COPD maintenance trials use heterogeneous outcomes and measurement instruments: a systematic literature review

Vi Truong[1], Sulaiman Bin Safar[1], Sachin Ananth[2,3], Markus Fally[4,5], Augusta Beech[1,6], Sara Souto-Miranda[7,8], Nawar Diar Bakerly[9,10], Sarah L. Gorst[11], Dave Singh[1,6], Jørgen Vestbo[1,12], Lies Lahousse[13,14] and Alexander G. Mathioudakis[1,15]

[1]Division of Immunology, Immunity to Infection and Respiratory Medicine, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester, UK. [2]Department of Respiratory Medicine, Imperial College Healthcare NHS Trust, London, UK. [3]National Heart and Lung Institute, Imperial College London, London, UK. [4]Section for Respiratory Medicine, Department of Internal Medicine, Zealand University Hospital – Roskilde, Roskilde, Denmark. [5]Department of Clinical Medicine, Faculty of Health and Medical Sciences, Copenhagen University, Copenhagen, Denmark. [6]Medicines Evaluation Unit – An IQVIA Business, Manchester, UK. [7]Department of Physiotherapy, School of Health, Polytechnic University of Setúbal, Setúbal, Portugal. [8]Comprehensive Health Research Center (CHRC), NOVA University Lisbon, Lisbon, Portugal. [9]Department of Respiratory Medicine, Salford Royal Hospital, Salford, UK. [10]School of Biological Sciences, Manchester Metropolitan University, Manchester, UK. [11]Department of Health Data Science, University of Liverpool, Liverpool, UK. [12]Department of Medicine, Section of Respiratory Medicine, Copenhagen University Hospital – Herlev and Gentofte, Copenhagen, Denmark. [13]Department of Bioanalysis, Ghent University, Ghent, Belgium. [14]Department of Epidemiology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands. [15]North West Lung Centre, Wythenshawe Hospital, Manchester University NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester, UK.

Corresponding author: Alexander G. Mathioudakis (Alexander.Mathioudakis@Manchester.ac.uk)

Shareable abstract (@ERSpublications)
COPD maintenance management trials measure heterogeneous outcomes and often overlook those that matter most to patients. This review of 240 RCTs underscores the urgent need for a globally representative, patient-centred core outcome set. https://bit.ly/42sUr2A

Cite this article as: Truong V, Bin Safar S, Ananth S, *et al*. COPD maintenance trials use heterogeneous outcomes and measurement instruments: a systematic literature review. *Eur Respir Rev* 2026; 35: 250300 [DOI: 10.1183/16000617.0300-2025].

Received: 19 Nov 2025
Accepted: 12 April 2026

Abstract

Background Randomised controlled trials (RCTs) of COPD management assess heterogeneous outcomes with diverse instruments and often omit those important to patients and healthcare professionals, limiting interpretability and comparability. This review aimed to identify the outcomes and instruments used in phase III/IV COPD maintenance management RCTs and assess their consistency.

Methods We systematically reviewed all phase III/IV RCTs registered on ClinicalTrials.gov between 2010 and 2025 evaluating COPD maintenance management. Outcomes and measurement instruments were extracted from registry entries and categorised using the COMET (Core Outcome Measures in Effectiveness Trials) taxonomy. Registered outcomes from a random 10% sample were compared with the corresponding publications to assess concordance. Outcome frequencies were summarised across intervention types and sponsor categories.

Results 43 unique outcomes were identified across 240 eligible RCTs. Physiological (89.5%), clinical (85.0%) and life impact outcomes (63.8%) were most frequently assessed, whereas resource use (39.6%), safety (36.7%) and mortality (16.7%) outcomes were less commonly reported. Only lung function (76.3%) and health-related quality of life (57.5%) appeared in over half of the trials. Exacerbations were reported in 40.4% of studies, while several patient-prioritised outcomes, particularly activities of daily living (5.4%) and exercise tolerance (18.3%), were infrequently assessed. Industry-sponsored RCTs more often reported lung function, resource use and adverse events; non-industry trials more frequently included biomarkers. Concordance between registered and published outcomes was acceptable (79.7%), although safety outcomes and instruments were sometimes under-reported.

Conclusion COPD maintenance management RCTs show substantial heterogeneity and incomplete assessment of patient-prioritised outcomes. An internationally representative, multi-stakeholder core outcome set is urgently needed to improve consistency and patient-centred evaluation in future trials.

Introduction

The management of COPD, a leading cause of death and disability globally [1, 2], remains suboptimal and patients often experience an unacceptable disease burden and adverse outcomes even if they receive optimal available treatments [3, 4]. Therefore, there is an urgent need for the development of novel, effective treatments to alter the natural history of COPD and improve its outcomes. Such novel treatments require meticulous testing through randomised controlled trials (RCTs), representing the gold standard studies for high quality and rigour [5].

Several challenges complicate the design of clinical trials of COPD maintenance management. One of the main challenges is that COPD is complex and heterogeneous, with patients experiencing varying components and symptoms of the disease, both among individuals and over time in the same individual [6–8]. To generate meaningful evidence, RCTs must measure outcomes that capture the impact of treatments across the full spectrum of disease burden. This requires selecting a balanced set of outcomes that address all key aspects of COPD. Trials also need to assess the same key outcomes, using the same outcome measurement instruments, to facilitate comparing, contrasting and pooling them together in meta-analyses [9, 10].

Unfortunately, clinical trials report heterogeneous outcomes, often omitting those that are most relevant to patients, clinicians and other stakeholders, sometimes limiting their value and ability to inform clinical practice guidelines and the care of patients [9]. This issue is often identified and highlighted in systematic reviews [11–14] and clinical guidelines [3, 15, 16] addressing maintenance management of COPD. Traditionally, regulatory authorities have also focused primarily on lung function and exacerbations as end-points for registration trials. A core outcome set, an international, multi-stakeholder consensus on a group of the most critical outcomes that should be addressed in all future clinical trials of COPD management, could address this issue [9, 10, 17].

Here, we present the results of a methodological systematic review of the outcomes assessed in clinical trials of COPD maintenance management. The objectives of this review were twofold: 1) to describe the range of outcomes and measurement instruments currently assessed in COPD maintenance RCTs, thereby providing an empirical basis to inform outcome selection in future trials, and 2) to inform the development of a core outcome set, a minimal group of critically important outcomes that should be assessed consistently across future studies.

Methods

This methodological systemic review was prospectively registered with the International Prospective Register of Systematic Reviews (PROSPERO: CRD42023466048) [18]. It was conducted and reported in accordance with the Cochrane [19], COMET (Core Outcome Measures in Effectiveness Trials) Initiative [9, 17] and PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) [20] guidelines.

Database search

We utilised the US National Library of Medicine registry of clinical trial protocols (ClinicalTrials.gov) to identify phase III/IV RCTs that assess the efficacy/effectiveness and/or safety of pharmacological and non-pharmacological interventions for maintenance management of stable COPD. We used the search terms “COPD” or “Chronic obstructive pulmonary disease”, either in the “condition/disease” or the “other terms” field of ClinicalTrials.gov. We included all studies registered between January 2010 and March 2025, with no language restrictions. The complete dataset containing all register information from all potentially eligible trials was retrieved from ClinicalTrials.gov.

