Supplementary materials

Definitions of non-response and response to biological therapy for severe asthma: a systematic review.

Ekaterina Khaleva, Anna Rattu, Chris Brightling, Andrew Bush, Arnaud Bourdin, Apostolos Bossios, Kian Fan Chung, Rekha Chaudhuri, Courtney Coleman, Ratko Djukanovic, Sven-Erik Dahlén, Andrew Exley, Louise Fleming, Stephen J Fowler, Atul Gupta, Eckard Hamelmann, Gerard H. Koppelman, Erik Melen, Vera Mahler, Paul Seddon, Florian Singer, Celeste Porsbjerg, Valeria Ramiconi, Franca Rusconi, Valentyna Yasinska, Graham Roberts on behalf of the 3TR and COMSA working group.

Authors and Affiliations

Ekaterina Khaleva: Clinical and Experimental Sciences and Human Development and Health, Faculty of Medicine, University of Southampton, Southampton, UK. ORCID: 0000-0002-2220-7745

Anna Rattu: Clinical and Experimental Sciences and Human Development and Health, Faculty of Medicine, University of Southampton, Southampton, UK.

Chris Brightling: Institute for Lung Health, Leicester NIHR BRC, University of Leicester, UK. ORCID: 0000-0002-9345-4903

Andrew Bush: Centre for Paediatrics and Child Health and National Heart and Lung Institute, Imperial College; Royal Brompton Hospital, London, UK ORCID: 0000-0001-6756-9822

Arnaud Bourdin: PhyMedExp, University of Montpellier, INSERM U1046, CNRS UMR 9214. 34295 Montpellier cedex 5, France

Apostolos Bossios: Department of Respiratory Medicine and Allergy, Karolinska University Hospital, Huddinge and Department of Medicine, Huddinge, Karolinska Institutet, Stockholm, Sweden. ORCID: 0000-0002-0494-2690.

Kian Fan Chung: National Heart & Lung Institute, Imperial College London, London. UK. ORCID: 0000-0001-7101-1426

Rekha Chaudhuri: Institute of Infection, Immunity & Inflammation, University of Glasgow, Glasgow, UK

Courtney Coleman: European Lung Foundation, Sheffield, UK

Ratko Djukanovic: NIHR Southampton Biomedical Research Centre, University Hospital Southampton, Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Sir Henry Wellcome Laboratories, Southampton, UK. ORCID: 0000-0001-6039-5612

Sven-Erik Dahlén: Department of Respiratory Medicine and Allergy, Karolinska University Hospital, Huddinge and Department of Medicine, Huddinge, Karolinska Institutet, Stockholm, Sweden.

Andrew Exley: Adept Biologica Consulting Limited, London, UK. ORCID: 0000-0002-2628-6129

Louise Fleming: National Heart and Lung Institute, Imperial College, London ORCID: 0000-0002-7268-7433

Stephen J Fowler: Faculty of Biology, Medicine and Health, School of Biological Sciences, Division of Infection, Immunity & Respiratory Medicine, The University of Manchester, and NIHR Manchester Biomedical Research Unit and Manchester University NHS Foundation Trust, Manchester, UK

Atul Gupta: Department of Paediatric Respiratory Medicine, King's College Hospital, London, UK. ORCID 0000-0002-1610-0335

Eckard Hamelmann: Children's Center Bethel, Department of Pediatrics, University Bielefeld, Bielefeld, Germany

Gerard H. Koppelman: University of Groningen, University Medical Center Groningen, Beatrix Children's Hospital, Department of Pediatric Pulmonology and Pediatric Allergology, Groningen, the Netherlands; University of Groningen, University Medical Center Groningen, Groningen Research Institute for Asthma and COPD (GRIAC), Groningen, the Netherlands. ORCID: 0000-0001-8567-3252.

Erik Melén: Department of Clinical Science and Education Södersjukhuset, Karolinska Institutet, Stockholm, Sweden. ORCID: 0000-0002-8248-0663

Vera Mahler: Paul-Ehrlich-Institut, Federal Institute for Vaccines and Biomedicines, Division of Allergology, Langen, Germany. ORCID: 0000-0001-6471-1811

Paul Seddon: Respiratory Care, Royal Alexandra Children's Hospital, Brighton, UK. ORCID: 0000-0003-2136-958X

Florian Singer: 1. Department of Respiratory Medicine, University Children's Hospital Zurich and Childhood Research Center, Zurich, Switzerland. 2. Division of Paediatric Pulmonology and Allergology, Department of Paediatrics and Adolescent Medicine, Medical University of Graz, Austria. ORCID: 0000-0003-3471-5664

Celeste Porsbjerg: Department of Respiratory Medicine, Respiratory Research Unit, Bispebjerg Hospital, Copenhagen, Denmark, ORCID: 0000-0003-4825-9436

Valeria Ramiconi: European Federation of Allergy and Airways Diseases Patients' Associations, Brussels, Belgium.

Franca Rusconi: Department of Mother and Child Health, Azienda USL Toscana Nord Ovest, Pisa, Italy. ORCID: 0000-0002-9544-6472

Valentyna Yasinska, Department of Respiratory Medicine and Allergy, Karolinska University Hospital, Huddinge and Department of Medicine, Huddinge, Karolinska Institutet, Stockholm, Sweden.

Graham Roberts: Clinical and Experimental Sciences and Human Development and Health, University of Southampton, University Road, Highfield, Southampton, UK. NIHR Southampton Biomedical Research Centre, University Hospital Southampton NHS Foundation Trust, Southampton, UK.

Correspondence Address: Professor Graham Roberts, Paediatric Allergy and Respiratory Medicine, University Child Health (MP803), University Hospital, Southampton NHS Foundation Trust, Tremona Road, Southampton SO16 6YD, UK. Tel.: +44 (0) 2380796160 E-mail: g.c.roberts@soton.ac.uk

Table of Contents

Appendix 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist.

