EUROPEAN RESPIRATORY journal

FLAGSHIP SCIENTIFIC JOURNAL OF ERS

Early View

Task force report

ERS Statement: A core outcome set for clinical trials evaluating the management of chronic obstructive pulmonary disease (COPD) exacerbations

Alexander G. Mathioudakis, Fekri Abroug, Alvar Agusti, Sachin Ananth, Per Bakke, Konstantinos Bartziokas, Bianca Beghe, Andras Bikov, Thomas Bradbury, Guy Brusselle, Cordula Cadus, Courtney Coleman, Marco Contoli, Alexandru Corlateanu, Olga Corlateanu, Gerard J. Criner, Balazs Csoma, Alexander Emelyanov, Rosa Faner, Gustavo Fernandez Romero, Zeineb Hammouda, Peter Horváth, Arturo Huerta Garcia, Michael Jacobs, Christine Jenkins, Guy Joos, Olga Kharevich, Konstantinos Kostikas, Elena Lapteva, Zsofia Lazar, Joerg D. Leuppi, Carol Liddle, John Linnell, Alejandra López-Giraldo, Vanessa M. McDonald, Rune Nielsen, Alberto Papi, Isabel Saraiva, Galina Sergeeva, Agni Sioutkou, Pradeesh Sivapalan, Elizabeth Stovold, Hao Wang, Fuqiang Wen, Janelle Yorke, Paula R. Williamson, Jørgen Vestbo, Jens-Ulrik Jensen

Please cite this article as: Mathioudakis AG, Abroug F, Agusti A, *et al*. ERS Statement: A core outcome set for clinical trials evaluating the management of chronic obstructive pulmonary disease (COPD) exacerbations. *Eur Respir J* 2021; in press (https://doi.org/10.1183/13993003.02006-2021).

This manuscript has recently been accepted for publication in the *European Respiratory Journal*. It is published here in its accepted form prior to copyediting and typesetting by our production team. After these production processes are complete and the authors have approved the resulting proofs, the article will move to the latest issue of the ERJ online.

ERS Statement: A core outcome set for clinical trials evaluating the management of chronic obstructive pulmonary disease (COPD) exacerbations.

Alexander G. Mathioudakis[1,2*], Fekri Abroug[3], Alvar Agusti[4,5], Sachin Ananth[6], Per Bakke[7], Konstantinos Bartziokas[8], Bianca Beghe[9], Andras Bikov[1,2], Thomas Bradbury[10], Guy Brusselle[11,12], Cordula Cadus[13,14], Courtney Coleman[15], Marco Contoli[16], Alexandru Corlateanu[17], Olga Corlateanu[18], Gerard J. Criner[19], Balazs Csoma[20], Alexander Emelyanov[21], Rosa Faner[4], Gustavo Fernandez Romero[19], Zeineb Hammouda[3], Peter Horváth[20], Arturo Huerta Garcia[4,22], Michael Jacobs[19], Christine Jenkins[10], Guy Joos[12], Olga Kharevich[23], Konstantinos Kostikas[8], Elena Lapteva[23], Zsofia Lazar[20], Joerg D. Leuppi[13,14], Carol Liddle[15], John Linnell[15], Alejandra López-Giraldo[4,5], Vanessa M. McDonald[24,25,26], Rune Nielsen[7,27], Alberto Papi[16], Isabel Saraiva[15], Galina Sergeeva[21], Agni Sioutkou[8], Pradeesh Sivapalan[28,29], Elizabeth Stovold[30], Hao Wang[31], Fuqiang Wen[31], Janelle Yorke[32,33], Paula R. Williamson[34], Jørgen Vestbo[1,2#], Jens-Ulrik Jensen[28,35*], on behalf of the DECODE-NET.

*AGM and J-UJ are the co-chairs of the COS-AECOPD ERS Task Force.

[#] Corresponding author.

Affiliations:

[1] Division of Infection, Immunity and Respiratory Medicine, School of Biological Sciences, The University of Manchester, Manchester Academic Health Science Centre, UK.

[2] The North West Lung Centre, Wythenshawe Hospital, Manchester University NHS Foundation Trust, Manchester, UK

[3] Hospital Fattouma Bourguiba, Monastir, Tunisia.

[4] Respiratory Institute, Hospital Clinic and Catedra de Salud Respiratoria, University of Barcelona, Spain.

[5] Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), National Spanish Network for Respiratory Research (CIBERES), Barcelona, Spain.

[6] West Hertfordshire Hospital NHS Trust, Watford, UK.

[7] Department of Clinical Science, University of Bergen, Bergen, Norway.

[8] Respiratory Medicine Department, University of Ioannina School of Medicine, Ioannina, Greece.

[9] Section of Respiratory Diseases, Department of Medicine, University of Modena and Reggio Emilia, Modena, Italy.

[10] The George Institute for Global Health, University of New South Wales, Sydney, Australia.

[11] Departments of Epidemiology and Respiratory Medicine, Erasmus Medical Center, Rotterdam, the Netherlands.

[12] Department of Respiratory Medicine, Ghent University Hospital, Ghent, Belgium.

[13] University Department of Medicine, Cantonal Hospital Basell and Liestal, Switzerland.

[14] Faculty of Medicine, University of Basel, Basel, Switzerland.

[15] European Lung Foundation, Sheffield, UK.

[16] Research Center on Asthma and COPD, Department of Medical Sciences, University of Ferrara, Ferrara, Italy.

[17] Department of Respiratory Medicine, State University of Medicine and Pharmacy "Nicolae Testemitanu", Chisinau, Moldova.

[18] Department of Internal Medicine, State University of Medicine and Pharmacy "Nicolae Testemitanu", Chisinau, Moldova.

[19] Department of Thoracic Medicine and Surgery, Lewis Katz School of Medicine at Temple University, Philadelphia, Pennsylvania, USA.

[20] Department of Pulmonology, Semmelweis University, Budapest, Hungary.

[21] Department of Respiratory Medicine and Allergy, North-Western Medical University St. Petersburg, Russian Federation.

[22] Respiratory Intensive Care Division, Clinica Sagrada Familia, Barcelona, Spain.

[23] Department of Pulmonology and Tuberculosis, Belarusian Medical Academy of Postgraduate Education.

[24] Department of Respiratory and Sleep Medicine, Medical and Interventional Services, John Hunter Hospital, Hunter New England Local Health District, New Lambton Heights, New South Wales, Australia.

[25] Priority Research Centre for Healthy Lungs, University of Newcastle, Callaghan, New South Wales, Australia.

[26] School of Nursing and Midwifery, University of Newcastle, Callaghan, New South Wales, Australia.

[27] Department of Thoracic Medicine, Haukeland University Hospital, Bergen, Norway.

[28] Section of Respiratory Medicine, Department of Internal Medicine, Herlev-Gentofte Hospital, Hellerup, Denmark.

[29] Department of Internal Medicine, Zealand University Hospital, Roskilde, Denmark.

[30] Cochrane Airways Group, Population Health Research Institute, St George's University of London, London, UK.

[31] Department of Respiratory and Critical Care Medicine, West China Hospital/West China School of Medicine, Sichuan University, Chengdu, China.

[32] School of Health Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK.

[33] Christie Patient Centred Research, The Christie NHS Foundation Trust, Manchester, UK.

[34] MRC/NIHR Trials Methodology Research Partnership, Department of Health Data Science, University of Liverpool (a member of Liverpool Health Partners), Liverpool, UK.

[35] Institute of Clinical Medicine, Faculty of Health Sciences, University of Copenhagen, Denmark.

Corresponding author:

Prof. Jørgen Vestbo DMSc, FRCP, FERS, FMedSci

Professor of Respiratory Medicine.

Division of Infection, Immunity and Respiratory Medicine,

School of Biological Sciences,

The University of Manchester, UK

Jorgen.Vestbo@Manchester.ac.uk

Keywords: Chronic Obstructive Pulmonary Disease, COPD, COPD exacerbations, Outcomes, Endpoints, Core outcome set, Randomized controlled trials, Systematic reviews, consensus, clinical trial methods.

Author contribution: AGM, JV and J-UJ conceived this study. AGM, JY, PRW, JV and J-UJ contributed to study design. AGM, PRW, JV and J-UJ designed, conducted, and analysed the Delphi Survey. AGM, PRW, JV and J-UJ designed the consensus meetings. All authors contributed to the dissemination of the Delphi survey and the development of consensus. AGM, PW, JV and J-UJ prepared the initial draft of the manuscript. All authors contributed to critical revision of the paper for intellectual content.

Conflicts of Interest: This study is funded by the European Respiratory Society (TF-2019-12). All authors have completed the ICMJE uniform disclosure form detailing any conflicts of interest outside the submitted work that they may have. None of the authors have conflicts directly related to this work. In the conduct of this study, potential conflicts of interest were declared and handled in accordance with ERS policy.

Funding: This study is funded by the European Respiratory Society (ERS TF-2019-12). AGM, AB and JV are supported by the NIHR Manchester Biomedical Research Centre (BRC). PRW is supported by the Medical Research Council (MRC) Trials Methodology Research Partnership (grant reference MR/S014357/1). The research salary for PS was provided by Herlev-Gentofte University Hospital.

Acknowledgement: We are thankful to the ERS and the European Lung Foundation (ELF) for supporting the conduct of all parts of this study (special thanks to Courtney Coleman and Jessica Denning); to the COPD Patient Advisory Group of the ELF for reviewing the plain language

material of our study; and to the COMET team, for providing input in the development of the patient interview questions; to Sara Brookes (Institute of Cancer and Genomic Sciences, University of Birmingham) for acting as an impartial moderator for our consensus meeting. We are grateful to all experts and lay respondents who completed our Delphi survey and to all the patients, carers of patients with COPD and patients' representatives who have contributed to our consensus meetings (listed in the online appendix). We are also thankful to the following organizations for distributing the Delphi Survey to their memberships and/or through their social media: Allergy and Asthma Network, Alpha-1 Netherlands, Alpha-1 Spain, Alpha-1Plus (Belgium), Asian Pacific Society of Respirology (APSR), the Association of Pulmonologists of Greece, Australian Lung Foundation, Brazilian Respiratory Society, British Lung Foundation (BLF), COPD Canada, COPD Foundation (USA), COPD Ireland, Danish Lung Association, Dutch Lung Foundation, Georgian Respiratory Association, Global Allergy & Airways Patient Platform, Greek Association of General Practitioners, Hellenic Thoracic Society, Hungarian Respiratory Society, Indonesian Respiratory Society, Irish Thoracic Society, Jedra Organisation to Help Those Suffering of Lung Cancer and Other Lung Diseases (Croatia), Kazakhstan Respiratory Society, Lovexair (Spain), National Institute for Health Research (NIHR) BEAT Respiratory Disease, NTM Info & Research (USA), Pan African Thoracic Society (PATS), Philippine College of Chest Physicians, Respiriamo Insieme (Italy), Russian Respiratory Society, Swiss Society of Pulmonology, Sociedad Española de Neumología y Cirugía Torácica (SEPAR), Swedish Heart and Lung Association, Task Force for Lung Health, Thoracic Society of Australia and New Zealand (TSANZ), US COPD Coalition, Turkish Respiratory Society.

Abstract

Clinical trials evaluating the management of acute exacerbations of COPD assess heterogeneous outcomes, often omitting those that are clinically relevant or more important to patients. We have developed a core outcome set, a consensus-based minimum set of important outcomes that we recommend are evaluated in all future clinical trials on exacerbations management, to improve their quality and comparability.

COPD exacerbations outcomes were identified through methodological systematic reviews and qualitative interviews with 86 patients from 11 countries globally. The most critical outcomes were prioritized for inclusion in the core outcome set through a two-round Delphi survey that was completed by 1,063 participants (256 patients, 488 health professionals and 319 clinical academics) from 88 countries in 5 continents. Two global, multi-stakeholder, virtual consensus meetings were conducted to (i) finalize the core outcome set and (ii) prioritize a single measurement instrument to be used for evaluating each of the prioritized outcomes. Consensus was informed by rigorous methodological systematic reviews. The views of patients with COPD were accounted for in all stages of the project.

Survival, treatment success, breathlessness, quality of life, activities of daily living, need for higher level of care, arterial blood gases, disease progression, future exacerbations and hospital admissions, treatment safety and adherence were all included in the core outcome set. Focused methodological research was recommended to further validate and optimize some of the selected measurement instruments. The panel did not consider the prioritized set of outcomes and associated measurement instruments burdensome for patients and health professionals to use.

Take home message: @EuroRespSoc Statement: A core outcome set and outcome measurement instruments for #ClinicalTrials evaluating #COPD exacerbations management was developed, based on evidence-informed, global, multi-stakeholder consensus.

Background

Acute exacerbations punctuate the natural history of chronic obstructive pulmonary disease (COPD) and are largely responsible for the adverse disease outcomes [1-3]. Every year, approximately a third of those diagnosed with COPD experience at least one moderate or severe exacerbation, while 9-16% experience these events even more frequently [4-7]. More importantly, every year, one in twenty unselected patients with COPD and one in four of those monitored in secondary care for their COPD experience severe exacerbations [6], which are associated with a ninety-day mortality that approximates 15% [8-10].

While novel maintenance treatments have reduced the occurrence of exacerbations [11], their management remains suboptimal and has not changed for decades [8, 12, 13]. However, over recent years, the complexity and heterogeneity of exacerbations, as well as their underlying mechanisms are increasingly being understood [3, 14-17]. In addition, the clinical validation of promising biomarkers paves the way for the introduction of precision medicine interventions, that could revolutionize the approaches to managing exacerbations [18-21]. Therefore, it is anticipated that an increased number of clinical trials will be conducted in the coming years, to evaluate novel treatments, including precision medicine interventions.

However, the design and conduct of clinical trials on managing COPD exacerbations are complicated by methodological and practical challenges [22]. Selection and consistent use of relevant, comparable, well-defined, and patient-important outcomes represent a critical challenge. A recent meta-epidemiological study revealed remarkable heterogeneity in the outcomes evaluated and reported in COPD exacerbation trials, as well as the definition of these outcomes and instruments used to assess them [23, 24]. This has led recent relevant systematic reviews and meta-analyses to report limited certainty in the available evidence [19, 25, 26].

To address this issue, the European Respiratory Society (ERS) formed this task force:

(i) **To develop a core outcome set for clinical trials evaluating the management of COPD exacerbations.** A core outcome set is an agreed minimum set of critically

important outcomes that should be evaluated in all future trials in a specific area of health care, aiming to improve their quality and comparability [27].

(ii) **To prioritize a single instrument for measuring each of the core outcomes.** The core outcome measurement instruments describe how each of the core outcomes should be evaluated in clinical trials [28].

The outputs of this project were based on global, multi-stakeholder consensus.

Methods

Detailed methodology of the COS-AECOPD (Core outcome set for the management of acute exacerbations of chronic obstructive pulmonary disease) ERS Task Force was prospectively registered with the COMET database (www.comet-initiative.com; ID: 1325), published [29] and is available in the online appendix 3. This study was conducted and reported following the methodology recommended by the COMET initiative (the COMET handbook) [27], the Core Outcome Set STAndards for Development (COS-STAD) [30] and STAndards for Reporting (COS-STAR) [31]. This project consisted of three components.

First, we developed a comprehensive list of all outcomes related to COPD exacerbations. Through a methodological systematic review, we identified outcomes that were evaluated in 123 randomized controlled trials and 38 systematic reviews on the management of COPD exacerbations [23, 24]. This list was enriched with additional outcomes considered important by patients, that have not been evaluated in trials so far. These were identified through a focused systematic review of qualitative studies [32-35], complemented by a focus group and individual interviews with a total of 86 patients from 11 countries globally. After removing duplicate entries, the list included 47 unique outcomes. This list was further enriched by the respondents of the subsequent Delphi survey.

Next, prioritization of the most critical outcomes for inclusion in the core outcome set was facilitated by a Delphi survey and a consensus panel. An online, two-stage, global,

multistakeholder Delphi survey was employed, that was developed in plain language and was available in 10 languages, to facilitate global participation [36]. Three stakeholder groups were invited to participate in the survey: (a) Patients diagnosed with COPD, who had experienced exacerbations, and personal caregivers or representatives of such patients (e.g., patient organisations); (b) Health professionals caring for patients (e.g., doctors, nurses or physiotherapists); and (c) Clinical researchers (health professionals who care for patients but are also involved in designing research studies). After the second round of the survey, consensus was assessed based on prospectively selected thresholds for inclusion or exclusion, considering responses of the three stakeholder groups separately and using data from respondents who completed both survey rounds.

Prioritization was finalized during the first consensus meeting (April 21st, 2021). Outcomes with an inconclusive survey result, that were prioritized for inclusion in the core outcome set by at least one, but not all stakeholder groups were discussed in detail. Participants were classified in two groups (a) health professionals or researchers and (b) patients diagnosed with COPD and their representatives. Thorough discussion where both groups were invited to share their views about the importance of each of these outcomes was followed by polls. Only outcomes that were rated as critical by at least 70% of the participants in both groups were added to the core outcome set.

The final component of this project consisted of the selection of a single, optimal instrument for measuring every core outcome, to ensure consistency and comparability across trials. Evidence-informed consensus was achieved during a second panel meeting (April 28th, 2021), where a pragmatic methodology was followed for prioritizing measurement instruments. Instruments that are already in use were identified through our methodological systematic review [23]. Since our aim was to promote consistency, for outcomes that are often evaluated by the same instrument, that instrument was considered for prioritization during the consensus meeting, upon evaluating its strengths and methodological limitations. For other outcomes, including all patient reported outcomes, we conducted focused literature searches of Medline/PubMed and the COSMIN database, to identify studies evaluating the quality and measurement properties of the different

instruments. The panel reviewed available evidence, which was circulated in advance of the consensus meeting via email and developed consensus on a simple instrument for each outcome considering (a) the frequency that each instrument is used in clinical trials; (b) the time and resources required to use each instrument; and (c) available data on their measurement properties, as described by COSMIN recommendations [37]. After discussion, a single instrument was selected for every core outcome and participants were asked to vote for (a) a strong recommendation, (b) an interim recommendation along with research agenda, a research agenda without a recommendation, or (c) for an alternative recommendation or the need for additional data to make an informed decision. Due to the more technical nature of this assignment, only two patients with COPD and a representative of the European Lung Foundation (ELF), with previous experience in COPD research, joined the consensus meeting, and therefore, the voting was not stratified by stakeholder group. Prespecified voting thresholds are described in the online appendix.

Feedback was sought by all participants of the consensus meetings to explore whether they felt they were offered the opportunity to share their views and that they were able to cast well-informed votes.

Changes from the prospectively registered protocol are summarized and justified in online appendix 3.

Results

The core outcome set development process is summarized in figure 1.

i. Delphi survey

The first round of the Delphi survey was available online between May 2nd and June 27th, 2020, and the second round between July 21st and October 30th. Of 1,201 individuals who started a

registration at the Delphi survey website, 1,063 (88.5%) from 88 countries in Africa, Americas, Asia, Europe and Oceania (figure 2) completed the first round of the survey and comprised our study population. These included 256 (24.1%) patients or patient representatives, 488 (45.9%) health professionals and 319 (30.0%) researchers. Baseline characteristics of the participants are described in tables 1-3. Six unique, additional outcomes were proposed by the respondents during the first round of the Delphi survey and were introduced in the second round (table 4).

Among all participants, 896 (84.3%) also completed the second survey round. Visual inspection of the distribution of first-round participant average outcome rating did not reveal differences between those who did or did not complete the second round of the survey. After the second round of the survey, 15 and 29 outcomes met the thresholds for inclusion in and exclusion from the core outcome set, respectively, while the ratings of 9 outcomes were inconclusive. These nine outcomes were further considered during the first consensus meeting. The results of the Delphi survey are presented in detail in online appendix 4 and summarized in table 4. Only a minority of the participants (3.1%) reported relevant conflicts of interest and the exclusion of their responses did not alter the survey results.

Visualisation of the responses of participants from (a) low or lower-middle (LMICs), (b) upper-middle, and (c) high income countries did not reveal any difference in the ratings among these groups. Moreover, for every outcome, the average (median) ratings of each of these groups were very similar (maximum difference = 1).

ii. Consensus meetings

The first consensus meeting was attended by a global panel including 17 patients or patient representatives, 22 health professionals and/or clinical researchers with relevant expertise, and two methodologists with expertise in core outcomes development (online appendix 8). The methodologists did not vote in the polls but provided methodological input during the discussion.

Nine outcomes with inconclusive ratings in the Delphi survey were discussed in the consensus meeting and three of them were prioritized for inclusion in the core outcome set (table 4).

The second meeting was attended by a global panel involving two patients and a patient representative (ELF), 21 health professionals and/or clinical researchers with relevant expertise, and one methodologist with expertise in core outcomes development (online appendix 8). The structure of the core outcome set was finalized (Box 1). Permanent deterioration in lung function was originally prioritized as a core outcome in the Delphi survey. However, during the consensus process, it became clear that this is a way of measuring the outcome disease progression and was, therefore, reclassified. For each of the core outcomes, a single, optimal measurement instrument was prioritized and recommended (table 5). Strong recommendations were issued for only four of the core outcomes, while for the remaining outcomes an interim instrument was recommended, along with a call for relevant methodological research (table 6).

Feedback was collected from all consensus meeting participants. All participants felt that their views were heard, and the consensus was well-informed.

Box 1. *Core Outcome Set for Clinical Trials Evaluating the Management of COPD Exacerbations. Detailed descriptions of the outcomes are available in section iii and online appendix 6.*

1. **Death**
    a. Death from any cause
    b. Death from a COPD exacerbation
2. **Treatment success**
3. **Need for higher level of care**
    a. Need for hospital admission for the presenting exacerbation
    b. Need for admission to the intensive care unit for the exacerbation
4. **Levels of oxygen and carbon dioxide in the blood (arterial blood gases)**
5. **Patient reported outcomes**
    a. Breathlessness
    b. Health related quality of life
    c. Activities of daily living
    d. Worsening of symptoms after the initial treatment
6. **Future Impact**
    a. Disease progression
    b. Future exacerbations
    c. Future hospital admissions
7. **Safety**
    a. Serious adverse events from treatments
    b. Development of resistant bacteria
    c. Development of pneumonia
8. **Treatment adherence**

iii. Considerations around the selection of outcome measurement instruments

The recommended outcome measurement instruments and relevant research recommendations are summarized in table 5 and appendices 5-7. Next, we summarize pertinent additional data and discussion points considered by the panel about some of the measurement instruments. A more detailed version of this section, focusing on all instruments is available in the online appendix 6.

a. Death from a COPD exacerbation

Death from COPD exacerbation is rarely evaluated in exacerbation trials. COPD exacerbations are often complicated by events such as ventricular arrhythmia, massive pulmonary embolism, acute myocardial infarction, or pneumonia [41]. As a result, the determination of the cause of death during an exacerbation is complex and often inconsistent across different centres and countries. For this reason the panel opted for a pragmatic approach based on the documented primary cause registered in the death certificate. If this is COPD exacerbation or an event considered to be an immediate complication of the exacerbation, then the death should be attributed to the exacerbation. It was recognized that ideally the cause of death should be confirmed by a well-informed and blinded adjudication committee; however, this may not always be feasible.

b. Treatment success

Treatment success is more frequently defined as cure of the exacerbation and more specifically as the “Complete resolution of all signs and symptoms of the exacerbation” [23]. However, the recovery period of an exacerbation varies significantly and may be very prolonged. Large observational studies have shown wide variability in the duration of exacerbation recovery, revealing that 25% of patients still experience symptoms associated with the exacerbation 25 or even 35 days after the onset of the exacerbation [42, 43]. Longer periods may be required until patients recover their previous exercise capacity or ADL levels [44, 45]. Moreover, exacerbations accelerate disease progression; therefore, the clinical condition after recovery from an exacerbation may be characterized by a greater symptomatic burden, compared to the previous baseline [46]. As a result, this definition of cure was considered problematic. For this reason, a more pragmatic, yet still suboptimal interim instrument was recommended by the panel: Treatment success is defined as sufficient improvement of the signs and symptoms of the exacerbation, such that no additional systemic treatments (antibiotics or systemic corticosteroids) are prescribed. While still subjective, the decision of the clinician to prescribe additional systemic

treatments better reflects daily clinical practice and it is often used in trials to determine treatment success or failure.

c. Need for higher level of care

This broad category encompasses (i) the need for hospital admission, and (ii) the need for admission to the intensive care unit (ICU), for the presenting exacerbation. These outcomes are frequently evaluated in clinical trials. However, the indications for hospital or ICU admission vary across different centres and countries.

Hospital at home and telemonitoring options introduce heterogeneity in the criteria for hospital admission and length of stay [47]. This outcome is also impacted by non-clinical factors, such as social reasons for admission, discharge planning delays [48], the availability of hospital beds, or travel distance. As a result, the panel recommends that the need for hospital admission should be defined pragmatically as a clinical need to admit a patient to the hospital, or to offer equivalent intensification of the monitoring or care, that may be provided in other settings (including the patients' home). Admissions for non-medical (e.g., social) reasons should be reported separately. Appreciating the heterogeneity in the design of health services, the panel recommends that trialists should prospectively and transparently define in detail the reasons for a need for hospital admission in the context of each trial.

Indications for ICU admission also vary. Characteristically, while in most centres non-invasive ventilation is now delivered in a respiratory ward or a high dependency unit, in some centres it is still delivered in the ICU [49]. Availability of ICU beds may also impact the decision to admit, and the duration of ICU stay. On the other hand, patients with COPD with poor functional status and underlying multi-morbidity are often not offered an ICU admission or invasive mechanical ventilation, due to futility [50]. The criteria used to support such decisions vary across centres and countries, according to local policies and availability of resources. Acknowledging that the main, consistent indication for ICU admission in this group of patients is the need for invasive

mechanical ventilation, the panel recommended that trials should record the need for invasive mechanical ventilation, defined as: (i) persistent or deteriorating respiratory acidosis, despite optimized medical treatment and delivery of non-invasive ventilation (NIV), (ii) persistent or deteriorating respiratory acidosis despite optimized medical treatment and a contra-indication for the use of NIV, for example due to upper airway obstruction, facial burns or severe facial deformities, where fitting a mask is impossible, or (iii) respiratory arrest or peri-arrest situations, unless there is a rapid recovery from manual ventilation or provision of NIV [50]. The decision to focus on the need for invasive mechanical ventilation rather than the receipt of ventilation was based on the earlier observation that often, while these criteria are fulfilled, patients are not offered invasive ventilation, due to futility.

d. Levels of oxygen and carbon dioxide in the blood (arterial blood gases).

This was considered a setting and intervention specific outcome. Firstly, it may not be feasible to be assessed in studies recruiting in an outpatient clinic. The panel agreed that the value of measuring blood levels of oxygen and carbon dioxide in this setting may be limited.

e. Breathlessness

The most frequently used validated scales for measuring breathlessness in trials focusing on the management of COPD exacerbations were the Borg scales (original or modified version) and the modified Medical Research Council (mMRC) scale, while this symptom was also frequently quantified as part of the COPD assessment test (CAT) [23]. More specifically, mMRC does not directly assess breathlessness, as it is a measure of activity limitation due to breathlessness. Use of mMRC during an exacerbation was considered by the panel less sensitive, since most patients with moderate or severe exacerbations would cluster in Grade 4 (“Too breathless to leave the house or breathless when dressing or undressing”), thus limiting the discriminant validity of the scale in this context. CAT is a multidimensional health status tool measuring several symptoms

and therefore does not provide a focus on breathlessness [51]. The modified Borg scale is easy to complete, and broadly used in clinical practice and research. Clinically validated translations are available in many languages. Its measurement properties have been thoroughly and positively assessed [52] (online appendix 6.7). As a result, the modified Borg scale is recommended for evaluating breathlessness. It should be measured at approximately the same time every day. It can be self-completed.

f. Health-related quality of life.

CAT is the most frequently used validated tool for assessing health related quality of life in trials on the management of exacerbations, followed by the Saint George's Respiratory Questionnaire (SGRQ) and the Chronic COPD Questionnaire (CCQ) [23]. A systematic review using the COSMIN methodology for evaluating the measurement properties of 23 instruments used to assess quality of life in COPD recommended the use of CAT, Chronic Respiratory Questionnaire (CRQ), the Saint George's Respiratory Questionnaire (SGRQ) or the Living with Chronic Obstructive Pulmonary Disease (LCOPD) Questionnaire [53]. While these tools have similar measurement properties (online appendix 6.8), CAT can be completed within 1-3 minutes while the other tools are more complex and time consuming. Given that CAT is already the most frequently used tool for evaluating health-related quality of life, it was recommended by the panel. A comparison with a baseline estimate of the health-related quality of life prior to the exacerbation would be preferable, but in larger randomized studies, balance in the baseline characteristics of participants in the study groups will usually suffice.

g. Activities of daily living (ADL)

This outcome is rarely evaluated in exacerbation trials [23]. ADL are classified as basic and instrumental [54]. Basic ADL are simple activities that are essential for independent life, such as self-care (showering, dressing, or grooming) and basic mobility, while instrumental ADL

encapsulate more complex activities, requiring higher functioning, such as preparing meals, home maintenance, shopping, handling finances, and travelling alone [55]. Instrumental ADL are less relevant during an exacerbation, especially during severe exacerbations, while patients are admitted in the hospital and may not be able to undertake such complex activities; however, they are pertinent to quantify the overall impact of an exacerbation on a patient's ADL. For this reason, the panel decided to recommend a tool focusing on basic ADL, to be evaluated during the exacerbation and a second tool, assessing both basic and instrumental ADL for longer-term follow-up.

The psychometric properties of instruments used to quantify ADL in patients with COPD have been evaluated in two methodological systematic reviews [55, 56]. Five of the identified instruments focused on basic ADL, of which the Katz Activities of Daily Living Scale, the Barthel index and the motor subscale of the functional independence measure (FIM) were not disease specific and included domains that are less relevant to COPD patients (e.g., control of bladder and bowels). While the Glittre index is disease specific, it focuses on exercise capacity and includes a simple exercise component, which many patients may find challenging to complete during an exacerbation. Finally, the Capacity of Daily Living during the Morning (CDLM) Questionnaire [57] is a simple, disease specific questionnaire, with measurement properties adequately evaluated with favourable findings (online appendix 6.9). For this reason, the CDLM tool was recommended for quantifying basic ADL during an exacerbation.

The identified methodological reviews revealed eight disease-specific tools assessing a combination of instrumental and basic ADL [55, 56]. Responsiveness to change in a patient's clinical condition, a crucial characteristic required for evaluating the impact of exacerbation on ADL, has only been confirmed for three of these tools: the Manchester Respiratory Activities of Daily Living Questionnaire (MRADL) [58], the COPD Activity Rating Scale (CARS) [59], and the 11-items Pulmonary Functional Status Scale (PFSS-11) [60]. While all three tools were considered valid options, the performance characteristics of the MRADL questionnaire were more thoroughly validated compared to CARS, while it was also considered simpler to complete,

compared to the PFSS-11 tool (online appendix 6.9). For promoting consistency, the panel recommends that the MRADL questionnaire be used to evaluate both basic and instrumental ADL at recovery from COPD exacerbations. A comparison with a baseline estimate of the ADL prior to the exacerbation would be beneficial and could potentially be captured retrospectively during recruitment. Recall bias is anticipated to be limited, since in most cases, the duration of the acute event at recruitment would rarely exceed a week and the questions refer to some of the most critical activities of daily living.

h. Disease progression

This outcome was suggested by patients during the qualitative research studies that preceded the Delphi survey. Acute exacerbations are known to accelerate disease progression in patients with COPD [46, 61, 62]. Several parameters have been used as potential measures of disease progression, including symptom burden, health status, exercise capacity, blood biomarkers, pulmonary function decline, or radiologic progression revealed in computed tomography (CT) of the chest [62-66].

There was agreement within the panel that evaluation of disease progression as an outcome in exacerbation trials is only meaningful as change from baseline; therefore, a baseline measurement is required. To achieve that, participants would have to be recruited while the disease is stable, in anticipation of developing an exacerbation. However, such a study design requires significantly more resources and prolonged follow-up periods or a patient database with recent measurement taken during periods of clinical stability.

Not surprisingly, disease progression is only rarely evaluated as an outcome in exacerbation trials using objective tests [23]. Change from baseline in pulmonary function was only assessed in two of the trials included in the methodological systematic review, while imaging was not used in any of the studies as an estimate of disease progression. Symptoms and quality of life are evaluated frequently, but not as change from baseline (see respective outcomes).

Change in FEV_1 over time is the most established instrument for evaluating COPD progression in clinical trials and observational studies evaluating the management of disease longitudinally and for this reason, the panel recommends that it should also be used for evaluating the impact of exacerbations on disease progression. Acknowledging the limitations of this study design, the panel recommends that this outcome only be considered core for long-term studies where baseline values can be captured. Acknowledging the limitations of this study design, the panel recommends that this outcome only be considered core for long-term studies where baseline values can be captured.

i. Development of resistant bacteria

Antimicrobial resistance is often explored as part of a composite microbiological response outcome in trials involving antibiotics as interventions [23]. Bacterial growth and resistance are usually evaluated in spontaneous sputum, while in the absence of sputum bacterial eradication is presumed and is not further assessed. The panel adopts this approach and recommends that trials evaluating antimicrobials, antimicrobial stewardship strategies, novel immune modifiers or other interventions that may affect bacterial resistance should test bacterial resistance in spontaneous sputum. As a minimum, resistance should be explored at baseline and within a week after completion of the treatment. Sputum induction may provide additional information when spontaneous sputum is not available. However, in each study, researchers should consider the balance between the added value of sputum induction, compared to the risk, a participant's discomfort and required resources.

j. Development of pneumonia

Development of pneumonia as a safety outcome is often evaluated in exacerbation trials [23]. Methodology is consistent and was adopted by this task force. Pneumonia should be confirmed by the presence of new consolidation in the chest X-ray or other imaging modalities of the chest,

in the presence of consistent clinical signs and symptoms. When possible, chest imaging should be acquired at baseline, to assess for the presence of pneumonia. However, this may not be possible for trials recruiting patients outside the hospital setting. Follow-up chest imaging should be driven by the clinical need.

Discussion

Based on a rigorous methodology, recommended by the COMET initiative, this task force developed a core outcome set for clinical trials assessing pharmacological and non-pharmacological interventions in COPD exacerbations. In addition, it recommended a single optimal measurement instrument for evaluating each core outcome and prioritized methodological research for further optimizing some of these instruments in the future. This work was informed by systematic reviews, qualitative research involving 86 patients from 11 countries globally, an extensive, multi-stakeholder two-stage Delphi survey that was completed by 1,063 participants from 88 countries and two multi-stakeholder consensus meetings with global representation. Uptake of the core outcome set by clinical trials and other clinical research studies is a critical measure of success and for this reason we have developed a dissemination strategy that is summarized in the online appendix 9.4.