Eligibility criteria

We included phase III and IV RCTs evaluating maintenance management of stable COPD in adults, including RCTs focusing on specific pulmonary treatable traits (*e.g.* airway eosinophilic inflammation in patients with COPD or pulmonary hypertension secondary to COPD). We excluded RCTs evaluating exclusively the management of COPD exacerbations (assessed previously [21]), extrapulmonary manifestations or comorbidities of COPD (such as cachexia or cardiovascular diseases in patients with COPD). No language restrictions were applied, as all protocols registered in ClinicalTrials.gov are available in English.

Data handling/extraction

A dataset encompassing all potentially suitable studies was obtained from ClinicalTrials.gov in CSV format. From this file, we extracted the following information: study characteristics (main eligibility

criteria, study sample, funder type, start and completion date, and participating countries), as well as all the reported outcomes (primary, secondary or other outcomes), measurement instruments and their measurement time-points.

We randomly selected 15% of the RCTs and cross-referenced their published reports in PubMed using their ClinicalTrials.gov registration number. More specifically, we used Microsoft Excel to randomly select studies until 36 with published reports were identified; studies without published reports were excluded from this analysis and replaced. We accessed the published reports of these manuscripts for extractions. Then, we assessed how accurately trial publications reported outcomes compared with their original study protocols, in this random subgroup of studies.

Two independent reviewers (among V. Truong, S. Bin Safar and A.G. Mathioudakis) screened each trial to determine eligibility. They subsequently completed data extraction using a predetermined customised extraction form and cross-verified each other's work for accuracy. Disagreement at any stage was addressed through discussion among all three reviewers. After reviewing the first 20 studies, we developed a mutually exclusive list of outcomes evaluated in these trials. Further trial outcomes were classified according to this list. New outcomes were added to the list based on consensus among three authors (V. Truong, S. Bin Safar and A.G. Mathioudakis). We extracted the outcomes, outcome measurement instruments and time-points of evaluation. The outcomes were classified using the COMET taxonomy into mortality/survival, clinical, physiological (objective biological, structural or physiological function measures), life impact, resource use or adverse events outcomes [22].

We quantified the proportion of trials that examined each outcome and measurement instrument, and specifically examined how frequently outcomes prioritised in three previous initiatives prioritising COPD outcomes were evaluated. Our results are presented in narrative and tabulated formats. Statistical summaries were generated using R software (version 4.3 or newer; R Foundation for Statistical Computing, Vienna, Austria).

Results

Included studies

Our search strategy identified 332 phase III and IV studies, of which 240 were deemed eligible (the selection process is summarised in figure 1). Among these, 104 RCTs (43.3%) were completed with published results, and most of the remaining were ongoing. The duration of follow-up was from 1 to 104 weeks. Across the RCTs, the median number of participants recruited (or planned for recruitment) was 257 participants, ranging between seven and 20 000 participants (interquartile range (IQR) 70–805). The interventions were inhaled drugs in 150 (62.5%) trials, systemic drugs in 62 (25.8%) trials and non-pharmacological interventions in 28 (11.7%) trials (table 1). The most frequently assessed interventions were long-acting β_2-agonists, long-acting muscarinic antagonists and inhaled corticosteroids

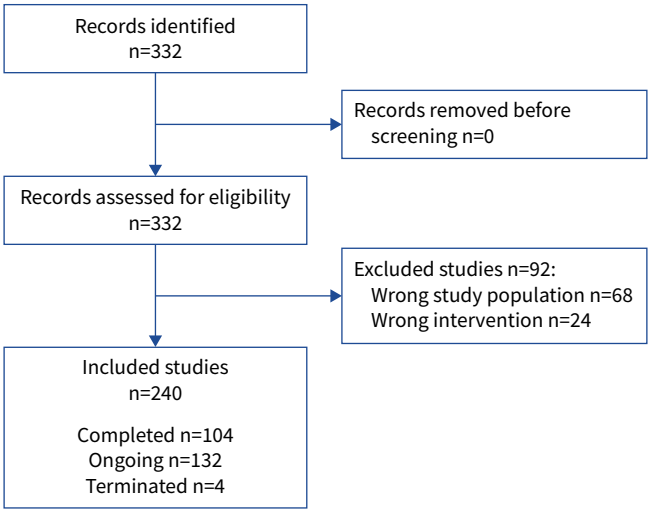

FIGURE 1 PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart.

TABLE 1 Characteristics of the included studies, stratified by intervention type (inhaled, systemic or non-pharmacological), sponsor type (industry *versus* non-industry) and phase (III *versus* IV)

	All studies	Inhaled	Systemic	Non-pharmacological	Industry sponsored	Non-industry sponsored	Phase III	Phase IV
Studies	240	150	62	28	148	92	143	97
Results available in ClinicalTrials.gov	104 (43.3)	89 (59.3)	12 (19.4)	3 (10.7)	97 (65.5)	7 (7.6)	73 (51.1)	31 (32.0)
Trial participants	256.5 (69.5–805)	390.5 (91.8–922.5)	230 (65.8–1035)	89.5 (33.8–180.5)	554.5 (189.8–1060)	80 (32–186.5)	511 (181–1021.5)	80 (38–346)
Phase								
Phase III	143 (59.6)	95 (63.3)	34 (54.8)	14 (50.0)	111 (75.0)	32 (34.8)		
Phase IV	97 (40.4)	55 (36.7)	28 (45.2)	14 (50.0)	37 (25.0)	60 (65.2)		
Interventions								
Antibiotics	12 (5)	3 (2.0)	9 (14.5)	0 (0.0)	0 (0.0)	12 (13.0)	8 (5.6)	4 (4.1)
Biologics	22 (9.2)	1 (0.7)	21 (33.9)	0 (0.0)	20 (13.5)	2 (2.2)	21 (14.7)	1 (1.0)
Chinese/alternative	6 (2.5)	0 (0.0)	6 (9.7)	0 (0.0)	3 (2.0)	3 (3.3)	6 (4.2)	0 (0.0)
LABA	37 (15.4)	37 (24.8)	0 (0.0)	0 (0.0)	29 (19.6)	8 (8.7)	23 (16.1)	14 (14.4)
LABA+ICS	46 (19.2)	46 (20.7)	0 (0.0)	0 (0.0)	36 (24.3)	10 (10.9)	28 (19.6)	18 (18.6)
LABA+LAMA	64 (26.7)	64 (42.7)	0 (0.0)	0 (0.0)	55 (37.2)	9 (9.8)	50 (35.0)	14 (14.4)
LABA+LAMA+ICS	28 (11.7)	28 (18.7)	0 (0.0)	0 (0.0)	24 (16.2)	4 (4.4)	18 (12.6)	10 (10.3)
LAMA	78 (32.5)	78 (52.0)	0 (0.0)	0 (0.0)	69 (14.6)	9 (9.8)	53 (37.1)	25 (25.8)
PDE inhibitors	28 (11.7)	8 (5.3)	20 (32.3)	0 (0.0)	15 (10.1)	13 (14.1)	16 (11.2)	12 (12.4)
Pulmonary rehabilitation	9 (3.8)	0 (0.0)	0 (0.0)	9 (32.1)	0 (0.0)	9 (9.8)	9 (6.3)	0 (0.0)
Other	69 (28.8)	16 (10.7)	27 (43.6)	26 (92.9)	18 (12.2)	51 (55.4)	26 (18.2)	49 (50.5)
Participating continents								
Africa	18 (7.5)	11 (7.3)	7 (11.3)	0 (0)	15 (10.1)	3 (3.3)	15 (10.5)	3 (3.1)
Asia	79 (32.9)	46 (30.7)	27 (43.6)	6 (21.4)	59 (39.9)	20 (21.7)	55 (38.5)	24 (24.7)
Europe	125 (52.1)	75 (50)	38 (61.3)	12 (42.9)	83 (56.1)	42 (45.7)	76 (53.2)	49 (50.5)
North America	96 (40.0)	69 (46.0)	20 (32.3)	7 (25.05)	77 (52.0)	19 (20.7)	71 (49.7)	25 (25.8)
Oceania	19 (7.9)	9 (6.0)	9 (14.5)	1 (3.6)	16 (10.8)	3 (3.3)	18 (12.6)	1 (1.0)
South America	34 (14.2)	20 (13.3)	13 (21.05)	1 (3.6)	31 (21.0)	3 (3.3)	26 (18.2)	8 (8.3)
Industry sponsored	148 (61.7)	119 (79.3)	25 (40.3)	4 (14.3)			111 (77.6)	37 (38.4)