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	Title
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	Abstract
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	Introduction
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	Introduction
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	Methods- inclusion criteria; exclusion criteria
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	Methods-Search strategy
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	Supplementary materials
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	Methods-Study selection
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	Methods- Data extraction, risk of bias assessment, quality, and synthesis of the results
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	Appendix 3.
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	Appendix 3.
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	Data extraction, risk of bias assessment, quality, and synthesis of the results
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	NA
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	NA
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	NA

Section and Topic	Item #	Checklist item	Location where item is reported
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	NA
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	NA
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	NA
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	NA
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	Appendix 3.
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	Methods- Data extraction, risk of bias assessment, quality, and synthesis of the results
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	Figure 1
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	Figure 1
Study characteristics	17	Cite each included study and present its characteristics.	Table 1; 2; S3; S6;
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	Table S7
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	NA
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	Results- Development and quality of definitions of non-response and response; Development and content validity of the outcome measures; Risk of bias and quality of evidence for validation studies of outcome measures
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	NA
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	NA
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	NA
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	NA
Certainty of	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	Results-Risk of bias and

Section and Topic	Item #	Checklist item	Location where item is reported
evidence			quality of evidence for validation studies of outcome measures
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	Discussion
	23b	Discuss any limitations of the evidence included in the review.	Discussion: Strengths and limitations
	23c	Discuss any limitations of the review processes used.	Discussion: Strengths and limitations
	23d	Discuss implications of the results for practice, policy, and future research.	Discussion: Policy implications and next steps
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	Methods
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	Methods
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	Methods
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	Funding
Competing interests	26	Declare any competing interests of review authors.	Conflict of interests
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	Supplementary materials

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71

Appendix 2. Search strategies

I. Search strategy in EMBASE (OVID)

1. asthma/ or allergic asthma/ or aspirin exacerbated respiratory disease/ or asthmatic state/ or exercise induced asthma/ or experimental asthma/ or extrinsic asthma/ or intrinsic asthma/ or mild intermittent asthma/ or mild persistent asthma/ or moderate persistent asthma/ or nocturnal asthma/ or occupational asthma/ or severe persistent asthma/

2. asthma*.ti,ab.

3. 1 or 2

4. omalizumab.mp. or exp omalizumab/

5. mepolizumab.mp. or exp mepolizumab/

6. reslizumab.mp. or reslizumab/

7. benralizumab.mp. or exp benralizumab/

8. dupilumab.mp. or exp dupilumab/

9. tralokinumab.mp. or exp tralokinumab/

10. lebrikizumab.mp. or exp lebrikizumab/

11. tezepelumab.mp. or exp tezepelumab/

12. brodalumab.mp. or exp brodalumab/

13. ligelizumab.mp. or exp ligelizumab/

14. Pitrakinra.mp. or pitrakinra/

15. exp biological product/ or exp biological therapy/ or biologic*.mp.

16. (biologic* adj1 (treatment* or therap* or medicine* or drug* or agent* or product*)).mp.

17. monoclonal antibod*.mp. or exp monoclonal antibody/

18. 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17

19. drug response/ or exp treatment response/ or partial drug response/

20. (responsive* or response or respond* or nonrespon*).mp.

21. treatment outcome/ or outcome assessment/

22. minimal clinically important difference/ or meaningful change.mp.

23. (Minimal* adj1 (clinical* or important or real or significant) adj1 (change or difference)).mp.

24. (Minimal* adj1 clinical* adj1 (important or significant) adj1 (change or difference)).mp.

25. (MCID or MID or MIC).mp.

26. 19 or 20 or 21 or 22 or 23 or 24 or 25

27. editorial/ or review/ or case report/ or case report*.mp.

28. editorial*.mp.

29. conference abstract*.mp.

30. conference paper*.mp. or conference paper/ or conference abstract/

31. ((systematic or narrative) adj2 review*).mp. or "systematic review"/

32. ((("semi-structured" or semistructured or unstructured or informal or "in-depth" or indepth or "face-to-face" or structured or guide) adj3 (interview* or discussion* or questionnaire*)) or ("focus group*" or qualitative or ethnograph* or fieldwork or "field work" or "key informant")).ti,ab. or survey*.ti.

33. interview/ or information processing/ or verbal communication/ or qualitative research/ or exp short survey/ or exp health care survey/ or exp health survey/

34. 27 or 28 or 29 or 30 or 31 or 32 or 33

35. 3 and 18 and 26

36. 35 not 34

37. 36 not ((exp animal/ or nonhuman/) not exp human/)

38. limit 37 to english language

II. Search strategy in MEDLINE (OVID)

1. exp Asthma, Aspirin-Induced/ or exp Asthma, Exercise-Induced/ or exp Asthma/ or exp Asthma, Occupational/ or asthma*.ti,ab.

2. omalizumab.mp. or Omalizumab/

3. mepolizumab.mp.

4. reslizumab.mp.

5. benralizumab.mp.

6. dupilumab.mp.

7. tralokinumab.mp.

8. lebrikizumab.mp.

9. tezepelumab.mp.

10. brodalumab.mp.

11. ligelizumab.mp.

12. Pitrakinra.mp.

13. biological product/ or biological therapy/ or biologic*.mp.

14. (biologic* adj1 (treatment* or therap* or medicine* or drug* or agent* or product*)).mp.

15. monoclonal antibod*.mp. or antibodies, monoclonal/ or antibodies, monoclonal, humanized/

16. 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15

17. (responsive* or response or respond* or nonrespon*).mp.

18. treatment outcome/ or Outcome Assessment, Health Care/

19. Minimal Clinically Important Difference/ or meaningful change.mp.

20. (Minimal* adj1 (clinical* or important or real or significant) adj1 (change or difference)).mp.

21. (Minimal* adj1 clinical* adj1 (important or significant) adj1 (change or difference)).mp.

22. (MCID or MID or MIC).mp.

23. 17 or 18 or 19 or 20 or 21 or 22

24. editorial/ or review/ or case report/ or case report*.mp.

25. (editorial* or conference abstract* or conference paper*).mp.

26. ((systematic or narrative) adj2 review*).mp. or "systematic review"/

27. ((("semi-structured" or semistructured or unstructured or informal or "in-depth" or indepth or "face-to-face" or structured or guide) adj3 (interview* or discussion* or questionnaire*)) or (focus group* or qualitative or ethnograph* or fieldwork or "field work" or "key informant")).ti,ab. or survey*.ti.

