A systematic review of how patients value COPD outcomes

Yuan Zhang[1], Rebecca L. Morgan[1], Pablo Alonso-Coello[1,2], Wojtek Wiercioch[1], Małgorzata M. Bała[3], Rafał R. Jaeschke[4], Krzysztof Styczeń[4], Hector Pardo-Hernandez[2], Anna Selva[5,6], Housne Ara Begum[1], Gian Paolo Morgano[1], Marcin Waligóra[7], Arnav Agarwal[1,8], Matthew Ventresca[1], Karolina Strzebońska[7], Mateusz T. Wasylewski[7], Lídia Blanco-Silvente[9], Janna-Lina Kerth[10], Mengxiao Wang[1], Yuqing Zhang[1], Saiprasad Narsingam[11], Yutong Fei[12], Gordon Guyatt[1,13] and Holger J. Schünemann[1,13]

Affiliations: [1]Dept of Health Research Methods, Evidence and Impact, McMaster University, Hamilton, ON, Canada. [2]Iberoamerican Cochrane Centre, CIBERESP-IIB Sant Pau, Barcelona, Spain. [3]Dept of Hygiene and Dietetics, Faculty of Medicine, Jagiellonian University Medical College, Krakow, Poland. [4]Section of Affective Disorders, Dept of Psychiatry, Jagiellonian University Medical College, Krakow, Poland. [5]Clinical Epidemiology and Cancer Screening, Corporació Sanitària Parc Taulí, Sabadell, Spain. [6]Research Network on Health Services in Chronic Diseases (REDISSEC), Spain. [7]REMEDY, Research Ethics in Medicine Study Group, Dept of Philosophy and Bioethics, Jagiellonian University Medical College, Krakow, Poland. [8]School of Medicine, University of Toronto, Toronto, ON, Canada. [9]TransLab Research Group, Dept of Medical Sciences, University of Girona, Girona, Spain. [10]Dept for Medical Didactics and Curricular Development, Medical Faculty RWTH Aachen University, Aachen, Germany. [11]Dept of Medicine, Dartmouth Medical School, Dartmouth-Hitchcock Medical Center, Lebanon, NH, USA. [12]Centre for Evidence-Based Chinese Medicine, Beijing University of Chinese Medicine, Beijing, China. [13]Dept of Medicine, McMaster University, Hamilton, ON, Canada.

Correspondence: Holger J Schünemann, Dept of Health Research Methods, Evidence and Impact, McMaster University, 1280 Main Street W., Hamilton, ON, L8S 4K1, Canada. E-mail: schuneh@mcmaster.ca

@ERSpublications
Systematic review of the importance placed by patients on COPD outcomes informs the trade-off between benefits and harms http://ow.ly/l5De30kgD9g

Cite this article as: Zhang Y, Morgan RL, Alonso-Coello P, *et al.* A systematic review of how patients value COPD outcomes. *Eur Respir J* 2018; 52: 1800222 [https://doi.org/10.1183/13993003.00222-2018].

ABSTRACT Our objective was to summarise systematically all research evidence related to how patients value outcomes in chronic obstructive pulmonary disease (COPD).

We conducted a systematic review (systematic review registration number CRD42015015206) by searching PubMed, Embase, PsycInfo and CINAHL, and included reports that assessed the relative importance of outcomes from COPD patients' perspective. Two authors independently determined the eligibility of studies, abstracted the eligible studies and assessed risk of bias. We narratively summarised eligible studies, meta-analysed utilities for individual outcomes and assessed the certainty of evidence using the Grading of Recommendations, Assessment, Development and Evaluations approach.

We included 217 quantitative studies. Investigators most commonly used utility measurements of outcomes (n=136), discrete choice exercises (n=13), probability trade-off (n=4) and forced choice techniques (n=46). Patients rated adverse events as important but on average, less so than symptom relief. Exacerbation and hospitalisation due to exacerbation are the outcomes that COPD patients rate as most important. This systematic review provides a comprehensive registry of related studies.

This article has supplementary material available from erj.ersjournals.com

Received: Jan 21 2018 | Accepted after revision: May 21 2018

Introduction

Considering patient values and preferences regarding the benefits and harms of a health intervention is essential for clinical evidence-based decision-making [1–4]. The Grading of Recommendations Assessment, Development and Evaluation (GRADE) working group has recently operationalised patient values and preferences as "the relative importance patients place on outcomes" [3, 5]. Information about the relative importance of outcomes is critical to weigh the health benefits and harms of interventions and test strategies [5], including those recommended in clinical practice guidelines. Indeed, numerous studies have addressed how patients value chronic obstructive pulmonary disease (COPD) outcomes but to appropriately inform practice and guidelines, this evidence should be summarised in systematic reviews that allow retrieving and summarising the best evidence from individual studies on health outcomes [2, 6–9]. Considering the disease burden of COPD [10], such a review would inform decision-making for a large patient community globally. We therefore conducted this systematic review to summarise all research evidence that addressed the question, what is the relative importance patients place on COPD-related outcomes [3, 5]?

Methods

Protocol and registration

We conducted this systematic review of the literature in accordance with the Preferred Reporting in Systematic Reviews and Meta-Analyses guidelines [11] and registered the review protocol on PROSPERO (registration number CRD42015015206).

Information sources

We searched Medline (through PubMed), Embase, PsycInfo and CINAHL from inception date to October 15, 2017, using an extensive search strategy developed for retrieving this type of evidence (supplementary material) [12], including reference lists of identified studies.

Study selection

Two authors independently determined the eligibility of studies by reviewing titles and abstracts and, for potentially eligible studies, through review of full-text articles with a standardised and piloted screening form. Reviewers resolved disagreement by discussion or through third-party adjudication. Eligible studies reported patient values and preferences of COPD patients, with no limits on the type of study design, language or treatments. Studies with the following characteristics were eligible for reporting the relative importance of outcomes [4].

1) Patient utility and health state value studies: studies that examined how patients value alternative health states and experiences with treatment. The eligible measurement techniques were: standard gamble, time trade-off (TTO), visual analogue scale (VAS), or mapping results based on either generic (EuroQol-5D (EQ-5D) or SF-36) [13] or specific measurement (*i.e.* Chronic Respiratory Questionnaire) of health-related quality of life. We expected one major category of eligible studies to be "utility" studies. Utilities represent the strength of an individual's preferences for different outcomes. They are expressed on a scale from 0 indicating dead to 1 indicating perfect health (for some variations of the scale, the upper bound may be 100). The higher the utility is (the closer the estimate is to perfect health), the more value patients will place on the outcome.

2) Direct choice studies: studies that examined patients' choice when they were presented with a description of hypothetical states or during decision making for their own actual health states (*i.e.* forced choice when presented with a decision aid, probabilistic trade-off techniques, discrete choice, willingness to pay, randomised controlled trials (RCTs) for preferences, *etc.*).

3) Other quantitative studies on outcome importance: studies that quantitatively examined the patients' views, attitudes or preferences on outcome importance through self-developed questionnaires or instruments that were not utility measurement techniques.

We included only quantitative studies reporting COPD as a comorbidity if they reported COPD relative importance of outcomes information separately. We excluded non-original studies such as clinical practice guidelines, reviews, commentaries, letters or viewpoints. We also excluded case reports, case series and health economic evaluation studies without original utility elicitation. Qualitative studies that explored patients' views, attitudes or preferences related to different treatment options were excluded from this review but included and reported in a subsequent review.

Data collection and certainty of evidence

Two authors independently recorded data: principal author, publication year, participant demographics (sample size, age, sex, *etc.*), survey techniques or methodologies used, relative importance of outcome results and risk of bias assessments.

Since there is no accepted risk of bias or study quality assessment tool for value and preference studies, we used an approach that we developed, validated and reported in a separate project [14]. The key items to assess the risk of bias include sample selection, response rate (or attrition rate if participants were followed up), choice and administration of the instrument, outcome (or health state) presentation, participants' understanding of the methodology, and data analysis (if applicable). We then used the GRADE approach to rate the certainty of the overall body of evidence for outcome importance [14, 15]. The GRADE approach classifies certainty of evidence as high, moderate, low or very low based on domains of risk of bias, inconsistency, indirectness, imprecision, publication bias and upgrading domains.

Data analysis

A priori, we set the disease severity following the Global Initiative for Chronic Obstructive Lung Disease (GOLD) criteria, indicating the severity of airflow obstruction, as a potential subgroup factor to consider [16]. We used the severity of airflow obstruction categories of mild (forced expiratory volume in 1 s (FEV_1) ⩾80% predicted), moderate (FEV_1 ⩾50% to <80% predicted), severe (FEV_1 ⩾30% to <50% predicted) and very severe (FEV_1 <30% predicted) reported by authors to determine subgroups. Information on the relative importance of outcomes exists in a variety of formats, including the utility of outcomes or disease stages, proportion of choice, rankings or scores on a scale. For the sake of simplicity, we report all estimates using the descriptive term "utility" to indicate the health status values elicited from standard gamble, TTO, VAS and results from indirect utility measurements [17]. We conducted meta-analyses to synthesise the utility results for same outcomes using a random-effects inverse variance method in Stata 11.0 (StatCorp, College Station, TX, USA) [18]. For consistency, we presented the results on a 0–1 scale even if they had been elicited on a 0–100 scale. For nonutility results regarding patient values and preferences, we narratively summarised the results.

Results

Study selection and study characteristics

Of 54 598 records, after excluding duplicates, 41 781 titles and abstracts remained; 3154 articles proved potentially eligible and underwent full-text screening. Of these, 217 quantitative studies reporting patient values and preferences on COPD outcomes proved eligible (figure 1 and references of all included studies in the supplementary material).