A key objective of the panel was to develop a pragmatic core outcome set, that would not require excessive resource commitment and would be feasible to be evaluated in all clinical trials, to promote implementation. While the final core outcome set includes more outcomes than some of the other sets, most of the selected outcomes are simple to assess, routinely collected, and do not require excessive resources. Moreover, when possible, the panel favoured the selection of simple and pragmatic measurement instruments, taking into consideration the time and resources required for capturing them. Recognizing that disease progression can only be evaluated in trials of a longer-term and resource intensive design, the panel recommended that this outcome should only be assessed in this subgroup of trials. However, the importance of disease progression as

an outcome should not be underestimated, and trialists are encouraged to consider appropriate study designs to capture it.

Several of the prioritized outcomes are currently only evaluated infrequently in relevant clinical trials [23]. Moreover, variability was observed in the instruments used to measure many of the core outcomes. These observations confirm that this work was indeed needed and can improve the consistency, quality and comparability of clinical trials on the management of exacerbations. While the panel was able to recommend one optimal instrument for consistently evaluating each of the core outcomes, most of these were considered interim recommendations, paired with a research agenda. Due to the variability in the instruments used in trials by now, adequate validation and information on the measurement properties of the instruments in the context of exacerbation trials to support strong recommendations was lacking. However, the recommendations of instruments were assessed based on currently available evidence, including data on the frequency that each instrument is used in exacerbation trials, but also previously conducted rigorous systematic reviews evaluating the measurement properties of all recommended patient reported outcomes [52, 53, 55, 56]. Still, trialists are encouraged to embed in their trials methodological research studies that could facilitate further optimization of the measurement instruments. Similar challenges with the selection of outcomes and measurement instruments to be used have been identified in trials assessing the management of other acute respiratory events, including pneumonia, acute bronchitis, and the coronavirus disease 2019 (COVID-19) [69-71]. Crosstalk among these fields could be beneficial.

COPD exacerbations represent an acute condition that can be successfully managed. Therefore, the timing of outcomes evaluation is a crucial parameter that should be optimized and standardized. This is especially so for the precise time when the overall treatment outcome (treatment success) is assessed. However, our methodological systematic review did not conclude on the optimal measurement timepoint due to significant clinical and methodological heterogeneity of the included studies [23]. Consequently, further data is needed to inform the optimal timepoint for evaluating treatment success and our panel was not able to produce

informed recommendations. Moreover, the duration of follow-up is trial specific, and the panel opted not to recommend a minimum duration of follow-up. However, to promote consistency and comparability, it is suggested that longer-term outcomes should be evaluated at three and six months from recruitment if the selected follow-up duration includes one or both timepoints. Moreover, it is suggested that the outcomes should be evaluated at specific timepoints, rather than at discharge or at symptom relief, since such “mobile” timepoints might introduce bias.

While this core outcome set and measurement instruments were developed for clinical trials in COPD exacerbations, it would be important to be captured also in observational studies as this could facilitate the validation and optimization of the measurement instruments recommended for each outcome. While this is the first formal core outcome set for COPD exacerbations trials, COPD exacerbations outcomes have been prioritized by two other initiatives. The eo-Drive trial group (Eosinophil-driven corticotherapy for patients hospitalized for COPD Exacerbations, NCT04234360) prioritized outcomes for their clinical trial [72] and the CICERO (The Collaboration In COPD ExaceRbatiOns ERS Clinical Research Collaboration) developed standards for clinical assessment, management and follow-up of hospitalized exacerbations [73]. While this core outcome set is broader than the outputs of the previous initiatives, as described in the online appendix 9.1, all previously prioritized outcomes are included in our core outcome set and that could further promote consistency.

A potential limitation of this work is that it did not fully follow the methodology proposed by the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) recommendations for selecting the recommended outcome instruments. COSMIN recommends de novo conduct of methodological systematic reviews to evaluate the measurement properties of all available instruments that could be used to assess an outcome and is particularly relevant for patient reported outcomes. While it was not feasible to complete these as part of an ERS Task Force, we identified relevant high-quality methodological systematic reviews, evaluating the available instruments for all patient reported outcomes that were included in the core outcome set, which were used to inform our recommendations. Despite our best effort, the Delphi survey

was somewhat limited by the lack of respondents from low-income countries. Lack of access or engagement represent a recognized problem, limiting the participation of people from low-income countries to such online surveys [76]. Given the wide geographic distribution and multi-stakeholder involvement of our sample, and the similar responses across lower-middle, upper-middle and high income countries, we do not believe that significantly limits the generalizability of our findings. The prospectively published, transparent protocol represent a major strength of this study. Unfortunately, we had to deviate from the protocol on two occasions; these deviations are described and justified in detail in the online appendices 3.6 and 9.3.

In summary, this task force developed a core outcome set for trials in acute exacerbations of COPD and recommended an optimal instrument for measuring each of the core outcomes, aiming to improve the consistency, quality and comparability of future relevant clinical trials.

Figure and Table Legends

Figure 1: Study flowchart summarizing the main steps of the COS development process.

Figure 2: Geographic distribution of the Delphi survey participants. The colour of each country represents the number of participants (see colour scheme).

Table 1: Baseline characteristics of the Delphi Survey Participants. Reported as N (% of the participants in the corresponding stakeholder group).

Table 2: Additional baseline characteristics of patients with COPD who completed the Delphi survey.

Table 3: Additional baseline characteristics of expert respondents (health professionals and researchers) of the Delphi survey.

Table 4: Summary of the selection process of the core outcomes from the longlist. Percentages refer to the proportion of participants that consider a particular outcome critical. Background colour coding: First column: Grey, blue, purple colours signify outcomes identified through the

methodological systematic reviews, qualitative interviews, or the Delphi survey, respectively. In the remaining columns, background colour refers to the results of the outcome selection process at each stage; green, yellow and red colours signify inclusion, inconclusive result or exclusion of the respective outcome.

Table 5: Outcome measurement instrument recommendations. Green and yellow background colours signify a strong recommendation and an interim recommendation with associated research agenda, respectively

References

1. Wedzicha JA, Seemungal TA: **COPD exacerbations: defining their cause and prevention.** *Lancet* 2007, **370:**786-796.
2. Singh D, Agusti A, Anzueto A, Barnes PJ, Bourbeau J, Celli BR, Criner GJ, Frith P, Halpin DMG, Han M, et al: **Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Lung Disease: the GOLD science committee report 2019.** *The European respiratory journal* 2019, **53:**1900164.
3. Mathioudakis AG, Janssens W, Sivapalan P, Singanayagam A, Dransfield MT, Jensen JUS, Vestbo J: **Acute exacerbations of chronic obstructive pulmonary disease: In search of diagnostic biomarkers and treatable traits.** *Thorax* 2020, **In press**.
4. Hastie AT, Martinez FJ, Curtis JL, Doerschuk CM, Hansel NN, Christenson S, Putcha N, Ortega VE, Li X, Barr RG, et al: **Association of sputum and blood eosinophil concentrations with clinical measures of COPD severity: an analysis of the SPIROMICS cohort.** *Lancet Respir Med* 2017, **5:**956-967.
5. Gayle A, Dickinson S, Morris K, Poole C, Mathioudakis AG, Vestbo J: **What is the impact of GOLD 2017 recommendations in primary care? - a descriptive study of patient classifications, treatment burden and costs.** *Int J Chron Obstruct Pulmon Dis* 2018, **13:**3485-3492.
6. Rothnie KJ, Mullerova H, Smeeth L, Quint JK: **Natural History of Chronic Obstructive Pulmonary Disease Exacerbations in a General Practice-based Population with Chronic Obstructive Pulmonary Disease.** *Am J Respir Crit Care Med* 2018, **198:**464-471.
7. Kardos P, Vogelmeier C, Worth H, Buhl R, Lossi NS, Mailander C, Criee CP: **A two-year evaluation of the 'real life' impact of COPD on patients in Germany: The DACCORD observational study.** *Respir Med* 2017, **124:**57-64.
8. Wedzicha JAEC-C, Miravitlles M, Hurst JR, Calverley PM, Albert RK, Anzueto A, Criner GJ, Papi A, Rabe KF, Rigau D, et al: **Management of COPD exacerbations: a European Respiratory Society/American Thoracic Society guideline.** *Eur Respir J* 2017, **49**.
9. Han MK, Quibrera PM, Carretta EE, Barr RG, Bleecker ER, Bowler RP, Cooper CB, Comellas A, Couper DJ, Curtis JL, et al: **Frequency of exacerbations in patients with chronic obstructive pulmonary disease: an analysis of the SPIROMICS cohort.** *Lancet Respir Med* 2017, **5:**619-626.
10. Mullerova H, Shukla A, Hawkins A, Quint J: **Risk factors for acute exacerbations of COPD in a primary care population: a retrospective observational cohort study.** *BMJ Open* 2014, **4:**e006171.

11. Wedzicha JA, Calverley PMA, Albert RK, Anzueto A, Criner GJ, Hurst JR, Miravitlles M, Papi A, Rabe KF, Rigau D, et al: **Prevention of COPD exacerbations: a European Respiratory Society/American Thoracic Society guideline.** *Eur Respir J* 2017, **50**.
12. Dobler CC, Morrow AS, Beuschel B, Farah MH, Majzoub AM, Wilson ME, Hasan B, Seisa MO, Daraz L, Prokop LJ, et al: **Pharmacologic Therapies in Patients With Exacerbation of Chronic Obstructive Pulmonary Disease: A Systematic Review With Meta-analysis.** *Ann Intern Med* 2020, **172:**413-422.
13. Dobler CC, Morrow AS, Farah MH, Beuschel B, Majzoub AM, Wilson ME, Hasan B, Seisa MO, Daraz L, Prokop LJ, et al: **Nonpharmacologic Therapies in Patients With Exacerbation of Chronic Obstructive Pulmonary Disease: A Systematic Review With Meta-Analysis.** *Mayo Clin Proc* 2020, **95:**1169-1183.
14. McDonald VM, Osadnik CR, Gibson PG: **Treatable traits in acute exacerbations of chronic airway diseases.** *Chron Respir Dis* 2019, **16:**1479973119867954.
15. Kefala AM, Fortescue R, Alimani GS, Kanavidis P, McDonnell MJ, Magiorkinis E, Megremis S, Paraskevis D, Voyiatzaki C, Mathioudakis GA, et al: **Prevalence and clinical implications of respiratory viruses in stable chronic obstructive pulmonary disease (COPD) and exacerbations: a systematic review and meta-analysis protocol.** *BMJ Open* 2020, **10:**e035640.
16. Papi A, Bellettato CM, Braccioni F, Romagnoli M, Casolari P, Caramori G, Fabbri LM, Johnston SL: **Infections and airway inflammation in chronic obstructive pulmonary disease severe exacerbations.** *Am J Respir Crit Care Med* 2006, **173:**1114-1121.
17. Bafadhel M, McKenna S, Terry S, Mistry V, Reid C, Haldar P, McCormick M, Haldar K, Kebadze T, Duvoix A, et al: **Acute exacerbations of chronic obstructive pulmonary disease: identification of biologic clusters and their biomarkers.** *Am J Respir Crit Care Med* 2011, **184:**662-671.
18. Sivapalan P, Lapperre TS, Janner J, Laub RR, Moberg M, Bech CS, Eklof J, Holm FS, Armbruster K, Sivapalan P, et al: **Eosinophil-guided corticosteroid therapy in patients admitted to hospital with COPD exacerbation (CORTICO-COP): a multicentre, randomised, controlled, open-label, non-inferiority trial.** *Lancet Respir Med* 2019, **7:**699-709.
19. Mathioudakis AG, Chatzimavridou-Grigoriadou V, Corlateanu A, Vestbo J: **Procalcitonin to guide antibiotic administration in COPD exacerbations: a meta-analysis.** *Eur Respir Rev* 2017, **26:**160073.
20. Butler CC, Gillespie D, White P, Bates J, Lowe R, Thomas-Jones E, Wootton M, Hood K, Phillips R, Melbye H, et al: **C-Reactive Protein Testing to Guide Antibiotic Prescribing for COPD Exacerbations.** *N Engl J Med* 2019, **381:**111-120.
21. Agusti A, Faner R, Celli B, Rodriguez-Roisin R: **Precision medicine in COPD exacerbations.** *Lancet Respir Med* 2018, **6:**657-659.
22. Mathioudakis AG, Sivapalan P, Papi A, Vestbo J, on behalf of the DECODE-NET Investigators: **The DisEntangling Chronic Obstructive pulmonary Disease Exacerbations clinical trials NETwork: DECODE-NET. Rationale and vision.** *Eur Respir J* 2020, **In press**.
23. Mathioudakis AG, Moberg M, Janner J, Alonso-Coello P, Vestbo J: **Outcomes reported on the management of COPD exacerbations: a systematic survey of randomised controlled trials.** *ERJ Open Res* 2019, **5:**00072-02019.
24. Mathioudakis AG, Janner J, Moberg M, Alonso-Coello P, Vestbo J: **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:**00136-02019.
25. Vollenweider DJ, Frei A, Steurer-Stey CA, Garcia-Aymerich J, Puhan MA: **Antibiotics for exacerbations of chronic obstructive pulmonary disease.** *Cochrane Database Syst Rev* 2018, **10:**CD010257.
26. Walters JA, Tan DJ, White CJ, Wood-Baker R: **Different durations of corticosteroid therapy for exacerbations of chronic obstructive pulmonary disease.** *Emergencias* 2020, **32:**363-364.
27. Williamson PR, Altman DG, Bagley H, Barnes KL, Blazeby JM, Brookes ST, Clarke M, Gargon E, Gorst S, Harman N, et al: **The COMET Handbook: version 1.0.** *Trials* 2017, **18:**280.

28. t Hoen LA, Groen J, Scheepe JR, Reuvers S, Diaz DC, Fernandez BP, Del Popolo G, Musco S, Hamid R, Ecclestone H, et al: **A Quality Assessment of Patient-Reported Outcome Measures for Sexual Function in Neurologic Patients Using the Consensus-based Standards for the Selection of Health Measurement Instruments Checklist: A Systematic Review.** *Eur Urol Focus* 2017, **3:**444-456.
29. Mathioudakis AG, Abroug F, Agusti A, Bakke P, Bartziokas K, Beghe B, Bikov A, Bradbury T, Brusselle G: **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, **In press**.
30. Kirkham JJ, Davis K, Altman DG, Blazeby JM, Clarke M, Tunis S, Williamson PR: **Core Outcome Set-STAndards for Development: The COS-STAD recommendations.** *PLoS Med* 2017, **14:**e1002447.
31. Kirkham JJ, Gorst S, Altman DG, Blazeby JM, Clarke M, Devane D, Gargon E, Moher D, Schmitt J, Tugwell P, et al: **Core Outcome Set-STAndards for Reporting: The COS-STAR Statement.** *PLoS Med* 2016, **13:**e1002148.
32. Zhang Y, Morgan RL, Alonso-Coello P, Wiercioch W, Bala MM, Jaeschke RR, Styczen K, Pardo-Hernandez H, Selva A, Ara Begum H, et al: **A systematic review of how patients value COPD outcomes.** *Eur Respir J* 2018, **52**.
33. Kouijzer M, Brusse-Keizer M, Bode C: **COPD-related fatigue: Impact on daily life and treatment opportunities from the patient's perspective.** *Respir Med* 2018, **141:**47-51.
34. Rosa F, Bagnasco A, Ghirotto L, Rocco G, Catania G, Aleo G, Zanini M, Dasso N, Hayter M, Sasso L: **Experiences of older people following an acute exacerbation of chronic obstructive pulmonary disease: A phenomenological study.** *J Clin Nurs* 2018, **27:**e1110-e1119.
35. Kessler R, Stahl E, Vogelmeier C, Haughney J, Trudeau E, Lofdahl CG, Partridge MR: **Patient understanding, detection, and experience of COPD exacerbations: an observational, interview-based study.** *Chest* 2006, **130:**133-142.
36. **DelphiManager. 2018.** [https://www.comet-initiative.org/delphimanager/index.html]
37. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, de Vet HC: **The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study.** *Qual Life Res* 2010, **19:**539-549.
38. Dodd S, Clarke M, Becker L, Mavergames C, Fish R, Williamson PR: **A taxonomy has been developed for outcomes in medical research to help improve knowledge discovery.** *J Clin Epidemiol* 2018, **96:**84-92.
39. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schunemann HJ, Group GW: **GRADE: an emerging consensus on rating quality of evidence and strength of recommendations.** *BMJ* 2008, **336:**924-926.
40. Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G, Alderson P, Glasziou P, Falck-Ytter Y, Schunemann HJ: **GRADE guidelines: 2. Framing the question and deciding on important outcomes.** *J Clin Epidemiol* 2011, **64:**395-400.
41. Zvezdin B, Milutinov S, Kojicic M, Hadnadjev M, Hromis S, Markovic M, Gajic O: **A postmortem analysis of major causes of early death in patients hospitalized with COPD exacerbation.** *Chest* 2009, **136:**376-380.
42. Seemungal TA, Donaldson GC, Bhowmik A, Jeffries DJ, Wedzicha JA: **Time course and recovery of exacerbations in patients with chronic obstructive pulmonary disease.** *Am J Respir Crit Care Med* 2000, **161:**1608-1613.
43. Aaron SD, Donaldson GC, Whitmore GA, Hurst JR, Ramsay T, Wedzicha JA: **Time course and pattern of COPD exacerbation onset.** *Thorax* 2012, **67:**238-243.
44. McAuley HJC, Harvey-Dunstan TC, Craner M, Richardson M, Singh SJ, Steiner MC, Greening NJ: **Longitudinal changes to quadriceps thickness demonstrate acute sarcopenia following**

admission to hospital for an exacerbation of chronic respiratory disease. *Thorax* 2021, **76:**726-728.
45. Spruit MA, Gosselink R, Troosters T, Kasran A, Gayan-Ramirez G, Bogaerts P, Bouillon R, Decramer M: **Muscle force during an acute exacerbation in hospitalised patients with COPD and its relationship with CXCL8 and IGF-I.** *Thorax* 2003, **58:**752-756.
46. Anzueto A: **Impact of exacerbations on COPD.** *Eur Respir Rev* 2010, **19:**113-118.
47. Jeppesen E, Brurberg KG, Vist GE, Wedzicha JA, Wright JJ, Greenstone M, Walters JA: **Hospital at home for acute exacerbations of chronic obstructive pulmonary disease.** *Cochrane Database Syst Rev* 2012**:**CD003573.
48. Bai AD, Dai C, Srivastava S, Smith CA, Gill SS: **Risk factors, costs and complications of delayed hospital discharge from internal medicine wards at a Canadian academic medical centre: retrospective cohort study.** *BMC Health Serv Res* 2019, **19:**935.
49. Trethewey SP, Edgar RG, Turner AM, Mukherjee R: **Ward-Based Non-Invasive Ventilation in Acute Exacerbations of COPD: A Narrative Review of Current Practice and Outcomes in the UK.** *Healthcare (Basel)* 2018, **6**.
50. Davidson AC, Banham S, Elliott M, Kennedy D, Gelder C, Glossop A, Church AC, Creagh-Brown B, Dodd JW, Felton T, et al: **BTS/ICS guideline for the ventilatory management of acute hypercapnic respiratory failure in adults.** *Thorax* 2016, **71 Suppl 2:**ii1-35.
51. Jones PW, Harding G, Berry P, Wiklund I, Chen WH, Kline Leidy N: **Development and first validation of the COPD Assessment Test.** *Eur Respir J* 2009, **34:**648-654.
52. Dorman S, Byrne A, Edwards A: **Which measurement scales should we use to measure breathlessness in palliative care? A systematic review.** *Palliat Med* 2007, **21:**177-191.
53. Weldam SW, Schuurmans MJ, Liu R, Lammers JW: **Evaluation of Quality of Life instruments for use in COPD care and research: a systematic review.** *Int J Nurs Stud* 2013, **50:**688-707.
54. **International Classification of Functioning, Disability and Health (ICF, updated 2018)** [https://www.who.int/standards/classifications/international-classification-of-functioning-disability-and-health]
55. Janaudis-Ferreira T, Beauchamp MK, Robles PG, Goldstein RS, Brooks D: **Measurement of activities of daily living in patients with COPD: a systematic review.** *Chest* 2014, **145:**253-271.
56. Liu Y, Li H, Ding N, Wang N, Wen D: **Functional Status Assessment of Patients With COPD: A Systematic Review of Performance-Based Measures and Patient-Reported Measures.** *Medicine (Baltimore)* 2016, **95:**e3672.
57. Partridge MR, Miravitlles M, Stahl E, Karlsson N, Svensson K, Welte T: **Development and validation of the Capacity of Daily Living during the Morning questionnaire and the Global Chest Symptoms Questionnaire in COPD.** *Eur Respir J* 2010, **36:**96-104.
58. Yohannes AM, Roomi J, Winn S, Connolly MJ: **The Manchester Respiratory Activities of Daily Living questionnaire: development, reliability, validity, and responsiveness to pulmonary rehabilitation.** *J Am Geriatr Soc* 2000, **48:**1496-1500.
59. Morimoto M, Takai K, Nakajima K, Kagawa K: **Development of the chronic obstructive pulmonary disease activity rating scale: reliability, validity and factorial structure.** *Nurs Health Sci* 2003, **5:**23-30.
60. Chen YJ, Narsavage GL, Culp SL, Weaver TE: **The development and psychometric analysis of the short-form Pulmonary Functional Status Scale (PFSS-11).** *Res Nurs Health* 2010, **33:**477-485.
61. Anzueto A, Leimer I, Kesten S: **Impact of frequency of COPD exacerbations on pulmonary function, health status and clinical outcomes.** *Int J Chron Obstruct Pulmon Dis* 2009, **4:**245-251.
62. Kerkhof M, Voorham J, Dorinsky P, Cabrera C, Darken P, Kocks JW, Sadatsafavi M, Sin DD, Carter V, Price DB: **Association between COPD exacerbations and lung function decline during maintenance therapy.** *Thorax* 2020, **75:**744-753.
63. Decramer M, Gosselink R, Rutten-Van Molken M, Buffels J, Van Schayck O, Gevenois PA, Pellegrino R, Derom E, De Backer W: **Assessment of progression of COPD: report of a workshop held in Leuven, 11-12 March 2004.** *Thorax* 2005, **60:**335-342.

64. Jones PW, Agusti AG: **Outcomes and markers in the assessment of chronic obstructive pulmonary disease.** *Eur Respir J* 2006, **27:**822-832.
65. Shaw JG, Vaughan A, Dent AG, O'Hare PE, Goh F, Bowman RV, Fong KM, Yang IA: **Biomarkers of progression of chronic obstructive pulmonary disease (COPD).** *J Thorac Dis* 2014, **6:**1532-1547.
66. Ash SY, San Jose Estepar R, Fain SB, Tal-Singer R, Stockley RA, Nordenmark LH, Rennard S, Han MK, Merrill D, Humphries SM, et al: **Relationship between Emphysema Progression at CT and Mortality in Ever-Smokers: Results from the COPDGene and ECLIPSE Cohorts.** *Radiology* 2021, **299:**222-231.
67. **International Conference on Harmonisation of Technical Requirments for Registration of Pharmaceuticals for Human Use. ICH Harmonised Tripartite Guideline. Clinical Safety Data Management: Definitions and Standards for Expedited Reporting** [https://database.ich.org/sites/default/files/E2A_Guideline.pdf]
68. Elvers KT, Wilson VJ, Hammond A, Duncan L, Huntley AL, Hay AD, van der Werf ET: **Antibiotic-induced changes in the human gut microbiota for the most commonly prescribed antibiotics in primary care in the UK: a systematic review.** *BMJ Open* 2020, **10:**e035677.
69. Weiss E, Zahar JR, Alder J, Asehnoune K, Bassetti M, Bonten MJM, Chastre J, De Waele J, Dimopoulos G, Eggimann P, et al: **Elaboration of Consensus Clinical Endpoints to Evaluate Antimicrobial Treatment Efficacy in Future Hospital-acquired/Ventilator-associated Bacterial Pneumonia Clinical Trials.** *Clin Infect Dis* 2019, **69:**1912-1918.
70. Mathioudakis AG, Fally M, Hashad R, Kouta A, Hadi AS, Knight SB, Bakerly ND, Singh D, Williamson PR, Felton T, Vestbo J: **Outcomes Evaluated in Controlled Clinical Trials on the Management of COVID-19: A Methodological Systematic Review.** *Life (Basel)* 2020, **10**.
71. Mathioudakis AG, Fally M, Hashad R, Knight S, Felton T, Vestbo J: **COVID-19 Clinical Trials: Unraveling a Methodological Gordian Knot.** *Am J Respir Crit Care Med* 2020, **202:**635-637.
72. Suehs CM, Zysman M, Chenivesse C, Burgel PR, Couturaud F, Deslee G, Berger P, Raherison C, Devouassoux G, Brousse C, et al: **Prioritising outcomes for evaluating eosinophil-guided corticosteroid therapy among patients with acute COPD exacerbations requiring hospitalisation: a Delphi consensus study.** *BMJ Open* 2020, **10:**e035811.
73. Ramakrishnan S, Janssens W, Burgel PR, Contoli M, Franssen FME, Greening NJ, Greulich T, Gyselinck I, Halner A, Huerta A, et al: **Standardisation of Clinical Assessment, Management and Follow-Up of Acute Hospitalised Exacerbation of COPD: A Europe-Wide Consensus.** *Int J Chron Obstruct Pulmon Dis* 2021, **16:**321-332.
74. Rodriguez-Roisin R: **Toward a consensus definition for COPD exacerbations.** *Chest* 2000, **117:**398S-401S.
75. Anthonisen NR, Manfreda J, Warren CP, Hershfield ES, Harding GK, Nelson NA: **Antibiotic therapy in exacerbations of chronic obstructive pulmonary disease.** *Ann Intern Med* 1987, **106:**196-204.
76. Thakur N, Lovinsky-Desir S, Appell D, Bime C, Castro L, Celedon JC, Ferreira J, George M, Mageto Y, Mainous Iii AG, et al: **Enhancing Recruitment and Retention of Minority Populations for Clinical Research in Pulmonary, Critical Care, and Sleep Medicine: An Official American Thoracic Society Research Statement.** *Am J Respir Crit Care Med* 2021, **204:**e26-e50.

Identification of COPD exacerbations outcomes

Systematic reviews

What outcomes are tested in clinical trials?

What outcomes are important to patients?

- ✓ 123 Clinical trials
- ✓ 38 Systematic reviews
- ✓ 4 studies and one systematic review evaluating patients preferences

Qualitative research

- ✓ Focus group(s) and individual interviews with 86 patients from 11 countries, globally

What outcomes are important to patients all-over the world?

Prioritization of the core outcomes

Delphi Survey

Prioritization of critical outcomes to be included in the core outcome set.

Scoring:	Limited importance			Important, Not critical			Critical			
	1	2	3	4	5	6	7	8	9	Unable to score
Outcome #1	o	o	o	o	o	o	✓	o	o	o

- ✓ Global
- ✓ Multi-stakeholder
- ✓ 2-stage Delphi survey

First consensus meeting

- ✓Presentation of the findings
- ✓Discussion
- ✓Global, multi-stakeholder representation.

Finalisation of the core outcome set

Measurement instruments

Second consensus meeting

- ✓ Focused literature reviews
- ✓ Identified relevant high-quality methodological systematic reviews evaluating the PRO instruments
- ✓ Selected a single, optimal instrument for every core outcome

Global, evidence-informed consensus

Vision

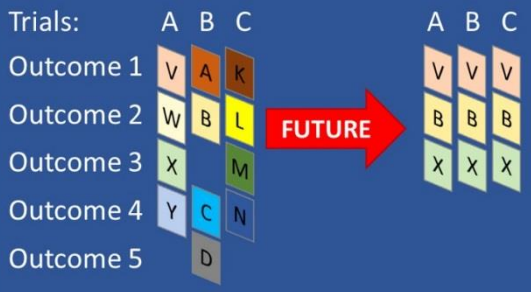
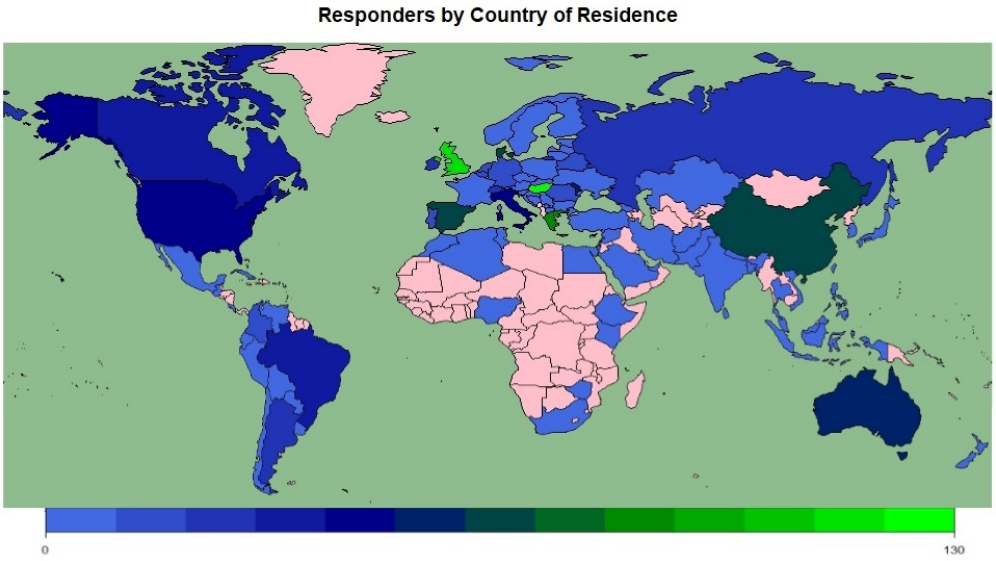


Figure: Outcomes 1,2,3: Core outcomes. Instruments V, B, X: Core measurement instruments.

Future trials will evaluate a minimum number of outcomes consistently, using the same measurement instruments.



Geographic distribution of the Delphi survey participants. The colour of each country represents the number of participants (see colour scheme).

Table 1: Baseline characteristics of the Delphi Survey Participants. Reported as N (% of the participants in the corresponding stakeholder group).

	Patients & Representatives	Health Professionals (HP)	Researchers
Study participants	**256** Patients: 229 Caregivers: 22 Representatives: 5	**488** Doctors: 399 Nurses: 53 Physiotherapists: 17 Other HP: 19	**319** Doctors: 230 Nurses: 13 Physiotherapists: 34 Other HP: 7 Others*: 35
Completed 2nd round	197 (77.0%)	398 (81.6%)	291 (91.2%)
Declared potential conflicts of interest	3 (1.2%)	13 (2.7%)	17 (5.3%)
Age (years):			
21 - 30	4 (1.6%)	74 (15.2%)	31 (9.7%)
31 - 40	10 (3.9%)	132 (27.0%)	91 (28.5%)
41 - 50	17 (6.6%)	109 (22.3%)	82 (25.8%)
51 - 60	56 (21.9%)	97 (19.9%)	64 (20.1%)
61 - 70	93 (36.3%)	62 (12.7%)	45 (14.1%)
71 - 80	66 (25.8%)	12 (2.5%)	6 (1.9%)
81 - 90	9 (3.5%)	2 (0.4%)	0 (0.0%)
>90	1 (0.4%)	0 (0.0%)	0 (0.0%)
Female (%)	112 (43.8%)	277 (56.8%)	140 (43.9%)
Continent:			
Africa	1 (0.4%)	6 (1.2%)	12 (3.8%)
Americas	44 (17.2%)	78 (16.0%)	51 (16.0%)
Asia	14 (5.5%)	68 (13.9%)	32 (10.0%)
Europe	175 (68.4%)	325 (66.6%)	201 (63.0%)
Oceania	22 (8.6%)	11 (2.3%)	23 (7.2%)
Economy:**			
Low	0 (0.0%)	1 (0.2%)	3 (0.9%)
Lower middle	12 (4.7%)	59 (12.1%)	19 (6.0%)
Upper middle	20 (7.8%)	125 (25.6%)	57 (17.9%)
High	170 (66.4%)	246 (50.4%)	175 (54.9%)
Conducting Research	2 (0.8%)	187 (38.3%)	283 (88.7%)
Designing Research studies	0 (0.0%)	0 (0.0%)	319 (100%)
Predominantly working on research	0 (0.0%)	21 (4.3%)	59 (18.5%)
Development of Guidelines	0 (0.0%)	95 (19.5%)	161 (50.5%)

*Others: Researchers and not health professionals; policy makers; regulators. HP: Health professionals. ** Economy of the participants' country, according to the World Bank Classification 2021.

Table 2: Additional baseline characteristics of patients with COPD who completed the Delphi survey.

Highest level of Education		
Primary education	23 (10.0%)	
Secondary education	111 (48.5%)	
University education	82 (35.8%)	
Not reported	13 (5.7%)	
Employment status		
Currently studying	1 (0.4%)	
Currently working	45 (19.7%)	
Currently unemployed	13 (5.7%)	
Early retirement	45 (19.7%)	
Retirement	117 (51.1%)	
Not reported	8 (3.5%)	
Years since COPD diagnosis		
Up to 5	66 (28.8%)	
6-10	68 (29.7%)	
11-15	43 (18.8%)	
16-20	28 (12.2%)	
Over 20	15 (6.6%)	
Not reported	9 (3.9%)	
Exacerbations history	**Any exacerbation**	**Severe (hospitalized) exacerbation**
None	55 (24.0%)	163 (71.2%)
1	49 (21.4%)	34 (14.8%)
2	36 (15.7%)	18 (7.9%)
3	31 (13.5%)	6 (2.6%)
4	12 (5.2%)	2 (0.9%)
More than 4	41 (17.9%)	2 (0.9%)
Not reported	5 (2.2%)	4 (1.7%)
Previous NIV use or ICU admission		
Yes	43 (18.8%)	
No	182 (79.5%)	
Not reported	4 (1.7%)	

Table 3: Additional baseline characteristics of expert respondents (health professionals and researchers) of the Delphi survey.

