Article:
Megaritis, D., Long, M., de las Heras, M. et al. (2026) The construct validity of real-world digital mobility outcomes in people with chronic obstructive pulmonary disease. ERJ Open Research. ISSN: 2312-0541

https://doi.org/10.1183/23120541.00993-2025

ERJ open research
THE BEST IN OPEN ACCESS BASIC, TRANSLATIONAL & CLINICAL RESPIRATORY RESEARCH

Early View

Original Research Article

The Construct Validity of Real-World Digital Mobility Outcomes in People with Chronic Obstructive Pulmonary Disease

Dimitrios Megaritis, Michael Long, Martí de las Heras, Victoria Alcaraz-Serrano, Paula Alvarez, Clemens Becker, Julia Braun, Joren Buekers, Sara Buttery, Brian Caulfield, Andrea Cereatti, Nikolaos Chynkiamis, Silvia Del Din, Laura Delgado-Ortiz, Heleen Demeyer, Anja Frei, Elena Gimeno-Santos, Nicholas S Hopkinson, Anisoara Ionescu, Carl-Philipp Jansen, Alícia Josa-Culleré, Anne Kirsten, Sarah Koch, Jorge Lemos, Keir EJ Philip, Lynn Rochester, Basil Sharrack, David Singleton, Beatrix Vereijken, Ioannis Vogiatzis, Henrik Watz, Vita Lanfranchi, Thierry Troosters, Judith Garcia-Aymerich

Please cite this article as: Megaritis D, Long M, de las Heras M, *et al*. The Construct Validity of Real-World Digital Mobility Outcomes in People with Chronic Obstructive Pulmonary Disease. *ERJ Open Res* 2026; in press (https://doi.org/10.1183/23120541.00993-2025).

The Construct Validity of Real-World Digital Mobility Outcomes in People with Chronic Obstructive Pulmonary Disease

Dimitrios Megaritis*[1,2], Michael Long*[3], Martí de las Heras[4,5,6], Victoria Alcaraz-Serrano[4,5,6,7], Paula Alvarez[4,5,6], Clemens Becker[8,9], Julia Braun[10], Joren Buekers[4,5,6], Sara Buttery[11], Brian Caulfield[12,13], Andrea Cereatti[14], Nikolaos Chynkiamis[15,16], Silvia Del Din[2,17], Laura Delgado-Ortiz[4,18], Heleen Demeyer[19,20], Anja Frei[10], Elena Gimeno-Santos[4,5,6], Nicholas S Hopkinson[11], Anisoara Ionescu[21], Carl-Philipp Jansen[8,9], Alícia Josa-Culleré[4,5,6], Anne Kirsten[22], Sarah Koch[4,5,6,18], Jorge Lemos[4,5,6], Keir EJ Philip[11], Lynn Rochester[2,17,23], Basil Sharrack[24], David Singleton[12,13], Beatrix Vereijken[25], Ioannis Vogiatzis[1], Henrik Watz[22], Vita Lanfranchi[3], Thierry Troosters[19,26], Judith Garcia-Aymerich[4,5,6]

1. Faculty of Health and Life Sciences, Northumbria University Newcastle, Newcastle upon Tyne, UK
2. Translational and Clinical Research Institute, Newcastle University, Newcastle upon Tyne, United Kingdom
3. School of Computer Science, The University of Sheffield, Sheffield, UK
4. Barcelona Institute for Global Health (ISGlobal), Barcelona, Spain
5. Universitat Pompeu Fabra (UPF), Barcelona, Spain
6. CIBER Epidemiología y Salud Pública (CIBERESP), Madrid, Spain
7. School of Health Sciences, Ramon Llull University, Barcelona, Spain
8. Robert Bosch Foundation for Medical Research, Stuttgart, Germany
9. Geriatric Center, Medical Faculty Heidelberg, Heidelberg University, Heidelberg, Germany
10. Epidemiology, Biostatistics and Prevention Institute, University of Zurich, Switzerland
11. National Heart and Lung Institute, Imperial College London, London, England
12. Insight Centre for Data Analytics, University College Dublin, Dublin, Ireland
13. School of Public Health, Physiotherapy and Sports Science, University College Dublin, Dublin, Ireland
14. Department of Electronics and Telecommunications, Politecnico di Torino, Turin, Italy
15. Rehabilitation Unit, 1st University Department of Respiratory Medicine, "Sotiria" Hospital, Medical School, National and Kapodistrian University of Athens, Greece
16. Thorax Research Foundation, Athens, Greece
17. National Institute for Health and Care Research (NIHR) Newcastle Biomedical Research Centre (BRC), Newcastle University
18. Department of Sport, Exercise and Health, University of Basel, Switzerland
19. Department of Rehabilitation Sciences, KU Leuven, Leuven, Belgium
20. Department of Rehabilitation Sciences, Ghent University, Gent, Belgium
21. Laboratory of Movement Analysis and Measurement, Ecole Polytechnique Federale de Lausanne, Lausanne, Switzerland.
22. Velocity Clinical Research Germany, formerly Pulmonary Research Institute at Lungen Clinic Grosshansdorf, Airway Research Center North (ARCN), German Center for Lung Research (DZL), Grosshansdorf, Germany
23. The Newcastle upon Tyne NHS Foundation Trust, Newcastle upon Tyne, United Kingdom
24. Department of Neuroscience and Sheffield NIHR Translational Neuroscience BRC, Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK
25. Department of Neuromedicine and Movement Science, Norwegian University of Science and Technology, Trondheim, Norway.
26. Pulmonary Rehabilitation, University Hospital Leuven, Leuven, Belgium

(*) shared first authorship

Corresponding Author:
Professor Judith Garcia-Aymerich
ISGlobal, Barcelona Institute for Global Health
Doctor Aiguader, 88, 08003 Barcelona, Spain
judith.garcia@isglobal.org

word count: 3,659

Abstract

Background: Recent advances in wearable technologies make it possible to accurately quantify real-world mobility performance through technically validated digital mobility outcomes (DMOs).

Aim: To evaluate the construct validity (convergent, divergent, and known-groups validity) of 24 DMOs quantifying walking activity (amount and pattern) and gait (pace, rhythm, and bout-to-bout variability) in people with chronic obstructive pulmonary disease (COPD).

Methods: Part of the Mobilise-D observational cohort study, people with COPD, recruited from seven European sites, wore an activity monitor for seven days during daily life. Functional capacity, health status, dyspnoea, lung function, quadriceps torque, and experience of difficulty with physical activity were used as constructs for convergent validity testing (Pearson/Spearman correlation coefficients). Diastolic blood pressure was used as an unrelated construct for divergent validity (criterion: $|r|<0.2$). Known-groups validity was evaluated across GOLD stages (I-IV), GOLD ABE, and mMRC dyspnoea grades (linear models with p-for-trend).

Results: 549 participants (37% females), had mean(SD) age of 68(8) years, post-bronchodilator FEV_1 54(20) %pred and 6-minute walk distance 416(119) m. Convergent validity was supported for the majority of DMOs (17 out of 24) with correlation coefficients meeting or exceeding the *a priori* hypotheses by clinical experts. All DMOs supported divergent validity. Twenty-two out of 24 DMOs successfully distinguished between disease severity groups. Expert consensus supported construct validity of 17 DMOs.

Conclusions: Construct validity was supported for all walking activity (amount and pattern) DMOs, and most of the gait (pace, rhythm, and bout-to-bout variability) DMOs, indicating the clinical utility of these measures.

Abstract word count: 244/250

Keywords: Pulmonary Disease, Chronic Obstructive; Epidemiology; Gait; Digital Mobility Outcomes; Wearable Devices, Accelerometry

Introduction

The assessment of mobility performance in people with chronic obstructive pulmonary disease (COPD) during daily life has predominantly been limited to the amount that they walk (e.g., steps/day, daily walking time)[1, 2]. While these are well-established prognostic markers in COPD[1, 3], they do not capture the full spectrum of mobility impairments present in this population[3-5]. Research in other long-term conditions[2] and our preliminary evidence in COPD[6-8] suggest that spatio-temporal mobility outcomes such as stride length, cadence, and walking speed provide additional insights, and are valid and sensitive to disease progression-induced mobility impairment. However, due to technical complexity, the validation of such outcomes has so far been limited to laboratory settings[2], hindering translation to the real-world.

Walking-related digital mobility outcomes (DMOs) are technically valid measurements of a person's mobility performance[9-11]. Wearable devices equipped with inertial measurement units enable the quantification of mobility[9, 12-14], allowing a detailed evaluation of real-world walking performance. The Mobilise-D project developed a processing pipeline to calculate DMOs encompassing both walking activity (in the domains of amount and pattern) and gait (in the domains of pace, rhythm and bout-to-bout variability) across diverse health conditions, including COPD[13]. The resulting DMOs exhibited good to excellent criterion validity in people with COPD[13, 14]. However, their construct validity–whether DMOs accurately reflect the specific aspects of health they are intended to reflect–has only been reported for walking activity amount, using analytical approaches relying on proprietary software without established criterion validity [1, 2]. Construct validity is still to be determined for novel DMOs such as pattern (e.g., walking bout (WB) characteristics), pace (e.g., walking speed), rhythm (e.g., cadence), and bout-to-bout variability (e.g., walking speed bout-to-bout variability).

Therefore, the aim of the present study was to assess the construct validity of real-world walking activity and gait DMOs derived from a recently proposed, open-source, and technically valid computational pipeline[9, 13] in a large, heterogeneous, international sample of people with COPD across Europe. We pre-specified a set of a priori hypotheses, based on pilot data and structured expert input, regarding expected correlations between DMOs and clinical constructs, in line with regulatory guidance that emphasises hypothesis-driven evaluation of construct validity for biomarkers and digital health technologies [15-17].

Methods

Study design

This cross-sectional analysis is part of the Mobilise-D Clinical Validation Study (ISRCTN:12051706), a multicentre observational cohort study aiming to validate novel DMOs in COPD, Parkinson's disease, multiple sclerosis, and proximal femur fracture[12]. This paper reports on COPD results from database version 7.0 of the study.

Participants

COPD participants were recruited between April 2021 and May 2022 in seven cities across six countries: Leuven (Belgium); Barcelona (Spain); Newcastle and London (United Kingdom); Zurich (Switzerland); Großhansdorf (Germany); and Athens (Greece). Individuals with diagnosis-confirmed COPD (post-bronchodilator forced expiratory volume in the first second(FEV_1)/forced vital capacity(FVC)<0.70[18], as per GOLD guidelines), smoking history of ≥10 pack-years, clinical stability (at least 4 weeks after onset of the last exacerbation prior to inclusion), able to walk 4 meters independently with or without walking aids, and willingness to wear a single wearable device for seven days, were eligible for inclusion. Individuals having substantial limitation in mobility due to factors unrelated to COPD, having a diagnosis of active lung cancer or primary respiratory disease other than COPD, having undergone major lung surgery or lung volume reduction surgery within 6 months prior to inclusion, were excluded as previously reported[12]. The study was approved by all local ethical committees and all participants provided written informed consent. Information on the baseline characteristics of the cohort were published previously[8, 19].

DMOs

Walking activity and gait were measured for seven days using one of two metrologically equivalent single wearable devices positioned on the lower back at the level of the lumbar spine (L4-5): the MoveMonitor+ (McRoberts B.V., The Hague, The Netherlands), worn on a belt around the waist, only removed for bathing/swimming; or the Axivity AX6 (Axivity Ltd, Newcastle Upon Tyne, UK), fixed to the participants' skin at the lower back using a patch. Both were equipped with a 3-axial accelerometer (±8g, 1mg resolution) and a 3-axial gyroscope (±2000dps, 70m°/s resolution), collecting data at 100Hz. We removed participants who did not have a minimum of >12 hours of daily wear time (07:00–22:00hours) across ≥3 days[20]. Using Mobilise-D algorithms, we first identified WBs (walking sequence containing at least two consecutive strides of both feet[21]), and obtained WB-level DMOs, which were previously validated against gold and silver standards for their criterion validity[13, 14]. We then calculated weekly-aggregated DMOs, as outlined previously[22]. A total of 24 DMOs were obtained, encompassing walking activity (amount and pattern) and gait (pace, rhythm, and bout-to-bout variability) domains. DMO definitions, and units have been previously published [8] and a concise description is presented in Table S2. Briefly, in the walking activity category, the amount domain captured daily walking duration and WB step count, while the pattern domain captured the daily number of WBs of different durations. In the gait category, the pace domain captured walking speed and stride length; the rhythm domain captured cadence and stride duration; and the bout-to-bout variability domain captured the variability of DMOs across WBs within a day. These DMOs were measured and reported in a diverse range of WB durations (all WBs, >10s, 10-30s, >30s). Daily DMO values were first computed as the average of all WB–level values within each day or, for peak metrics, as the 90th percentile of the daily WB values. Weekly values were then derived by averaging the corresponding daily values over seven consecutive days.

Constructs

Clinical constructs included: (i) functional exercise capacity, using the 6-minute walk test (6MWT), according to the ERS-ATS standard[23]; (ii) health status, measured using the COPD assessment test (CAT); (iii) dyspnoea, using the modified Medical Research Council (mMRC) scale; (iv) lung function, including post-bronchodilator FEV_1, and FVC according to the ERS standards[18]; (v) quadriceps maximum voluntary contraction (QMVC), using isokinetic dynamometer at 90° knee/hip angles (torque=force x limb length (Nm)); (vi) experienced difficulty with physical activity, using the Clinical visit PROactive Physical Activity tool

(C-PPAC)[24]; (vii) number of moderate-to-severe COPD exacerbations in the last 12 months from medical history and self-report; (viii) GOLD I-IV, GOLD ABE[25]; and (ix) diastolic blood pressure.

Statistical Analysis

Assuming an alpha level of 0.05 and a power of 80%, we estimated that detecting statistically significant correlations of ≥0.5, ≥0.4, and ≥0.2 would require 29, 46, and 193 participants, respectively, using Fisher's z-test. Hence, our available sample size, was deemed to provide sufficient power for the planned analyses.