28. interviews as topic/ or focus groups/ or narration/ or qualitative research/ or health care surveys/ or health surveys/

29. 24 or 25 or 26 or 27 or 28

30. 1 and 16 and 23

31. 30 not 29

32. 31 not (Animals/ not (Animals/ and Humans/))

33. limit 32 to english language

III. Search strategy in CINAHL (EBSCOhost)

1. (MH "Asthma+") OR (MH "Asthma, Occupational") OR (MH "Asthma, Exercise-Induced") OR TI asthma* OR AB asthma*

2. "omalizumab" OR "mepolizumab" OR "reslizumab" OR "benralizumab" OR "dupilumab" OR "tralokinumab" OR "lebrikizumab" OR "tezepelumab" OR "brodalumab" OR "ligelizumab" OR "Pitrakinra" (MH "Biological Therapy") OR (MH "Antibodies, Monoclonal+") OR ((biologic*) N1 (treatment* OR therap* OR medicine* OR drug* OR agent* OR product*)) OR "biologic*" OR "monoclonal antibod*"

3. "responsive*" OR "response" OR "respond*" OR "nonrespon*" OR (MH "Treatment Outcomes") OR (MH "Outcome Assessment")

4. "MCID" OR "MID" OR "MIC" OR "meaningful change" OR (Minimal* N1 (clinical* OR important OR real OR significant) N1 (change OR difference)) OR (Minimal* N1 clinical* N1 (important OR significant) N1 (change OR difference))

5. TI (("semi-structured" OR semistructured OR unstructured OR informal OR "in-depth" OR indepth OR "face-to-face" OR structured OR guide) N3 (interview* OR discussion* OR questionnaire*)) OR TI ("focus group*" OR qualitative OR ethnograph* OR fieldwork OR "field work" OR "key informant"))

6. AB (("semi-structured" OR semistructured OR unstructured OR informal OR "in-depth" OR indepth OR "face-to-face" OR structured OR guide) N3 (interview* OR discussion* OR questionnaire*)) OR AB ("focus group*" OR qualitative OR ethnograph* OR fieldwork OR "field work" OR "key informant")

7. (MH "Qualitative Studies") OR (MH "Focus Groups") OR (MH "Narratives") OR (MH "Interviews") OR (MH "Surveys") OR TI Survey*

8. (MH "Literature Review") OR (MH "Scoping Review") OR PT "Systematic Review" OR PT review OR PT editorial OR PT proceedings

9. S3 OR S4

10. S5 OR S6 OR S7 OR S8

11. S1 AND S2 AND S9

12. S11 NOT S10

13. (MH "Animals+") NOT (MH "Human")

14. S12 NOT S13 Limiters - English Language

IV. Search strategy in Web of science

1. TS=(asthma*)

2. TS=(omalizumab) OR TS=(mepolizumab) OR TS=(reslizumab) OR TS=(benralizumab) OR TS=(dupilumab) OR TS=(tralokinumab) OR TS=(lebrikizumab) OR TS=(tezepelumab) OR TS=(brodalumab) OR TS=(ligelizumab) OR TS=(Pitrakinra)

3. TS=((biologic*) NEAR/1 (treatment* OR therap* OR medicine* OR drug* OR agent* OR product*)) OR TS=("monoclonal antibod*") OR TS=("biologic*")

4. TS=("responsive*") OR TS=("response") OR TS=("respond*") OR TS=("nonrespon*") OR TS=("outcome assessment*") OR TS=("treatment outcome*") OR TS=("meaningful change") OR TS=(Minimal* NEAR/1 (clinical* OR important OR real OR significant) NEAR/1 (change OR difference)) OR TS=(Minimal* NEAR/1 clinical* NEAR/1 (important OR significant) NEAR/1 (change OR difference)) OR TS=("MCID") OR TS=("MID") OR TS=("MIC")

5. #3 OR #2

6. (#1 AND #4 AND #5) NOT TS=("interview*") NOT TS=("focus group*") NOT TS=(narration) NOT TS=("qualitative research") NOT TI=(survey*)

7. #6 NOT TS=((("semi-structured" OR semistructured OR unstructured OR informal OR "in-depth" OR indepth OR "face-to-face" OR structured OR guide) NEAR/3 (interview* OR discussion* OR questionnaire*)) OR (focus group* OR qualitative OR ethnograph* OR fieldwork OR "field work" OR "key informant"))

8. (#7 NOT TS=((animal*) NOT (human* OR patient*))) AND LANGUAGE: (English)

9. (#7 NOT TS=((animal*) NOT (human* OR patient*))) AND LANGUAGE: (English)

Refined by: [excluding] DOCUMENT TYPES: (PROCEEDINGS PAPER OR EDITORIAL MATERIAL OR REVIEW OR MEETING ABSTRACT)

Appendix 3. Detailed methods

Data extraction, risk of bias assessment, quality, and synthesis of the results.

Data extraction was based on the COSMIN (COnsensus-based Standards for the selection of Measurement Instruments) guideline[1] for outcome measures. Data about study design; population characteristics and subgroups including sample size; asthma definition and severity; intervention and comparator (where appropriate); follow-up period; methodological approach to defining therapeutic response; definition of response and non-response (sole or composite outcome measures), development data, data on measurement properties (including: reliability (internal consistency, reliability, measurement error), validity (content, construct validity, responsiveness to change)) and characteristics of the outcome measurements were extracted into a template form independently by two reviewers (EK, AR). Any discrepancies were resolved by discussion or by a third reviewer (GR). The final extraction was cross-checked. Authors of included studies were contacted to provide additional data if needed.

Two reviewers (EK,AR) independently assessed the Risk of Bias (RoB) in individual studies using the COSMIN checklist for PROMs[2,3] and composite outcome measures (COSMIN RoB for non-Patient Reported Outcomes)[4]. Criterion validity was not evaluated as no gold standard exists in severe asthma.

First, development of the outcome measures was assessed based on relevance, comprehensiveness, and comprehensibility according to ten criteria.[3] Each criterion was rated as positive (+), negative (-), or indeterminate (?). The overall rating was provided as sufficient (+), insufficient (-), or inconsistent (±) which were based on the results from developmental and content validity studies as well as reviewers rating. If the developmental process for an outcome measure was not reported, then the overall rating was based only on the reviewer rating.

Second, we assessed RoB for each measurement property in the validation studies and rated it as very good, adequate, doubtful, or inadequate. The overall rating per measurement property was determined by the lowest rating for each standard.[1,2] The RoB assessment of response definitions was not undertaken as it is not part of the COSMIN RoB checklist.

Furthermore, we applied quality criteria. Each measurement property was rated as either sufficient (+), insufficient (-), or indeterminate (?) based on the predefined criteria for good measurement properties (GMP).[1] For construct validity and responsiveness, the review team formulated *a priori* hypotheses about the expected relationships between an outcome measure and comparator instruments. Overall, ≥75% of the results were expected to meet the criteria to be classified as sufficient.[1] Criteria for GMP are listed in **Table S2**.