Of the 217 eligible studies, 136 reported utility or health state values for COPD outcomes, of which: 69 utilised the feeling thermometer or VAS, including the EQ-5D VAS; eight the standard gamble; and six the TTO. For indirect measurements, 82 studies reported EQ-5D utilities, 14 SF-6D utilities, seven Health Utility Index (HUI), seven 15D and three the Quality of Well-Being utilities. Of 65 direct choice studies, 46 used forced choice techniques, 13 discrete choice exercise/conjoint analysis or willingness to pay, four probability trade-off, and three ranking methods (supplementary table 1). Regarding study design, 127 were cross-sectional studies, 21 cohort studies, 11 repeated surveys, 51 RCTs and seven quasirandomised trials.

The outcomes studied typically included exacerbation or hospitalisation due to exacerbation, adverse events, symptom relief, and different severities of COPD. Table 1 summarised this type of evidence (also see supplementary table 1). Despite the large number of eligible studies, few reported the relative importance of outcome information on the same outcomes. Meta-analyses were restricted to studies focusing on exacerbation, and different COPD severities measured with VAS and EQ-5D utility. We found no compelling evidence of publication bias.

Supplementary table 2 summarises the risk of bias assessment. Studies suffered from serious risk of bias related to limitations in the validity and reliability of the measurement tools (68 studies directly asking participants to choose among a set of options) and use of a convenience sampling strategy or a volunteer sample (14 studies); or response rates <50% (32 studies). For other risk of bias considerations, we classified most studies as low risk of bias (supplementary table 2).

Importance of exacerbation

The measurements used to elicit the importance of exacerbation or hospitalisation due to exacerbation include VAS (including the EQ-5D VAS) (10 studies [19–27, 47]). TTO (one study [27]) and the EQ-5D utility (seven studies [21–24, 39, 40, 43]). We observed variations in the description of "exacerbation". Three studies utilised clinical diagnoses, without a specific definition of an exacerbation [24, 43, 47]. All remaining studies defined exacerbation as worsened symptoms [19–27, 39, 40, 43, 47, 48]. Of these, three studies explicitly reported a category for exacerbation, with one using the definition of the British Thoracic Society [21] and the two others the American Thoracic Society/European Respiratory Society definition [23, 40]. The other reports varied in defining "exacerbation," with three reports focusing on exacerbations needing hospitalisation [19, 20, 26] and another one specifying the length of symptoms [22].

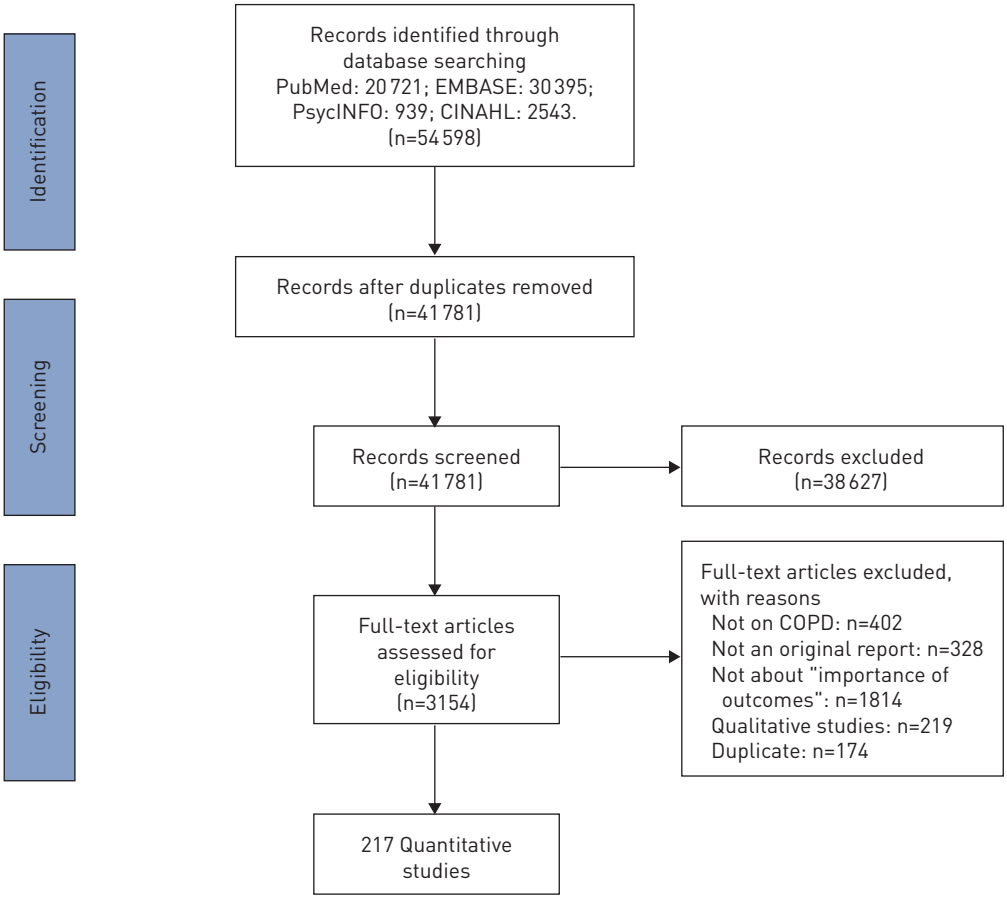

FIGURE 1 Flow diagram for systematic review on chronic obstructive pulmonary disease patients' values and preferences.

The estimates varied from 0.259 to 0.580 on the VAS, and 0.430 to 0.740 with the EQ-5D utility. We conducted meta-analysis using the inverse variance method to pool the estimates based on VAS and EQ-5D, yielding utility of exacerbation of 0.462 (95% CI 0.453–0.471, I^2=98.2%, p<0.001 for the test of heterogeneity) on the VAS and 0.519 (95% CI 0.502–0.537, I^2=95.5%, p<0.001 for the test of heterogeneity) with the EQ-5D utility. Of the eight studies included in the meta-analysis, six recruited patient populations with a mean age between 66 and 69 years [19–21, 24–26]; of those, four were from the UK [21, 24–26], two from other European countries [19, 20], one from the USA [22], and another was a multicentre study conducted in countries including the USA, UK and other countries [23]. The study populations were similar across the studies regarding age and setting. We could not explain the large degree of inconsistency and, thus, rated down the certainty of evidence as moderate (table 1). For studies that used the EQ-5D utility measurement, we further rated down for indirectness given the indirect measurement tool used (*i.e.* the patients participating did not themselves place a value on exacerbations, but merely reported the consequences on EQ-5D items). One study used a more granular approach to addressing the importance of exacerbations: Rutten van Molken *et al.* [27] reported the values of different severities of exacerbations. The authors described serious and nonserious exacerbations according to the severity of increase in respiratory and nonrespiratory symptoms, impact on daily activities, and response to treatment. To summarise, for a nonserious exacerbation, patients will experience mild-to-moderate worsening of breathlessness and cough, and the symptoms interfere with daily activities; while patients with a serious exacerbation will experience severe-to-very severe worsening of breathlessness and cough, and the symptoms will completely disrupt daily activities. Based on VAS and TTO measurements, respectively, the disutility (defined as a reduction in utility) for one nonserious exacerbation was 0.037 (VAS) and 0.010 (TTO); for two nonserious exacerbations, 0.068 and 0.021; for one serious exacerbation, 0.090 and 0.042; for one serious exacerbation and one nonserious, 0.130 and 0.088 (table 2). The certainty of this evidence is high. Other studies suggested patients have lower utility as the exacerbations became more frequent or more severe [40, 43].

