



Physical activity measurements in individuals with interstitial lung disease: a systematic review and meta-analysis

Masahiro Iwakura ¹, Atsuyoshi Kawagoshi¹, Akira Tamaki², Yutaro Oki ³, Yohei Oshima⁴ and Martijn A Spruit ^{5,6}

¹Department of Rehabilitation, Akita City Hospital, Akita City, Japan. ²School of Rehabilitation, Hyogo Medical University, Hyogo, Japan. ³Department of Public Health, Kobe University Graduate School of Health Sciences, Hyogo, Japan. ⁴Rehabilitation Unit, Kyoto University Hospital, Kyoto, Japan. ⁵Department of Research and Development, CIRO, Horn, The Netherlands. ⁶Department of Respiratory Medicine, Maastricht University Medical Centre (MUMC+) NUTRIM School of Nutrition and Translational Research in Metabolism, Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, The Netherlands.

Corresponding author: Masahiro Iwakura (masa.iwaku.91@gmail.com)



Shareable abstract (@ERSpublications)

Reporting quality of physical activity measurement in ILD can be low due to the severe heterogeneity of the measurement methods and metrics used. Standardisation of measurement and improvement in reporting quality is essential to compare studies. <https://bit.ly/3GRyq3p>

Cite this article as: Iwakura M, Kawagoshi A, Tamaki A, *et al.* Physical activity measurements in individuals with interstitial lung disease: a systematic review and meta-analysis. *Eur Respir Rev* 2023; 32: 220165 [DOI: 10.1183/16000617.0165-2022].

Copyright ©The authors 2023

This version is distributed under the terms of the Creative Commons Attribution Non-Commercial Licence 4.0. For commercial reproduction rights and permissions contact permissions@ersnet.org

Received: 22 Aug 2022
Accepted: 11 April 2023

Abstract

Background Physical activity (PA) measurements are becoming common in interstitial lung disease (ILD); however, standardisation has not been achieved. We aimed to systematically review PA measurement methods, present PA levels and provide practical recommendations on PA measurement in ILD.

Methods We searched four databases up to November 2022 for studies assessing PA in ILD. We collected information about the studies and participants, the methods used to measure PA, and the PA metrics. Studies were scored using 12 items regarding PA measurements to evaluate the reporting quality of activity monitor use.

Results In 40 of the included studies, PA was measured using various devices or questionnaires with numerous metrics. Of the 33 studies that utilised activity monitors, a median of five out of 12 items were not reported, with the definition of nonwear time being the most frequently omitted. The meta-analyses showed that the pooled means (95% CI) of steps, time spent in moderate to vigorous PA, total energy expenditure and sedentary time were 5215 (4640–5791) steps·day⁻¹, 82 (58–106) min·day⁻¹, 2130 (1847–2412) kcal·day⁻¹ and 605 (323–887) min·day⁻¹, respectively, with considerable heterogeneity.

Conclusion The use of activity monitors and questionnaires in ILD lacks consistency. Improvement is required in the reporting quality of PA measurement methods using activity monitors.

Introduction

Physical activity (PA) means any bodily movement produced by skeletal muscles resulting in energy expenditure [1]. PA is a complex behaviour described according to the type of PA, movement intensity, movement duration or a combination thereof. The PA guideline for Americans recommends a minimum of 150 min·week⁻¹ of at least moderate-intensity PA to gain health benefits across the adult population, including adults with chronic diseases [2]. The benefits of regular PA include reducing the risk of all-cause and cardiovascular disease mortality, cardiovascular disease, hypertension, type 2 diabetes and other chronic diseases [2].

Participation in regular PA is also crucial in patients with interstitial lung disease (ILD). PA is reduced in patients with ILD compared to healthy controls [3, 4]. Greater dyspnoea and exercise intolerance are associated with lower PA [5]. Reduced PA is one of the strong risk factors for hospitalisation and all-cause mortality in patients with ILD [6, 7]. Therefore, the number of studies regarding PA has been increasing.



Currently, there is no systematic review of PA measurements in patients with ILD. Some researchers have used questionnaires [4, 6–9], while others used activity monitors [3–5, 10]. Questionnaires have multiple limitations, including recall bias, missing data and less precision [11]. Thus, activity monitors containing accelerometers are preferred. Furthermore, different activity monitors and collecting and processing data methods were used in previous studies. Inaccurate assessment of PA can adversely impact the advancement of PA research in ILD. Therefore, understanding how activity monitors or questionnaires have been used in ILD is crucial.

This systematic review, therefore, aimed 1) to explore the types of activity monitors or questionnaires used for collecting PA data in patients with ILD, 2) to evaluate activity monitor-based or questionnaire-based metrics used for assessing PA, 3) to examine the quality of reporting on data collection and processing using activity monitors, 4) to describe PA levels using each metrics, and 5) to provide practical recommendations on how to measure PA and sedentary time (ST).

Methods

This review was conducted following the Preferred Reporting Items for Systematic Reviews and Meta-Analysis statement (supplementary material 1). The review protocol was registered on the International Prospective Register of Systematic Reviews (CRD42021264114).

Eligibility criteria

Study designs

Observational and interventional studies were included. Observational studies were cross-sectional studies, cohort studies or case-control studies. Interventional studies were studies that investigated the effects of interventions. Case series, case reports and grey literature (*e.g.* conference abstracts) were excluded.

Participants

We included studies examining adults with ILD of any origin, diagnosed according to investigator definitions. Participants with exacerbation histories in the preceding 4 weeks [12] were excluded to minimise the influence of exacerbations. There were no restrictions by a history of pulmonary rehabilitation because it is a standard treatment for ILD [13].

PA measurements

We included studies that used activity monitors or questionnaires to measure PA. Additionally, studies had to report on at least the characteristics of patients with ILD separately.

Setting

There were no restrictions on the type of setting.

Language

We included studies reported in English.

Information sources and search strategy

We searched PubMed, CENTRAL, PEDro and OTseeker from the inception of the databases up to 28 November 2022. Search strategies were developed using medical subject headings and text words related to ILD and PA. No study design, date or language limits was imposed on the search. The full search strategy is presented in supplementary material 1.

Selection process and data collection

Two review pairs (M.I. and A.K.; Y.O. and Y.O.) independently performed the first screening (titles and abstracts) and the second screening (full text). Reviewers independently extracted data from each study that met inclusion criteria. Reviewers resolved disagreements by discussion and an arbitrator (A.T. or M.A.S.) adjudicated unresolved disputes.

Data items and outcomes

We extracted the following data:

- patient characteristics – age, sex, forced vital capacity (FVC), diffusing capacity of the lungs for carbon monoxide (D_{LCO}), modified Medical Research Council (mMRC) dyspnoea score [14], 6-min walk distance (6MWD) and use of long-term oxygen therapy (LTOT);
- types of activity monitors or questionnaires – brand, model, sensor type and sensor location;
- PA data collection – period of wear (requested days wear time, weekend/weekday wear requirements and overnight wear) and number of hours wear for a valid day;

- PA data management – valid days requirement, rules for exclusion of days and a method for nonwear time detection;
- PA metrics – steps, time spent in the specific intensity of PA (*e.g.* time spent in moderate to vigorous PA (MVPA)), energy expenditure (EE) (*e.g.* total EE (TEE) and activity-related EE (AEE)), ST and any other types of metrics.

We extracted data from baseline assessments for cohort or interventional studies to avoid influences of exposures or treatments on their PA. Additionally, we extracted the same data from healthy controls in included studies.

We assessed reporting quality, in the objectified measurement of PA, following the checklist by MONTJOYE *et al.* [15]. This checklist consists of 12 questions on accelerometer information, data collection and processing. Reviewers gave a “+1” score for a sufficiently reported item and a “–1” score for an insufficiently documented item. The number of “–1” scores is summed and each study was given a score of 0–12, with scores closer to 0 indicating complete reporting.

Risk of bias assessment

The two pairs independently evaluated the risk of bias following the Joanna Briggs Institute (JBI) critical appraisal checklist for an analytical cross-sectional study [16]. We used the JBI checklist for all studies regardless of these designs because only baseline data were collected. A judgment on the possible risk of bias on criteria 1–7 was made from the extracted information, rated as “yes”, “no”, “unclear” or “not applicable”. We did not use criterion 8 (was appropriate statistical analysis used?) because we focused on baseline descriptive data. If there was insufficient detail, we judged the risk of bias as “unclear”.

Data synthesis process

We performed all statistical analyses using RStudio (version 1.2.5001). Each metric was combined and calculated using the “meta” package [17].

Assessment of heterogeneity

We tested the clinical heterogeneity by considering the variability in participant factors, types of activity monitors or questionnaires, and PA data collection and management. Statistical heterogeneity was tested using the I^2 statistic.

Dealing with missing data

After the author contacts, the missing data were excluded using listwise deletion.

Data synthesis

If high heterogeneity existed among the studies ($I^2 \geq 50\%$ or $p < 0.1$), we conducted meta-analyses using a random-effects model. Calculating a pooled mean and 95% confidence interval of each metric was performed using the inverse variance method. If a study reported only a single median with an interquartile range or range, we estimated a mean and SD using sample size, median and interquartile range or range following the method reported by WAN *et al.* [18].

Investigation of heterogeneity

Subgroup analyses using aggregate data were performed to explore possible sources of heterogeneity based on the types and locations of activity monitors and patients' characteristics (ILD subtype, age, FVC, D_{LCO} , 6MWD and LTOT). We divided the included studies into two groups as follows: not old (<65 years) and old (≥ 65 years); preserved FVC ($\geq 65\%$ predicted) and low FVC (<65% predicted) [19]; preserved D_{LCO} ($\geq 45\%$ predicted) and low D_{LCO} (<45% predicted) [19]; and preserved 6MWD (≥ 350 m) and low 6MWD (<350 m) [20]. Regarding LTOT, studies were divided into three groups: studies including patients with and without LTOT (mix), only without LTOT (without) and only with LTOT (with). Moreover, studies were divided by the MVPA or ST definition. Pooled mean, 95% confidence interval and I^2 for each subgroup were estimated if a subgroup included ≥ 3 studies.