Statistical analysis was planned a priori and conducted using R(v4.4.0). The analysis used a pairwise complete-case approach, including only observations where both the DMO and the corresponding construct measure were available. Participant characteristics and DMOs were described using mean and standard deviation (SD) or median and P25–P75, depending on their distribution.

Pearson (r) or Spearman (ρ) coefficients with their 95% CIs were used between the 24 DMOs and all constructs. Each DMO–construct pair was considered to hold convergent validity if it met the expected correlation coefficients anticipated from pilot study results[7] and expert consultation (Table S3) or exceeded them up to a maximum of 0.9, indicating that the DMO provides the same information as existing constructs[26]. Sub-group analysis of convergent validity was conducted for age tertiles (≤64, 65–72, and ≥73 years), sex, recruitment site (collapsed into three categories: Mediterranean (Athens, Barcelona), Oceanic (London, Leuven, and Newcastle) and Continental (Zurich, Großhansdorf), and occurrence of previous exacerbations. No strata were excluded since they met the minimum sample size requirements.

Divergent validity was assessed by testing the correlations of each of the DMOs with diastolic blood pressure, a priori expected to have no relationship with any of the DMOs. Divergent validity was supported when $|r/\rho|<0.2$.

Known-groups validity was assessed by testing the distribution of DMOs according to the groups defined by GOLD I-IV, GOLD ABE, and mMRC, and supported when p-for-trend from linear regression models was ≤0.05.

Finally, construct validity was established through expert consensus. A group of nine experts from diverse backgrounds (respiratory medicine, epidemiology, physiotherapy, exercise physiology, and data science) assessed each DMO individually, evaluating their validity based on the predefined assumptions described above. Each expert independently assessed the validity of each DMO and final decisions on whether a DMO demonstrated construct validity were made in a consensus meeting. During the initial stage, each expert independently assessed each DMO in a blinded manner, without access to other experts' responses. During the consensus meeting, discrepancies between experts were discussed to identify the reasoning behind differing opinions and ensure transparency, after which a consensus was reached, supported by all experts. The distribution of votes for each DMO is reported to reflect the level of agreement among experts.

Results

From a total of 607 recruited participants, 38 did not have any DMO data and 20 did not meet the above DMO data inclusion criteria with respect to either the number of needed days or the required hours per day; the remaining 549(90%) were included in the analysis. There were no differences between included and excluded participants except for lower FVC%predicted, and higher walking aid use in the included participants (Table S1). The included sample (Table 1) consisted of 37% females. The overall mean(SD) age was 68(8) years, and the mean(SD) FEV1%predicted was 54(20). Participants walked a median of 6561 steps per day, during a mean of 296 walking bouts, with a 90th percentile (P90) walking speed in longer walking bouts of 0.99 m/s, a cadence of 85 steps/minute, and walking speed bout-to-bout variability in longer walking bouts of 17.2%.

Convergent validity

The walking activity amount domain DMOs provided moderate-to-strong correlations with related constructs (|ρ|=0.20–0.67) (Table 2) consistent with the anticipated coefficients (Table S3). Walking pattern DMOs provided weak-to-strong associations with related constructs (|ρ|=0.08–0.64), with most coefficients matching the expected values.

For pace domain DMOs, correlations with related constructs were small-to-strong (|r|=0.07–0.65), with the vast majority matching the expected coefficients. Rhythm DMOs provided small-to-strong coefficients (|ρ/r|=0.06–0.67), matching expectations for most constructs except for QMVC torque and CAT. Walking speed and stride length bout-to-bout variability provided weak-to-moderate correlations (|r|=0.15–0.58), but stride duration and cadence variability produced weak correlations (|r|=0.00–0.35) (Table 2). Correlation coefficients were found to be comparable across age groups, sex, history of exacerbations and sites (Figures 1, S1-5).

Divergent validity

Weak correlations for all 24 DMOs were established with diastolic blood pressure (|ρ|=0.01–0.17) (Table 2).

Known-groups validity

Twenty-two DMOs (all but stride duration in all WBs and stride duration bout-to-bout variability) were statistically differentiated across groups (GOLD I-IV, GOLD ABE, and mMRC), suggesting known-groups validity (Figures S6-8).

Expert consensus on construct validity

There was consensus agreement regarding seventeen DMOs that met construct validity. For ten of them, there was unanimous agreement (100% voting), while for seven DMOs, there was a slight disagreement, and, after additional discussion, all were deemed to meet construct validity, with no justification to suggest otherwise based on the constructs used. Seven DMOs were considered not to meet construct validity (Table 3).

Discussion

This is the first multi-site study evaluating the construct validity of real-world DMOs, in a large COPD sample. Consensus identified 17 of 24 DMOs supporting construct validity. Specifically: (1) all amount and pattern DMOs were supported by all validity analyses; (2) most of the pace and rhythm DMOs showed construct validity, with correlation magnitudes varying between WB duration categorisation; and (3) half of the bout-to-bout variability DMOs supported all types of validity. Accordingly, DMOs capture the intended health aspects in COPD, providing reliable insights into patient health status. Notably, seven DMOs did not meet construct validity: the walking bout duration, four derived from short walking bouts alone (walking speed, cadence, stride length, stride duration), and two of bout-to-bout variability (cadence, stride duration), highlighting that the DMOs not meeting construct validity were those measured in or including short walking bouts.

Walking activity DMOs

Amount

Similar outcomes to the technically validated Mobilise-D walking activity-amount DMOs have been widely applied in research and clinical settings for people with COPD[2]. They produced strong correlations for all related constructs except for QMVC torque, and were robust to stratifications by age group, sex, and occurrence of exacerbations, consistent with previous research in COPD[27, 28]. They also exhibited good divergent and known-groups validity. Based on expert consensus, both walking duration and WB step count were deemed to exhibit construct validity.

Pattern

We assessed seven DMOs related to walking activity pattern. Four of these DMOs correspond to the number of WBs of different durations (overall, longer than 10s, 30s and 60s). These are new parameters not previously tested in relation to other constructs. Longer WBs may identify outdoor walking and indicate both the ability and the behaviour to sustain longer WBs despite COPD symptomatology. These DMOs exhibited moderate-to-strong correlations with the 6MWD, the mMRC and $FEV_{1\%predicted}$, and weak correlations with CAT, C-PPAC difficulty, and QMVC torque. However, WB duration bout-to-bout variability and WB duration correlation coefficients were low. Convergent validity results were found to be consistent across subgroups tested. Average WB duration and bout-to-bout variability matched four out of six hypotheses, although with weak-to-small correlations, potentially not supporting convergent validity. This is possibly related to most WBs being short due to behavioural patterns and living environments, making average WB duration less reflective of an individual's functional capacity and walking performance. All pattern DMOs supported divergent and known-groups validity. Expert consensus supported the construct validity of all pattern DMOs except for average WB duration and WB duration bout-to-bout variability. In addition, the number of WBs longer than 30s appears to perform best and would be the preferred DMO within the walking activity pattern domain.

Gait DMOs

Pace

Real-world walking speed and stride length are two prominent metrics whose construct validity has not yet been assessed in people with COPD. Weak-to-strong correlation coefficients were observed, with most walking pace-related DMOs meeting the expected ones, also in subgroup analyses, and all of them fulfilled the expected divergent and known-groups expectations. However, convergent, divergent, and known-groups validity of walking speed in shorter (10-30s) WBs was not consistently met. There were consistently higher correlation coefficients between the DMOs in longer bouts and the related constructs, compared to shorter bouts. This is consistent with recent evidence suggesting that differences in DMOs between COPD and healthy counterparts are intensified during longer WBs[5], likely due to capacity limitations such as exertional breathlessness and peripheral muscle dysfunction being amplified during longer WBs[5]. Furthermore, shorter bouts likely reflect indoor walking activity. According to expert consensus, all pace DMOs, except walking speed and stride length in shorter (10–30s) WBs, fulfilled construct validity. By examining the correlation coefficients, the P90 walking speed in longer (>30s) WBs appears to be the preferred DMO to represent the pace domain.

Rhythm

Cadence and stride duration reflect the rhythmical and temporal walking patterns of an individual and emerging evidence has supported the clinical relevance of real-world cadence in COPD[6]. Our findings support convergent validity for average and P90 cadence, and stride duration in longer WB, but not in all WBs. Furthermore, stride duration in all WBs did not provide known-groups validity. The C-PPAC score, a patient-reported outcome (PRO) capturing symptoms experienced during physical activity and need for adaptations, showed low to medium correlations with DMOs. This likely reflects that DMOs specifically capture walking performance and may not reflect domains encompassed by the C-PPAC. In addition, QMVC torque, representing the maximum torque generated by the quadriceps muscles, may not be directly related to rhythmical gait patterns, and therefore showed no correlations with cadence, stride duration, or their variability measures. Hence, like the pace DMOs, we suggest that rhythm DMOs are more appropriate and exhibit greater clinical validity when measured in longer bouts. As suggested by expert consensus, the walking rhythm DMOs exhibiting construct validity are cadence, P90 cadence, and stride duration in longer WBs. The preferred DMOs to represent rhythm seems to be P90 cadence in longer WBs.

Bout-to-bout variability

This domain provides information on the consistency of gait performance and has not been researched in people with COPD nor in other respiratory or cardiovascular conditions. Convergent validity was supported for walking speed and stride length variability, which was consistent across subgroups. Those DMOs also provided known-groups validity, with higher bout-to-bout variability values present in people with milder disease status and lower dyspnoea (Figures S6-S8). These results suggest that people with less severe COPD can adjust their walking behaviour, e.g. to different surfaces as well as environmental and social factors, while people with more severe COPD may lack the capacity to adjust their walking patterns or simply move in an environment that requires less variability (indoors). Hence, according to expert consensus, walking speed and stride length bout-to-bout variability exhibited construct validity but need further research in COPD to identify their clinical relevance and potential applications. Since the experts were unable to fully interpret the clinical meaning of bout-to-bout variability DMOs and the constructs that they represent, we suggest considering that this construct validity evidence be considered exploratory.

Implications

Our results support the potential use of DMOs for monitoring and management of people with COPD. A total of seventeen Mobilise-D DMOs (pace, rhythm) relate to gait capacity and performance, while others (amount) reflect overall physical activity behaviour (Table 3). This study constitutes the first published evidence of clinical validity for Mobilise-D DMOs [29] in COPD and provides regulatory-grade evidence for these digital biomarkers, aligning with the EMA's recognition of their potential as monitoring biomarkers [30]. Further analysis of their predictive capacity for clinical events and establishment their minimal clinically important difference is needed before recommending their use in real-world settings. These insights may

guide personalised treatment and early detection. Potential clinical applications of DMOs include supporting disease management via objective, continuous outcome measures to help clinicians tailor pharmacological or non-pharmacological (rehabilitation) interventions and detect early signs of disease exacerbation. They may also enable early detection of health and functional decline, and facilitate remote monitoring. This includes monitoring outside of clinical visits, where changes in DMOs may serve as early indicators of deteriorating health status and prompt timely clinical assessment. However, additional clinical validation is essential before these tools can be routinely implemented. DMO-based monitoring may reduce the potential need for frequent clinic visits, which would be particularly beneficial for people with mobility challenges. The finding that seventeen DMOs correlate with relevant constructs of functional and physical capacity, as well as health status in individuals with COPD, suggests that they may offer complementary insights into disease severity and health status.

At the research level, the present study is the first step in clinically validating DMOs in COPD. Important next steps include determining their predictive capacity for clinical events, and establishing their minimal clinically important difference, even for those that did not meet construct validity, before being definitively excluded as relevant DMOs for COPD. Finally, the sensitivity of DMOs following pharmacological and non-pharmacological interventions will need to be established through randomised controlled trials[31]. This would allow a holistic and comprehensive identification of the most clinically valid DMOs [29], and those that are suitable as outcome measures in clinical trials. After clinical validation is complete and the most robust DMOs are identified, these measures could be integrated as endpoints in clinical trials evaluating non-pharmacological interventions, such as pulmonary rehabilitation, as well as pharmacological treatments. It will be important to ensure that the regulatory-grade evidence supporting these DMOs is extensively published and collected to facilitate their acceptance for as secondary or primary endpoints.

Strengths and limitations

The multicentre design enhances generalisability across age, sex, and disease severity. The rigorous protocol for data quality and predefined thresholds for valid wear time guarantees high data reliability, while the combination of a priori hypotheses, pilot study findings, and expert consensus enforces the validity of the results. Construct validity was assessed using three types of statistical validity alongside clinical relevance from expert consensus, with the majority of DMOs (17 out of 24) meeting 8 to 10 out of the 10 predefined hypotheses. While we acknowledge that some subjectivity is inherent in the expert judgement, the multidisciplinary nature of the panel and the inclusion of health practitioners adds methodological rigor and clinical relevance to the decision. This study presents a practical framework for assessing construct validity in digital mobility biomarkers by combining predefined hypotheses with carefully conducted expert consensus, grounded in existing regulatory frameworks [15-17]. Although the study population was primarily recruited from European urban settings, which may affect the absolute levels of DMOs and constructs, the clinical characteristics were comparable to other large European COPD cohorts, supporting generalisability [32-36]. Recruitment during the COVID-19 pandemic may have influenced attendance, possibly favouring participants with milder disease. However, the construct validity is expected to remain similar in other settings, as any variation in the constructs would likely be reflected proportionally in the associated DMOs. Thus, the coefficients should be comparable, even if the levels differ across rural or non-European COPD populations. While the magnitude of correlation coefficients varied across geographical clusters (Mediterranean, Oceanic, Continental), the 95% confidence intervals overlapped, indicating consistent convergent validity across sites. We used typical constructs related to disease severity that can be measured in multicentre studies. Other criterion measures could have included overall energy expenditure, such as using doubly labelled water to comprehensively capture physical activity, although this would have not been appropriate since our focus was on walking activity only.