TABLE 1 Summary of findings

Health state/outcome (categories of values and preferences)	Estimates of outcome importance range across studies/pooled mean (95% CI)	Participants/ studies	Certainty of evidence	Interpretation of findings
Exacerbation (utility[#] measured with visual analogue scale[¶])	0.259–0.580/0.462 (0.453–0.471)[####]	1991/8[####]	+++– Moderate certainty due to inconsistency[####]	Most people find exacerbation of COPD probably has a large impact on their lives. There is probably no important variability for this assessment.
Exacerbation (EQ-5D utility[+])	0.430–0.683/0.519 (0.502–0.537)[¶¶¶¶]	927/3[¶¶¶¶]	++–– Low certainty due to inconsistency and indirectness[¶¶¶¶.######]	Most people find exacerbation of COPD probably has a large impact on their lives. There is probably no important variability for this assessment.
Exacerbation (disutility)[§]	Visual analogue scale 1 nonserious exacerbation: −0.037 (0.005) 2 nonserious exacerbations: −0.068 (0.005) 1 serious exacerbation: −0.090 (0.007) 1 nonserious and 1 serious exacerbation: −0.130 (0.007) Time trade off 1 nonserious exacerbation: −0.010 (0.007) 2 nonserious exacerbations: −0.021 (0.007) 1 serious exacerbation: −0.042 (0.009) 1 nonserious and 1 serious exacerbation: −0.088 (0.009)	239/1	++++ High certainty	Most people find exacerbation of COPD has an impact on their lives, which grows larger as the severity of exacerbation progresses. There is probably no important variability for this assessment.
Level 1 dyspnoea (utility measured with visual analogue scale)[f]	0.751	146/1[f]	++++ High certainty[f]	Most people find level 1 dyspnoea has a small to moderate impact on lives. There is probably no important variability for this assessment.
Level 2 dyspnoea (utility measured with visual analogue scale)[f]	0.656	45/1[f]	+++– Moderate certainty due to imprecision[f]	Most people find level 2 dyspnoea probably has a moderate impact on lives. There is probably no important variability for this assessment.
Level 3 dyspnoea (utility measured with visual analogue scale)[f]	0.526	7/1[f]	++–– Low certainty due to very serious imprecision[f]	Most people find level 3 dyspnoea probably has a large impact on lives. There is probably no important variability for this assessment.
Adverse events (discrete choice)[##]	Two studies suggested that patients consider adverse events as important outcomes. One study suggested adverse events were more important than onset time of medicine, ease of use, rescue medicine use. Another suggested adverse events were more important than costs of treatment, extent to which the patient sees the same doctor each time and extent to which the doctor treats the patient as an entire person. Both studies concluded symptom relief to be more important than adverse events.	564/2	+++– Moderate certainty due to risk of bias[##]	People probably consider adverse events as an important outcome. There is likely no important variability for this assessment.
Extent of symptom relief (discrete choice)[##]	Two studies compared extend of symptom relief with other outcomes. Extent of symptom relief was considered the most important outcome in these two studies.	564/2	+++– Moderate certainty due to risk of bias[##]	Most people probably find symptom relief as important outcome. There is probably no important variability for this assessment.
Extent of symptom relief (forced choice)[¶¶]	In a survey on expectation of treatment, greater symptomatic relief was chosen by 82.3% of the participants, thus the most important outcome. Extent of symptom relief was considered the second most important outcome in one cross-sectional study (less preferred "not to be kept alive on life support when there is little hope for a meaningful recovery"). Another study reported 58.0% of the participants would prefer treatment focusing on relieving pain and discomfort rather than extending life.	1640/3	+++– Moderate certainty due to risk of bias[¶¶]	Most people probably find symptom relief as important outcome. There is probably no important variability for this assessment.
Very severe COPD (utility measured with visual analogue scale)[++]	0.321–0.651/0.345 (0.335–0.354)[++++]	746/7	++–– Low certainty due to risk of bias[++++++] and inconsistency[++]	Most people find very severe COPD seems to have a large impact on lives. There is probably important variability for this assessment.
Very severe COPD (EQ-5D utility)[§§]	0.520–0.740/0.681 (0.667–0.694)[§§§§]	898/10	+++– Moderate certainty due to inconsistency[§§§§]	Most people find very severe COPD probably has a large impact on lives. There is probably important variability for this assessment.
Severe COPD (utility measured with visual analogue scale)[ff]	0.446–0.689/0.508 (0.501–0.515)[ffff]	4683/8	++–– Low certainty due to risk of bias[++++++] and inconsistency [ffff]	Most people find severe COPD probably has a moderate to large impact on lives. There is probably important variability for this assessment.

Continued

TABLE 1 Continued

Health state/outcome (categories of values and preferences)	Estimates of outcome importance range across studies/pooled mean (95% CI)	Participants/ studies	Certainty of evidence	Interpretation of findings
Severe COPD (EQ-5D utility)[###]	0.620–0.810/0.741 (0.734–0.749)[#####]	4352/11	+++– Moderate certainty due to inconsistency[#####]	Most people find severe COPD probably has a moderate to large impact on lives. There is probably important variability for this assessment.
Moderate COPD (utility measured with visual analogue scale)[¶¶¶]	0.589–0.726/0.639 (0.635–0.642)[¶¶¶¶¶]	9664/10	++–– Low certainty due to risk of bias[++++++] and inconsistency[¶¶¶¶¶]	Most people find moderate COPD probably has a moderate impact on lives. There is probably important variability for this assessment.
Moderate COPD (EQ-5D utility)[+++]	0.680–0.890/0.821 (0.815–0.826)[+++++]	4620/9	+++– Moderate certainty due to inconsistency[+++++]	Most people find moderate COPD probably has a moderate impact on lives. There is probably important variability for this assessment.
Mild COPD (utility measured with visual analogue scale)[§§§]	0.680–0.811/0.738 (0.732-0.746)[§§§§§]	3623/8	++–– Low certainty due to risk of bias[++++++] and inconsistency[§§§§§]	Most people find moderate COPD probably has a small to moderate impact on lives. There is likely important variability for this assessment.
Mild COPD (EQ-5D utility)[fff]	0.770–0.900/0.873 (0.863–0.883)[fffff]	2067/7	+++– Moderate certainty due to inconsistency [fffff]	Most people find moderate COPD probably has a small to moderate impact on lives. There is probably important variability for this assessment.

Grading of Recommendations Assessment, Development and Evaluation Working Group grades of evidence: here we assessed the certainty of evidence on mean outcome importance. We use "certainty of evidence", "certainty in estimates", "quality of evidence" and "strength of evidence" interchangeably. High certainty: we are very confident that the true value of outcome importance lies close to that of the estimate. Moderate certainty: we are moderately confident in the estimate; the true value of outcome importance is likely to be close to the estimate but there is a possibility that it is substantially different. Low certainty: our confidence in the estimate is limited; the true value of outcome importance may be substantially different from the estimate. Very low certainty: we have very little confidence in the estimate; the true value of outcome importance is likely to be substantially different from the estimate EQ-5D: EuroQol-5D; COPD: chronic obstructive pulmonary disease. [#]: utilities represent the value individuals place on different outcomes; they are measured on an interval scale, with 0 reflecting states of health equivalent to death/worst imaginable health and 1 (or 100 in some cases) reflecting perfect health/best imaginable health. [¶]: eight studies [19–26] used EQ-5D visual analogue scale to elicit health state values on exacerbation of COPD. [+]: three studies [21–23] used EQ-5D utility to elicit the importance of outcome. [§]: Rutten van Molken *et al.* [27] reported the disutility due to the exacerbations; the measurement tools included visual analogue scale and time trade off; the researchers estimated the disutility due to exacerbation using random effects regression analysis. [*f*]: Kim *et al.* [28] reported the utility of dyspnoea, according to the levels of breathlessness (level 1, short of breath during strenuous activities; level 2, stopping to catch breath after a few minutes walking; level 3, breathless when dressing or washing); in a total sample of 200, the numbers of participants experiencing level 1, 2 and 3 breathlessness were 146, 45 and seven, respectively; due to small sample size, we downgraded the certainty of evidence by one level for the estimates of level 2 breathlessness and two levels for level 3 breathlessness. [##]: Bulcun *et al.* [29] compared extent of symptom relief with extent to which the doctor gives sufficient time to listen to the patient, possibility of experiencing adverse effects from treatment, costs of treatment, extent to which the patient sees the same doctor each time and extent to which the doctor treats the patient as an entire person; Kawata *et al.* [30] recruited 515 patients for an online voluntary survey on the comparison of importance of symptom relief, speed of symptom relief, rescue medicine use and side-effects; participants' eligibility and their answers were considered as having serious risk of bias. [¶¶]: three studies [31–33] asked directly what participants would prefer in facing a COPD treatment decision; the questions included expectation of treatment, reasons to continue or not continue with treatment and preferred treatment characteristics; the assessment was at risk of bias due to unclear reliability and validity features. [++]: [27, 28, 34–37]. [§§]: [27, 28, 35–42]. [*ff*]: [27, 28, 34–37, 43, 44]. [###]: [28, 35–40, 43, 44]. [¶¶¶]: [27, 28, 34–37, 43, 44–46]. [§§§]: [27, 28, 34–37, 43, 45]. [*fff*]: [28, 35–37, 42, 43, 45]. [####]: across eight included studies, the point estimates range from 0.259 to 0.580; using inverse-variance method to pool the estimates, the I^2 (95.5%) and statistical test (<0.001) suggest potential heterogeneity across studies; the difference in study populations cannot explain the source of heterogeneity (the participants in the studies were exacerbation patients, exacerbation patients not needing hospitalisation, ambulatory patients and hospitalised patients due to exacerbation). [¶¶¶¶]: across three included studies, the point estimates range from 0.430 to 0.683; using inverse-variance method to pool the estimates, the I^2 (95.5%) and statistical test (<0.001) suggest potential heterogeneity across studies; the difference in study populations cannot explain the source of heterogeneity (the participants in the three studies were exacerbation patients, exacerbation patients not needing hospitalisation and ambulatory patients). [++++]: the point estimates range from 0.321 to 0.651; using inverse-variance method to pool the estimates, the I^2 (98.5%) and statistical test (<0.001) suggest potential heterogeneity across studies; the difference in study population cannot explain the source of heterogeneity. [§§§§]: the point estimates range from 0.520 to 0.740; using inverse-variance method to pool the estimates, the I^2 (80.2%) and statistical test (<0.001) suggest potential heterogeneity across studies; the difference in study population cannot explain the source of heterogeneity. [*ffff*]: the point estimates range from 0.446 to 0.689; using inverse-variance method to pool the estimates, the I^2 (98.8%) and statistical test (<0.001) suggest potential heterogeneity across studies; the difference in study population cannot explain the source of heterogeneity. [#####]: the point estimates range from 0.620 to 0.810; using inverse-variance method to pool the estimates, the I^2 (94.5%) and statistical test (<0.001) suggest potential heterogeneity across studies; the difference in study population cannot explain the source of heterogeneity. [¶¶¶¶¶]: the point estimates range from 0.589 to 0.726; using inverse-variance method to pool the estimates, the I^2 (97.9%) and statistical test (<0.001) suggest potential heterogeneity across studies; the difference in study population cannot explain the source of heterogeneity. [+++++]: the point estimates range from 0.680 to 0.890; using inverse-variance method to pool the estimates, the I^2 (97.8%) and statistical test (<0.001) suggest potential heterogeneity across studies; the difference in study population cannot explain the source of heterogeneity. [§§§§§]: the point estimates range from 0.680 to 0.811; using inverse-variance method to pool the estimates, the I^2 (88.0%) and statistical test (<0.001) suggest potential heterogeneity across studies; the difference in study population cannot explain the source of heterogeneity. [*fffff*]: the point estimates range from 0.770 to 0.900; using inverse-variance method to pool the estimates, the I^2 (91.3%) and statistical test (<0.001) suggest potential heterogeneity across studies; the difference in study population cannot explain the source of heterogeneity. [######]: we rated down the quality of evidence for indirectness because an indirect measurement tool (EQ-5D) was used to elicit the utility of outcomes. [++++++]: we downgraded the certainty in evidence because of low response rate observed by Lin *et al.* [35] and the potentially biased sampling strategy by asking physicians to provided recruited patients by Boros and Lubinski [34].