Lastly, the certainty of evidence was assessed using the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) approach.[1,3,5] Quality of evidence was rated as 'high', 'moderate', 'low' or

'very low' for four factors (RoB, inconsistency, imprecision, and indirectness) for 'validity' studies while for 'developmental' studies rating was done according to three (RoB, inconsistency, and indirectness) by two reviewers (EK, AR). Papers describing development of the outcome measure were eligible for inclusion regardless of severity of asthma but subsequently downgraded for indirectness. Only inconsistency, imprecision and indirectness were assessed for the definitions of response as per the COSMIN guideline.[1] GRADE was not assessed in studies with indeterminate (?) rating based on GMP.[1] Any disagreements were resolved through the consultation with a third reviewer (GR). A descriptive synopsis with summary data tables were produced, and results were summarized using narrative synthesis.

Table S1. COSMIN definitions of domains, measurement properties, and aspects of measurement properties.

Term			Definition
Domain	**Measurement Property**	**Aspect of a Measurement Property**	
Reliability			The degree to which the measurement is free from measurement error
Reliability (extended definition)			The extent to which scores for patients who have not changed are the same for repeated measurement under several conditions: e.g. using different sets of items from the same health related-patient reported outcomes (HR-PRO; internal consistency); over time (test-retest); by different persons on the same occasion (inter-rater); or by the same persons (i.e. raters or responders) on different occasions (intra-rater)
	Internal consistency		The degree of the interrelatedness among the items
	Reliability		The proportion of the total variance in the measurements which is due to ‘‘true’’ differences between patients
	Measurement error		The systematic and random error of a patient’s score that is not attributed to true changes in the construct to be measured
Validity			The degree to which an HR-PRO instrument measures the construct(s) it purports to measure
	Content validity		The degree to which the content of an HR-PRO instrument is an adequate reflection of the construct to be measured
		Face validity	The degree to which (the items of) an HR-PRO instrument indeed looks as though it is an adequate reflection of the construct to be measured
	Construct validity		The degree to which the scores of an HR-PRO instrument are consistent with hypotheses (for instance with regard to internal relationships, relationships to scores of other instruments, or differences between relevant groups) based on the assumption that the HR-PRO instrument validly measures the construct to be measured
		Structural validity	The degree to which the scores of an HR-PRO instrument are an adequate reflection of the dimensionality of the construct to be measured
		Hypotheses	Idem construct validity

		testing	
		Cross-cultural validity	The degree to which the performance of the items on a translated or culturally adapted HR-PRO instrument are an adequate reflection of the performance of the items of the original version of the HR-PRO instrument
	Criterion validity		The degree to which the scores of an HR-PRO instrument are an adequate reflection of a ‘‘gold standard’’
Responsiveness			The ability of an HR-PRO instrument to detect change over time in the construct to be measured
	Responsiveness		Idem responsiveness
Interpretability			Interpretability is the degree to which one can assign qualitative meaning—that is, clinical or commonly understood connotations—to an instrument’s quantitative scores or change in scores.

COSMIN, COnsensus-based Standards for the selection of health Measurement INstruments; HR PRO, health related-patient reported outcomes. Taken from Mokkink LB et al.[6]

Table S2. COSMIN criteria for good measurement properties.

Measurement property (definition)	Rating	Criteria
Structural validity	+	**CTT** CFA: CFI or TLI or comparable measure > 0.95 OR RMSEA < 0.06 OR SRMR < 0.08[a] **IRT/Rasch** No violation of unidimensionality[b]: CFI or TLI or comparable measure > 0.95 OR RMSEA < 0.06 OR SRMR < 0.08 AND no violation of local independence: residual correlations among the items after controlling for the dominant factor < 0.20 OR Q3's < 0.37 AND no violation of monotonicity: adequate looking graphs OR item scalability > 0.30 AND adequate model fit IRT: χ^2 > 0.001 Rasch: infit and outfit mean squares ≥ 0.5 and ≤ 1.5 OR Z-standardized values > −2 and < 2
	?	CTT: not all information for '+' reported IRT/Rasch: model fit not reported
	-	Criteria for '+' not met
Internal consistency	+	At least low evidence[c] for sufficient structural validity[d] AND Cronbach's alpha(s) ≥ 0.70 for each unidimensional scale or subscale[e]
	?	Criteria for "At least low evidence[c] for sufficient structural validity[d]" not met
	-	At least low evidence[c] for sufficient structural validity[d] AND Cronbach's alpha(s) < 0.70 for each unidimensional scale or subscale[e]
Reliability	+	ICC or weighted Kappa ≥ 0.70
	?	ICC or weighted Kappa not reported
	-	ICC or weighted Kappa < 0.70
Measurement error	+	SDC or LoA < MIC[d]
	?	MIC not defined
	_	SDC or LoA > MIC[d]

Hypotheses testing for construct validity	+	The result is in accordance with the hypothesis[f]
	?	No hypothesis defined (by the review team)
	-	The result is not in accordance with the hypothesis[f]
Responsiveness to change	+	The result is in accordance with the hypothesis[f] OR AUC ≥ 0.70
	?	No hypothesis defined (by the review team)
	-	The result is not in accordance with the hypothesis[f] OR AUC < 0.70

AUC, area under the curve; CFA, confirmatory factor analysis; CFI, comparative fit index; CTT, classical test theory; DIF, differential item functioning; ICC, intraclass correlation coefficient; IRT, item response theory; LoA, limits of agreement; MIC, minimal important change; RMSEA, root mean square error of approximation; SEM standard error of measurement; SDC, smallest detectable change; SRMR, standardized root mean residuals; TLI, Tucker–Lewis index. Taken from COSMIN, COnsensus-based Standards for the selection of health Measurement INstruments[1].

"+" = sufficient, "–" = insufficient, "?" = indeterminate

[a]To rate the quality of the summary score, the factor structures should be equal across studies

[b]Unidimensionality refers to a factor analysis per subscale, while structural validity refers to a factor analysis of a (multidimensional) patient reported outcome measure

[c]As defined by grading the evidence according to the GRADE approach

[d]This evidence may come from different studies

[e]The criteria 'Cronbach alpha < 0.95' was deleted, as this is relevant in the development phase of a PROM and not when evaluating an existing PROM

[f]The results of all studies should be taken together and it should then be decided if 75% of the results are in accordance with the hypotheses

Table S3. Approach to development of outcome measures.

