Coherence between self-reported and objectively measured physical activity in patients with chronic obstructive lung disease: a systematic review

Abstract: The beneficial effects of physical activity (PA) in patients with COPD, as well as the methods of their assessment, are well known and described. As objective measures of PA, such as the use of motion sensors, video recordings, exercise capacity testing, and indirect calorimetry, are not easily obtained in the daily clinical life, the reliability of the more accessible self-reported measurements of PA is important. In this review, we systematically identified original studies involving COPD patients and at least one parameter of self-reported and objective exercise testing, and analyzed every article for coherence between the objectively and self-reported measured PA. The studies are few, small, and very diverse, both in their use of questionnaires and objective measurements. Self-reported assessments were found to generally overestimate the level of PA compared to measurements made objectively by activity monitors; however, more studies are needed to rely solely on the use of PA questionnaires in COPD patients. The most accurate and valid questionnaires appear to be the self-completed Physical Activity Scale for the Elderly and the interviewer-completed Stanford Seven-Day Physical Activity Recall Questionnaire, but the ideal questionnaire still awaits construction. The motion sensors are accurate and validated in this patient group, especially SenseWear™, but not easily accessible in clinical practice, as they have various technical and adhesive difficulties.

Keywords: respiratory medicine, COPD, physical activity monitoring, questionnaire, motion sensor

Introduction

The level of PA in patients with COPD is inversely correlated with disease severity: lung function decline, disease burden, frequency of hospitalization, and mortality. PA is defined as any bodily movement produced by skeletal muscles that results in energy expenditure. In COPD patients, the amount of PA is reduced compared to matched controls, as a result of – among other things – decreased threshold for exercise-induced dyspnea due to the irreversible airflow obstruction. PA is efficacious in alleviating the vicious circle of deconditioning, and is the cornerstone of pulmonary rehabilitation.

In COPD, correct assessment of PA in daily life is crucial to allocate patients to correct interventions: self-training, tele-health-based training, community-based or hospital-based pulmonary rehabilitation. In addition, knowledge of PA levels can motivate patients to improve, thereby reducing the morbidity and mortality by maintaining a proper PA level. Unfortunately, the correct assessment of a patient’s PA levels is difficult, and no GOLD standard exists to quantify the amount of PA.

PA levels can be assessed objectively or subjectively (self-reported), prospectively or retrospectively, and in experimental or real-life settings. Exercise capacity tests
(objective/prospective experimental setting), eg, the six-minute walk test (6MWT) or the incremental shuttle walk test (ISWT), are routinely used in rehabilitation programs and then a patient’s daily PA levels are extrapolated from these tests, although a recent study showed a moderate-to-weak correlation to PA levels when the number of steps were measured by a pedometer. Other objective assessments often used in research settings or in athletes are the measurements of VO2 max and energy expenditure with indirect calorimetry (DLW method), which are considered as golden standard methods. Yet, assessment of PA energy expenditure does not quantify the duration, frequency, and intensity of the PA performed. Video recordings or activity monitors are examples of objective, prospective assessments in real-life settings (see “Objective methods for quantification of daily PA” section below), and the importance of the latter has recently been elaborately reviewed by Liao et al. Self-reported PA is more easily accessible and cheaper in both clinical life and epidemiological studies compared to video recordings or activity monitors – although often not very reliable. The most commonly used methods are self-reported, retrospective questionnaires on PA, whereas daily entries in a PA diary are less often used. Few questionnaires have been validated in COPD populations, which are characterized by much lower levels of PA compared to the general population (see “Objective methods for quantification of daily PA” section below).

Therefore, we aimed at performing a systematic review to investigate the agreement between self-reported and objectively measured PA in daily life, as well as the reliability and validity of each method and most importantly, the clinical application thereof. The following two sections describe the different methods followed for both subjective/self-reported and objective assessments.