TABLE 2 Utility of exacerbation or hospitalisation due to exacerbation

Study	Instrument	Report format	Results
Alcazar *et al.* [19]	EQ-5D VAS	Mean±SD	Hospitalised patients: 0.551±0.197
Antoniu *et al.* [20]	EQ-5D VAS	Mean±SD	Hospitalised patients: 0.279±0.252
Bourbeau *et al.* [47]	EQ-5D VAS	Mean±SD	Change from baseline: −0.126±0.190
Cross *et al.* [21]	EQ-5D VAS	Mean±SD	Exacerbation of COPD: manual chest physiotherapy arm 0.450±0.210; no manual chest physiotherapy arm 0.466±0.214
	EQ-5D utility	Mean±SD	Exacerbation of COPD: manual chest physiotherapy arm 0.450±0.320; no manual chest physiotherapy arm 0.430±0.360
Goossens *et al.* [22]	EQ-5D VAS	Mean±SD	Exacerbation (at enrolment): 0.367±0.252
	EQ-5D utility	Mean±SD	Exacerbation (at enrolment): 0.683±0.209
Menn *et al.* [39]	EQ-5D utility	Mean±SD	EQ-5D Admission Stage III: 0.620±0.260 EQ-5D Admission Stage IV: 0.600±0.260
	SF-6D utility	Mean±SD	SF-12-SF-6D Admission Stage III: 0.610±0.130 SF-12-SF-6D Admission Stage IV: 0.540±0.080
Miravitlles *et al.* [23]	EQ-5D utility	Mean±SD	EQ-5D index baseline (exacerbation): 0.540±0.230
	EQ-5D VAS	Mean±SD	EQ VAS baseline (exacerbation): 0.344±0.274
O'Reilly *et al.* [24]	EQ-5D utility	Mean±SD	Hospital admission: −0.077±0.397
	EQ-5D VAS		Hospital admission: 0.259±0.170
Punekar *et al.* [43]	EQ-5D utility	Mean (95% CI)	1–2 exacerbations in primary care physician setting: 0.740 (0.720–0.770) >3 exacerbations in primary care physician setting: 0.610 (0.590–0.640) 1–2 exacerbations in respiratory specialist setting: 0.730 (0.710–0.760) >3 exacerbations in respiratory specialist setting: 0.570 (0.540–0.600)
Rutten van Molken *et al.* [27]	VAS	Regression coefficient±SEM	1 nonserious exacerbation: −0.037±0.005 2 nonserious exacerbations: −0.068±0.005 1 serious exacerbation: −0.090±0.007 1 nonserious and 1 serious exacerbation: −0.130±0.007
	TTO		1 nonserious exacerbation: −0.010±0.007 2 nonserious exacerbations: −0.021±0.007 1 serious exacerbation: −0.042±0.009 1 nonserious and 1 serious exacerbation: −0.088±0.009
Seymour *et al.* [25]	EQ-5D VAS	Mean±SD	COPD baseline in usual care group: 0.540±0.170 COPD baseline in post-exacerbation pulmonary rehabilitation group: 0.580±0.180
Solem *et al.* [40]	EQ-5D utility	Mean±SD	Patients recently experiencing a severe exacerbation: 0.627±0.210 Patients recently experiencing a moderate exacerbation. 0.698±0.197 Patients who had experienced ⩾3 exacerbations in the previous year: 0.638±0.212 Patients who had experienced 2 exacerbations in the previous year: 0.684±0.204 Patients who had experienced 1 exacerbation in the previous year: 0.727±0.175 Current health (last exacerbation): 0.552±0.283 Thought back, patients experiencing a severe exacerbation (last exacerbation): 0.471±0.313 Thought back, patients experiencing a moderate exacerbation (last exacerbation): 0.595±0.257 Very severe COPD (last exacerbation): 0.494±0.312 Severe COPD (last exacerbation): 0.590±0.256 Patients who had experienced ⩾3 exacerbations in the previous year (last exacerbation): 0.520±0.282 Patients who had experienced 2 exacerbations in the previous year (last exacerbation): 0.552±0.306 Patients who had experienced one exacerbation in the previous year (last exacerbation): 0.610±0.254
Torrance *et al.* [48]	HUI	Mean±SD	For the first acute exacerbation of chronic bronchitis: ciprofloxacin group: 0.720±0.200); usual care group 0.680±0.190 For the remaining acute exacerbation of chronic bronchitis: ciprofloxacin group: 0.740±0.180; usual care group 0.690±0.220
Wildman *et al.* [26]	EQ-5D VAS	Mean±SD Median (IQR)	0.549±0.195 0.500 (0.400–0.700)

EQ-5D: EuroQol-5D; VAS: visual analogue scale; TTO: time trade-off; HUI: Health Utility Index; IQR: interquartile range; COPD: chronic obstructive pulmonary disease.

Importance of dyspnoea

Few studies explored the importance that patients place on dyspnoea. Three studies reported utilities related to dyspnoea. Kim *et al.* [38] reported the utilities measured by VAS by levels of breathlessness: 0.751, 0.656 and 0.526 for level 1 (short of breath during strenuous activities), level 2 (stopping to catch breath after a few minutes walking) and level 3 (breathless when dressing or washing) breathlessness, respectively. The estimates were based on a small sample, so we downgraded the certainty for the estimates by one level for level 2 and two levels for level 3 breathlessness due to concerns about imprecision. Two other reports corroborated that the more severe the dyspnoea symptom, the lower utility patients place on their health, though the specific levels of breathlessness were described differently (table 3) [43, 49]. Other structured surveys, without reporting utility values, also suggested dyspnoea as burdensome and a very important consideration in COPD-related decision-making [31, 50–58].

Adverse events

Table 4 summarises the results related to the importance of adverse events. One of the two included discrete-choice studies compared the "possibility of adverse effects" with "the extent to which treatment seems to relieve symptoms", "the extent to which the doctor gives sufficient time to listen to the patient", "costs of treatment", "the extent to which the patient sees the same doctor each time", and "the extent to which the doctor treats the patient as an entire person" [29]. The extent of symptom relief was deemed to be more important than adverse effects but the possibility of adverse effects was more important than other outcomes. Another discrete choice study suggested symptom relief to be the most important outcome, while the possibility of adverse events was considered more important than the timing and use of (rescue) medicine [30]. The latter study was an online voluntary study in 515 participants, which we rated as having serious risk of bias due to selection bias and limited validity of the instrument. None of the studies explicitly described the outcome of "adverse events." The overall certainty of evidence about the importance of adverse events, based on these two discrete choice studies, is moderate due to serious risk of bias.

Symptom relief

In general, patients considered symptom relief important. In one survey, 46.6% of patients considered relief of symptoms (*i.e.*, chest pain due to coughing, shortness of breath, nausea, *etc.*) as extremely important (ranking second after "not to be kept alive on life support when there is little hope for a meaningful recovery") [61]. For the extent of symptom relief, two discrete-choice studies suggested the extent of symptom relief as more important than adverse effects, the doctor giving sufficient time to listen to the patient, costs of treatment, seeing the same doctor each time, being treated as an entire person, onset time of medication, ease of medication use and use of rescue medication [29, 30]. A large proportion of the study participants were recruited through an online survey, and the eligibility of the participants and the accuracy of their answers were in question. For these reasons, we classified this study at high risk of bias and downgraded the certainty of evidence as moderate (table 1). Three other forced-choice studies corroborated this result [31–33]. For example, in a survey addressing expectation of treatment, 82.3% of the respondents chose greater symptomatic relief as the most important outcome [33]. Because the instruments in these surveys lacked evidence of validity, we rated down the certainty of evidence for risk of bias (moderate certainty evidence).

Utility of COPD

Most studies addressing the utility of the experience of COPD itself were based on EQ-5D, HUI and 15D. Table 5 summarises the utilities based on various instruments across the airflow obstruction levels. Based on the EQ-5D only, we observed a gradient of disutility across GOLD stages: pooled estimates for EQ-5D measurements of mild COPD 0.873 (95% CI 0.863–0.883, I^2=91.3%, p<0.001 for heterogeneity) [28, 35–37, 42, 43, 45], moderate 0.821 (95% CI 0.815–0.826, I^2=97.8%, p<0.001 for heterogeneity) [28, 35–37, 41–45]; severe 0.741 (95% CI 0.734–0.749, I^2=94.5%, p<0.001 for heterogeneity) [28, 35–37, 39–42, 44] and very severe 0.681 (95% CI 0.667–0.694, I^2=80.2%, p<0.001 for heterogeneity) [28, 35–37, 39, 40–42, 44], respectively (figure 2). We rated down the certainty of evidence for these utilities due to unexplained inconsistency and for indirectness of the measurement tool (EQ-5D) (low-certainty evidence); we also observed a similar trend with VAS results (table 1).

Other results

We also identified studies reporting importance on other outcomes (supplementary table 3); for example, intubation and speed of symptom relief.