Data are presented as n, n (%) or median (interquartile range). LABA: long-acting β_2-agonist; ICS: inhaled corticosteroid; LAMA: long-acting muscarinic antagonist; PDE: phosphodiesterase.

as monocomponents or combinations, followed by phosphodiesterase inhibitors and biologics, while the most frequently assessed non-pharmacological treatment was pulmonary rehabilitation (table 1). Less frequently assessed therapies included Chinese or alternative medicine interventions, dietary supplements, opioids, vaccines mucolytics, mechanical ventilation, app- or website-based approaches, tai chi and other behavioural, educational or complex non-pharmacological interventions. Pharmaceutical companies sponsored 148 (61.7%) studies. Most trials recruited participants in Europe (125 (52.1%)) and North America (96 (40%)), with a smaller number of studies taking place in Asia (79 (32.9%)), Oceania (19 (7.9%)) and Africa (18 (7.5%)) (figure 2). Study names and registration identifiers are listed in supplementary table S5.

Outcomes assessed

Our review identified 44 outcomes. Details of the outcomes along with their reporting frequencies in the RCTs are listed in figure 3 and table 2. Most trials reported on physiological (215 (89.5%)), clinical (204 (85.0%)) and life impact outcomes (153 (63.8%)), while fewer trials reported on resource use (95 (39.6%)), adverse event (88 (36.7%)) and mortality (40 (16.7%)) outcomes. Lung function (76.3%), health-related quality of life (57.5%), exacerbations (40.4%) and dyspnoea (30.0%) were the most frequently assessed of all identified outcomes. Lung function and COPD exacerbations emerged as the most widely adopted primary outcomes, chosen in 43.7% and 17.5% of the studies, respectively.

There were notable differences in the selection of outcomes across studies evaluating different types of interventions. Life impact outcomes were least often assessed in inhaled treatment trials (54.0%) compared with 77.4% and 85.7% in systemic and non-pharmacological studies, respectively. Among life impact outcomes, health-related quality of life appeared less frequently in inhaled therapy trials (48.7%) than in the other two groups (71.0% and 75.0%), whereas physical activity showed the opposite trend, being a

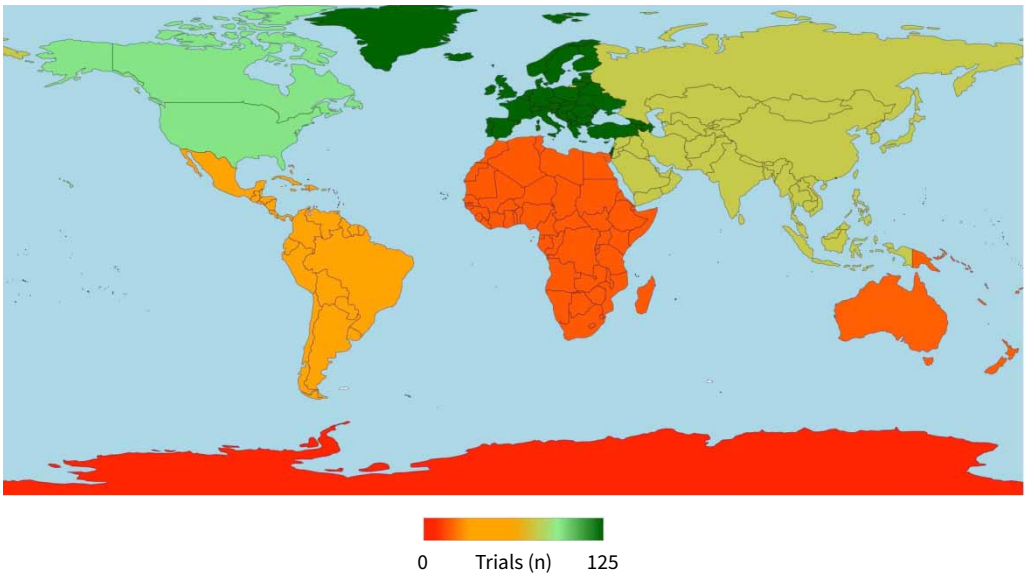

FIGURE 2 Geographic distribution of included COPD trials by continent. The map illustrates the number of trials recruiting participants in each continent. Colours represent the total number of trials conducted, with darker shades indicating higher numbers. Some trials recruited participants in multiple continents, and these were counted in each relevant region.

focus in 28.6% of non-pharmacological interventions but in only 2.0% and 1.6% of inhaled and systemic drug trials. Lung function was predominantly assessed in inhaled and systemic drug trials (88.7% and 67.7% *versus* 28.6%), while biomarkers were most frequently measured in systemic treatment trials (54.8% *versus* <20% elsewhere).

Clear distinctions were also observed between industry- and non-industry-sponsored trials. Resource use outcomes were reported nearly twice as often in industry studies (48.0% *versus* 26.1%), largely due to the frequent inclusion of rescue medication use (41.9% *versus* 7.6%). Exacerbations were also more commonly assessed in industry trials (44.6% *versus* 33.7%). Lung function was more frequently reported in industry-sponsored trials (85.8% *versus* 60.9%), whereas non-industry studies more often included biomarkers (45.7% *versus* 13.5%). Adverse events were also reported more frequently in industry trials (43.2% *versus* 26.1%), with serious events documented in 16.2% compared with 6.5% of non-industry studies.

Differences in outcome selection between phase III and phase IV trials were modest overall. Phase III trials more frequently reported resource use outcomes (48.3% *versus* 26.8%), largely driven by greater assessment of rescue medication use (40.6% *versus* 11.3%). In contrast, phase IV trials more often assessed biomarkers (35.1% *versus* 19.5%) and exercise tolerance (24.7% *versus* 14.0%).

Each of the included studies reported between one and 16 unique outcomes (median (IQR) 4 (2–7)) (figure 4a). It should be noted that some studies initially reported a larger variety of (related) outcomes; however, for the purposes of this analysis, outcomes were reclassified and merged according to the 44 uniquely identified outcomes. For instance, studies assessing multiple health-related quality of life instruments as distinct outcomes were considered to contribute a single outcome within that category. There were no significant differences between the number of outcomes reported among studies sponsored by the pharmaceutical industry or academic sponsors. There was no association between the study population and number of outcomes assessed (figure 4b).