	Doctors	Nurses	Physiotherapists	Other health professionals	Researchers and not health professionals
Study participants	629	66	51	26	30
Completed 2nd round	522	63	50	23	27
Declared potential conflicts of interest	20 (3.2%)	1 (1.5%)	0 (0.0%)	5 (19.2%)	4 (13.3%)
Primary employment setting:					
Primary care	60 (9.5%)	5 (7.6%)	5 (9.8%)	4 (15.4%)	0 (0.0%)
Secondary hospital	121 (19.2%)	14 (21.2%)	2 (3.9%)	0 (0.0%)	0 (0.0%)
Tertiary/University hospital	348 (55.3%)	17 (25.8%)	30 (58.8%)	9 (34.6%)	2 (6.7%)
Clinical trials, methodology or epidemiology unit	1 (0.2%)	3 (4.5%)	0 (0.0%)	1 (3.8%)	1 (3.3%)
Health technology Assessment or guidelines development organization	3 (0.5%)	0 (0.0%)	0 (0.0%)	1 (3.8%)	2 (6.7%)
Governmental Organization	2 (0.3%)	0 (0.0%)	1 (2.0%)	0 (0.0%)	1 (3.3%)
Research funding organization/Charity	1 (0.2%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (3.3%)
Patients' organization	0 (0.0%)	0 (0.0%)	1 (2.0%)	1 (3.8%)	0 (0.0%)
Pharmaceutical industry	4 (0.6%)	0 (0.0%)	0 (0.0%)	4 (15.4%)	0 (0.0%)
Other	26 (4.1%)	3 (4.5%)	7 (13.7%)	1 (3.8%)	7 (23.3%)
Not reported	63 (10.0%)	24 (36.4%)	4 (7.8%)	5 (19.2%)	16 (53.3%)
COPD patients assessed during the previous year					
None	16 (2.5%)	4 (6.1%)	6 (11.8%)	5 (19.2%)	5 (16.7%)
1-250	283 (45.0%)	25 (37.9%)	29 (56.9%)	15 (57.7%)	2 (6.7%)
251-500	154 (24.5%)	8 (12.1%)	10 (19.6%)	0 (0.0%)	0 (0.0%)
501-750	58 (9.2%)	3 (4.5%)	4 (7.8%)	1 (3.8%)	0 (0.0%)
751-1000	30 (4.8%)	1 (1.5%)	1 (2.0%)	0(0.0%)	0 (0.0%)
>1000	35 (5.6%)	1 (1.5%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Not reported	53 (8.4%)	24 (36.4%)	1 (2.0%)	5 (19.2%)	23 (76.7%)
Research activity:					
Involved in conducting research	369 (58.7%)	29 (43.9%)	40 (78.4%)	16 (61.5%)	13 (43.3%)
Involved in designing research	230 (36.6%)	13 (19.7%)	34 (66.7%)	9 (34.6%)	11 (36.7%)
Devote >50% of their working time to research	45 (7.2%)	11 (16.7%)	11 (21.6%)	6 (23.1%)	7 (23.3%)
Involved in developing guidelines	210 (33.4%)	18 (27.3%)	17 (33.3%)	5 (19.2%)	5 (16.7%)

Table 4: Summary of the selection process of the core outcomes from the longlist. Percentages refer to the proportion of participants that consider a particular outcome critical. Background colour coding: First column: Grey, blue, purple colours signify outcomes identified through the methodological systematic reviews, qualitative interviews, or the Delphi survey, respectively. In the remaining columns, background colour refers to the results of the outcome selection process at each stage; green, yellow and red colours signify inclusion, inconclusive result or exclusion of the respective outcome.

COPD exacerbations outcomes considered		**Delphi Survey Results**			**Consensus meeting**	
Sources of outcomes Methodological SRs Qualitative interviews Delphi survey (Round 1)	**Outcomes' selection results** Included Inconclusive Excluded	Patients & Patient representatives	Health professionals	Researchers	Patients & Patient representatives	Health professionals & Researchers
Death outcomes						
Death from COPD Exacerbation		81.8%	94.5%	96.9%		
Death from any cause		68.5%	74.8%	84.0%	100%	100%
Clinical and Physiological Outcomes						
Anxiety		35.5%	27.0%	28.3%		
Breathlessness		79.3%	93.3%	94.9%		
Chest discomfort		15.8%	5.8%	8.2%		
Fatigue		54.2%	46.3%	44.7%		
Cough		49.3%	54.3%	53.6%		
Coughing up blood (haemoptysis)		62.1%	58.3%	46.8%		
Production of dark-coloured sputum		56.7%	58.5%	53.6%		
Sputum amount		38.4%	42.0%	35.5%		
Sputum thickness (ease of expectoration)		40.4%	41.8%	29.0%		
Wheeze		39.4%	46.8%	35.2%		
Appetite		24.6%	17.5%	14.0%		
Change in weight		33.5%	25.8%	23.9%		
Respiratory muscle strength		65.5%	58.8%	47.8%		
Low mood/ depression		41.9%	35.5%	40.6%		
Sleep quality		51.7%	38.3%	35.5%		
Early morning symptoms		36.5%	32.0%	25.6%		
Night time symptoms		45.8%	50.3%	41.3%		

Treatment success (or failure)	80.3%	87.8%	89.1%		
Worsening of symptoms after the initial treatment	71.9%	78.5%	77.1%		
Disease progression	83.7%	88.8%	86.7%		
Future exacerbations	75.9%	89.3%	90.4%		
Lung function during and immediately after the exacerbation	71.4%	54.3%	43.0%	7.7%	11.1%
Permanent deterioration in lung function	87.7%	88.5%	82.3%		
Levels of oxygen and carbon dioxide in the blood (arterial blood gases)	76.4%	80.3%	75.4%		
Development of pneumonia	76.4%	86.8%	83.6%		
Development of resistant bacteria	73.4%	80.8%	70.6%		
Damage of lung cells and lung tissue	81.3%	71.5%	57.3%	38.5%	22.2%
Infection by bacteria (bugs) or viruses	72.4%	68.0%	64.8%	92.9%	68.4%
Inflammation in the lungs/airways	73.4%	61.5%	49.1%	50.0%	27.8%
Adverse event outcomes					
Adverse events of treatments	60.6%	58.3%	65.9%		
Serious adverse events from treatments	76.8%	89.5%	93.5%		
Development and/or progression of other diseases (e.g. heart attack)	67.5%	69.5%	69.6%		
Resources use outcomes					
Need for hospital admission for the presenting exacerbation	69.0%	84.6%	90.8%	100%	100%
Length of hospital stay for the exacerbation	45.3%	62.3%	68.3%		
Future hospital admissions	52.2%	70.5%	76.5%	71.4%	77.8%
Need for non-invasive ventilation (NIV) use for the exacerbation	64.0%	83.5%	81.9%	61.5%	78.6%
Length of non-invasive ventilation (NIV) use for the exacerbation	58.1%	60.25%	57.0%		
Need for admission to the intensive care unit for the exacerbation	71.9%	86.8%	88.7%		
Length of stay in the intensive care unit for the exacerbation	63.1%	72.8%	71.0%	38.5%	50%
Need for additional medications to achieve symptoms control	64.5%	59.5%	57.3%		
Need for long-term administration of supplemental oxygen after the exacerbation	58.6%	62.8%	66.9%		
Need for long-term use of non-invasive ventilation (NIV) after the exacerbation	55.7%	69.5%	65.5%		
Life impact outcomes					
Ability to exercise	57.6%	51.0%	60.4%		
Physical strength	48.8%	38.3%	35.5%		
Walking distance	57.6%	67.3%	68.3%		
Activities of daily living	70.4%	82.5%	84.6%		

Health related quality of life	75.4%	82.5%	87.7%		
Social engagement/ isolation	54.2%	50.5%	50.5%		
Treatment adherence	72.4%	83.8%	84.6%		
Impact of family members and caregivers	56.7%	50.3%	47.4%		
Impact on sexual function	36.0%	36.3%	37.5%		

Table 5: Outcome measurement instrument recommendations. Green and yellow background colours signify a strong recommendation and an interim recommendation with associated research agenda, respectively.

Death from any cause. Death from any cause during study period. Record date of death.
Death from COPD exacerbation. Consider the immediate cause of death as documented in the death summary. In cases of death due to an immediate complication of an exacerbation, such as a ventricular arrhythmia, massive pulmonary embolism, or myocardial infarction, the exacerbation should be considered the cause of death. Ideally, cause of death will need to be confirmed by a blinded adjudication committee. However, this may not always be feasible.
Treatment success. Treatment success defined as sufficient improvement of the signs and symptoms of the exacerbation that no additional systemic treatments (antibiotics or systemic corticosteroids) are required.
Need for hospital admission for the presenting exacerbation. A clinical need to admit a patient to the hospital, or equivalent intensification of the monitoring or care that may be provided in other settings (including patients' home). Admissions for social reasons should be reported separately. For evaluating this outcome investigators should record whether a patient required admission at any timepoint and whether they still required hospital admission at a specific follow-up timepoint.
Need for admission to the intensive care unit (ICU) for the presenting exacerbation. Need for ICU admission should be evaluated on the basis of the need for invasive mechanical ventilation, defined as (i) persistent or deteriorating respiratory acidosis despite optimized medical treatment and delivery of non-invasive ventilation (NIV); (ii) persistent or deteriorating respiratory acidosis despite optimized medical treatment and a contra-indication for the use of NIV, for example due to severe facial deformity where fitting a mask is impossible, upper airway obstruction, or facial burns; (iii) respiratory arrest or peri-arrest situations unless there is a rapid recovery from manual ventilation or provision of NIV. For evaluating this outcome investigators should record whether a patient required admission at any timepoint and whether they still require ICU admission at a specific follow-up timepoint.
Levels of oxygen and carbon dioxide in the blood (arterial blood gases). A setting and intervention specific outcome. A baseline and at least one follow-up measurement are required with a clear indication of whether or not the patient was receiving oxygen at the time of the measurement, and if yes, how much. It may not be feasible for studies evaluating outpatients.

Breathlessness. Breathlessness should be evaluated using the modified Borg's scale. It should be measured at approximately the same time every day. It can be self-completed.
Health related quality of life. The COPD Assessment Test (CAT) should be used for assessing health related quality of life.
Activities of daily living. The Capacity of Daily Living in the Morning Questionnaire (CDLM) should be used for evaluating basic activities of daily living during the exacerbation. The Manchester Activities of Daily Living Questionnaire (MRADL) should be used for evaluating basic and instrumental activities of daily living, during recovery (long-term impact of the exacerbation).
Worsening of symptoms after the initial treatment. The modified Borg's scale and the COPD assessment test (CAT) should be used to detect symptoms worsening after the initial treatment.
Disease progression. Permanent deterioration in lung function should be used to evaluate the impact of exacerbations on disease progression. Two pulmonary function tests during stable clinical condition are needed: One within 6 months prior to the index exacerbation, and one within 2-6 months afterwards. Change from baseline in forced expiratory volume in 1 second (FEV_1), forced vital capacity (FVC) and FEV_1/FVC ratio should be noted. The number of exacerbations experienced between the two measurements should be noted. Ideally, only the index exacerbation should be included between the two measurements. Disease progression as a core outcome is only relevant for longer-term studies that recruit participants during stable disease state, in anticipation of an exacerbation.
Future exacerbations. Future exacerbations, noting whether they are moderate or severe, after treatment success is confirmed.
Future hospital admissions. Future hospital admissions for any medical reason, or equivalent intensification of the monitoring or care that may be provided in other settings, after treatment success is confirmed.
Serious adverse events from treatments. Following the definition of the International Council for Harmonisation. Serious adverse event is any untoward medical occurrence in a patient or clinical investigation subject administered a pharmaceutical product and which does not necessarily have to have a causal relationship with the treatment, that fulfils any of the following: (a) Results in death; (b) Is life threatening; (c) Requires inpatient hospitalisation or prolongation of existing hospitalisation; (d) Results in persistent or significant disability / incapacity; (e) Is a congenital anomaly or birth defect.

Suspected unexpected serious adverse reactions (SUSARs) should also be reported.			
Development of resistant bacteria. Trials evaluating antimicrobials, antimicrobial stewardship strategies, novel immune modifiers or other interventions that may affect bacterial resistance should evaluate bacterial resistance to the administered antibiotics in spontaneous sputum. As a minimum, resistance should be evaluated at baseline and within a week after treatment completion. Sputum induction may provide additional information. However, in each study, researchers should consider the balance between the added value compared to the risk, participants discomfort and required resources.			
Development of pneumonia. Pneumonia confirmed by the presence of new consolidation in the chest x-ray or other imaging modalities of the chest, in the presence of consistent clinical signs and symptoms. When possible, chest imaging should be acquired at baseline, to assess for the presence of pneumonia. This may not be possible for trials recruiting patients outside the hospital setting. Follow-up chest imaging should be driven by clinical need.			
Treatment adherence. An intervention specific outcome. Methods for assessing treatment adherence should be clearly reported.			
	Strong recommendation		**Interim recommendation, with research agenda**

Online Appendix

ERS Statement: A core outcome set for clinical trials evaluating the management of chronic obstructive pulmonary disease exacerbations.

Alexander G. Mathioudakis, Fekri Abroug, Alvar Agusti, Sachin Ananth, Per Bakke, Konstantinos Bartziokas, Bianca Beghe, Andras Bikov, Thomas Bradbury, Guy Brusselle, Cordula Cadus, Courtney Coleman, Marco Contoli, Alexandru Corlateanu, Olga Corlateanu, Gerard Criner, Balazs Csoma, Alexander Emelyanov, Rosa Faner, Gustavo Fernandez Romero, Zeineb Hammouda, Peter Horváth, Arturo Huerta Garcia, Michael Jacobs, Christine Jenkins, Guy Joos, Olga Kharevich, Konstantinos Kostikas, Elena Lapteva, Zsofia Lazar, Joerg D. Leuppi, Carol Liddle, John Linnell, Alejandra López-Giraldo, Vanessa M. McDonald, Rune Nielsen, Alberto Papi, Isabel Saraiva, Galina Sergeeva, Agni Sioutkou, Pradeesh Sivapalan, Elizabeth Stovold, Hao Wang, Fuqiang Wen, Janelle Yorke, Paula R. Williamson, Jørgen Vestbo, Jens-Ulrik Jensen, on behalf of the DECODE-NET.

Appendices

The DECODE-NET investigators

Aaron Shawn, Abroug Fekri, Agusti Alvar, Ananth Sachin, Attaway Amy, Bakke Per, Baldomero Arianne, Barjaktarevic Igor, Bartziokas Konstantinos, Beghe Bianca, Bhat Surya P, Bikov Andras, Billings Joanne, Bolton Charlotte, Bourdin Arnaud, Bradbury Thomas, Brusselle Guy, Burkes Robert, Cadus Cordula, Cassady Steven, Contoli Marco, Corlateanu Alexandru, Corlateanu Olga, Criner Gerard, Csoma Balázs, Dean Victoria, Demeyere Benedicte, Diaz Philip, Dobler Claudia C, Dransfield Mark T, Duffy Sean, Echevarria Carlos, Eklöf Josephin, Emelyanov Alexander, Faner Rosa, Fernandez Romero Gustavo, Fisk Marie, Gaeckle Nate, Gangemi Andrew, Gkrepi Georgia, Hammouda Zeineb, Han MeiLan K, Hill Nicholas S, Horvath Peter, Huerta Arturo, Hurst John, Iyer Anand, Jacobs Michael, Jenkins Christine, Jensen Jens-Ulrik, Joos Guy, Katibnikova Elena, Kharevich Olga, Kostikas Konstantinos, Krishnan Jerry A, Kunisaki Ken, Kyriakopoulos Christos, Labaki Wassim, LaBedz Stephanie, Lafon David, Lahousse Lies, Lapteva Elena, Lazar Zsofia, Leuppi Jörg, Liddle Carol, Linnell John, Lopez-Giraldo Alejandra, MacIntyre Neil, Marchetti Nathaniel, Martinez Fernando D, Martinez Fernando J, Mathioudakis Alexander G, McDonald Vanessa, McLoughlin Rebecca, Melzer Anne, Mkorombindo Takudzwa, Mosher Christopher, Moy Marilyn, Sunita Mulpuru, Nielsen Rune, Papi, Alberto, Panos Ralph, Parekh Trisha, Pavord Ian, Pinto-Plata Victor, Potonos Dimitris, Quint Jennifer, Reed Robert M, Rommes Jean, Saraiva Isabel, Sargeeva Galina, Sciurba Frank C, Singh Dave, Sivapalan Pradeesh, Soler Nestor, So Jennifer, Stavila Ecaterina, Stolz Daiana, Strange Charlie B, Suehs Carey, Torsani Fransesca, Vestbo Jørgen, Wang Hao, Wells James, Wen Fuqiang, Wendt Chris, Williamson Paula, Yorke Janelle, Zantah Massa.

2 List of professional and patient organizations that facilitated the Delphi survey dissemination

We are very thankful to the following organizations for distributing the Delphi Survey to their memberships and/or through their social media.

- ✓ Allergy and Asthma Network,
- ✓ Alpha-1 Netherlands,
- ✓ Alpha-1 Spain,
- ✓ Alpha-1Plus (Belgium),
- ✓ Asian Pacific Society of Respirology (APSR)
- ✓ Association of Pulmonologists of Greece,
- ✓ Australian Lung Foundation,
- ✓ Brazilian Respiratory Society,
- ✓ British Lung Foundation (BLF),
- ✓ COPD Canada,
- ✓ COPD Foundation (USA),
- ✓ COPD Ireland,
- ✓ Danish Lung Association,
- ✓ Dutch Lung Foundation,
- ✓ Georgian Respiratory Association,
- ✓ Global Allergy & Airways Patient Platform,
- ✓ Greek Association of General Practitioners,
- ✓ Hellenic Thoracic Society,
- ✓ Hungarian Respiratory Society,
- ✓ Indonesian Respiratory Society,
- ✓ Irish Thoracic Society,
- ✓ Jedra Organisation to Help Those Suffering of Lung Cancer and Other Lung Diseases (Croatia),
- ✓ Kazakhstan Respiratory Society,

- ✓ Lovexair (Spain),
- ✓ National Institute for Health Research (NIHR) BEAT Respiratory Disease,
- ✓ NTM Info & Research (USA),
- ✓ Pan African Thoracic Society (PATS),
- ✓ Philippine College of Chest Physicians,
- ✓ Respiriamo Insieme (Italy),
- ✓ Russian Respiratory Society,
- ✓ Swiss Society of Pulmonology,
- ✓ Sociedad Española de Neumología y Cirugía Torácica (SEPAR),
- ✓ Swedish Heart and Lung Association,
- ✓ Task Force for Lung Health,
- ✓ Thoracic Society of Australia and New Zealand (TSANZ),
- ✓ US COPD Coalition,
- ✓ Turkish Respiratory Society.

3 Detailed methodology of the COS-AECOPD ERS Task Force

Detailed methodology of the COS-AECOPD (Core outcome set for the management of acute exacerbations of chronic obstructive pulmonary disease) ERS Task Force was prospectively registered with the COMET database (www.comet-initiative.com; ID: 1325) and published [1]. This study was conducted and reported following the methodology recommended by the COMET initiative (the COMET handbook) [2], the Core Outcome Set STAndards for Development (COS-STAD)[3] and STAndards for Reporting (COS-STAR) [4].

This section describes methodology of this task force in more detail and summarizes the findings of the systematic reviews and qualitative data that informed the development of the initial long list of COPD exacerbations outcomes (reported in detail separately).

3.1 Study oversight

The Task Force was co-chaired by Alexander G. Mathioudakis and Jens-Ulrik Jensen. A steering committee was formed consisting of the Task Force co-chairs and Jørgen Vestbo (clinical researchers with expertise in clinical trials in COPD), Carol Liddle and Isabel Saraiva (patient representatives) and Paula Williamson (chair of the COMET initiative). The steering committee was responsible for the management and co-ordination of the study and met regularly (face-to-face or via teleconference) to review the study progress, ensure the study complied with good clinical practice principles, relevant regulations, and adhered to the study protocol. Feedback from the ERS Task Force panel (consisting of clinical researchers with expertise in the management of COPD exacerbations, methodologists, and patient representatives; the authors of this document) was sought regularly via email. The recommendations about the core outcomes and their measurement instruments were finalized in two virtual consensus meetings on April 21st and 28th, 2021 and were attended by panel members (consisting of health professionals, researchers, methodologists and patient representatives) and additional patient representatives.

3.2 Management of the conflicts of interest

Potential conflicts of interest of the panel members and all consensus meeting participants were reported and managed in line with the ERS policies (available here: https://www.ersnet.org/ science-and-research/development-programme/). None of the panel members or consensus meeting participants reported any conflicts directly related to this project, but in the event such conflicts had been reported, our plan was to ask members with such conflicts to abstain from the respective polls.

3.3 Identification of COPD exacerbations outcomes

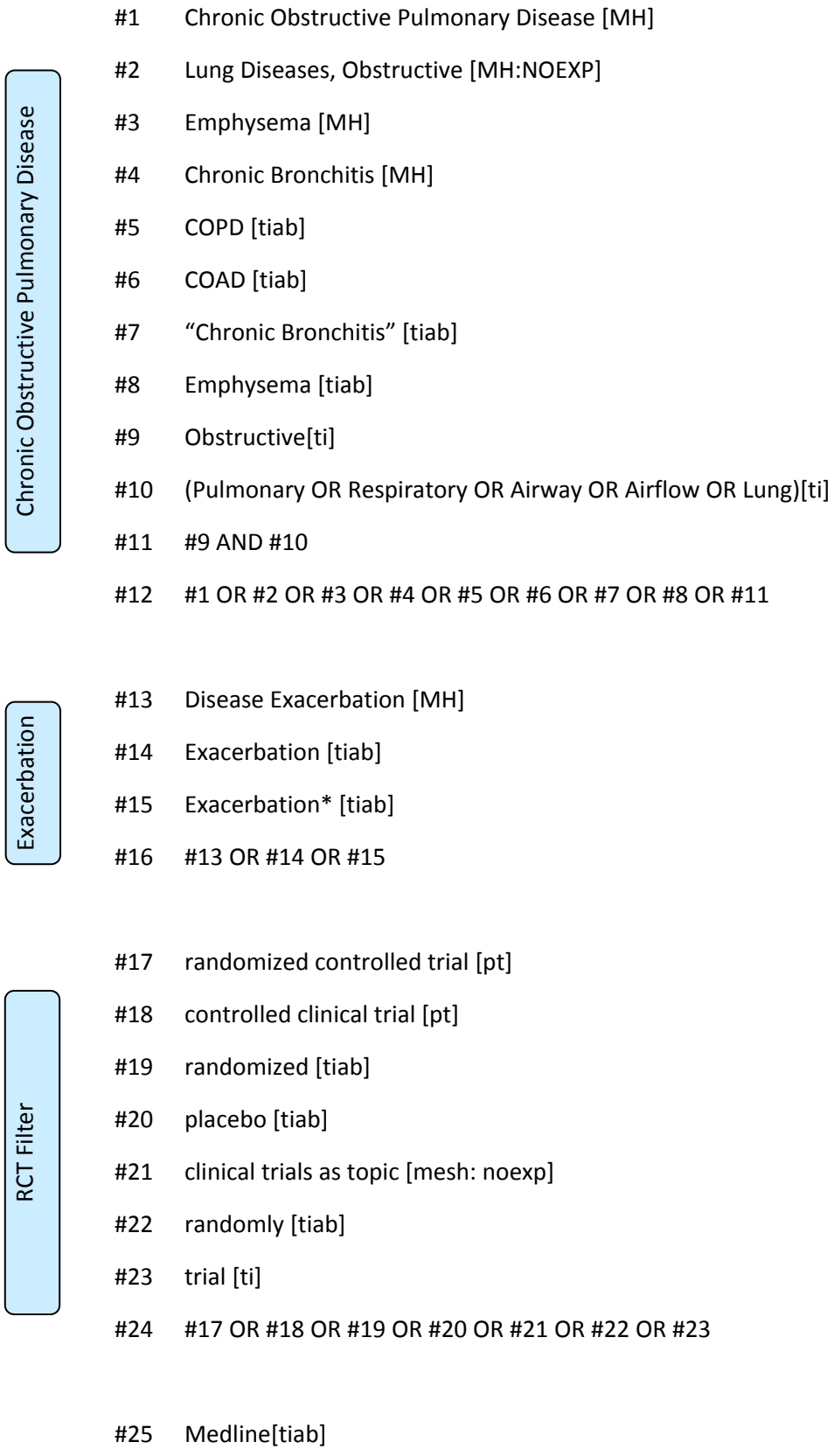
For the development of this core outcome set, in line with recommendations by the COMET initiative, we first developed a comprehensive list of all outcomes related to COPD exacerbations. This list was informed by (i) a methodological systematic review to capture the outcomes evaluated in randomized controlled trials (RCTs) and systematic reviews on the management of COPD exacerbations, (ii) a focused systematic review of qualitative studies exploring outcomes considered important by patients and their caregivers, and (iii) qualitative research consisting of a focus group and individual interviews with patients with COPD from 11 countries globally.

3.3.1 Systematic review of outcomes evaluated in RCTs and SRs on COPD exacerbations management

This methodological systematic review has been reported separately [5, 6]. In brief, we searched Medline/ PubMed for RCTs and systematic reviews of RCTs evaluating pharmacological and non-pharmacological interventions for the management of COPD exacerbations, published between 2006-2018. Detailed search strategy is presented in figure S1 and the PRISMA flowchart in figure S2. Two authors screened the titles and abstracts of all studies yielded by the search and the full text of all potentially eligible studies based on the initial screening. Main characteristics of the included studies and details about the outcomes evaluated and measurement instruments used were extracted in a structured excel form by

one author and cross-checked by a second author. In each step of this process, disagreement was resolved by consensus among the authors.

Figure S1. Search strategy (reproduced from [5])



SR Filter

#26 Systematic[tiab] and (review[tiab])

#27 Meta analysis[publication type]

#28 Meta-analysis[tiab]

#29 Metaanalysis[tiab]

#30 #25 OR #26 OR #27 OR #28 OR #29

Final selection

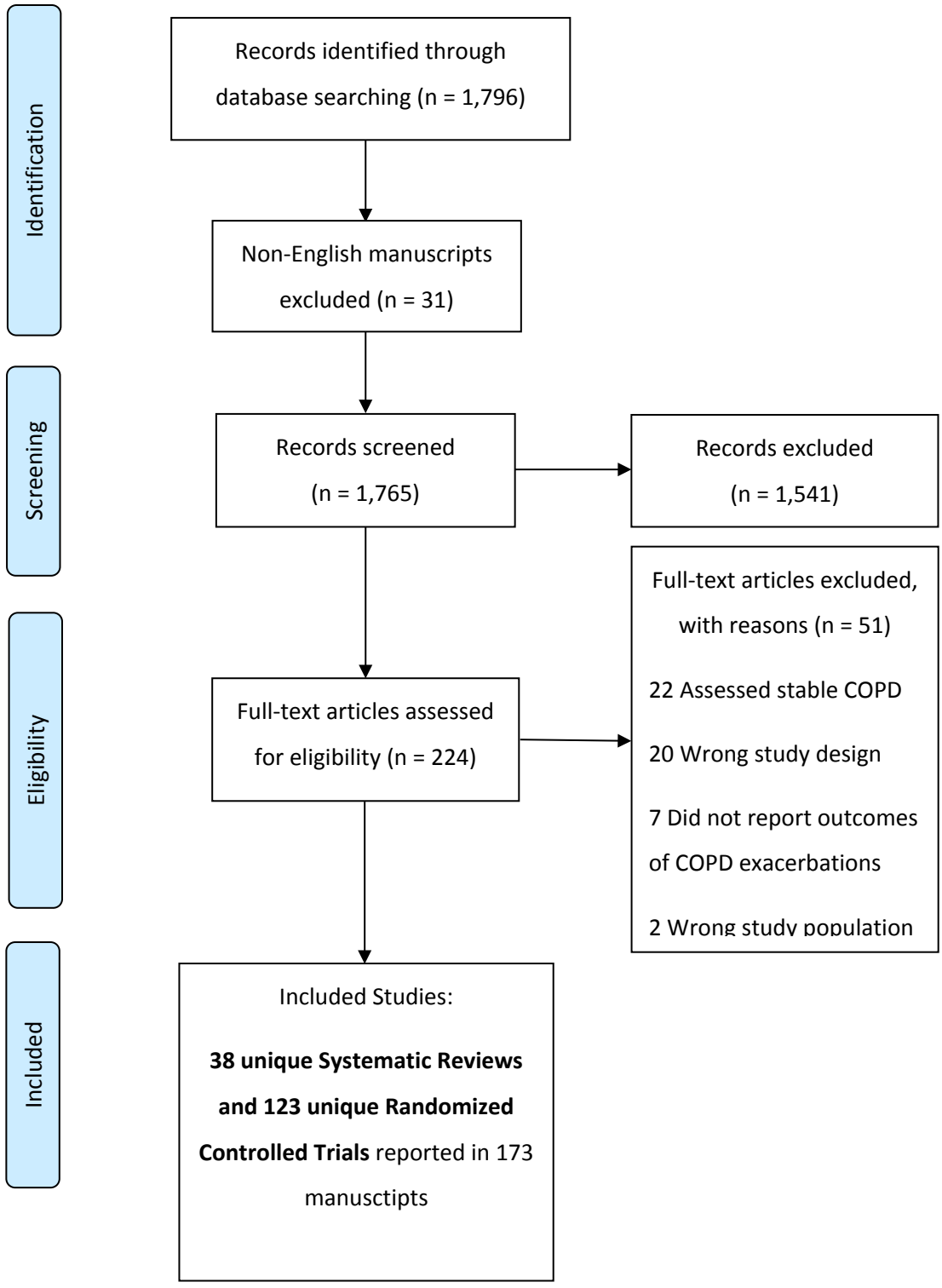
#31 Search ("2006"[Date - Publication] : "2017"[Date - Publication])

#32 animals [mh] NOT humans [mh]

#33 #12 AND #16 AND #31 AND (#24 OR #30)

#34 #33 NOT #32

Figure S2. PRISMA Flowchart (reproduced from [5])



We identified 123 eligible RCTs and 38 systematic reviews. Upon deduplication, we identified 39 unique outcomes that are summarized in table 4 of the main text. The outcomes that were most frequently evaluated in the included studies are summarized in table S1.

Table S1. Frequency that different outcomes were reported in the 123 randomised controlled trials (RCTs) and 38 systematic reviews (SRs) included in the methodological review (reproduced from [5])

Outcomes	Frequency of reporting	
	RCTs n (%)	**SRs n (%)**
Patient important Outcomes		
Mortality	101 (82%)	29 (76%)
Treatment success or failure	77 (63%)	29 (76%)
Adverse effects	73 (59%)	26 (68%)
Health status, symptoms & quality of life	73 (59%)	17 (45%)
Duration of exacerbations	42 (34%)	20 (53%)
Re-exacerbation, re-hospitalization	33 (27%)	16 (42%)
Exercise capacity	14 (11%)	1 (3%)
Anxiety and depression	6 (5%)	1 (3%)
Surrogate, Physiological and Laboratory Outcomes		
Lung function	58 (47%)	18 (47%)
Arterial blood gases and oxygen saturation	40 (33%)	5 (13%)
Microbiological response	16 (13%)	7 (18%)
Biomarkers	32 (26%)	2 (5%)
Medication use	18 (15%)	3 (8%)

3.3.2 Systematic review of qualitative studies exploring outcomes considered important by patients and their caregivers.

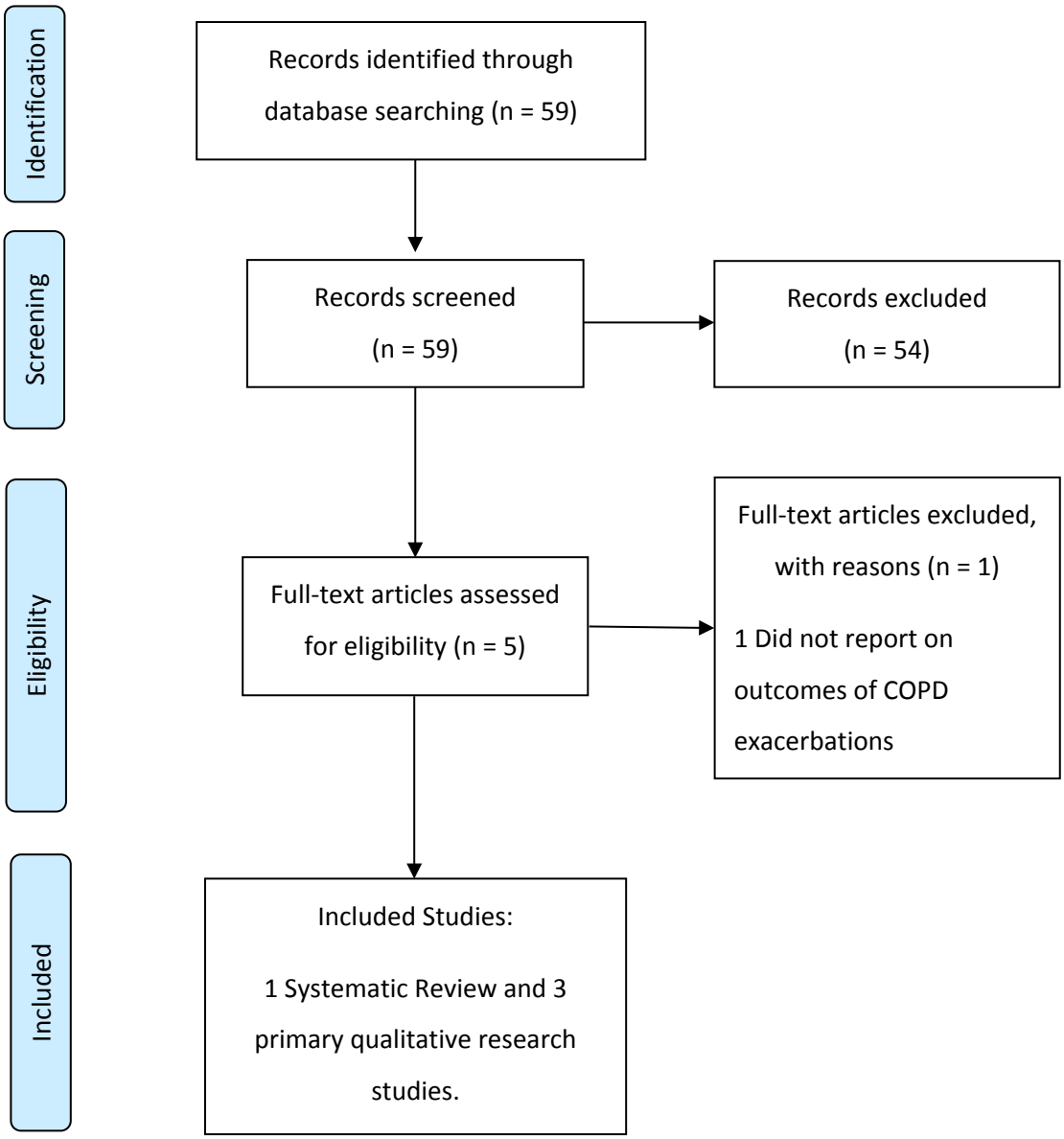
To enrich the list of COPD exacerbation outcomes, we conducted a systematic review aiming to identify qualitative studies evaluating the experiences, views and preferences of patients with COPD and their caregivers around the management of COPD exacerbations. We searched Medline/ PubMed using a filter for qualitative studies on the outcomes of diseases that was developed by the COMET group [7]. Detailed search strategy is presented in figure S3 and the PRISMA flowchart in figure S4. Titles and abstracts and -when required- full texts were screened by two authors independently for eligibility. One author identified all outcomes of COPD exacerbations that were described in the included studies and a second author cross-checked for accuracy. Disagreement was resolved by consensus among the authors.