TABLE 3 Importance on breathlessness, shortness of breath, or dyspnoea

Study	Instrument	Report format	Results
Gruenberger *et al.* [49]	SF-6D utility	Mean	SF-6D health utilities were 0.060 points lower in higher dyspnoea patients (modified Medical Research Council score ⩾2) than in lower dyspnoea patients
Kim *et al.* [28]	EQ-5D VAS	Mean±SEM	EQ-5D utility Level 1 breathlessness (short of breath during strenuous activities): 0.870±0.020 Level 2 breathlessness (stopping to catch breath after a few minutes walking): 0.740±0.030 Level 3 breathlessness (breathless when dressing or washing): 0.540±0.060
	EQ-5D utility	Mean±SEM	EQ-VAS Level 1 breathlessness (short of breath during strenuous activities): 0.751±0.026 Level 2 breathlessness (stopping to catch breath after a few minutes walking): 0.656±0.035 Level 3 breathlessness (breathless when dressing or washing): 0.526±0.071
Punekar *et al.* [43]	EQ-5D utility	Mean (95% CI)	All in primary care physician setting: 0.700 (0.680–0.710) Breathlessness after exercising heavily in primary care physician setting: 0.880 (0.860–0.900) Breathlessness when hurrying on level ground in primary care physician: 0.790 (0.770–0.810) Too breathless to leave house in primary care physician: 0.170 (0.110–0.240) All in respiratory specialist setting: 0.680 (0.660–0.690) Breathlessness after exercising heavily in respiratory specialist setting: 0.880 (0.850–0.900) Breathlessness when hurrying on level ground in respiratory specialist setting: 0.790 (0.770–0.810) Too breathless to leave house in respiratory specialist setting: 0.290 (0.220–0.350)
Braido *et al.* [50]	Uncategorised survey	Choice or proportion of choice	Breathlessness as most troublesome symptom: 64.6% (ranking first; chronic cough: 13.9%; sputum production: 11.0%; exacerbation: 8.3%)
Downey *et al.* [51]	Uncategorised survey: End of life Priority Score (the highest priority aspect of the end-of-life period)	Mean±SD	In a survey on end-of-life priority score measured by rank order (out of 5), breathing comfort was considered as priority: 1.27±1.83 (ranking third, only after time with family and friends, and pain under control)
Haughney *et al.* [52]	Conjoint analysis/discrete choice analysis	Mean	Breathlessness was considered important for patients. Of all the attributes, it was after "impact on everyday life", "need for medical care" and "number of future attacks". It is more important than speed of recovery, productive cough, social impact, sleep disturbance and impact on mood.
Hernández *et al.* [53]	Impact of shortness of breath	Choice or proportion of choice	Shortness of breath is an important outcome, because 6.0% of participants stated the impact on activities of daily living was extreme, 29.0% "very much", 24.0% "a little" and 13.0% "not at all"
Miravitlles *et al.* [54]	Ideal characteristics of a COPD therapy	Choice or proportion of choice	37.0% of the participants chose "increased shortness of breath" as the symptom that has a high impact on wellbeing (ranking second; increased coughing 42.0%, increased fatigue 37.0%, increased production of sputum 35.0%, increased frequency of chest pains 20.0% and fever 13.0%)

Continued

TABLE 3 Continued

Study	Instrument	Report format	Results
Pisa *et al.* [55]	Direct choice: relative importance of COPD attributes (%, higher proportion indicating more importance)	Choice or proportion of choice	Dyspnoea was considered the most important COPD attribute Relative importance of COPD attributes Dyspnoea: 36.0% Performance capability (bodily resilience) due to COPD: 19.0% Sleep quality due to COPD: 19.0% Onset of action of the medication: 3.0% Frequency of administration of the medication: 6.0% Health state after awakening (day start) due to COPD: 13.0% Emotional state due to COPD base medication: 4.0% Effect of attribute levels on health state preference: part-worth utilities (higher value indicating more importance): Dyspnoea Never dyspnoea, except on strong exertion: 115.80 Dyspnoea on exertion: 38.20 Dyspnoea at normal walking pace: −6.60 Dyspnoea on slight effort: −10.10 Dyspnoea even at rest: −137.40
Polati *et al.* [56]	Uncategorised survey: expectation of treatment	Choice or proportion of choice	120 (24.1%) patients would like to have more ease with "breathing" due to treatment; if they were doctors, 215 (43.3%) patients would like to first heal shortness of breath. For both questions, breathing problems were considered most important compared with other symptoms.
Reinke *et al.* [57]	Forced choice: treatment	Choice or proportion of choice	Preferences about death and dying questionnaire. 52.6% of 357 patients chose "being able to breath comfortably in the last 7 days of life" as preferred characteristics of treatment.
Rocker *et al.* [31]	Uncategorised survey: reasons to continue (or not) with opioids	Choice or proportion of choice	I would strongly prefer when followed up at 2 months, 8 (23.5%) and 1 (2.9%) patient claimed would "strongly prefer" and "prefer" to continue on opioids because they provide significant relief from dyspnoea; while at 4–6 months, 12 (29.3%) and 7 (17.1%) patients claimed would "strongly prefer" and "prefer" to continue on opioids because they provide significant relief from dyspnoea.
Wilson *et al.* [58]	Importance of mechanical ventilation: scales for the specific questions about mechanical ventilation	Median (IQR)	On a scale of 1–4 (0: not at all important; 1: a little; 2: quite a bit; 3: very much; 4: extremely important), the score for easing breathlessness was 2.5 (1.8–3.0) for those forego mechanical ventilation, and 3.0 (2.8–4.0) for those uncertain/accept mechanical ventilation.

EQ-5D: EuroQol-5D; VAS: visual analogue scale; COPD: chronic obstructive pulmonary disease; IQR: interquartile range.

TABLE 4 Importance of adverse events

Study	Instrument	Reported format	Results
Bulcun *et al.* [29]	Conjoint analysis/discrete choice analysis	Influence or contribution or weight of certain aspects/ attributes	Possibility of experiencing adverse effects from treatment 20%: −0.90 10%: −0.06 4%: 1.00 Difference between highest and lowest utility levels: 8.20 The sequence of attributes from most important to least important: extent to which the doctor gives sufficient time to listen to the patient, possibility of experiencing adverse effects from treatment, costs of treatment, extent to which the patient sees the same doctor each time, and extent to which the doctor treats the patient as an entire person.
Kawata *et al.* [30]	Willingness to pay, conjoint analysis/discrete choice analysis	Mean (95% CI)	Utility score Mild side-effects (no side-effects as reference): −0.29 (−0.33– −0.24) Moderate to severe side-effects (no side-effects as reference): −1.13 (−1.18– −1.09) Willingness to pay Mild side-effects (no side-effects as reference): $14.81 ($12.40–17.22) Moderate to severe side-effects (no side-effects as reference): $58.69 ($56.28–61.11) Adverse event was considered important for COPD treatment. It was the second most important, only after "complete symptom relief", and more important than "rarely use rescue medication", "quick and easy to use inhaler" and "feeling medication work quickly".
Miravitlles *et al.* [54]	Ideal characteristics of a COPD therapy	Choice or proportion of choice	Ideal characteristics of a COPD therapy as listed by survey respondents Fewer side-effects 36.0% The sequence of ideal characteristics from most important to least important: quick symptom relief, longer intervals between flare-ups, fewer side-effects, better ability to cope with daily chores again, lower costs of treatment, better doses.
Patridge *et al.* [59]	Uncategorised survey: perception of disease severity	Choice or proportion of choice	30.6% of participants expressed concern regarding potential medication side-effects, and on average, patients considered that explaining clearly what are the possible side-effects and risks of the products was very important (9.0 out of 10 on a scale with 1 indicating not at all important and 10 indicating extremely important).
Sharafkhaneh *et al.* [60]	Primary disadvantages of nebulisation therapy	Choice or proportion of choice	Question: what do you see as the main negatives or disadvantages of nebulisation? No negatives: 86 (21.5%) Side-effects: 46 (11.5%) The sequence of disadvantages from most important to least important: device immobile/bulky/cumbersome, time-consuming, side-effects, inconvenient, don't like doing it, having to use it several times a day, care and clean-up after use, too expensive.

COPD: chronic obstructive pulmonary disease.

Discussion

We have conducted the most comprehensive systematic review to date of how COPD patients value outcomes. The identified studies were highly variable in their designs, measurement instruments used and outcomes addressed. Patients rated exacerbations of COPD or hospitalisation due to exacerbations as very important. Studies, primarily using the EQ-5D, consistently reported that the utility associated with living with COPD decreases as the disease progresses. Patients considered symptom relief important and more important than adverse events from treatment.