Limited assessment of key outcomes previously prioritised for COPD trials

Key outcomes consistently prioritised for assessment in COPD trials remain infrequently assessed in the included COPD maintenance management RCTs (supplementary tables S1–S3). Among the outcomes prioritised by the 2008 European Respiratory Society (ERS)/American Thoracic Society (ATS) statement on COPD trial outcomes [23], only lung function and health-related quality of life were assessed in more

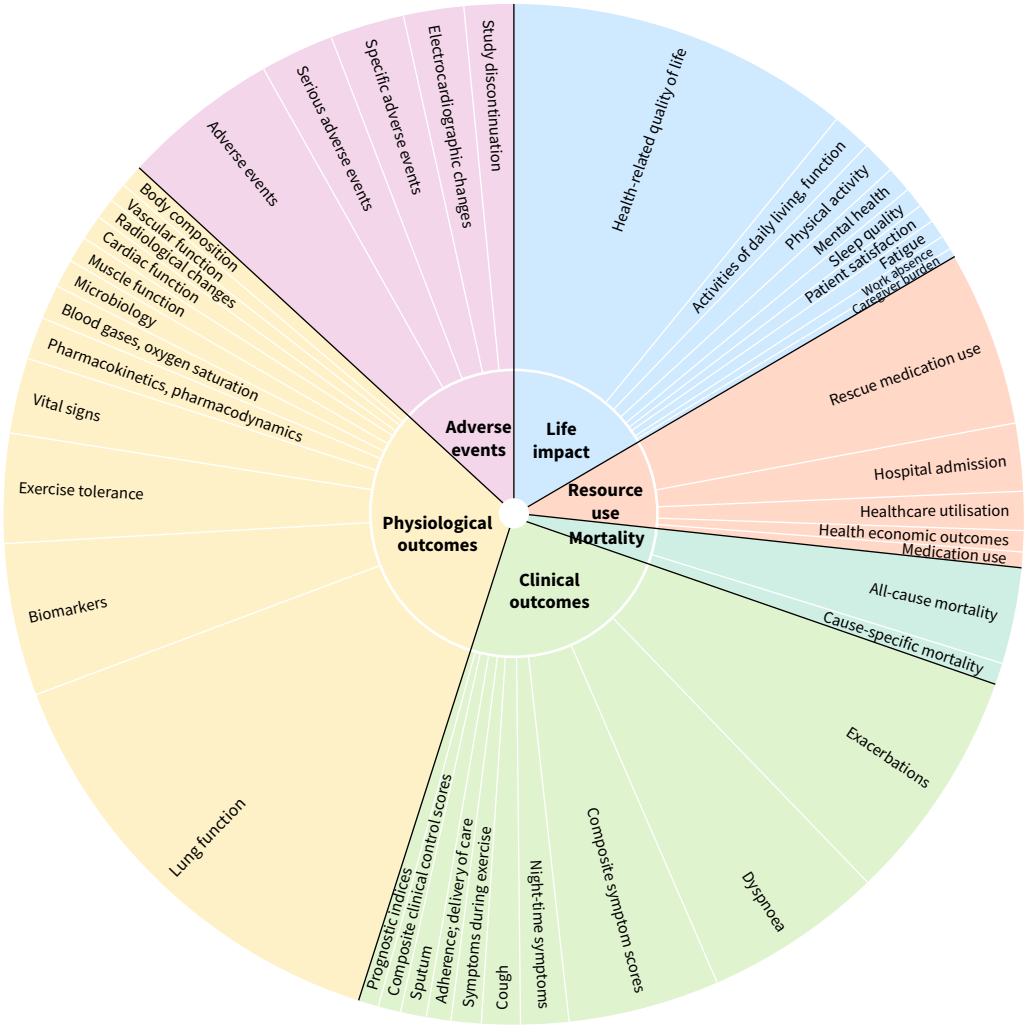

FIGURE 3 Sunburst plot illustrating the frequency of outcomes assessed in included COPD randomised controlled trials. The figure depicts the distribution and relative frequency of outcomes assessed across included trials, grouped by COMET (Core Outcome Measures in Effectiveness Trials) outcome domains (inner ring). Segment size corresponds to the number of trials reporting each outcome.

than half of the trials, while exacerbations, breathlessness and exercise tolerance were only assessed in 40.4%, 30.0% and 18.3% of trials, respectively. Functional status (corresponding to our "activities of daily living" outcome) and symptoms during exercise were only assessed in 5.4% and 3.8% of trials, respectively.

Similarly, three outcomes prioritised in the core outcome set for studies assessing pulmonary rehabilitation in COPD [24] (activities of daily living, muscle strength and fatigue) and four outcomes prioritised in the ERS COPD exacerbations core outcome set [21, 25, 26] (activities of daily living, blood gas levels, treatment adherence and death from COPD) were each assessed in <10% of trials.

Measurement instruments

Details of the measurement instruments used to assess each outcome are provided in supplementary table S4. Among the most frequently evaluated patient-reported outcome measures, certain instruments were used more often than others, demonstrating a degree of consistency in measurement across studies. The St George's Respiratory Questionnaire and the COPD Assessment Test were the most commonly used tools for assessing health-related quality of life, employed in 80.4% and 29.0% of relevant trials, respectively. Dyspnoea was most frequently measured using the Transitional Dyspnoea Index (44.4%) or the modified Medical Research Council dyspnoea scale (25.0%). In contrast, composite symptom scores were assessed

TABLE 2 Frequency of assessment of outcomes across the included COPD trials, stratified by intervention type (inhaled, systemic or non-pharmacological), sponsor type (industry *versus* non-industry) and phase (III *versus* IV)