One systematic review [8] and three primary qualitative research studies [9-11] were selected for inclusion. Overall, this review yielded two additional outcomes that were incorporated in the long-list: (i) Anxiety and (ii) Fatigue.

Figure S3. Search strategy

Chronic Obstructive Pulmonary Disease	#1	Chronic Obstructive Pulmonary Disease [MH]
	#2	Lung Diseases, Obstructive [MH:NOEXP]
	#3	Emphysema [MH]
	#4	Chronic Bronchitis [MH]
	#5	COPD [tiab]
	#6	COAD [tiab]
	#7	"Chronic Bronchitis" [tiab]
	#8	Emphysema [tiab]
	#9	Obstructive[ti]
	#10	(Pulmonary OR Respiratory OR Airway OR Airflow OR Lung)[ti]
	#11	#9 AND #10
	#12	#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #11
Exacerbation	#13	Disease Exacerbation [MH]
	#14	Exacerbation [tiab]
	#15	Exacerbation* [tiab]
	#16	flare* [tiab]
	#17	#13 OR #14 OR #15 OR #16
Qualitative	#18	qualitative [tiab]
	#19	themes [tiab]
	#20	#18 or #19
Disease outcomes	#21	symptom* [tiab]
	#22	treatment* [tiab]
	#23	living with [tiab]
	#24	patient* [tiab]
	#25	#21 OR #22 OR #23 OR #24
	#26	#12 AND #17 AND #20 AND #25

Figure S4. PRISMA Flowchart



3.3.3 Qualitative research

To further complement the long-list of outcomes of COPD exacerbations, we conducted qualitative research to identify outcomes that patients deem important that might have not been captured by our systematic reviews. We conducted a focus group (n=8 participants, UK) and individual interviews with a total of 86 purposefully selected patients with COPD from 11 countries globally (Australia, Belarus, China, Denmark, Greece, Hungary, Moldova, Russia, Spain, Tunisia and the United Kingdom). We involved patients with a history of a recent hospitalised exacerbation, patients with frequent moderate exacerbations (treated in the community) and patients with a history of exacerbations with concomitant type 2 respiratory failure, requiring non-invasive ventilation. We included both male and female patients and sought to involve different age groups, geographic, cultural, and socioeconomic backgrounds. Preselected open-ended questions/ discussion topics were used to prompt participants to describe their experience of COPD exacerbations, to describe how exacerbations affect them and, -most importantly- what specific effects they would like a new treatment for COPD exacerbations to have on them. The aim of these questions was to encourage patients to describe outcomes of COPD exacerbations. A detailed list of the questions is available in table S2. At the end of the focus group and interviews, participants were asked to review the plain language description of the long-list of outcomes [either in their language or the English language if a translation in their language was unavailable]. To help us improve the plain language descriptions of the outcomes, they were asked to provide feedback on the simplicity and clarity of descriptions and to explain their understanding of each outcome. The focus group and interviews were audio-recorded. Each recruiting centre retained the recordings (to preserve patients' data and confidentiality) and the local investigators extracted quotes potentially describing COPD exacerbations outcomes.

Six additional outcomes were identified and added to our long-list of COPD exacerbations outcomes: (i) Appetite, (ii) Sleep quality, (iii) Early morning symptoms, (iv) Night-time symptoms, (v) Disease progression, and (iv) Social engagement / isolation. Additional details on the interviews will be reported separately.

Table S2: Qualitative research – List of open ended questions

1.	**Ask about the experience of having a COPD exacerbation.** "How has your experience of your last exacerbation/flare-up been?" "Can you tell me about your experience of having a flare-up of your COPD?" "How did your last exacerbation affect you?"
2.	**Ask about the impact of exacerbations on patients' health and well-being:** "What have been the challenges from COPD exacerbations/flare-ups to your health and wellbeing?" Prompt specifically about physical/ mental/ social wellbeing. "When you have a flare-up of your COPD, how does this impact your life? "How is your life different while your COPD is stable compared to when you have a flare-up?"
3.	**Ask about the treatments that are offered for COPD exacerbations.** "During your previous exacerbations, when did you decide that you needed treatment?" "What treatments were you offered / did you use for your recent exacerbations?" "When was the last time you had a discussion with a doctor or nurse about the treatments you receive for exacerbations? What factors did you consider when deciding to try or not try a treatment?"
4.	**Ask about their expectations from treatments of COPD exacerbations. Ask specifically about pharmacotherapy and non-invasive ventilation (for patients who have used it).** "To what extent the effects of treatments you had for your exacerbations matched your expectations?" "What specifically have you hoped for from the treatments for your COPD exacerbations?" Prompt specifically about physical/ mental/ social wellbeing.
5.	**Ask about the effects that COPD exacerbations treatments have:** "How medicines for your flare-ups make you feel?". Also ask for NIV. "What do you consider to be the most beneficial effects of treatments? "What are the most concerning effects (called side effects) of medications for your exacerbations, for you?" Prompt for specific areas such as physical/ mental/ social impact
6.	**Ask about concerns for future COPD exacerbations:** "What concerns do you have about your future COPD exacerbations/flare-ups?" "What are the most concerning effects of exacerbations in your life?"
7.	**"If a new treatment became available, what specific effects would you like it to have on you?"** Prompt for details on physical/ mental/ social impact. "Cure" is not an acceptable response here.
8.	**After making sure the participants understand what an outcome is, ask explicitly which outcomes they think are important to be evaluated.** **Plain English Language definition of outcomes:** To help patients, doctors and other health professionals make decisions about treatments, we need evidence about what works best. Treatments are developed and tested by researchers to make sure they work and are safe. To do this, researchers need to look at the effects those treatments have on patients. Researchers do this by

	measuring an 'outcome'. For example, in a study of how well a new asthma treatment works, 'outcomes' might include: - Night time wheeze - Quality of life measures (Can describe instead outcomes of COPD exacerbations that the patients have already mentioned) "Which outcomes do you think are important to be evaluated?" Prompt for details on outcomes related to physical/ mental/ social wellbeing. Ask why they think those outcomes are more important (and document participants' quotes): "You 've said X outcome is important, what makes you think that?" Avoid "why" questions as those can make people feel put on the spot.
9.	**Ask whether they think their perspective on what is important has changed over time.** "Do you think anything has altered your perspective regarding your exacerbations and their outcomes?" "Were there any outcomes that you considered important previously that were not mentioned during this interview?"
10.	**Plain English Version of the outcomes:** Discuss each of the outcomes described in the following table. Ask patients to describe them in their own language. Do they understand the outcomes correctly? At the end, ask again the patients if they think any other important outcomes are missing from our list. You should clearly highlight outcomes that were volunteered by patients earlier, compared to the outcomes that were discussed by the interviewer later.

3.3.4 Finalization of the long-list of outcomes

After deduplication, the long-list of outcomes of COPD exacerbations management included 47 unique outcomes. Of these, 39 originated from the first methodological systematic review, two from the systematic review of qualitative research studies and six from the qualitative research that we conducted. This list was further enriched by the respondents of the Delphi survey, as described in the next section.

Following the COMET taxonomy, all identified outcomes were grouped in five areas: Mortality or Survival outcomes, Physiological or Clinical, Life impact, Resource Use, and Adverse Events or adverse effects outcomes [12].

3.4 Prioritization of outcomes for inclusion in the core outcome set.

Prioritization of the most critical outcomes for inclusion in the core outcome set was facilitated by an online, two-stage, global, multistakeholder, modified Delphi survey and a consensus

meeting involving patient representatives, clinicians and clinical researchers with relevant expertise and global representation.

3.4.1 Modified Delphi survey

The modified Delphi survey, along with detailed instructions and description of the research project were developed in plain language with input from the European Lung Foundation (ELF) and lay members of the ELF's COPD Patient Advisory Group. It was translated in 10 languages (Chinese simplified, Danish, English, German, Greek, Hungarian, Italian, Portuguese, Russian, Spanish). Translations were validated using two-way translations by native speakers. The survey was conducted using DelphiManager, a secure, online software developed by the COMET initiative [13].

Three stakeholder groups were invited to participate in the survey: (a) Patients diagnosed with COPD, who had experienced exacerbations, and personal caregivers or representatives of such patients (e.g., patient organisations); (b) Health professionals caring for patients (e.g., doctors, nurses or physiotherapists); and (c) Clinical researchers (health professionals who care for patients but are also involved in designing research studies).

The survey was disseminated broadly, to health professionals, members of the ERS with a documented interest in airway diseases, as well as members of other national and international scientific societies. It was also disseminated to patients with COPD and their caregivers through the ELF's network of local, national, and international organisations representing patients across the world. The complete list of professional and patient organisations that disseminated the survey is available in online appendix 2. Finally, the survey was publicized through social media (Twitter and Facebook); it was shared by the panel members and the previously mentioned professional and patient organizations.

In the first round of the Delphi survey, after completing their baseline characteristics and declaring potential conflicts of interest, participants were presented with a list of 47 unique outcomes identified through the previously described systematic reviews and qualitative

research studies. Participants were asked to rate the importance of each outcome for clinical decision making on a scale from 1 to 9, following the Grading of Recommendations, Assessment and Evaluation (GRADE) guidance [14, 15]. Scores between 1-3, 4-6 and 7-9 signified outcomes of limited importance, important but not critical, and critical outcomes, respectively. Finally, respondents were encouraged to suggest additional outcomes they considered important that had not been included in the survey.

Only participants who completed the first round of the survey by providing ratings for at least 80% of the outcomes were included in the analyses and invited to participate in the second round. In the second survey round, participants were presented with graphical displays of the distribution of scores submitted from each stakeholder group during the first round of the survey. The outcomes list was supplemented by additional, new outcomes identified during the first survey round. Respondents were asked to re-consider their ratings taking into account how the different stakeholder groups rated each of the outcomes, clarifying that they should not feel under any pressure to change their ratings if they did not want to.

After the second Delphi round, consensus was assessed using data from respondents who completed the second round by providing ratings for at least 80% of the outcomes. Outcomes rated critical (between 7-9) by at least 70% in all three stakeholder groups, and of limited importance (between 1-3) by less than 15% of all participants, in all stakeholder groups, were included in the core outcome set. Outcomes that were not prioritized by any of the stakeholder groups (based on the previous criteria), were excluded, while those that were prioritized by some but not all groups were selected for further evaluation during the consensus meeting.

3.4.2 First consensus meeting: Core Outcome Set Completion

Two consensus meetings were organized as part of this project (April 21st and April 28th, 2021). The Core Outcome Set was finalized during the first virtual meeting, while the second was devoted to the selection of the optimal measurement instrument for each of the core outcomes. To empower patients, who had an active role in both meetings, we offered training about the

research project rationale, aims and methods and their role during the consensus meetings. The active involvement of patient representatives necessitated that the two meetings were moderated by experienced and impartial facilitators (Sara Brookes and Paula Williamson). The facilitators ensured relevant data were presented objectively and in a plain language, and that all participants had the opportunity to share their views and cast a well-informed and independent vote.

During the first consensus meeting, the results of the Delphi survey were presented and the inclusion or exclusion of outcomes that had reached the respective thresholds in the Delphi survey were confirmed. Outcomes with an inconclusive survey result, that were prioritized by at least one, but not all stakeholder groups were discussed in detail. Thorough discussion where both health professionals/ researchers and patients were invited to share their views about the level of importance of each of these outcomes was followed by a poll. Each participant was asked to re-rate the outcomes considering their previous ratings, the Delphi survey results and the preceding discussion. Participants were classified in two groups (a) health professionals or researchers and (b) patients diagnosed with COPD and patient representatives. Only outcomes that were rated as critical by at least 70% of the participants in both groups were added to the core outcome set.

3.5 Core outcome measurement instrument selection

The aim of this component of our study was to select and recommend a single, optimal instrument to measure every core outcome, to ensure consistency and comparability across clinical trials. This was achieved through evidence-informed consensus, during the second consensus meeting of our task force. The methodology followed is summarized in figure S5

A pragmatic methodology was followed for prioritizing measurement instruments. Our aim during this process was to select methodologically sound outcomes, while promoting consistency. For this reason, we first identified instruments that are already in use through our methodological systematic review [5]. In line with our prospectively published protocol,

outcomes that are often evaluated by the same instrument (in >40% of trials evaluating that outcome), this instrument was considered established and was preselected for prioritization, unless important methodological issues were raised by any of the panel members during the second consensus meeting. In case members of the panel had raised such concerns, we had plans to further evaluate instruments used to measure the specific outcome with the aim to develop consensus in a third meeting – however that was not necessary at the end.

3.5.1 Focused literature reviews

For other outcomes, not consistently measured using the same instrument, we conducted focused literature searches of Medline/ PubMed and the COSMIN database. Detailed search strategies and study selection process are summarized in section 6 of the supplement. All searches were updated on April 22nd, 2021. At first, we searched for systematic reviews evaluating the quality and measurement properties of different instruments. In the absence of high-quality methodological systematic reviews, we searched for primary methodological studies formally assessing measurement properties. Alternatively, we looked for previous position or consensus documents or studies of any design that could inform the panel's decision. These literature searches were launched after the first round of the Delphi survey and initially focused on the outcomes which were clearly considered critical by the respondents already from that stage.

3.5.2 Second consensus meeting.

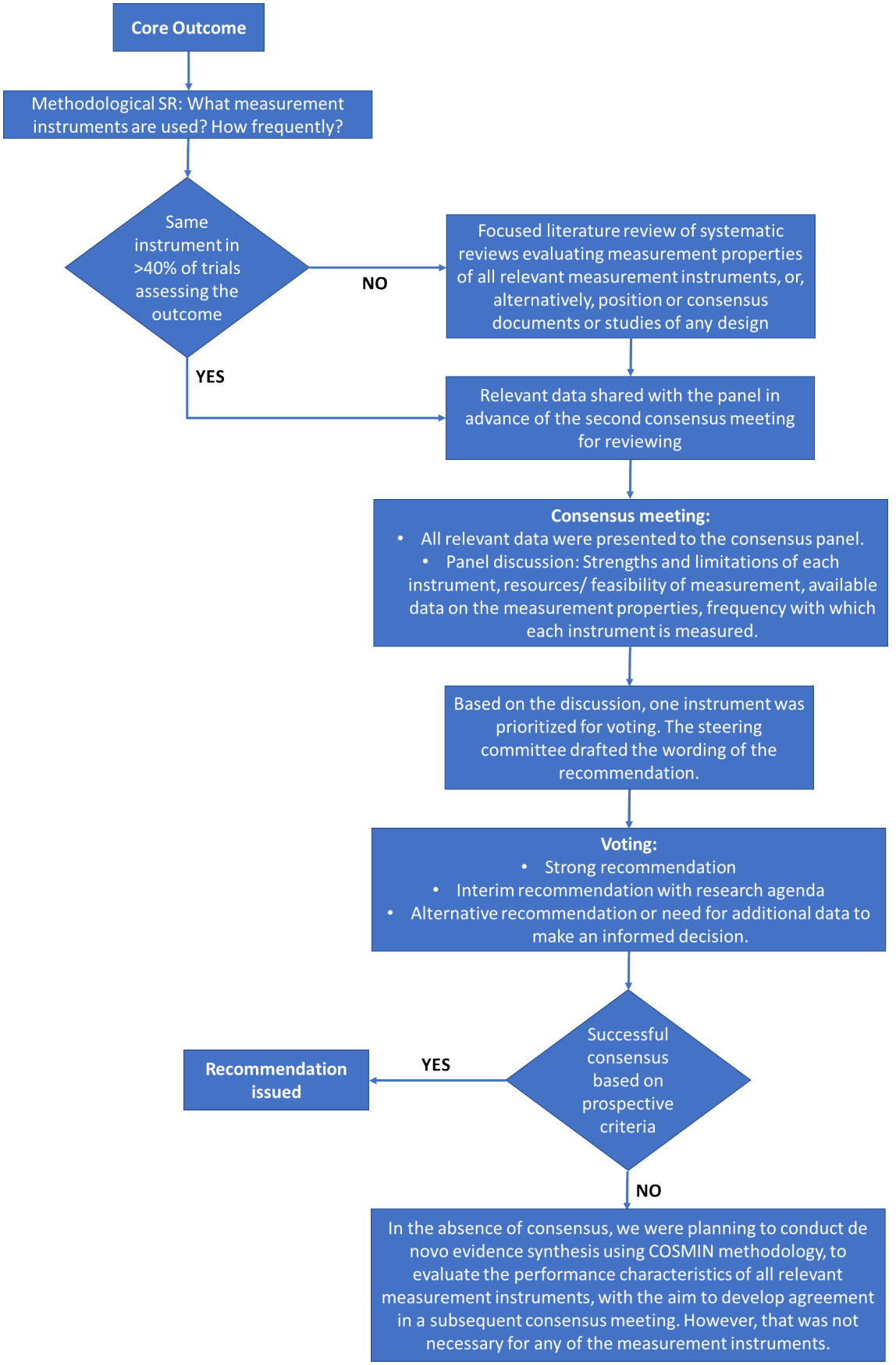
The objective of the second consensus meeting was to select and recommend a single measurement instrument for every core outcome, to ensure consistency and comparability across clinical trials. Each of the outcomes were discussed during the consensus meeting. The panel reviewed available evidence, which was circulated in advance via email (after the first consensus meeting, when the selection of the core outcomes was finalized) and developed consensus on a simple instrument for each outcome after considering (a) the frequency with which each instrument has been used in clinical trials; (b) the time and resources required to use each instrument; and (c) available data on their measurement properties, as described by COSMIN recommendations [16]. After discussion, a single

instrument was selected for every core outcome and participants were asked to vote. Due to the more technical nature of this assignment, only two patients and a representative of the ELF, with previous experience in COPD research, joined the consensus meeting, and therefore the voting was not stratified by stakeholder group. Voting options included: (a) a strong recommendation, (b) an interim recommendation along with research agenda, a research agenda without a recommendation, or (c) for an alternative recommendation or the need for additional data to make an informed decision. A strong recommendation for a specific instrument was issued if at least 70% of the participants voted that option. If less than 70% considered a strong recommendation appropriate but at least 70% voted for the first or second options, then an interim recommendation was issued, along with a recommendation for further research for this core outcome. The prespecified threshold for making a research recommendation without an interim instrument was also 70%; in any other case we were planning on re-voting in a future consensus meeting, after further discussion and data acquisition; that was not necessary as consensus was developed for all core outcomes.

Feedback was sought by all participants of the consensus meeting to explore whether they felt they were offered the opportunity to share their views and that they were able to cast well-informed votes.

Changes from the prospectively registered protocol are summarized and justified in the next section.

Figure S5: Flowchart summarizing the methodology used for selecting core outcome measurement instruments.



3.6 Deviations from the study protocol

3.6.1 Delphi survey stakeholder groups.

We were planning on including a fourth stakeholder group in the Delphi survey, consisting of regulators, policy makers, guideline methodologists or those working in health technology assessment organizations. However, we did not manage to attract adequate responses in order to consider them independently.

This stakeholder group was represented in the consensus meetings.

3.6.2 Change in the threshold for excluding outcomes based on the Delphi survey results.

When interpreting the Delphi survey results, we were planning to exclude outcomes that were considered non-critical by at least 50% of the Delphi survey participants from each stakeholder group. However, due to the coronavirus disease 19 (COVID-19) pandemic, we had to switch our planned face-to-face consensus meeting to two virtual meetings. Conducting virtual multi-stakeholder consensus meetings involving lay participants is challenging and time-consuming. Drawing on the experience amassed by the COMET initiative while facilitating similar, virtual consensus meetings during the pandemic, we decided to further consider during the consensus meetings only outcomes that had been rated as critical by at least one stakeholder group. This approach allowed a more thorough and constructive discussion and more confident consensus decisions for the outcomes that were considered. In parallel, reassurance was offered by our methodologist that based on the initiative's prior experience selection of outcomes that have not been prioritized by any stakeholder groups within the Delphi survey for inclusion in the core outcome set is unlikely.

4 Final results of the Delphi survey

Table S3: Summary of the Delphi survey results. The proportion of participants that considered a particular outcome critical (both rounds).

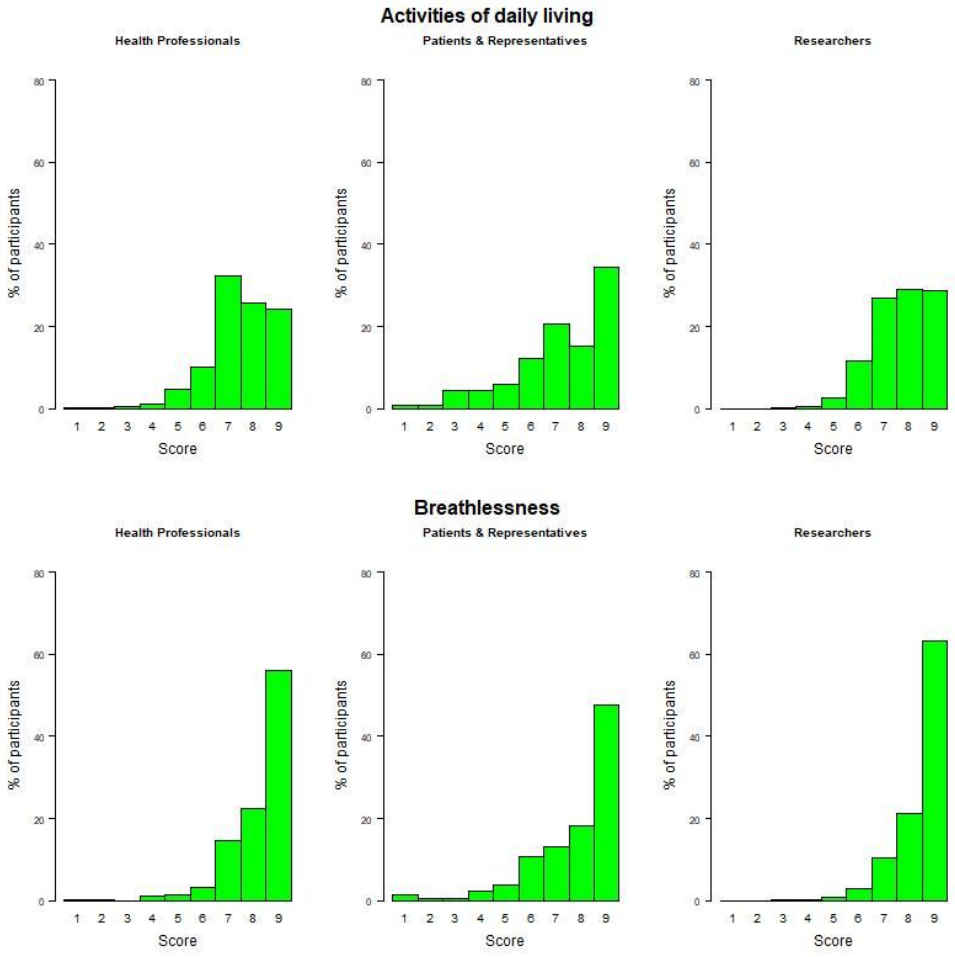
COPD exacerbations outcomes considered		Round 1			Round 2			Final set
Sources of outcomes Methodological SR Qualitative interviews Delphi survey (Round 1)	**Outcomes' selection results (for each round)** Prioritized by all groups Prioritized by 1-2 groups Not prioritized by any group	Patients & Patient representatives	Health professionals	Researchers	Patients & Patient representatives	Health professionals	Researchers	
Death outcomes								
Death from COPD Exacerbation		93.0%	82.3%	93.1%	81.8%	94.5%	96.9%	
Death from any cause		63.6%	64.6%	68.9%	68.5%	74.8%	84.0%	
Clinical and Physiological Outcomes								
Anxiety		29.8%	33.6%	34.7%	35.5%	27.0%	28.3%	
Breathlessness		91.2%	75.0%	84.3%	79.3%	93.3%	94.9%	
Chest discomfort		49.7%	49.8%	40.6%	15.8%	5.8%	8.2%	
Fatigue		47.6%	53.8%	45.9%	54.2%	46.3%	44.7%	
Cough		52.3%	49.4%	53.1%	49.3%	54.3%	53.6%	
Coughing up blood (haemoptysis)		56.7%	62.0%	43.3%	62.1%	58.3%	46.8%	
Production of dark-coloured sputum		60.2%	52.5%	50.2%	56.7%	58.5%	53.6%	
Sputum amount		52.3%	38.3%	40.8%	38.4%	42.0%	35.5%	
Sputum thickness (ease of expectoration)		46.7%	39.3%	36.1%	40.4%	41.8%	29.0%	
Wheeze		52.3%	40.3%	42.0%	39.4%	46.8%	35.2%	
Appetite		21.6%	25.7%	20.5%	24.6%	17.5%	14.0%	
Change in weight		29.8%	29.8%	30.2%	33.5%	25.8%	23.9%	
Respiratory muscle strength					65.5%	58.8%	47.8%	
Low mood/ depression		32.4%	39.2%	39.3%	41.9%	35.5%	40.6%	
Sleep quality		37.1%	51.2%	38.6%	51.7%	38.3%	35.5%	
Early morning symptoms		39.1%	33.9%	34.0%	36.5%	32.0%	25.6%	
Night time symptoms		50.7%	41.5%	42.3%	45.8%	50.3%	41.3%	
Treatment success (or failure)		77.2%	67.5%	74.7%	80.3%	87.8%	89.1%	
Worsening of symptoms after the initial treatment		69.6%	64.0%	66.6%	71.9%	78.5%	77.1%	
Disease progression		80.1%	78.1%	72.5%	83.7%	88.8%	86.7%	
Future exacerbations		80.7%	78.1%	72.5%	75.9%	89.3%	90.4%	
Lung function during and immediately after the exacerbation		56.2%	70.3%	46.5%	71.4%	54.3%	43.0%	
Permanent deterioration in lung function		80.5%	82.6%	67.8%	87.7%	88.5%	82.3%	

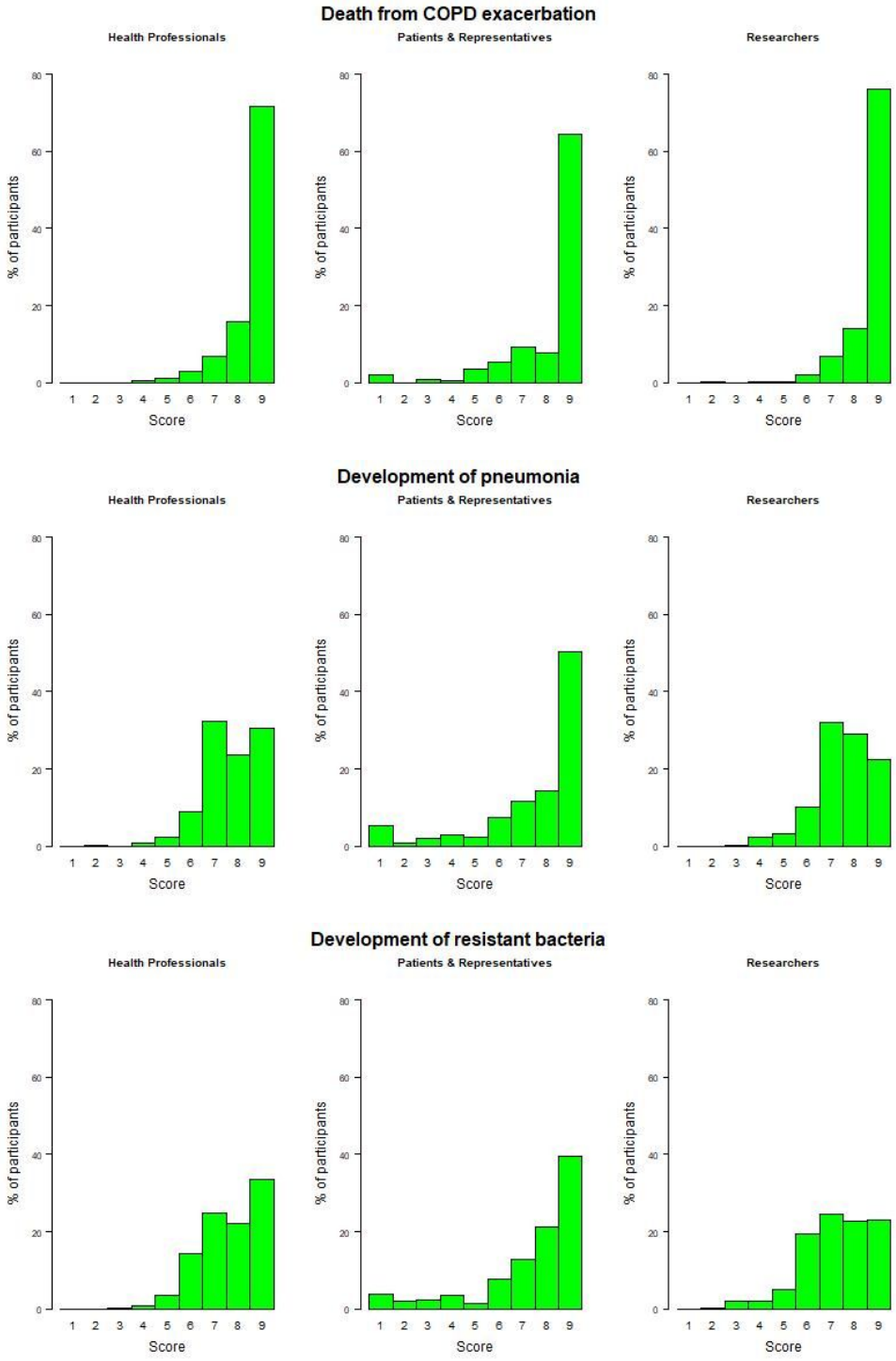
Levels of oxygen and carbon dioxide in the blood (arterial blood gases)	71.8%	70.6%	64.6%	76.4%	80.3%	75.4%	
Development of pneumonia	78.0%	73.3%	70.4%	76.4%	86.8%	83.6%	
Development of resistant bacteria	73.1%	71.4%	61.5%	73.4%	80.8%	70.6%	
Damage of lung cells and lung tissue	64.7%	78.2%	51.9%	81.3%	71.5%	57.3%	
Infection by bacteria (bugs) or viruses	64.2%	69.8%	57.1%	72.4%	68.0%	64.8%	
Inflammation in the lungs/airways	59.9%	70.2%	47.0%	73.4%	61.5%	49.1%	
Adverse event outcomes							
Adverse events of treatments	56.9%	58.4%	61.4%	60.6%	58.3%	65.9%	
Serious adverse events from treatments	84.1%	75.0%	89.0%	76.8%	89.5%	93.5%	
Development and/or progression of other diseases (e.g. heart attack)				67.5%	69.5%	69.6%	
Resources use outcomes							
Need for hospital admission for the presenting exacerbation	76.4%	56.0%	85.2%	69.0%	84.6%	90.8%	
Length of hospital stay for the exacerbation	57.7%	47.0%	64.3%	45.3%	62.3%	68.3%	
Future hospital admissions	63.4%	47.8%	69.6%	52.2%	70.5%	76.5%	
Need for non-invasive ventilation (NIV) use for the exacerbation	74.9%	62.6%	67.9%	64.0%	83.5%	81.9%	
Length of non-invasive ventilation (NIV) use for the exacerbation	58.6%	60.0%	52.8%	58.1%	60.25%	57.0%	
Need for admission to the intensive care unit for the exacerbation	78.4%	71.1%	72.7%	71.9%	86.8%	88.7%	
Length of stay in the intensive care unit for the exacerbation	64.9%	65.2%	59.8%	63.1%	72.8%	71.0%	
Need for additional medications to achieve symptoms control	59.2%	61.2%	53.8%	64.5%	59.5%	57.3%	
Need for long-term administration of supplemental oxygen after the exacerbation				58.6%	62.8%	66.9%	
Need for long-term use of non-invasive ventilation (NIV) after the exacerbation				55.7%	69.5%	65.5%	
Life impact outcomes							
Ability to exercise	53.0%	53.4%	57.2%	57.6%	51.0%	60.4%	
Physical strength	42.8%	47.6%	39.8%	48.8%	38.3%	35.5%	
Walking distance	64.9%	56.4%	64.0%	57.6%	67.3%	68.3%	
Activities of daily living	72.6%	61.8%	73.7%	70.4%	82.5%	84.6%	
Health related quality of life	75.0%	69.6%	79.3%	75.4%	82.5%	87.7%	
Social engagement/ isolation	50.9%	49.4%	47.7%	54.2%	50.5%	50.5%	
Treatment adherence	76.3%	64.2%	73.9%	72.4%	83.8%	84.6%	
Impact of family members and caregivers				56.7%	50.3%	47.4%	
Impact on sexual function				36.0%	36.3%	37.5%	