TABLE 5 Utility of different chronic obstructive pulmonary disease (COPD) severities

Study	Instrument	Reported format	GOLD classifications			
			Mild COPD[#]	Moderate COPD[¶]	Severe COPD[+]	Very severe COPD[§]
Boros and Lubinski [34]	VAS	Mean (95% CI, SD)	0.730 (0.722–0.739, 0.164)	0.626 (0.621–0.630, 0.164)	0.446 (0.439–0.452, 0.161)	0.321 (0.302–0.339, 0.171)
Chen *et al.* [38]	EQ-5D utility	Mean			0.686	0.565
	SF-6D utility[f]	Mean			0.646	0.597
Hong *et al.* [45]	EQ-5D utility	Mean±SD	0.900±0.140	0.890±0.140	0.840±0.150	
	EQ-5D VAS	Mean±SD	0.730±0.186	0.708±0.191	0.609±0.234	
Kim *et al.* [28]	EQ-5D utility	Mean±SD (adjusted mean ±SEM)	0.830±0.170 (0.830±0.040)	0.880±0.120 (0.880±0.020)	0.820±0.160 (0.810±0.030)	0.610±0.260 (0.600±0.040)
	EQ-5D VAS	Mean±SD (adjusted mean ±SEM)	0.719±0.189 (0.739±0.054)	0.719±0.178 (0.751±0.029)	0.650±0.206 (0.689±0.033)	0.609±0.139 (0.651±0.056)
Lin *et al.* [35]	EQ-5D utility	Mean±SD	0.810±0.140	0.810±0.140	0.760±0.170	0.740±0.150
	EQ-5D VAS	Mean±SD	0.766±0.175	0.726±0.191	0.657±201	0.611±0.197
Menn *et al.* [39]	EQ-5D utility	Mean±SD			0.620±0.260	0.600±0.260
	SF-6D utility	Mean±SD			0.610±0.130	0.540±0.080
Pickard *et al.* [36]	EQ-5D utility[##]	Mean±SD	0.800±0.130	0.700±0.210	0.720±0.190	0.720±0.160
	EQ-5D utility[¶¶]	Mean±SD	0.730±0.190	0.590±0.320	0.630±0.250	0.630±0.240
	EQ-5D VAS	Mean±SD	0.743±0.163	0.662±0.200	0.601±0.184	0.587±0.158
Punekar *et al.* [43]	EQ-5D utility	Mean (95% CI)	0.770 (0.730–0.810) in primary care setting	0.680 (0.620–0.740)	0.620 (0.560–0.680)	
	EQ-5D utility	Mean (95% CI)	0.680 (0.640–0.720) in respiratory specialist care setting	0.720 (0.690–0.750)	0.640 (0.610–0.720)	
Rodriguez Gonzalez-Moro *et al.*[46]	EQ-5D VAS	Mean (95% CI)		0.589 (0.581–0.599)	0.459 (0.449–0.467)	
Rutten van Molken *et al.* [44]	EQ-5D VAS	Mean (95% CI)		0.677 (0.665–0.690)	0.625 (0.610–0.639)	0.578 (0.545–0.612)
	EQ-5D utility[¶¶]	Mean (95% CI)		0.787 (0.771–0.802)	0.750 (0.731–0.768)	0.647 (0.598–0.695)
	EQ-5D utility[##]	Mean (95% CI)		0.832 (0.821–0.843)	0.803 (0.790–0.816)	0.731 (0.699–0.762)
Rutten van Molken *et al.* [27]	VAS	Mean±SEM	Mild COPD: 0.811 ±0.011	Disutility of moderate COPD in relation to mild COPD: −0.133±0.006	Disutility of severe COPD in relation to mild COPD: −0.354±0.006	Disutility of very severe COPD in relation to mild COPD: −0.508±0.006
	TTO	Mean±SEM	Mild COPD: 0.974 ±0.017	Disutility of moderate COPD in relation to mild COPD: −0.045±0.008	Disutility of severe COPD in relation to mild COPD: −0.257±0.008	Disutility of very severe COPD in relation to mild COPD: −0.452±0.008
Scharf *et al.* [62]	HUI utility	Mean±SD	0.400±0.330	0.580±0.360	0.530±0.350	0.390±0.510
Solem *et al.* [40]	EQ-5D utility	Mean±SD			0.707±0.174	0.623±0.234
Stahl *et al.* [37]	EQ-5D VAS	Mean±SD	0.730±0.210	0.650±0.240	0.620±0.210	0.370±0.280
	EQ-5D utility	Mean±SD	0.840±0.150	0.730±0.230	0.740±0.250	0.520±0.260
Starkie *et al.* [41]	EQ-5D utility	Mean±SD		Observed utility for moderate COPD: 0.752 ±0.220	Observed utility for severe COPD: 0.708 ±0.230	Observed utility for very severe COPD: 0.672±0.220
Szende *et al.* [42]	EQ-5D utility	Mean±SD	0.850±0.160	0.730±0.210	0.740±0.240	0.530±0.280
	SF-6D utility	Mean±SD	0.800±0.130	0.730±0.130	0.730±0.140	0.620±0.150

GOLD: Global Initiative for Chronic Obstructive Lung Disease; VAS: visual analogue scale; EQ-5D: EuroQol-5D; TTO: time trade-off; HUI: Health Utility Index. [#]: forced expiratory volume in 1 s (FEV_1) ⩾80% predicted; [¶]: FEV_1 <80–⩾50% predicted; [+]: FEV_1 <50–⩾30% predicted; [§]: FEV_1 <30% predicted; [f]: Hong Kong value set; [##]: US value set; [¶¶]: UK value set.

Several aspects distinguish our work from previous published literature reviews [63–67]. Our work yielded more studies because of the broad definition focusing on the importance of outcomes, and including all types of relevant studies and measurement tools. For example, our work is more comprehensive than the work Moayeri *et al.* [63] who evaluated EQ-5D utilities of COPD stages, though the results of our pooled EQ-5D utilities proved similar. Two other reviews included only multiattribute utility results [63, 64]. Brooker *et al.* [67] identified 10 studies on patient preferences for mechanical ventilation in COPD, most of them cross-sectional surveys with forced choice questions. A second aspect in which our work differs is the critical assessment, both at the individual-study level for risk of bias, and at the body of evidence level with the GRADE approach and the associated summary of findings table [68].

Our study has some limitations. First, because of the paucity of evidence based on standard gamble and TTO, we were only able to conduct meta-analysis across severity levels of EQ-5D utility and VAS measurements. For the same reason, we were unable to quantitatively explore the study population characteristics as potential sources of inconsistency through approaches such as metaregression. Second,

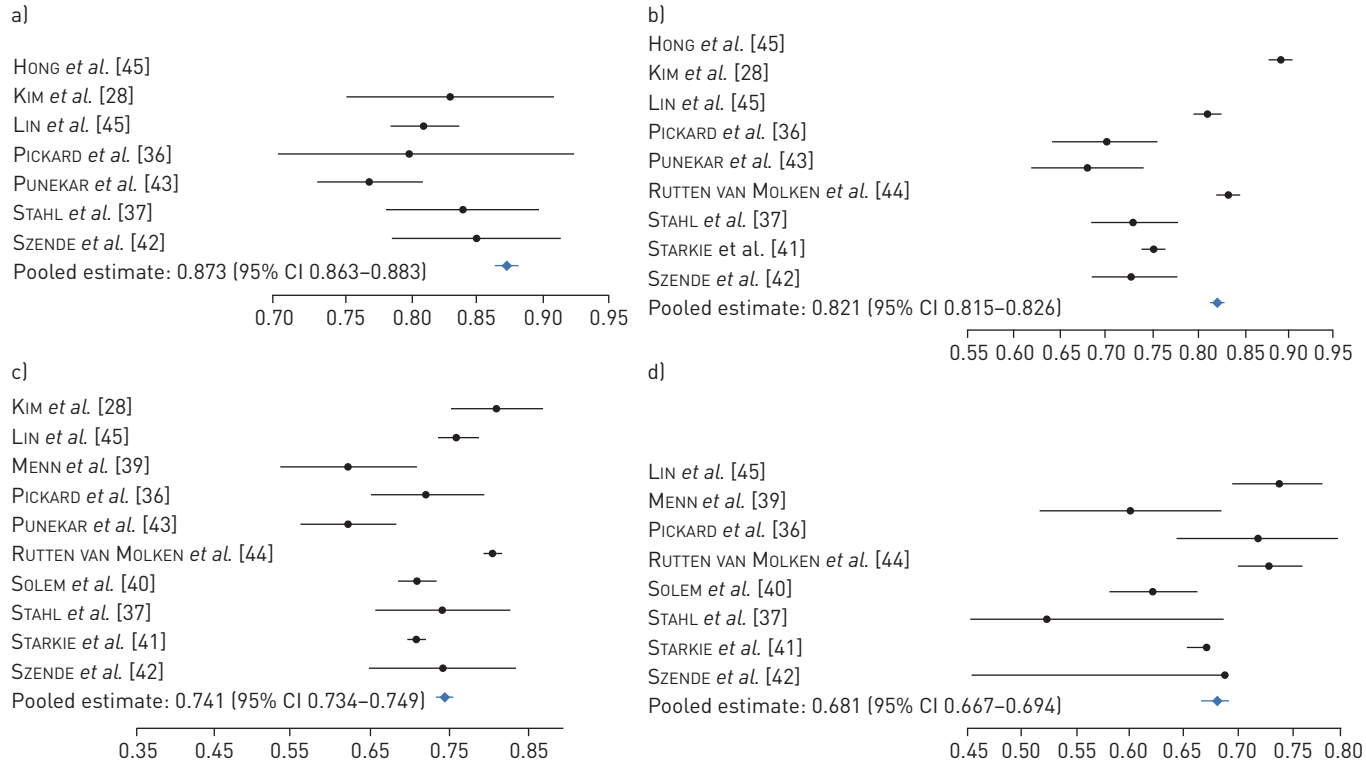

FIGURE 2 Forest plots for EuroQol-5D (EQ-5D) utility of chronic obstructive pulmonary disease patients with forced expiratory volume in 1 s a) ⩾80%, b) <80–⩾50%, c) <50–⩾30% and d) <30% predicted.

we identified a relatively small number of discrete choice and probability trade-off studies. These studies could provide information on the threshold for a change in decision [69] and have the merit of allowing customisation of the methodology according to the study objectives. The few probability trade-off and discrete choice exercise studies reported only a limited range of attributes and levels of attributes [70–72]. Lastly, given the lack of empirical knowledge in what manner and to what extent publication bias may affect our systematic review results, our assessment of publication bias is limited.

