. The first searched LitCovid, Scholar and Google using the terms ‘step test or field test’ and ‘hypoxia or exertional desaturation’. From these results, promising articles were then used to find additional documents via two methods—forward and backward citation matching, and searching for related articles (also using Microsoft Academic Search).

The second search covered the following databases: AMED, CINAHL, EMBASE MEDLINE and PubMed using the over-arching question ‘Very short exercise desaturation tests for use in the emergency department’. Search terms were:

- (quick OR short) AND oxygen) AND exercise) AND (desaturation OR saturation)) AND test*(‘emergency department*’ AND oxygen) AND exercise) AND desaturation) AND test*)
- (‘emergency department*’ AND oxygen) AND exercise) AND saturation) AND test*)
- (‘emergency department*’ AND oxygen) AND exercise) AND desaturation)
- (‘emergency department*’ AND oxygen) AND exercise) AND saturation)
- (1-min sit-to-stand test) or (‘1 min sit to stand test’) or (‘one minute sit to stand’)

The third search involved searching the Cochrane Database of Systematic Reviews (CDSR), the Cochrane Central Register of Controlled Trials (CENTRAL) and the Cochrane COVID-19 study register.

Search strings for different searches are listed below.

Search of Cochrane Library—Issue 10 of 12, October 2020:

- #1 ((‘step test’ OR ‘walk test’ OR ‘field test’ OR ‘chair rise’ OR ‘sit to stand’ OR ‘exercise test’ OR ‘exercise testing’):ti,ab,kw OR (exertional NEAR/2 (desaturation OR hypoxia)):ti,ab,kw
- #2 MeSH descriptor: [Coronavirus Infections] this term only
- #3 (coronavirus OR covid*):ti,ab,kw
- #4 #2 or #3
- #5 #1 AND #4

Search of Cochrane COVID-19 study register:

- ‘step test’ OR ‘walk test’ OR ‘field test’ OR ‘chair rise’ OR ‘sit to stand’ OR ‘exercise test’ OR ‘exercise testing’

Our inclusion criteria used a framework based on the PRISMA-DTA standards [25] as follows:

- *Target condition*: Individuals with suspected Covid-19 or another lung disease with or without symptoms.
- *Index test*: Any form of rapid exercise test performed at home or in a healthcare setting.
- *Reference standards*: 6MWT or CPET (which are the most commonly used in exercise testing), or a diagnostic test to diagnose the disease in question such as bronchoalveolar lavage to diagnose pneumocystis pneumonia.
- *Study designs*: clinical practice guidelines and systematic reviews addressing desaturation in exercise tests in lung disease, using the Cochrane definition of a Systematic Review [27]. Primary human studies of all designs (e.g. experimental studies, quasi-experimental studies, diagnostic accuracy studies and observational studies), excluding case series, that involved patients with COVID-19 or a lung disease undergoing rapid exercise testing were included. Editorials and letters around the subject of Covid-19 and rapid exercise testing were included to provide background and surface hypotheses on this new disease.
- *Outcome*. Pulse oximetry or arterial blood gas measurement, and association with any adverse outcome, e.g. hospital admission, need for organ support, death.
- *Time periods*: All periods of time and duration of follow up.

No other limitations were imposed on the search or study selection process. Both peer-reviewed and preprint papers were potentially eligible for inclusion. We planned to seek translation of any relevant papers published in languages other than English, but since none were found, the review was limited to English-language papers. Screening, data abstraction and quality appraisal of full-text papers were completed independently by two reviewers including a topic expert and a review expert. Disagreements were resolved by a consensus-based discussion.

Data elements were extracted from each study onto an Excel spreadsheet, including information on study population and size, study design, inclusion and exclusion criteria, research question, exercise tests used, results and authors' conclusion. Markers of study quality from each of the articles were identified in the literature search. All extracted data were included in a series of detailed tables and a summary table. There was no requirement to contact authors for missing data.

Risk of bias appraisal

The risk of bias in individual studies was assessed using the QUADAS 2 tool. Using the signalling questions, AK rated each potential source of bias as high, low or unclear; these ratings were checked independently by TG (see Table 1) [28].

Diagnostic accuracy measures

Where relevant, diagnostic accuracy measures are reported as sensitivity, specificity, positive and negative predictive values of the new test (index test) in relation to the gold standard (reference test).

Synthesis

Given the heterogeneity of the evidence included, formal statistical synthesis was not possible. Data from the included studies were tabulated in summary tables of origins, methods and results, and then summarised narratively.

Results

Overview of dataset

The study flowchart is shown in Fig. 1. From approximately 900 potentially relevant articles, we identified 47 relevant papers of which 15 were empirical studies and 32 were narrative reviews, editorials and commentaries. Of the empirical studies, only two (published in 2020) related to patients with Covid-19; neither was a validation study (both considered whether exertional desaturation predicted outcome). Two small studies (published in 1988 and 1994) considered exertional desaturation as a predictive test for pneumocystis pneumonia in people with HIV; we include them because of physiological parallels with Covid-19 discussed below. The other 11 studies were attempts at validation of one or more exercise desaturation tests in chronic lung diseases; they were published between 2007 and 2020 and included between 15 and 107 participants.