Sensitivity analysis

Sensitivity analysis was performed to explore the impact of reporting quality of PA measurements on the heterogeneity. We divided included studies into adequate reporting quality (–6–0 points) and low reporting quality (–12––7 points), based on the quartile 1 of –6.5 points in this review. Pooled mean, 95% confidence interval and I^2 for each subgroup were estimated if a subgroup included ≥ 3 studies.

We conducted meta-analysis and subgroup analyses on TEE by omitting KHOR *et al.* [21] and PRASAD *et al.* [22] to test the robustness of the estimations because the means of TEE in those studies are approximately 3.5 times higher than the other studies included in the meta-analysis.

Narrative synthesis

A narrative synthesis was provided to summarise and explain the characteristics and PA metrics following the Centre for Reviews and Dissemination guidance [23].

Confidence in cumulative evidence

The quality of evidence on PA metrics, included in meta-analyses, was assessed across the domains of risk of bias [24], inconsistency [25], indirectness [26] and imprecision [27] following the Grading of Recommendations Assessment, Development and Evaluation working group methodology [28]. Quality was adjudicated as high, moderate, low or very low.

Results

Description of studies

Details of the included studies are available in table 1.

Results of the search

The search yielded 15 407 citations and ended with 40 studies [3–5, 7–10, 21, 22, 29–59] from 49 citations (figure 1).

HUR *et al.* [60] and HUR *et al.* [8] used the same cohort data. All PA metrics reported by HUR *et al.* [60] were shown in the other study [8]. Thus, we excluded HUR *et al.* [60]. Three studies by VAINSELBOIM *et al.* [6, 7, 61] used the same participant data (Registration No. NCT01499745). A PA metric (overall PA calculated in metabolic equivalent of tasks (METs) min·week⁻¹) shown in two studies [6, 61] was not reported in the other one [7]. Therefore, we combined the two studies [6, 61] with VAINSELBOIM *et al.* [7] and excluded them [6, 61]. Two studies by DALE *et al.* [44, 62] used the same cohort data (Registration No. ACTRN12608000147381). We excluded the latter one [62]. Two reports by BAHMER *et al.* [5, 63] used the same participant data (Registration No. DRKS00006170) and BAHMER *et al.* [5] reported complete data at baseline. Thus, we excluded BAHMER *et al.* [63]. We excluded four studies that did not show the participant characteristics with ILD separately from other participants [64–67].

Included studies

Table 1 shows the characteristics of the included studies. Of the 40 studies, 11 studies [9, 21, 34, 35, 37, 46, 49–52, 58] were intervention studies, 13 studies [3, 7, 8, 22, 29, 30, 33, 36, 41, 43, 47, 54, 56] were cohort studies and 16 studies [4, 5, 10, 31, 32, 38–40, 42, 44, 45, 48, 53, 55, 57, 59] were cross-sectional studies. Of the 16 cross-sectional studies, one was a validation study of a PA questionnaire [8].

15 studies included patients with ILD due to any causes [8, 21, 30–34, 37, 46–48, 51, 53, 55, 59], 15 studies included only idiopathic pulmonary fibrosis (IPF) [5, 7, 9, 22, 29, 40–43, 45, 49, 50, 54, 56, 58], seven studies included only sarcoidosis [4, 35, 36, 38, 39, 52, 57] and three studies included another ILD subtype [3, 10, 44]. The total sample size of each study ranged from 13 [46] to 629 [57]. The mean or median age was more than 65 years in 26 studies (65%).

PA measurement

Of 40 studies, 31 (78%) studies used only activity monitors [3, 5, 10, 21, 22, 29–32, 34–45, 48–52, 55–59], seven (18%) used only questionnaires [7, 9, 33, 46, 47, 53, 54] and two (5%) used both [4, 8]. Thus, 33 (83%) of 40 studies used activity monitors and nine (23%) used questionnaires.

Table 2 summarises the activity monitor-based PA measurements. Of 33 studies, 14 studies used SenseWear Armband [3, 5, 21, 22, 35–37, 39, 40, 44, 49, 50, 55, 56], nine studies used ActiGraph [8, 10, 34, 38, 48, 51, 52, 58, 59] and four studies used Lifecorder [29, 32, 42, 45].

12 studies did not report the sensor location [21, 31, 32, 35, 37, 38, 40, 43, 50, 52, 57, 58], nine studies located the monitors at the arm (upper arm: seven studies [3, 22, 36, 39, 44, 55, 56]; not specified: two studies [5, 49]), three studies at the wrist [30, 34, 51], four studies at the waist [29, 42, 45, 59], two studies at the hip [41, 48], one study at the upper thigh [4], one study at the wrist and waist [8], and one study wrist or waist [10].

TABLE 1 Characteristics of included studies (40 studies)

Study	Study design	ILD subtype	Group	Sample size, n	Age, years	Males, n	BMI, kg·m ⁻²	FVC, % pred	D _{LCO} , % pred	mMRC	6MWD, m	LTOT, yes/no
CHU <i>et al.</i> [57]	Cross-sectional	Sarcoidosis	Whole	629	51 (11)	199	32.1 (7.8)	NA	NA	NA	NA	NR
KING <i>et al.</i> [51] [¶]	RCT	ILD	Intervention	30	69 (10)	15	32.8 (7.6)	60 (16)	36 (14)	NA	273 (81)	30/0
			Control	14	71 (5)	10	29.4 (4.1)	61 (15)	35 (8)	NA	279 (76)	14/0
FAVERIO <i>et al.</i> [54]	Cohort	IPF	Whole	90	73 (7)	71	27.6 (4.0)	87 (21)	54 (18)	NA	408 (110)	Unclear
CERDÁN-DE-LAS-HERAS <i>et al.</i> [52]	RCT	Sarcoidosis	Intervention	15	52 (13)	9	27.3 (4.9)	85 (18)	64 (17)	NA	528 (163)	NR
			Control	15	56 (14)	10	28.7 (4.5)	93 (19)	66 (16)	NA	499 (119)	NR
BREULS <i>et al.</i> [55] [¶]	Cross-sectional	ILD	Whole	45	66 (1)	31	27.8 (0.7)	81 (3)	44 (2)	NA	480 (18)	9/36
LABRECQUE <i>et al.</i> [53] [¶]	Cross-sectional	ILD	Whole	36	70 (7)	28	28 (5)	70 (14)	51 (20)	NA	476 (94)	1/35
BADENES-BONET <i>et al.</i> [56]	Cohort	IPF	Whole	40	71 (7)	30	27.4 (4.6)	79 (19)	45 (14)	1 (0.25–2.0)	452 (93)	6/34
CERDÁN-DE-LAS-HERAS <i>et al.</i> [58]	RCT	IPF	Intervention	15	70 (9)	13	NA	77 (16)	46 (11)	NA	462 (115)	3/12
			Control	14	72 (8)	8	NA	91 (17)	55 (14)	NA	446 (64)	0/13
PRASAD <i>et al.</i> [22]	Cohort	IPF	Whole	54	68 (7.5)	37	29.4 (4.6)	70 (17)	46 (17)	NA	434 (130)	18/36
AGUIAR <i>et al.</i> [59] [¶]	Cross-sectional	ILD	Whole	30	59 (10)	15	27.4 (5.3)	73 (61–80)	49 (35–67)	3 (2–4)	469 (100)	NR
SHINGAI <i>et al.</i> [29]	Cohort	IPF	Whole	87	68 (63–72)	73	23.7 (21.1–26.0)	80 (66–92)	60 (48–77)	22/33/22/5/4 ⁺	559 (463–613)	0/87
WIJSENBECK <i>et al.</i> [30]	Cohort	IPF and non-IPF	IPF	68	69.5 (65–75)	49	NA	2.8 (2.1–3.4) L [#]	NA	NA	449 (370–510)	NR
			non-IPF	62	65.5 (56–74)	30	NA	2.6 (1.8–3.3) L [#]	NA	NA	449 (317–531)	NR
ALEXANDRE <i>et al.</i> [31] [¶]	Cross-sectional	ILD	Whole	21	63 (11)	7	27.5 (5.3)	75 (21)	51 (15)	Mean 1 (min. 1, max. 1)	437 (90)	1/20
NOLAN <i>et al.</i> [50]	RCT	IPF	Whole	22	76 (74–82)	16	24.4 (22.4–29.1)	62 (50–75)	26 (22–38)	4 (4–4)	289 (149–360)	4/18
JAROSCH <i>et al.</i> [49]	RCT	IPF	Intervention	34	68 (9)	25	27.2 (4.4)	74 (19)	44 (15)	NA	415 (101)	7/27
			Control	17	65 (10)	13	27.8 (5.1)	72 (20)	37 (19)	NA	405 (109)	6/11
HIRABAYASHI <i>et al.</i> [32] [¶]	Cross-sectional	ILD	Whole	51	74 (7.5)	35	23.7 (3.5)	2.6 (0.7) [#]	74 (26)	19/13/10/6/3 ⁺	429 (124)	6/51
MONTGOMERY <i>et al.</i> [33] [¶]	Cohort	ILD	Whole	100	59 (7)	82	28.0 (4.0)	60 (18)	25 (9)	NA	NA	71/29
KHOR <i>et al.</i> [21] [¶]	RCT	ILD	Whole	30	72 (8)	22	28.1 (4.4)	71 (14)	42 (12)	NA	NA	0/30
NATHAN <i>et al.</i> [34] [¶]	RCT	ILD	Intervention	23	69	16	NA	56%	31	NA	294 (88)	23/0
			Control	13	66	13	NA	60%	30	NA	271 (91)	18/0
WALLAERT <i>et al.</i> [35]	RCT	Sarcoidosis	Intervention	20	57.5 (48–63.5)	10	28.4 (23.7–31.1)	81 (18)	57 (16)	2 (1.0–2.5)	430 (372–505)	NR
			Control	18	57.5 (49–65)	7	27.3 (23.4–31.2)	81 (18)	63 (19)	1 (0–1)	456 (380–510)	NR
CHO <i>et al.</i> [4]	Cross-sectional	Sarcoidosis	Whole	15	53 (16)	4	27.7 (4.4)	79 (18)	58 (14)	3 (1)	375 (59)	NR
FROIDURE <i>et al.</i> [36]	Cohort	Sarcoidosis	Whole	53	59 (9.4)	29	28.1 (6.6)	79 (20)	56 (18)	2 (1)	423 (115)	NR
HUR <i>et al.</i> [8]	Cohort	ILD	Whole	111	70 (9)	69	29.0 (5.0)	77 (19)	51 (16)	NA	NA	NR
PEREZ-BOGERD <i>et al.</i> [37] [¶]	RCT	ILD	Intervention	30	64 (13)	22	28.0 (4.0)	77 (21)	45 (16)	3 (0.1)	462 (123)	NR
			Control	30	64 (8)	15	26.0 (5.0)	79 (23)	41 (13)	2.6 (0.1)	491 (95)	NR