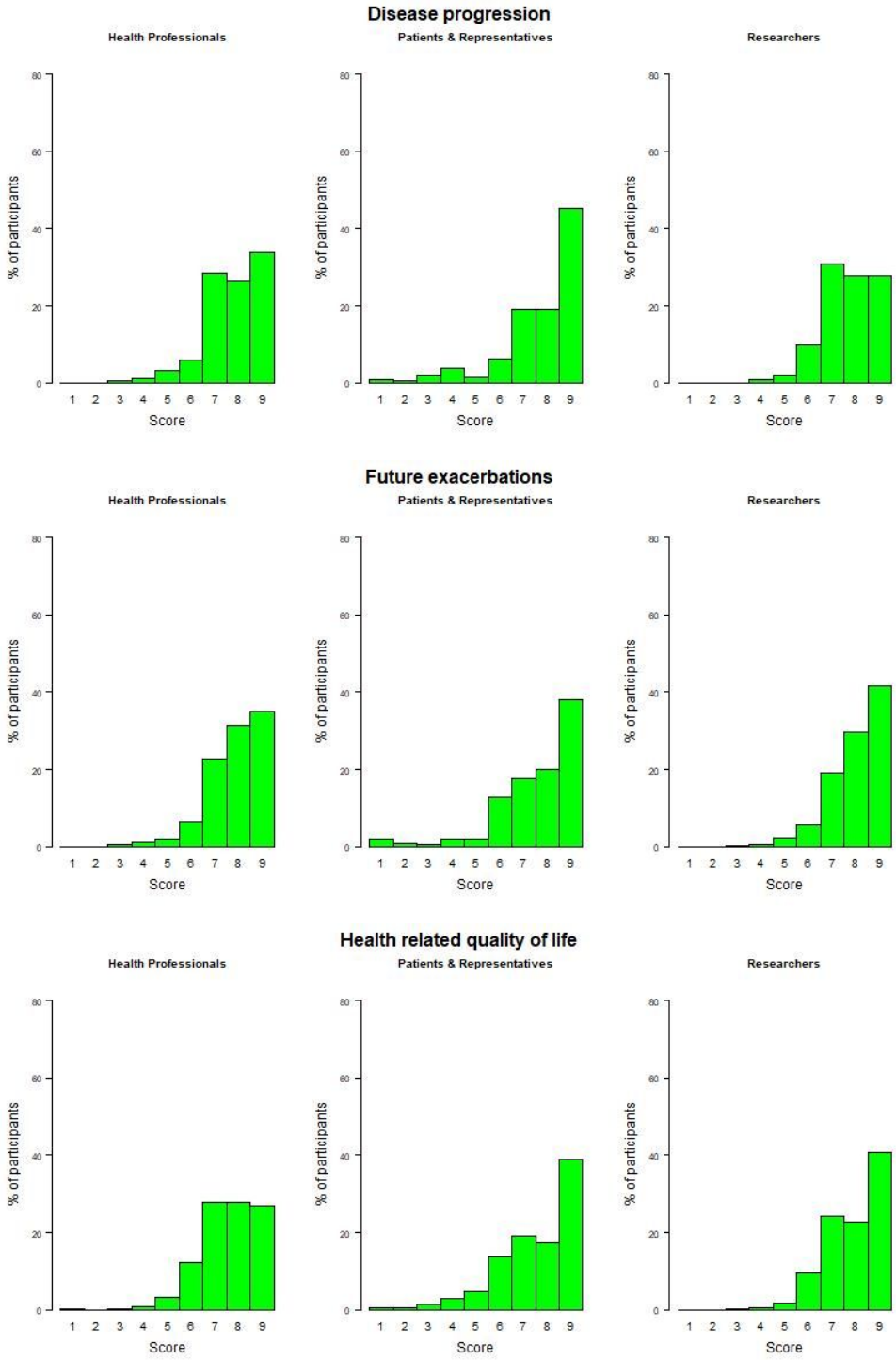
Figure S65. Detailed results of the second round of the Delphi survey. (53 panels)

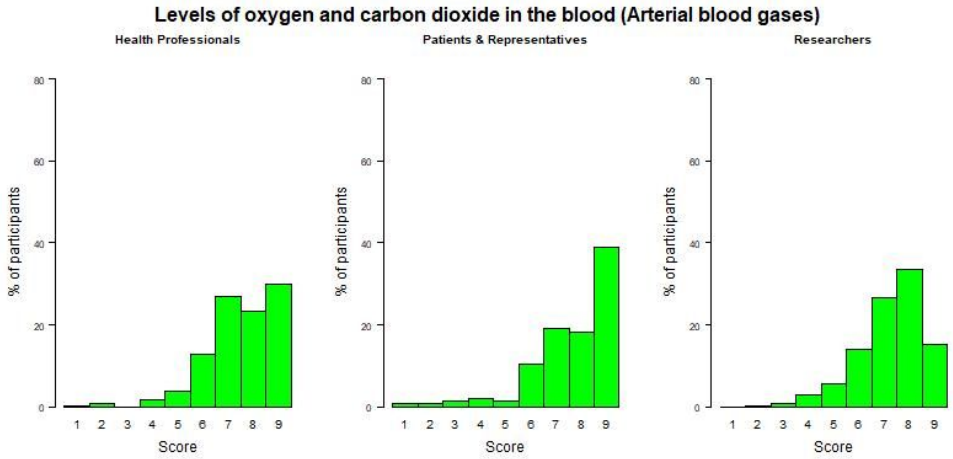
Colour coding:

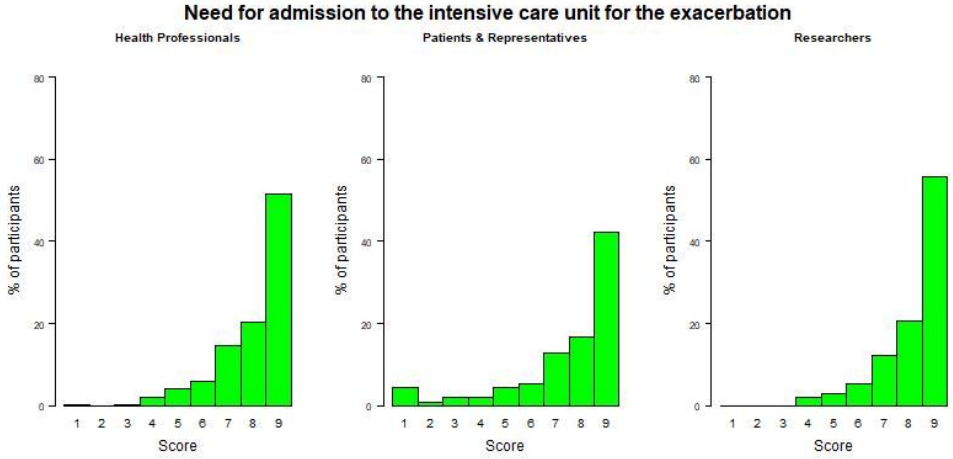
- ✓ Green: The outcome was considered a priority by the respondents group. More specifically, it was rated between 7-9 (critical) by ≥70% and between 1-3 (of limited importance) by ≤15% of all participants from that stakeholder group.
- ✓ Red: The outcome was considered of limited importance by the respondents group. It was rated between 7-9 (critical) by ≤50% of all participants from that stakeholder group.
- ✓ Orange: The ratings were intermediate and did not fulfil either of the previously described thresholds.

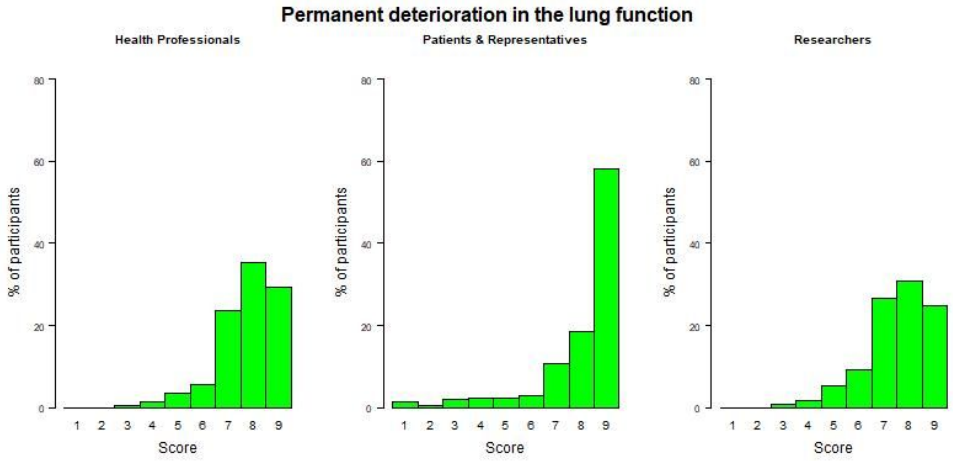


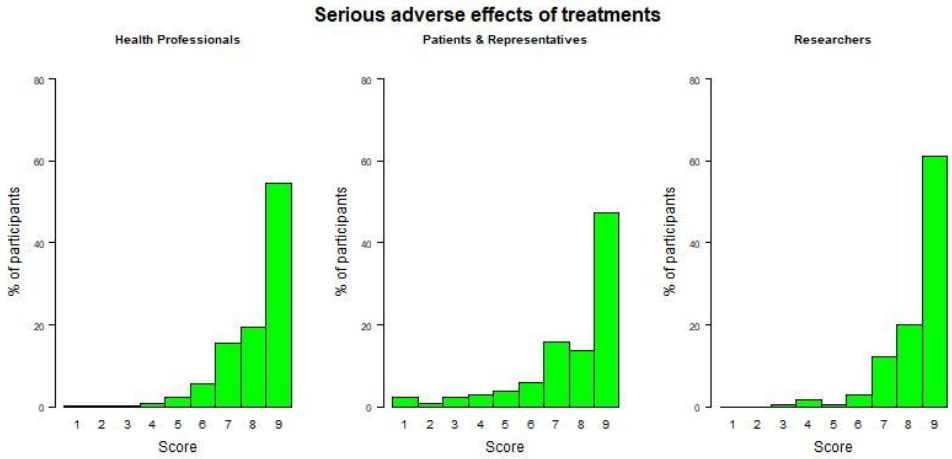

Death from COPD exacerbation
Health Professionals
Patients & Representatives
Researchers
% of participants
Score
Development of pneumonia
Health Professionals
Patients & Representatives
Researchers
% of participants
Score
Development of resistant bacteria
Health Professionals
Patients & Representatives
Researchers
% of participants
Score

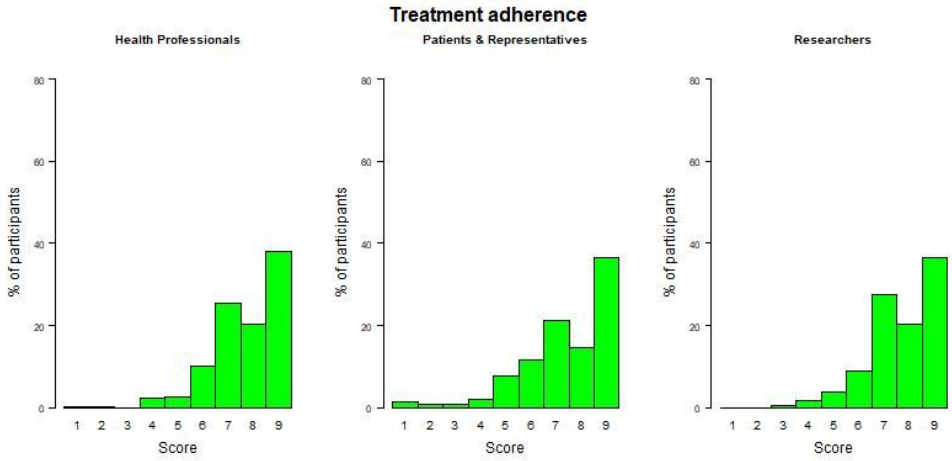

Disease progression
Health Professionals
Patients & Representatives
Researchers
% of participants
Score
Future exacerbations
Health Professionals
Patients & Representatives
Researchers
% of participants
Score
Health related quality of life
Health Professionals
Patients & Representatives
Researchers
% of participants
Score

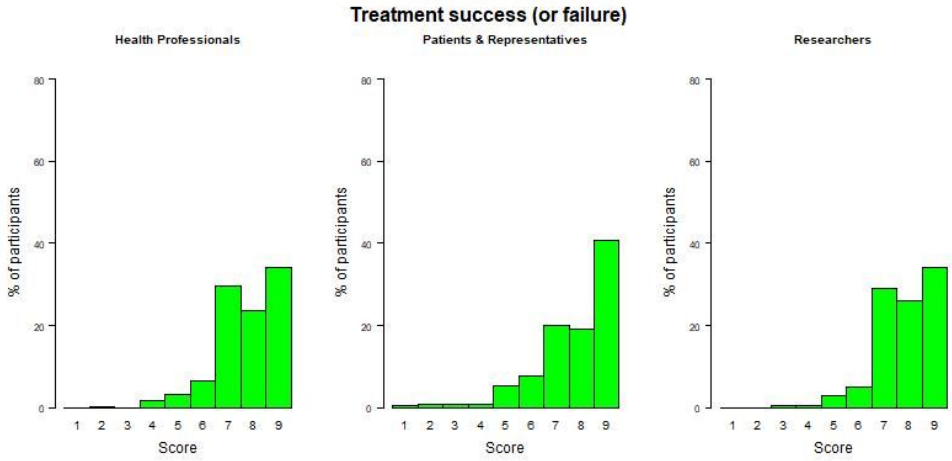

Levels of oxygen and carbon dioxide in the blood (Arterial blood gases)
Health Professionals
Patients & Representatives
Researchers
% of participants
Score

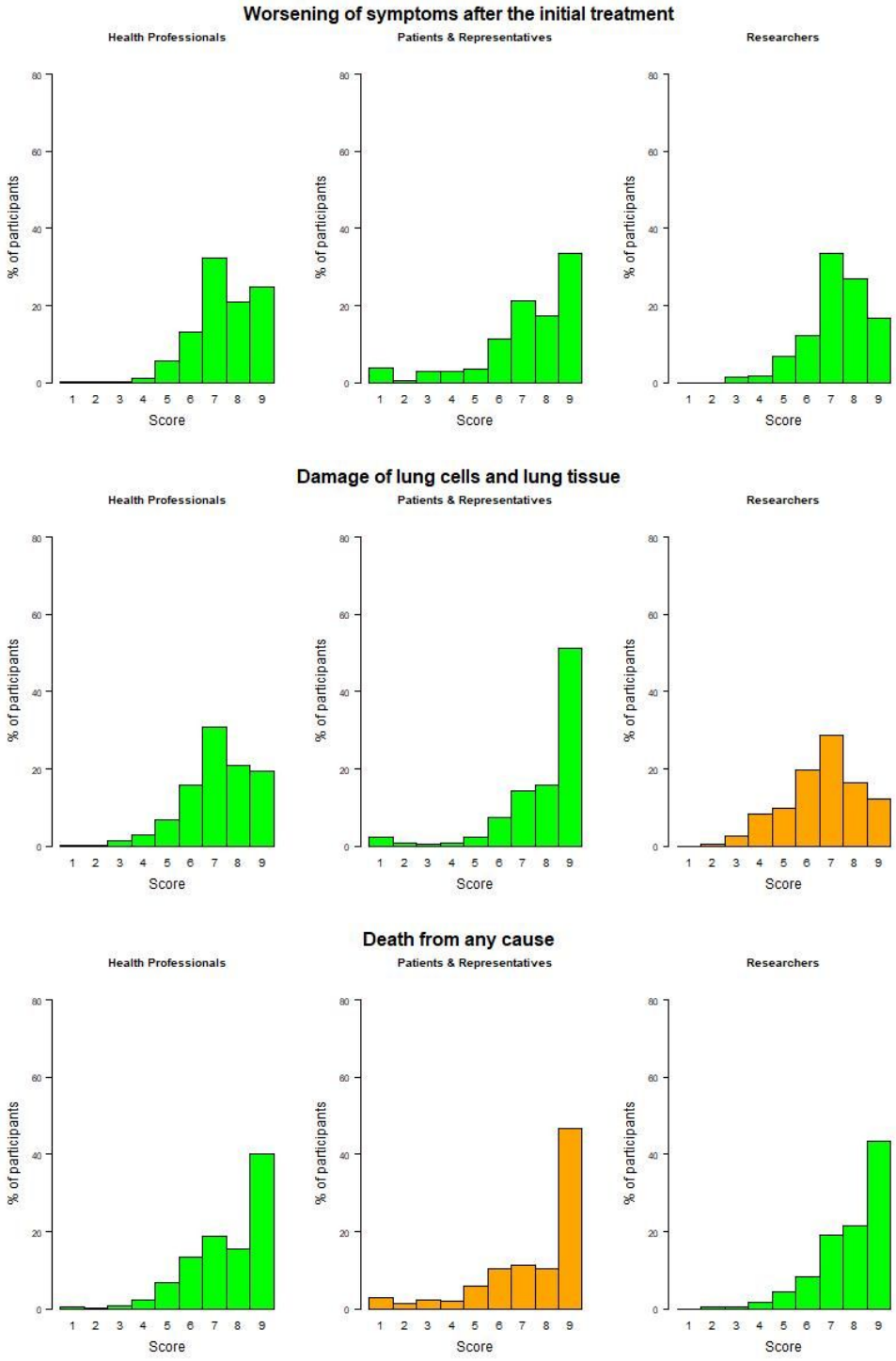

Need for admission to the intensive care unit for the exacerbation
Health Professionals
Patients & Representatives
Researchers
% of participants
Score

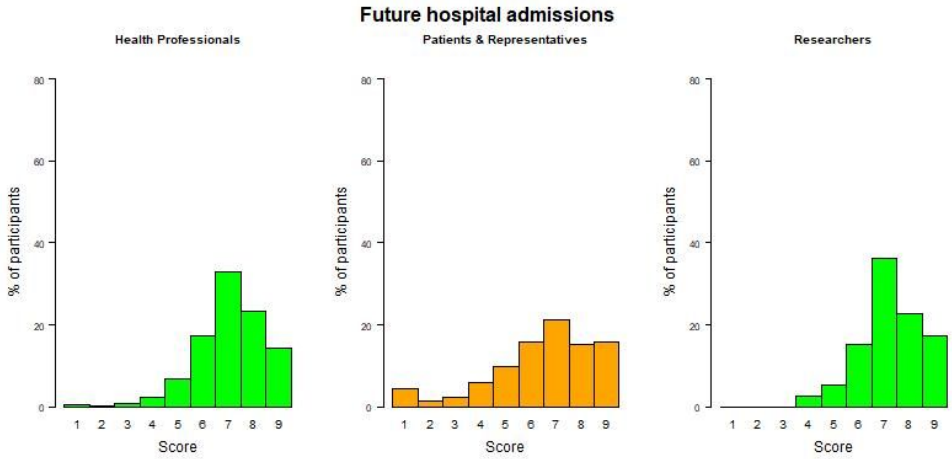

Permanent deterioration in the lung function
Health Professionals
Patients & Representatives
Researchers
% of participants
Score

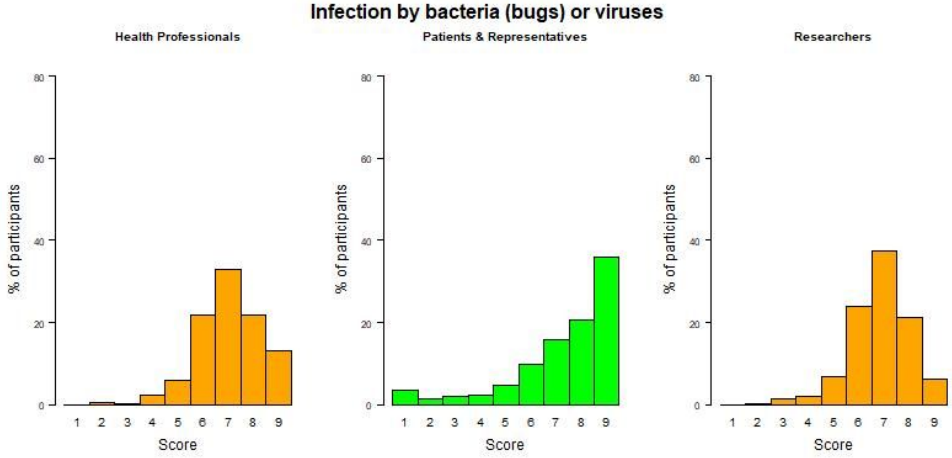

Serious adverse effects of treatments
Health Professionals
Patients & Representatives
Researchers
% of participants
Score

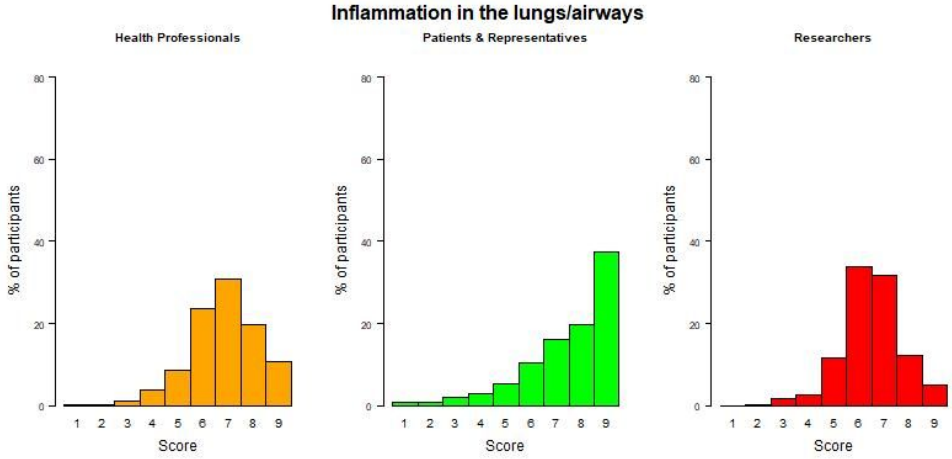

Treatment adherence
Health Professionals
Patients & Representatives
Researchers
% of participants
Score

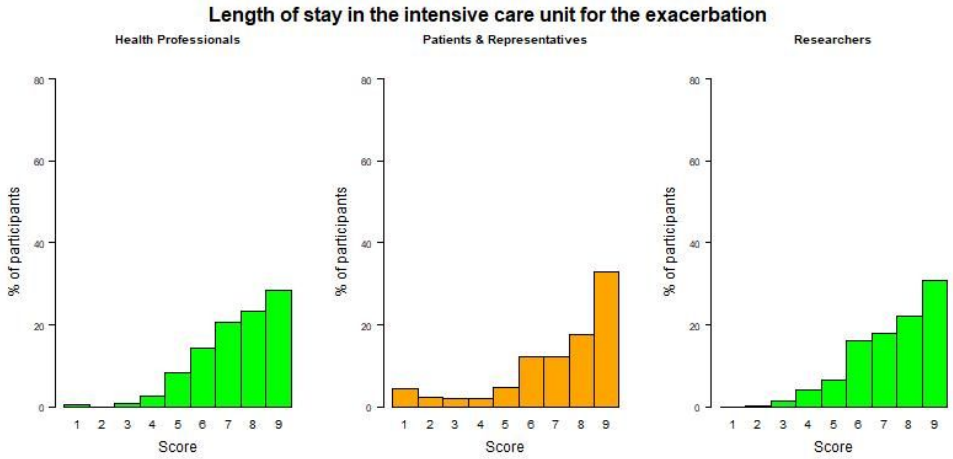

Treatment success (or failure)
Health Professionals
Patients & Representatives
Researchers
% of participants
Score

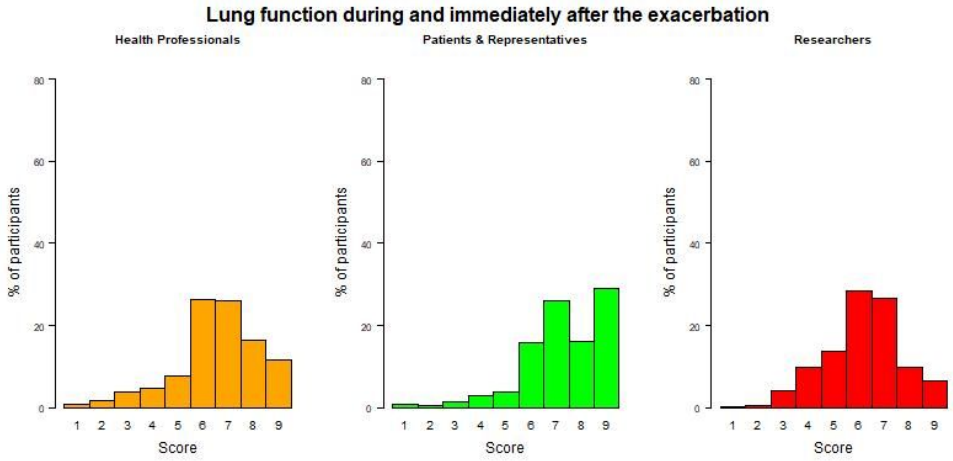

Worsening of symptoms after the initial treatment
Health Professionals
Patients & Representatives
Researchers
% of participants
Score
Damage of lung cells and lung tissue
Health Professionals
Patients & Representatives
Researchers
% of participants
Score
Death from any cause
Health Professionals
Patients & Representatives
Researchers
% of participants
Score

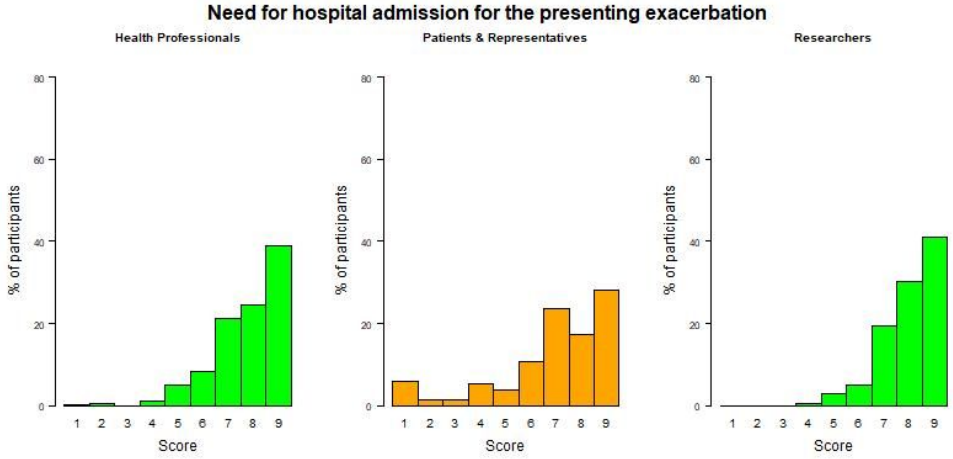

Future hospital admissions
Health Professionals
Patients & Representatives
Researchers
% of participants
Score

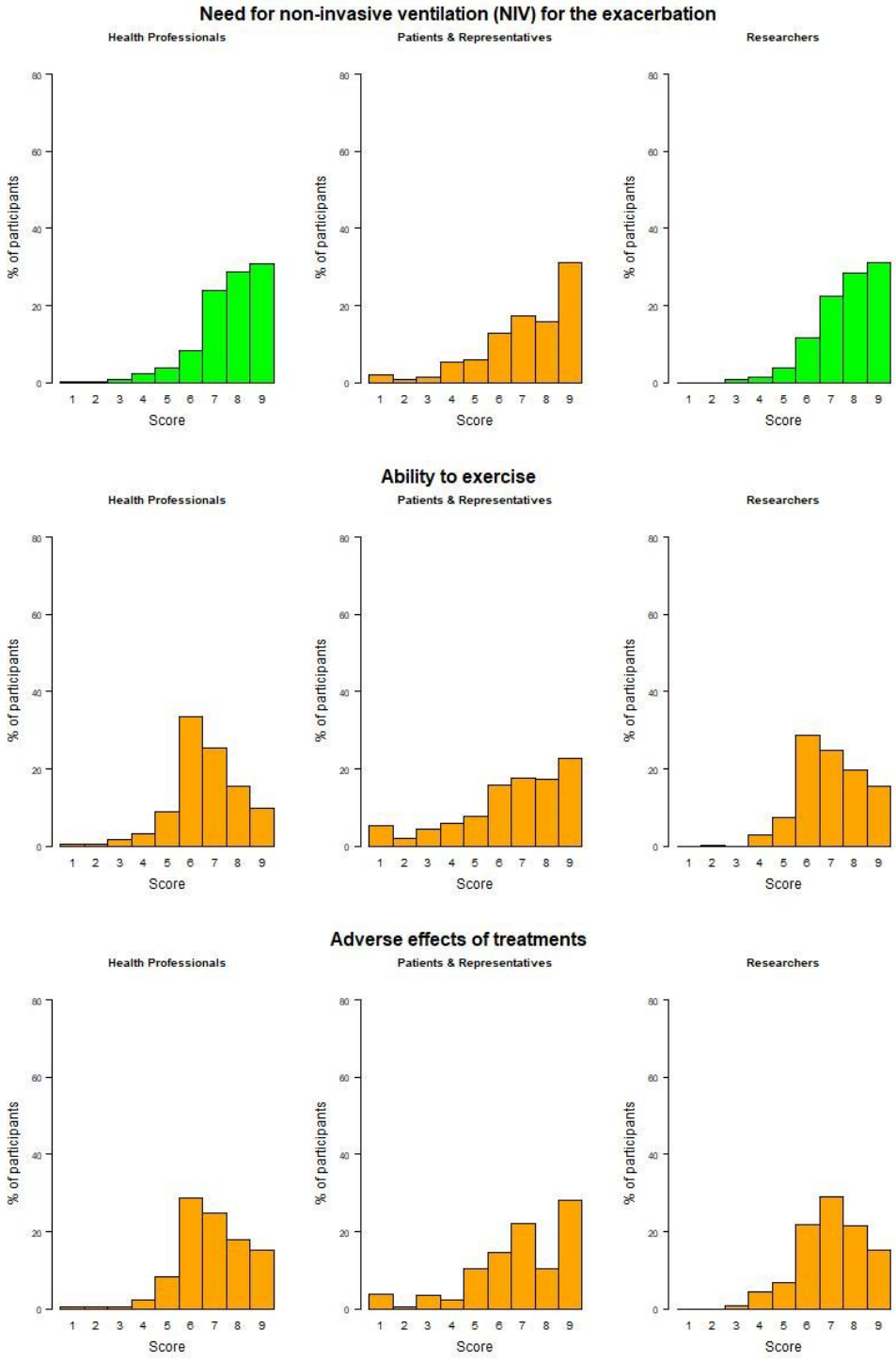

Infection by bacteria (bugs) or viruses
Health Professionals
Patients & Representatives
Researchers
% of participants
Score

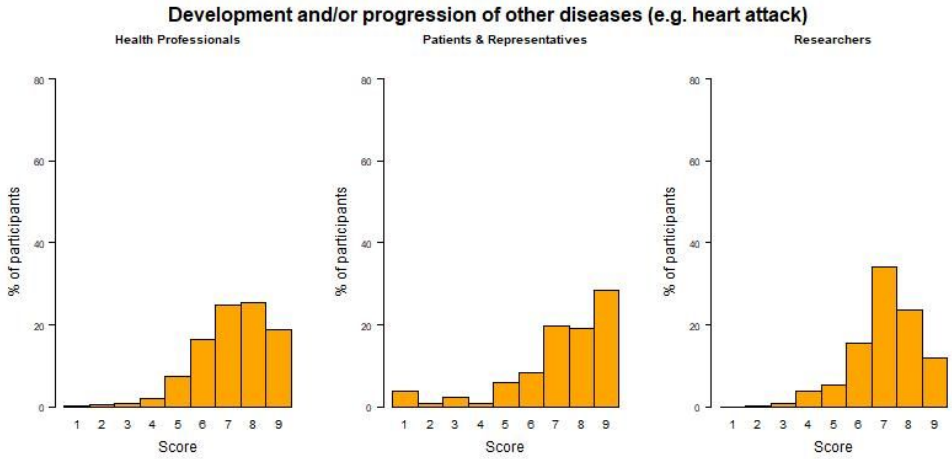

Inflammation in the lungs/airways
Health Professionals
Patients & Representatives
Researchers
% of participants
Score

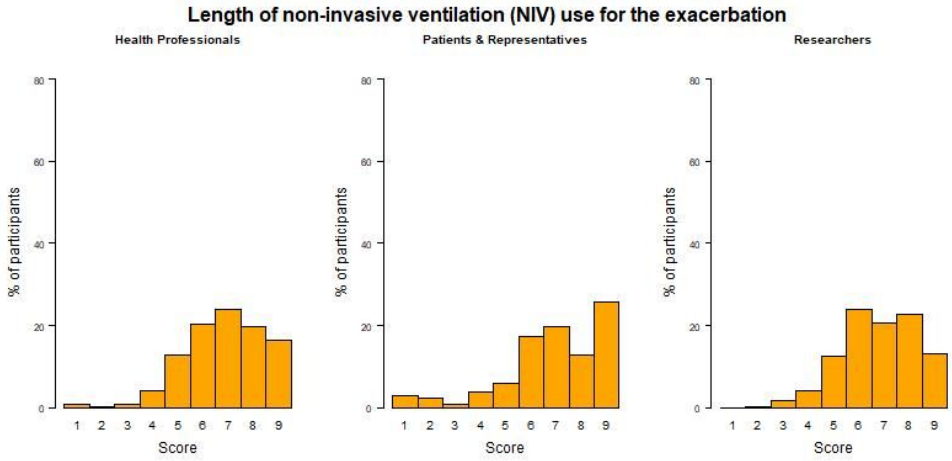

Length of stay in the intensive care unit for the exacerbation
Health Professionals
Patients & Representatives
Researchers
% of participants
Score

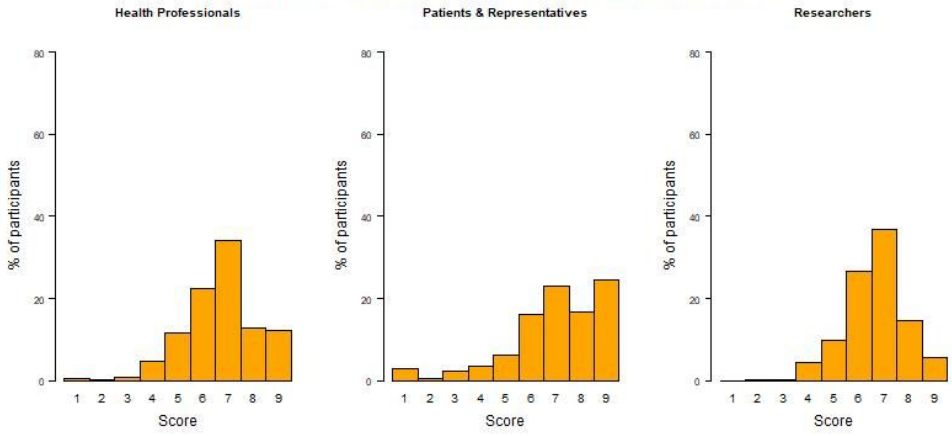

Lung function during and immediately after the exacerbation
Health Professionals
Patients & Representatives
Researchers
% of participants
Score