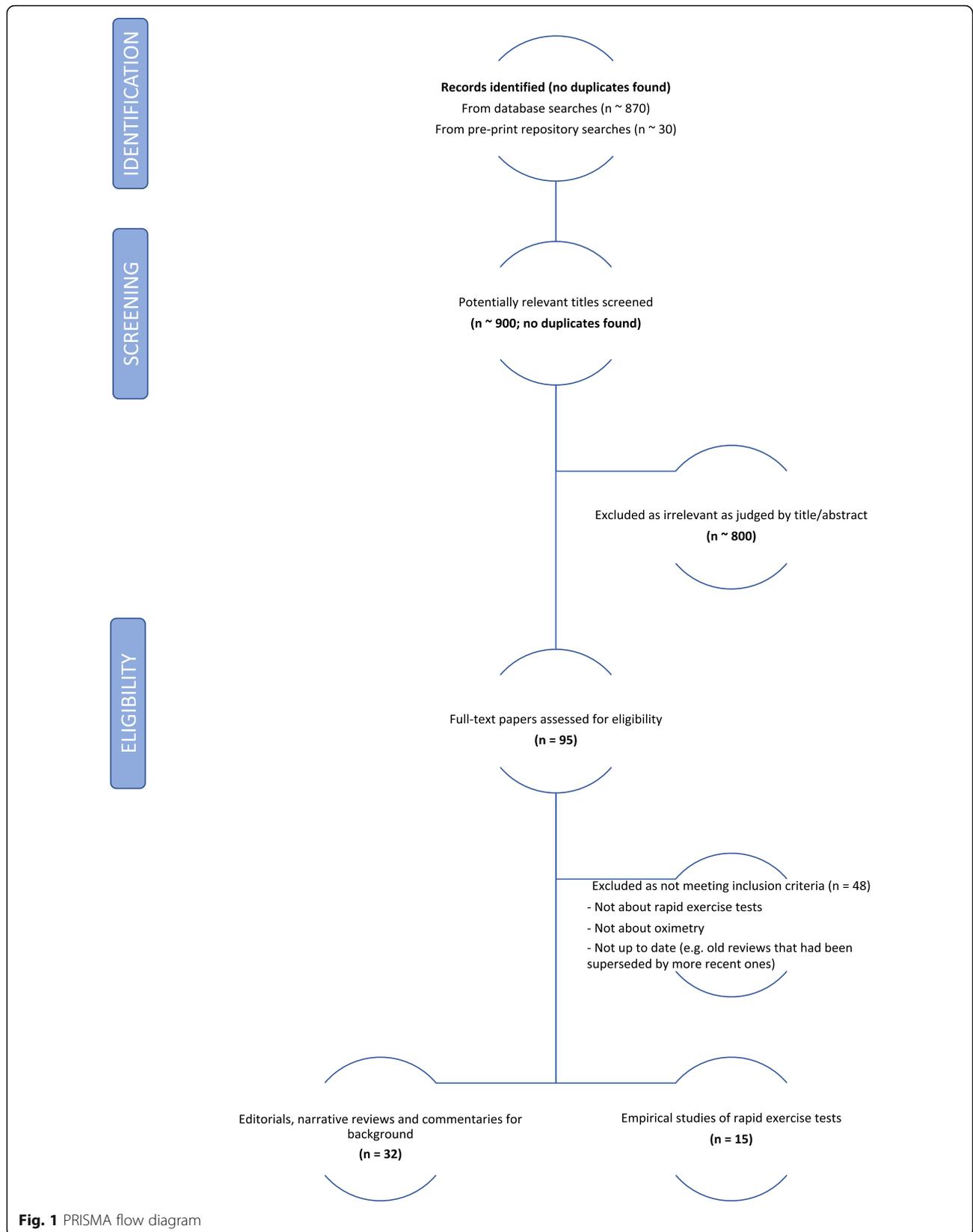
The empirical studies are summarised briefly in Table 1 and in more detail in Supplementary Tables S3–S7, which include a detailed risk of bias and applicability assessment. Overall, the methodological quality of most studies as assessed by the QUADAS 2 tool was uncertain. Of the 11 validation studies, none scored 'high quality' on all 7 dimensions in the QUADAS2 tool. One (Briand et al 2018 [29]) scored 'high quality' in 6 of the 7 dimensions but across the other 8 studies there were many aspects of methodological quality that scored poorly or could not be assessed confidently from the information supplied in the paper. In the sections below, we have placed more emphasis on the studies scoring higher on the QUADAS2 tool and on those which were undertaken on participants with perfusion defects.

Table 1 Summary of included empirical studies

Author / Year / Country	Study design	Risk of bias / applicability*	Sample and size	Inclusion / exclusion criteria	Research question	Tests used	Main results	Authors' conclusion
STUDIES OF EXERTIONAL DESATURATION TESTS IN COVID-19								
Fuglebjerg 2020 Denmark	Prospective	⊕ ⊕ ⊕ ⊕ ? ⊕ ⊕	"Discharge-ready" Covid-19 patients; n= 26	SPO2 >= 94% Excluded chronic lung disease	Can a 6MWT elicit silent hypoxia in Covid-19 patients?	6MWT	13/26 (50%) developed hypoxia (SPO2 <90%), of which 4 had pulmonary embolism	6MWT is useful in detection of silent hypoxia in Covid-19
Goodacre 2020 UK	Retrospective	⊕ ⊕ ⊕ ⊕ ⊕ ⊕ ⊕	Patients with Covid-19 who had a rapid exercise test; n= 817	Test done as part of routine care	Can a rapid exercise test predict outcome in Covid-19?	Not stated	Adverse outcome (need for organ support, death) correlated with test result	Rapid exercise test result has "modest" predictive value
STUDIES OF EXERTIONAL DESATURATION TESTS IN PATIENTS WITH HIV AND SUSPECTED PNEUMOCYSTIS PNEUMONIA								
Sauleda 1994 Spain	Prospective	⊕ ⊕ ? ? ⊕ ⊕ ⊕	HIV positive patients with clinical signs of PCP; n= 45	Excluded HIV negative, previous respiratory or cardiovascular disease	Can an exertional desaturation test be used for the initial diagnosis of PCP in HIV positive patients?	Pedalling motions in the air from lying on bed	Significantly greater desaturation in patients with PCP. Sensitivity was 77% and specificity 91%.	Exercise oximetry is a valuable tool in the early diagnosis of PCP in HIV positive patients.
Smith 1988 UK	Prospective	? ⊕ ? ? ⊕ ⊕ ⊕ ⊕	HIV positive patients with suspected PCP; n= 39	Exclusion criteria not specified	Can an exertional desaturation test distinguish PCP from other causes of pneumonia in HIV positive patients?	10MCT	Significantly greater exertional desaturation was found in patients with PCP than in those with other causes of pneumonia.	Exercise oximetry is potentially valuable tool in the early diagnosis of PCP in HIV positive patients.
VALIDATION STUDIES OF EXERTIONAL DESATURATION TESTS IN CONDITIONS OTHER THAN COVID-19- CHRONIC LUNG DISEASE								
Briand et 2018 France	Prospective	⊕ ⊕ ⊕ ⊕ ⊕ ⊕ ⊕	Interstitial lung disease; n= 107	Excluded infection or unstable disease	Can 1MSTST be used instead of 6MWT?	1MSTST 6MWT	Tests were strongly correlated (r = 0.9). In 90 of 107 participants both tests gave same result	1MSTST can be used as an alternative to 6MWT
Gephine 2020 Canada	Prospective	? ? ⊕ ⊕ ⊕ ⊕ ⊕ ⊕	Severe COPD; n=14 and healthy subjects; n=12	Excluded difficulty with exercise testing, recent exacerbation, long term steroid use and asthma.	How does the cardiorespiratory response compare between 1MSTST and CPET in patients with COPD and controls?	1MSTST CPET	Tests were strongly correlated including the decrease in SpO2 pre and post exercise.	The 1MSTST induced a similar cardiorespiratory response to that of CPET in patients with COPD.
Gloeckl 2020 Germany	Prospective	⊕ ? ? ? ⊕ ⊕ ⊕ ⊕	COPD; n= 26	Excluded severe COPD exacerbations and respiratory failure	Can the 2MWT be used instead of the 6MWT?	2MWT 6MWT	Strong correlation between tests (including minimum SpO2)	2MWT can be used instead of the 6MWT
Viera 2020 Brazil	Prospective	? ? ? ? ⊕ ⊕ ⊕ ⊕	Pulmonary hypertension; n= 20	Excluded long term oxygen therapy, limitations to perform exercise testing	Can IST be used as an alternative to CPET?	IST CEPT	Greater desaturation and higher VO2PEAK in IST vs CPET. Clinically relevant but not statistically significant difference.	IST is a useful tool in the clinical evaluation of patients with pulmonary hypertension.
Labrecque 2020 Canada	Prospective	? ? ? ⊕ ⊕ ⊕ ⊕ ⊕	Interstitial lung disease; n= 15	Excluded significant concomitant respiratory disease and recent ILD exacerbation	Can 1MSTST be validated against 6MWT and CPET?	1MSTST 6MWT CPET	Peak VO2 was lower in 1MSTST but nadir SpO2 similar in all tests. Peak values were achieved in recovery phase of 1MSTST but at end of test in 6MWT and CPET.	1MSTST produces a substantial but submaximal cardiorespiratory response; can be an alternative to 6MWT and CPET for exercise-induced hypoxia.
Rusanov 2007 Israel	Prospective	? ? ? ⊕ ⊕ ⊕ ⊕ ⊕	Interstitial lung disease; n= 51	Excluded cardiac or other pulmonary or infective disease	Can 15SCT be used to assess exercise-induced hypoxemia and functional capacity?	15SCT 6MWT CPET	Desaturations measured by the 15SCT are comparable to those measured by the CPET and the 6MWT	The 15SCT can be used as an alternative to the 6MWT.
Gruet 2020 France	Prospective	? ? ? ⊕ ⊕ ⊕ ⊕ ⊕	Cystic fibrosis; n=25	Excluded patients with major co-morbidities or infection.	Can 1MSTST replace 6MWT or CPET to find maximum exercise capacity (main q) and exertional hypoxia?	1MSTST 6MWT CPET	Strong correlation in oxygen desaturation between 1MSTST and CPET (P<0.001).	1MSTST can be used as an alternative to 6MWT or CPET to detect exertional desaturation.
Morita 2020 Brazil	Prospective	⊕ ⊕ ? ? ? ⊕ ⊕ ⊕	COPD; n= 23	Excluded severe co-morbidities or inability to perform tests.	(1) How do 5STS, 30secSTS and 1MSTST compare with each other? (2) how do they compare with other tests in use?	5STST 30secSTST 1MSTST 6MWT 4MGST ISWT 1-rep max of quads	1MSTST correlated with 6MWT. 5STST and 30secSTST were associated with 4MGST. Bigger changes were found in 1MSTST than 5STST and 30secSTST including a higher haemodynamic demand.	Of the tests evaluated in this study, 1MSTST seems to be the best protocol to evaluate subjects with COPD.
Kohlbrenner 2020 Switzerland	Retrospective	⊕ ⊕ ⊕ ⊕ ? ? ? ⊕	Lung transplant candidates; n= 38	Excluded patients who couldn't perform all 3 tests	Can 1MSTST be used to assess exercise capacity (instead of 6MWT) and knee extensor strength?	1MSTST 6MWT Knee extension strength test	SpO2 fell during 1MSTST and 6MWT (P < 0.001). 1MSTST caused similar HR response but less desaturation (p<0.001).	1MSTST cannot replace 6MWT in lung transplant candidates but is a safe alternative when time and space are limited.
Crook 2017 Switzerland	Prospective	⊕ ⊕ ? ? ? ⊕ ⊕ ⊕	COPD; n= 21 (1MSTST & 5STST) and n= 15 (all 3 tests)	Data extracted from longitudinal studies of COPD patients.	Can 1MSTST be used instead of 6MWT or 5MSTST?	1MSTST 5STST 6MWT	Strong correlation between 1MSTST and 6MWT. Desaturation continued beyond 1st minute after 1MSTST.	1MSTST can be used as an alternative to 6MWT in COPD patients
Azzi 2020 France	Retrospective	? ⊕ ⊕ ? ? ⊕ ⊕ ⊕	Non small cell lung cancer; n= 36	Tests done as part of routine pre-operative assessment.	Can the 3MCRT be used instead of the CPET to predict VO2peak?	3MCRT CPET	Strong correlation between the 3MCRT and CPET with regards to change in SpO2.	No conclusion drawn about use of 3MCRT to assess exertional desaturation