Continued

TABLE 1 Continued

Study	Study design	ILD subtype	Group	Sample size, n	Age, years	Males, n	BMI, kg·m ⁻²	FVC, % pred	D _{LCO} , % pred	mMRC	6MWD, m	LTOT, yes/no
PILZAK <i>et al.</i> [38]	Cross-sectional	Sarcoidosis	Whole	17	47 (9)	10	28.0 (4.3)	86 (36)	81 (28)	NA	544 (46)	NR
ROOT <i>et al.</i> [10] [¶]	Cross-sectional	Pulmonary fibrosis	Whole	194	66 (10)	108	28.3 (7.1)	74 (15) n=154	63 (15) n=142	NA	NA	0/194
BAHMER <i>et al.</i> [39]	Cross-sectional	Sarcoidosis	Whole	57	50 (11)	32	27.4 (4.8)	91 (21)	73 (15)	NA	525 (85)	NR
CAPPARELLI <i>et al.</i> [40]	Cross-sectional	IPF	Whole	23	72 (7)	18	NA	69 (16)	40 (18)	3 (1–4)	379 (92)	NR
NISHIYAMA <i>et al.</i> [41]	Cohort	IPF	Whole	31	72 (5)	23	23.1 (3.1)	74 (18)	61 (17)	NA	410 (78)	0/31
VAINSELBOIM <i>et al.</i> [7]	Cohort	IPF	Whole	34	Mean 68 (min. 50, max. 81)	22	Mean 29.0 (min. 22.0, max. 37.0)	68 (37–109)	Mean 50 (min. 23, max. 91)	1/14/8/10/1 ⁺	505 (130–749)	13/21
MORINO <i>et al.</i> [42]	Cross-sectional	IPF	Whole	38	71 (8)	26	23.9 (3.0)	88 (20)	48 (17)	1 (0.7)	444 (99)	5/33
ATKINS <i>et al.</i> [43]	Cohort	IPF	Whole	39	75 (8)	24	NA	83 (18)	52 (14)	5/12/11/5/2 ⁺	325 (127)	0/39
BAHMER <i>et al.</i> [5]	Cross-sectional	IPF	Whole	48	67 (8)	36	27.9 (4.5)	75 (23)	43 (15)	1.5 (1–2)	355 (140)	11/36
DALE <i>et al.</i> [44]	Cross-sectional	Dust-related ILD and ARPD	Dust-related ILD	10	72 (10)	10	27.0 (3.0)	98 (19)	49 (13)	1 (1)	462 (102)	0/10
			ARPD	22	71 (5)	22	28 (3.0)	82 (20)	59 (13)	1 (1)	483 (68)	0/22
NAKAYAMA <i>et al.</i> [45]	Cross-sectional	IPF	Whole	31	68 (6)	22	24.2 (2.9)	89 (21)	79 (22)	17/12/2/0/0 ⁺	436 (82)	2/29
KEYSER <i>et al.</i> [46]	non-RCT	ILD	Whole	13	57 (9)	5	28.3 (4.4)	NA	40 (15)	NA	433 (93)	5/8
GAUNAURD <i>et al.</i> [9]	RCT	IPF	Intervention	11	71 (6)	NA	NA	60 (11)	44 (11)	NA	361 (55)	NR
			Control	10	66 (7)	NA	NA	61 (14)	43 (11)	NA	339 (109)	NR
RYERSON <i>et al.</i> [47] [¶]	Cohort	ILD	Whole	54	69 (11)	26	NA	69 (21)	47 (13)	NA	366 (120)	19/35
WALLAERT <i>et al.</i> [3] [¶]	Cohort	f-IIPs and NSIP	Whole	50	64 (10)	29	26.3 (4.8)	71 (21)	37 (13)	NA	347 (109)	12/38
WICKERSON <i>et al.</i> [48] [¶]	Cross-sectional	ILD	Whole	24	62 (53–65)	NA	25.5 (3.5)	49 (14)	43 (14)	NA	344 (67)	24/0

Data are expressed as mean (SD), or median (25th–75th percentiles), unless otherwise stated. ARPD: asbestos-related pleural disease; BMI: body mass index; D_{LCO}: diffusing capacity of the lung for carbon monoxide; f-IIP: fibrotic idiopathic interstitial pneumonia; FVC: forced vital capacity; IIP: idiopathic interstitial pneumonia; ILD: interstitial lung disease; IPF: idiopathic pulmonary fibrosis; LTOT: long-term oxygen therapy; max.: maximum; min.: minimum; mMRC: modified Medical Research Council dyspnoea score; NA: not applicable; NR: not reported; NSIP: nonspecific interstitial pneumonia; RCT: randomised controlled trial; 6MWD: 6-min walk distance. [‡]: Unit is L; ⁺: presented as grade 0/1/2/3/4; [¶]: LABRECQUE *et al.* [53] included IPF, NSIP, unclassifiable IIP and connective tissues disease-related ILD. BREULS *et al.* [55] included IPF, idiopathic NSIP, cryptogenic organising pneumonia and IIP. KING *et al.* [51] included IPF, idiopathic NSIP, unclassifiable IIP and others. AGUIAR *et al.* [59] included IPF, pulmonary fibrosis related to diseases of the connective tissue, diseases related to inhalation of particles and NSIP. ALEXANDRE *et al.* [31] included ILD secondary to rheumatic disease, IPF, ILD due to hypersensitivity, usual interstitial pneumonia and no definite cause for ILD. HIRABAYASHI *et al.* [32] included IPF, cardiovascular disease-associated ILD, dermatomyositis-associated ILD, rheumatoid arthritis-associated ILD, anti-neutrophil cytoplasmic antibody-associated ILD, systemic sclerosis-associated ILD, Sjögren's syndrome-associated ILD and other (biopsy not performed or non-IPF). MONTGOMERY *et al.* [33] included IPF, hypersensitivity pneumonitis, NSIP, connective tissue disease and others. KHOR *et al.* [21] included IPF, chronic hypersensitivity pneumonia, connective tissue disease-related ILD, interstitial pneumonia with autoimmune features, sarcoidosis and unclassifiable ILD. NATHAN *et al.* [34] included IPF, hypersensitivity pneumonitis, NSIP and unclassifiable IIP. PEREZ-BOGERD *et al.* [37] included IPF, NSIP, desquamative interstitial pneumonia, chronic hypersensitivity pneumonitis, asbestosis, drug-induced ILD, connective tissue disease-related ILD and unclassifiable ILD. ROOT *et al.* [10] included IPF, connective tissue disease, chronic hypersensitivity pneumonitis, familial pulmonary fibrosis and others. RYERSON *et al.* [47] included IPF, unclassifiable ILD, fibrotic NSIP and others. WALLAERT *et al.* [3] included f-IIPs and NSIP. WICKERSON *et al.* [48] included interstitial idiopathic pneumonia, collagen vascular disease, sarcoidosis and extrinsic allergic alveolitis.

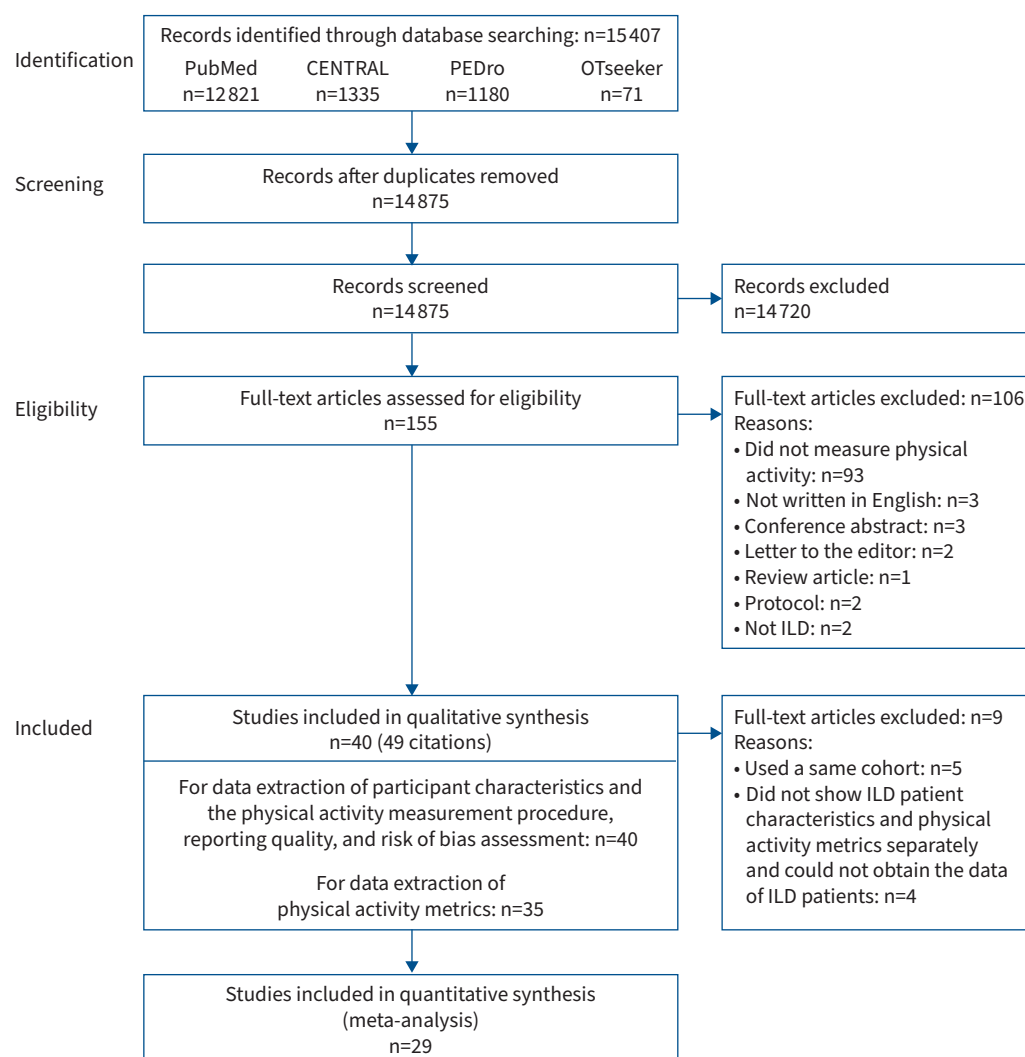


FIGURE 1 Flowchart of literature review. ILD: interstitial lung disease.