Conclusion

A total of 17 walking activity and gait DMOs, representing domains of amount, pattern, pace, rhythm and bout-to-bout variability, are objective and valid outcomes in COPD, as supported by convergent, divergent, and known-groups validity after rigorous statistical testing and expert consensus evaluation. Specifically, all amount and most pattern DMOs demonstrate robust construct validity, while pace, rhythm and bout-to-

bout variability DMOs show stronger clinical validity when derived from longer WBs. DMOs provide a seamless and cost-effective solution exhibiting great measurement and clinimetric properties and reflecting a wide range of clinical outcomes such as disease severity, functional capacity, health status, and perception of physical activity. Finally, this study offers a template for rigorous construct validation of digital mobility biomarkers.

Funding

This work was supported by the Mobilise-D project that has received funding from the Innovative Medicines Initiative 2 Joint Undertaking (JU) under grant agreement No. 820820. This JU receives support from the European Union's Horizon 2020 research and innovation program and the European Federation of Pharmaceutical Industries and Associations (EFPIA). Content in this publication reflects the authors' view and neither IMI nor the European Union, EFPIA, or any Associated Partners are responsible for any use that may be made of the information contained herein.

Declaration of interest

D. Megaritis reports support for the present study from Mobilise-D project, funded by the Innovative Medicines Initiative 2 Joint Undertaking (JU) under grant agreement No. 820820.

V. Lanfranchi and M. Long were supported by the National Institute for Health and Care Research (NIHR) Sheffield Biomedical Research Centre (NIHR203321). The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

M. de las Heras was funded by European Union-NextGenerationEU, under the Program Investigo (INVESTIGO 2022, AGAUR- ref. BDNS 608313, file 2022 INV-1 00046 - Code 100046TC19).

V. Alcaraz-Serrano has nothing to disclose.

P. Alvarez has nothing to disclose.

C. Becker has nothing to disclose.

J. Braun has nothing to disclose.

J. Buekers (FJC2021-046458-I) received funding from the Juan de la Cierva Formación financed by MICIU/AEI/10.13039/501100011033 and by the European Union NextGenerationEU/PRTR.

S. Buttery has nothing to disclose.

B. Caulfield reports support for the present study from European Commission – IMI2 Programme – Mobilise-D project.

A. Cereatti reports support for the present study from the Mobilise-D IMI project.

N. Chynkiamis has nothing to disclose

S. Del Din and L. Rochester were also supported by the IDEA-FAST project that has received funding from the Innovative Medicines Initiative 2 Joint Undertaking (JU) under grant agreement No. 853981. S. Del Din and

L. Rochester were also supported by the National Institute for Health Research (NIHR) Newcastle Biomedical Research Centre (BRC) based at The Newcastle upon Tyne Hospital NHS Foundation Trust, Newcastle University and the Cumbria, Northumberland and Tyne and Wear (CNTW) NHS Foundation Trust. S. Del Din and L. Rochester were also supported by the NIHR/Wellcome Trust Clinical Research Facility (CRF) infrastructure at Newcastle upon Tyne Hospitals NHS Foundation Trust. S. Del Din was supported by the UK Research and Innovation (UKRI) Engineering and Physical Sciences Research Council (EPSRC) (Grant Ref: EP/X031012/1 and Grant Ref: EP/X036146/1). S. Del Din reports consultancy activity with Hoffmann-La Roche Ltd. outside of this study. L. Rochester reports receiving consulting fees from the Michael J. Fox Foundation for serving on the Endpoints Advisory Committee. The content in this publication reflects the authors' view, and neither IMI nor the European Union, EFPIA, NHS, NIHR or any associated partners are responsible for any use that may be made of the information contained herein.

L. Delgado-Ortiz has nothing to disclose.

H. Demeyer is a post-doctoral research fellow of the Flemish Research Foundation (FWO Flanders, #12ZW822N).

A. Frei has nothing to disclose.

E. Gimeno-Santos has nothing to disclose.

N. S. Hopkinson has nothing to disclose.

A. Ionescu has nothing to disclose.

C. P. Jansen has nothing to disclose.

A. Josa-Culleré acknowledges the Spanish Ministry of Science and Innovation through the Ayudas para la Formación de Profesorado Universitario (FPU) 2020–2024 doctoral funding (FPU21/03336).

A. Kirsten has nothing to disclose.

S. Koch has nothing to disclose.

J. Lemos has nothing to disclose

K.E.J. Philip reports travel support from GSK (Travel Award Winner, 2024); non-promotional speaker fees from Chiesi; and salary and research support from the National Heart and Lung Institute, Imperial College, through the Clinical Lecturer scheme.

B. Sharrack has nothing to disclose.

D. Singleton has nothing to disclose

B. Vereijken reports funding from the Innovative Medicines Initiative Joint Undertaking 2 (IMI-JU2) under grant agreement No. 820820.

I. Vogiatzis has nothing to disclose

H. Watz has nothing to disclose

T. Troosters reports support for the present study from the Mobilise-D project, funded by the Innovative Medicines Initiative 2 Joint Undertaking (JU) under grant agreement No. 820820. T. Troosters also reports consulting fees from Roche and honoraria for lectures or presentations from AstraZeneca.

J. Garcia-Aymerich reports support for the present study from Mobilise-D project, funded by the Innovative Medicines Initiative 2 Joint Undertaking (JU) under grant agreement No. 820820. ISGlobal acknowledges support from the grant CEX2023-0001290-S funded by MCIN/AEI/10.13039/501100011033, and support from the Generalitat de Catalunya through the CERCA Programme.

Reference list

1. Demeyer, D. Mohan, C. Burtin, et al., *Objectively Measured Physical Activity in Patients with COPD: Recommendations from an International Task Force on Physical Activity.* Chronic Obstr Pulm Dis, 2021.
2. Polhemus, L.D. Ortiz, G. Brittain, et al., *Walking on common ground: a cross-disciplinary scoping review on the clinical utility of digital mobility outcomes.* npj Digital Medicine, 2021. **4**(1): p. 149.
3. Buttery, S., P. Williams, S. Alghamdi, et al., *Investigating the prognostic value of digital mobility outcomes in patients with chronic obstructive pulmonary disease: a systematic literature review and meta-analysis.* European Respiratory Review, 2023. **32**(170): p. 230134.
4. Gimeno-Santos, E., A. Frei, C. Steurer-Stey, et al., *Determinants and outcomes of physical activity in patients with COPD: a systematic review.* Thorax, 2014. **69**(8): p. 731-9.
5. Buekers, J., D. Megaritis, S. Koch, et al., *Laboratory and free-living gait performance in adults with COPD and healthy controls.* ERJ Open Research, 2023: p. 00159-2023.
6. Delgado-Ortiz, L., S. Ranciati, A. Arbillaga-Etxarri, et al., *Real-world walking cadence in people with COPD.* ERJ Open Research, 2024: p. 00673-2023.
7. Megaritis, D., J. Buekers, T. Bonci, et al., *Impact of symptoms and disease severity on digital mobility outcomes in COPD.* European Respiratory Journal, 2023. **62**(suppl 67): p. PA1592.
8. Delgado-Ortiz, L., J. Buekers, N. Chynkiamis, et al., *How do people with COPD walk? A European study on digitally measured real-world gait.* European Respiratory Journal, 2025: p. 2402303.
9. Kirk, C., A. Küderle, M.E. Micó-Amigo, et al., *Mobilise-D insights to estimate real-world walking speed in multiple conditions with a wearable device.* Scientific Reports, 2024. **14**(1): p. 1754.
10. Mazzà, C., L. Alcock, K. Aminian, et al., *Technical validation of real-world monitoring of gait: a multicentric observational study.* BMJ Open, 2021. **11**(12): p. e050785.
11. Scott, K., T. Bonci, F. Salis, et al., *Design and validation of a multi-task, multi-context protocol for real-world gait simulation.* J Neuroeng Rehabil, 2022. **19**(1): p. 141.
12. Mikolaizak, A.S., L. Rochester, W. Maetzler, et al., *Connecting real-world digital mobility assessment to clinical outcomes for regulatory and clinical endorsement-the Mobilise-D study protocol.* PLoS One, 2022. **17**(10): p. e0269615.
13. Micó-Amigo, M.E., T. Bonci, A. Paraschiv-Ionescu, et al., *Assessing real-world gait with digital technology? Validation, insights and recommendations from the Mobilise-D consortium.* Journal of NeuroEngineering and Rehabilitation, 2023. **20**(1): p. 78.
14. Salis, F., S. Bertuletti, T. Bonci, et al., *A multi-sensor wearable system for the assessment of diseased gait in real-world conditions.* Frontiers in Bioengineering and Biotechnology, 2023. **11**.
15. FDA. *Digital health technologies for remote data acquisition in clinical investigations*. 2022; Available from: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/digital-health-technologies-remote-data-acquisition-clinical-investigations.
16. FDA. *Biomarker Qualification Program*. 2025; Available from: https://www.fda.gov/drugs/drug-development-tool-ddt-qualification-programs/biomarker-qualification-program.
17. FDA. *Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims*. 2009; Available from: https://www.fda.gov/regulatory-information/search-fda-

guidance-documents/patient-reported-outcome-measures-use-medical-product-development-support-labeling-claims.
18. Miller, M.R., J. Hankinson, V. Brusasco, et al., *Standardisation of spirometry.* European Respiratory Journal, 2005. **26**(2): p. 319.
19. Blondeel, A., H. Demeyer, V. Alcaraz-Serrano, et al., *Validation of the Late-Life Function and Disability Instrument in People Living with COPD.* Ann Am Thorac Soc, 2024.
20. Buekers, J., J. Chernova, J. Marchena, et al., *Reliability of real-world walking activity and gait assessment in people with COPD. How many hours and days are needed?* European Respiratory Journal, 2024. **64**(suppl 68): p. PA791.
21. Kluge, F., S. Del Din, A. Cereatti, et al., *Consensus based framework for digital mobility monitoring.* PLoS One, 2021. **16**(8): p. e0256541.
22. Koch, S., J. Buekers, I. Cobo, et al., *From high-resolution time series to a single, clinically-interpretable value - considerations for the aggregation of real world walking speed assessed by wearable sensors in patients with chronic obstructive pulmonary disease (COPD).* European Respiratory Journal, 2023. **62**(suppl 67): p. PA1595.
23. Holland, A.E., M.A. Spruit, T. Troosters, et al., *An official European Respiratory Society/American Thoracic Society technical standard: field walking tests in chronic respiratory disease.* Eur Respir J, 2014. **44**(6): p. 1428-46.
24. Garcia-Aymerich, J., M.A. Puhan, S. Corriol-Rohou, 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**(3): p. 228.
25. GOLD, *GLOBAL STRATEGY FOR PREVENTION, DIAGNOSIS AND MANAGEMENT OF COPD: 2023 Report.* 2023.
26. Pesudovs, K., J.M. Burr, C. Harley, et al., *The development, assessment, and selection of questionnaires.* Optom Vis Sci, 2007. **84**(8): p. 663-74.
27. Troosters, T., F. Sciurba, S. Battaglia, et al., *Physical inactivity in patients with COPD, a controlled multi-center pilot-study.* Respir Med, 2010. **104**(7): p. 1005-11.
28. Watz, H., B. Waschki, C. Boehme, et al., *Extrapulmonary effects of chronic obstructive pulmonary disease on physical activity: a cross-sectional study.* Am J Respir Crit Care Med, 2008. **177**(7): p. 743-51.
29. Rochester, L., C. Mazzà, A. Mueller, et al., *A Roadmap to Inform Development, Validation and Approval of Digital Mobility Outcomes: The Mobilise-D Approach.* Digit Biomark, 2020. **4**(Suppl 1): p. 13-27.
30. EMA. *Letter of support for Mobilise-D digital mobility outcomes as monitoring biomarkers*. 2020; Available from: https://www.ema.europa.eu/en/documents/other/letter-support-mobilise-d-digital-mobility-outcomes-monitoring-biomarkers_en.pdf?fbclid=IwAR0bxuPK5P6ZiKe7uE7-Pk00rhUNqHQB3-u7VPRDqTANrQ7PPEbNI6kmEVA.
31. Megaritis, D., E. Hume, N. Chynkiamis, et al., *Effects of pharmacological and non-pharmacological interventions on physical activity outcomes in chronic respiratory diseases: a systematic review and meta-analysis.* European Respiratory Journal, 2022. **60**(suppl 66): p. 557.
32. Rambod, M., J. Porszasz, B.J. Make, et al., *Six-minute walk distance predictors, including CT scan measures, in the COPDGene cohort.* Chest, 2012. **141**(4): p. 867-875.
33. Agusti, A., P.M. Calverley, B. Celli, et al., *Characterisation of COPD heterogeneity in the ECLIPSE cohort.* Respir Res, 2010. **11**(1): p. 122.
34. Han, M.K., P.M. Quibrera, E.E. Carretta, et al., *Frequency of exacerbations in patients with chronic obstructive pulmonary disease: an analysis of the SPIROMICS cohort.* Lancet Respir Med, 2017. **5**(8): p. 619-626.
35. Keene, J.D., S. Jacobson, K. Kechris, et al., *Biomarkers Predictive of Exacerbations in the SPIROMICS and COPDGene Cohorts.* Am J Respir Crit Care Med, 2017. **195**(4): p. 473-481.
36. Alter, P., C. Stoleriu, K. Kahnert, et al., *Characteristics of Current Smokers versus Former Smokers with COPD and Their Associations with Smoking Cessation Within 4.5 Years: Results from COSYCONET.* Int J Chron Obstruct Pulmon Dis, 2023. **18**: p. 2911-2923.

Table 1. Sociodemographic and clinical characteristics of 549 people with COPD, recruited from 7 sites in 6 European countries, 2021-2022.

mMRC: Modified Medical Research Council Dyspnoea Score; CAT: COPD Assessment Test; FEV_1 (% predicted): Forced Expiratory Volume in 1 Second (Percentage of Predicted); FVC: Forced Vital Capacity; C-PPAC: Clinical Visit-PROactive Physical Activity Tool. BMI: Body Mass Index.