6MWT 6-minute walk test, 2MWT 2-minute walk test, 5STST 5-repetition sit to stand test, 30secSTST 30 s sit to stand test, 1MSTST 1-minute STS test, ISWT incremental shuttle walking test, 1-rep max of quads 1-repetition maximum of quadriceps muscle, CPET cardiopulmonary exercise testing, 15SCT 15-steps climbing test, IST incremental step test, 10MCT 10-minutes cycling test, 4MGST 4-minute gait speed test, 3MCRT 3-minute chair to rise test, PCP pneumocystis pneumonia, IPF idiopathic pulmonary fibrosis (IPF)

*Risk of bias column summarises the following assessments in this order: Risk of bias: patient selection/index test/reference standard/flow and timing, Concerns about applicability: patient selection/index test/reference standard: ⊕ Low risk of bias; ⊕ high risk of bias; ? unclear risk of bias



The remainder of our results section is divided into four sub-sections. First, we set the context for our review with a narrative summary of the use of exertional desaturation tests in general lung disease. Next, we describe a sparse literature (two studies) in which the results of exertional desaturation tests were correlated with clinical outcomes in Covid-19, followed by an equally sparse literature in which exertional desaturation tests were correlated with the type of pneumonia in HIV-positive patients. Finally, we describe and critique a somewhat larger literature on the validation of selected tests for exertional desaturation in various chronic lung diseases.

Use of exertional desaturation tests in general lung disease

Tests of exertion in lung disease have mostly addressed the monitoring over time of *chronic* lung disease and have been oriented to measuring exercise capacity. A helpful narrative review by Lee et al. (which drew on an earlier systematic review by European Respiratory Society and American Thoracic Society [30]) lists, for example, a 30-min walk test, a 4-min walk test, a stair climb power test (10 flights), a more moderate stair climb test (un-standardised but based on the patient's own home stairs), 6-min and 3-min step tests, a 15-step test (step up and down on a 25 cm platform 15 times as fast as possible), Chester step test (an incremental protocol on a 20 cm platform with 2-min phases commencing with 15 steps per minute and increasing by 2 per minute till terminated by dyspnoea or fatigue) and modified Chester step test (starting at 10 steps per minute) [31]. These authors also describe three different sit-to-stand tests: five repetition sit-to-stand (5STS: the patient stands up fully and sits down 5 times as quickly as they can); 1-min sit to stand test (1MSTS: patient stands up fully and sits down as many times as they can in 1 min) and the 30-s and 2-min variants of this [31]. They also review tasks based on activities of daily living such as a semi-standardised 'grocery shelving task' [31].

All the exercises described by Lee et al. in the above review were designed mainly for longitudinal monitoring of the severity of chronic lung disease, and several have been shown to correlate with survival [31]. The tests combine an assessment of lung function with that of general physical fitness and muscle strength—a useful composite measure in patients with (for example) chronic obstructive pulmonary disease. They were not originally designed with assessment of acute breathlessness in mind, but as described below, some have subsequently been evaluated for that purpose.

A systematic review considered the validity of the 1MSTST in measuring exercise capacity in patients with chronic lung disease [32]. The main focus of that review was on (a) whether the test correlated with severity of

lung disease (broadly, it did), the test-retest reliability of the test (it was high) and whether the test score correlated with the gold standard 6-min walk test (it did). They concluded that 'The 1-min STS appears to be a practical, reliable, valid, and responsive alternative for measuring exercise capacity, particularly where space and time are limited' [32]. However, these authors did not look at the 1MSTST in the assessment of exertional desaturation.