	Trials assessing each outcome							
	All studies	Type of intervention			Sponsor		Phase	
		Inhaled	Systemic	Non-pharmacological	Industry	Non-industry	Phase III	Phase IV
Studies	240	150	62	28	148	92	143	97
Mortality	40 (16.7)	22 (14.7)	13 (21.0)	5 (17.9)	24 (16.2)	16 (17.4)	25 (17.5)	15 (15.5)
All-cause mortality	40 (16.7)	22 (14.7)	13 (21.0)	5 (17.9)	24 (16.2)	16 (17.4)	25 (17.5)	15 (15.5)
Cause-specific mortality	5 (2.1)	2 (1.3)	2 (3.2)	1 (3.6)	3 (2.0)	2 (2.2)	4 (2.8)	1 (1.0)
Life impact	153 (63.8)	81 (54.0)	48 (77.4)	24 (85.7)	87 (68.8)	66 (71.7)	94 (65.7)	59 (60.8)
Health-related quality of life	138 (57.5)	73 (48.7)	44 (71.0)	21 (75.0)	79 (53.4)	59 (64.1)	85 (59.4)	53 (54.6)
Activities of daily living, physical function	13 (5.4)	10 (6.7)	2 (3.2)	1 (3.6)	11 (7.4)	2 (2.2)	9 (6.3)	4 (4.1)
Physical activity	12 (5.0)	3 (2.0)	1 (1.6)	8 (28.6)	3 (2.0)	9 (9.8)	8 (5.6)	4 (4.1)
Mental health	7 (2.9)	4 (2.7)	2 (3.2)	1 (3.6)	4 (2.7)	3 (3.3)	3 (2.1)	4 (4.1)
Sleep quality	4 (1.7)	2 (1.3)	1 (1.6)	1 (3.6)	2 (1.4)	2 (2.2)	2 (1.3)	2 (2.1)
Patient satisfaction	4 (1.7)	0 (0.0)	2 (3.2)	2 (7.1)	2 (1.4)	2 (2.2)	1 (0.6)	1 (1.1)
Fatigue	3 (1.3)	0 (0.0)	2 (3.2)	1 (3.6)	0 (0.0)	3 (3.3)	2 (1.3)	1 (1.1)
Work absence	1 (0.4)	1 (0.7)	0 (0.0)	0 (0.0)	1 (0.7)	0 (0.0)	1 (0.6)	0 (0.0)
Caregiver burden	1 (0.4)	1 (0.7)	0 (0.0)	0 (0.0)	1 (0.7)	0 (0.0)	1 (0.6)	0 (0.0)
Resource use	95 (39.6)	59 (39.3)	25 (40.3)	11 (39.3)	71 (48.0)	24 (26.1)	69 (48.3)	26 (26.8)
Rescue medication use	69 (28.8)	54 (36.0)	15 (24.2)	0 (0.0)	62 (41.9)	7 (7.6)	58 (40.6)	11 (11.3)
Hospital admission	27 (12.1)	8 (5.3)	13 (21.0)	6 (21.4)	13 (8.8)	14 (15.2)	14 (9.8)	13 (13.4)
Healthcare utilisation	15 (6.3)	8 (5.3)	5 (8.1)	2 (7.1)	9 (6.1)	6 (6.5)	9 (6.3)	6 (6.2)
Health economic outcomes	6 (2.7)	0 (0.0)	4 (6.5)	2 (7.1)	1 (0.7)	5 (5.4)	4 (2.8)	1 (1.0)
Medication use	2 (0.8)	1 (0.7)	1 (1.6)	0 (0.0)	1 (0.7)	1 (1.1)	2 (1.3)	0 (0.0)
Clinical outcomes	204 (85.0)	124 (82.7)	57 (91.9)	23 (82.1)	128 (86.5)	76 (82.6)	126 (88.1)	78 (80.4)
Exacerbations	97 (40.4)	58 (38.7)	31 (50.0)	8 (28.6)	66 (44.6)	31 (33.7)	62 (43.4)	35 (36.1)
Dyspnoea	72 (30)	40 (26.7)	23 (37.1)	9 (32.1)	43 (29.1)	29 (31.5)	49 (34.3)	23 (23.7)
Composite symptom scores	62 (25.8)	41 (27.3)	14 (22.6)	7 (25.0)	41 (27.7)	21 (22.8)	38 (26.6)	24 (24.7)
Night-time symptoms	19 (7.9)	11 (7.3)	6 (9.7)	2 (7.1)	11 (7.4)	8 (8.7)	8 (5.6)	11 (11.3)
Cough	11 (4.6)	6 (4.0)	4 (6.5)	1 (3.6)	7 (4.7)	4 (4.3)	7 (4.9)	4 (4.1)
Symptoms during exercise	9 (3.8)	5 (3.3)	3 (4.8)	1 (3.6)	5 (3.4)	4 (4.3)	4 (2.8)	5 (5.2)
Adherence; delivery of care	8 (3.3)	0 (0.0)	3 (4.8)	5 (17.9)	3 (2.0)	5 (5.4)	3 (2.1)	5 (5.2)
Sputum	6 (2.5)	3 (2.0)	2 (3.2)	1 (3.6)	4 (2.7)	2 (2.2)	4 (2.8)	2 (2.1)
Composite clinical control scores	5 (2.1)	2 (1.3)	2 (3.2)	1 (3.6)	3 (2.0)	2 (2.2)	2 (1.3)	3 (3.1)
Prognostic indices	5 (2.1)	2 (1.3)	2 (3.2)	1 (3.6)	3 (2.0)	2 (2.2)	1 (0.6)	4 (4.1)

Continued

TABLE 2 Continued

	Trials assessing each outcome							
	All studies	Type of intervention			Sponsor		Phase	
		Inhaled	Systemic	Non-pharmacological	Industry	Non-industry	Phase III	Phase IV
Physiological outcomes	215 (89.5)	139 (92.7)	55 (88.7)	21 (75.0)	136 (91.9)	79 (85.9)	129 (90.2)	86 (88.7)
Lung function	183 (76.3)	133 (88.7)	42 (67.7)	8 (28.6)	127 (85.8)	56 (60.9)	108 (75.5)	75 (77.3)
Biomarkers	62 (25.8)	23 (15.3)	34 (54.8)	5 (17.9)	20 (13.5)	42 (45.7)	28 (19.5)	34 (35.1)
Exercise tolerance	44 (18.3)	12 (8.0)	14 (22.6)	18 (65.3)	8 (5.4)	36 (39.1)	20 (14.0)	24 (24.7)
Vital signs	31 (12.9)	21 (14.0)	9 (14.5)	1 (3.6)	23 (15.5)	8 (8.7)	18 (12.6)	13 (13.4)
Pharmacokinetics, pharmacodynamics	15 (6.3)	2 (1.3)	13 (21.0)	0 (0.0)	13 (8.8)	2 (2.2)	13 (9.1)	2 (2.1)
Blood gases, oxygen saturation	10 (4.2)	5 (3.3)	4 (6.5)	1 (3.6)	3 (2.0)	7 (7.6)	6 (4.2)	4 (4.1)
Microbiology	8 (3.3)	5 (3.3)	3 (4.8)	0 (0.0)	0 (0.0)	8 (8.7)	5 (3.5)	3 (3.1)
Muscle function	8 (3.3)	4 (2.7)	3 (4.8)	1 (3.6)	5 (3.4)	3 (3.3)	3 (2.1)	5 (5.2)
Cardiac function	6 (2.5)	3 (2.0)	2 (3.2)	1 (3.6)	4 (2.7)	2 (2.2)	5 (3.5)	1 (1.0)
Radiological changes	6 (2.5)	3 (2.0)	2 (3.2)	1 (3.6)	4 (2.7)	2 (2.2)	4 (2.8)	2 (2.1)
Vascular function	5 (2.1)	2 (1.3)	2 (3.2)	1 (3.6)	3 (2.0)	2 (2.2)	3 (2.1)	2 (2.1)
Body composition	5 (2.1)	2 (1.3)	2 (3.2)	1 (3.6)	3 (2.0)	2 (2.2)	1 (0.6)	4 (4.1)
Adverse events	88 (36.7)	58 (38.7)	24 (38.7)	6 (21.4)	64 (43.2)	24 (26.1)	56 (39.2)	32 (33.0)
Adverse events	65 (27.1)	46 (30.7)	15 (24.2)	4 (14.3)	49 (33.1)	16 (17.4)	42 (29.4)	23 (23.7)
Serious adverse events	30 (12.5)	26 (17.3)	4 (6.5)	0 (0.0)	24 (16.2)	6 (6.5)	23 (16.1)	7 (7.2)
Specific adverse events	30 (12.5)	18 (12.0)	9 (14.5)	3 (10.7)	17 (11.5)	13 (14.1)	21 (14.7)	9 (9.3)
Electrocardiographic changes	24 (10)	21 (14.0)	3 (4.8)	0 (0.0)	21 (14.2)	3 (3.3)	19 (13.3)	5 (5.2)
Study discontinuation	22 (9.2)	15 (10.0)	7 (11.3)	0 (0.0)	17 (11.5)	5 (5.4)	17 (11.9)	5 (5.2)
Adverse reactions	2 (0.1)	1 (0.7)	1 (1.6)	0 (0.0)	1 (0.7)	1 (1.1)	2 (1.3)	0 (0.0)

Data are presented as n or n (%).