	All n = 549	**GOLD 1 n = 62 (11%)**	**GOLD 2 n = 235 (43%)**	**GOLD 3 n = 178 (33%)**	**GOLD 4 n = 74 (13%)**
Age, mean (SD)	68 (8)	67 (8)	68 (8)	68 (8)	65 (7)
Sex, female n(%)	202 (37%)	28 (45%)	84 (36%)	60 (34%)	30 (41%)
Recruitment site, n(%)					
Athens	48 (9%)	2 (3%)	20 (9%)	15 (8%)	11 (15%)
Barcelona	148 (27%)	20 (32%)	60 (26%)	49 (28%)	19 (26%)
Grosshansdorf	132 (24%)	18 (29%)	65 (28%)	38 (21%)	11 (15%)
Leuven	109 (20%)	5 (8%)	53 (23%)	37 (21%)	14 (19%)
London	24 (4%)	3 (5%)	9 (4%)	8 (4%)	4 (5%)
Newcastle	47 (9%)	5 (8%)	15 (6%)	22 (12%)	5 (7%)
Zurich	41 (7%)	9 (15%)	13 (6%)	9 (5%)	10 (14%)
BMI (kg/m^2), mean (SD)*	28 (5)	27 (4)	29 (5)	27 (5)	26 (6)
Packs years, mean (SD)	53 (29)	49 (29)	52 (26)	55 (30)	55 (31)
Dyspnoea (mMRC grade 0-4), median (Q1 - Q3)*	2 (1 - 2)	1 (1 - 1)	1 (1 - 2)	2 (1 - 3)	2 (2 - 3)
CAT Score (0-40), median (Q1-Q3)*	14 (9 - 19)	10 (6 - 17)	13 (8 - 18)	16 (11 - 20)	16 (12 - 23)
Isometric quadriceps muscle torque (N·m), mean (SD)*	121 (55)	135 (63)	126 (54)	115 (55)	108 (45)
6-min walk distance (m), mean (SD)*	416 (119)	483 (103)	444 (100)	395 (124)	320 (108)
FEV1 (% pred), mean (SD)	54 (20)	90 (9)	64 (8)	40 (5)	25 (4)
FVC (% pred), mean (SD)	85 (20)	111 (11)	91 (14)	77 (14)	60 (15)
GOLD ABE*					
Group A - Low symptom severity, low exacerbation risk	109 (20%)	25 (40%)	61 (26%)	20 (11%)	3 (4%)
Group B - High symptom severity, low exacerbation risk	322 (59%)	31 (50%)	138 (59%)	111 (62%)	42 (57%)
Group E - High exacerbation risk	107 (19%)	6 (10%)	28 (12%)	45 (25%)	28 (38%)
Inhaled corticosteroids (ICS), n (%)	328 (60%)	24 (39%)	119 (51%)	125 (70%)	60 (81%)
Long-acting bronchodilator (LAMA and/or LABA), n (%)	472 (86%)	41 66%)	192 (82%)	170 (96%)	69 (93%)
Short-acting bronchodilator (SAMA and/or SABA), n (%)	296 (54%)	22 (35%)	115 (49%)	109 (61%)	50 (68%)
Long-acting muscarinic agonist (LAMA) monotherapy, n (%)	20 (4%)	6 (10%)	12 (5%)	2 (1%)	0 (0%)
Long-acting beta agonist (LABA) and ICS, n (%)	306 (56%)	22 (35%)	108 (46%)	120 (67%)	56 (75%)
Triple therapy (LAMA + LABA + ICS), n (%)	263 (47%)	9 (16%)	92 (39%)	110 (61%)	52 (70%)
Wearable device used, n(%)					
Dynaport MoveMonitor	452 (82%)	41 (66%)	201 (86%)	146 (82%)	64 (86%)
Axivity AX6	97 (18%)	21 (34%)	34 (14%)	32 (18%)	10 (14%)
Walking aids (yes), n(%)	44 (8%)	1 (2%)	12 (5%)	17 (10%)	14 (19%)
Physical Activity Experience (C-PPAC Scores, 0-100), mean (SD)*					
Amount	64 (19)	73 (17)	69 (17)	61 (19)	49 (19)
Difficulty	72 (16)	81 (13)	76 (14)	68 (15)	60 (14)
Total Score	68 (15)	77 (12)	73 (13)	64 (15)	55 (14)
Diastolic blood pressure (mm Hg), mean (SD)	80 (13)	83 (10)	80 (12)	80 (14)	79 (12)
Occurrence of moderate-to-severe exacerbations 12 month prior to study inclusion, n (%)	178 (32%)	11 (6%)	52 (29%)	71 (40%)	44 (25%)

*Some constructs had missing values: 6 in 6-min walk distance; 15 in CAT; 25 in mMRC; 34 in Isometric quadriceps muscle torque; 7 in CPPAC; 13 in GOLD ABE; 2 in diastolic blood pressure; 7 in BMI.

Table 2. Distribution and convergent and divergent validity (correlation coefficients with corresponding 95% CIs) of DMOs in 549 people with COPD. Coefficients in bold represent those for which convergent or divergent validity is suggested.

DMO	Mean (SD) or Median (p25 - p75)*	Convergent						Divergent
		6-min walk distance n = 543	CAT n = 534	mMRC n = 524	FEV_1 % pred n = 549	QMVC Torque n = 515	C-PPAC Difficulty n = 542	Diastolic Blood Pressure n = 547
Walking activity								
Amount								
Walking Duration (min/day)	70 (44 - 104)	**0.66 (0.60, 0.71)**	**-0.34 (-0.41, -0.26)**	**-0.48 (-0.54, -0.40)**	**0.32 (0.24, 0.40)**	**0.20 (0.12, 0.28)**	**0.43 (0.36, 0.50)**	**0.10 (0.02, 0.18)**
WB Step Count (steps/day)	6578 (4023 - 9744)	**0.67 (0.61, 0.72)**	**-0.34 (-0.42, -0.26)**	**-0.48 (-0.55, -0.41)**	**0.33 (0.24, 0.40)**	**0.20 (0.11, 0.27)**	**0.44 (0.37, 0.50)**	**0.10 (0.02, 0.18)**
Pattern								
Number of WBs (n)	295.7 (137.9)	**0.53 (0.47, 0.59)**	**-0.29 (-0.37, -0.21)**	**-0.42 (-0.50, -0.35)**	**0.29 (0.19, 0.37)**	0.08 (0.00, 0.16)	**0.37 (0.29, 0.44)**	**0.09 (0.01, 0.17)**
Number of WBs > 10s (n)	121 (83 - 171)	**0.59 (0.53, 0.65)**	**-0.32 (-0.39, -0.24)**	**-0.45 (-0.52, -0.38)**	**0.31 (0.22, 0.39)**	0.19 (0.11, 0.28)	**0.40 (0.33, 0.47)**	**0.12 (0.03, 0.20)**
Number of WBs > 30s (n)	19 (11 - 33)	**0.64 (0.58, 0.69)**	**-0.33 (-0.40, -0.25)**	**-0.47 (-0.53, -0.40)**	**0.28 (0.20, 0.36)**	**0.25 (0.17, 0.32)**	**0.42 (0.36, 0.50)**	**0.15 (0.07, 0.22)**
Number of WBs > 60s (n)	6 (2 - 12)	**0.57 (0.50, 0.62)**	**-0.29 (-0.37, -0.20)**	**-0.41 (-0.48, -0.32)**	**0.26 (0.18, 0.34)**	0.19 (0.10, 0.26)	**0.35 (0.27, 0.43)**	**0.15 (0.07, 0.23)**
WB duration (s)	8.9 (8.2 - 9.6)	**0.31 (0.24, 0.39)**	-0.15 (-0.23, -0.07)	**-0.22 (-0.30, -0.14)**	0.12 (0.03, 0.2)	**0.28 (0.20, 0.36)**	**0.21 (0.14, 0.29)**	**0.10 (0.02, 0.19)**
P90 WB duration (s)	24.9 (21.3 - 30.1)	**0.49 (0.41, 0.55)**	**-0.26 (-0.33, -0.18)**	**-0.35 (-0.42, -0.26)**	0.17 (0.07, 0.25)	**0.27 (0.19, 0.35)**	**0.32 (0.25, 0.40)**	**0.14 (0.06, 0.22)**
WB duration bout to bout variability (%)	127.0 (93.7 - 179.1)	**0.51 (0.44, 0.57)**	**-0.22 (-0.30, -0.14)**	**-0.34 (-0.42, -0.26)**	**0.24 (0.16, 0.32)**	0.16 (0.09, 0.25)	**0.30 (0.22, 0.38)**	**0.09 (0.01, 0.17)**
Gait								
Pace								
Walking speed in shorter (10-30s) WBs (m/s)	0.67 (0.07)	0.36 (0.28, 0.43)	-0.11 (-0.20, -0.03)	-0.27 (-0.34, -0.18)	**0.20 (0.12, 0.28)**	0.19 (0.12, 0.27)	**0.22 (0.14, 0.3)**	**0.17 (0.08, 0.25)**
Walking speed in longer (>30s) WBs (m/s)	0.83 (0.12)	**0.54 (0.46, 0.59)**	**-0.21 (-0.29, -0.13)**	-0.36 (-0.43, -0.28)	**0.23 (0.15, 0.31)**	**0.25 (0.18, 0.33)**	**0.32 (0.24, 0.4)**	**0.15 (0.07, 0.23)**
P90 walking speed in WBs >10s (m/s)	0.9 (0.13)	**0.60 (0.54, 0.65)**	**-0.22 (-0.31, -0.14)**	**-0.40 (-0.47, -0.32)**	**0.24 (0.16, 0.32)**	**0.28 (0.21, 0.36)**	**0.35 (0.28, 0.43)**	**0.16 (0.07, 0.24)**

P90 walking speed in longer (>30s) WBs (m/s)	0.99 (0.18)	**0.67 (0.61, 0.71)**	**-0.24 (-0.32, -0.17)**	**-0.44 (-0.51, -0.37)**	**0.27 (0.19, 0.35)**	**0.29 (0.22, 0.36)**	**0.40 (0.32, 0.46)**	**0.16 (0.08, 0.24)**
Stride length in shorter (10-30s) WBs (m)	0.91 (0.09)	**0.28 (0.20, 0.36)**	-0.06 (-0.13, 0.03)	**-0.21 (-0.29, -0.12)**	0.14 (0.06, 0.22)	**0.23 (0.16, 0.31)**	0.14 (0.06, 0.23)	**0.16 (0.07, 0.24)**
Stride length in longer (>30s) WBs (m)	1.05 (0.12)	**0.40 (0.31, 0.47)**	-0.14 (-0.23, -0.06)	**-0.27 (-0.34, -0.19)**	0.16 (0.07, 0.25)	**0.26 (0.18, 0.33)**	**0.21 (0.13, 0.29)**	**0.14 (0.05, 0.21)**
Rythm								
Cadence in all WBs (steps/min)	84.6 (4.1)	0.12 (0.03, 0.20)	-0.07 (-0.15, 0.02)	-0.08 (-0.17, 0.01)	0.12 (0.04, 0.19)	-0.06 (-0.14, 0.01)	0.10 (0.01, 0.18)	**0.01 (-0.07, 0.09)**
Cadence in longer (>30s) WBs (step/min)	91.5 (6.6)	**0.39 (0.32, 0.45)**	-0.18 (-0.25, -0.11)	**-0.23 (-0.31, -0.14)**	**0.20 (0.12, 0.27)**	0.07 (-0.01, 0.15)	**0.26 (0.19, 0.34)**	**0.07 (-0.01, 0.15)**
P90 cadence in longer (>30s) WBs (steps/min)	100.1 (8.6)	**0.53 (0.48, 0.59)**	**-0.22 (-0.30, -0.14)**	**-0.31 (-0.39, -0.24)**	**0.24 (0.17, 0.32)**	0.10 (0.02, 0.18)	**0.35 (0.27, 0.43)**	**0.08 (0.00, 0.17)**
Stride duration in all WBs (s)	1.30 (0.06)	-0.13 (-0.21, -0.04)	0.06 (-0.03, 0.14)	0.11 (0.02, 0.18)	-0.07 (-0.15, 0.01)	0.00 (-0.09, 0.08)	-0.07 (-0.14, 0.01)	**0.01 (-0.07, 0.09)**
Stride duration in longer (>30s) WBs (s)	1.26 (0.09)	**-0.34 (-0.41, -0.26)**	0.16 (0.08, 0.25)	**0.25 (0.18, 0.33)**	-0.19 (-0.28, -0.11)	-0.06 (-0.14, 0.02)	**-0.24 (-0.32, -0.16)**	**-0.04 (-0.13, 0.04)**
Bout-to-bout variability								
Walking speed bout to bout variability in longer (>30s) WBs (%)	17.2 (5.2)	**0.58 (0.52, 0.63)**	-0.16 (-0.24, -0.08)	**-0.35 (-0.42, -0.27)**	0.19 (0.10, 0.27)	**0.25 (0.16, 0.34)**	**0.34 (0.26, 0.41)**	**0.08 (0.00, 0.16)**
Stride length bout to bout variability in longer (>30s) WBs (%)	11.9 (3.7)	**0.48 (0.42, 0.54)**	-0.19 (-0.26, -0.11)	**-0.30 (-0.38, -0.23)**	0.15 (0.06, 0.23)	**0.23 (0.15, 0.31)**	**0.29 (0.22, 0.37)**	**0.04 (-0.05, 0.12)**
Cadence bout to bout variability (%)	12.2 (1.3)	**0.35 (0.26, 0.42)**	-0.12 (-0.21, -0.04)	**-0.23 (-0.31, -0.14)**	0.11 (0.04, 0.19)	0.05 (-0.02, 0.13)	**0.24 (0.15, 0.32)**	**0.06 (-0.02, 0.14)**

Stride duration bout to bout variability (%)	14.1 (1.8)	0.00 (-0.1, 0.09)	0.05 (-0.04, 0.13)	0.00 (-0.08, 0.09)	-0.01 (-0.10, 0.08)	-0.04 (-0.12, 0.04)	0.01 (-0.07, 0.10)	**-0.08 (-0.16, 0.00)**

* Some variables have missing values: 4 values in walking speed bout-to-bout variability in longer (>30s) WBs, and 4 values in stride length bout-to-bout variability in longer (>30s); WBs: walking bouts.