The cardiopulmonary exercise test (CPET) has long been used to derive important variables that are known to be good predictors of prognosis in many cardiorespiratory conditions (including chronic obstructive pulmonary disease, interstitial lung disease, pulmonary arterial hypertension, congestive heart failure, cystic fibrosis and chronic thromboembolic pulmonary hypertension) [33]. Several studies have confirmed that peak VO_2 is the preferred method for risk stratification and for the prognostic evaluation of patients with end-stage lung disease such as COPD and cystic fibrosis [33]. However, field tests (such as the 6MWT and 1MSTST) are more commonly used in clinical practice since they do not require specialist lung function facilities [34].

Fox et al. explored the use of oximetry along with step climbing tests in the detailed assessment of pulmonary capacity, using area under the curve of a continuous oximetry reading [35]. Whilst these authors found that oximetry thus measured correlated with severity of disease and survival, this cannot automatically be applied to the remote assessment of the acutely breathless or hypoxic patient. Similarly, several studies of oximetry in the 6MWT showed strong correlation with disease severity and survival in patients with idiopathic pulmonary fibrosis [36], but these findings may not apply to Covid-19.

Exertional tests for measuring desaturation in Covid-19

Our search identified no validation studies of exertional tests for hypoxia in patients with Covid-19. We found two studies which sought to correlate the results of an exertional desaturation test with clinical outcomes in Covid-19.

In a small study of 26 COVID-19 patients assessed prior to discharge from hospital, Fuglebjurg et al. used the 6MWT to assess the degree of exertional hypoxia; symptoms of subjective dyspnoea were noted [37]. Thirteen patients developed exercise-induced hypoxia (defined as $\text{SpO}_2 < 90\%$) during the 6MWT, of whom four had pulmonary embolism (a perfusion defect). Covid-19 patients experienced less hypoxia-related dyspnoea during the 6MWT compared with historical idiopathic pulmonary fibrosis controls (none of whom were documented as having pulmonary embolism). The authors concluded that the 6MWT is a potentially useful tool in the diagnosis of asymptomatic exercise-induced

hypoxia in hospitalised COVID-19 patients prior to discharge. Whilst interesting, the study does not have direct relevance to the question of exertional desaturation tests in a less select Covid-19 population in the acute phase, nor does it tell us anything about the briefer tests currently in use in community settings.

Goodacre et al. conducted a retrospective observational cohort study (a methodologically weak study design) across 70 emergency departments in the UK during the first wave of the Covid-19 pandemic [38]. Eight hundred seventeen patients out of the 22,000 who were assessed had an exertional test recorded on their record. Of these, 30 had an adverse outcome (defined narrowly as requiring organ support in intensive care) and 9 died. Whilst the positive 1.78 (1.25 to 2.53) and negative 0.67 (0.46 to 0.98) likelihood ratios of a 3% or more desaturation just achieved statistical significance, the authors concluded that exertional desaturation was not a significant predictor of adverse outcome when baseline clinical assessment was taken into account ($p = 0.37$). The study specifically did not assess whether patients with exertional desaturation alone would otherwise have fulfilled criteria for hospital admission. It is possible that if adverse outcome had been more broadly defined (e.g. the need to be admitted to hospital, receive supplemental oxygen or in terms of subsequent health-care usage), the test may have proved a useful predictor. It is noteworthy that only 3% of the cohort had an exertional test and were not randomly assigned and the reasons for exertional tests being performed/not performed were not analysed. Additional information would have been gained if all patients with a particular baseline oximeter reading had been tested for exertional desaturation and followed up for adverse outcomes. In short, little can be concluded from this retrospective study of a highly selected sample.

Exertional desaturation as a predictor of acute lung disease

There are some important clinical parallels between pneumocystis pneumonia and the respiratory manifestations of acute Covid-19. Like those with acute Covid-19, patients with pneumocystis pneumonia may be hypoxic (and even cyanotic) on initial presentation, but less severe cases can be normoxaemic initially and become more hypoxic as the disease progresses [39, 40].

In two studies in people with HIV, the value of an exertional desaturation test to discriminate pneumocystis pneumonia from other causes of acute pneumonia was tested. Here, desaturation during the exercise tests was correlated with subsequent results from a bronchoalveolar lavage (and in some cases biopsy) which confirmed or rejected its diagnosis. We describe these studies briefly below.

In a study we rated as high quality, Sauleda and colleagues assessed 45 HIV-positive subjects with pneumonia who were admitted to the emergency department performed pedalling motions in the air for 2 min on the stretcher bed [39]. Oxygen saturations were monitored throughout the test. During exercise, the mean SaO₂ fell in patients with pneumocystis pneumonia from 88 to 84% ($p < 0.01$), whilst it improved slightly in patient with non-pneumocystis pneumonia from 91 to 93% ($p < 0.05$).

In a similar study (which we rated as lower methodological quality) from 1988 of 39 patients with pneumocystis pneumonia (all HIV-positive young men), exertional desaturation was demonstrated in most of them (including many who had normal saturation at rest) using a 10-min cycling test, whereas patients who presented with other acute lung conditions including bacterial pneumonia, tuberculosis and pulmonary candidiasis were significantly less likely to show exertional desaturation [40].

Validation of tests for exertional hypoxia in conditions other than Covid-19

We could not find studies comparing different modalities of testing for exertional desaturation in Covid-19. However, we found 11 studies which described attempts to compare and establish non-inferiority of different exercise tests to assess exertional hypoxia in various chronic lung conditions (including chronic obstructive pulmonary disease, interstitial lung disease, advanced lung disease requiring a lung transplant and pulmonary hypertension). Whilst caution must be exercised in applying data from studies in patients with a variety of chronic lung diseases, with differential impacts upon ventilation and diffusion, the data regarding evaluation of tests to detect exertional hypoxia in a variety of conditions and settings is relevant.

In these studies, measured physiological variables from various exercise tests (such as the 1MSTST, 5STST, 2MWT, IST) were compared to those measured with a 6MWT and/or CPET. Variables such as SpO₂, heart rate and respiratory rate were measured during (and occasionally after) the various exercise tests. Five studies looked at the 1MSTST and showed that it correlated well with the 6MWT and/or the CPET. We describe below the studies in this group which we scored as high quality.

In the study we rated as highest quality, Briand et al. compared the nadir SpO₂ measured by oximetry on the 6MWT and 1MSTST (performed on the same day) in a clinic population of 107 patients with chronic interstitial lung disease [29]. There was high correlation between the two tests ($r = 0.9$; $p < 0.0001$). The authors also found that the correlation between the tests in terms of desaturation appeared to hold at lower levels of SpO₂. No

adverse events were described in this study. Table 2 shows the distribution of findings in Briand et al's study.

Using the data in Table 2, and taking the 6MWT as the gold standard, the 1MSTST appears to have a sensitivity of 88% [95% CI impossible to calculate because numbers too small], a specificity of 81% [95% CI 71–91%], a positive predictive value of 79% [95% CI 68–90%] and a negative predictive value of 89% [95% CI impossible to calculate because numbers too small] [29].

Gephine et al. (2020) compared the 1MSTST with the CPET in 14 people with severe COPD and 12 healthy participants [41]. In the COPD group, the fall in SpO₂ from pre-exercise to peak exercise was similar in the COPD groups with both the 1MSTST (mean –5% SD 4%) and the CPET (mean –6%, SD 6%); differences were not statistically significant. In the healthy control group, there was very little fall in SpO₂ with either the 1MSTST (–1%, SD 2%) or the CPET (–1%, SD 1%).