Of nine studies that used questionnaires, five studies used the International Physical Activity Questionnaire (IPAQ) (long form: one study [8]; short form: three studies [4, 7, 9]; not reported: one study [54]), four studies used the Human Activity Profile [46], the Rapid Assessment of Physical Activity Questionnaire [47], a part of frailty assessment by the Fried Frailty Criteria [33], and the Minnesota Leisure-Time Physical Activity Questionnaire short form [53] (table S1 in supplementary material 2).

Reporting quality of PA measurements

Table 3 summarises the reporting quality of PA measurements using activity monitors (33 studies). Accelerometer brand and model, the number of days of data collected, PA metrics, and the number of people not meeting wear-time criteria were well reported in most studies. However, epoch length, placement of the accelerometer (especially on the side of the body), the number of participants receiving the accelerometer, distribution method of the accelerometer, criteria for defining nonwear time, minutes requirement for a valid day, and the number of valid days needed were poorly reported (table S2 in supplementary material 2).

Risk of bias

Table S3 in supplementary material 2 shows the results of the risk of bias. The domains with the highest risks of bias were item 7 regarding the method of PA measurements and items 5 and 6 regarding confounding factors.

TABLE 2 Details of activity monitor-based physical activity measurement in included studies (33 studies)

Study	Name (model)	Sensor type	Sensor location		Distribution	Wear period, days	Wear time, h·day ⁻¹ #	WD/WE	Definition of valid day, h·day ⁻¹	Min. valid days	Exclusion rule	Definition of nonwear time	Number of excluded
			Site	Side									
CHU <i>et al.</i> [57]	Sarcoidosis app (smartphone app)	NR	NR	NR	Download	NR	NR	NR	NR	NR	NR	NR	NR
KING <i>et al.</i> [51]	ActiGraph (GT9X)	T-Ac	Wrist	NR	Hand over	Whole of study period	24	Both	10	14	NR	NR	0
CERDÁN-DE-LAS-HERAS <i>et al.</i> [52]	ActiGraph (wGT3x-BT)	T-Ac	NR	NR	NR	7	NR	Both	NR	NR	NR	NR	0
BREULS <i>et al.</i> [55]	SenseWear (Pro)	NR	Upper arm	NR	NR	≥7	≥8 (07:00–22:00)	Both	≥8	4	Yes	NR	0
BADENES-BONET <i>et al.</i> [56]	SenseWear (Pro2)	Multi-sensor	Upper arm	Left	Hand over	7	23 (personal hygiene)	NR	70% of the daytime (AM8-PM10)	≥3	Yes	NR	0
CERDÁN-DE-LAS-HERAS <i>et al.</i> [58]	ActiGraph (wGT3x-BT)	T-Ac	NR	NR	NR	7	NR	NR	NR	NR	NR	NR	0
PRASAD <i>et al.</i> [22]	SenseWear (Pro)	Multi-sensor	Upper arm	Nondominant	NR	7	24 (high risk of water damage)	Both	22	≥4 WD and ≥1 WE	NR	NR	0
AGUIAR <i>et al.</i> [59]	Actigraph (wGT3x-BT)	T-Ac	Waist	NR	NR	6	24	Both	8	4	NR	NR	0
SHINGAI <i>et al.</i> [29]	Kenz Lifecorder (GS)	U-Ac	Waist	NR	NR	7	24 (bath and sleep)	Both	8	7	Yes	NR	0
WIJSENBEEK <i>et al.</i> [30]	WITHINGS (Steel HR)	Multi-sensor	Wrist	NR	NR	7	24 (bath and sleep)	Both	NR	NR	No	NR	4
ALEXANDRE <i>et al.</i> [31]	DynaPort (MiniMod)	T-Ac	NR	NR	NR	2	12	NR	NR	2	NR	NR	2
NOLAN <i>et al.</i> [50]	SenseWear (NR)	Multi-sensor	NR	NR	NR	7	24	Both	22.5	5	NR	No	0
JAROSCH <i>et al.</i> [49]	SenseWear (NR)	Multi-sensor	Arm	NR	NR	7	≥23	Both	23	NR	No	No	2
HIRABAYASHI <i>et al.</i> [32]	Kenz Lifecorder (GS)	U-Ac	NR	NR	NR	7	NR	NR	NR	5	No	NR	1
KHOR <i>et al.</i> [21]	SenseWear (NR)	Multi-sensor	NR	NR	NR	7	NR	Both	NR	NR	NR	NR	7
NATHAN <i>et al.</i> [34]	Actigraph (GT9X)	T-Ac	Wrist	NR	Hand over	NR	24	Both	NR	NR	NR	NR	4
WALLAERT <i>et al.</i> [35]	SenseWear (Pro Armband)	Multi-sensor	NR	NR	NR	5	24 (shower or bath)	Both	NR	NR	NR	NR	0
CHO <i>et al.</i> [4]	ActivPal™ (ActivPal3™)	T-Ac	Upper thigh	NR	Hand over	8	24 (personal hygiene)	Both	8	NR	Yes	Yes	0

Continued

TABLE 2 Continued

Study	Name (model)	Sensor type	Sensor location		Distribution	Wear period, days	Wear time, h·day ⁻¹ #	WD/WE	Definition of valid day, h·day ⁻¹	Min. valid days	Exclusion rule	Definition of nonwear time	Number of excluded
			Site	Side									
FROIDURE <i>et al.</i> [36]	SenseWear (Pro Armband)	Multi-sensor	Upper arm	Right	NR	5	24 (bath or swim)	Both	NR	NR	NR	NR	0
HUR <i>et al.</i> [8]	ActiGraph (wGT3X-BT)	T-Ac	Wrist and waist	Nondominant	NR	7	24 (bath or swim)	Both	8	7	Yes	Yes	0
PEREZ-BOGERD <i>et al.</i> [37]	SenseWear (NR)	Multi-sensor	NR	NR	NR	7	Wake (07:00–20:00)	WD	8	2	Yes	NR	0
PILZAK <i>et al.</i> [38]	Actigraph (GT3X+)	T-Ac	NR	NR	NR	7	15	Both	NR	NR	NR	NR	NR
ROOT <i>et al.</i> [10]	Actigraph (GT3X+)	T-Ac	Wrist or waist	NR	Mailing	7	24	Both	10	≥4	Yes	Yes	1
BAHMER <i>et al.</i> [39]	SenseWear (Armband)	Multi-sensor	Upper arm	Left	NR	7	24 (personal hygiene)	Both	22.5	≥5	Yes	NR	2
CAPPARELLI <i>et al.</i> [40]	SenseWear (Armband)	Multi-sensor	NR	NR	NR	6	24 (personal hygiene)	Both	NR	NR	NR	NR	0
NISHIYAMA <i>et al.</i> [41]	Actical (NR)	Om-d AC	Hip	NR	NR	7	24 (bath and sleep)	Both	NR	NR	NR	NR	0
MORINO <i>et al.</i> [42]	Kenz Lifecorder (NR)	U-Ac	Waist	NR	NR	15	Wake (bath)	Both	10	NR	Yes	No	0
ATKINS <i>et al.</i> [43]	GENEActiv actiwatch (NR)	T-Ac	Wrist	Nondominant	NR	≥7	NR	Both	16	≥2 WD and ≥2 WE	No	No	4
BAHMER <i>et al.</i> [5]	SenseWear (NR)	T-Ac	Arm	NR	NR	7	24 (personal hygiene)	Both	22.5	6	Yes	No	3
DALE <i>et al.</i> [44]	SenseWear (Pro 3 Armband)	Multi-sensor	Upper arm	Right	NR	9	24 (shower, swim, or adverse event)	Both	≥20	≥4	No	No	0
NAKAYAMA <i>et al.</i> [45]	Kenz Lifecorder (GS)	U-Ac	Waist	NR	NR	1 month	NR	Both	≥10	NR	Yes	No	0
WALLAERT <i>et al.</i> [3]	SenseWear (Pro Armband)	Multi-sensor	Upper arm	Right	NR	4	24 (shower or bath)	Both	NR	4	NR	No	0
WICKERSON <i>et al.</i> [48]	Actigraph (GT3X)	T-Ac	Hip	NR	NR	7	Wake	Both	8	NR	No	No	1

Min.: minimum; NR: not reported; Om-d Ac: omni-dimensional accelerometer; T-Ac: tri-axial accelerometer; U-Ac: uni-axial accelerometer; WE: weekend; WD: weekday. #: () represents “except for”.

TABLE 3 Reporting quality of accelerometer-based physical activity measurement (33 studies)

Criteria	Studies meeting each criterion, %
1) Brand of accelerometer used	97
2) Model of accelerometer used	82
3) Epoch length used	27
4) Placement of accelerometer (must indicate location and side of the body)	24
5) Number of participants enrolled at study start receiving accelerometers	52
6) How the accelerometers were distributed	21
7) Days of data collected	97
8) Criteria for defining nonwear of accelerometer	12
9) How many minutes of accelerometer data needed to be considered a valid day	58
10) Number of valid days of accelerometer data needed	52
11) What were physical activity outcome variables	88
12) Reported the number of people not meeting wear-time criteria	91
Total score, median (IQR)	−5 [−6.5—3]
Total score ranged from −12 to 0. A higher score indicates better reporting quality of accelerometer-based physical activity measurement. IQR: interquartile range.	