Table 3. Expert consensus decisions about each DMO meeting construct validity or not.

DMO	Individual evaluation results	Discussion	Consensus decision: meeting construct validity
Walking Activity			
Amount			
Walking Duration	100% in favour		**YES**
WB Step Count	100% in favour		**YES**
Pattern			
Number of WBs	100% in favour		**YES**
Number of WBs > 10s	100% in favour		**YES**
Number of WBs > 30s	100% in favour		**YES**
Number of WBs > 60s	100% in favour		**YES**
WB duration	66% against	• Correlations with CAT and FEV_1 weaker than expected correlations • Correlations with mMRC, a clinically important and potentially related construct, close to lower bound of expected correlations range • Correlations with C-PPAC difficulty close to lower bound of expected correlations range • Weaker correlations than other pattern DMOs	**NO**
P90 WB duration	100% in favour		**YES**
WB duration bout-to-bout variability	66% in favour	• Only correlations with QMVC weaker than expected correlations • Construct validity was confirmed, but the expert group expressed caution in their interpretation, as clinical meaning and relevance of this DMO remains unclear	**YES**
Gait			
Pace			
Walking speed in shorter (10-30s) WBs	77% against	• Correlations with CAT, mMRC and QMVC torque weaker than expected correlations • Correlations with FEV_1 and C-PPAC difficulty close to lower bound of expected correlations range • Likely represents WBs performed indoors, which may reduce its clinical meaning	**NO**
Walking speed in longer (>30s) WBs	100% in favour		**YES**
P90 walking speed in WB > 10 s	100% in favour		**YES**
P90 walking speed in longer (>30s) WB	100% in favour		**YES**
Stride length in shorter (10-30s) WBs	77% against	• Correlations with CAT, FEV_1 and C-PPAC difficulty weaker than expected correlations • Correlations with mMRC and QMVC torque close to lower bound of expected correlations range • Likely represents WBs performed indoors, which may reduce its clinical meaning. • Correlations for divergent validity relatively high	**NO**

Stride length in longer (>30s) WBs	88% in favour	• Some of the hypothesised related constructs may not have been ideal • Performed better than stride length in shorter (10-30s) WBs	**YES**
Rhythm			
Cadence in all WBs	66% against	• All correlations for convergent validity weaker than expected correlations • After joint revision of results all agreed to reject	**NO**
Cadence in longer (>30s) WBs	88% in favour	• Correlations with CAT close to lower bound of expected correlations range, and cadence likely not expected to correlate with QMVC torque • Other convergent, divergent and known-groups expectations were met	**YES**
P90 cadence in longer (>30s) WB	88% in favour	• Cadence likely not expected to correlate with QMVC torque • Other convergent, divergent and known-groups expectations were met • After joint revision of results all agreed to accept	**YES**
Stride duration in all WBs	77% against	• All correlations for convergent validity weaker than expected correlations • After joint revision of results all agreed to reject	**NO**
Stride duration in longer (>30s) WBs	88% in favour	• Stride duration likely not expected to correlate with QMVC torque • Correlations with CAT and FEV_1 close to lower bound of expected correlations range • Divergent and known-groups expectations were met • Rest of correlations are good	**YES**
Bout-to-bout variability			
Walking speed bout-to-bout variability in longer (>30s) WBs	66% in favour	• Correlations with CAT and FEV_1 close to lower bound of expected correlations range • Other convergent, divergent and known-groups expectations were met • Construct validity is confirmed but the expert group expressed caution in their interpretation, as clinical meaning and relevance of this DMO remains unclear	**YES**
Stride length bout-to-bout variability in longer (>30s) WBs	66% in favour	• Correlations with CAT and FEV_1 close to lower bound of expected correlations range • Other convergent, divergent and known-groups expectations were met • Construct validity is confirmed but the expert group expressed caution in their interpretation, as clinical meaning and relevance of this DMO remains unclear	**YES**
Cadence bout-to-bout variability	66% against	• Correlations with CAT, FEV_1 and QMVC torque weaker than expected correlations • Correlations with mMRC and C-PPAC difficulty close to lower bound of expected correlations range	**NO**
Stride duration bout-to-bout variability	66% against	• All correlations for convergent validity weaker than expected correlations	**NO**

		• After joint revision of results all agreed to reject	

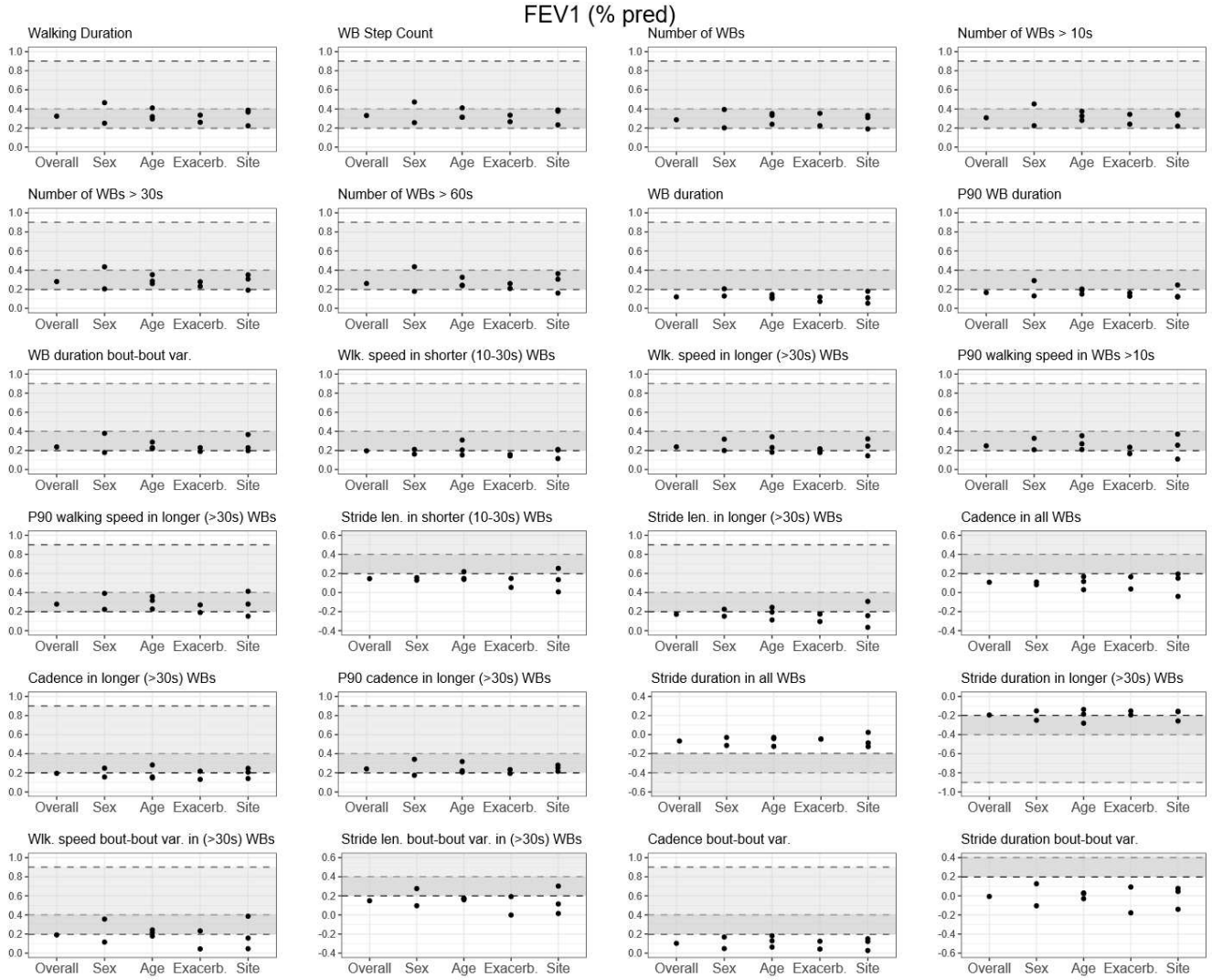

Figure 1. Convergent validity with FEV_1 %predicted across subgroups of age, sex, history of moderate-to-severe exacerbation in the last 12 months, and site Mediterranean, Continental, and Oceanic, see methods). The dark grey area represents *a priori* defined correlations, while the light grey area indicates exceeded correlations.

Supplementary Material to:

The Construct Validity of Real-World Digital Mobility Outcomes in People with Chronic Obstructive Pulmonary Disease

Authors: Dimitrios Megaritis, Michael Long, Martí de las Heras, Victoria Alcaraz-Serrano, Paula Alvarez, Clemens Becker, Julia Braun, Joren Buekers, Sara Buttery, Brian Caulfield, Andrea Cereatti, Nikolaos Chynkiamis, Silvia Del Din, Laura Delgado-Ortiz, Heleen Demeyer, Anja Frei, Elena Gimeno-Santos, Nicholas S Hopkinson, Anisoara Ionescu, Carl-Philipp Jansen, Alícia Josa-Culleré, Anne Kirsten, Sarah Koch, Jorge Lemos, Keir EJ Philip, Lynn Rochester, Basil Sharrack, David Singleton, Beatrix Vereijken, Ioannis Vogiatzis, Henrik Watz, Vita Lanfranchi, Thierry Troosters, Judith Garcia-Aymerich

Table of Contents

Table S1. Comparison of sociodemographic and clinical characteristics between included and excluded individuals with COPD.

	Excluded n = 58 (11%)	**Included n = 549 (89%)**	**p-value**
Age, mean (SD)	68 (9)	68 (8)	0.488
Sex, female n(%)	20 (34%)	202 (37%)	0.773
Recruitment site, n(%)			<0.001
Athens	2 (3%)	48 (9%)	
Barcelona	8 (14%)	148 (27%)	
Grosshansdorf	8 (14%)	132 (24%)	
Leuven	30 (52%)	109 (20%)	
London	6 (10%)	24 (4%)	
Newcastle	3 (5%)	47 (9%)	
Zurich	1 (2%)	41 (7%)	
BMI (kg/m2), mean (SD)	27 (5)	28 (5)	0.425
Pack years, mean (SD)	47 (25)	53 (29)	0.096
Dyspnoea (mMRC grade 0-4), median (P25-P75)	1 (1 - 2)	2 (1 - 2)	0.276
CAT Score 0-40, median (P25-P75)	14 (10 - 18)	14 (9 - 19)	0.810
Isometric quadriceps muscle torque (N·m), mean (SD)	126 (63)	121 (55)	0.848
6-min walk distance (m), mean (SD)	414 (125)	416 (119)	0.920
FEV1 (% pred), mean (SD)	58 (17)	54 (20)	0.069
FVC (% pred), mean (SD)	92 (19)	85 (20)	**0.005**
GOLD ABE			0.400
Group A - Low symptom severity, low exacerbation risk	10 (17%)	109 (20%)	
Group B - High symptom severity, low exacerbation risk	32 (55%)	322 (59%)	
Group E - High exacerbation risk	14 (24%)	107 (19%)	
Inhaled corticosteroids (ICS), n (%)	30 (52%)	328 (60%)	0.298
Long-acting bronchodilator (LAMA and/or LABA), n (%)	44 (76%)	472 (86%)	0.063
Short-acting bronchodilator (SAMA and/or SABA), n (%)	26 (45%)	296 (54%)	0.238
Long-acting muscarinic agonist (LAMA) monotherapy, n (%)	4 (7%)	20 (4%)	0.393
Long-acting beta agonist (LABA) and ICS, n (%)	24 (41%)	306 (56%)	0.051
Triple therapy (LAMA + LABA + ICS), n (%)	20 (34%)	263 (47%)	0.070
Wearable device used, n(%)			0.039
Dynaport MoveMonitor	53 (91%)	452 (82%)	
Axivity AX6	4 (7%)	97 (18%)	
Walking aids (yes), n(%)	11 (19%)	44 (8%)	**0.009**
Physical Activity Experience (C-PPAC Scores, 0-100), mean (SD)			
PPAC Amount (0-less amount- to 100-more amount)	62 (26)	64 (19)	0.612
PPAC Difficulty (0-more difficulty to 100- no difficulty)	73 (17)	72 (16)	0.464
PPAC Total Score (0-worst experience to 100-best experience)	66 (19)	68 (15)	0.665
Diastolic blood pressure (mm Hg), mean (SD)	78 (12)	80 (13)	0.227
Occurrence of moderate-to-severe exacerbations 12 month prior to study inclusion, n (%)	22 (38%)	178 (32%)	0.480

Table S2. Description of the 24 digital mobility outcomes (DMOs) as presented in Delgado-Ortiz (2025), Eur Respir J