During the 1STST, a ≥4% SpO₂ fall was seen in seven people with COPD, amongst which nadir SpO₂ values were reached during the recovery period in five patients. For these patients, the lowest value of saturation was reached 33 ± 12 s after the end of the exercise. In comparison, 10 people with COPD exhibited an SpO₂ fall of similar magnitude during the CPET. In five of them, the SpO₂ values occurred during the recovery period, 51 ± 16 s after the end of exercise.

The authors concluded that (i) the 1STST elicited a similar peak physiological response to the CPET; (ii) people with COPD showed a nadir SpO₂ during the recovery period of the 1STST, therefore highlighting the relevance of monitoring this crucial phase of exercise. This study did not, however, report a formal validation exercise of the kind shown in Table 2 (perhaps because numbers were small).

Gloeki et al. found fairly high correlation ($r = 0.81$) between desaturation levels on the 6MWT and those on a shorter 2MWT in a small sample of 26 patients with COPD [42]. Oxygen saturation fell from a mean of 93.8% (95% confidence interval 92.8–94.7) to 83.2% (80.8–85.5) on the 2MWT compared with 93.3% (92.4–94.3) to 82.0% (79.8–84.3) on the 6MWT. Differences in nadir and percentage drop were not statistically

significant between the two tests. The authors concluded that 'the decline in oxygen saturation [is] very similar during the 2MWT and the 6MWT [and] that the short duration of a 2MWT is sufficient to induce a similar oxygen desaturation under room air conditions in patients with severe COPD as the 6MWT' (page 260) [42].

Rusanov et al. [43] compared the 15SCT against the 6MWT test in 51 patients with idiopathic pulmonary fibrosis (IPF), along with a CPET. SpO₂ fell from 95% (SD 3) to 86% (SD 7) in the 15SCT and from 94% (SD 3) to 86% (SD 8) in the 6MWT. The nadir of hypoxaemia was very similar on the CPET (88%, SD 6) and showed high correlation with the 15SCT ($r = 0.85$; $p < 0.0001$). The fall in SpO₂ and nadir SpO₂ was also highly correlated between 15SCT and 6MWT. The authors concluded that the desaturation measured by the 15SCT test is comparable to the desaturation measured by the CPET and the 6MWT test, making the 15SCT a reliable tool for monitoring disease progression in IPF and for evaluating the need for oxygen supplementation. Another paper by the same authors reports duplicate findings [44]. Again, however, this study only measured *correlation*, without *validation* against the 6MWT and the CPET.

Another study done in patients with a perfusion defect was by Vieira et al., who looked at the usefulness of the IST in 20 patients with pulmonary hypertension [45]. They found a high correlation between desaturation levels on the IST and CPET in patients walking on a treadmill: in the CPET, SpO₂ fell from 96% (SD3) to 92% (SD 6); in the IST, it fell from 96% (SD 3) to 89% (SD 8)—a difference that was not statistically significant but which may have been clinically significant. The authors concluded that the IST is a useful tool in the evaluation of patients with pre-capillary pulmonary hypertension.

In a further small study of 15 participants with pulmonary fibrosis, Labrecque et al. compared the 1MSTST (done twice to assess reproducibility), 6MWT and CPET [46]. The main aim of the study was to validate the 1MSTST not as a test of exertional desaturation but as a test of exertion which consistently produces a cardiorespiratory stress. 1MSTST, 6MWT and CPET all produced a similar fall in SpO₂ (10% SD 5, 12% SD 4 and 8% SD 4 respectively). There was no significant difference between the nadir SpO₂ reached (88% SD4, 85% SD 4 and 87% SD4 respectively), though these differences may have reflected clinically significant differences in desaturation. Perhaps the most important observation from the Labrecque study for this review was that the 1MSTST (an intensive burst of exercise over 1 min) was shown to be considerably more strenuous from a cardiorespiratory perspective than the 6MWT and the CPET (which was a longer but less intensive period of exercise lasting 8–9 min). As the authors noted, 'Coping with such a surge in physiological demand during the 1STS

Table 2 Validation study comparing 1MSTST with 6MWT in chronic lung disease

Change in SpO ₂ ≥ 4% in 1MSTST	Change in SpO ₂ ≥ 4% in the 6MWT (gold standard)		
	Yes	No	Total
Yes	42	11	53
No	6	48	54
Total	48	59	107

Data from Briand et al [29]

[test] was demanding for people with ILD [interstitial lung disease]' (page 15).

We placed less weight on the final five studies—as the methodological quality was uncertain, or because methodological quality was considered to be poor on our risk of bias and applicability tool, though the findings from all these additional studies were similar to the above ones. These studies were Gruet et al. (a prospective correlation study from France in 25 patients with cystic fibrosis, which concluded that the 1MSTST may be used as an alternative to the 6MWT and CPET for assessing exertional desaturation) [47]; Morita et al. (a study in 23 patients with COPD which showed good correlation between various exercise desaturation tests) [48]; Kohlbrenner et al. (a retrospective study from Switzerland in 38 lung transplant candidates which concluded that the 1MSTST is a safe alternative to 6MWT in such patients despite lower desaturation nadirs) [49]; Crook et al. (a prospective study in 21 COPD patients which concluded that 1MSTST is a safe and accurate alternative to 6MWT); and Azzi et al. (a retrospective study in 36 lung cancer patients whose main aim was to compare a 3-min chair-to-rise test, 3MCTRT, with the CPET in terms of maximal exercise capacity but which also found high correlation with the level of oxygen desaturation achieved) [50].

The test that is most used in acute practice is the 40-step walk test—the patient is asked to walk 40 steps on the flat and oximetry repeated. We found no research studies on this at all.

Discussion

Summary of key findings

This rapid review has produced several key findings relevant to the assessment of exertional desaturation in patients with suspected Covid-19. First, we identified no published studies which had compared the performance of different brief exercise tests in a cohort of patients with (or suspected of having) Covid-19.

Second, in all but one of 11 studies presented as 'validation studies' of brief exercise tests for assessing exertional desaturation in other lung diseases, methodological quality was poor or impossible to fully assess. Furthermore, whilst the authors of all 11 studies had *correlated* the level of exertional desaturation on a range of exercise tests with one or both accepted gold standard tests (6MWT and CPET) in various acute or chronic lung conditions, none of these studies had been designed as a formal diagnostic test validation study. Rather, the focus had been on comparing the average SpO₂ in the same group of patients who underwent both tests and showing no statistical difference between the two measurements. It is reassuring that there was high correlation between the 6MWT or CPET and the shorter 1MSTST, but we would expect the results of

these tests to be *correlated* (as, for example, a person's height will be correlated with their weight). Correlation alone, however, does not *validate* the test [51].

Only one of the 11 studies (Briand et al. [29]) contained sufficient raw data for us to calculate the sensitivity (88%), specificity (81%), and positive and negative predictive value (79% and 89% respectively) of the new test in relation to the gold standard (6MWT), and the authors had not calculated these values themselves. This study suggests that the accuracy of the 1MSTST is far from perfect: 16 patients in every 100 will be misclassified [29]—a finding which underscores the need to interpret them in the context of a full clinical assessment.