PA metrics

Of 40 studies, 35 studies (activity monitor only: 29 studies [3, 5, 10, 21, 22, 29–32, 35–45, 48–52, 55, 56, 58, 59]; questionnaire only: four studies [7, 46, 53, 54]; both: two studies [4, 8]) displayed at least one PA metric of patients with ILD. The PA metrics used were heterogeneous (tables S4 and S5 in supplementary material 2).

Of 29 studies that used activity monitors, 28 studies (97%) reported steps [3–5, 8, 10, 22, 29–32, 35–42, 44, 45, 48–50, 52, 55, 56, 58, 59], 16 studies (55%) MVPA [3, 5, 8, 10, 21, 22, 31, 35, 37, 41, 44, 48, 50, 51, 55, 56], eight studies (20%) TEE [3, 21, 22, 31, 35, 36, 44, 45], four studies (10%) AEE [8, 22, 41, 45] and nine studies (23%) ST [8, 10, 21, 31, 41, 43, 50, 55, 56]. Definitions of MVPA and ST were heterogeneous (table S4 in supplementary material 2). For example, seven of 16 studies defined moderate intensity of PA as >3 METs [5, 22, 31, 37, 44, 55, 56], two studies defined it as >2.5 METs [3, 35] and others used acceleration magnitude [8, 51] or EE [41] for defining MVPA.

Of nine studies that used questionnaires, four studies presented IPAQ overall EE (MET min·week^{−1}) [4, 7, 8, 54], two studies IPAQ walking EE (MET min·week^{−1}) [4, 7] and ST (min·day^{−1}) [7, 8], and one study IPAQ MVPA EE (MET min·week^{−1}) [8] (table S4 in supplementary material 2). Other metrics measured by questionnaires are described in table S5 in supplementary material 2.

Estimation of PA levels in steps, MVPA, TEE and ST

The pooled mean (95% CI) of steps was 5215 (4640–5791) steps·day^{−1} ($I^2=97$ (97–98) %) (figure 2a). Subgroup analyses found that people with IPF or ILD took fewer steps than those with sarcoidosis. Activity monitors worn on the wrist or upper arm showed higher steps than those worn on the waist or lower extremity. Additionally, people with lower FVC, D_{LCO} , 6MWD and LTOT exhibited fewer steps (figures S1–S8 in supplementary material 3). However, I^2 only slightly improved in subgroup analyses by the ILD subtype.

The pooled mean of MVPA was 82 (58–106) min·day^{−1} ($I^2=99$ (99–99) %) (figure 2b). Subgroup analyses revealed that people with sarcoidosis spent more time in MVPA than the other ILD subtypes (figures S9–S17 in supplementary material 3). MVPA measured by SenseWear (91 (61–121) min·day^{−1}) appeared to be higher than by ActiGraph (49 (3–96) min·day^{−1}), but it was not significant. Patients with lower FVC or LTOT spent a shorter time in MVPA than those with preserved FVC or without LTOT (all $p<0.05$). A mean of MVPA, defined as 2.5 METs or more, was approximately 2.5 times longer than that defined as 3.0 METs (152 (123–181) min·day^{−1} versus 67 (51–82) min·day^{−1}) with a slight improvement in heterogeneity.

The pooled mean of TEE was 3574 (1684–5464) kcal·day^{−1} ($I^2=100$ (100–100) %) (figure S18 in supplementary material 3). We did not perform a subgroup analysis by FVC because no study was classified into the low FVC group. Activity monitors worn on the upper extremity showed higher TEE

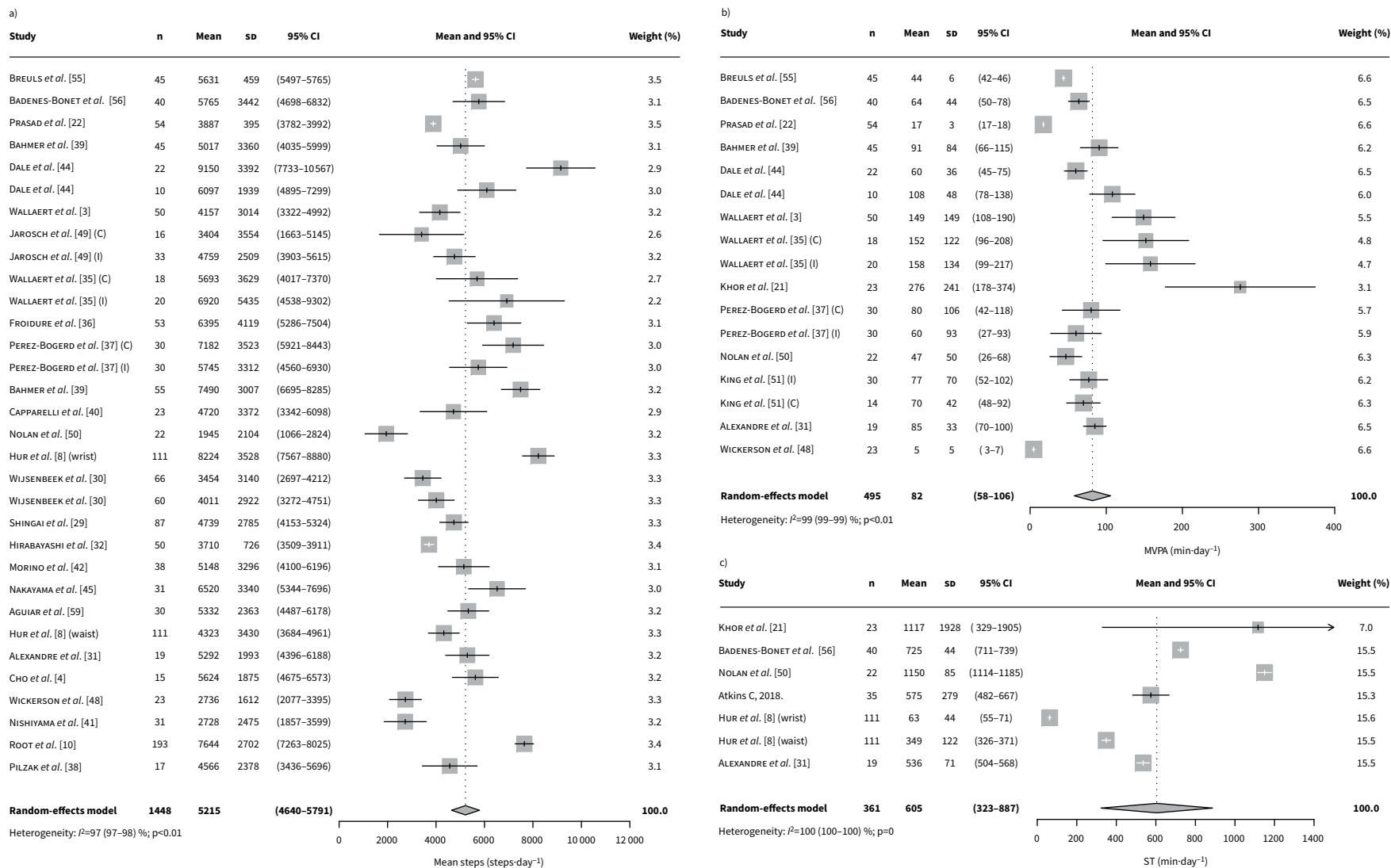


FIGURE 2 Forest plots of estimation of overall means of **a)** steps (steps·day⁻¹), **b)** moderate to vigorous physical activity (MVPA) (min·day⁻¹) and **c)** sedentary time (ST) (min·day⁻¹). Weight is calculated by the random-effects model. Wrist and waist refer to the activity monitor location. C: control; I: intervention.

than those worn on the waist or lower extremity (figures S19–S25 in supplementary material 3). Heterogeneity was improved only in subgroup analysis by the ILD subtype.

The pooled mean of ST was 605 (323–887) $\text{min}\cdot\text{day}^{-1}$ ($I^2=100$ (100–100) %) (figure 2c). Types of activity monitors showed significant differences in ST. People with worse FVC or D_{LCO} exhibited longer ST than those with better FVC or D_{LCO} . ST significantly differed between studies that used different definitions of ST (figures S26–S34 in supplementary material 3). There was no improvement in I^2 in all subgroup analyses.

Sensitivity analysis

Subgroup analyses by reporting quality of PA measurements found that studies with low reporting quality reported longer MVPA than those with adequate reporting quality ($p=0.02$) (figures S35–S38 in supplementary material 4). Substantial heterogeneities were observed in all PA metrics.

After omitting two studies [21, 22], the pooled mean value of TEE was 2130 (1847–2412) $\text{kcal}\cdot\text{day}^{-1}$ (figure S39 in supplementary material 4). I^2 was unchanged (96 (94–97) %). Subgroup analyses showed similar results (figures S40–S47 in supplementary material 4).

Reference data of healthy controls

Of 40 studies, seven studies recruited healthy controls (tables S6 and S7 and figures S48 and S49 in supplementary material 5). The pooled mean of steps from six studies was 10 167 (8433–11 901) $\text{steps}\cdot\text{day}^{-1}$ ($I^2=88$ (76–94) %). The pooled mean of TEE from three studies was 2618 (2505–2730) $\text{kcal}\cdot\text{day}^{-1}$ ($I^2=15$ (0–91) %). The means (SD) of MVPA in three studies were 261 (118), 132 (72) and 86 (8) $\text{min}\cdot\text{day}^{-1}$. No study reported ST.

Quality of the body of evidence

The quality of the body of evidence on steps, MVPA, TEE and ST was very low, mainly due to severe inconsistency and imprecision (table S8 in supplementary material 6).

Discussions

This systematic review revealed that 1) measurement procedures varied tremendously between studies, 2) reporting quality of PA measurements was poor in most studies, 3) types and definitions of PA metrics were heterogeneous and influenced the PA metrics values, and 4) use of PA questionnaires is limited in patients with ILD. Additionally, there was very low-quality evidence in the pooled means of steps, MVPA, TEE and ST. Therefore, clinicians and researchers should improve the quality of PA measurements.