DMOs	Definition	Unit
Amount		
Walking duration	Weekly mean of time spent walking per day	min/day
WB Step Count	Weekly mean of the number of steps per day	steps/day
Pattern		
Number of WBs	Weekly mean of the sum of walking bouts per day	WBs/day
Number of WBs > 10s	Weekly mean of the sum of walking bouts per day including bouts longer than 10 seconds	WBs/day
Number of WBs > 30s	Weekly mean of the sum of walking bouts per day including bouts longer than 30 seconds	WBs/day
Number of WBs > 60s	Weekly mean of the sum of walking bouts per day including bouts longer than 60 seconds	WBs/day
WB duration	Weekly mean of the daily mean of walking bout duration	s
P90 WB duration	Weekly mean of the daily 90th percentile of walking bout duration	s
WB duration bout to bout variability	Weekly mean of the daily bout to bout variability of walking bout duration	%
Gait		
Pace		
Walking speed in shorter (10-30s) WBs	Weekly mean of the daily average walking speed, assessed in walking bouts between 10 and 30 seconds	m/s
Walking speed in longer (>30s) WBs	Weekly mean of the daily average walking speed, assessed in walking bouts longer than 30 seconds	m/s
P90 walking speed in WB > 10 s	Weekly mean of the daily 90th percentile of the walking speed, assessed in walking bouts of more than 10 seconds	m/s
P90 walking speed in longer (>30s) WB	Weekly mean of the daily 90th percentile of the walking speed, assessed in walking bouts of longer than 30 seconds	m/s
Stride length in shorter (10-30s) WBs	Weekly mean of the daily length of two consecutive steps, assessed during walking bouts between 10 and 30 seconds	cm
Stride length in longer (>30s) WBs	Weekly mean of the daily average length of two consecutive steps, assessed during walking bouts longer than 30 seconds	cm
Rhythm		
Cadence in all WB	Weekly mean of the daily average of the steps frequency during a period of time (minutes), calculated in all walking bouts	steps/min
Cadence in longer (>30s) WBs	Weekly mean of the daily average of the steps frequency during a period of time (minutes),	steps/min

	calculated in walking bouts longer than 30 seconds	
P90 cadence in longer (>30s) WB	Weekly mean of the daily 90th percentile of the steps frequency during a period of time (minutes), calculated in walking bouts longer than 30 seconds	steps/min
Stride duration in all WBs	Weekly mean of the daily average of the time elapsed between the initial contacts of two consecutive footfalls of the same foot, assessed in all walking bouts	s
Stride duration in longer (>30s) WBs	Weekly mean of the daily average of the time elapsed between the initial contacts of two consecutive footfalls of the same foot, assessed in walking bouts longer than 30 seconds	s
Bout-to-bout variability		
Walking speed bout-to-bout variability in longer (>30s) WBs	Weekly mean of the daily bout-to-bout variability of walking speed, assessed in walking bouts of longer than 30 seconds	%
Stride length bout-to-bout variability in longer (>30s) WBs	Weekly mean of the daily bout-to-bout variability of stride length, assessed in walking bouts of longer than 30	%
Cadence bout-to-bout variability	Weekly mean of the daily bout-to-bout variability of cadence, assessed in all walking bouts	%
Stride duration bout-to-bout variability	Weekly mean of the daily bout-to-bout variability of stride duration, assessed in all walking bouts	%

Table S3: Expected correlation coefficient ranges between DMOs and related constructs derived from previous research, pilot testing and expert consultation

DMO	6MWD	HRQoL (CAT)	Dyspnoea(mMRC)	FEV1 (%)	QMVC Torque	C-PPAC Difficulty (0-100)
Walking duration (h/day)	0.4 to 0.6	-0.4 to -0.2	-0.6 to -0.4	0.2 to 0.4	0.2 to 0.4	0.4 to 0.6
Number of steps (steps/day)	0.4 to 0.6	-0.4 to -0.2	-0.6 to -0.4	0.2 to 0.4	0.2 to 0.4	0.4 to 0.6
Number of WB (n)	0.4 to 0.6	-0.4 to -0.2	-0.4 to -0.2	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
Number of WB > 10s (n)	0.4 to 0.6	-0.4 to -0.2	-0.4 to -0.2	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
Number of WB > 30s (n)	0.4 to 0.6	-0.4 to -0.2	-0.4 to -0.2	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
Number of WB > 60s (n)	0.4 to 0.6	-0.4 to -0.2	-0.4 to -0.2	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
WB duration (s), m (SD)	0.2 to 0.4	-0.4 to -0.2	-0.4 to -0.2	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
Maximum WB duration (s)	0.2 to 0.4	-0.4 to -0.2	-0.4 to -0.2	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
WB duration variability (unit)	0.2 to 0.4	-0.4 to -0.2	-0.4 to -0.2	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
Walking speed in shorter (10-30s) WB (m/s)	0.4 to 0.6	-0.4 to -0.2	-0.4 to -0.6	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
Walking speed in longer (>30s) WB (m/s)	0.4 to 0.6	-0.4 to -0.2	-0.4 to -0.6	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
Maximum walking speed in WB > 10 s (m/s)	0.4 to 0.6	-0.4 to -0.2	-0.4 to -0.6	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
Maximum walking speed in longer (>30s) WB (m/s)	0.4 to 0.6	-0.4 to -0.2	-0.4 to -0.6	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
Stride length in shorter (10-30s) WB (cm)	0.2 to 0.4	-0.4 to -0.2	-0.4 to -0.2	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
Stride length in longer (>30s) WB (cm)	0.2 to 0.4	-0.4 to -0.2	-0.4 to -0.2	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
Cadence in all WB (steps/min)	0.2 to 0.4	-0.4 to -0.2	-0.4 to -0.2	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
Cadence in longer (>30s) WB (steps/min)	0.2 to 0.4	-0.4 to -0.2	-0.4 to -0.2	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
Maximum cadence in longer (>30s) WB (steps/min)	0.2 to 0.4	-0.4 to -0.2	-0.4 to -0.2	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
Stride duration in all WB (s)	-0.4 to -0.2	0.2 to 0.4	0.2 to 0.4	-0.4 to -0.2	-0.4 to -0.2	-0.4 to -0.2
Stride duration in longer (>30s) WB (s)	-0.4 to -0.2	0.2 to 0.4	0.2 to 0.4	-0.4 to -0.2	-0.4 to -0.2	-0.4 to -0.2
Walking speed variability in longer (>30s) WBs (%)	0.2 to 0.4	-0.4 to -0.2	-0.4 to -0.2	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
Stride length variability in longer (>30s) WBs (%)	0.2 to 0.4	-0.4 to -0.2	-0.4 to -0.2	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
Cadence variability (%)	0.2 to 0.4	-0.4 to -0.2	-0.4 to -0.2	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
Stride duration variability (%)	0.2 to 0.4	-0.4 to -0.2	-0.4 to -0.2	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4

DMO	6MWD	CAT	mMRC	FEV1 (%)	QMVC Torque	C-PPAC Difficulty (0-100)
Walking Duration (min/day)	0.4 to 0.6	-0.4 to -0.2	-0.6 to -0.4	0.2 to 0.4	0.2 to 0.4	0.4 to 0.6
WB Step Count (steps/day)	0.4 to 0.6	-0.4 to -0.2	-0.6 to -0.4	0.2 to 0.4	0.2 to 0.4	0.4 to 0.6
Number of WB (WB/day)	0.4 to 0.6	-0.4 to -0.2	-0.4 to -0.2	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
Number of WB > 10s (WB/day)	0.4 to 0.6	-0.4 to -0.2	-0.4 to -0.2	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
Number of WB > 30s (WB/day)	0.4 to 0.6	-0.4 to -0.2	-0.4 to -0.2	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
Number of WB > 60s (WB/day)	0.4 to 0.6	-0.4 to -0.2	-0.4 to -0.2	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
WB duration (s)	0.2 to 0.4	-0.4 to -0.2	-0.4 to -0.2	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
P90 WB duration (s)	0.2 to 0.4	-0.4 to -0.2	-0.4 to -0.2	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
WB duration variability (unit)	0.2 to 0.4	-0.4 to -0.2	-0.4 to -0.2	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
Walking speed in shorter (10-30s) WB (m/s)	0.4 to 0.6	-0.4 to -0.2	-0.4 to -0.6	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
Walking speed in longer (>30s) WB (m/s)	0.4 to 0.6	-0.4 to -0.2	-0.4 to -0.6	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
P90 walking speed in WB > 10 s (m/s)	0.4 to 0.6	-0.4 to -0.2	-0.4 to -0.6	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
P90 walking speed in longer (>30s) WB (m/s)	0.4 to 0.6	-0.4 to -0.2	-0.4 to -0.6	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
Stride length in shorter (10-30s) WB (cm)	0.2 to 0.4	-0.4 to -0.2	-0.4 to -0.2	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
Stride length in longer (>30s) WB (cm)	0.2 to 0.4	-0.4 to -0.2	-0.4 to -0.2	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
Cadence in all WB (steps/min)	0.2 to 0.4	-0.4 to -0.2	-0.4 to -0.2	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
Cadence in longer (>30s) WB (steps/min)	0.2 to 0.4	-0.4 to -0.2	-0.4 to -0.2	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
P90 cadence in longer (>30s) WB (steps/min)	0.2 to 0.4	-0.4 to -0.2	-0.4 to -0.2	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
Stride duration in all WB (s)	-0.4 to -0.2	0.2 to 0.4	0.2 to 0.4	-0.4 to -0.2	-0.4 to -0.2	-0.4 to -0.2
Stride duration in longer (>30s) WB (s)	-0.4 to -0.2	0.2 to 0.4	0.2 to 0.4	-0.4 to -0.2	-0.4 to -0.2	-0.4 to -0.2
Walking speed bout to bout variability in longer (>30s) WBs (%)	0.2 to 0.4	-0.4 to -0.2	-0.4 to -0.2	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
Stride length bout to bout variability in longer (>30s) WBs (%)	0.2 to 0.4	-0.4 to -0.2	-0.4 to -0.2	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
Cadence bout to bout variability (%)	0.2 to 0.4	-0.4 to -0.2	-0.4 to -0.2	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4
Stride duration bout to bout variability (%)	0.2 to 0.4	-0.4 to -0.2	-0.4 to -0.2	0.2 to 0.4	0.2 to 0.4	0.2 to 0.4

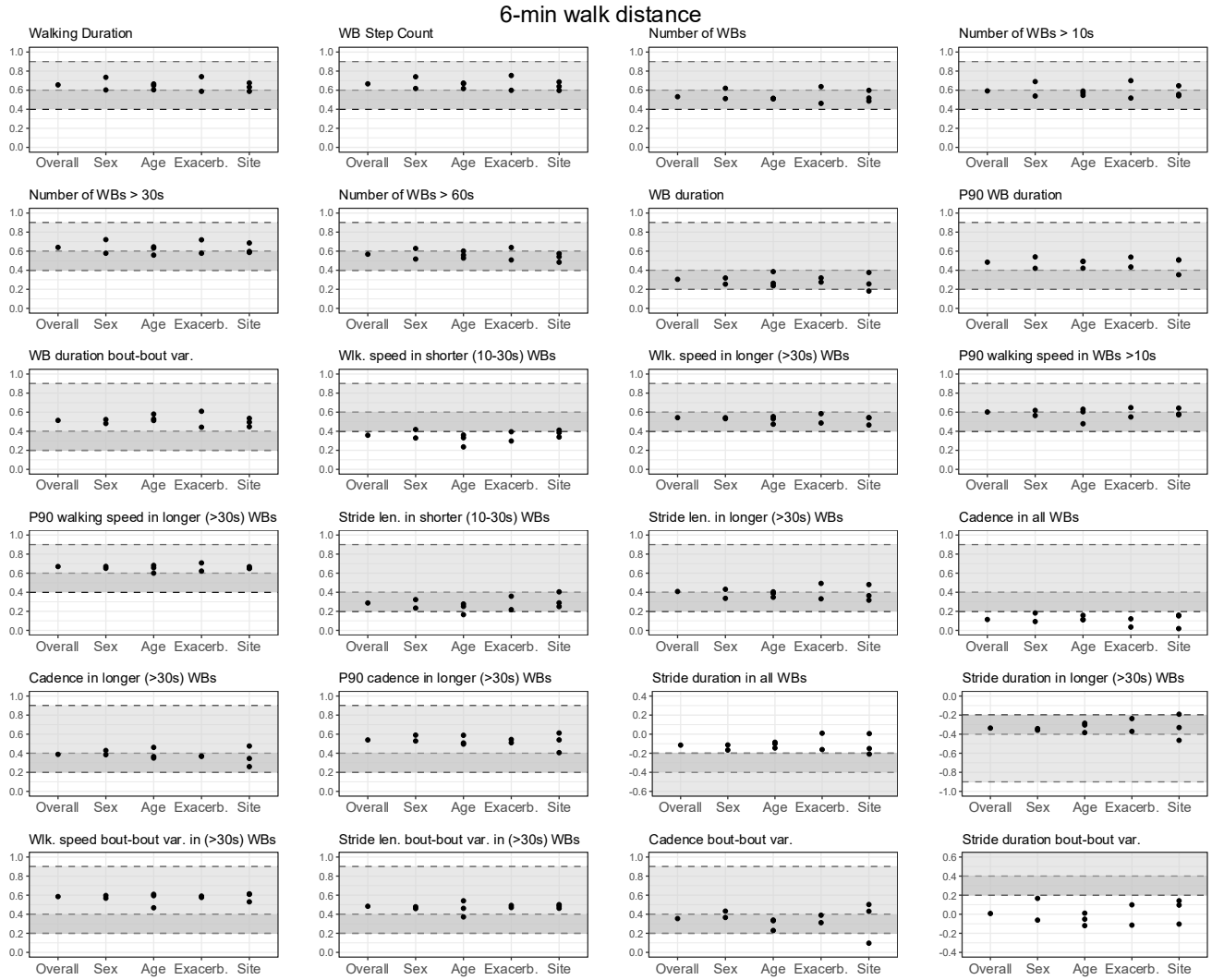

Figure S1. Convergent validity with the 6-min walk distance across subgroups of age, sex, history of moderate-to-severe exacerbation in the last 12 months, and site (Mediterranean, Continental, and Oceanic, see methods). The dark grey area represents *a priori* defined correlations, while the light grey area indicates exceeded correlations.