Objective methods for quantification of daily PA

The Tuscan Renaissance artist and inventor Leonardo da Vinci drew the first sketches of a motion sensor in the 15th century: a pedometer held by both hands and looking like a wheelbarrow. Today, motion sensors are body-worn and divided into three categories: accelerometer (uniaxial, biaxial, triaxial), multisensor device, and pedometer. Accelerometers detect activity through changes in velocity as well as frequency, intensity, and duration of the activity. Their cost and technical complexity are higher than regular pedometers that only collect steps by vertical accelerations. The multisensor device combines the triaxial accelerometer with sensors (heat flux, galvanic skin response and skin temperature). Within each category, motion sensors differ in economic costs, technical complexity, and placement on the body (hip, arm, ankle, etc.). This is often a practical challenge in patients using walking aids, such as a walking stick or a walker, in addition to the use of motion systems in different environments such as in the water to record swimming or bicycling as examples.

Questionnaires for quantification of daily PA

Self-reported measurements include diaries, logbooks, and questionnaires. Overall, questionnaires are classified into three categories: global, recall, and quantitative. The global questionnaire is short, self-administered, and usually consists of 2–4 questions, and can be used to assess whether a patient fulfills the international daily PA requirement of 30 minutes. The short recall questionnaire commonly includes 7–12 items and can be either interviewer- or self-administered. The quantitative questionnaire is longer, typically includes 20–60 questions and provides in greater detail the PA level over a longer period of time.

The following validated questionnaires were used in the papers included in the study: PASE, PAR, Baechle Physical Activity Questionnaire, ZPAQ, and MARQ. Other validated questionnaires for the COPD population include Yale Physical Activity Survey and Minnesota Leisure Time Physical Activity Questionnaire, among others. A summary of the questionnaires featured in this review is presented in Table 1. The questionnaires are divided into four categories: multi- or single-item questionnaire, and validated or nonvalidated questionnaire in COPD populations.

Methods

A literature search was performed using PubMed on March 5, 2016, using the terms “Physical activity”, “Monitoring” or “Measuring”, “COPD”, “Self-reported”, “objective”, and “correlation” or “coherence”. We had no limitations concerning the publication year. Inclusion criteria were the following: 1) studies including patients with known COPD, 2) original studies involving PA, 3) objective measurement of the level of daily PA, 4) self-reported measurement of the level of daily PA, 5) peer-reviewed papers published in English. Exclusion criteria were the following: 1) studies including patients with diagnoses other than COPD, and without clear data separation between COPD and non-COPD groups and 2) studies reporting only subjective or objective PA measurement.

All identified abstracts were independently screened by two persons (MT and UB) and discussed with an experienced pulmonologist in COPD (Dr Anders Løkke) in case of disagreement.
Table 1 Overview of the questionnaires

<table>
<thead>
<tr>
<th>Study</th>
<th>Questionnaire</th>
<th>Administration form</th>
<th>No of questions</th>
<th>Recall time</th>
<th>Classification of physical activity</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>PAR</td>
<td>Interviewer</td>
<td>4 including complete minute to minute diary</td>
<td>7 days</td>
<td>Moderate, hard, very hard</td>
<td>Minutes</td>
</tr>
<tr>
<td>18</td>
<td>Baecke</td>
<td>Self</td>
<td>16</td>
<td>12 months</td>
<td>Light, moderate, hard</td>
<td>Minutes</td>
</tr>
<tr>
<td>18,18</td>
<td>Zutphen</td>
<td>Self</td>
<td>17,32</td>
<td>1 month, 7 days</td>
<td>Light, moderate, very hard, recreation</td>
<td>Hours -&gt; score</td>
</tr>
<tr>
<td>17</td>
<td>SBAS</td>
<td>Self</td>
<td>2</td>
<td>1 day</td>
<td>Inactive, light, moderate, hard, very hard</td>
<td>Score</td>
</tr>
<tr>
<td>19</td>
<td>MARQ</td>
<td>Self</td>
<td>–</td>
<td>–</td>
<td>All activities</td>
<td>Energy expenditure</td>
</tr>
<tr>
<td>23</td>
<td>SQUASH</td>
<td>Self</td>
<td>13</td>
<td>7 days</td>
<td>Light, moderate, intense</td>
<td>Hours</td>
</tr>
</tbody>
</table>