Reference, year	Scale	Approach to development of outcome measurements
Composite outcome measures		
Fitzpatrick, 2020[7]	ASSESS	Adapted from the CASI by clinicians only: removed daytime symptoms and night time symptom dimensions and replaced with the total ACT score (weighted at 30%), modified ranges for FEV_1, medications, and length for assesment of exacerbations.
Wildfire 2012[8]	CASI	Developed by physicians only. 1. Determining independent dimensions of asthma severity via factor analysis. 2. Delphi exercise: clinical weighting of the dimensions of asthma severity. 3. Scale properties of the Composite Asthma Severity Index. 4. External validation.
De Llano, 2021[9]	FEOS	Developed by physicians only. 1. Systematic literature review. 2. Selection of domains and measurement tools: Delphi exercise. 3.Weighted of selected domains: multicriteria decision analysis. 4.Face validity.
Asthma symptom outcome measures		
Shen, 2021[10] Revicki, 1998[11]	ASUI	1. Literature review, patient interviews (including ranking order the relative importance of the items) and discussion with physicians. 2. Determination of a scoring algorithm using visual analog scale and standard gamble techniques, subsequently using multi-attribute utility function.
Shen, 2021[10]	ASI	Modified version of the ASUI which includes the 4 asthma symptoms, but excludes questions about assessment of medication side effects (eg, "how many days were you bothered by side effects of your asthma medication during the past 2 weeks?," "if 1 day or more what side effects did you have?," and "on average, how severe were the side effects of your asthma medication during the past 2 weeks?").
Globe, 2015 [12] Globe, 2019[13]	ASD	1. Concept elicitation interviews in 34 adults (38.9 years (13.0), 61.8% females, ACQ≥3 in 20.6%) and 16 adolescents (15.2 years (1.6), 56.3% males, ACQ≥3 in 31.3%) with clinical diagnosis of persistent asthma. 2. Cognitive interviews in 15 adults (30.7 years (9.7), 86.7% females, ACQ≥3 in 20.0%) and 9 adolescents (14.1 years (2.2), 77.8% males, ACQ≥3 in 11.1%) with a clinical diagnosis of persistent asthma.
Asthma control outcome measures		
Lloyd, 2007[14]	GETE	Developed by physicians only
Asthma quality of life measures		
Hyland, 2018[15]	SAQ	1. Identification of domains of an instrument. 2.Focus group to seek feedback about draft instrument: patient with severe asthma defined by BTS guideline (n=16) between 24-69 y.o; mean age of 47 (SD = 13.53); female (n=12).

ACQ, Asthma Control Questionnaire; BTS, British Thoracic Society; GETE, Global Evaluation of Treatment Effectiveness; ASSESS, Asthma Severity Scoring System; ASUI, Asthma Symptom Utility Index; ASI, Asthma Symptom Index; ASD, Asthma Symptom Diary; ACT, Asthma Control Test; CASI, Composite Asthma Severity Index; FEOS, FEV1, Exacerbations, Oral Corticosteroids, Symptoms Score; SAQ, Severe Asthma Questionnaire; NR, Not reported; FEV1, Forced Expiratory Volume in 1 second.

Table S4. Summary of characteristics of the outcome measures.

Instrument (year)	Mode of administration	(Sub)scale(s) (No. of Items)	Type of response categories	Intended context of use	Target population	Time to complete (minutes)	Patient/carer report	Original language
Composite outcome measures								
Fitzpatrick, 2020[7] ASSESS	Interviewer administered, paper form (ACT[16-18]: self, at-home paper, phone, mail)	4 items: ACT (5 items), FEV_1, current medications, exacerbations.	Multiple choice questions	Clinical trials and routine clinical practice	Adolescents (≥12 years) and adults	Not reported (ACT: 2 min)	Patient and clinician	English
Wildfire,2012[8] CASI	Interviewer administered, paper form, online calculator available	5 domains: day symptoms and albuterol use, night symptoms and albuterol use, controller treatment, lung function measures, and exacerbations.	Multiple choice questions	Intervention studies and clinical practice	Children ≥ 6 years and adolescents*	Not reported	Patient and clinician	English
de Llano, 2021[9] FEOS	Paper (ACT[16-18] : self, at-home paper, phone, mail)	4-items (OCS, severe exacerbations, ACT, FEV_1)	Multiple choice questions	Clinical trials, patient monitoring	Adults	Not reported (ACT: 2 min)	Patient and clinician	English
Asthma symptom outcome measures								
Revicki, 1998[11] ASUI	Interviewer administered, paper form	11 items [four symptoms (cough, wheeze, shortness of breath, and awakening at night) and two dimensions (frequency and severity] and side effect of medications	4-point Likert scale	Clinical trials and cost effectiveness studies	Adults	Not reported	Patient	English (for the USA). Italian, French
Shen, 2021[10] ASI	Interviewer administered, paper	8 items [four symptoms (cough, wheeze, shortness of breath, and awakening at night) and two dimensions (frequency and severity]	4-point Likert scale	Clinical trials, patient monitoring	Adults	Not reported	Patient	English, Italian, French
Globe,2015[12] ASD	Self-complete, electronic device	10-items (5 morning and 5 evening)	5-point Likert scale	Clinical research	Adolescents (≥ 12 years) and adults	Not reported	Patient	English
Asthma control outcome measures								
Llyod, 2007[14] GETE	Interviewer administered, paper form	2 items	5-point Likert scale	Clinical trials and routine clinical practice	Adolescents and adults	Not reported	Patient and clinician	English
Asthma quality of life measures								
Hyland, 2018[15] SAQ	Self-complete, paper form	SAQ: 16 items SAQ-global: 1 item	7-point Likert scale	Clinical research, patient monitoring	Adults 16–78 years (reading age 11-12 years)	3-6 minutes	Patient	English (UK), Portuguese

ACT, Asthma Control Test; ASUI, Asthma Symptom Utility Index; ASI, Asthma Symptom Index; ASD, asthma symptom diary; ASSESS, Asthma Severity Scoring System; CASI, Composite Asthma Severity Index; GETE, Global Evaluation of Treatment Effectiveness; SAQ, Severe Asthma Questionnaire. FEOS, FEV1, Exacerbations, Oral Corticosteroids, Symptoms Score; FEV_1, Forced Expiratory Volume in 1 second; OCS, Oral Corticosteroids. * CASI is also validated in adults with asthma based on a conference abstract.[19]

Table S5. Summary of data for measurement properties of outcome measures.