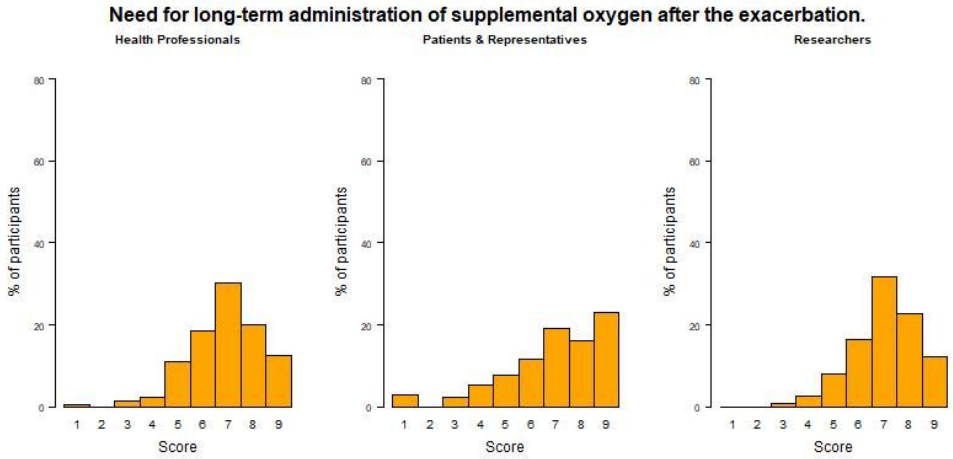

Need for hospital admission for the presenting exacerbation
Health Professionals
Patients & Representatives
Researchers
% of participants
Score

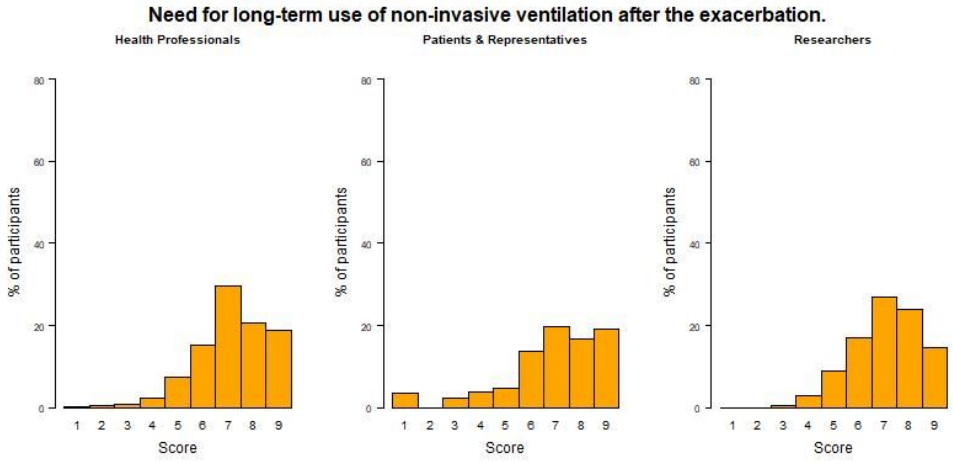

Need for non-invasive ventilation (NIV) for the exacerbation
Health Professionals
Patients & Representatives
Researchers
% of participants
Score
Ability to exercise
Health Professionals
Patients & Representatives
Researchers
% of participants
Score
Adverse effects of treatments
Health Professionals
Patients & Representatives
Researchers
% of participants
Score

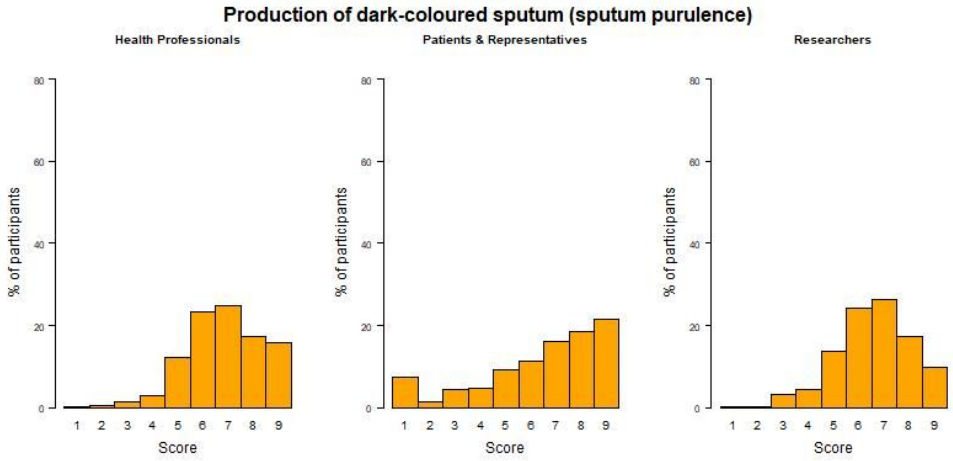

Development and/or progression of other diseases (e.g. heart attack)
Health Professionals
Patients & Representatives
Researchers
% of participants
Score

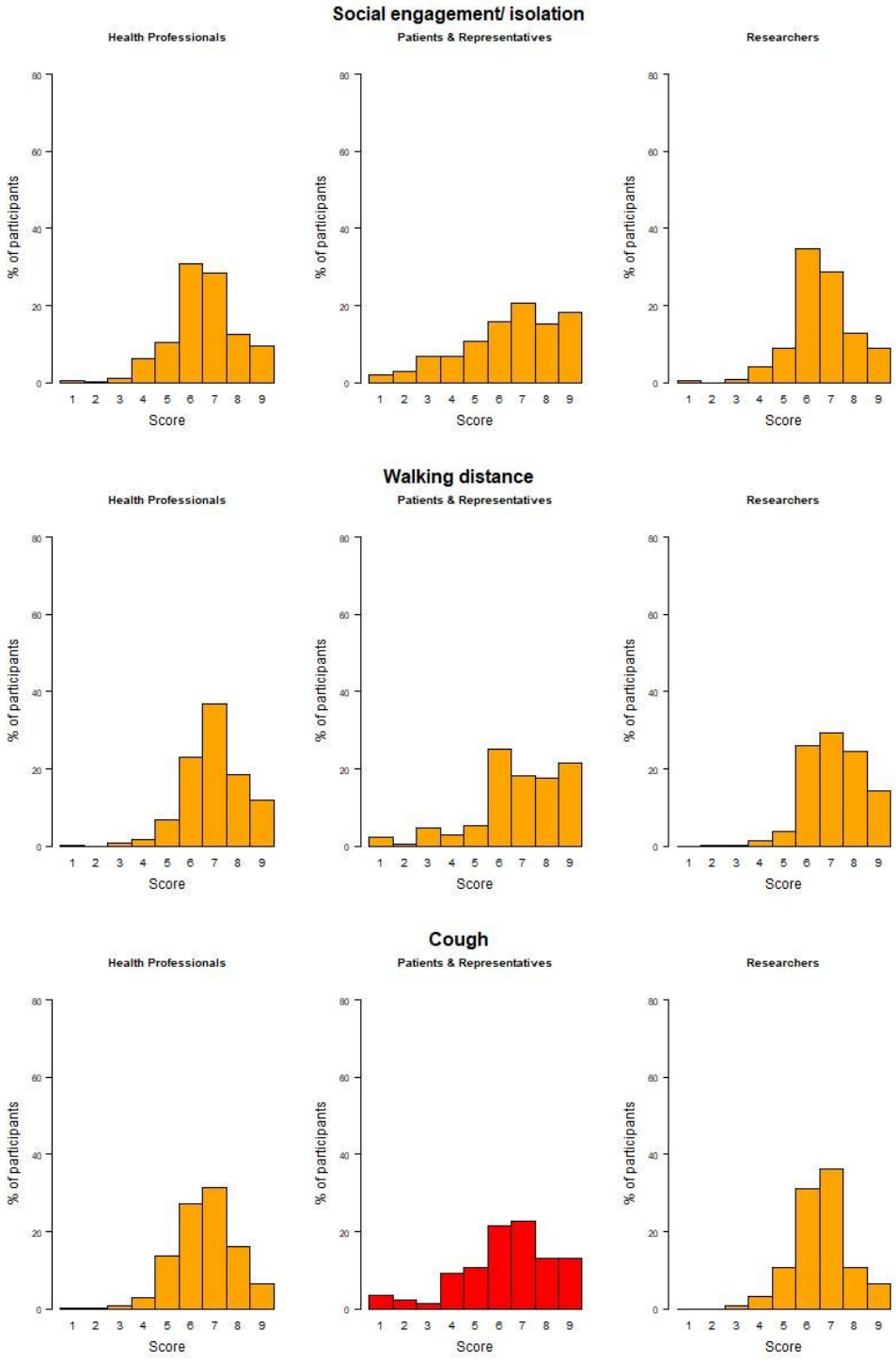

Length of non-invasive ventilation (NIV) use for the exacerbation
Health Professionals
Patients & Representatives
Researchers
% of participants
Score

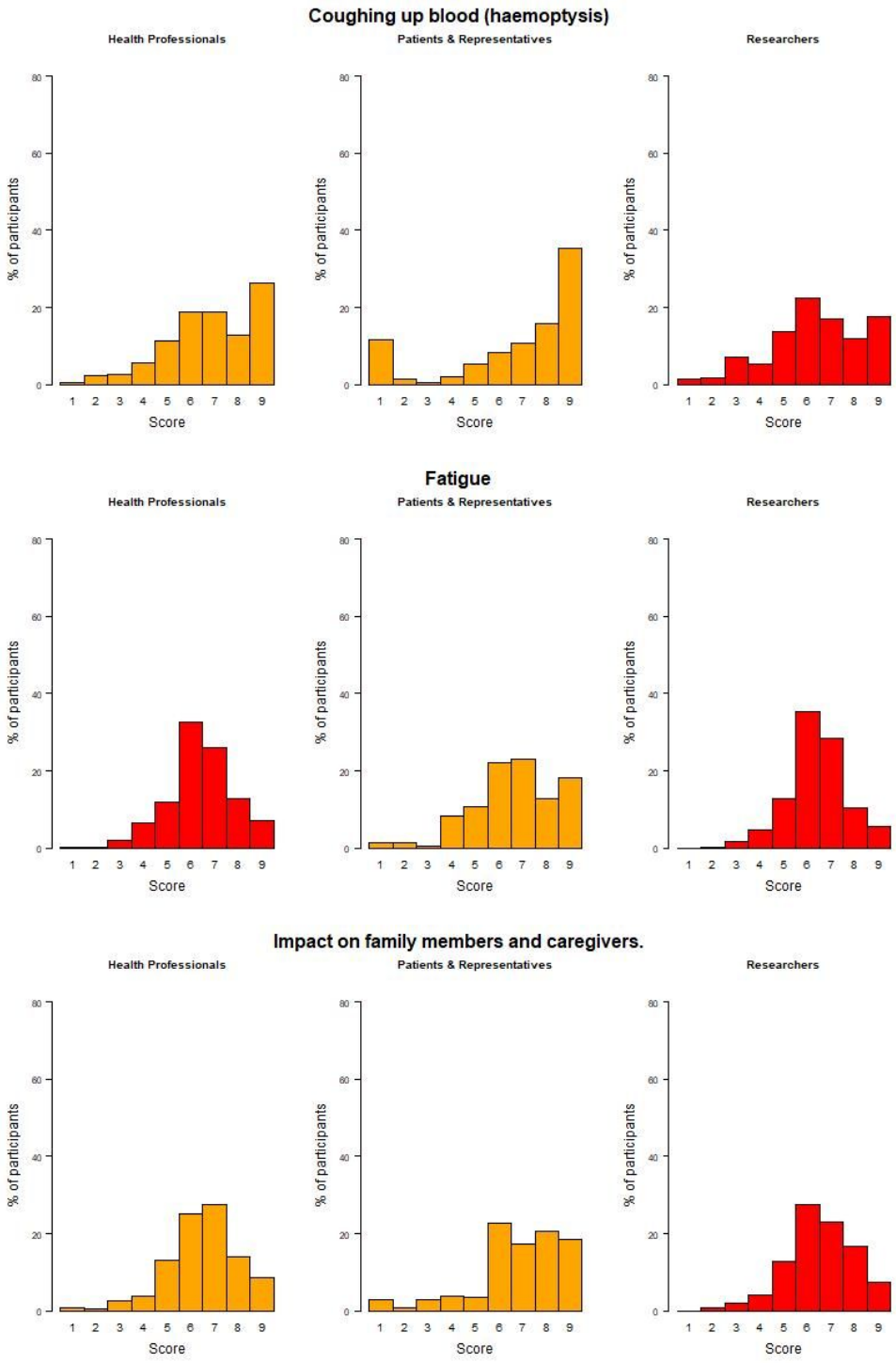

Need for additional medications to achieve symptoms control
Health Professionals
Patients & Representatives
Researchers
% of participants
Score

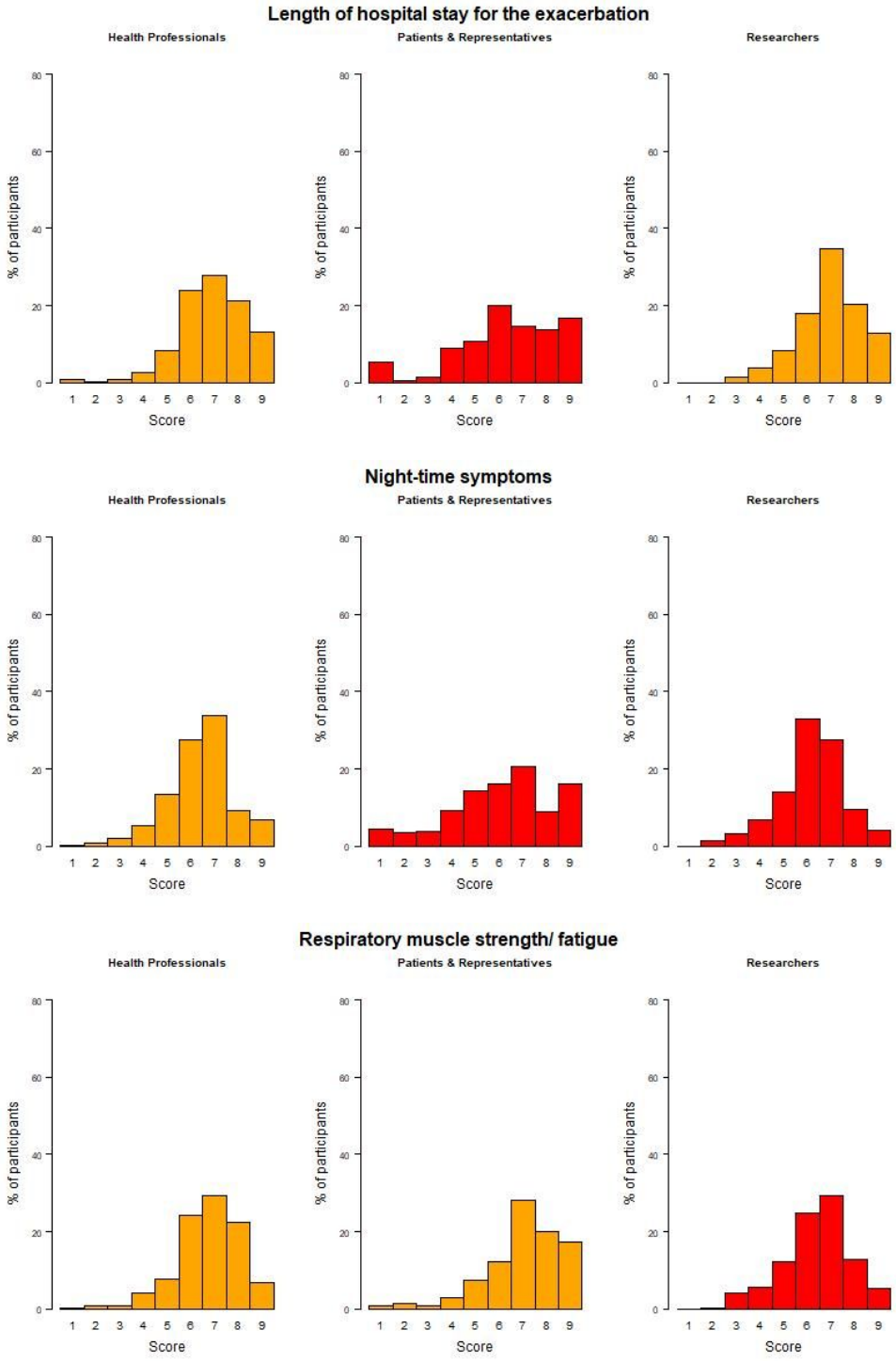

Need for long-term administration of supplemental oxygen after the exacerbation.
Health Professionals
Patients & Representatives
Researchers
% of participants
Score

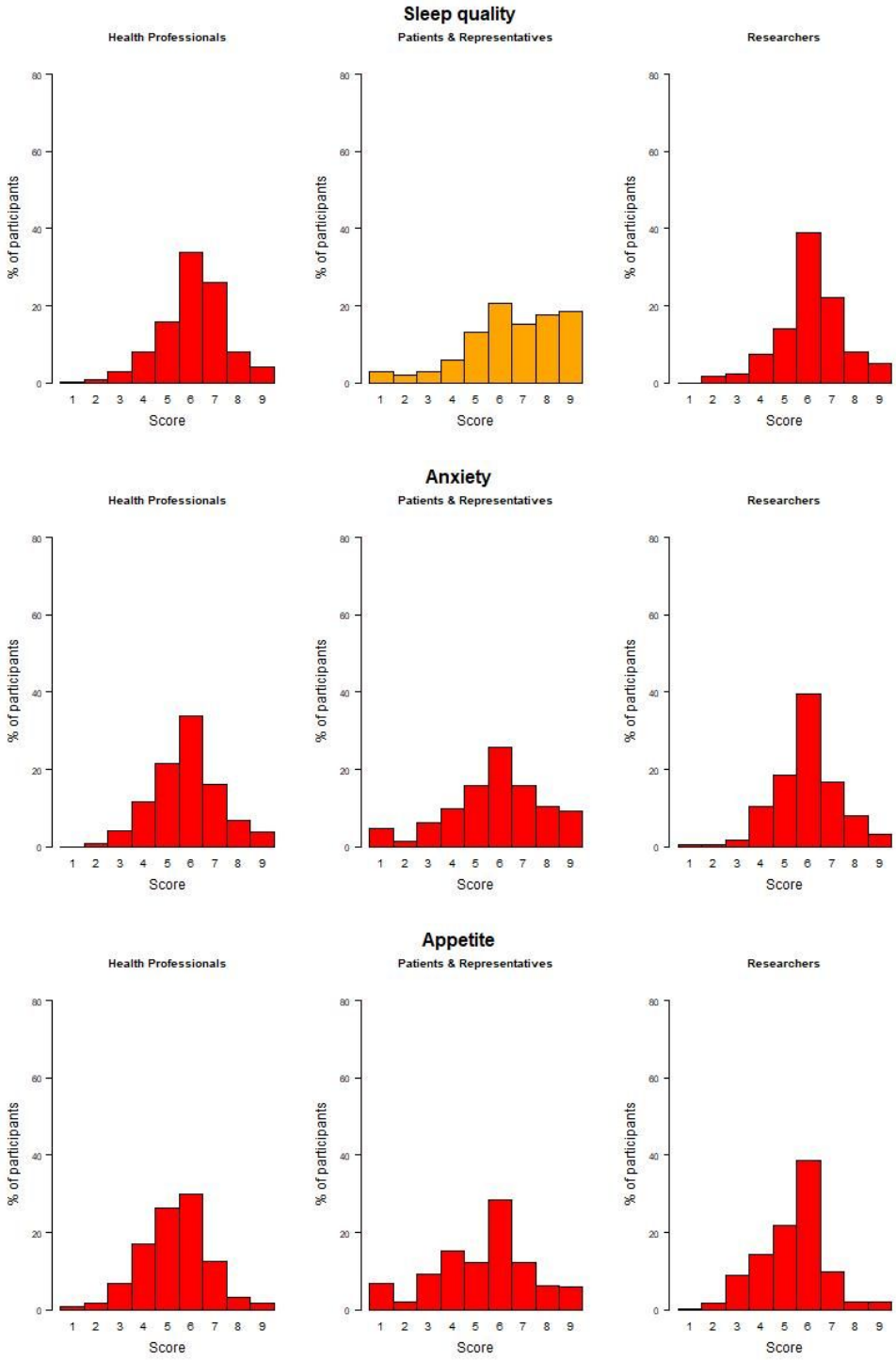

Need for long-term use of non-invasive ventilation after the exacerbation.
Health Professionals
Patients & Representatives
Researchers
% of participants
Score

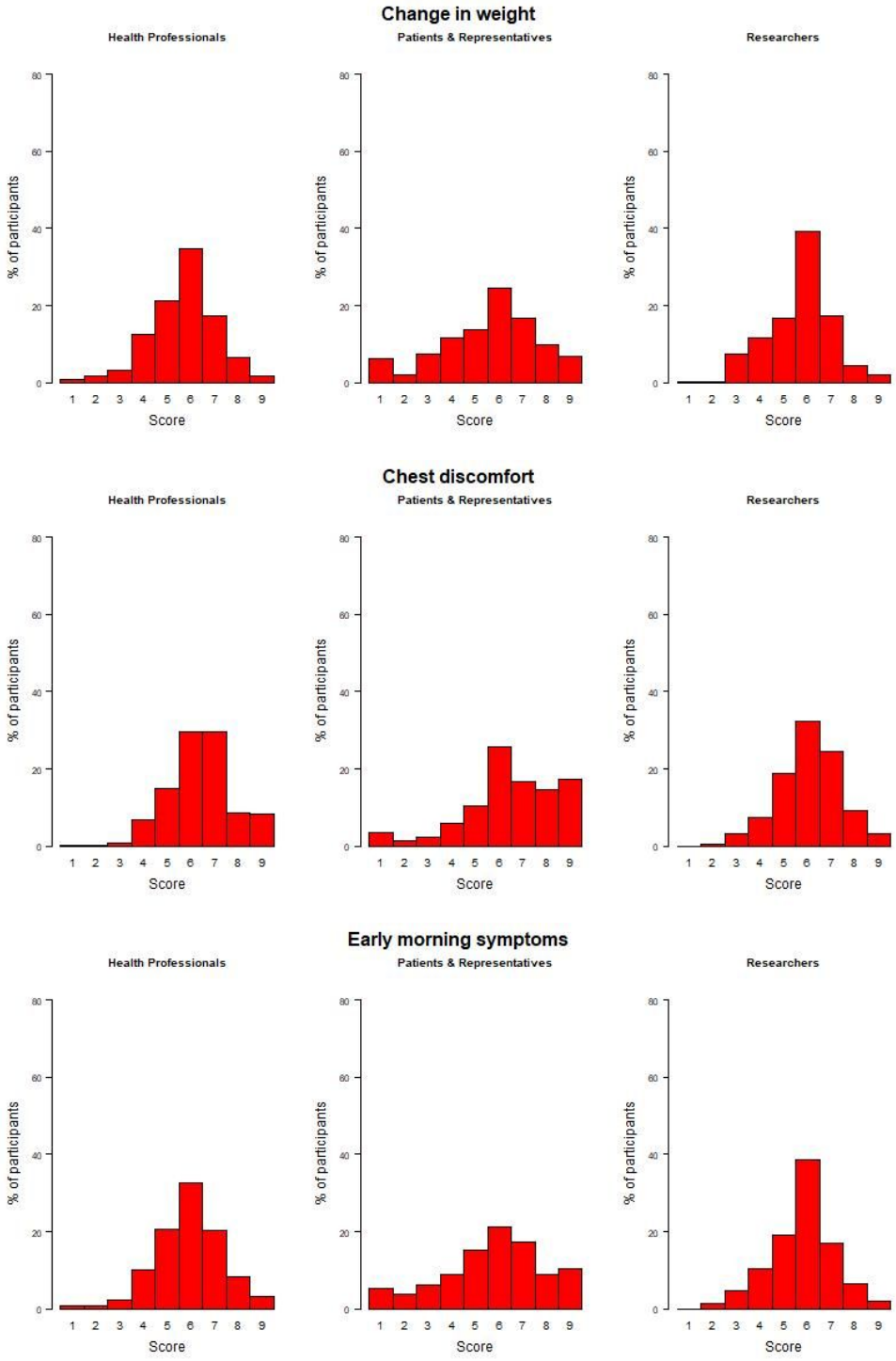

Production of dark-coloured sputum (sputum purulence)
Health Professionals
Patients & Representatives
Researchers
% of participants
Score

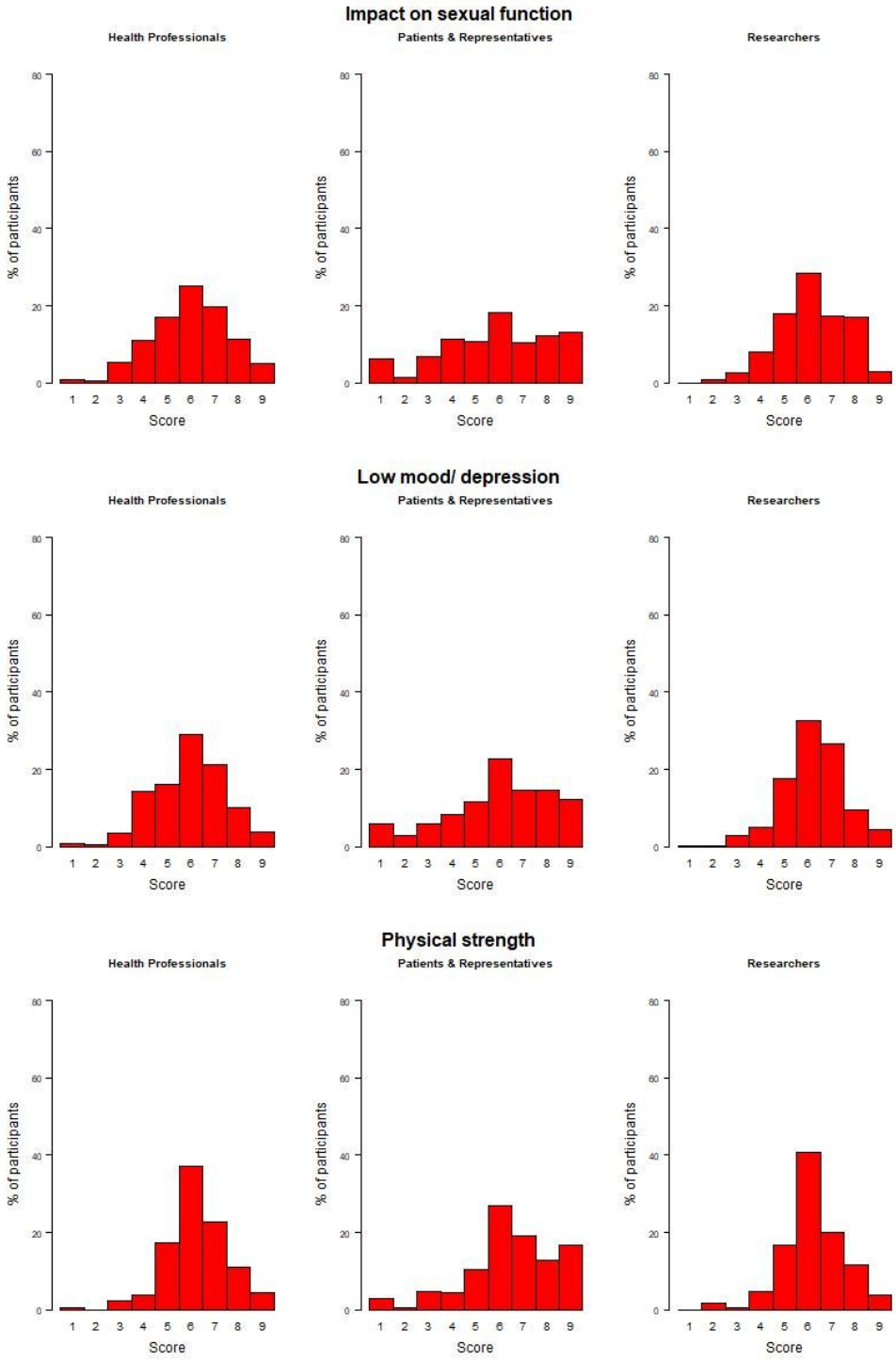

Social engagement/ isolation
Health Professionals
Patients & Representatives
Researchers
% of participants
Score
Walking distance
Health Professionals
Patients & Representatives
Researchers
% of participants
Score
Cough
Health Professionals
Patients & Representatives
Researchers
% of participants
Score

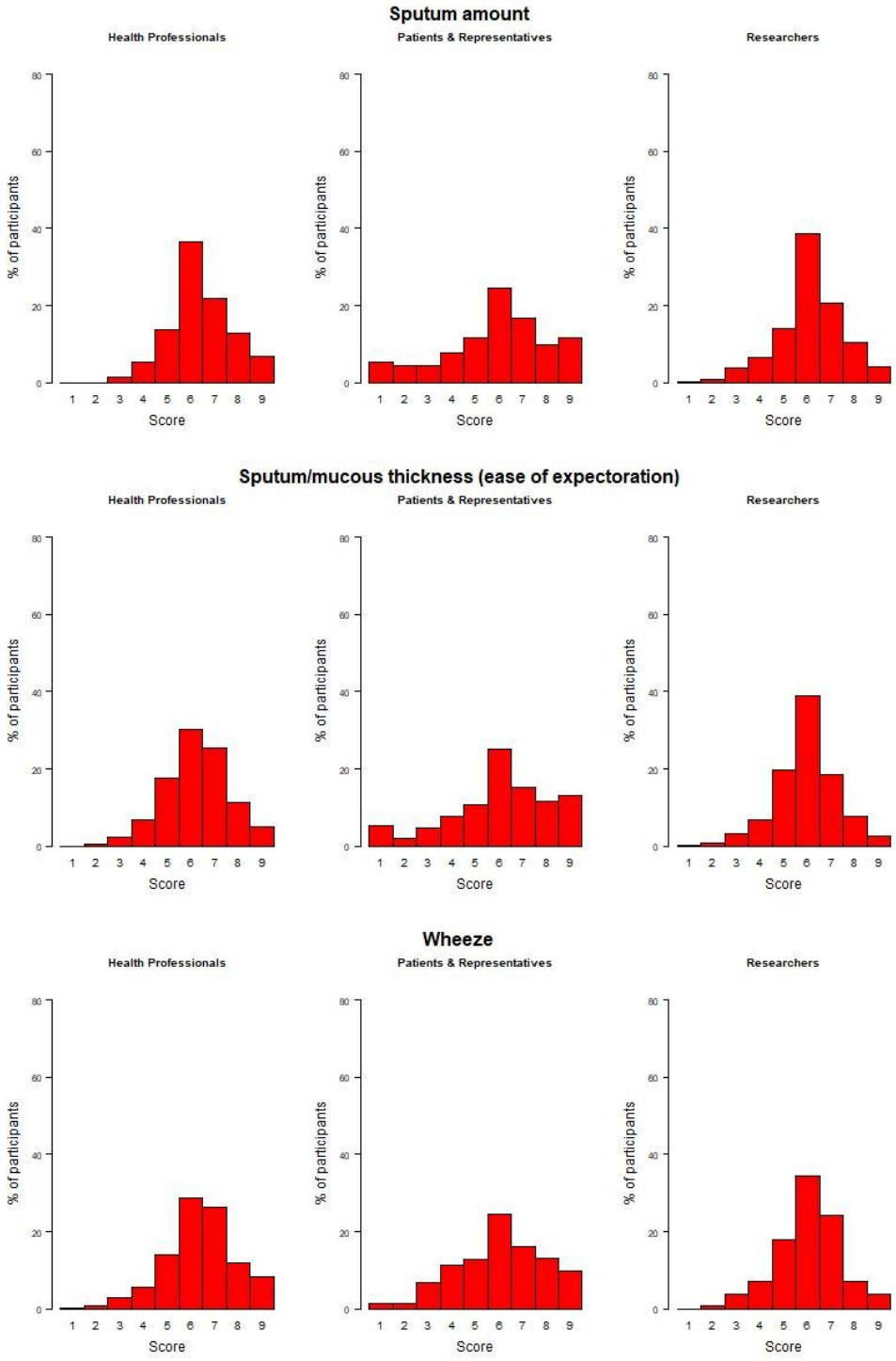

Coughing up blood (haemoptysis)
Health Professionals
Patients & Representatives
Researchers
% of participants
Score
Fatigue
Health Professionals
Patients & Representatives
Researchers
% of participants
Score
Impact on family members and caregivers.
Health Professionals
Patients & Representatives
Researchers
% of participants
Score


Length of hospital stay for the exacerbation
Health Professionals
Patients & Representatives
Researchers
% of participants
Score
Night-time symptoms
Health Professionals
Patients & Representatives
Researchers
% of participants
Score
Respiratory muscle strength/ fatigue
Health Professionals
Patients & Representatives
Researchers
% of participants
Score


Sleep quality
Health Professionals
Patients & Representatives
Researchers
% of participants
Score
Anxiety
Health Professionals
Patients & Representatives
Researchers
% of participants
Score
Appetite
Health Professionals
Patients & Representatives
Researchers
% of participants
Score


Change in weight
Health Professionals
Patients & Representatives
Researchers
% of participants
Score
Chest discomfort
Health Professionals
Patients & Representatives
Researchers
% of participants
Score
Early morning symptoms
Health Professionals
Patients & Representatives
Researchers
% of participants
Score


Impact on sexual function
Health Professionals
Patients & Representatives
Researchers
% of participants
Score
Low mood/ depression
Health Professionals
Patients & Representatives
Researchers
% of participants
Score
Physical strength
Health Professionals
Patients & Representatives
Researchers
% of participants
Score


Sputum amount
Health Professionals
Patients & Representatives
Researchers
% of participants
Score
Sputum/mucous thickness (ease of expectoration)
Health Professionals
Patients & Representatives
Researchers
% of participants
Score
Wheeze
Health Professionals
Patients & Representatives
Researchers
% of participants
Score

5 Measurement instruments selection: Second consensus meeting voting results

After discussion, consensus meeting participants were asked to vote on the most favoured outcome. Voting options included (a) Strong recommendation; (b) Interim recommendation with research agenda; (c) Research agenda without a recommendation; and (d) Other instrument or further data is needed to decide. Based on prespecified criteria, a strong recommendation for a specific instrument was issued if at least 70% of the participants considered a strong recommendation appropriate. Alternatively, if at least 70% of the participants voted for the first or the second options, then an interim recommendation was issued, with research agenda. The prespecified threshold for recommending further research without an interim instrument was also 70%. In any other case, we were planning on re-voting in a future consensus meeting, after further discussion and data acquisition. However, that was not necessary as consensus was developed for all outcomes.