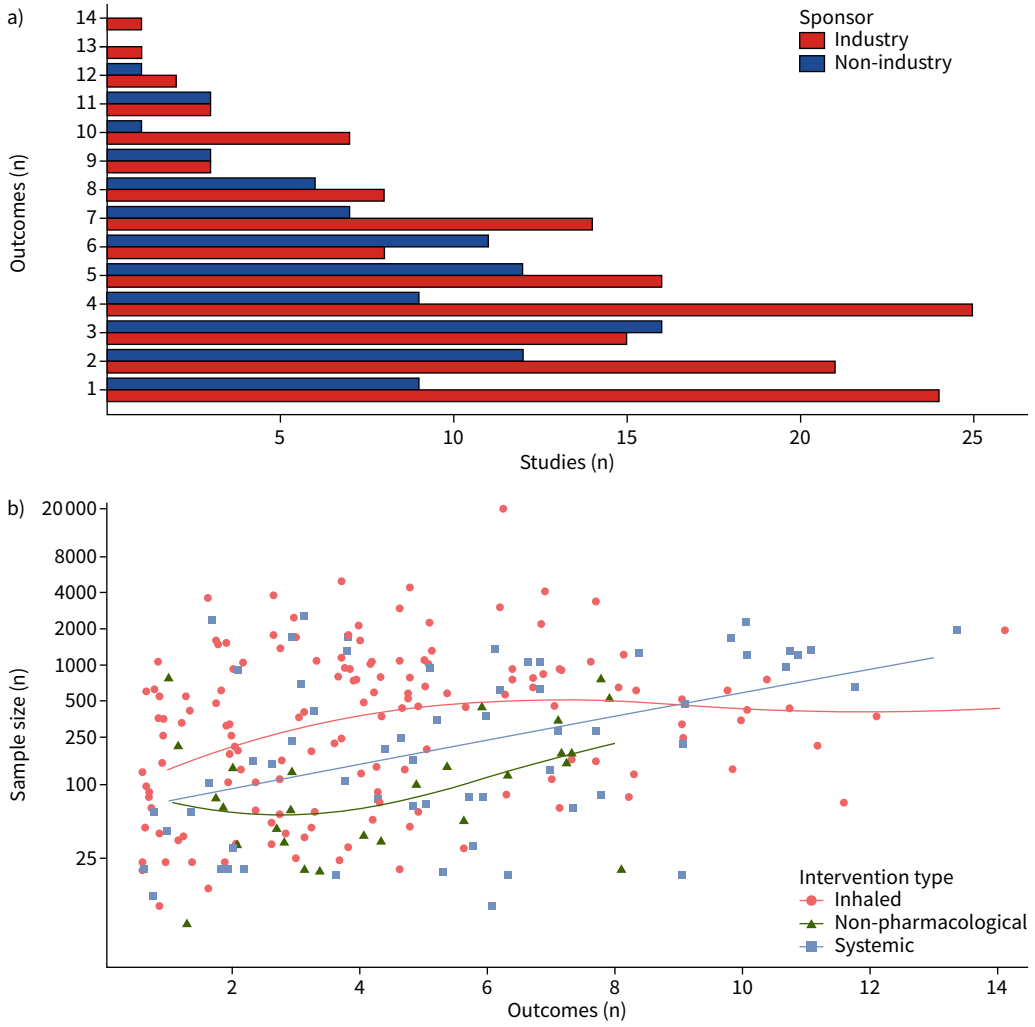

FIGURE 4 Relationship between the number of outcomes assessed and study characteristics. a) Distribution of number of outcomes reported per trial, stratified by sponsor type. b) Association between sample size and number of outcomes, stratified by intervention type.

inconsistently, often with non-validated measures, and similar heterogeneity was observed for instruments used to assess less frequently reported patient-reported outcomes.

Comparison of registered and published outcomes

We assessed 36 RCTs to compare registered and published outcomes. Among 187 outcomes reported across these trials, 14 (7.5%) were reported only in the publications, whereas 24 (12.8%) were registered in the protocols but not published. 12 trials had registered additional outcomes that were not reported in their corresponding publications. 11 trials reported additional non-registered outcomes, predominantly safety outcomes (adverse events, mortality and treatment discontinuation (n=9)), with only four trials reporting additional non-registered efficacy outcomes (such as symptoms and exacerbations) and one trial reporting pharmacokinetic data only in the publication. Furthermore, nine trials introduced additional measurement instruments for outcomes that were already registered, reflecting multiple approaches to assessing the same construct, including seven trials that reported additional lung function parameters and two trials that used additional composite symptom scores. No new outcomes or measurement instruments were identified in the publications that were not already captured in other trial protocols. Overall, a potential gap was identified, primarily in the registration of safety outcomes.

Discussion

This methodological systematic review provides the first comprehensive synthesis of outcomes and measurement instruments used in phase III and IV RCTs evaluating the maintenance management of

COPD. Most trials assessed physiological (89.5%), clinical (85.0%) and life impact (63.8%) outcomes, whereas resource use, safety and mortality outcomes were evaluated in <40%. The most frequently reported outcomes were lung function, health-related quality of life, exacerbations and dyspnoea, with only the first two measured in over half of all trials. Moreover, although instruments for measuring key patient-reported outcomes, such as health-related quality of life and breathlessness, were relatively consistent, considerable variability persisted for less frequently evaluated patient-reported and physiological outcomes. This lack of standardisation limits comparability across studies, constrains evidence synthesis, and reduces the ability of clinical trials to inform guidelines and clinical practice. These findings highlight the pressing need for the development and adoption of a standardised core outcome set for COPD maintenance management and the identification of optimal measurement instruments for each outcome.

Exacerbations are debilitating events associated with increased risks of hospitalisation and death, leaving a lasting impact on patients' quality of life and function [8]. Although guidelines, including the Global Initiative for Chronic Obstructive Lung Disease (GOLD) guidelines [3], prioritise their prevention as a key therapeutic objective, only 40% of the included trials assessed exacerbations as an outcome. Diagnostic criteria were rarely specified, and no studies reported using standardised methods to exclude common mimics such as heart failure or pulmonary embolism [27]. Mild exacerbations, despite their recognised role in accelerating disease progression, were seldom evaluated [27]. Furthermore, although exacerbations are known to be heterogeneous and to respond differently to treatment [8] (*e.g.* inhaled corticosteroids predominantly prevent non-infective events [28, 29]), none of the trials characterised exacerbation subtypes or analysed them separately. The principal barrier remains the substantial resources required for clinical assessment of exacerbations within clinical studies. Nonetheless, such efforts are essential to fully evaluate the efficacy of anti-inflammatory therapies, including inhaled corticosteroids and biologics. The increasing availability of point-of-care biomarker testing may offer a practical solution for improving exacerbation identification and phenotyping in future research [8].

Despite the availability of several outcome prioritisation exercises [23, 24, 26], our review shows that many outcomes identified by patients and health professionals as critical, particularly activities of daily living and exercise capacity, remain rarely assessed in RCTs of maintenance COPD management. Two of these prioritisation initiatives were published only recently [24, 26], which may partly explain the limited alignment observed. However, the earliest (2008 ERS/ATS statement [23]) predates the trials included in our analysis. Similarly, clinical guidelines such as GOLD [3] emphasise both exacerbation prevention and improvement in health status as key therapeutic objectives, yet the latter remains inconsistently assessed across COPD maintenance trials. The persistence of this gap underscores the need for further efforts to harmonise outcome selection. Experience from other disease areas demonstrates that the implementation of internationally representative, multi-stakeholder core outcome sets can substantially improve the consistency and relevance of outcome evaluation in clinical research [10, 30, 31]. Developing and adopting such a framework for COPD would ensure that future trials address outcomes most meaningful to patients, clinicians and policy makers alike.