Reporting quality of PA measurements

Four systematic reviews have assessed the reporting quality of PA measurements using accelerometers in the general population [15], cancer survivors [68] and patients with chronic heart failure [69] or COPD [70]. Although direct comparisons of our results to these reviews are difficult due to the differences in methodology of reporting quality assessment, we selected the same tool used in two reviews [15, 69], enabling us to compare the reporting quality in patients with ILD to other populations.

More than 50% of studies included in this review failed to report six of 12 items related to data collecting and processing. Specifically, 88% of the included studies failed to report the criteria for defining nonwear, compared with 69% in general populations [15], 49% in cancer survivors [68] and 80% in heart failure [69]. Moreover, included studies also failed to report how accelerometers were distributed to participants (79%), placement of accelerometer (76%) and epoch length used (73%). The percentages are worse than those in the general population (69, 51 and 36%) [15], cancer survivors (46, 7 and 52%) [68] and heart failure (0, 64 and 60%) [69]. In contrast, 48% of the included studies did not report the number of valid days needed and 42% failed to report how many minutes were needed to be considered a valid day. The percentages are similar to those in the general population (52 and 50%) [15], cancer survivors (43 and 38%) [68] and heart failure (76 and 78%) [69]. BURTIN *et al.* [70] revealed that only 37 of 110 (34%) studies with COPD patients fulfilled the following minimal preferred methodologic quality of PA assessment: measurement period ≥ 7 days; minutes needed to be considered a valid day $\geq 8 \text{ h}\cdot\text{day}^{-1}$; ≥ 4 consecutive or nonconsecutive valid days; and invalid days excluded from analysis. In our review, only 36% met the criteria. We are reluctant to use these results to claim that the reporting quality in studies with ILD is similar or inferior to that in other populations. However, we believe that reporting on PA measurements should be improved because these factors are crucial for replicating and comparing studies. Therefore, we encourage clinicians and researchers to report data collecting and processing methods following the checklist by MONTROYE *et al.* [15], as shown in table 3 and table S2 in supplementary material 2.

We have developed a template for adequate reporting quality (table S9 in supplementary material 7), which can be used to report required information as supplementary materials.

Risk of bias

The studies included in this review recruited well-defined patients with ILD using objective measures following international guidelines. In contrast, the included studies did not adequately define and deal with confounding factors. Difficulties in recruiting enough patients with ILD may be a possible reason for this, because a large sample size is required to control confounding factors. Furthermore, most studies failed to measure PA validly and reliably. Therefore, defining and dealing with confounding factors and PA measurements are at high risk of bias, leading to the heterogeneity of the PA metric values.

PA metrics

This review showed the pooled mean values of four commonly used PA metrics (steps, MVPA, TEE and ST) in patients with ILD. However, we need to be cautious when interpreting these values due to the substantial heterogeneity caused by differences in the activity monitors used, PA measurement methods and participant characteristics. For example, although most studies used validated activity monitors and PA metrics in COPD [71–74], there was no validation study of them in patients with ILD. However, we believe that these values are valid because PA metrics are broadly distributed in the general population or people with chronic diseases.

First, the pooled mean value of steps in ILD (5231 (95% CI 4577–5885) steps·day⁻¹) was similar to that in patients with COPD (5723 (SD, 3768) steps·day⁻¹) [75] and about two times lower than that in their healthy counterparts (10 195 (95% CI 8023–12 367) steps·day⁻¹). The mean value is slightly higher than a physical inactivity threshold (5000 steps·day⁻¹) [76], but the 95% CI includes 5000 steps·day⁻¹. Moreover, subgroup analyses revealed that type of activity monitor, sensor location, ILD subtype, FVC, D_{LCO} , 6MWD and use of LTOT are associated with steps, which aligns with previous findings. Although caution is required when interpreting the values due to variations in how activity monitors count steps [77], about half or more patients with ILD are inactive in terms of walk-related PA.

Second, the pooled mean value of MVPA in ILD was 97 (95% CI 64–130) min·day⁻¹. This value is lower than healthy controls in included studies (132 and 261 min·day⁻¹) and a previous study recruited similar-aged people (156 min·day⁻¹) [78]. Subgroup analysis showed that the mean was 152 (95% CI 123–181) min·day⁻¹ in studies which defined MVPA as time spent in PA of ≥ 2.5 METs, while the mean of MVPA defined as >3.0 METs was 67 (95% CI 51–82) min·day⁻¹. Patients with COPD spent similar times in MVPA, defined as >3.0 METs (65 (SD 11) min·day⁻¹) [55]. These results suggest that using a consistent definition of MVPA across studies and populations helps to ensure valid comparisons.

The mean values of MVPA in patients with ILD and healthy controls were dramatically higher than the international guideline recommendation of 150 min·week⁻¹ (approximately 21 min·day⁻¹) [2]. Similar values were reported in other populations (e.g. COPD, healthy controls) [55]. This highlights the need for a new PA recommendation based on MVPA measured by activity monitors. Moreover, calibration studies are needed to establish ILD-specific cut points to distinguish between various PA intensity levels using indirect calorimetry for the following reasons. First, cut points are drastically different between populations (e.g. children *versus* adults *versus* elderly) [79, 80]. Second, using cut points derived from healthy people in patients with exercise intolerance (e.g. COPD, ILD) may be flawed.

Third, the pooled mean of TEE was 2917 (95% CI 1370–4463) kcal·day⁻¹. This value was higher than that in healthy controls in included studies (2618 kcal·day⁻¹) and older adults (2501 kcal·day⁻¹) [81]. After omitting two studies with outliers, the mean was changed to 2130 (95% CI 1847–2412) kcal·day⁻¹. TEE of 2130 kcal·day⁻¹ could be referred to as a more valid value of TEE because the means of TEE in the omitted studies might be outliers.

Fourth, the pooled mean value of ST was 587 (95% CI 253–920) min·day⁻¹. Subgroup analyses suggested that possible sources of heterogeneity were the differences in ST definition, sensor types and D_{LCO} . Additionally, differences in calculation method and wear-time requirements may affect ST. For example, the means of ST in HUR *et al.* [8] and ALEXANDRE *et al.* [31] are 349 and 536 min·day⁻¹. HUR *et al.* [8] measured PA 24 h·day⁻¹ and calculated daytime ST by excluding sleeping time. ALEXANDRE *et al.* [31] measured daytime (12 h·day⁻¹) PA and calculated daytime ST. Interestingly, similar ST was observed in healthy older people in studies that measured daytime ST [82, 83]. Thus, standardisation of definition, wear-time requirement and calculation method are required to compare studies.

Finally, the use of questionnaires is limited in ILD. The most common questionnaire was IPAQ. Concurrent validity, internal consistency and responsiveness of IPAQ long-form were acceptable, and they estimated a minimally important difference [8]. Thus, IPAQ long-form can be used in combination with activity monitors or as an alternative in situations where activity monitors are unavailable.

Quality of the body of evidence

The quality of a body of evidence on steps, MVPA, TEE and ST was very low due to serious inconsistency and imprecision. Their sources were not only participant characteristics but also PA data collection and processing. Thus, future studies which measure PA using a standardised method with high reporting quality will change the pooled mean values of PA metrics.

Limitations

There are limitations in this review. First, the authors did not have access to Embase, which may reduce its comprehensiveness. Although the *Cochrane Handbook for Systematic Reviews of Interventions* encourages researchers to search Embase if accessible, Cochrane regularly searches Embase for trial reports and includes them in CENTRAL [84]. Thus, our review included at least relevant interventional studies, contributing to the comprehensive search. Second, ILD subtypes were inconsistent among included studies and affected the pooled value of PA metrics. Third, non-English studies were excluded during the selection process, potentially causing language bias. However, language restrictions during the selection process appeared to have little impact on language bias [85]. Fourth, studies that did not report separate participant characteristics and PA metrics for ILD patients were excluded. Fifth, multiple subgroup analyses were performed to investigate the sources of heterogeneity. Multiple subgroup analyses could reduce statistical power, leading to the potential of overlooking significant differences between subgroups [86]. Finally, we used several estimated mean values of individual studies not reporting the mean and SD of PA metrics to estimate the pooled means. These limitations could lead to biased results.

Practical recommendations

- 1) Researchers and clinicians are encouraged to report details of PA measurements following the checklist by MONTÖYE *et al.* [15] or the template presented in table S9 in supplementary material 7.
- 2) Unfortunately, there is no standardised method for measuring PA in patients with ILD. Following expert consensus on objectively measured PA in COPD [87, 88] may be crucial to facilitate interpretation, pooling of PA data and comparisons with COPD.
- 3) The same MVPA and ST definitions should be used. We propose time spent in ≥ 3.0 METs PA as a definition of MVPA and time spent in ≤ 1.5 METs as a definition of ST, following the World Health Organization guidelines [89].
- 4) For identifying nonwear time, using activity monitors with automated algorithms for calculating nonwear time [90, 91] is encouraged if available. Alternatively, a wear diary could be used.
- 5) Patient characteristics, including ILD subtype, the severity of ILD, pulmonary function, exercise capacity and the number of patients with LTOT should be reported because these variables influence PA.
- 6) Validation studies are necessary to ensure the accuracy of major activity monitors and PA metrics in patients with ILD. Additionally, calibration studies are required to establish an ILD-specific cut-off for counting steps and distinguishing different intensities of PA.

Conclusions

In this systematic review of 40 studies using activity monitors or questionnaires in patients with ILD, we found severe heterogeneity in the methods used for PA data collection, processing and definition of PA metrics. The heterogeneity makes it difficult to interpret and pool PA data and compare results in ILD with other diseases. Therefore, we encourage researchers and clinicians to improve the quality of PA measurements. Our recommendations could be helpful when measuring and reporting PA in patients with ILD.

Points for clinical practice and questions for future research

We propose that PA measurements in patients with ILD should be conducted using validated activity monitors and PA metrics following expert consensus on objectively measured PA in COPD. Therefore, clinicians and researchers are encouraged to report details of PA measurements based on the checklist by MONTÖYE *et al.* [15] or the template shown in table S9 in supplementary material 7. In addition, validation and calibration studies are required for more accurate measurements of PA in patients with ILD.

Provenance: Submitted article, peer reviewed.

Acknowledgements: The authors would like to thank all authors of included studies for contributing to the advancement of PA measurement in ILD.