The third key finding of this review is that the 1MSTST produced a high cardiorespiratory stress (and indeed, can be exhausting for those who are fully fit), and patients with lung disease continued to show a further drop in SpO₂ levels even after the test had been completed [46, 52]. In the context of assessing patients remotely who may have acute Covid-19 (i.e. where the clinician is geographically distant from the patient and a full clinical assessment is impossible), the 1MSTST may therefore be risky in unsupervised settings.

Finally, we identified a significant gap in the literature, namely the lack of validation studies (or indeed any relevant research) on the less strenuous 40-step test, which features in local and national guidance for Covid-19 [18] and is in widespread use [19]. This test is likely to be less risky than the 1MSTST but we can currently say nothing about its accuracy. Intuitively, we would expect the specificity to be high, but the sensitivity to be low due to the lower level of physiological exertion required. This is, however, speculative.

Comparison with previous literature

As the studies listed in Table 1 illustrate, there is a good (though by no means perfect) evidence base in the use of exertional tests for a variety of different pulmonary disease (including airway and interstitial lung diseases). However, the main focus in most such studies is on performance as a surrogate of VO₂ (for example, noting the number of sit-stands done), rather than oxygen saturation per se. Given Covid-19's physiological affinity with other perfusion-defect lung conditions such as interstitial lung disease and pneumocystis pneumonia, it follows that the assessment of exertional desaturation would be useful.

As none of the exertional tests have yet been validated to show that assessing exertional desaturation is better than routine clinical assessment at demonstrating increased risk, their use remains pragmatic. The resulting evidence gap in this regard will most likely be addressed in the near future given the continuing incidence of Covid-19.

Implications for practice

Our review findings, combined with what is known about the pathophysiology of acute Covid-19 [8, 12, 13, 37], suggest a conservative and risk adverse approach to exercise desaturation testing, especially in the home or remote assessment setting. The levels of desaturation observed with brief exercise tests in patients with chronic lung disease [29, 41–43, 45, 46, 48, 49, 52] may be even more marked in those with acute Covid-19. For this reason, we suggest that even a small desaturation on exercise should alert the clinician of the need for further evaluation (such as a face to face consultation) and a drop of 3% should be cause for prompt assessment, regardless of the amount of exercise needed to produce it. There is not yet good evidence that exertional desaturation should prompt a change in treatment (e.g. earlier use of steroids), but this could be the subject of further research.

Current evidence therefore supports the recommendation to undertake a thorough clinical assessment and evaluation including a history, assessment of risk factors, pulse, respiratory rate and oxygen saturations. There is a physiological basis for adding graded assessment into the process of clinical evaluation as a positive test would raise significance.

The '40 steps around the room test' or its alternative '40 steps on the spot test' (in a patient able to stand safely unaided and whose resting saturation is 96% or above) is the lowest level of exertion of any test either in the literature or in clinical practice. An alternative would be the '40 steps on the spot test' with the patient standing in front of a chair in case of needing to sit down. This is appropriate for the home environment (no clinician on hand to guide or resuscitate) given the high-risk group. This test will likely have a low sensitivity but using basic physiological principles, if positive it is likely to be highly specific and warrant urgent assessment.

If the 1MSTS is used, it should be followed by monitoring for at least one minute to observe for desaturation. Indeed, a short test means that it can take time to utilise the free oxygen in the blood stream and therefore witness the impact of desaturation. Such a test should only be done with another person (preferably a clinical staff member) in attendance or nearby. The patient should be kept on continuous oximetry monitoring for at least 1 min following the test; the 1MSTST should never be attempted with a patient home alone.

Our recommendation with regards to oxygen saturation monitoring would therefore be to measure baseline saturations, followed by the 40-step test, followed by the 1 MSTST (if judged safe and there is clinical supervision). If oxygen saturations show a decrease at any of these stages, then the next step should not be attempted.

Oximeters in smartphone apps are unreliable [53], so an approved and tested medical-grade oximeter should be supplied to (or purchased by) the patient. This is happening in some UK localities as part of 'virtual ward' arrangements.

Risk to the patient from exercise tests should be considered. Patients should be advised to terminate promptly if they develop any adverse symptoms (severe breathlessness, chest pain, dizziness) [33]. Tests involving climbing a flight of stairs should be avoided, since a staircase is a dangerous place to collapse. A formal 6MWT is not necessary in the home environment but may be useful in following up patients in a clinic or inpatient setting. When doing these more strenuous exertion tests, carefully observe the patient and also make a clinical judgement based on severe fatigue and tachypnoea.

The 40-step test (which is in widespread clinical use but has not been validated) and the 1MSTST are considered the least demanding tests and therefore may be the most appropriate for recommending to patients at home (perhaps modified so the patient is *not* advised to complete 'as many as you can' in the 1MSTST). We hypothesise that they are specific but not sensitive (that is, a positive test is serious cause for concern but a negative test should not necessarily reassure), though there is currently no hard data on this.

Conclusion

A subset of patients with Covid-19 present with marked hypoxia. Earlier identification of these patients to allow closer monitoring and treatment may improve outcome. Exertional desaturation tests are hypothesised to identify these patients before hypoxia at rest occurs. Whilst the evidence base provides some limited support for the equivalence of 1MSTST to the 6MWT or CPET in identifying exertional hypoxia and desaturation, and some comparisons can be drawn to exertional desaturation in PCP, we cannot draw conclusions on the use of these tests within the context of acute Covid-19.

More research is needed on the prognostic value and clinical utility of exertional desaturation tests in all settings (GP, emergency department, ambulance), in the context of Covid-19. Furthermore, an understanding of how best to ask the patient about breathlessness on exertion, and how this correlates with exertion oximetry, could also help in the assessment of hypoxia in Covid-19.

Finally, it is important to remember that a pulse oximeter reading, whether at rest or in an exertional test, does not replace thorough clinical assessment. A normal test does not necessarily mean the patient can be reassured that all is fine. Nevertheless, exertional hypoxia of 3% or more is a significant finding and always warrants further assessment and investigation.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13643-021-01620-w>.

Additional file 1.

Additional file 2.

Additional file 3.

Acknowledgements

We thank health informaticians and librarians: Helen F Williams, Jon Brassey and Nia Roberts, and 3 anonymous reviewers for helpful comments on a previous version of this paper.

Authors' contributions

This review built on a rapid review that had been undertaken early in the pandemic, to which TG, BJ, MK and MI-K contributed equally with librarian support from Helen Williams and Jon Brassey. To extend that review, AK undertook systematic searches with librarian support from Nia Roberts. MK, MI—K and BJ contributed additional sources from their clinical and academic knowledge. AK undertook data extraction on all papers and prepared first drafts of the detailed supplementary tables, and wrote the first draft of the paper. All other authors checked data extraction on some papers. TG developed the first draft of the summary table, which AK refined. All authors provided feedback on the draft paper and approved the final version.

Funding

AK is funded by an NIHR Academic Clinical Fellowship. TG's research is funded from the following sources: National Institute for Health Research (BRC-1215-20008), ESRC (ES/V010069/1), and Wellcome Trust (WT104830MA).