Abbreviations: MARQ, Modified Activity Recall Questionnaire; PAR, the Stanford Seven-Day Physical Activity Recall Questionnaire; PASE, Physical Activity Scale for the Elderly; SBAS, Stanford Brief Activity Scale; SQUASH, Short Questionnaire to Assess Health-enhancing physical activity.

Important definitions

SPI: the PAL is obtained by dividing the total daily energy expenditure by the resting energy expenditure, measured by an activity monitor. SPI is defined as PAL <1.40.17

MET: a physiological measure expressing the energy cost of PA, as measured by an activity monitor. MET score is defined as follows: <3 light activities, 3–6 moderate activities, and >6 vigorous activities.18

Outcome

Primary outcome: difference in PA. Self-reported PA as a fraction of measured PA, possibly reported in the original study as Pearson’s correlation.

Secondary outcomes: Percentage of SPI by questionnaire compared to objective measurements.

Results

The search process resulted in identification of 39 abstracts, and after review, 7 original papers (Figure 1) were selected for the study. During the review process, most studies were discarded due to a lack of either an objective or a subjective measure of self-reported PA (n=30). A few studies were published in non-English languages (n=2). Table 2 depicts differences in sample size, PA and methods for objective and subjective measurements of PA in the various studies. All included papers involved motion sensors. The study could not find any relevant RCTs on the subject, only cohort studies.

Risk of bias

According to the NOS, we assessed the quality of the included studies, as given in Table 3.22 Agreement between the reviewers (MT and UB) was calculated and expressed as a percentage and quantified by kappa statistics; the results revealed an agreement of 81% and a kappa value of 0.74.

Validated multi-item questionnaire

In an English study, PAR was compared with a SenseWear™ (Bodymedia, Pittsburgh, PA, USA) multisensor and validated using three other validated PA questionnaires: Baecke, PASE, and ZPAQ (Table 2).13,18

The study included 43 COPD patients who wore SenseWear™ for 7 days. A significant correlation between

---

Figure 1 Flow diagram of search results.
### Table 2: Overview of included studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Number</th>
<th>COPD population: FEV₁, %</th>
<th>Accelerometer</th>
<th>Placement</th>
<th>Type</th>
<th>Questionnaire</th>
<th>Self-reported/measured PA, correlation coefficient (R)</th>
<th>Ratio self-reported/measured PA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pitta et al⁷⁷</td>
<td>Cross sectional</td>
<td>Study A = 10 Study B = 13</td>
<td>40±16</td>
<td>Dynaport activity monitor</td>
<td>3 sensors, two worn on the waist and one on the left leg</td>
<td>Accelerometer, triaxial</td>
<td>Diary</td>
<td>0.71 (P&lt;0.01)</td>
<td>&gt;1</td>
</tr>
<tr>
<td>Singh et al⁹⁶</td>
<td>Cross sectional</td>
<td>11 COPD, 9 healthy</td>
<td>1.16±0.56 L (no other data available)</td>
<td>Z80-32k V1 INT (Gaehwiler Electronics, Hombrechtikon, Switzerland)</td>
<td>Waist</td>
<td>Uniaxial</td>
<td>Diary</td>
<td>--</td>
<td>&gt;1</td>
</tr>
<tr>
<td>Garfield et al¹⁰</td>
<td>Cross sectional</td>
<td>43</td>
<td>46±22</td>
<td>Sensewear</td>
<td>Left upper arm</td>
<td>Multisensor</td>
<td>PAR, PASE, Baecke, Zutphen, diary</td>
<td>0.54 (P&lt;0.001)</td>
<td>&gt;1</td>
</tr>
<tr>
<td>Steele et al¹⁹</td>
<td>Cross sectional</td>
<td>47</td>
<td>37±16</td>
<td>Tritac R3D</td>
<td>Hip</td>
<td>Multisensor</td>
<td>MARQ</td>
<td>0.14 (NS)</td>
<td>&gt;1</td>
</tr>
<tr>
<td>DePew et al¹⁷</td>
<td>Cross sectional</td>
<td>67</td>
<td>46.7±19.98</td>
<td>Sensewear</td>
<td>Left upper arm</td>
<td>Multisensor</td>
<td>PASE, SBAS</td>
<td>0.62 (P&lt;0.001)</td>
<td>&gt;1</td>
</tr>
<tr>
<td>Moore et al²⁴</td>
<td>Cross sectional</td>
<td>76</td>
<td>46.3±18.5</td>
<td>Yamax Digiwalker sw-700</td>
<td>Hip</td>
<td>Pedometer</td>
<td>Diary</td>
<td>0.37 (P=0.001)</td>
<td>&gt;1</td>
</tr>
<tr>
<td>Bossenbroek et al²²</td>
<td>Cross sectional</td>
<td>15 vs 47</td>
<td>Candidates: 20±5.2 Recipients: 93±28.4</td>
<td>Yamax Digiwalker sw-200</td>
<td>Waist</td>
<td>Pedometer</td>
<td>SQUASH</td>
<td>--</td>
<td>&gt;1</td>
</tr>
</tbody>
</table>