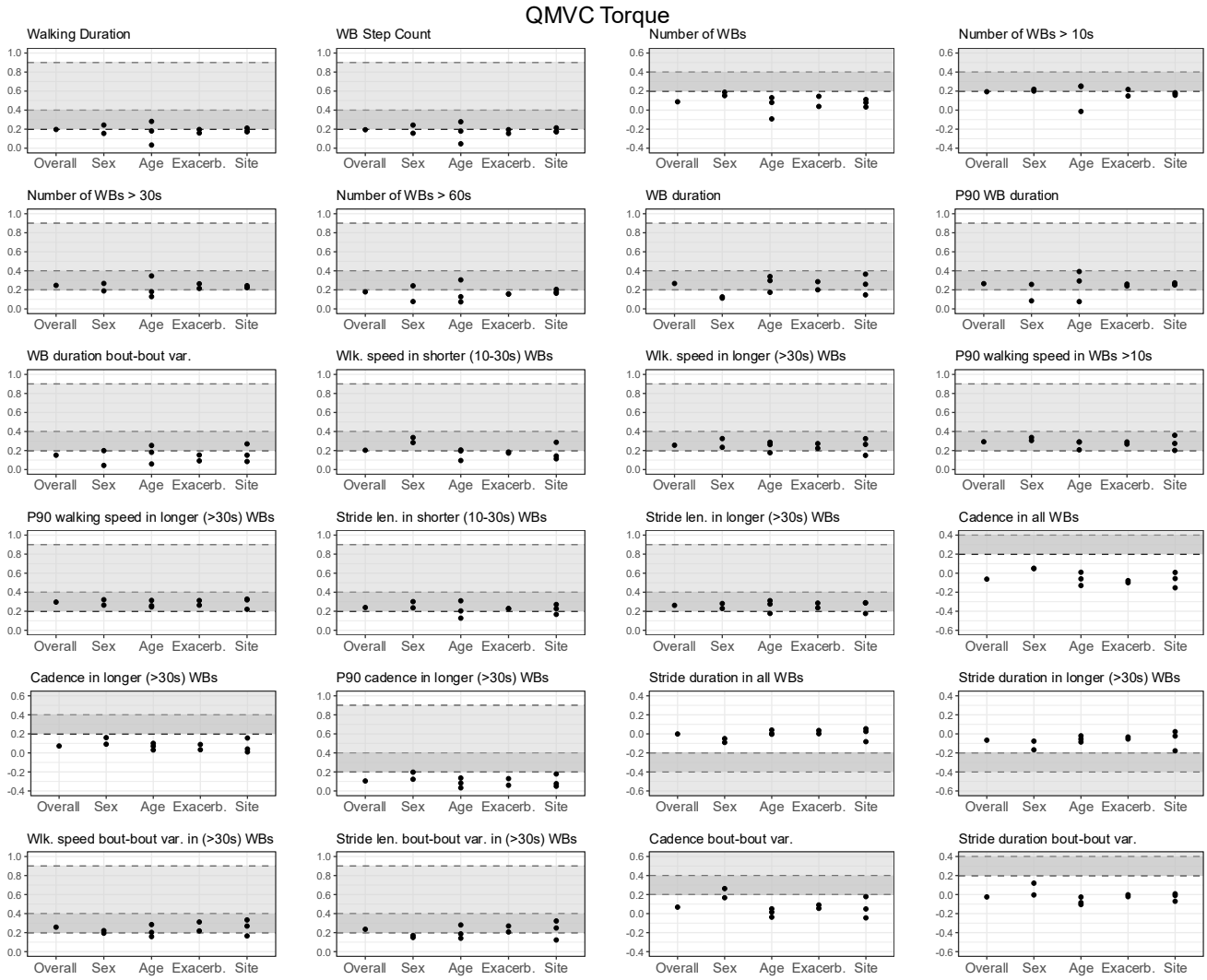

Figure S2. Convergent validity with isometric quadriceps muscle (QMVC) torque across subgroups of age, sex, history of moderate-to-severe exacerbation in the last 12 months, and site (Mediterranean, Continental, and Oceanic, see methods). The dark grey area represents a priori defined correlations, while the light grey area indicates exceeded correlations.

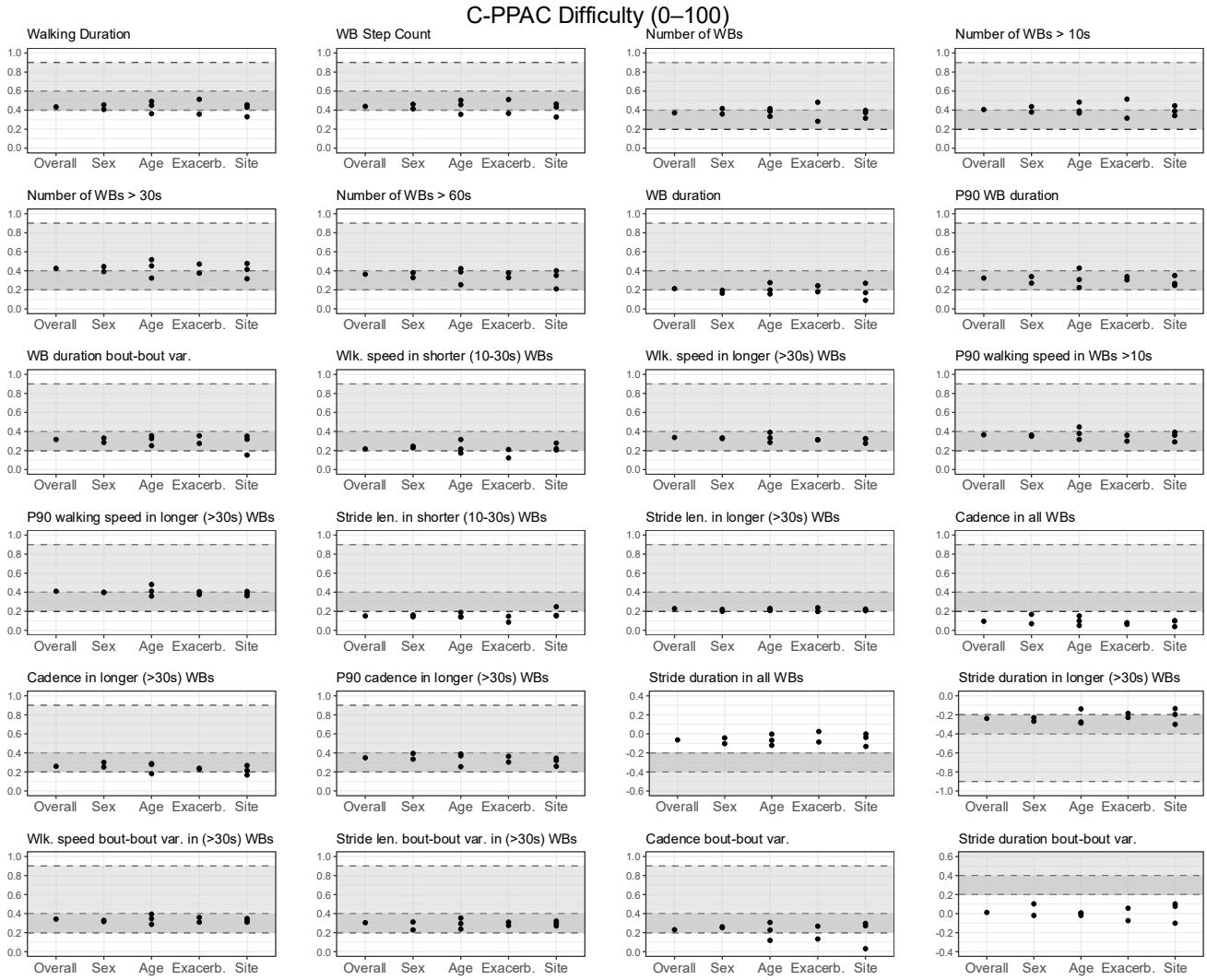

Figure S3. Convergent validity with C-PPAC Difficulty domain across subgroups of age, sex, history of moderate-to-severe exacerbation in the last 12 months, and site (Mediterranean, Continental, and Oceanic, see methods). The dark grey area represents a priori defined correlations, while the light grey area indicates exceeded correlations.

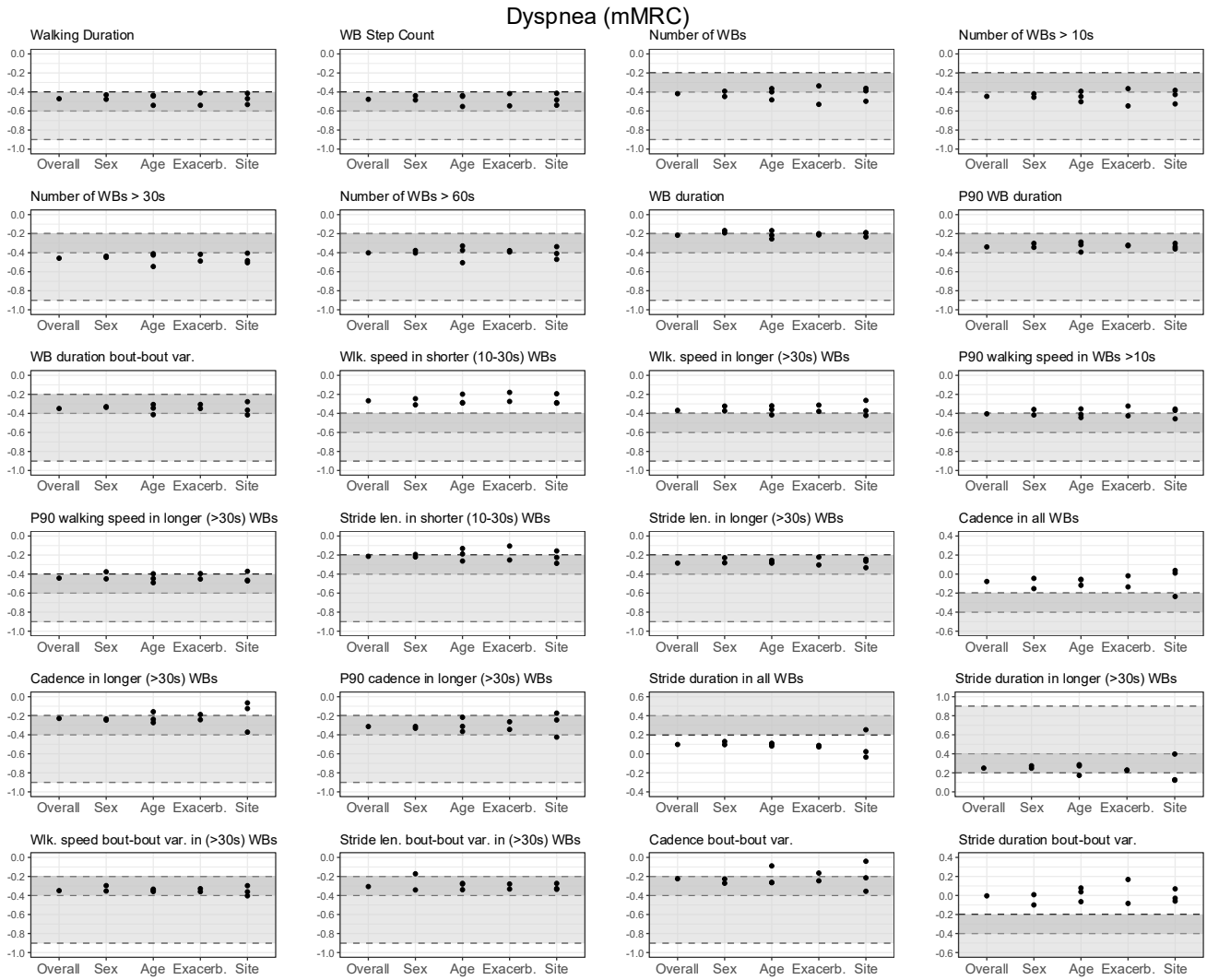

Figure S4. Convergent validity with Dyspnoea (modified Medical Research Council, mMRC) across subgroups of age, sex, history of moderate-to-severe exacerbation in the last 12 months, and site (Mediterranean, Continental, and Oceanic, see methods). The dark grey area represents a priori defined correlations, while the light grey area indicates exceeded correlations.