Table S4. Second consensus meeting: Voting results

	Voting responses			
Outcome	**Strong**	**Interim + Research agenda**	**Research agenda only**	**Other instrument or further data needed**
Death from any cause	100%	0%	0%	0%
Death from a COPD exacerbation	55%	39%	6%	0%
Treatment success	0%	88%	6%	6%
Need for hospital admission for the presenting exacerbation	22%	78%	0%	0%
Need for admission to the intensive care unit for the presenting exacerbation	18%	82%	0%	0%
Levels of oxygen and carbon dioxide in the blood (arterial blood gases)	47%	47%	0%	6%

Breathlessness	28%	66%	6%	0%
Health related quality of life	65%	35%	0%	0%
Activities of daily living	0%	76%	0%	24%
Worsening of symptoms after the initial treatment	0%	100%	0%	0%
Disease progression	7%	80%	13%	0%
Future exacerbations	50%	44%	6%	0%
Future hospital admissions	53%	33%	7%	7%
Serious adverse events from treatments	86%	7%	7%	0%
Development of resistant bacteria	64%	22%	14%	0%
Development of pneumonia	71%	29%	0%	0%
Treatment adherence	100%	0%	0%	0%

6 Selection of outcome measurement instruments: Evidence

The recommended outcome measurement instruments and relevant research recommendations are summarized in table 5 and appendix 7, respectively. This section describes the additional data considered by the panel and the main discussion points from the second consensus meeting.

6.1 Death from any cause.

Data from the methodological systematic review [5]: Mortality was evaluated in 101 (82%) of all included RCTs. 100/101 studies evaluated number of deaths in each treatment group during a specific follow-up period, or during hospital or ICU stay.

Literature review: Not performed since this outcome is evaluated consistently.

Panel discussion summary: Death from any cause is the most frequently evaluated mortality outcomes in clinical trials and it is evaluated consistently. The panel agreed to adopt this approach.

Recommendation: Death from any cause should be measured as the number of deceased patients in each treatment group by a specific timepoint (**Strong Recommendation**).

6.2 Death from COPD exacerbation

Data from the methodological systematic review [5]: Only one trial evaluated death from COPD exacerbation as an outcome. The methodology used to determine the cause of death and whether a death was caused by an exacerbation was not described.

Literature review:

Search terms describing the outcome	*COPD and Exacerbations terms – see figure S1* **AND**

	((Cause of death [MH]) or (death [ti]) or (mortality [ti])) **AND** ((Treatment Outcome [MH]) or (Outcome Assessment, Health Care [MH]) or (instrument* [tiab]) or (outcome* [tiab]) or (endpoint* [tiab]) or (adjudic* [tiab]))
Number of titles screened	232
Relevant studies	Methodological systematic reviews: 0 Other references: 3 [17-19]

We did not identify methodological studies evaluating outcome measurement instruments for assessing this outcome in COPD exacerbation trials. Three studies described the rules used for determining the cause of death in TORCH and UPLIFT, two clinical trials evaluating the management of stable COPD [17-19]. Both adjudication committees described that if the final illness was precipitated by a recent COPD exacerbation, then the final cause of death should be considered COPD exacerbation, regardless of the subsequent fatal events, such as pneumonia, sepsis, respiratory, renal, or multi-organ failure, myocardial infarction. Both adjudication committees also highlighted inconsistency between the cause of death described in the death certificate and issued by the adjudication committee.

Panel discussion summary: Death from COPD exacerbation is rarely evaluated in exacerbation trials. COPD exacerbations are often complicated by events such as ventricular arrhythmia, massive pulmonary embolism, acute myocardial infarction, or pneumonia [20]. As a result, the determination of the cause of death during an exacerbation is complex and often inconsistent across different centres and countries. The panel agreed that if a death is caused by an immediate complication of the exacerbation, then the exacerbation should be considered the cause of death. Given the inconsistencies observed in the determination of the cause of death, the panel agreed that ideally, cause of death should be confirmed by a well-informed and blinded adjudication committee. However, such committees are resource

intensive and may not always be feasible. For this reason, a pragmatic approach based on the documented primary cause registered in the death certificate was adopted by the panel.

Recommendation: Consider the immediate cause of death as documented in the death summary. In cases of death due to an immediate complication of an exacerbation, such as a ventricular arrhythmia, massive pulmonary embolism, or myocardial infarction, the exacerbation should be considered the cause of death.

Ideally, cause of death will need to be confirmed by a blinded adjudication committee. However, this may not always be feasible. (**Interim Recommendation with research agenda**).

6.3 Treatment success

Data from the methodological systematic review [5]: Treatment success or treatment failure was evaluated in 77 (63%) of the trials included in our systematic review. More specifically, 21 (17%) studies reported data on both treatment success and failure rates, while 27 (22%) and 29 (24%) studies only reported on treatment failure, or treatment success, respectively. The instruments used to evaluate this outcome varied significantly. In the absence of existing methodological study to inform our decision-making process (see next section), we conducted a meta-epidemiological systematic review. The methods of this systematic review were prospectively registered (PROSPERO ID: CRD42020222287) and the results will be reported separately. In brief, using the search strategies that were employed in our original methodological systematic review, we searched PubMed/ Medline and the Cochrane Airways Trial Register on November 12th, 2020. In this meta-epidemiological study we explored:

(i) The instruments used to measure treatment success/ failure or cure and how frequently each instrument is measured.

(ii) Which is the most sensitive instrument? We assessed the magnitude of treatment effect observed in studies compared the addition of an active intervention versus placebo or no intervention, stratified by the instruments used to evaluate the outcome of interest.

(iii) Which is the optimal timepoint? We assessed the magnitude of treatment effect observed across different timepoints of evaluation of treatment success in studies comparing the addition of an active intervention versus placebo or no intervention.

We identified a total of 176 ongoing or completed RCTs evaluating the management of COPD exacerbations, of which 56 (31.8%) assessed the overall outcome of the index exacerbation (treatment success or treatment failure). For the purposes of this study, we defined treatment success/ failure, or cure of the exacerbation as a dichotomous measure of the outcome of the exacerbation. We excluded continuous measures evaluating change in variables without pre-specified thresholds of success or failure. We used a stricter definition compared to our original methodological SR [5] and for this reason, we found a lower proportion of studies assessing this outcome.

In brief, two broad categories of instruments were used to describe this outcome.

The first category, that was used in 24 RCTs described treatment failure as a composite outcome consisting of different unfavourable outcomes. Most frequently used components were (i) death, (ii) need for hospital admission or re-admission, (iii) need for endotracheal intubation or mechanical ventilation, and (iv) persistence or deterioration of the symptoms and signs.

The second category, that was assessed in 33 RCTs consisted of qualitative or semi-quantitative descriptions of the clinical status of the patient. Four states were described: Cure, marked improvement, improvement, and treatment failure. RCTs frequently used more than one states to describe the outcome. In trials evaluating both favourable and unfavourable states, the definition of treatment failure was usually complimentary to the definition of one of the favourable outcomes. The most frequently described favourable outcomes were: (i)

Complete resolution of all signs and symptoms of the exacerbation [reported in 8 RCTs], and (ii) Sufficient improvement of the signs and symptoms, such that no additional systemic treatments were prescribed. The most frequently utilized definitions for treatment failure were (i) Lack of resolution of signs and symptoms, requiring additional treatment, or death; [reported in 7 RCTs], and (ii) Persistence or worsening of signs or symptoms, or death [reported in 7 RCTs]. All these definitions are based on the clinicians' opinion around the exacerbation's status.

Literature review:

Search terms describing the outcome	*COPD and Exacerbations terms – see figure S1* **AND** ((Treatment failure [MH]) or (cure [tiab]) or (treatment success [tiab]) or (treatment failure [tiab])) **AND** ((Treatment Outcome [MH]) or (Outcome Assessment, Health Care [MH]) or (instrument* [tiab]) or (outcome* [tiab]) or (endpoint* [tiab]))
Number of titles screened	269
Relevant studies	Methodological systematic reviews: 0 Other references: 0

This focused literature review did not reveal any methodological study evaluating the measurement properties of different instruments used to assess treatment success or cure of a COPD exacerbation. For this reason, we conducted a more thorough meta-epidemiological study to inform the selection of this outcome (see previous section).

Panel discussion summary: Our systematic reviews revealed significant variability in the definitions and/or instruments used to evaluate treatment success or failure. Some trials used composite endpoints consisting of several adverse outcomes of an exacerbation, such as death, need for treatment intensification, or need for hospital admission, together defining an

overall unfavourable outcome. However, these include components that have very different impact (utility) on patients, while the relative frequency of these outcomes may differ across COPD exacerbations subgroups. Moreover, most of these components are included as independent outcomes in the core set, anyway. For these reasons, the panel did not consider that the evaluation of such composite endpoints would add value to the trials.

In most trials, evaluation of treatment success was based on qualitative or semi-quantitative descriptions of the status of the exacerbation, such as cure, improvement, or treatment failure, based on the extent of patients' symptoms and signs. The main limitation of these instruments is the subjectivity of the assessments of the severity of symptoms and signs by patients and clinicians. Therefore, these outcomes may be susceptible to performance and detection bias.

Treatment success was more frequently defined as cure of the exacerbation and more specifically as the "Complete resolution of all signs and symptoms of the exacerbation". However, the recovery period of an exacerbation varies significantly and may be very prolonged. Large observational studies have shown wide variability in the duration of exacerbation recovery, revealing that 25% of patients still experience symptoms associated with the exacerbation 25 or even 35 days after the onset of the exacerbation [21, 22]. Longer periods may be required until patients recover their previous exercise capacity or ADL levels [23, 24]. Moreover, exacerbations accelerate disease progression; therefore, the clinical condition after recovery from an exacerbation may be characterized by a greater symptomatic burden, compared to the previous baseline [25]. As a result, this definition of cure was considered problematic. The second most frequently used definition of treatment success "Sufficient improvement of the signs and symptoms, such that no additional systemic treatments were prescribed" was considered more pragmatic and was endorsed by the panel as an interim instrument. While still subjective, the decision of the clinician to prescribe additional systemic treatments better reflects daily clinical practice and it is often used in trials.

Recommendation: Treatment success defined as sufficient improvement of the signs and symptoms of the exacerbation that no additional systemic treatments (antibiotics or systemic corticosteroids) are required (**Interim Recommendation with research agenda**).

6.4 Need for hospital admission for the presenting exacerbation

Data from the methodological systematic review [5]: This outcome has two components: whether a patient required hospital admission at any timepoint and whether they still required hospital admission at a specific follow-up timepoint. The former and latter components are more relevant for RCTs evaluating moderate and severe exacerbations, respectively. In our methodological systematic review, 33 (27%) studies evaluated length of hospital stay and three studies need for hospital admission for the index exacerbation. This outcome was assessed consistently by recording whether a participant was admitted to the hospital (at a specific timepoint or daily until discharge).

Literature review: Not performed since this outcome is evaluated consistently.

Panel discussion summary: Hospital at home and telemonitoring options introduce heterogeneity in the criteria for hospital admission and length of stay [26]. This outcome is also impacted by non-clinical factors, such as social reasons, discharge planning delays [27], the availability of hospital beds, or travel distance. These issues should be accounted for when evaluating duration of hospital stay.

Recommendation: A clinical need to admit a patient to the hospital, or equivalent intensification of the monitoring or care that may be provided in other settings (including patients' home). Admissions for social reasons should be reported separately.

For evaluating this outcome investigators should record whether a patient required admission at any timepoint and whether they still require hospital admission at a specific follow-up timepoint (**Interim Recommendation with research agenda**).

6.5 Need for admission to the intensive care unit (ICU) for the presenting exacerbation.

Data from the methodological systematic review [5]: Similar to the outcome need for hospital admission, this outcome has two components: whether a patient required admission to the ICU at any timepoint and whether they still required ICU admission at a specific follow-up timepoint. The former and latter components are more relevant for RCTs evaluating severe (hospitalized) and critical (admitted to the ICU) exacerbations, respectively. In our methodological systematic review, 10 (8%) studies evaluated length of ICU admission, 10 (8%) length of invasive mechanical ventilation, two the need for ICU admission and two the need for invasive mechanical ventilation. As described in the following section, invasive mechanical ventilation could be used as a measure of the need for ICU admission. This outcome was assessed consistently by recording whether a participant was admitted to the ICU or were invasively ventilated (at a specific timepoint or daily until discharge).

Literature review: Not performed since this outcome is evaluated consistently.

Panel discussion summary: Indications for admission to the ICU vary significantly. Characteristically, while in most centres non-invasive ventilation is now delivered in a respiratory ward or a high dependency unit, in some centres it is still delivered in the ICU [28]. Availability of ICU beds may also impact the decision to admit, and the duration of ICU stay. On the other hand, patients with COPD with poor functional status and underlying multi-morbidity are often not offered an ICU admission or invasive mechanical ventilation, due to futility [29]. The criteria used to support such decisions vary across centres and countries, according to local policies and availability of resources.

Acknowledging that the main, consistent indication for ICU admission in this group of patients is the need for invasive mechanical ventilation, the panel recommended that trials should record the need for invasive mechanical ventilation. A clear definition for the need for invasive mechanical ventilation for adult patients with acute hypercapnic respiratory failure was identified in the BTS/ICS guideline for the ventilatory management of acute hypercapnic respiratory failure in adults (see next section; a focused literature review did not reveal any

other recent guidelines addressing indications for invasive ventilation in this patient group) [29]. The decision to focus on the need for invasive mechanical ventilation rather than the receipt of ventilation was based on the earlier observation that often, while these criteria are fulfilled, patients are not offered invasive ventilation, due to futility.

Recommendation: Need for ICU admission should be evaluated on the basis of the need for invasive mechanical ventilation, defined as (i) persistent or deteriorating respiratory acidosis despite optimized medical treatment and delivery of non-invasive ventilation (NIV); (ii) persistent or deteriorating respiratory acidosis despite optimized medical treatment and a contra-indication for the use of NIV, for example due to severe facial deformity where fitting a mask is impossible, upper airway obstruction, or facial burns; (iii) respiratory arrest or peri-arrest situations unless there is a rapid recovery from manual ventilation or provision of NIV.

For evaluating this outcome investigators should record whether a patient required admission at any timepoint and whether they still require ICU admission at a specific follow-up timepoint (**Interim Recommendation with research agenda**).

6.6 Levels of oxygen and carbon dioxide in the blood (arterial blood gases).

Data from the methodological systematic review [5]: Forty (33%) RCTs reported on arterial blood gases (pH, oxygen tension, carbon dioxide tension, and/or oxygen saturation measured by pulse oximetry). In all studies, arterial blood was sampled for evaluating blood gases.

Literature review: Not performed since this outcome is evaluated consistently.

Panel discussion summary: This was considered a setting and intervention specific outcome. Firstly, it may not be feasible to be assessed in studies recruiting in an outpatient clinic. The panel agreed that the value of measuring blood levels of oxygen and carbon dioxide in this setting may be limited.

On the other hand, evaluating arterial (and not venous) blood gases as an outcome in hospitalized patients is crucial both for clinical purposes, but also as a research outcome.

While a single measurement might be sufficient in clinical practice, at least two measurements are required in the context of a research study, in order to evaluate the magnitude of change from baseline in response to treatment. For this reason, the panel recommends that a baseline and at least one follow-up measurement are required. However, more intensive monitoring of the arterial blood gases may be required for specific interventions, such as non-invasive ventilation or modes of oxygen delivery.

Recommendation: A setting and intervention specific outcome. A baseline and at least one follow-up measurement are required with a clear indication of whether or not the patient was receiving oxygen at the time of the measurement, and if yes, how much. It may not be feasible for studies evaluating outpatients (**Interim Recommendation with research agenda**).

6.7 Breathlessness

Data from the methodological systematic review [5]: Breathlessness was evaluated using the Borg's scale in 13 (11%) of the included studies, the modified Medical Research Council (mMRC) Dyspnoea Scale in 6 (5%) trials, the Baseline and Transitional Dyspnoea Index in 1 trial and other, non-validated Scales in 11 (9%) trials. Moreover, it was assessed as part of multidimensional symptoms/ severity scores, mainly the COPD Assessment Test (CAT). Other scores evaluated less frequently included the EXAcerbation of Chronic Obstructive pulmonary disease tool – Patient Reported Outcome (EXACT-PRO), the Clinical COPD questionnaire (CCQ), Chronic Respiratory Questionnaire (CRQ) and the BODE index.

Literature review:

Search terms describing the outcome	*COPD and Exacerbations terms – see figure S1* **AND** ((Dyspnea [MH]) or (dyspnea [ti]) or (dyspnoea [ti]) or (breathlessness [ti])) **AND**

	((Treatment Outcome [MH]) or (Outcome Assessment, Health Care [MH]) or (instrument* [tiab]) or (outcome* [tiab]) or (endpoint* [tiab]))
Number of titles screened	269
Relevant studies	Methodological systematic reviews: 3 [30-32] Other references: Not considered – Adequate SR data

Our focused literature review revealed three methodological systematic reviews evaluating the performance characteristics of instruments used to evaluate breathlessness. Oliveira and Marques only included studies focusing on the measurement properties of instruments used to assess breathlessness specifically during pulmonary rehabilitation in patients with acute exacerbations, therefore, it was not informed by adequate data [32]. Jadad and colleagues did not formally evaluate measurement properties of the identified instruments [30]. For these reasons the panel discussion was mainly informed by the review conducted by Dorman and colleagues, that focused on the evaluation of breathlessness in palliative care, with a specific focus on COPD and the identified measurement properties were considered applicable to our work (although indirect) [31].

Panel discussion summary: The mMRC Scale does not directly assess breathlessness, as it is a measure of activity limitation due to breathlessness. Moreover, use of the mMRC during an exacerbation was considered by the panel less sensitive, since most patients with moderate or severe exacerbations would cluster in Grade 4 (“Too breathless to leave the house or breathless when dressing or undressing”), thus limiting the discriminant validity of the scale in this context. CAT is a multidimensional tool measuring several symptoms and health status and therefore does not provide a focus on breathlessness [33]. CAT will be captured anyway, as it is recommended for evaluating health-related quality of life.

The modified Borg Scale is easy to complete, and broadly used in clinical practice and research. Clinically validated translations are available in many languages. Its measurement

properties have been thoroughly and positively assessed [31] (table S5). As a result, the modified Borg Scale was recommended by the panel.

Table S5. Psychometric properties of the Borg's Scale (Data source: [31])

Psychometric characteristics	Confirmation
Face validity	Confirmed
Content validity	-
Factor analysis	N/A
Construct validity	Confirmed
Discriminant validity	Confirmed
Test-retest	? Variability identified
Internal consistency	N/A
Responsiveness	No data*
Acceptability	Confirmed
Time to complete	Confirmed – Very quick

*Responsiveness was not confirmed in this methodological SR, that was not specific to COPD exacerbations. However, numerous trials using the scale as an outcome for COPD exacerbations demonstrate treatment response, suggesting good responsiveness.

Recommendation: Breathlessness should be evaluated using the modified Borg's scale. It should be measured at approximately the same time every day. It can be self-completed (**Interim Recommendation with research agenda**).

6.8 Health-related quality of life.

Data from the methodological systematic review [5]: Comprehensive health status and quality of life questionnaires were used in 34 (28%) of the included studies. COPD assessment test (CAT) was used in 11 (9%) studies, the Saint George's Respiratory Questionnaire in 8 studies, the Clinical COPD Questionnaire (CCQ) in 6 studies, the Chronic Respiratory Questionnaire (CCQ), the Euroqol-5D and the 36-Item Short Form Survey in 5 studies each (some studies assessed more than one instruments). Other instruments were used less frequently.

Literature review:

Search terms describing the outcome	*COPD and Exacerbations terms – see figure S1* **AND** ((Quality of life [MH]) or (quality of life [ti]) or (health status[ti])) **AND** ((Treatment Outcome [MH]) or (Outcome Assessment, Health Care [MH]) or (instrument* [tiab]) or (outcome* [tiab]) or (endpoint* [tiab]) or (questionnaire* [tiab]))
Number of titles screened	1,018
Relevant studies	Methodological systematic reviews: 4 [32, 34-36] Other references: Not considered – Adequate SR data

Our focused literature review revealed two methodological systematic reviews evaluating the performance characteristics of instruments used to evaluate health related quality of life in COPD. Oliveira and Marques only included studies focusing on the measurement properties of instruments used to assess quality of life specifically during pulmonary rehabilitation in patients with acute exacerbations, therefore, it was not informed by adequate data [32]. As a result, the panel discussion was mainly informed by Weldam and colleagues, a systematic review that evaluated the performance characteristics of Quality of Life instruments for use in COPD [37]. While this methodological systematic review was not specifically focused on

COPD, it was considered appropriate for informing out work. Further information about CAT and the CCQ were sourced by two other systematic reviews by Gupta et al [35] and Zhou et al [36], focusing on the performance characteristics of these tools, respectively.

Panel discussion summary:

CAT is the most frequently used validated tool for assessing health related quality of life in trials on the management of exacerbations, followed by the Saint George's Respiratory Questionnaire (SGRQ) and the Chronic COPD Questionnaire (CCQ) [5]. A systematic review using the COSMIN methodology for evaluating the measurement properties of 23 instruments used to assess quality of life in COPD recommended the use of CAT, Chronic Respiratory Questionnaire (CRQ), the Saint George's Respiratory Questionnaire (SGRQ) or the Living with Chronic Obstructive Pulmonary Disease (LCOPD) Questionnaire [34]. While these tools have similar measurement properties (summarized in table S6), CAT can be completed within 1-3 minutes while the other tools are more complex and time consuming. Given that CAT is already the most frequently used tool for evaluating health-related quality of life, it was recommended by the panel. A comparison with a baseline estimate of the health-related quality of life prior to the exacerbation would be beneficial, but in larger randomized studies, balance in the baseline characteristics of participants in the study groups can usually be trusted to randomization.

Table S6. Measurement properties of instruments used to assess quality of life in COPD. Summary of the (i) judgements on the quality of the available methodological studies and (ii) their findings around whether the instruments fulfil each criterion. Judgement of the methodological quality was based on the study with the best methodological quality, among those concluding more favourable properties for each of the instruments. Scale: Poor, Fair, Good, Excellent. Findings: Sufficient (+), Indeterminate (?), Insufficient (-). (Data source: [34])

	CAT	CRQ	SGRQ	LCOPD	CCQ
Disease specific	Yes	Yes	Yes	Yes	Yes
Content validity	Excellent +	Excellent +		Excellent +	
Criterion validity					
Structural validity	Excellent +	Excellent +	Excellent +		
Cross-cultural validity		Poor +	Poor ?	Poor ?	Poor ?
Internal consistency	Excellent +	Excellent +	Good +	Good +	Poor +
Reliability	Good +	Good +	Excellent +	Good +	Good +
Measurement error					
Responsiveness	Good +	Good +	Good +		Good +
Ease of completion	1-3 mins	15-25 mins	25 mins	10 mins	1-3 mins

Recommendation: The COPD Assessment Test (CAT) should be used for assessing health related quality of life (Interim Recommendation with research agenda).

6.9 Activities of daily living (ADL)

Data from the methodological systematic review [5]: Activities of daily living as an outcome is rarely evaluated in COPD exacerbations trials. More specifically only two of the included studies evaluated this outcome. One used the Activity of Daily Living Dyspnoea Scale (ADL-D scale) and the other the Barthel's index.

Literature review:

Search terms describing the outcome	*COPD and Exacerbations terms – see figure S1* **AND**

	((Activities of daily living [MH]) or (Functional Status [MH]) or ((activities [ti]) and ((life[ti]) or (living[ti]))) or ((function* [ti]) and (status [ti]))) **AND** ((Treatment Outcome [MH]) or (Outcome Assessment, Health Care [MH]) or (instrument* [tiab]) or (outcome* [tiab]) or (endpoint* [tiab]) or (questionnaire* [tiab]))
Number of titles screened	221
Relevant studies	Methodological systematic reviews: [32, 38, 39] Other references: Not considered – Adequate SR data

This focused systematic review revealed three methodological systematic reviews evaluating thee performance characteristics of instruments used to evaluate activities of daily living in COPD. Oliveira and Marques only included studies focusing on the measurement properties of instruments used specifically during pulmonary rehabilitation in patients with acute exacerbations, therefore, it was not informed by adequate data [32]. Two systematic reviews by Janaudis-Ferreira [38] and by Liu [39] assessed ADL in COPD. While they were not focused specifically on exacerbations, they were considered appropriate for informing our work.

Panel discussion summary:

This outcome is rarely evaluated in exacerbation trials. ADL are classified as basic and instrumental [40]. Basic ADL are simple activities that are essential for independent life, such as self-care (showering, dressing, or grooming) and basic mobility, while instrumental ADL encapsulate more complex activities, requiring higher functioning, such as preparing meals, home maintenance, shopping, handling finances, and travelling alone [38]. Instrumental ADL are less relevant during an exacerbation, especially during severe exacerbations, while patients are admitted in the hospital and may not be able to undertake such complex activities; but they are pertinent to quantify the overall impact of an exacerbation on a patient's ADL. For

this reason, the panel decided to recommend a tool focusing on basic ADL, to be evaluated during the exacerbation and a second tool, assessing both basic and instrumental ADL for longer-term follow-up.

The psychometric properties of instruments used to quantify ADL in patients with COPD have been evaluated in two methodological systematic reviews [38, 39]. Five of the identified instruments focused on basic ADL, of which the Katz Activities of Daily Living Scale, the Barthel index and the motor subscale of the functional independence measure (FIM) were not disease specific and included domains that are less relevant to COPD patients (e.g., control of bladder and bowels). While the Glittre index is disease specific, it focuses on exercise capacity and includes a simple exercise component, which many patients may find challenging to complete during an exacerbation. Finally, the Capacity of Daily Living during the Morning (CDLM) Questionnaire [41] is a simple, disease specific questionnaire, whose measurement properties have been adequately evaluated with favourable findings (table S7). For this reason, the CDLM tool was recommended for quantifying basic ADL during an exacerbation.

The identified methodological reviews revealed eight disease-specific tools assessing a combination of instrumental and basic ADL [38, 39]. Responsiveness to change in a patient's clinical condition, a crucial characteristic required for evaluating the impact of exacerbation on ADL, has only been confirmed for three of these tools: the Manchester Respiratory Activities of Daily Living Questionnaire (MRADL) [42], the COPD Activity Rating Scale (CARS) [43], and the 11-items Pulmonary Functional Status Scale (PFSS-11) [44]. While all three tools were considered valid options, the performance characteristics of the MRADL questionnaire were more thoroughly validated compared to CARS, while it was also considered simpler to complete, compared to the PFSS-11 tool (table S7). For promoting consistency, the panel recommends that the MRADL questionnaire be used to evaluate both basic and instrumental ADL at recovery from COPD exacerbations. A comparison with a baseline estimate of the ADL prior to the exacerbation would be beneficial and could potentially be captured retrospectively during recruitment. Recall bias is anticipated to be limited, since in most cases, the duration

of the acute event at recruitment would rarely exceed a week and the questions refer to some of the most critical activities of daily living.

Table S7. Measurement properties of instruments used to assess activities of daily living in COPD. Summary of the (i) judgements on the quality of the available methodological studies and (ii) their findings around whether the instruments fulfil each criterion. Judgement of the methodological quality was based on the study with the best methodological quality, among those concluding more favourable properties for each of the instruments. Scale: Poor, Fair, Good, Excellent. Findings: Sufficient (+), Indeterminate (?), Insufficient (-). (Data source: [38, 39]).

	CDLM	Glittre	MRADL	CARS	PFSS-11
Disease specific	YES	YES	X	X	X
Content validity	Good (+)	Poor (?)	Fair (+)	Poor (?)	Fair (?)
Criterion validity					
Structural validity				Fair (+)	Good (+)
Hypothesis testing	Fair (+)	Fair (-)	Good (+)	Fair (+)	Fair (+)
Cross-cultural validity					
Internal consistency	Poor (?)		Good (+)	Fair (+)	Good (+)
Reliability	Fair (+)	Good (+)	Good (+)		Poor (?)
Measurement error					
Responsiveness	Fair (?)	Fair (?)	Fair (+)		Fair (+)
Interpretability	X				
Ease of completion	Yes	Not during AECOPD	X	X	X

Recommendation: The Capacity of Daily Living in the Morning Questionnaire (CDLM) should be used for evaluating basic activities of daily living during the exacerbation (**Interim Recommendation with research agenda**).

The Manchester Activities of Daily Living Questionnaire (MRADL) should be used for evaluating basic and instrumental activities of daily living, during recovery (long-term impact of the exacerbation) (**Interim Recommendation with research agenda**).

6.10 Worsening of symptoms after the initial treatment

Data from the methodological systematic review [5]: Changes in symptoms was evaluated using symptom scores and scales, quality of life and/or health status instruments in 73 (59%) trials. 41 (33%) of the studies assessed symptoms progression using simple symptom scores, such as visual analogue scales or Likert scales. 34 (28%) of the studies utilized comprehensive health status and quality of life questionnaires, mostly the COPD assessment test (CAT), the Saint George's Respiratory Symptoms Questionnaire and the Clinical COPD questionnaire (CCQ).

Literature review: Not performed. The discussion for this instrument was informed by the focused systematic reviews conducted for the outcomes (i) Breathlessness and (ii) Quality of Life.

Panel discussion summary: The panel considered that this outcome can be evaluated using the Borg's scale and CAT test, that have already been recommended as measures of breathlessness and health related quality of life, respectively. Moreover, it was highlighted that three PROs have already been recommended for regular assessment during the exacerbation (Borg's scale, CAT test and the CDLM scale). There were concerns that a recommendation for additional daily PROs could limit the feasibility and uptake of the core outcome set.

Recommendation: The modified Borg's scale and the COPD assessment test (CAT) should be used to detect symptoms worsening after the initial treatment (**Interim Recommendation with research agenda**).

6.11 Disease progression

Data from the methodological systematic review [5]: The definition of this outcome is available in the panel discussion summary section. Four studies recruited patients at stable clinical disease and could therefore captured their baseline status. However, only two of them attempted to evaluate disease progression by comparing forced expiratory volume in 1 second (FEV_1) before and after the exacerbation. No other studies evaluated disease progression.

Literature review:

Search terms describing the outcome	*COPD terms – see figure S1* **AND** ((Disease Progression [MH]) or (progression [ti])) **AND** ((Treatment Outcome [MH]) or (Outcome Assessment, Health Care [MH]) or (instrument* [tiab]) or (outcome* [tiab]) or (endpoint* [tiab]))
Number of titles screened	1530
Relevant studies	Methodological systematic reviews: 0 Other references: 19 [25, 45-62].

We did not identify methodological systematic reviews or studies evaluating the measurement properties of instruments used to evaluate disease progression in COPD. Such studies would be challenging and resource intense to conduct, as large study populations and prolonged

follow-up would be needed to formally assess instruments for evaluating disease progression. We identified one consensus document attempting to define disease progression as an outcome [45] and several studies aiming to identify variables that could be used to assess this outcome [46-58]. The consensus document described several instruments for evaluating disease progression: Decline in FEV_1, exercise capacity, or health status health status, assessment of progression by CT scanning, increase in healthcare utilization and costs. The list of studies aiming to identify variables that could be used to assess disease progression is not exhaustive, since the search strategy aimed to identify methodological studies. These studies assessed the association of numerous laboratory tests and biomarkers as predictors of disease progression. Interestingly, they used decline in FEV_1 and progression by CT scanning as gold-standards for evaluating disease progression.

Panel discussion summary: This outcome was suggested by patients during the qualitative research studies that preceded the Delphi survey. Acute exacerbations are known to accelerate disease progression in patients with COPD [25, 59, 60]. Several parameters have been used as potential measures of disease progression, including symptom burden, health status, exercise capacity, blood biomarkers, pulmonary function decline, or radiologic progression revealed in computed tomography (CT) of the chest [45, 58, 60-62].

There was agreement within the panel that evaluation of disease progression as an outcome in exacerbation trials is only meaningful as change from baseline; therefore, a baseline measurement is required. To achieve that, participants would have to be recruited while the disease is stable, in anticipation of developing an exacerbation. However, such a study design requires significantly more resources and prolonged follow-up periods or a patient database with recent measurement taken during periods of clinical stability.

Not surprisingly, disease progression is only rarely evaluated as an outcome in exacerbation trials using objective tests [5]. Change from baseline in pulmonary function was only assessed in two of the trials included in the methodological systematic review, while imaging was not

used in any of the studies as an estimate of disease progression. Symptoms and quality of life are evaluated frequently, but not as change from baseline (see respective outcomes).

Change in FEV_1 over time is the most established instrument for evaluating COPD progression in clinical trials and observational studies evaluating the management of disease longitudinally and for this reason, the panel recommends that it should also be used for evaluating the impact of exacerbations on disease progression. Acknowledging the limitations of this study design, the panel recommends that this outcome only be considered core for long-term studies where baseline values can be captured.

Recommendation: Permanent deterioration in lung function should be used to evaluate the impact of exacerbations on disease progression. Two pulmonary function tests during stable clinical condition are needed: One within 6 months prior to the index exacerbation, and one within 2-6 months afterwards. Change from baseline in forced expiratory volume in 1 second (FEV1), forced vital capacity (FVC) and FEV1/FVC ratio should be noted. The number of exacerbations experienced between the two measurements should be noted. Ideally, only the index exacerbation should be included between the two measurements.

Disease progression as a core outcome is only relevant for longer-term studies that recruit participants during stable disease state, in anticipation of an exacerbation (**Interim Recommendation with research agenda**).

6.12 Future exacerbations

Data from the methodological systematic review [5]: Future exacerbations were evaluated in 28 (23%) clinical trials. Exacerbations during follow-up were noted and many trials also noted whether these were moderate or severe. Analytical methodology varied (number of patients with at least one exacerbation, mean/median number of exacerbations, time to next exacerbation). However, analytical methodology is beyond the scope of this document.

Literature review: Not performed since this outcome is evaluated consistently.

Panel discussion: It is crucial that treatment success or cure of the index exacerbation should be clearly defined, to allow for the distinction between prolonged symptoms due to a single exacerbation and new exacerbations.

Recommendation: The number of future exacerbations during follow-up should be recorded, noting whether they are moderate or severe (**Interim Recommendation with research agenda**).

6.13 Future hospital admissions

Data from the methodological systematic review [5]: Future hospital admissions were evaluated in 14 (11%) clinical trials. All trials evaluating this outcome noted hospitalisations for any reason during follow-up.

Literature review: Not performed, since this outcome is evaluated consistently.

Panel discussion summary: Similar to the outcome need for admission for the presenting exacerbation, concerns were raised regarding (i) social admissions and (ii) the variability in the indications for future hospital admission, for example due to hospital-at-home and telemonitoring options. For this reason, it was decided that the outcome “Future hospital admission” should incorporate equivalent intensification of the monitoring or care that may be provided in another setting. Trialists need to prospectively record available hospital-at-home and telemonitoring options and the thresholds for considering “equivalent intensification of the monitoring or care” in their setting. Hospital admissions for social reasons should not be counted.