**Abbreviations:** FEV₁, forced expiratory volume in one second; MARQ, Modified Activity Recall Questionnaire; NS, non significant; PA, physical activity; PAR, the Stanford Seven-Day Physical Activity Recall Questionnaire; PASE, Physical Activity Scale for the Elderly; SBAS, Stanford Brief Activity Scale; SQUASH, Short Questionnaire to Assess Health-enhancing physical activity.
the TEE derived by PAR and SenseWear™ was observed in patients with moderate PA (MET ≥3, r=0.83). Furthermore, PAR correctly predicted patients at each end of the PA scale as very inactive and very active patients. In addition, the patients completed a 6MWT that did not predict the active patients, according to the PA detected by SenseWear™.

Likewise, DePew et al applied both the PASE questionnaire and SenseWear™. PASE demonstrated a moderate correlation with the PAL measured by SenseWear™ in 67 COPD patients with a positive correlation coefficient (r=0.62). Using a cut-off PASE score of <111, it predicted SPI measured by SenseWear™, with a negative predictive value of 0.65 and a positive predictive value of 0.83, indicating an underestimation of SPI utilizing a PASE questionnaire.

The best predictor of SPI was the mMRC ≥3, although PASE and mMRC together achieved an even better correlation (concordance correlation coefficient 0.833).

Another study compared the MARQ, a self-reported PA questionnaire which assesses energy expenditure, with the TriTrac-R3D accelerometer in 47 COPD patients. No significant relationship was found between reported and measured PA, and the low correlation (r=0.4) suggests profound overestimation of PA measured by the MARQ. In the same study, PA observed during 6MWT correlated significantly with that measured by TriTrac-R3D (r=0.74).