Availability of data and materials

All sources cited in this review are publicly available.

Declarations

Ethics approval and consent to participate

Not applicable (desk research).

Consent for publication

Not applicable (desk research).

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Primary Health Care Sciences, University of Oxford, Oxford, UK. ²Division of Experimental Medicine, University of California, Berkeley, CA, USA. ³West Hertfordshire Hospitals NHS Trust, Vicarage Rd, Watford, Hertfordshire WD18 0HB, UK. ⁴Hampshire Hospitals NHS Foundation Trust, Basingstoke, UK.

Received: 9 November 2020 Accepted: 23 February 2021

Published online: 16 March 2021

References

- To KK-W, Tsang OT-Y, Leung W-S, Tam AR, Wu T-C, Lung DC, Yip CC-Y, Cai J-P, Chan JM-C, Chik TS-H. Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. *Lancet Infect Dis*. 2020;20(5):565–74.
- Pandit R, Vaity C, Mulakavulipil B, Matthew A, Sabnis K, Joshi S. Unmasking Hypoxia in COVID 19 - Six Minute Walk Test. *J Assoc Physicians India*. 2020; 50–1.
- Knight SR, Ho A, Pius R, Buchan I, Carson G, Drake TM, Dunning J, Fairfield CJ, Gamble C, Green CA, et al. Risk stratification of patients admitted to hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: development and validation of the 4C Mortality Score. *BMJ*. 2020; 370(6490):m3339.
- Xie J, Covassin N, Fan Z, Singh P, Gao W, Li G, Kara T, Somers VK. Association between hypoxemia and mortality in patients with COVID-19. In: *Mayo Clinic Proceedings*: 2020: Elsevier; 2020.
- National Institute for Health and Clinical Excellence: COVID-19 rapid guideline: managing suspected or confirmed pneumonia in adults in the community. London: NICE; 2020. Accessed 30th October 2020 at <https://www.nice.org.uk/guidance/ng165>.
- Mantha S, Tripuraneni SL, Roizen MF, Fleisher LA: Proposed Modifications in the 6-Minute Walk Test for Potential Application in Patients With Mild COVID-19: A Step to Optimize Triage Guidelines. *Anesthesia and Analgesia* 2020, e-pub ahead of print
- Gupta RK, Harrison EM, Ho A, Docherty AB, Knight SR, van Smeden M et al. Development and validation of the ISARIC 4C Deterioration model for adults hospitalised with COVID-19: a prospective cohort study. *Lancet Respir Med*. 2021. e-pub ahead of print.
- Shenoy N, Luchtel R, Gulani P. Considerations for target oxygen saturation in COVID-19 patients: are we under-shooting? *BMC Med*. 2020;18(1):1–6.
- O'Driscoll B, Howard L, Earis J, Mak V. BTS guideline for oxygen use in adults in healthcare and emergency settings. *Thorax*. 2017;72(Suppl 1):ii1–ii90.
- Berezin L, Zhabokritsky A, Andany N, Chan AK, Gershon A, Lam PW, Leis JA, MacPhee S, Mubareka S, Simor AE. The diagnostic accuracy of subjective dyspnea in detecting hypoxemia among outpatients with COVID-19. In: medRxiv; 2020. <https://doi.org/10.1101/2020.08.10.20172262>.
- Dhont S, Derom E, Van Braeckel E, Depuydt P, Lambrecht BN. The pathophysiology of 'happy/hypoxemia in COVID-19. *Respir Res*. 2020;21(1):1–9.
- Couzin-Frankel J. The mystery of the pandemic's 'happy hypoxia. *Science*. 2020;368(6490):455–6.
- Tobin MJ. Basing respiratory management of COVID-19 on physiological principles. *Am J Respir Crit Care Med*. 2020;201:1319–20.
- Richardson S, Hirsch JS, Narasimhan M, Crawford JM, McGinn T, Davidson KW, Barnaby DP, Becker LB, Chelico JD, Cohen SL. Presenting characteristics, comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the New York City area. *Jama*. 2020;323:2052–9.
- Guan W-J, Ni Z-Y, Hu Y, Liang W-H, Ou C-Q, He J-X, Liu L, Shan H, Lei C-L, Hui DS. Clinical characteristics of coronavirus disease 2019 in China. *New England J Med*. 2020;382:1708–20.
- Goyal D, Donnelly H, Kussner A, Neil J, Bhatti S, Mansab F: Oxygen and mortality in COVID-19 pneumonia: a comparative analysis of supplemental oxygen policies and health outcomes across 26 countries. Available at SSRN 3633151 2020.
- NHS England and NHS Improvement: Pulse oximetry to detect early deterioration of patients with COVID-19 in primary and community care settings. London: NHSE/I. Accessed 17th October 2020 at <https://www.england.nhs.uk/coronavirus/publication/pulse-oximetry-to-detect-early-deterioration-of-patients-with-covid-19-in-primary-and-community-care-settings/>; 2020 (updated 7th October).
- NHS England: Reference guide for emergency medicine. London: NHS England. Accessed 17th October 2020 at <https://www.england.nhs.uk/coronavirus/wp-content/uploads/sites/52/2020/03/CO261-specialty-guide-emergency-medicine-v5-22-April.pdf>; 2020 (20th April).
- NHS England: Principles of safe video consulting in general practice during COVID-19. London: NHS England. Accessed 27th October 2020 at <https://www.guidelines.co.uk/public-health/nhs-video-consulting-guideline/455510/article>; 2020.
- Greenhalgh T, Thompson P, Wieringa S, Neves AL, Husain L, Dunlop M, Rushforth A, Nunan D, de Lusignan S, Delaney B. What items should be included in an early warning score for remote assessment of suspected COVID-19? qualitative and Delphi study. *BMJ Open*. 2020;10(11):e042626.
- Noormohammadpour P, Abolhasani M. Besides other signs, can a 6-min walk test be applied as a criterion for going to the hospital with a diagnosis of COVID-19? *Adv J Emerg Med*. 2020;4(2s):e42.
- British Thoracic Society: BTS Guideline for oxygen use in healthcare and emergency settings. London: British Thoracic Society; 2017 (updated 2019).
- Gupta R, Ruppel GL, Espiritu JRD. Exercise-induced oxygen desaturation during the 6-minute walk test. *Med Sci*. 2020;8(1):8.
- Garrity C GG, Kamel C, King VJ, Nussbaumer-Streit B, Stevens A, Hamel C, Affengruber L: Cochrane rapid reviews. Interim Guidance from the Cochrane Rapid Reviews Methods Group. March 2020: Cochrane 2020. Accessed 15th April at https://methods.cochrane.org/rapidreviews/sites/methods.cochrane.org/rapidreviews/files/public/uploads/cochrane_rr_-_guidance-23mar2020-v1.pdf.