Conflict of interest: M.A. Spruit reports grants from Lung Foundation Netherlands and Stichting Astma Bestrijding, grants and fees from Boehringer Ingelheim, AstraZeneca, Chiesi and TEVA, outside the submitted work. All grants and fees were paid to the author's institute. All other authors have no conflict of interest to report.

References

- 1 Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public Health Rep* 1985; 100: 126–131.
- 2 Piercy KL, Troiano RP, Ballard RM, *et al.* The physical activity guidelines for Americans. *JAMA* 2018; 320: 2020–2028.
- 3 Wallaert B, Monge E, Le Rouzic O, *et al.* Physical activity in daily life of patients with fibrotic idiopathic interstitial pneumonia. *Chest* 2013; 144: 1652–1658.
- 4 Cho PSP, Vasudevan S, Maddocks M, *et al.* Physical inactivity in pulmonary sarcoidosis. *Lung* 2019; 197: 285–293.
- 5 Bahmer T, Kirsten A-M, Waschki B, *et al.* Clinical correlates of reduced physical activity in idiopathic pulmonary fibrosis. *Respir Int Rev Thorac Dis* 2016; 91: 497–502.
- 6 Vainshelboim B, Kramer MR, Izhakian S, *et al.* Physical activity and exertional desaturation are associated with mortality in idiopathic pulmonary fibrosis. *J Clin Med* 2016; 5: E73.
- 7 Vainshelboim B, Oliveira J, Izhakian S, *et al.* Lifestyle behaviors and clinical outcomes in idiopathic pulmonary fibrosis. *Respir Int Rev Thorac Dis* 2018; 95: 27–34.
- 8 Hur SA, Guler SA, Khalil N, *et al.* Minimal important difference for physical activity and validity of the international physical activity questionnaire in interstitial lung disease. *Ann Am Thorac Soc* 2019; 16: 107–115.
- 9 Gaunaurd IA, Gómez-Marín OW, Ramos CF, *et al.* Physical activity and quality of life improvements of patients with idiopathic pulmonary fibrosis completing a pulmonary rehabilitation program. *Respir Care* 2014; 59: 1872–1879.
- 10 Root ED, Graney B, Baird S, *et al.* Physical activity and activity space in patients with pulmonary fibrosis not prescribed supplemental oxygen. *BMC Pulm Med* 2017; 17: 154.
- 11 Limb ES, Ahmad S, Cook DG, *et al.* Measuring change in trials of physical activity interventions: a comparison of self-report questionnaire and accelerometry within the PACE-UP trial. *Int J Behav Nutr Phys Act* 2019; 16: 10.
- 12 Seemungal T, Harper-Owen R, Bhowmik A, *et al.* Respiratory viruses, symptoms, and inflammatory markers in acute exacerbations and stable chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2001; 164: 1618–1623.
- 13 Spruit MA, Singh SJ, Garvey C, *et al.* An official American Thoracic Society/European Respiratory Society statement: key concepts and advances in pulmonary rehabilitation. *Am J Respir Crit Care Med* 2013; 188: e13–e64.
- 14 Mahler DA, Wells CK. Evaluation of clinical methods for rating dyspnea. *Chest* 1988; 93: 580–586.
- 15 Montoye AHK, Moore RW, Bowles HR, *et al.* Reporting accelerometer methods in physical activity intervention studies: a systematic review and recommendations for authors. *Br J Sports Med* 2018; 52: 1507–1516.
- 16 Moola S, Munn Z, Tufanaru C, *et al.* Chapter 7: Systematic reviews of etiology and risk. In: Aromataris E, Munn Z, eds. *JBIM Manual for Evidence Synthesis*, Joanna Briggs Institute, 2020. <https://jbi-global-wiki.refined.site/space/MANUAL>
- 17 Schwarzer G. meta: general package for meta-analysis. 2021. <https://CRAN.R-project.org/package=meta>
- 18 Wan X, Wang W, Liu J, *et al.* Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC Med Res Methodol* 2014; 14: 135.
- 19 du Bois RM, Weycker D, Albera C, *et al.* Ascertainment of individual risk of mortality for patients with idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2011; 184: 459–466.
- 20 du Bois RM, Albera C, Bradford WZ, *et al.* 6-minute walk distance is an independent predictor of mortality in patients with idiopathic pulmonary fibrosis. *Eur Respir J* 2014; 43: 1421–1429.
- 21 Khor YH, Holland AE, Goh NSL, *et al.* Ambulatory oxygen in fibrotic interstitial lung disease: a pilot, randomized, triple-blinded, sham-controlled trial. *Chest* 2020; 158: 234–244.
- 22 Prasad JD, Paul E, Holland AE, *et al.* Physical activity decline is disproportionate to decline in pulmonary physiology in IPF. *Respirology* 2021; 26: 1152–1159.
- 23 Tacconelli E. Systematic reviews: CRD's guidance for undertaking reviews in health care. *Lancet Infect Dis* 2010; 10: 226.
- 24 Guyatt GH, Oxman AD, Vist G, *et al.* GRADE guidelines: 4. Rating the quality of evidence—study limitations (risk of bias). *J Clin Epidemiol* 2011; 64: 407–415.

- 25 Guyatt GH, Oxman AD, Kunz R, *et al.* GRADE guidelines: 7. Rating the quality of evidence—inconsistency. *J Clin Epidemiol* 2011; 64: 1294–1302.
- 26 Guyatt GH, Oxman AD, Kunz R, *et al.* GRADE guidelines: 8. Rating the quality of evidence—indirectness. *J Clin Epidemiol* 2011; 64: 1303–1310.
- 27 Guyatt GH, Oxman AD, Kunz R, *et al.* GRADE guidelines 6. Rating the quality of evidence—imprecision. *J Clin Epidemiol* 2011; 64: 1283–1293.
- 28 Guyatt G, Oxman AD, Akl EA, *et al.* GRADE guidelines: 1. Introduction—GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol* 2011; 64: 383–394.
- 29 Shingai K, Matsuda T, Kondoh Y, *et al.* Cutoff points for step count to predict 1-year all-cause mortality in patients with idiopathic pulmonary fibrosis. *Respir Int Rev Thorac Dis* 2021; 100: 1151–1157.
- 30 Wijsenbeek MS, Bendstrup E, Valenzuela C, *et al.* Disease behaviour during the peri-diagnostic period in patients with suspected interstitial lung disease: the STARLINER study. *Adv Ther* 2021; 38: 4040–4056.
- 31 Alexandre HF, Cani KC, Araújo J, *et al.* Reliability and validity of the Glittre-ADL test to assess the functional status of patients with interstitial lung disease. *Chron Respir Dis* 2021; 18: 14799731211012962.
- 32 Hirabayashi R, Takahashi Y, Nagata K, *et al.* The validity and reliability of four-meter gait speed test for stable interstitial lung disease patients: the prospective study. *J Thorac Dis* 2020; 12: 1296–1304.
- 33 Montgomery E, Macdonald PS, Newton PJ, *et al.* Frailty as a predictor of mortality in patients with interstitial lung disease referred for lung transplantation. *Transplantation* 2020; 104: 864–872.
- 34 Nathan SD, Flaherty KR, Glassberg MK, *et al.* A randomized, double-blind, placebo-controlled study of pulsed, inhaled nitric oxide in subjects at risk of pulmonary hypertension associated with pulmonary fibrosis. *Chest* 2020; 158: 637–645.
- 35 Wallaert B, Kyheng M, Labreuche J, *et al.* Long-term effects of pulmonary rehabilitation on daily life physical activity of patients with stage IV sarcoidosis: a randomized controlled trial. *Respir Med Res* 2020; 77: 1–7.
- 36 Froidure S, Kyheng M, Grosbois JM, *et al.* Daily life physical activity in patients with chronic stage IV sarcoidosis: a multicenter cohort study. *Health Sci Rep* 2019; 2: e109.
- 37 Perez-Bogerd S, Wuyts W, Barbier V, *et al.* Short and long-term effects of pulmonary rehabilitation in interstitial lung diseases: a randomised controlled trial. *Respir Res* 2018; 19: 182.
- 38 Pilzak K, Żebrowska A, Sikora M, *et al.* Physical functioning and symptoms of chronic fatigue in sarcoidosis patients. *Adv Exp Med Biol* 2018; 1040: 13–21.
- 39 Bahmer T, Watz H, Develaska M, *et al.* Physical activity and fatigue in patients with sarcoidosis. *Respir Int Rev Thorac Dis* 2018; 95: 18–26.
- 40 Capparelli I, Fernandez M, Saadia Otero M, *et al.* Translation to Spanish and validation of the specific Saint George's questionnaire for idiopathic pulmonary fibrosis. *Arch Bronconeumol* 2018; 54: 68–73.
- 41 Nishiyama O, Yamazaki R, Sano H, *et al.* Physical activity in daily life in patients with idiopathic pulmonary fibrosis. *Respir Investig* 2018; 56: 57–63.
- 42 Morino A, Takahashi H, Chiba H, *et al.* Daily physical activity affects exercise capacity in patients with idiopathic pulmonary fibrosis. *J Phys Ther Sci* 2017; 29: 1323–1328.
- 43 Atkins C, Baxter M, Jones A, *et al.* Measuring sedentary behaviors in patients with idiopathic pulmonary fibrosis using wrist-worn accelerometers. *Clin Respir J* 2018; 12: 746–753.
- 44 Dale MT, McKeough ZJ, Munoz PA, *et al.* Physical activity in people with asbestos related pleural disease and dust-related interstitial lung disease: an observational study. *Chron Respir Dis* 2015; 12: 291–298.
- 45 Nakayama M, Bando M, Araki K, *et al.* Physical activity in patients with idiopathic pulmonary fibrosis. *Respirol Carlton Vic* 2015; 20: 640–646.
- 46 Keyser RE, Christensen EJ, Chin LMK, *et al.* Changes in fatigability following intense aerobic exercise training in patients with interstitial lung disease. *Respir Med* 2015; 109: 517–525.
- 47 Ryerson CJ, Cayou C, Topp F, *et al.* Pulmonary rehabilitation improves long-term outcomes in interstitial lung disease: a prospective cohort study. *Respir Med* 2014; 108: 203–210.
- 48 Wickerson L, Mathur S, Helm D, *et al.* Physical activity profile of lung transplant candidates with interstitial lung disease. *J Cardiopulm Rehabil Prev* 2013; 33: 106–112.
- 49 Jarosch I, Schneeberger T, Gloeckl R, *et al.* Short-term effects of comprehensive pulmonary rehabilitation and its maintenance in patients with idiopathic pulmonary fibrosis: a randomized controlled trial. *J Clin Med* 2020; 9: E1567.
- 50 Nolan CM, Patel S, Barker RE, *et al.* Muscle stimulation in advanced idiopathic pulmonary fibrosis: a randomised placebo-controlled feasibility study. *BMJ Open* 2021; 11: e048808.
- 51 King CS, Flaherty KR, Glassberg MK, *et al.* A phase-2 exploratory randomized controlled trial of inopulse in patients with fibrotic interstitial lung disease requiring oxygen. *Ann Am Thorac Soc* 2022; 19: 594–602.
- 52 Cerdan de las Heras J, Balbino F, Catalán-Matamoros D, *et al.* Effect of a telerehabilitation program in sarcoidosis. *Sarcoidosis Vasc Diffuse Lung Dis* 2022; 39: e2022003.
- 53 Labrecque P-FT, Dion G, Saey D. Functional clinical impairments and frailty in interstitial lung disease patients. *ERJ Open Res* 2022; 8: 00144-02022.