Recommendation: Future hospital admissions for any medical reason, or equivalent intensification of the monitoring or care that may be provided in other settings, after treatment success is confirmed (**Interim Recommendation with research agenda**).

6.14 Serious adverse events from treatments

Data from the methodological systematic review [5]: Serious adverse events were captured in 73 (59%) of the included studies. This outcome is consistently captured following the definition and methodology proposed by the International Council for Harmonisation [63].

Literature review: Not performed since this outcome is evaluated consistently.

Panel discussion summary: This outcome is consistently evaluated universally following the definition and methodology proposed by the International Council for Harmonisation.

Recommendation: Following the definition of the International Council for Harmonisation. Serious adverse event is any untoward medical occurrence in a patient or clinical investigation subject administered a pharmaceutical product and which does not necessarily have to have a causal relationship with the treatment, that fulfils any of the following: (a) Results in death; (b) Is life threatening; (c) Requires inpatient hospitalisation or prolongation of existing hospitalisation; (d) Results in persistent or significant disability / incapacity; (e) Is a congenital anomaly or birth defect. Suspected unexpected serious adverse reactions (SUSARs) should also be reported. **(Strong recommendation)**.

6.15 Development of resistant bacteria

Data from the methodological systematic review [5]: Bacterial resistance was evaluated as part of the composite outcome microbiological response in 16 (13%) RCTs. Other trials reported the presence of new bacterial resistance as an adverse event. None of the included studies reported performing sputum induction and bacterial resistance results are based on spontaneous sputum.

Literature review: Not performed since this outcome is evaluated consistently.

Panel discussion summary: Antimicrobial resistance is often explored as part of a composite microbiological response outcome or as adverse event in trials involving antibiotics as interventions. Bacterial growth and resistance are usually evaluated in spontaneous sputum, while in the absence of sputum, bacterial eradication is presumed and is not further assessed.

The panel adopts this approach. Moreover, it was discussed that bacterial resistance may not be a relevant outcome for all interventions, but only for antimicrobials, antimicrobial stewardship strategies, novel immune modifiers, or other interventions that may affect bacterial resistance.

Recommendation: Trials evaluating antimicrobials, antimicrobial stewardship strategies, novel immune modifiers or other interventions that may affect bacterial resistance should evaluate bacterial resistance to the administered antibiotics in spontaneous sputum. As a minimum, resistance should be evaluated at baseline and within a week after treatment completion.

Sputum induction may provide additional information. However, in each study, researchers should consider the balance between the added value compared to the risk, participants discomfort and required resources (**Interim Recommendation with research agenda**).

6.16 Development of pneumonia

Data from the methodological systematic review [5]: Development of pneumonia is captured as an adverse event. Adverse events were captured in 73 (59%) of the included studies. Most of these studies described the frequency of the most prevalent adverse events, including pneumonia. Pneumonia was diagnosed by the presence of new consolidation in a chest X-ray or CT chest that was performed in response to consistent clinical signs and symptoms. Not surprisingly, none of the trials described asymptomatic screening for pneumonia during the follow-up.

Literature review: Not performed since this outcome is evaluated consistently.

Panel discussion summary: Development of pneumonia as a safety outcome is often evaluated in exacerbation trials. Methodology is consistent and was adopted by this task force. Pneumonia should be confirmed by the presence of new consolidation in the chest X-ray or other imaging modalities of the chest, in the presence of consistent clinical signs and

symptoms. A baseline chest x-ray would be helpful, but it may not be feasible for trials recruiting patients outside the hospital setting.

Recommendation: Pneumonia confirmed by the presence of new consolidation in the chest x-ray or other imaging modalities of the chest, in the presence of consistent clinical signs and symptoms. When possible, chest imaging should be acquired at baseline, to assess for the presence of pneumonia. This may not be possible for trials recruiting patients outside the hospital setting. Follow-up chest imaging should be driven by clinical need. **(Strong recommendation)**.

6.17 Treatment adherence

Data from the methodological systematic review [5]: Adherence was evaluated in 7 (6%) of the included trials. Methodology varied according to the intervention.

Literature review: Not performed since assessment of this outcome is treatment specific.

Panel discussion summary: This is an intervention specific outcome. Trialists should describe transparently the methodology used for evaluating treatment adherence.

Recommendation: his outcome was considered intervention specific. Methods for assessing treatment adherence should be clearly reported **(Strong recommendation)**.

7 Outcome measurement instruments: Research agenda

Death from COPD exacerbation. - Development and implementation of standardized methodology for determining the cause of death during an acute event, such as an acute exacerbation.
Treatment success. - Development of objective and accurate methods for confirming treatment success. - Development of objective and accurate methods for confirming cure. - Quantification of the duration of exacerbations and identification of timepoints when the evaluation of treatment success is sensitive to treatment effect.
Need for hospital admission for the presenting exacerbation. - There is a need for novel instruments that could consistently capture the need for monitoring or care intensification that is traditionally offered in a hospital setting.
Need for admission to the intensive care unit for the presenting exacerbation. - Standardization of the indications and contra-indications for (i) admission to the intensive care unit, and (ii) mechanical ventilation, of patients with COPD exacerbations.
Levels of oxygen and carbon dioxide in the blood (arterial blood gases). - Development of instruments that will allow for comparison of the levels of oxygen and carbon dioxide in the blood of patients receiving different levels of supplemental oxygen. - Development and validation of non-invasive methods for estimating the levels of carbon dioxide in the blood.
Breathlessness. - Formal evaluation/ comparison of measurement properties of instruments used to evaluate breathlessness during COPD exacerbations.
Health related quality of life. - Formal evaluation/ comparison of measurement properties of instruments used to evaluate quality of life during COPD exacerbations.
Activities of daily living. - Formal evaluation/ validation of the properties of instruments used to measure activities of daily living during and after a COPD exacerbation, using the COSMIN methodology. - The Capacity of Daily Living in the Morning (CDLM) questionnaire focuses on morning activities. Evaluation of other tools evaluating activities throughout the day. - Development of validated translations of the selected instruments and confirmation of cross- cultural validity.
Worsening of symptoms after the initial treatment.

- Formal evaluation of the measurement properties of tools that could be used to identify worsening of symptoms after the initial treatment (such as the EXACT-PRO [64]).
Disease progression. - Development novel and simple methods for evaluating the impact of exacerbations on disease progression. - Evaluation of the role of other pulmonary function parameters in evaluating the impact of exacerbations on disease progression (e.g. lung volumes, diffusion capacity). - Could change in the computed tomography (CT) of the chest compared to baseline reveal the impact of the exacerbation on disease progression (e.g. the extent of emphysema quantified by loss of lung density, or changes in the diameter of the pulmonary artery).
Future exacerbations. - Development of consistent methods for differentiating a prolonged exacerbation from the onset of a new exacerbation. - Development and validation of methodology for differentiating different types of COPD exacerbations.
Future hospital admissions. See: Need for hospital admission for the index exacerbation.
Development of resistant bacteria. - Assessment of the additional information offered by conducting sputum induction to assess for bacterial resistance in patients recovering from a COPD exacerbation. - Evaluation of the sensitivity of different types of samples (respiratory or non-respiratory) in evaluating bacterial resistance.

8 Consensus meeting participants

First consensus meeting: Finalization of the Core Outcome Set

1. Patients with COPD and patient representatives

Name (if consented)	Country
Arrowsmith, Christine	UK
Branch, Kay	UK
Bruce, Elaine	Ireland
Coleman, Courtney	UK (ELF representative)
Jessica Denning	UK (ELF representative)
Jensen, Bo Hammer	Denmark
Hood, David	UK
Janssen, Elly	Netherlands
Jelen, Tessa	UK
Linnell, John	USA
Jonsdottir, Aldis	Iceland
Meggitt, Richard	Australia
Preston, Allan	UK
Ratcliffe, John	Australia
Ruttle, John	Australia
Winders, Tonya	USA
Vinuela, Alfonso	Spain

2. Health professionals and clinical researchers

Name	Country
Agusti, Alvar	Spain
Bartziokas, Konstantinos	Greece

Bradbury, Thomas	Australia
Corlateanu, Alexandru	Moldova
Csoma, Balazs	Hungary
Emelyanov, Alexander	Russia
Fernandez Romero, Gustavo	USA
Jenkins, Christine	Australia
Jensen, Jens-Ulrik	Denmark
Kharevich, Olga	Belarus
Kostikas, Konstantinos	Greece
Lazar, Zsofia	Hungary
Lopez-Giraldo, Alejandra	Spain
Mathioudakis, Alexander	UK
McDonald, Vanessa	Australia
Papi, Alberto	Italy
Sergeeva, Galina	Russia
Sivapalan, Pradeesh	Denmark
Stovold, Elizabeth	UK
Vestbo, Jørgen	UK/ Denmark
Wang, Hao	China
Wen, Fuqiang	China

3. COMET representatives / methodologists

Name	**Role**
Brookes, Sara	Meeting facilitator
Williamson, Paula	Methodological input

Second consensus meeting: Selection of outcome measurement instruments

1. Patients with COPD and patient representatives

Name	**Country**
Coleman, Courtney	UK (ELF representative)
Linnell, John	USA
Saraiva, Isabel	Portugal

2. Health professionals and clinical researchers

Name	**Country**
Ananth, Sachin	UK
Bartziokas, Konstantinos	Greece
Beghe, Bianca	Italy
Bradbury, Thomas	Australia
Corlateanu, Alexandru	Moldova
Emelyanov, Alexander	Russia
Fernandez Romero, Gustavo	USA
Jenkins, Christine	Australia
Jensen, Jens-Ulrik	Denmark
Kostikas, Konstantinos	Greece
Lazar, Zsofia	Hungary
Mathioudakis, Alexander	UK
McDonald, Vanessa	Australia
Papi, Alberto	Italy
Sergeeva, Galina	Russia
Sioutkou, Agni	Greece
Sivapalan, Pradeesh	Denmark
Stovold, Elizabeth	UK

Vestbo, Jørgen	UK/ Denmark
Wang, Hao	China
Wen, Fuqiang	China

3. COMET representative / methodologist

Name	**Role**
Williamson, Paula	Methodological input & Meeting facilitator

9 Expanded discussion

9.1 Comparison with other outcome prioritization initiatives

While this is the first formal core outcome set for COPD exacerbations outcomes have been prioritized by two other initiatives.

First, COPD exacerbations outcomes have also been assessed and prioritized by the eo-Drive trial group (Eosinophil-driven corticotherapy for patients hospitalized for COPD Exacerbations, NCT04234360). Consensus was developed through a Delphi survey involving 21 French clinical academics with expertise in COPD exacerbation trials [65]. In general, the outcomes that were selected by that group were consistent with our core outcome set. Our panel included additional safety outcomes (serious adverse events and development of bacterial resistance), which may have been considered of less importance for the eo-Drive trial as the safety profile of systemic corticosteroids has been thoroughly evaluated in previous studies. Moreover, disease progression, activities of daily living and quality of life were not prioritized for evaluation in the eo-Drive study either. The lack of validated instruments for assessing some of these outcomes in the context of an exacerbation trial may have discouraged the eo-Drive group. Moreover, the eo-Drive trial will recruit participants upon presentation with an exacerbation; therefore, assessment of disease progression is not possible. On the other hand, the multi-stakeholder involvement and rigorous methodological research may have allowed our panel to identify additional outcomes that may be more relevant to patients. For example, ADL were not captured in the longlist of outcomes assessed by the French group.

While this core outcome set and measurement instruments were developed for clinical trials on the management of COPD exacerbations, it would be important to be captured in relevant systematic reviews, meta-analyses and, also, observational studies. Their adoption in observational studies would enhance the comparability with trial results and interpretability of the complete body of available evidence. Finally, well-conducted observational studies could facilitate the validation and optimization of the measurement instruments recommended for each outcome. The Collaboration In COPD ExaceRbatiOns (CICERO) ERS Clinical Research

Collaboration has recently developed standards for clinical assessment, management and follow-up of acute hospitalised exacerbations of COPD [66]. These also include research recommendations, about outcomes that should be measured in relevant observational studies [66]. These largely overlap with the core outcomes that were prioritized by this panel. The CICERO panel also recommended the evaluation of new or worsening comorbidities following the index exacerbation event (such as diabetes or osteoporosis) and increase in short-acting inhaled therapy. On the other hand, activities of daily living, disease progression, development of resistant bacteria and development of pneumonia were not considered by that initiative. There was agreement between the two groups in all other outcomes. These differences may result from the different scope of the two projects as the COS-AECOPD ERS Task Force developed a core outcome set for clinical trials evaluating the management of COPD exacerbations, while CICERO developed standards for clinical practice evaluating the management of severe (hospitalized) COPD exacerbations, that were also recommended to be captured in clinical research studies.

Moreover, CICERO did not recommend measurement instruments; therefore, adopting recommendations from this task force, could improve comparability across the spectrum of clinical research on COPD exacerbations. However, CICERO did recommend the use of mMRC dyspnoea index and COPD assessment test for assessing symptoms during a hospitalized exacerbation. Our panel recommended the Borg's scale instead. mMRC was not considered sensitive in this setting, since most patients, especially those with severe exacerbations, would cluster in Grade 4 ("Too breathless to leave the house or breathless when dressing or undressing").

9.2 *Other challenges in the design of COPD exacerbations RCTs.*

Selection and measurement of outcomes are not the only challenges researchers face when designing clinical research on the management of COPD exacerbations. The diagnostic, classification and severity grading criteria of exacerbations remain ill-defined, subjective, and suboptimal, revealing an urgent unaddressed research need [6, 67, 68]. More specifically, it

is increasingly understood that exacerbations of different aetiology or characteristics (e.g. those caused by bacterial or viral infections, triggered by eosinophilic inflammation, or associated with type 2 respiratory failure), represent distinct clinical entities with different outcomes, that require personalized management [37, 69-71]. These distinctions should be made both in clinical practice and trials, however, adequately validated diagnostic tests are still lacking. Extensive, well-designed studies and international collaboration are needed to address these issues.

9.3 Justification of the protocol deviations

The methodology of this task force was prospectively published and transparent. However, on two occasions we had to deviate from the protocol. While we were planning on including a fourth stakeholder group in the Delphi survey, consisting of regulators, policymakers, guideline methodologists or those working in health technology assessment organizations, we did not manage to attract adequate responses to consider this group independently. However, this stakeholder group was represented in the consensus meetings. In addition, we had to change the threshold for excluding outcomes based on the results of the Delphi survey. Initially, we had planned on excluding outcomes that were considered non-critical by at least 50% of the Delphi survey participants from each stakeholder group. However, due to the coronavirus disease 19 (COVID-19) pandemic, we had to switch our planned face-to-face consensus meeting to two virtual meetings. Conducting virtual multi-stakeholder consensus meetings involving lay participants is challenging and time-consuming. Drawing on the experience amassed by the COMET initiative while facilitating similar, virtual consensus meetings during the pandemic, we decided to further consider during the consensus meetings only outcomes that had been rated as critical by at least one stakeholder group. This approach allowed a more thorough and constructive discussion and more confident consensus decisions for the outcomes that were considered. Moreover, none of the consensus meeting participants suggested that any of the other outcomes should have been considered.

9.4 Dissemination strategy

Uptake in future, relevant clinical trials is a crucial challenge for core outcome sets and for this reason, we have developed an implementation strategy. Firstly, we attempted to engage in the development of this core set all relevant stakeholders globally, through the Delphi survey and the consensus meetings. Moreover, the resulting set is endorsed by the ERS, adopted by the DECODE-NET (DisEntangling Chronic Obstructive pulmonary Disease Exacerbations – an international clinical trials NETwork) [72], and registered with the COMET Initiative. We intend to disseminate this document to clinical researchers with similar research interests and sponsors of COPD exacerbations trials, that completed the Delphi survey, or were identified through our methodological systematic reviews. The document will also be disseminated to relevant professional organizations, health technology assessment and guideline development groups, policymakers and regulators. Finally, a plain English description of this document will be shared with patient organizations and the lay participants of the Delphi survey and consensus group meetings.

10 References

1. Mathioudakis AG, Abroug F, Agusti A, Bakke P, Bartziokas K, Beghe B, Bikov A, Bradbury T, Brusselle G: **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, **In press**.
2. Williamson PR, Altman DG, Bagley H, Barnes KL, Blazeby JM, Brookes ST, Clarke M, Gargon E, Gorst S, Harman N, et al: **The COMET Handbook: version 1.0.** *Trials* 2017, **18:**280.
3. Kirkham JJ, Davis K, Altman DG, Blazeby JM, Clarke M, Tunis S, Williamson PR: **Core Outcome Set-STAndards for Development: The COS-STAD recommendations.** *PLoS Med* 2017, **14:**e1002447.
4. Kirkham JJ, Gorst S, Altman DG, Blazeby JM, Clarke M, Devane D, Gargon E, Moher D, Schmitt J, Tugwell P, et al: **Core Outcome Set-STAndards for Reporting: The COS-STAR Statement.** *PLoS Med* 2016, **13:**e1002148.
5. Mathioudakis AG, Moberg M, Janner J, Alonso-Coello P, Vestbo J: **Outcomes reported on the management of COPD exacerbations: a systematic survey of randomised controlled trials.** *ERJ Open Res* 2019, **5:**00072-02019.
6. Mathioudakis AG, Janner J, Moberg M, Alonso-Coello P, Vestbo J: **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:**00136-02019.
7. Gorst SL, Young B, Williamson PR, Wilding JPH, Harman NL: **Incorporating patients' perspectives into the initial stages of core outcome set development: a rapid review of qualitative studies of type 2 diabetes.** *BMJ Open Diabetes Res Care* 2019, **7:**e000615.
8. Zhang Y, Morgan RL, Alonso-Coello P, Wiercioch W, Bala MM, Jaeschke RR, Styczen K, Pardo-Hernandez H, Selva A, Ara Begum H, et al: **A systematic review of how patients value COPD outcomes.** *Eur Respir J* 2018, **52**.
9. Kessler R, Stahl E, Vogelmeier C, Haughney J, Trudeau E, Lofdahl CG, Partridge MR: **Patient understanding, detection, and experience of COPD exacerbations: an observational, interview-based study.** *Chest* 2006, **130:**133-142.
10. Kouijzer M, Brusse-Keizer M, Bode C: **COPD-related fatigue: Impact on daily life and treatment opportunities from the patient's perspective.** *Respir Med* 2018, **141:**47-51.
11. Rosa F, Bagnasco A, Ghirotto L, Rocco G, Catania G, Aleo G, Zanini M, Dasso N, Hayter M, Sasso L: **Experiences of older people following an acute exacerbation of chronic obstructive pulmonary disease: A phenomenological study.** *J Clin Nurs* 2018, **27:**e1110-e1119.
12. Dodd S, Clarke M, Becker L, Mavergames C, Fish R, Williamson PR: **A taxonomy has been developed for outcomes in medical research to help improve knowledge discovery.** *J Clin Epidemiol* 2018, **96:**84-92.
13. **DelphiManager. 2018.** [https://www.comet-initiative.org/delphimanager/index.html]
14. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schunemann HJ, Group GW: **GRADE: an emerging consensus on rating quality of evidence and strength of recommendations.** *BMJ* 2008, **336:**924-926.
15. Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G, Alderson P, Glasziou P, Falck-Ytter Y, Schunemann HJ: **GRADE guidelines: 2. Framing the question and deciding on important outcomes.** *J Clin Epidemiol* 2011, **64:**395-400.
16. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, de Vet HC: **The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study.** *Qual Life Res* 2010, **19:**539-549.
17. McGarvey LP, John M, Anderson JA, Zvarich M, Wise RA, Committee TCE: **Ascertainment of cause-specific mortality in COPD: operations of the TORCH Clinical Endpoint Committee.** *Thorax* 2007, **62:**411-415.
18. Drummond MB, Wise RA, John M, Zvarich MT, McGarvey LP: **Accuracy of death certificates in COPD: analysis from the TORCH trial.** *COPD* 2010, **7:**179-185.

19. McGarvey LP, Magder S, Burkhart D, Kesten S, Liu D, Manuel RC, Niewoehner DE: **Cause-specific mortality adjudication in the UPLIFT(R) COPD trial: findings and recommendations.** *Respir Med* 2012, **106:**515-521.
20. Zvezdin B, Milutinov S, Kojicic M, Hadnadjev M, Hromis S, Markovic M, Gajic O: **A postmortem analysis of major causes of early death in patients hospitalized with COPD exacerbation.** *Chest* 2009, **136:**376-380.
21. Seemungal TA, Donaldson GC, Bhowmik A, Jeffries DJ, Wedzicha JA: **Time course and recovery of exacerbations in patients with chronic obstructive pulmonary disease.** *Am J Respir Crit Care Med* 2000, **161:**1608-1613.
22. Aaron SD, Donaldson GC, Whitmore GA, Hurst JR, Ramsay T, Wedzicha JA: **Time course and pattern of COPD exacerbation onset.** *Thorax* 2012, **67:**238-243.
23. McAuley HJC, Harvey-Dunstan TC, Craner M, Richardson M, Singh SJ, Steiner MC, Greening NJ: **Longitudinal changes to quadriceps thickness demonstrate acute sarcopenia following admission to hospital for an exacerbation of chronic respiratory disease.** *Thorax* 2021, **76:**726-728.
24. Spruit MA, Gosselink R, Troosters T, Kasran A, Gayan-Ramirez G, Bogaerts P, Bouillon R, Decramer M: **Muscle force during an acute exacerbation in hospitalised patients with COPD and its relationship with CXCL8 and IGF-I.** *Thorax* 2003, **58:**752-756.
25. Anzueto A: **Impact of exacerbations on COPD.** *Eur Respir Rev* 2010, **19:**113-118.
26. Jeppesen E, Brurberg KG, Vist GE, Wedzicha JA, Wright JJ, Greenstone M, Walters JA: **Hospital at home for acute exacerbations of chronic obstructive pulmonary disease.** *Cochrane Database Syst Rev* 2012**:**CD003573.
27. Bai AD, Dai C, Srivastava S, Smith CA, Gill SS: **Risk factors, costs and complications of delayed hospital discharge from internal medicine wards at a Canadian academic medical centre: retrospective cohort study.** *BMC Health Serv Res* 2019, **19:**935.
28. Trethewey SP, Edgar RG, Turner AM, Mukherjee R: **Ward-Based Non-Invasive Ventilation in Acute Exacerbations of COPD: A Narrative Review of Current Practice and Outcomes in the UK.** *Healthcare (Basel)* 2018, **6**.
29. Davidson AC, Banham S, Elliott M, Kennedy D, Gelder C, Glossop A, Church AC, Creagh-Brown B, Dodd JW, Felton T, et al: **BTS/ICS guideline for the ventilatory management of acute hypercapnic respiratory failure in adults.** *Thorax* 2016, **71 Suppl 2:**ii1-35.
30. Jadad A, Rizo C, Cubillos P, Stahl E: **Measuring symptom response to pharmacological interventions in patients with COPD: a review of instruments used in clinical trials.** *Curr Med Res Opin* 2004, **20:**1993-2005.
31. Dorman S, Byrne A, Edwards A: **Which measurement scales should we use to measure breathlessness in palliative care? A systematic review.** *Palliat Med* 2007, **21:**177-191.
32. Oliveira AL, Marques AS: **Outcome Measures Used in Pulmonary Rehabilitation in Patients With Acute Exacerbation of Chronic Obstructive Pulmonary Disease: A Systematic Review.** *Phys Ther* 2018, **98:**191-204.
33. Jones PW, Harding G, Berry P, Wiklund I, Chen WH, Kline Leidy N: **Development and first validation of the COPD Assessment Test.** *Eur Respir J* 2009, **34:**648-654.
34. Weldam SW, Schuurmans MJ, Liu R, Lammers JW: **Evaluation of Quality of Life instruments for use in COPD care and research: a systematic review.** *Int J Nurs Stud* 2013, **50:**688-707.
35. Gupta N, Pinto LM, Morogan A, Bourbeau J: **The COPD assessment test: a systematic review.** *Eur Respir J* 2014, **44:**873-884.
36. Zhou Z, Zhou A, Zhao Y, Chen P: **Evaluating the Clinical COPD Questionnaire: A systematic review.** *Respirology* 2017, **22:**251-262.
37. Bafadhel M, McKenna S, Terry S, Mistry V, Reid C, Haldar P, McCormick M, Haldar K, Kebadze T, Duvoix A, et al: **Acute exacerbations of chronic obstructive pulmonary disease: identification of biologic clusters and their biomarkers.** *Am J Respir Crit Care Med* 2011, **184:**662-671.

38. Janaudis-Ferreira T, Beauchamp MK, Robles PG, Goldstein RS, Brooks D: **Measurement of activities of daily living in patients with COPD: a systematic review.** *Chest* 2014, **145:**253-271.
39. Liu Y, Li H, Ding N, Wang N, Wen D: **Functional Status Assessment of Patients With COPD: A Systematic Review of Performance-Based Measures and Patient-Reported Measures.** *Medicine (Baltimore)* 2016, **95:**e3672.
40. **International Classification of Functioning, Disability and Health (ICF, updated 2018)** [https://www.who.int/standards/classifications/international-classification-of-functioning-disability-and-health]
41. Partridge MR, Miravitlles M, Stahl E, Karlsson N, Svensson K, Welte T: **Development and validation of the Capacity of Daily Living during the Morning questionnaire and the Global Chest Symptoms Questionnaire in COPD.** *Eur Respir J* 2010, **36:**96-104.
42. Yohannes AM, Roomi J, Winn S, Connolly MJ: **The Manchester Respiratory Activities of Daily Living questionnaire: development, reliability, validity, and responsiveness to pulmonary rehabilitation.** *J Am Geriatr Soc* 2000, **48:**1496-1500.
43. Morimoto M, Takai K, Nakajima K, Kagawa K: **Development of the chronic obstructive pulmonary disease activity rating scale: reliability, validity and factorial structure.** *Nurs Health Sci* 2003, **5:**23-30.
44. Chen YJ, Narsavage GL, Culp SL, Weaver TE: **The development and psychometric analysis of the short-form Pulmonary Functional Status Scale (PFSS-11).** *Res Nurs Health* 2010, **33:**477-485.
45. Decramer M, Gosselink R, Rutten-Van Molken M, Buffels J, Van Schayck O, Gevenois PA, Pellegrino R, Derom E, De Backer W: **Assessment of progression of COPD: report of a workshop held in Leuven, 11-12 March 2004.** *Thorax* 2005, **60:**335-342.
46. Tabberer M, Gonzalez-McQuire S, Muellerova H, Briggs AH, Rutten-van Molken M, Chambers M, Lomas DA: **Development of a Conceptual Model of Disease Progression for Use in Economic Modeling of Chronic Obstructive Pulmonary Disease.** *Med Decis Making* 2017, **37:**440-452.
47. Zemans RL, Jacobson S, Keene J, Kechris K, Miller BE, Tal-Singer R, Bowler RP: **Multiple biomarkers predict disease severity, progression and mortality in COPD.** *Respir Res* 2017, **18:**117.
48. Exuzides A, Colby C, Briggs AH, Lomas DA, Rutten-van Molken M, Tabberer M, Chambers M, Muellerova H, Locantore N, Risebrough NA, et al: **Statistical Modeling of Disease Progression for Chronic Obstructive Pulmonary Disease Using Data from the ECLIPSE Study.** *Med Decis Making* 2017, **37:**453-468.
49. Young AL, Bragman FJS, Rangelov B, Han MK, Galban CJ, Lynch DA, Hawkes DJ, Alexander DC, Hurst JR, Investigators CO: **Disease Progression Modeling in Chronic Obstructive Pulmonary Disease.** *Am J Respir Crit Care Med* 2020, **201:**294-302.
50. Halper-Stromberg E, Yun JH, Parker MM, Singer RT, Gaggar A, Silverman EK, Leach S, Bowler RP, Castaldi PJ: **Systemic Markers of Adaptive and Innate Immunity Are Associated with Chronic Obstructive Pulmonary Disease Severity and Spirometric Disease Progression.** *Am J Respir Cell Mol Biol* 2018, **58:**500-509.
51. Grace J, Leader JK, Nouraie SM, Pu J, Chandra D, Zhang Y, Sciurba FC, Bon J: **Mediastinal and Subcutaneous Chest Fat Are Differentially Associated with Emphysema Progression and Clinical Outcomes in Smokers.** *Respiration* 2017, **94:**501-509.
52. Dreyse J, Diaz O, Repetto PB, Morales A, Saldias F, Lisboa C: **Do frequent moderate exacerbations contribute to progression of chronic obstructive pulmonary disease in patients who are ex-smokers?** *Int J Chron Obstruct Pulmon Dis* 2015, **10:**525-533.
53. Musuamba FT, Teutonico D, Maas HJ, Facius A, Yang S, Danhof M, Della Pasqua O: **Prediction of disease progression, treatment response and dropout in chronic obstructive pulmonary disease (COPD).** *Pharm Res* 2015, **32:**617-627.
54. Coxson HO, Dirksen A, Edwards LD, Yates JC, Agusti A, Bakke P, Calverley PM, Celli B, Crim C, Duvoix A, et al: **The presence and progression of emphysema in COPD as determined by CT**

scanning and biomarker expression: a prospective analysis from the ECLIPSE study. *Lancet Respir Med* 2013, **1:**129-136.

55. Park HY, Man SF, Tashkin D, Wise RA, Connett JE, Anthonisen NA, Sin DD: **The relation of serum myeloperoxidase to disease progression and mortality in patients with chronic obstructive pulmonary disease (COPD).** *PLoS One* 2013, **8:**e61315.
56. Galban CJ, Han MK, Boes JL, Chughtai KA, Meyer CR, Johnson TD, Galban S, Rehemtulla A, Kazerooni EA, Martinez FJ, Ross BD: **Computed tomography-based biomarker provides unique signature for diagnosis of COPD phenotypes and disease progression.** *Nat Med* 2012, **18:**1711-1715.
57. Tanabe N, Muro S, Hirai T, Oguma T, Terada K, Marumo S, Kinose D, Ogawa E, Hoshino Y, Mishima M: **Impact of exacerbations on emphysema progression in chronic obstructive pulmonary disease.** *Am J Respir Crit Care Med* 2011, **183:**1653-1659.
58. Ash SY, San Jose Estepar R, Fain SB, Tal-Singer R, Stockley RA, Nordenmark LH, Rennard S, Han MK, Merrill D, Humphries SM, et al: **Relationship between Emphysema Progression at CT and Mortality in Ever-Smokers: Results from the COPDGene and ECLIPSE Cohorts.** *Radiology* 2021, **299:**222-231.
59. Anzueto A, Leimer I, Kesten S: **Impact of frequency of COPD exacerbations on pulmonary function, health status and clinical outcomes.** *Int J Chron Obstruct Pulmon Dis* 2009, **4:**245-251.
60. Kerkhof M, Voorham J, Dorinsky P, Cabrera C, Darken P, Kocks JW, Sadatsafavi M, Sin DD, Carter V, Price DB: **Association between COPD exacerbations and lung function decline during maintenance therapy.** *Thorax* 2020, **75:**744-753.
61. Jones PW, Agusti AG: **Outcomes and markers in the assessment of chronic obstructive pulmonary disease.** *Eur Respir J* 2006, **27:**822-832.
62. Shaw JG, Vaughan A, Dent AG, O'Hare PE, Goh F, Bowman RV, Fong KM, Yang IA: **Biomarkers of progression of chronic obstructive pulmonary disease (COPD).** *J Thorac Dis* 2014, **6:**1532-1547.
63. **International Conference on Harmonisation of Technical Requirments for Registration of Pharmaceuticals for Human Use. ICH Harmonised Tripartite Guideline. Clinical Safety Data Management: Definitions and Standards for Expedited Reporting** [https://database.ich.org/sites/default/files/E2A_Guideline.pdf]
64. Mackay AJ, Kostikas K, Murray L, Martinez FJ, Miravitlles M, Donaldson G, Banerji D, Patalano F, Wedzicha JA: **Patient-reported Outcomes for the Detection, Quantification, and Evaluation of Chronic Obstructive Pulmonary Disease Exacerbations.** *Am J Respir Crit Care Med* 2018, **198:**730-738.
65. Suehs CM, Zysman M, Chenivesse C, Burgel PR, Couturaud F, Deslee G, Berger P, Raherison C, Devouassoux G, Brousse C, et al: **Prioritising outcomes for evaluating eosinophil-guided corticosteroid therapy among patients with acute COPD exacerbations requiring hospitalisation: a Delphi consensus study.** *BMJ Open* 2020, **10:**e035811.
66. Ramakrishnan S, Janssens W, Burgel PR, Contoli M, Franssen FME, Greening NJ, Greulich T, Gyselinck I, Halner A, Huerta A, et al: **Standardisation of Clinical Assessment, Management and Follow-Up of Acute Hospitalised Exacerbation of COPD: A Europe-Wide Consensus.** *Int J Chron Obstruct Pulmon Dis* 2021, **16:**321-332.
67. Rodriguez-Roisin R: **Toward a consensus definition for COPD exacerbations.** *Chest* 2000, **117:**398S-401S.
68. Anthonisen NR, Manfreda J, Warren CP, Hershfield ES, Harding GK, Nelson NA: **Antibiotic therapy in exacerbations of chronic obstructive pulmonary disease.** *Ann Intern Med* 1987, **106:**196-204.
69. Papi A, Bellettato CM, Braccioni F, Romagnoli M, Casolari P, Caramori G, Fabbri LM, Johnston SL: **Infections and airway inflammation in chronic obstructive pulmonary disease severe exacerbations.** *Am J Respir Crit Care Med* 2006, **173:**1114-1121.

70. Mathioudakis AG, Chatzimavridou-Grigoriadou V, Corlateanu A, Vestbo J: **Procalcitonin to guide antibiotic administration in COPD exacerbations: a meta-analysis.** *Eur Respir Rev* 2017, **26:**160073.
71. Mathioudakis AG, Janssens W, Sivapalan P, Singanayagam A, Dransfield MT, Jensen JUS, Vestbo J: **Acute exacerbations of chronic obstructive pulmonary disease: In search of diagnostic biomarkers and treatable traits.** *Thorax* 2020, **In press**.
72. Mathioudakis AG, Sivapalan P, Papi A, Vestbo J, on behalf of the DECODE-NET Investigators: **The DisEntangling Chronic Obstructive pulmonary Disease Exacerbations clinical trials NETwork: DECODE-NET. Rationale and vision.** *Eur Respir J* 2020, **In press**.