Reference, year	Construct validity**	Reproducibility	Internal consistency	Responsiveness
Lloyd, 2007[14] GETE	1.Spearman rank-order correlation between GETE and AQLQ (physician GETE / patient GETE)*: • Activities score: –0.29 / –0.32 • Change from baseline in activities score: –0.35 / –0.37 • Emotions score: –0.36 / –0.37 • Change from baseline in emotions score: –0.31 / –0.35 • Environmental exposure score:–0.25 / –0.26 • Change from baseline in environmental exposure score: –0.27 / –0.30 • Symptom score –0.40 / –0.45 • Change from baseline in symptom score: –0.36 / –0.39 • Overall score: –0.38 /–0.41 • Change from baseline in overall score: –0.38 /–0.41 * All correlations were p<0.0001. 2. Spearman rank-order correlation between GETE and clinical characteristics (physician GETE / patient GETE)*: • Actual FEV1 value: –0.20/–0.14 • Total asthma symptom score: 0.32/ 0.34 • Change in total asthma symptom score: 0.26/ 0.31 • Nocturnal symptom score: 0.22/ 0.22 • Change in nocturnal symptom score: 0.21/ 0.23 • Daytime symptom score: 0.31/ 0.34 • Change in daytime symptom score: 0.24/ 0.29 • No. of puffs of rescue medication/day: 0.33 /0.33 • Change in no. of puffs of rescue medication/day: 0.26/ 0.29 * All correlations were p<0.0001. 3. Actual mean FEV1 (SD)/ mean total asthma symptom score (SD)/ mean nocturnal symptom score (SD) / mean daytime symptom score (SD) / mean n on puffs of rescue meds (SD)	NA	NA	NA

Reference, year	Construct validity**	Reproducibility	Internal consistency	Responsiveness
	Patient version • Complete control of asthma: 2.20 (824.58) / 1.49 (1.58) / 0.50 (0.63) / 0.68 (0.71) / 3.23 (4.49) • Marked improvement of asthma: 2.12 (776.94) / 2.14 (1.85) / 0.69 (0.81) / 1.02 (0.86) / 3.76 (4.99) • Discernible, but limited improvement in asthma: 2.07 (761.41) / 2.70 (1.99) / 0.91 (0.96) / 1.38 (0.98) / 5.47 (6.84) • No appreciable change in asthma: 2.03 (838.37) / 2.98 (2.21) / 1.01 (1.09) / 1.48 (1.05) / 5.20 (5.20) • Worsening of asthma: 1.82 (691.97) / 5.38 (3.39) / 2.06 (1.34) / 2.32 (1.46) / 13.23 (7.83) p values per clinical indicator: 0.37/ 0.0091/ <0.0001/ <0.0001/ 0.0002/ 0.0016 / < 0.0001 / 0.0009 / 0.0002 *Physician version* • Complete control of asthma: 2.37 (877.81) / 1.68 (1.73)/ 0.64 (0.70) / 0.74 (0.75) / 3.13 (4.17) • Marked improvement of asthma: 2.15 (790.23) / 2.01 (1.83) / 0.61 (0.81) / 1.00 (0.88) / 3.65 (5.66) • Discernible, but limited improvement in asthma: 2.08 (751.92) / 2.61 (1.90) / 0.83 (0.87)/ 1.27 (0.90)/ 4. 93 (5.66) • No appreciable change in asthma: 1.95 (751.86) / 3.15 (2.34) / 1.15 (1.13) / 1.58 (1.12) / 6.35 (5.98) • Worsening of asthma: 1.66 (445.85)/ 6.41 / 1.38 (1.95) / 2.63 / 16.12 (11.49) P values per clinical indicator: 0.0091; < 0.0001/ 0.0016/ <0.0001/ 0.0002 4. Data presented per GETE level by AQLQ mean activity score (SD)/Mean emotions score (SD) /Mean environment score (SD) / Mean symptoms score (SD) / Mean overall score (SD) *Patient version GETE* • Complete control of asthma: 5.74 (1.21) / 5.83 (1.19) / 5.52 (1.37) / 5.75 (1.07) / 5.73 (1.07)			

Reference, year	Construct validity**	Reproducibility	Internal consistency	Responsiveness
	• Marked improvement of asthma: 5.15 (1.21) / 5.29 (1.30) / 4.89 (1.34) / 5.15 (1.08) / 5.13 (1.06) • Discernible, but limited improvement in asthma: 4.76 (1.25) / 4.72 (1.43) / 4.56 (1.43) / 4.58 (1.13) / 4.64 (1.12) • No appreciable change in asthma: 4.45 (1.33) / 4.33 (1.47) / 4.43 (1.35) / 4.22 (1.17) / 4.31 (1.10) • Worsening of asthma: 4.40 (1.47) / 3.88 (1.57) / 4.33 (1.55) / 3.76 (1.24) / 4.03 (1.19) <u>*Physician version GETE*</u> • Complete control of asthma: 5.73 (1.22) / 5.85 (1.17) / 5.50 (1.38) / 5.72 (1.05) / 5.71 (1.06) • Marked improvement of asthma: 5.21 (1.25) / 5.38 (1.27) / 4.99 (1.35) / 5.23 (1.09) / 5.20 (1.07) • Discernible, but limited improvement in asthma: 4.79 (1.26) / 4.72 (1.49) / 4.59 (1.42) / 4.60 (1.21) / 4.67 (1.17) • No appreciable change in asthma: 4.56 (1.29) /4.54 (1.42) /4.48 (1.40) /4.37 (1.16) /4.45 (1.09) Worsening of asthma: 4.42 (1.40)/ 3.29 (1.32) /4.04 (1.46) / 3.70 (1.00) / 3.90 (1.10)			
Fitzpatrick, 2020[7] ASSESS	• AQLQ total score: r= -0.315** • AQLQ symptom: r= -0.387** • AQLQ activity: r= -0.244* • AQLQ emotion: r= -0.387** • AQLQ environment: r= -0.253* *P < .05 and **P < .01.	ICC (baseline/ 12mo; 12mo/24 mo; 24mo/36 mo) • *Entire sample* 0.764/ 0.768/ 0.813 • *12-17 ys*: 0.717/ 0.841/ 0.732 • *>18 y:* 0.768 / 0.766/ 0.816	Cronbach's alpha: *entire sample* 0.639 *12-17y:* 0.468 *≥18 y:* 0.662	<u>1. r values: AQLQ total score / symptom / activity / emotion / environment:</u> • *0-12 mo:* -0.550* / -0.579* / -0.453* / -0.488* / -0.300* • *12 - 24 mo:* -0.462* / -0.508* / -0.349* / -0.408* / -0.212* • *24 - 36 mo:* -0.468* / -0.481* / -0.396* / -0.368* / -0.265* *P < .001. <u>2. r values for changes: 0 and 12 months / 12 and 24 months/ 24 and 36 months:</u> • Change in ASSESS vs Change in ACT: -0.668* / -0.676* / -0.622