Given the breadth of findings, this systematic review has implications for healthcare providers, researchers including systematic review authors and guideline developers. This systematic review summarises current evidence to inform guideline developers about how important the benefits and harms of COPD treatment strategies are from the patients' perspective. The results will inform clinicians who make decisions with COPD patients. This systematic review provides empirical evidence to support using the relative importance of outcomes to inform values and preferences, and the methods can be used by systematic review authors who are interested in other disease topics. The utilities summarised serve as the parameter inputs for cost analyses. When guideline developers determine the balance between benefits and harms, they can take into consideration both the probability and the importance of benefits (*e.g.* symptom relief) and harms (*e.g.* adverse events) from this review. Additionally, the results of this review also help researchers identify research gaps for designing new studies.

Research gaps exist when there is no evidence, or the certainty of evidence is low or very low. For example, although there is evidence about the importance of adverse events, guideline developers need to know the exact types and probabilities of adverse events considered by patients. Researchers can use standard gamble, discrete choice and probability trade-off techniques to address the levels of adverse events, with the severities or probabilities directly relevant to the research questions [29]. Additionally, for better understanding and application of the findings, researchers also need to further explore the socioeconomic, cultural and disease-specific characteristics that influence patient values on the COPD outcomes.

There are still unanswered challenges related to the optimal strategy to elicit the outcome importance evidence. For considering the risk of bias, one concern is the merits of measurement tools involving a valuation of hypothetical scenarios in relation to measurements of an actual outcome that participants experience. If the participants value a health state specified by the investigators, barring only different interpretations or limited understanding, they value the same outcome; but if, for example, participants are asked to evaluate the

outcome "shortness of breath" they are experiencing, or having experienced in the past, the degree of shortness may vary a lot across participants. Further studies are also necessary to validate the search strategy for these types of studies. Our strategy, which is sensitive but not specific, led to a large number of hits [12], replication of which would place a substantial burden for systematic review authors and guideline panels (as it did for us).

Conclusion

Our systematic review showed that patients value the outcome of exacerbation or hospitalisation due to exacerbation as very important. We observed large variability in the utility associated with COPD severity across studies. We identified a gradient of disutility as the disease progresses, from both the direct utility instrument VAS and the indirect utility instrument EQ-5D. Quantitative approaches, including direct and indirect utility measurement of outcomes, discrete choice exercise, probability trade-off and forced choice, represent the predominant measurement instruments investigators have used to address the importance patients place on outcomes.

Although further studies are necessary to explore the unsolved methodological questions, through this systematic review process, we demonstrated the usefulness of systematic reviews as a potential strategy for summarising evidence in this field and informing decision makers, both in the context of health technology assessments and guidelines.

Acknowledgements: We are grateful to Amiram Gafni (McMaster University) for the comments on the manuscript, and Sean Doran (University of Missouri, Kansas City, MO, USA) for title and abstract screening.

Author Contributions: Y. Zhang and H.J. Schünemann designed the study; Y. Zhang, R.L. Morgan, P. Alonso-Coello, A. Selva, H. Ara Begum, G.P. Morgano, W. Wiercioch, M. Ventresca, M.M. Bała, R.R. Jaeschke, M.T. Wasylewski, K. Styczeń, H. Pardo-Hernandez, A. Agarwal, J-L. Kerth, L. Blanco-Silvente, M. Wang, Y. Zhang, S. Narsingam and Y. Fei screened the literature and abstracted the data; Y. Zhang, R.L. Morgan, P. Alonso-Coello, G. Guyatt and H.J. Schünemann drafted the manuscript; all authors read and approved the final manuscript; and H.J. Schünemann conceived of and funded the study.

Conflict of interest: H.J. Schünemann reports that he has no financial conflict of interest. He is Co-chair of the GRADE working group.

Support statement: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors but was supported by the MacGRADE center at McMaster University. This project does not have any sponsorship from industrial or governmental sources in its design, collection of data, analysis or interpretation of data, writing of the report, or decision to submit for publication. Funding information for this article has been deposited with the Crossref Funder Registry.

References

1 Sackett DL, Rosenberg WM, Gray JA, *et al.* Evidence based medicine: what it is and what it isn't. *BMJ* 1996; 312: 71–72.
2 MacLean S, Mulla S, Akl EA, *et al.* Patient values and preferences in decision making for antithrombotic therapy: a systematic review: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. *Chest* 2012; 141: Suppl., e1S–23S.
3 Schunemann HJ, Wiercioch W, Etxeandia I, *et al.* Guidelines 2.0: systematic development of a comprehensive checklist for a successful guideline enterprise. *CMAJ* 2014; 186: E123–E142.
4 Zhang Y, Coello PA, Brozek J, *et al.* Using patient values and preferences to inform the importance of health outcomes in practice guideline development following the GRADE approach. *Health Qual Life Outcomes* 2017; 15: 52.
5 Alonso-Coello P, Schünemann HJ, Moberg J, *et al.* GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. *BMJ* 2016; 353: i2016.
6 Bremner KE, Chong CA, Tomlinson G, *et al.* A review and meta-analysis of prostate cancer utilities. *Med Decis Making* 2007; 27: 288–298.
7 Pickard AS, Wilke CT, Lin H-W, *et al.* Health utilities using the EQ-5D in studies of cancer. *Pharmacoeconomics* 2007; 25: 365–384.
8 Sepucha K, Ozanne EM. How to define and measure concordance between patients' preferences and medical treatments: a systematic review of approaches and recommendations for standardization. *Patient Educ Couns* 2010; 78: 12–23.
9 Joy SM, Little E, Maruthur NM, *et al.* Patient preferences for the treatment of type 2 diabetes: a scoping review. *PharmacoEconomics* 2013; 31: 877–892.
10 Mannino DM, Buist AS. Global burden of COPD: risk factors, prevalence, and future trends. *Lancet* 2007; 370: 765–773.
11 Moher D, Shamseer L, Clarke M, *et al.* Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015; 4: 1.
12 Selva A, Sola I, Zhang Y, *et al.* Development and use of a content search strategy for retrieving studies on patients' views and preferences. *Health Qual Life Outcomes* 2017; 15: 126.
13 Gafni A, Birch S. Preferences for outcomes in economic evaluation: an economic approach to addressing economic problems. *Soc Sci Med* 1995; 40(6): 767–776.
14 Zhang Y, Alonso-Coello P, Guyatt GH, *et al.* GRADE Guidelines: 19. Assessing the certainty of evidence in the importance of outcomes or values and preferences - risk of bias and indirectness. *J Clin Epidemiol* 2018; in press [https://doi.org/10.1016/j.jclinepi.2018.01.013].