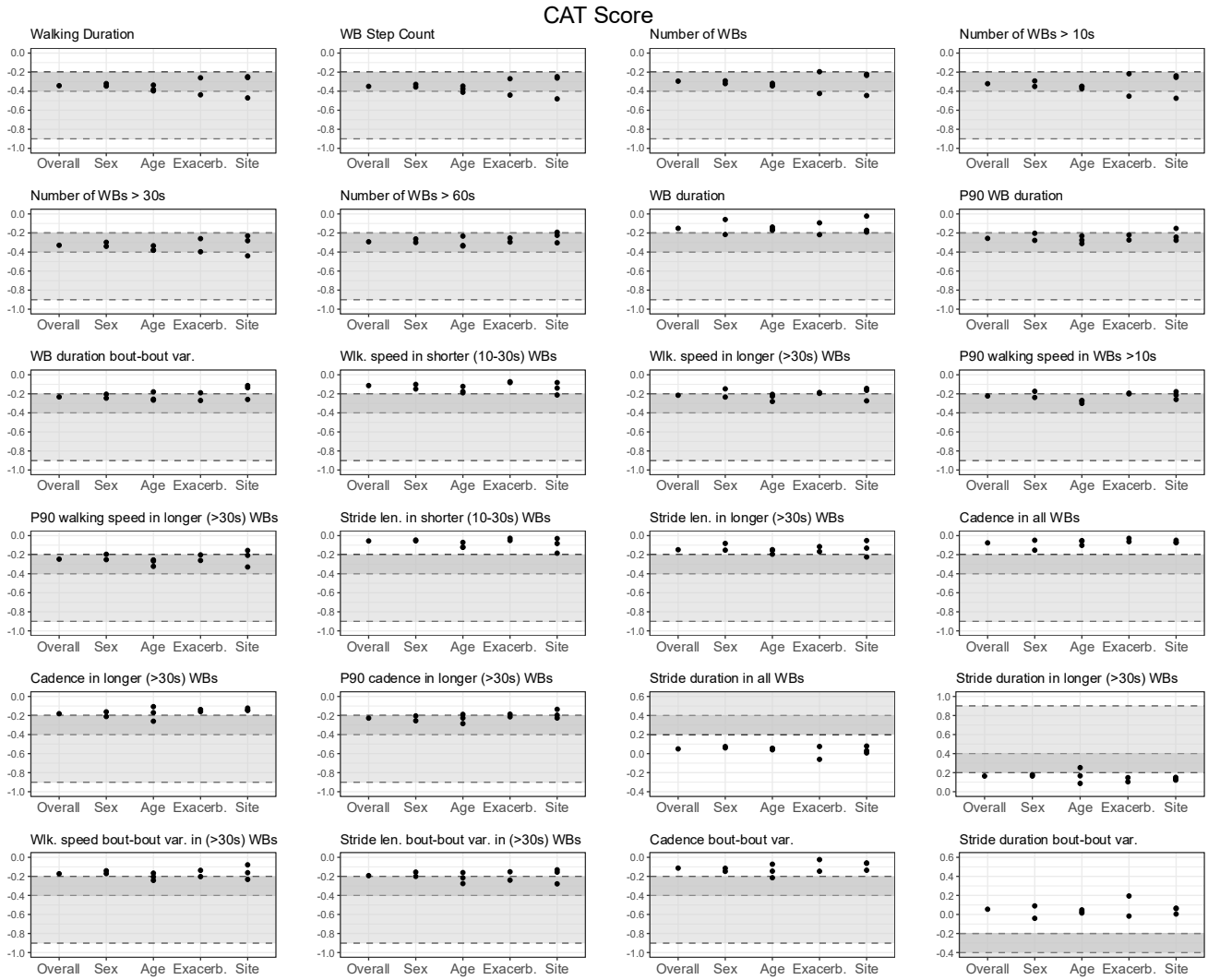

Figure S5. Convergent validity with COPD Assessment Test (CAT) scores across subgroups of age, sex, history of moderate-to-severe exacerbation in the last 12 months, and site (Mediterranean, Continental, and Oceanic, see methods). The dark grey area represents a priori defined correlations, while the light grey area indicates exceeded correlations.

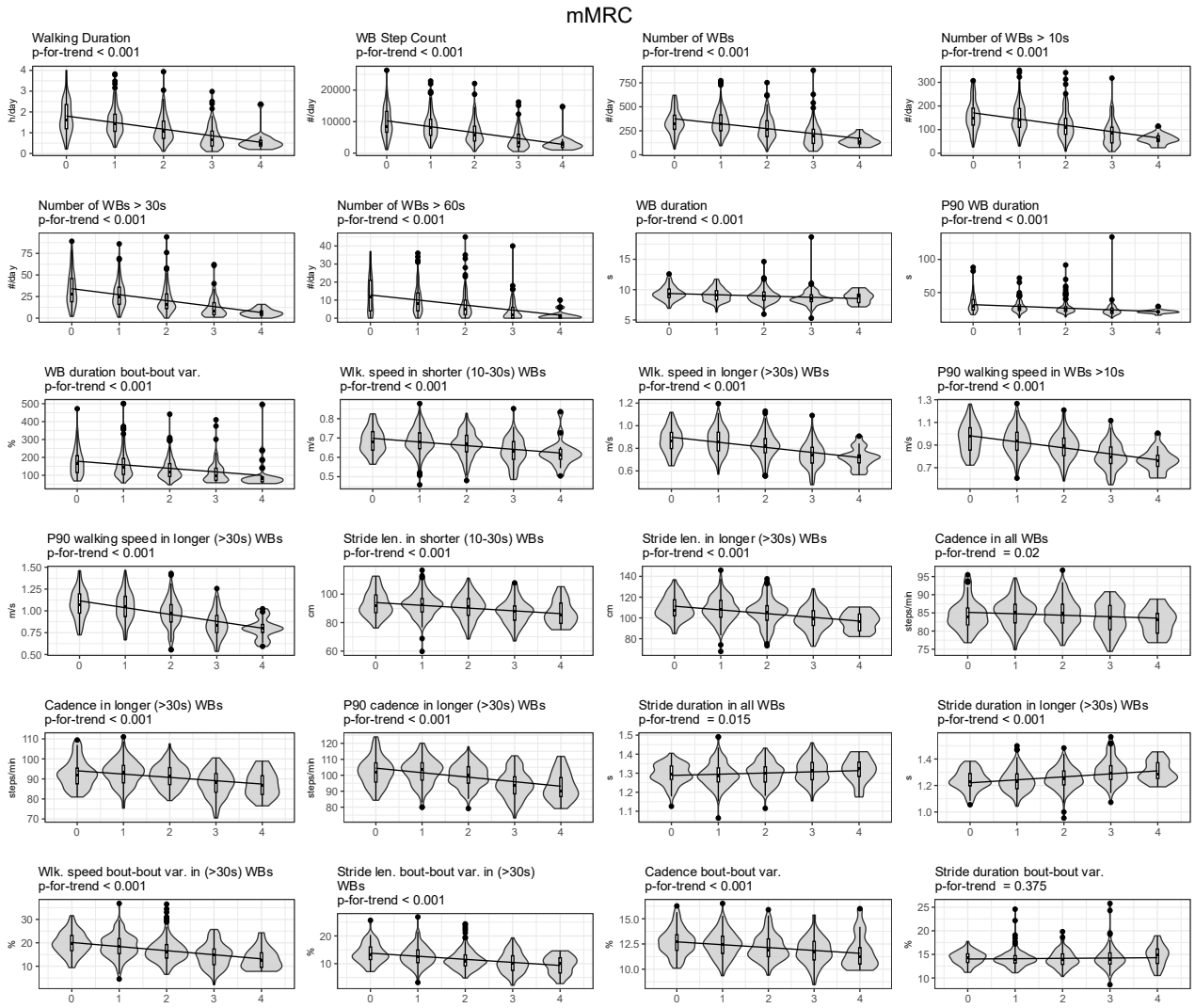

Figure S6. Distribution of DMOs across dyspnoea (mMRC) groups. P-values of linear trends were obtained using linear regression models.

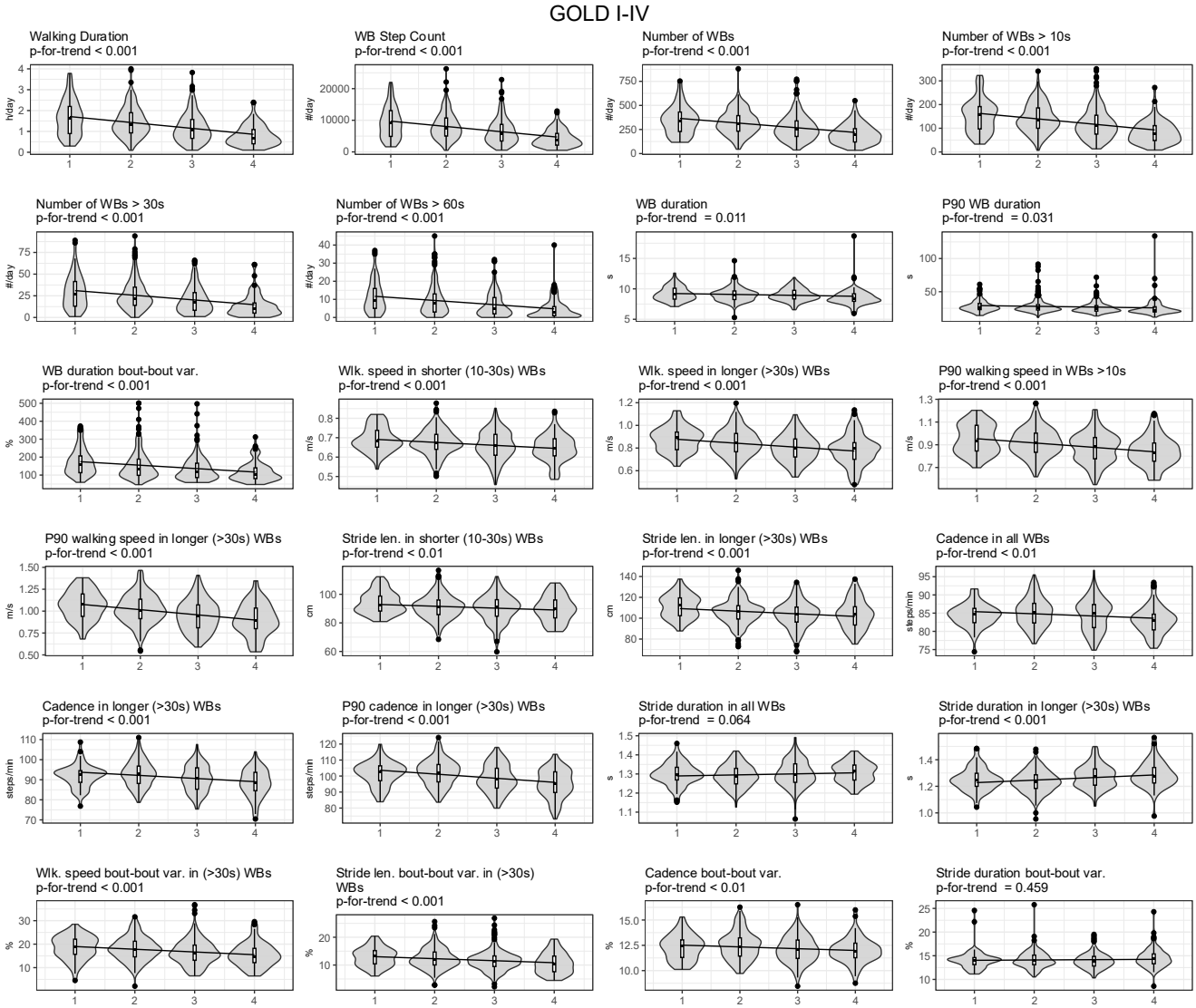

Figure S7. Distribution of DMOs across disease severity groups (GOLD I-IV). P-values of linear trends were obtained using linear regression models.

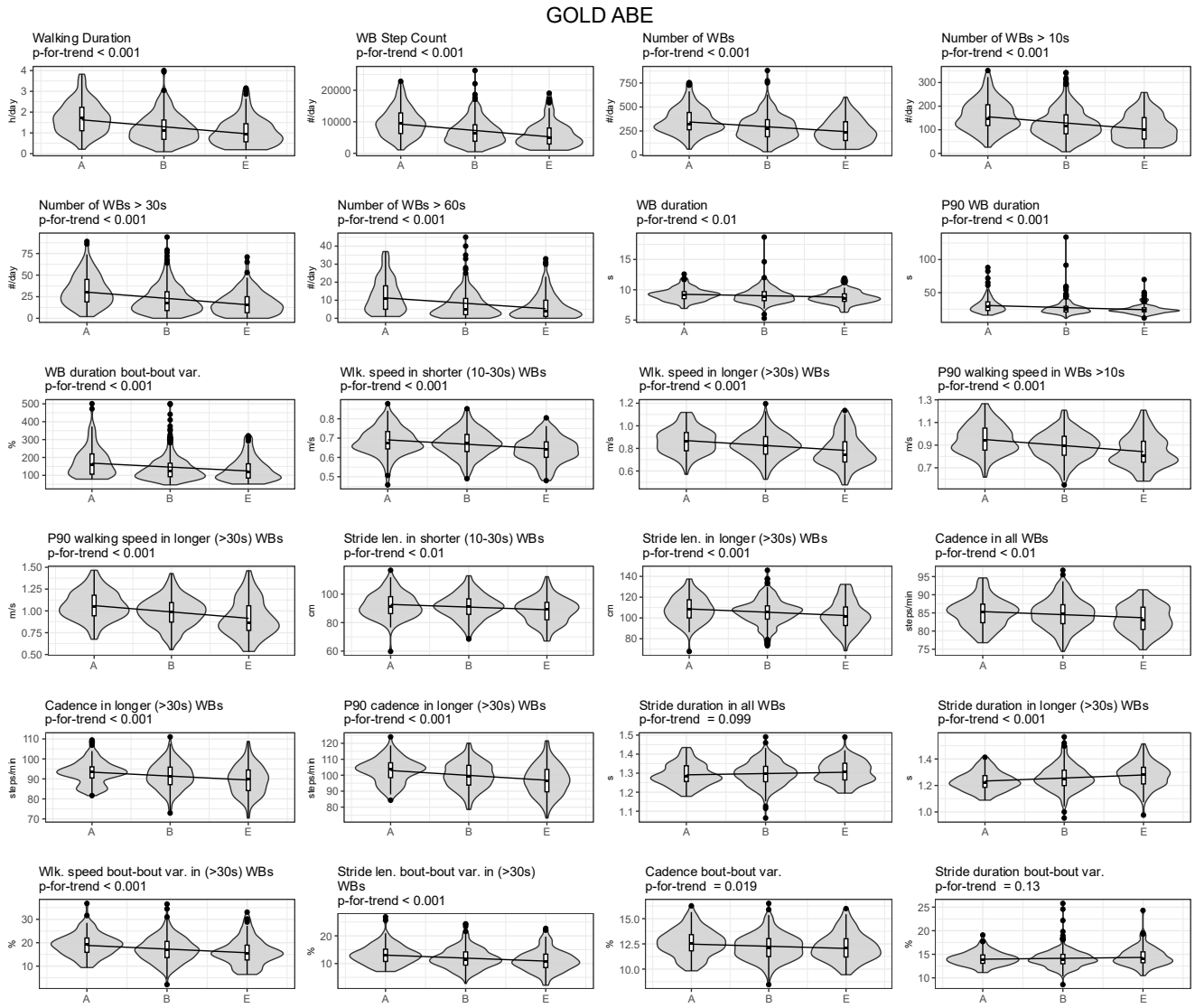

Figure S8. Distribution of DMOs across disease severity groups (GOLD ABE). P-values of linear trends were obtained using linear regression models.