Reference, year	Construct validity**	Reproducibility	Internal consistency	Responsiveness
				• Change in ASSESS vs Change in FEV1 absolute % difference: -0.395* / -0.369* / -0.372*.
Wildfire, 2012[8] CASI*				Intervention group showed improvement in CASI & symptom days (0.67 points & 0.48-day improvement; both P < .001). CASI: 32% greater magnitude of improvement (standardized effect size: 0.25 vs 0.17 for symptom days)
Shen, 2021[10] ASUI	1.ASUI baseline/ week 12: SGRQ score: -0.68 / -0.72 SGRQ Symptom: -0.78 / -0.81 SGRQ Impact: -0.46 / -0.56 SGRQ Activity: -0.60 / -0.66 ACQ-5 score: -0.78 / -0.85 EQ-5D index score: 0.51 / 0.52 EQ-5D VAS score: 0.44 / 0.56 % FEV1 pred.: 0.19 / 0.28 FEV1 (mL): 0.15 / 0.20 No. of exacerbations: -0.15 / -0.29 Global rating of activity limitation: -0.43 / -0.51 ASD Score: -0.54 / -0.53 2.Known group validity: Group with higher ACQ-5 scores (≥1.5 indicating poorly controlled asthma) tended to have lower ASUI scores (indicative of greater symptom burden) (p<0.0001). For % pred FEV1, group with lowest FEV1 function (≤ 60%) had the lowest ASUI scores (p<0.0001).	ICC=0.87-0.90	Cronbach's alpha: Baseline=0.87 Week 12 =0.90	1.ASUI change from baseline to week 4: ΔACQ-5 score: - 0.57 ΔSGRQ score: 0.50 ΔSGRQ Symptom: -0.53 ΔSGRQ Impact: -0.25 ΔSGRQ Activity: -0.41 Δ % predicted FEV1: 0.16 No. of asthma exacerbations during on-treatment phase: -0.02 2. ASUI change from baseline to week 12: ΔACQ-5 score: -0.67 ΔSGRQ score: -0.60 ΔSGRQ Symptom: -0.67 ΔSGRQ Impact: -0.42 ΔSGRQ Activity: -0.50 Δ % predicted FEV1: 0.25 No. of asthma exacerbations during on-treatment phase: -0.05
Shen, 2021[10] ASI	1.ASI (baseline/week 12): SGRQ score: 0.67/ 0.71 SGRQ Symptom: 0.80 / 0.82	ICC=0.87-0.90	Cronbach's alpha:	1.ASI change from baseline to week: 4: ΔACQ-5 score: 0.58

Reference, year	Construct validity**	Reproducibility	Internal consistency	Responsiveness
	SGRQ Impact: 0.46 / 0.55 SGRQ Activity: 0.59 / 0.65 ACQ-5 score: 0.79 / 0.85 EQ-5D index score: -0.49/ -0.49 EQ-5D VAS score: -0.43/ -0.55 % FEV1 pred.: -0.20/ -0.28 FEV1 (mL): -0.14/ -0.19 No. of exacerbations: 0.12 / 0.28 Global rating of activity limitation: 0.43 / 0.49 ASD Score: 0.54 / 0.52 / <u>2. Known group validity:</u> Group with higher ACQ-5 scores (≥1.5 indicating poorly controlled asthma) tended to have higher ASI scores (p<0.0001). For % pred FEV1, group with lowest FEV1 function (≤60%) had the highest ASI scores (p<0.0001).		Baseline=0.89, Week 12=0.93	ΔSGRQ score: 0.50 ΔSGRQ Symptom: 0.55 ΔSGRQ Impact: 0.27 ΔSGRQ Activity: 0.39 Δ % predicted FEV1: -0.18 No. of asthma exacerbations during on-treatment phase: 0.05 <u>2.ASI change from baseline to week 12</u>: ΔACQ-5 score: 0.69 ΔSGRQ score: 0.61 ΔSGRQ Symptom: 0.70 ΔSGRQ Impact: 0.45 ΔSGRQ Activity: 0.49 Δ % predicted FEV1: -0.28 No. of asthma exacerbations during on-treatment phase*: 0.09
Hyland, 2018[20] Masoli, 2021[21] Lanario, 2021[22] SAQ	<u>1. SAQ vs</u> miniAQLQ =0.76; ACT=0.68; EQ-5D-5L score=−0.76; EQ-5D-VAS= 0.71; SAQ-global scale= 0.72; FEV1 % predicted=0.27; BMI=−0.31 <u>2. SAQ-global vs</u> MiniAQLQ= 0.71; ACT total= 0.68; EQ-5D-5L= −0.71; EQ-5D-VAS= 0.76; FEV1 % predicted=0.26; BMI=−0.22 <u>3. Data for FEV1% predicted vs SAQ domains:</u> SAQ score: 0.23; SAQ My Life: 0.29; SAQ My Mind: 0.15; SAQ My Body: 0.15; SAQ global score: 0.28 <u>4.Data for cumulative prednisolone vs SAQ domains:</u> SAQ score: -0.34; SAQ My Life: − 0.35; SAQ My Mind: − 0.23; SAQ My Body: − 0.34; SAQ global score: − 0.37 <u>5.Data for Exacerbations in the last 12 mo requiring OCS vs SAQ domains:</u>	ICC= 0.93 (SAQ) ICC= 0.93 (SAQ-global)	Cronbach's alpha= 0.93.	Change scores for different degrees of global rating of change is available for SAQ, SAQ subscales and SAQ-global.