15 Zhang Y, Alonso-Coello P, Guyatt GH, *et al.* GRADE Guidelines: 20. Assessing the certainty of evidence in the importance of outcomes or values and preferences - inconsistency, imprecision, and other domains. *J Clin Epidemiol* 2018; in press [https://doi.org/10.1016/j.jclinepi.2018.05.011].
16 Global Initiative for Chronic Obstructive Lung Disease. Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Pulmonary Disease. Available from: www.goldcopd.org/
17 Birch S, Ismail AI. Patient preferences and the measurement of utilities in the evaluation of dental technologies. *J Dent Res* 2002; 81: 446–450.
18 Peasgood T, Brazier J. Is meta-analysis for utility values appropriate given the potential impact different elicitation methods have on values? *Pharmacoeconomics* 2015; 33: 1101–1105.
19 Alcazar B, Garcia-Polo C, Herrejon A, *et al.* Factores asociados a la hospitalización por exacerbación de la enfermedad pulmonar obstructiva crónica [Factors associated with hospital admission for exacerbation of chronic obstructive pulmonary disease]. *Arch Bronconeumol* 2012; 48: 70–76.
20 Antoniu SA, Puiu A, Zaharia B, *et al.* Health status during hospitalisations for chronic obstructive pulmonary disease exacerbations: the validity of the Clinical COPD Questionnaire. *Expert Rev Pharmacoeconomics Outcomes Res* 2014; 14: 283–287.
21 Cross J, Elender F, Barton G, *et al.* A randomised controlled equivalence trial to determine the effectiveness and cost-utility of manual chest physiotherapy techniques in the management of exacerbations of chronic obstructive pulmonary disease (MATREX). *Health Technol Assess* 2010; 14: 1–147.
22 Goossens LM, Nivens MC, Sachs P, *et al.* Is the EQ-5D responsive to recovery from a moderate COPD exacerbation? *Respir Med* 2011; 105: 1195–1202.
23 Miravitlles M, Izquierdo I, Herrejon A, *et al.* COPD severity score as a predictor of failure in exacerbations of COPD. The ESFERA study. *Respir Med* 2011; 105: 740–747.
24 O'Reilly JF, Williams AE, Rice L. Health status impairment and costs associated with COPD exacerbation managed in hospital. *Int J Clin Pract* 2007; 61: 1112–1120.
25 Seymour JM, Moore L, Jolley CJ, *et al.* Outpatient pulmonary rehabilitation following acute exacerbations of COPD. *Thorax* 2010; 65: 423–428.
26 Wildman MJ, Sanderson CF, Groves J, *et al.* Survival and quality of life for patients with COPD or asthma admitted to intensive care in a UK multicentre cohort: the COPD and Asthma Outcome Study (CAOS). *Thorax* 2009; 64: 128–132.
27 Rutten-van Molken MP, Hoogendoorn M, Lamers LM. Holistic preferences for 1-year health profiles describing fluctuations in health: the case of chronic obstructive pulmonary disease. *Pharmacoeconomics* 2009; 27: 465–477.
28 Kim SH, Oh YM, Jo MW. Health-related quality of life in chronic obstructive pulmonary disease patients in Korea. *Health Quality Life Outcomes* 2014; 12: 57.
29 Bulcun E, Ekici M, Ekici A. Assessment of patients' preferences regarding the characteristics associated with the treatment of chronic obstructive pulmonary disease. *Int J Chron Obstruct Pulm Dis* 2014; 9: 363–368.
30 Kawata AK, Kleinman L, Harding G, *et al.* Evaluation of patient preference and willingness to pay for attributes of maintenance medication for chronic obstructive pulmonary disease (COPD). *Patient* 2014; 7: 413–426.
31 Rocker GM, Simpson AC, Horton R, *et al.* Opioid therapy for refractory dyspnea in patients with advanced chronic obstructive pulmonary disease: patients' experiences and outcomes. *CMAJ Open* 2013; 1: E27–E36.
32 Claessens MT, Lynn J, Zhong Z, *et al.* Dying with lung cancer or chronic obstructive pulmonary disease: insights from SUPPORT. Study to Understand Prognoses and Preferences for Outcomes and Risks of Treatments. *J Am Geriatr Soc* 2000; 48: Suppl., S146–S153.
33 Kuyucu T, Guclu SZ, Saylan B, *et al.* A cross-sectional observational study to investigate daily symptom variability, effects of symptom on morning activities and therapeutic expectations of patients and physicians in COPD-SUNRISE study. *Tuberk Toraks* 2011; 59: 328–339.
34 Boros PW, Lubinski W. Health state and the quality of life in patients with chronic obstructive pulmonary disease in Poland: a study using the EuroQoL-5D questionnaire. *Pol Arch Med Wewn* 2012; 122: 73–81.
35 Lin FJ, Pickard AS, Krishnan JA, *et al.* Measuring health-related quality of life in chronic obstructive pulmonary disease: properties of the EQ-5D-5L and PROMIS-43 short form. *BMC Med Res Methodol* 2014; 14: 78.
36 Pickard AS, Yang Y, Lee TA. Comparison of health-related quality of life measures in chronic obstructive pulmonary disease. *Health Qual Life Outcomes* 2011; 9: 26.
37 Stahl E, Lindberg A, Jansson SA, *et al.* Health-related quality of life is related to COPD disease severity. *Health Quality Life Outcomes* 2005; 3: 56.
38 Chen J, Wong CK, McGhee SM, *et al.* A comparison between the EQ-5D and the SF-6D in patients with chronic obstructive pulmonary disease (COPD). *PLoS ONE* 2014; 9: e112389.
39 Menn P, Weber N, Holle R. Health-related quality of life in patients with severe COPD hospitalized for exacerbations – comparing EQ-5D, SF-12 and SGRQ. *Health Quality life Outcomes* 2010; 8: 39.
40 Solem CT, Sun SX, Sudharshan L, *et al.* Exacerbation-related impairment of quality of life and work productivity in severe and very severe chronic obstructive pulmonary disease. *Int J Chron Obstruct Pulm Dis* 2013; 8: 641–652.
41 Starkie HJ, Briggs AH, Chambers MG, *et al.* Predicting EQ-5D values using the SGRQ. *Value Health* 2011; 14: 354–360.
42 Szende A, Leidy NK, Stahl E, *et al.* Estimating health utilities in patients with asthma and COPD: evidence on the performance of EQ-5D and SF-6D. *Qual Life Res* 2009; 18: 267–272.
43 Punekar YS, Rodriguez-Roisin R, Sculpher M, *et al.* Implications of chronic obstructive pulmonary disease (COPD) on patients' health status: a western view. *Respir Med* 2007; 101: 661–669.
44 Rutten-van Molken MP, Oostenbrink JB, Tashkin DP, *et al.* Does quality of life of COPD patients as measured by the generic EuroQol five-dimension questionnaire differentiate between COPD severity stages? *Chest* 2006; 130: 1117–1128.
45 Hong JY, Kim SY, Chung KS, *et al.* Factors associated with the quality of life of Korean COPD patients as measured by the EQ-5D. *Qual Life Res* 2015; 24: 2549–2558.
46 Rodriguez Gonzalez-Moro JM, de Lucas Ramos P, Izquierdo Alonso JL, *et al.* Impact of COPD severity on physical disability and daily living activities: EDIP-EPOC I and EDIP-EPOC II studies. *Int J Clin Pract* 2009; 63: 742–750.

47 Bourbeau J, Ford G, Zackon H, *et al.* Impact on patients' health status following early identification of a COPD exacerbation. *Eur Respir J* 2007; 30: 907–913.
48 Torrance G, Walker V, Grossman R, *et al.* Economic evaluation of ciprofloxacin compared with usual antibacterial care for the treatment of acute exacerbations of chronic bronchitis in patients followed for 1 year. *Pharmacoeconomics* 1999; 16: 499–520.
49 Gruenberger JB, Vietri J, Keininger DL, *et al.* Greater dyspnea is associated with lower health-related quality of life among European patients with COPD. *Int J Chron Obstruct Pulm Dis* 2017; 12: 937–944.
50 Braido F, Baiardini I, Molinengo G, *et al.* Choose your outcomes: From the mean to the personalized assessment of outcomes in COPD. An exploratory pragmatic survey. *Eur J Intern Med* 2016; 34: 85–88.
51 Downey L, Engelberg RA, Curtis JR, *et al.* Shared priorities for the end-of-life period. *J Pain Symptom Manage* 2009; 37: 175–188.
52 Haughney J, Partridge MR, Vogelmeier C, *et al.* Exacerbations of COPD: quantifying the patient's perspective using discrete choice modellling. *Eur Respir J* 2005; 26: 623–629.
53 Hernandez P, Balter MS, Bourbeau J, *et al.* Canadian practice assessment in chronic obstructive pulmonary disease: respiratory specialist physician perception *versus* patient reality. *Can Respir J* 2013; 20: 97–105.
54 Miravitlles M, Anzueto A, Legnani D, *et al.* Patient's perception of exacerbations of COPD – the PERCEIVE study. *Respir Med* 2007; 101: 453–460.
55 Pisa G, Freytag S, Schandry R. Chronic obstructive pulmonary disease (COPD) patients' disease-related preferences : a study using conjoint analysis. *Patient* 2013; 6: 93–101.
56 Polatli M, Bilgin C, Saylan B, *et al.* A cross sectional observational study on the influence of chronic obstructive pulmonary disease on activities of daily living: the COPD-Life study. *Tuberk Toraks* 2012; 60: 1–12.
57 Reinke LF, Uman J, Udris EM, *et al.* Preferences for death and dying among veterans with chronic obstructive pulmonary disease. *Am J Hosp Palliat Care* 2013; 30: 768–772.
58 Wilson KG, Aaron SD, Vandemheen KL, *et al.* Evaluation of a decision aid for making choices about intubation and mechanical ventilation in chronic obstructive pulmonary disease. *Patient Educ Counsel* 2005; 57: 88–95.
59 Partridge MR, Dal Negro RW, Olivieri D. Understanding patients with asthma and COPD: insights from a European study. *Prim Care Respir J* 2011; 20: 315–323.
60 Sharafkhaneh A, Wolf RA, Goodnight S, *et al.* Perceptions and attitudes toward the use of nebulized therapy for COPD: patient and caregiver perspectives. *COPD* 2013; 10: 482–492.
61 Rocker GM, Dodek PM, Heyland DK. Toward optimal end-of-life care for patients with advanced chronic obstructive pulmonary disease: insights from a multicentre study. *Can Respir J* 2008; 15: 249–254.
62 Scharf SM, Maimon N, Simon-Tuval T, *et al.* Sleep quality predicts quality of life in chronic obstructive pulmonary disease. *Int J Chron Obstruct Pulm Dis* 2011; 6: 1–12.
63 Moayeri F, Hsueh YS, Clarke P, *et al.* Health state utility value in chronic obstructive pulmonary disease (COPD); the challenge of heterogeneity: a systematic review and meta-analysis. *COPD* 2016; 13: 380–398.
64 Petrillo J, van Nooten F, Jones P, *et al.* Utility estimation in chronic obstructive pulmonary disease: a preference for change? *Pharmacoeconomics* 2011; 29: 917–932.
65 Pickard AS, Wilke C, Jung E, *et al.* Use of a preference-based measure of health (EQ-5D) in COPD and asthma. *Respir Med* 2008; 102: 519–536.
66 Bereza BG, Troelsgaard Nielsen A, Valgardsson S, *et al.* Patient preferences in severe COPD and asthma: a comprehensive literature review. *Int J Chron Obstruct Pulm Dis* 2015; 10: 739–744.
67 Brooker AS, Carcone S, Witteman W, *et al.* Quantitative patient preference evidence for health technology assessment: a case study. *Int J Technol Assess Health Care* 2013; 29: 290–300.
68 Balshem H, Helfand M, Schunemann HJ, *et al.* GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol* 2011; 64: 401–406.
69 Janssen DJ, Spruit MA, Schols JM, *et al.* A call for high-quality advance care planning in outpatients with severe COPD or chronic heart failure. *Chest* 2011; 139: 1081–1088.
70 Ryan M. Discrete choice experiments in health care. *BMJ* 2004; 328: 360–361.
71 Ryan M, Farrar S. Using conjoint analysis to elicit preferences for health care. *BMJ* 2000; 320: 1530–1533.
72 Sculpher M, Bryan S, Fry P, *et al.* Patients' preferences for the management of non-metastatic prostate cancer: discrete choice experiment. *BMJ* 2004; 328: 382.