COPD is a heterogeneous condition, and many interventions are designed to target specific treatable traits or aspects of disease burden. Consequently, outcome selection in individual trials should reflect the mechanism of action and therapeutic objectives of the intervention under investigation. Differences in outcome selection across intervention categories may also reflect variation in trial populations and disease severity. For example, inhaled therapy trials often enrol patients with relatively milder or more stable disease and focus on regulatory end-points such as lung function or exacerbations, which may partly explain the lower frequency of life impact outcomes observed in these studies. We also observed that biomarkers were more frequently assessed in non-industry-sponsored trials. This may reflect differences in trial objectives, as academic studies often include biomarker assessments to explore disease mechanisms or treatment response, whereas industry-sponsored trials may prioritise clinically established end-points required for regulatory evaluation. Biomarker assessments may also be more feasible in smaller mechanistic studies than in large multicentre trials. Core outcome sets are intended to standardise only a minimal set of critically important outcomes that should be assessed in all trials, reflecting priorities shared by patients, health professionals and other stakeholders. This approach does not restrict investigators from selecting their preferred primary outcomes or evaluating additional end-points tailored to the mechanism of action and objectives of the intervention under study. Rather, it ensures that key outcomes essential for understanding treatment impact are consistently captured, facilitating meaningful comparison of benefits across trials and interventions, and supporting clinical decision making and health policy. Consistent assessment of core outcomes also facilitates evidence synthesis across trials and reduces the risk that outcomes important to patients and decision makers are selectively omitted.

The main limitation of our work is that outcomes were extracted from clinical trial registrations rather than from full study reports. Although registration has been mandatory for over two decades, discrepancies

between registered and reported outcomes remain common, albeit with evidence of gradual improvement over time [32]. In our exploratory comparison, we noted potential under-reporting of safety outcomes and measurement instruments, which should be considered when interpreting our findings. This finding highlights the importance of comprehensive and consistent trial registration, as incomplete reporting of safety outcomes in trial registries may reduce transparency and complicate efforts to assess protocol adherence and selective outcome reporting. Nonetheless, no novel outcomes or instruments were identified beyond those captured in our comprehensive review of 240 registered trials, supporting the robustness of our dataset.

Another limitation is that we included only trials registered in the US National Library of Medicine's ClinicalTrials.gov database, which may have led to an over-representation of studies conducted in Western countries or sponsored by the pharmaceutical industry [33]. Moreover, by restricting our inclusion to RCTs, we may have overlooked outcomes or measurement instruments more commonly used in real-world studies which are often better powered to evaluate more rare or safety outcomes. In addition, we may have missed studies in which key registry fields used for filtering, such as study phase, were not completed in ClinicalTrials.gov. Lastly, we did not assess the analysis frameworks (*e.g.* exacerbations rate *versus* time-to-first exacerbation), and these will need to be further assessed in future studies.

Overall, this methodological review revealed substantial heterogeneity in the outcomes and measurement instruments used across clinical trials of maintenance COPD management. It also showed that several outcomes regarded as most important by patients and health professionals, such as activities of daily living and exercise tolerance, are rarely assessed. These findings underscore the urgent need for an internationally agreed, multi-stakeholder core outcome set to ensure consistent evaluation of clinically meaningful outcomes in future trials [9].

Questions for future research

- What key outcomes should be assessed consistently in future COPD maintenance management RCTs, and how can the development and implementation of an internationally representative, multi-stakeholder core outcome set best support this?
- Which measurement instruments, particularly patient-reported outcomes, demonstrate sufficient performance to be recommended as the single core instrument for each key outcome?
- How can alignment between registered and published outcomes be strengthened to enhance transparency and reduce selective reporting?

Acknowledgements: ChatGPT (version 5; https://openai.com) was used solely for language editing; all intellectual content was developed by the authors.

Provenance: Submitted article, peer reviewed.

The systematic review protocol was registered at PROSPERO with identifier number CRD42023466048.

Author contributions: Study conception: V. Truong, A.G. Mathioudakis. Study design: V. Truong, A.G. Mathioudakis. Methodological expertise: A.G. Mathioudakis. Study selection: V. Truong, S. Bin Safar, A.G. Mathioudakis. Data extraction: V. Truong, S. Bin Safar, A.G. Mathioudakis. Analysis: V. Truong, A.G. Mathioudakis. Manuscript drafting: V. Truong, A.G. Mathioudakis. Critical revision of the manuscript for intellectual content: all authors.

Conflict of interest: The authors declare no conflicts of interest directly related to this work. V. Truong, S. Bin Safar, S. Ananth, M. Fally, A. Beech, S. Souto-Miranda and S.L. Gorst report no conflicts of interest. N. Diar Bakerly reports honoraria from GlaxoSmithKline, AstraZeneca, Sanofi, Orion, Chiesi, Teva and Boehringer Ingelheim, and funding for attending educational meetings from Chiesi, GlaxoSmithKline, AstraZeneca and Sanofi, not related to this work. D. Singh reports grants from AstraZeneca, and personal consulting fees from Advocate, Aerogen, Almirall, Apogee, Arrowhead, AstraZeneca, Bial, Boehringer Ingelheim, Chiesi, Cipla, Connect Pharma, Covis, CSL Behring, DevPro Biopharma, Elpen, Empirico, EpiEndo, Genentech, Generate Biomedicines, GlaxoSmithKline, Glenmark, Gossamerbio, Kamada, Kinaset Therapeutics, Kymera, Menarini, MicroA, Novartis, OM Pharma, Orion, Pieris Pharmaceuticals, Pulmatrix, Revolo, Roivant Sciences, Sanofi, Synairgen, Tetherex, Teva, Theravance Biopharm, Upstream and Verona, not related to this work. D. Singh is a member of the *European Respiratory Review* editorial board. J. Vestbo reports personal consulting fees from AstraZeneca and GlaxoSmithKline, not related to this work. L. Lahousse reports personal consulting fees from AstraZeneca, GlaxoSmithKline and Sanofi,

honoraria for lectures paid to her institution from IPSA vzw, Domus Medica, Chiesi, Johnson & Johnson Services Inc., and support for attending meetings from AstraZeneca and Menarini, not related to this work. A.G. Mathioudakis reports consulting fees from Sanofi, honoraria from GlaxoSmithKline, and stock options in ChestPal, not related to the submitted work.

Support statement: A. Beech, D. Singh, J. Vestbo and A.G. Mathioudakis were supported by the National Institute for Health and Care Research Manchester Biomedical Research Centre (NIHR Manchester BRC; NIHR203308). S. Ananth was supported by an NIHR Academic Clinical Fellowship in Respiratory Medicine. A.G. Mathioudakis was supported by an NIHR Clinical Lectureship in Respiratory Medicine (NIHR CL-2021-06-007) and a North West Lung Centre Charity Pickering Fellowship. The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care. Funding information for this article has been deposited with the Open Funder Registry.