- 54 Faverio P, Fumagalli A, Conti S, *et al.* Nutritional assessment in idiopathic pulmonary fibrosis: a prospective multicentre study. *ERJ Open Res* 2022; 8: 00443-02021.
- 55 Breuls S, Pereira de Araujo C, Blondeel A, *et al.* Physical activity pattern of patients with interstitial lung disease compared to patients with COPD: a propensity-matched study. *PLoS One* 2022; 17: e0277973.
- 56 Badenes-Bonet D, Rodó-Pin A, Castillo-Villegas D, *et al.* Predictors and changes of physical activity in idiopathic pulmonary fibrosis. *BMC Pulm Med* 2022; 22: 340.
- 57 Chu B, O'Connor DM, Wan M, *et al.* Quality of life and physical activity in 629 individuals with sarcoidosis: prospective, cross-sectional study using smartphones (Sarcoidosis App). *JMIR MHealth UHealth* 2022; 10: e38331.
- 58 Cerdán-de-las-Heras J, Balbino F, Løkke A, *et al.* Tele-rehabilitation program in idiopathic pulmonary fibrosis —a single-center randomized trial. *Int J Environ Res Public Health* 2021; 18: 10016.
- 59 Aguiar WF, Mantoani LC, Silva H, *et al.* Translation, cross-cultural adaptation, and measurement properties of the Brazilian-Portuguese version of the idiopathic pulmonary fibrosis-specific version of the Saint George's Respiratory Questionnaire (SGRQ-I) for patients with interstitial lung disease. *Braz J Phys Ther* 2021; 25: 794–802.
- 60 Hur SA, Guler SA, Khalil N, *et al.* Impact of psychological deficits and pain on physical activity of patients with interstitial lung disease. *Lung* 2019; 197: 415–425.
- 61 Vainshelboim B, Fox BD, Kramer MR, *et al.* Short-term improvement in physical activity and body composition after supervised exercise training program in idiopathic pulmonary fibrosis. *Arch Phys Med Rehabil* 2016; 97: 788–797.
- 62 Dale MT, McKeough ZJ, Munoz PA, *et al.* Exercise training for asbestos-related and other dust-related respiratory diseases: a randomised controlled trial. *BMC Pulm Med* 2014; 14: 180.
- 63 Bahmer T, Kirsten A-M, Waschki B, *et al.* Prognosis and longitudinal changes of physical activity in idiopathic pulmonary fibrosis. *BMC Pulm Med* 2017; 17: 104.
- 64 Fettes L, Bayly J, de Bruin LM, *et al.* Relationships between prolonged physical and social isolation during the COVID-19 pandemic, reduced physical activity and disability in activities of daily living among people with advanced respiratory disease. *Chron Respir Dis* 2021; 18: 14799731211035822.
- 65 Yoshida C, Ichiyasu H, Ideguchi H, *et al.* Four-meter gait speed predicts daily physical activity in patients with chronic respiratory diseases. *Respir Investig* 2019; 57: 368–375.
- 66 Layton AM, Armstrong HF, Baldwin MR, *et al.* Frailty and maximal exercise capacity in adult lung transplant candidates. *Respir Med* 2017; 131: 70–76.
- 67 DePew ZS, Karpman C, Novotny PJ, *et al.* Correlations between gait speed, 6-minute walk distance, physical activity, and self-efficacy in patients with severe chronic lung disease. *Respir Care* 2013; 58: 2113–2119.
- 68 Peddle-McIntyre CJ, Cavalheri V, Boyle T, *et al.* A review of accelerometer-based activity monitoring in cancer survivorship research. *Med Sci Sports Exerc* 2018; 50: 1790–1801.
- 69 Vetrovsky T, Clark CCT, Bisi MC, *et al.* Advances in accelerometry for cardiovascular patients: a systematic review with practical recommendations. *ESC Heart Fail* 2020; 7: 2021–2031.
- 70 Burtin C, Mohan D, Troosters T, *et al.* Objectively measured physical activity as a COPD clinical trial outcome. *Chest* 2021; 160: 2080–2100.
- 71 Furlanetto KC, Bisca GW, Oldenberg N, *et al.* Step counting and energy expenditure estimation in patients with chronic obstructive pulmonary disease and healthy elderly: accuracy of 2 motion sensors. *Arch Phys Med Rehabil* 2010; 91: 261–267.
- 72 Van Remoortel H, Raste Y, Louvaris Z, *et al.* Validity of six activity monitors in chronic obstructive pulmonary disease: a comparison with indirect calorimetry. *PLoS One* 2012; 7: e39198.
- 73 Rabinovich RA, Louvaris Z, Raste Y, *et al.* Validity of physical activity monitors during daily life in patients with COPD. *Eur Respir J* 2013; 42: 1205–1215.
- 74 Demeyer H, Mohan D, Burtin C, *et al.* Objectively measured physical activity in patients with COPD: recommendations from an international task force on physical activity. *Chronic Obstr Pulm Dis* 2021; 8: 528–550.
- 75 Garcia-Aymerich J, Puhan MA, Corriol-Rohou S, *et al.* Validity and responsiveness of the daily- and clinical visit-PROactive physical activity in COPD (D-PPAC and C-PPAC) instruments. *Thorax* 2021; 76: 228–238.
- 76 Tudor-Locke-Catrine C, Craig CL, Thyfault JP, *et al.* A step-defined sedentary lifestyle index: <5000 steps/day. *Appl Physiol Nutr Metab* 2012; 38: 100–114.
- 77 John D, Morton A, Arguello D, *et al.* “What is a step?” Differences in how a step is detected among three popular activity monitors that have impacted physical activity research. *Sensors* 2018; 18: 1206.
- 78 Madden KM, Ashe MC, Lockhart C, *et al.* Sedentary behavior and sleep efficiency in active community-dwelling older adults. *Sleep Sci* 2014; 7: 82–88.
- 79 Migueles JH, Cadenas-Sanchez C, Rowlands AV, *et al.* Comparability of accelerometer signal aggregation metrics across placements and dominant wrist cut points for the assessment of physical activity in adults. *Sci Rep* 2019; 9: 18235.

- 80 Migueles JH, Cadenas-Sanchez C, Tudor-Locke C, *et al.* Comparability of published cut-points for the assessment of physical activity: implications for data harmonization. *Scand J Med Sci Sports* 2019; 29: 566–574.
- 81 Calabro MA, Kim Y, Franke WD, *et al.* Objective and subjective measurement of energy expenditure in older adults: a doubly labeled water study. *Eur J Clin Nutr* 2015; 69: 850–855.
- 82 Amagasa S, Inoue S, Murayama H, *et al.* Associations of sedentary and physically-active behaviors with cognitive-function decline in community-dwelling older adults: compositional data analysis from the NEIGE study. *J Epidemiol* 2020; 30: 503–508.
- 83 Wilson JJ, McMullan I, Blackburn NE, *et al.* Associations of sedentary behavior bouts with community-dwelling older adults' physical function. *Scand J Med Sci Sports* 2021; 31: 153–162.
- 84 Lefebvre C, Glanville J, Briscoe S, *et al.* Searching for and selecting studies. In: Higgins JPT, Thomas J, Chandler J, *et al.*, eds. *Cochrane Handbook for Systematic Reviews of Interventions*. 2nd Edn. Chichester, Wiley, 2019; pp. 67–107.
- 85 Dobrescu AI, Nussbaumer-Streit B, Klerings I, *et al.* Restricting evidence syntheses of interventions to English-language publications is a viable methodological shortcut for most medical topics: a systematic review. *J Clin Epidemiol* 2021; 137: 209–217.
- 86 Cuijpers P, Griffin JW, Furukawa TA. The lack of statistical power of subgroup analyses in meta-analyses: a cautionary note. *Epidemiol Psychiatr Sci* 2021; 30: e78.
- 87 Byrom B, Rowe DA. Measuring free-living physical activity in COPD patients: deriving methodology standards for clinical trials through a review of research studies. *Contemp Clin Trials* 2016; 47: 172–184.
- 88 Demeyer H, Burtin C, Van Remoortel H, *et al.* Standardizing the analysis of physical activity in patients with COPD following a pulmonary rehabilitation program. *Chest* 2014; 146: 318–327.
- 89 Bull FC, Al-Ansari SS, Biddle S, *et al.* World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *Br J Sports Med* 2020; 54: 1451–1462.
- 90 Zhou S-M, Hill RA, Morgan K, *et al.* Classification of accelerometer wear and non-wear events in seconds for monitoring free-living physical activity. *BMJ Open* 2015; 5: e007447.
- 91 Choi L, Ward SC, Schnelle JF, *et al.* Assessment of wear/nonwear time classification algorithms for triaxial accelerometer. *Med Sci Sports Exerc* 2012; 44: 2009–2016.