**Nonvalidated single-item score**

A similar study was carried out by Moore et al in 76 COPD patients using a pedometer and a simple 7-day PA diary. A significant, but poor or moderate correlation (r=0.37) was demonstrated using a modified diary developed by Follick et al. It consisted of four categories of activities (sleeping, sitting, standing/walking, lying) and was completed a minimum of three times per day. The authors concluded that the pedometer used underestimated PA levels, and that the diary was a more accurate measure of PA, yet no other objective measure was established to confirm this assertion. A small study with 11 COPD patients and 9 healthy age-matched controls investigated if brisk walking was detected by an accelerometer worn for 48 hours. The COPD patients had recently completed pulmonary rehabilitation, and brisk walking was defined as 60% of the predicted peak oxygen consumption in ISWT. No data on how brisk walking was defined in the healthy age-matched controls were provided. The COPD patients recorded daily PA in 10 different categories: lying, sitting, sitting activity, standing, standing activity, personal, walking (brisk), walking (slow/intermittent), driving, and others. There was a significant correlation between the diary recorded brisk walking and that counted by the accelerometer, but there was a wide range in the standard deviation scores in all other categories. No results on whether the diary over- or underestimated PA according to the accelerometer were provided in the paper, but step count differed insignificantly between lying and all activities, except for brisk walking—suggesting overestimation of self-reported PA.

Pitta et al performed two unrelated studies: A and B. In study A, a small group of 10 patients with COPD correlated
patient estimations, DAM, and video recordings in a 1-hour standardized protocol, whereas study B, consisting of 13 other COPD patients, correlated DAM, and a self-report in a regular-day protocol. The self-report consisted of a logbook, reporting on an hourly basis the time spent on each activity or body position. In study A, patients significantly underestimated sitting time in the self-report. DAM and video recordings were correlated, even though the DAM systematically slightly overestimated walking and cycling time. There was a high disagreement between the self-report and the video recordings. In study B, patients significantly overestimated walking time \((r=0.71)\) and underestimated standing time in the self-report compared to the DAM.

**Discussion**

The review unveils the existence of few and heterogeneous studies using accelerometers and self-reported PA simultaneously, resulting in a low validity and test–retest reliability. Heterogeneity was observed in patient selection (ranging from mild to very severe COPD), intensity, and duration of PA. Steele et al confirmed the correlation between FEV\(_1\) and PA, as measured by the accelerometer and the 6MWT, highlighting the relationship between airway obstruction and maximum exercise capacity.\(^{19,23,28}\) A general overestimation of self-reported PA was demonstrated, and no reliable PA questionnaire was found.\(^{17,19,27,29}\) The most accurate tool for self-reported PA based on the papers assessed here is to keep a daily diary with a description of all activities.\(^{24}\)

However, two questionnaires predicted very low levels of daily PA: PASE and PAR. The PASE questionnaire was able to predict SPI in COPD patients, although, together with mMRC, it had a more accurate outcome.\(^{17}\) The PAR questionnaire was the only questionnaire to reach significant correlation with SenseWear\textsuperscript{TM}.\(^{15}\) However, this finding was only observed in the subgroup of COPD patients reaching a moderate PA level (MET \(\geq 3\)). The authors concluded that PAR predicted PA only as a binary outcome (active/inactive) with respect to PA levels. Since PAR was an interviewer-led questionnaire, it was difficult to use this in an epidemiological setting.

Various challenges are associated with PA recollection in COPD patients. An accurate PA questionnaire in the, mostly sedentary, COPD population should include various questions on the reasons for low PA. The PASE questionnaire was designed for an elderly population, whereas the validation of the SQUASH questionnaire was done in a subgroup including mostly men (70%), with a mean age of 44 years, and 50% had a good or high fitness level, which is far from that expected for the COPD population.\(^{17,30}\)

Questionnaires are either interviewer administered or self-completed by the patient. Of the questionnaires included in this review, only PAR was interviewer administered, with the risk of over-administration the answers. On the other hand, there might be a risk of missing data when using self-administered questionnaires, as well as misunderstanding of both content- and time-related questions.