25. Salameh J-P, Bossuyt PM, McGrath TA, Thombs BD, Hyde CJ, Macaskill P, Deeks JJ, Leeftang M, Korevaar DA, Whiting P. Preferred reporting items for systematic review and meta-analysis of diagnostic test accuracy studies (PRISMA-DTA): explanation, elaboration, and checklist. *BMJ*. 2020;370:m2632.
26. Kalin A, Javid B, Knight M, Inada-Kim M, Greenhalgh T. What is the efficacy and safety of rapid exercise tests for exertional desaturation in Covid-19: a rapid review protocol. *Medrxiv*. 2020;223453.
27. Lasserson TJ, Thomas J, Higgins JPT: Chapter 1: starting a review. In: *Cochrane Handbook* (online). edn. Edited by Higgins JPT, Thomas J. Oxford: Cochrane Collaboration. Accessed 31st October 2020 at <https://training.cochrane.org/handbook/current/chapter-01#section-1-1>; 2020: Section 1.1.
28. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA, Bossuyt PM. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011; 155(8):529–36.
29. Briand J, Behal H, Chenivresse C, Wémeau-Stervinou L, Wallaert B. The 1-minute sit-to-stand test to detect exercise-induced oxygen desaturation in patients with interstitial lung disease. *Ther Adv Respir Dis*. 2018;12: 1753466618793028.
30. Singh SJ, Puhan MA, Andrianopoulos V, Hernandez NA, Mitchell KE, Hill CJ, Lee AL, Camillo CA, Troosters T, Spruit MA. An official systematic review of the European Respiratory Society/American Thoracic Society: measurement properties of field walking tests in chronic respiratory disease. *Eur Respir J*. 2014. <https://doi.org/10.1183/09031936.00150414>.
31. Lee A, Harrison S, Beauchamp MK, Janaudis-Ferreira T, Brooks D. Alternative field exercise tests for people with respiratory conditions. *Curr Phys Med Rehabil Rep*. 2015;3(3):232–41.
32. Bohannon RW, Crouch R. 1-minute sit-to-stand test: systematic review of procedures, performance and clinimetric properties. *J Cardiopulmonary Rehabil Prevention*. 2019;39(1):2–8.
33. Laviolette L, Laveneziana P. Exercise testing in the prognostic evaluation of patients with lung and heart diseases. *Clinical Exercise Testing (ERS Monograph)* Sheffield, European Respiratory Society. 2018:222–34.
34. Myers J, Kokkinos P, Chan K, Dandekar E, Yilmaz B, Nagare A, Faselis C, Soofi M. Cardiorespiratory fitness and reclassification of risk for incidence of heart failure: the veterans exercise testing study. *Circ Heart Fail*. 2017;10(6): e003780.
35. Fox BD, Sheffy N, Vainshelboim B, Fuks L, Kramer MR. Step oximetry test: a validation study. *BMJ Open Respir Res*. 2018;5(1):e000320.
36. Lancaster LH. Utility of the six-minute walk test in patients with idiopathic pulmonary fibrosis. *Multidiscip Respir Med*. 2018;13(1):1–7.
37. Fuglebjerg NJU, Jensen TO, Hoyer N, Rysør CK, Madsen BL, Harboe ZB. Silent hypoxia in patients with SARS CoV-2 infection before hospital discharge. *Int J Infect Dis*. 2020;99:100–1.
38. Goodacre S, Thomas B, Lee E, Sutton L, Loban A, Waterhouse S, Simmonds R, Biggs K, Marincowitz C, Schutter J, Connelly S. Post-exertion oxygen saturation as a prognostic factor for adverse outcome in patients attending the emergency department with suspected COVID-19: a substudy of the PRIEST observational cohort study. *Emerg Med J*. 2021;38(2):88–93.
39. Sauleda J, Gea J, Aran X, Aguar MC, Orozco-Levi M, Broquetas JM. Simplified exercise test for the initial differential diagnosis of *Pneumocystis carinii* pneumonia in HIV antibody positive patients. *Thorax*. 1994;49(2):112–4.
40. Smith D, Wyatt J, McLuckie A, Gazzard B. Severe exercise hypoxaemia with normal or near normal X-rays: a feature of *Pneumocystis carinii* infection. *Lancet*. 1988;332(8619):1049–51.
41. Gephine S, Bergeron S, Tremblay-Labrecque P-F, Mucci P, Saey D, Maltais F. Cardiorespiratory response during the 1-min sit-to-stand test in chronic obstructive pulmonary disease. *Med Sci Sports Exerc*. 2020;52:1441–8.
42. Gloeckl R, Teschler S, Jarosch I, Christle JW, Hitzl W, Kenn K. Comparison of two- and six-minute walk tests in detecting oxygen desaturation in patients with severe chronic obstructive pulmonary disease—a randomized crossover trial. *Chron Respir Dis*. 2016;13(3):256–63.
43. Rusanov V, Shitrit D, Fox B, Amital A, Peled N, Kramer MR. Use of the 15-steps climbing exercise oximetry test in patients with idiopathic pulmonary fibrosis. *Respir Med*. 2008;102(7):1080–8.
44. Shitrit D, Rusanov V, Peled N, Amital A, Fuks L, Kramer MR. The 15-step oximetry test: a reliable tool to identify candidates for lung transplantation among patients with idiopathic pulmonary fibrosis. *J Heart Lung Transplant*. 2009;28(4):328–33.
45. Vieira E, Ota-Arakaki J, Dal Corso S, Ivanaga I, Fonseca A, Oliveira R, Rodrigues-Júnior J, Ferreira E, Nery LE, Ramos R. Incremental step test in patients with pulmonary hypertension. *Respir Physiol Neurobiol*. 2020;271: 103307.
46. Labrecque P, Harvey J, Nadreau É, Maltais F, Dion G, Saey D. Validation and cardiorespiratory response of the 1-min sit-to-stand test in interstitial lung disease. *Med Sci Sports Exerc*. 2020;52:2508–14.
47. Gruet M, Peyré-Tartaruga LA, Mely L, Vallier JM. The 1-minute sit-to-stand test in adults with cystic fibrosis: correlations with cardiopulmonary exercise test, 6-minute walk test, and quadriceps strength. *Respir Care*. 2016;61(12): 1620–8.
48. Morita AA, Bisca GW, Machado FV, Hernandez NA, Pitta F, Probst VS. Best protocol for the sit-to-stand test in subjects with COPD. *Respir Care*. 2018; 63(8):1040–9.
49. Kohlbrenner D, Benden C, Radtke T. The 1-minute sit-to-stand test in lung transplant candidates: an alternative to the 6-minute walk test. *Respir Care*. 2020;65(4):437–43.
50. Azzi M, Debeaumont D, Bonnevie T, Aguilaniu B, Cerasuolo D, Boujibar F, Cuvelier A, Gravier FE. Evaluation of the 3-minute chair rise test as part of preoperative evaluation for patients with non-small cell lung cancer. *Thorac Cancer*. 2020;11(9):2431–9.
51. Cleophas TJ, Droogendijk J, van Ouwkerk BM. Validating diagnostic tests, correct and incorrect methods, new developments. *Curr Clin Pharmacol*. 2008;3(2):70–6.
52. Crook S, Büsching G, Schultz K, Lehbert N, Jelusic D, Keusch S, Wittmann M, Schuler M, Radtke T, Frey M. A multicentre validation of the 1-min sit-to-stand test in patients with COPD. *Eur Respir J*. 2017;49(3):1601871.
53. Tarassenko L, Greenhalgh T. Should smartphone apps be used as oximeters? Oxford: COVID-19 Evidence Service; 2020.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://www.biomedcentral.com/submissions)