References

1 GBD 2021 Causes of Death Collaborators. Global burden of 288 causes of death and life expectancy decomposition in 204 countries and territories and 811 subnational locations, 1990–2021: a systematic analysis for the Global Burden of Disease Study 2021. *Lancet* 2024; 403: 2100–2132.
2 GBD 2021 Diseases and Injuries Collaborators. Global incidence, prevalence, years lived with disability (YLDs), disability-adjusted life-years (DALYs), and healthy life expectancy (HALE) for 371 diseases and injuries in 204 countries and territories and 811 subnational locations, 1990–2021: a systematic analysis for the Global Burden of Disease Study 2021. *Lancet* 2024; 403: 2133–2161.
3 Agusti A, Celli BR, Criner GJ, *et al.* Global Initiative for Chronic Obstructive Lung Disease 2023 Report: Gold Executive Summary. *Eur Respir J* 2023; 61: 2300239.
4 Stolz D, Matera MG, Rogliani P, *et al.* Current and future developments in the pharmacology of asthma and COPD: ERS seminar, Naples 2022. *Breathe* 2023; 19: 220267.
5 Guyatt G, Oxman AD, Akl EA, *et al.* GRADE guidelines: 1. Introduction – GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol* 2011; 64: 383–394.
6 Agusti A, Bafadhel M, Beasley R, *et al.* Precision medicine in airway diseases: moving to clinical practice. *Eur Respir J* 2017; 50: 1701655.
7 Agusti A, Bel E, Thomas M, *et al.* Treatable traits: toward precision medicine of chronic airway diseases. *Eur Respir J* 2016; 47: 410–419.
8 Mathioudakis AG, Janssens W, Sivapalan P, *et al.* Acute exacerbations of chronic obstructive pulmonary disease: in search of diagnostic biomarkers and treatable traits. *Thorax* 2020; 75: 520–527.
9 Williamson PR, Altman DG, Bagley H, *et al.* The COMET handbook: version 1.0. *Trials* 2017; 18: Suppl. 3, 280.
10 Mathioudakis AG, Khaleva E, Fally M, *et al.* Core outcome sets, developed collaboratively with patients, can improve the relevance and comparability of clinical trials. *Eur Respir J* 2023; 61: 2202107.
11 Janjua S, Mathioudakis AG, Fortescue R, *et al.* Prophylactic antibiotics for adults with chronic obstructive pulmonary disease: a network meta-analysis. *Cochrane Database Syst Rev* 2021; 1: CD013198.
12 Ammous O, Feki W, Lotfi T, *et al.* Inspiratory muscle training, with or without concomitant pulmonary rehabilitation, for chronic obstructive pulmonary disease (COPD). *Cochrane Database Syst Rev* 2023; 1: CD013778.
13 Lenferink A, Brusse-Keizer M, van der Valk PD, *et al.* Self-management interventions including action plans for exacerbations versus usual care in patients with chronic obstructive pulmonary disease. *Cochrane Database Syst Rev* 2017; 8: CD011682.
14 Miravitlles M, Urrutia G, Mathioudakis AG, *et al.* Efficacy and safety of tiotropium and olodaterol in COPD: a systematic review and meta-analysis. *Respir Res* 2017; 18: 196.
15 Dabscheck E, George J, Hermann K, *et al.* COPD-X Australian guidelines for the diagnosis and management of chronic obstructive pulmonary disease: 2022 update. *Med J Aust* 2022; 217: 415–423.
16 Nici L, Mammen MJ, Charbek E, *et al.* Pharmacologic management of chronic obstructive pulmonary disease. An Official American Thoracic Society Clinical Practice Guideline. *Am J Respir Crit Care Med* 2020; 201: e56–e69.
17 Kirkham JJ, Davis K, Altman DG, *et al.* Core Outcome Set-STAndards for Development: the COS-STAD recommendations. *PLoS Med* 2017; 14: e1002447.
18 Truong V, Bin Safar S, Felton T, *et al.* Systematic review of the outcomes and outcome measurement instruments reported in randomised controlled trials on the management of stable chronic obstructive pulmonary disease. PROSPERO. 2024. www.crd.york.ac.uk/PROSPERO/view/CRD42023466048 Date last accessed: 24 April 2026.
19 Higgins JPT, Thomas J, Chandler J, *et al.* Cochrane Handbook for Systematic Review of Interventions. Version 6.0 (updated July 2019). 2019. www.cochrane.org/handbook Date last accessed: 24 April 2026.
20 Page MJ, McKenzie JE, Bossuyt PM, *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *J Clin Epidemiol* 2021; 134: 178–189.

21 Mathioudakis AG, Moberg M, Janner J, *et al.* Outcomes reported on the management of COPD exacerbations: a systematic survey of randomised controlled trials. *ERJ Open Res* 2019; 5: 00072-2019.
22 Dodd S, Clarke M, Becker L, *et al.* A taxonomy has been developed for outcomes in medical research to help improve knowledge discovery. *J Clin Epidemiol* 2018; 96: 84–92.
23 Cazzola M, MacNee W, Martinez FJ, *et al.* Outcomes for COPD pharmacological trials: from lung function to biomarkers. *Eur Respir J* 2008; 31: 416–469.
24 Souto-Miranda S, Saraiva I, Spruit MA, *et al.* Core outcome set for pulmonary rehabilitation of patients with COPD: results of a modified Delphi survey. *Thorax* 2023; 78: 1240–1247.
25 Mathioudakis AG, Abroug F, Agusti A, *et al.* Core outcome set for the management of acute exacerbations of chronic obstructive pulmonary disease: the COS-AECOPD ERS Task Force study protocol. *ERJ Open Res* 2020; 6: 00193-2020.
26 Mathioudakis AG, Abroug F, Agusti A, *et al.* ERS statement: a core outcome set for clinical trials evaluating the management of COPD exacerbations. *Eur Respir J* 2022; 59: 2102006.
27 Mathioudakis AG, Janner J, Moberg M, *et al.* A systematic evaluation of the diagnostic criteria for COPD and exacerbations used in randomised controlled trials on the management of COPD exacerbations. *ERJ Open Res* 2019; 5: 00136-2019.
28 Mathioudakis AG, Bate S, Sivapalan P, *et al.* Rethinking blood eosinophils for assessing inhaled corticosteroids response in COPD: a post-hoc analysis from the FLAME trial. *Chest* 2024; 166: 987–997.
29 Martinez FJ, Vestbo J, Anderson JA, *et al.* Effect of fluticasone furoate and vilanterol on exacerbations of chronic obstructive pulmonary disease in patients with moderate airflow obstruction. *Am J Respir Crit Care Med* 2017; 195: 881–888.
30 Fally M, Mathioudakis AG, Wingfield Digby J, *et al.* Outcomes assessed in therapeutic randomized controlled trials in hospitalized patients with COVID-19: is the meta Core Outcome Set (meta-COS) adopted? *Clin Microbiol Infect* 2022; 28: 456–457.
31 Hughes KL, Clarke M, Williamson PR. A systematic review finds Core Outcome Set uptake varies widely across different areas of health. *J Clin Epidemiol* 2021; 129: 114–123.
32 Serpas VJ, Raghav KP, Halperin DM, *et al.* Discrepancies in endpoints between clinical trial protocols and clinical trial registration in randomized trials in oncology. *BMC Med Res Methodol* 2018; 18: 169.
33 Banno M, Tsujimoto Y, Kataoka Y. Studies registered in non-ClinicalTrials.gov accounted for an increasing proportion of protocol registrations in medical research. *J Clin Epidemiol* 2019; 116: 106–113.