Cognitive impairment in COPD has gained increasing attention.\(^{31}\) Recollecting the details of daily life activities, including PA, is evidently hampered in the most diseased COPD patients, and therefore reliability of the data collected has been questioned.\(^{27,32,33}\) The health benefits of normal PA are widely accepted.\(^{5,6,13}\) These influence the patients’ self-reporting measurements, resulting in a tendency to overestimate PA levels, which is found in this review as well.\(^{17,19,27,29,33}\)

Recall bias is also affected by time factors, such as when and how the questionnaire is to be completed, and the amount of time the patient has to recall. The Baecke questionnaire covers the preceding year, and previously we have shown that otherwise healthy patients with seasonal pollen allergy only recollected the single day with the most symptoms 4 months after the pollen season. Thus, recollection of PA during an entire year is unlikely to be reliable or correct.\(^{17,34}\)

A significant dropout rate is also a well-known problem in the COPD population, making it even harder to generalize the results to the entire COPD population as only the healthiest patients complete each study.\(^{28}\)

Accelerometers have very different ways of calculating and reporting energy expenditure and PA, and it is mandatory to convert the reported output for comparison with that of questionnaire-based self-reported PA, work capacity testing, or other activity monitors. A device’s test–retest reliability and validity is typically derived in a non-COPD population. Measuring PA in the relatively inactive COPD population results in a rather narrow PA range as most time is spent sedentary with a low PA level concerning both duration and intensity.\(^{12}\) This results in a demand for highly sensitive activity monitors that can detect differences in PA even in very inactive patients, including a measurement of PA that does not involve total body movement, such as weight lifting and stationary cycling.\(^{19,29}\)

In addition to the challenge of recollecting PA, Pitta et al reported adherence difficulties upon using an accelerometer in 19% of COPD subjects including body placement and technical issues, such as battery problems.\(^{29}\) Additionally, Moore et al demonstrated that 18% of COPD patients failed to use a simple pedometer correctly; furthermore, 5% completed the diary insufficiently.\(^{24}\) These failure rates are probably higher.
in unselected, everyday patients who do not feel obliged to please an investigator during a clinical trial.

The limitations of this study are the lack of high-quality, high-volume studies and not all the included studies published coherence between self-reported and objectively measured PA. Only two studies included >50 patients. This review unveils a critical need for a more accurate PA questionnaire in the COPD population, as objective measures are costly, time-consuming and prone to adherence difficulties as well as patient-related issues.

Conclusion
The PA level has an impact on both prognosis of the patient and the disease severity. Assessment of daily PA is warranted to optimize PA-directed intervention. In the surprisingly few and small studies available on measurements of PA in daily life, self-reported assessments generally overestimated the level of PA measured objectively by activity monitors. However, the level of evidence is hampered by significant inter-study heterogeneity precluding statistical evaluation by meta-analysis. The most accurate and valid questionnaires appear to be the self-completed PASE and the interviewer-completed PAR questionnaires, but the ideal questionnaire still awaits construction. Establishment of a well-defined golden standard of daily-life PA will improve the efforts in the development of a simple, accurate, and valid measurement of daily-life PA.

Abbreviations
PA, physical activity; COPD, chronic obstructive pulmonary disease; 6MWT, six-minute walking test; ISWT, incremental shuttle walk test; VO₂ max, maximum oxygen consumption; DLW, double labeled water; PASE, Physical Activity Scale for the Elderly; PAR, the Stanford Seven-Day Physical Activity Recall Questionnaire; ZPAQ, Zutphen Physical Activity Questionnaire; MET, metabolic equivalent of task; MARQ, Modified Activity Recall Questionnaire; SPI, severe physical inactivity; TEE, total energy expenditure; PAL, physical activity level; NOS, Newcastle-Ottawa Scale; SQUASH, Short QUestionnaire to ASsess Health-enhancing physical activity; DAM, DynaPort Activity Monitor; FEV₁, forced expiratory volume in one second; GOLD, Global Initiative for Chronic Obstructive Lung Disease; mMRC, Modified Medical Research Council Dyspnea Scale; RCT, randomized controlled trial.

Acknowledgment
The authors gratefully thank Anders Lokke, MD, Department of Respiratory Medicine, Aarhus University Hospital, Denmark, for agreeing to settle possible disagreement between the authors during the review process.

Disclosure
The authors report no conflicts of interest in this work.

References