Reference, year	Construct validity**	Reproducibility	Internal consistency	Responsiveness
	SAQ score: -0.37; SAQ My Life: − 0.37; SAQ My Mind: − 0.33; SAQ My Body: − 0.33; SAQ global score: − 0.36 6. Data for Hospital admissions in the last 12 mo vs SAQ domains: SAQ score: -0.17; SAQ My Life: − 0.16; SAQ My Mind: − 0.16; SAQ My Body: − 0.13; SAQ global score: − 0.23 7. EQ-5D-5L Index value/EQ-5D-5L item 5–Anxiety and Depression/EQ-5D VAS/ ACQ score/ACT total SAQ score:0.72/ -0.64 /0.73/ -0.75/0.71 SAQ My Life: 0.73/-0.54/0.74/-0.79/0.72 SAQ My Mind: 0.64/-0.73/0.63/ -0.62/ 0.62 SAQ My Body: 0.59/-0.56/0.62/-0.60/ 0.64 SAQ global score: 0.66/-0.50/ 0.79/ 0.77/ 0.68			
Globe, 2019[13] ASD				1. Responsiveness of the Average 7-Day ASD Score at Weeks 12 and 24 Data presented for Responders Mean (SE) Non-Responders/ Mean (SE) Difference P-Value. Effect size presented for responder / nonresponder Week 12 ACQ > 0.5: −0.49 (0.03) / 0.05 (0.03).Effect size: 0.82 / 0.08 ACQ > 1.0: −0.54 (0.03) / −0.13 (0.03).Effect size: 0.90 / 0.22 PGA: −0.48 (0.03) / −0.07 (0.03) Effect size: 0.80 / 0.12 Week 24: ACQ > 0.5: −0.59 (0.03) / −0.06 (0.03) / − 0.53. Effect size: 0.98 / 0.10 ACQ > 1.0: −0.68 (0.04) / −0.15 (0.03) / − 0.53.Effect size: 1.13 / 0.25

Reference, year	Construct validity**	Reproducibility	Internal consistency	Responsiveness
				PGA: −0.60 (0.03) / −0.10 (0.04) / −0.49.Effect size: 1.00 / 0.17 2. Responsiveness of ASD Symptomatic Days in a 7-Day Period at Weeks 12 and 24 Data presented for Responders Mean (SE) Non-Responders Mean (SE).Effect size presented for responder / nonresponder: Week 12: ACQ > 0.5: −2.21 (0.16) / −0.57 (0.18).Effect size: 0.73 / 0.19 ACQ > 1.0: −2.35 (0.20) / −0.90 (0.16).Effect size: 0.78 / 0.30 PGA: −2.34 (0.16) / −0.45 (0.17) Effect size 0.78 / 0.15 Week 24: ACQ > 0.5: −2.86 (0.18) / − 0.28 (0.28).Effect size 0.95 / 0.09 ACQ > 1.0: −3.21 (0.21) / −0.77 (0.20).Effect size 1.07 / 0.26 PGA: −2.97 (0.19) / −0.45 (0.23) Effect size 0.99 / 0.15 3. Spearman correlations between baseline to 12-week changes in ASD scores and baseline to 12-week changes in ACQ and PGA scores were 0.59 and 0.57, respectively. 4. Correlations between baseline to 24-week changes in ASD scores and

Reference, year	Construct validity**	Reproducibility	Internal consistency	Responsiveness
				baseline to 24-week changes in ACQ and PGA scores were 0.67 and 0.53, respectively.

ACQ, Asthma Control Questionnaire; ACT, Asthma Control Test; AQLQ, Asthma Quality of Life Questionnaire; ASSESS, Asthma Severity Scoring System; ASUI, Asthma Symptom Utility Index; ASI, Asthma Symptom Index; ASD, Asthma Symptom Diary; BMI, Body Mass Index; CASI, Composite Asthma Severity Index; EQ-5D-5L,EuroQol Questionnaire-5 Dimensions-5 Levels; EQ-5D-VAS, EuroQol Questionnaire-5 Dimensions Visual Analogue Scale; GETE, Global Evaluation of Treatment Effectiveness; FEV1, forced expiratory volume in 1 second; ICC, intraclass correlation coefficient; miniAQLQ, mini- Asthma Quality of Life Questionnaire; PGA, Patient's Global Assessment; SAQ, Severe Asthma Questionnaire; SGRQ, St George's Respiratory Questionnaire. *Only external validation data was used for analysis as it was performed in a study with biologics. **As there is no golden standard in asthma, data about criterion validity was combined with construct validity.

Table S6. Additional study characteristics for validation studies.

Reference, year	Scale	Study design	N	Age (years) Mean (IQR)	Patient characteristics	Asthma severity (severe %)	Definition of asthma	Biological drug
Hyland, 2018[20]	SAQ	Observational	160	51	F=66%; FEV_1 % predicted=72 (28–137)	Severe (100%)	ERS/ATS guidelines	Omalizumab =21% Mepolizumab=3%
Lanario, 2021[22]	SAQ	Cross-sectional	460	51 (50–53)	F=65%; FEV_1 % predicted, mean (CI): 71.75 (69.79–73.71) Prescribed maintenance OCS, n (%): 218 (47)	Severe (100%)	ERS/ATS guidelines	Different biologics=39%
Wildfire, 2012[8]*	CASI	RCT	419	10.8 (8-14)	F= 42%; FEV_1 % predicted (mean ± SD) = 92.1±17.1	Mild to severe (54%)	Physician-diagnosis of asthma	Omalizumab=50%

ATS, American Thoracic Society; CASI, Composite Asthma Severity Index; ERS, European Respiratory Society; FEV1, forced expiratory volume in 1 second; F, females; SAQ, Severe Asthma Questionnaire; IQR, interquartile range; CI, confidence interval; SD, standard deviation; OCS, oral corticosteroids; RCT, Randomised Control Trial. *Only external validation data was used for analysis as it was performed in a study with biologics.

Table S7. Risk of bias assessment.

	ASSESS [7]	CASI[8]*	FEOS[9]	ASUI[10,11]	ASI[10]	ASD[12,13]	GETE[14]*	SAQ[15,20-22]**
PROM development	I	I	I	D	I	D	I	V
Structural validity								
Internal consistency	I			D	V			
Cross-cultural validity								
Reliability	I			A	A			A
Measurement error	I			A	A			
Construct validity	A			D	D		V	D
Responsiveness	D	V		D	A	D		D

ASSESS, Asthma Severity Scoring System; ASUI, Asthma Symptom Utility Index; ASI, Asthma Symptom Index; ASD, Asthma Symptom Diary; CASI, Composite Asthma Severity Index; GETE, Global Evaluation of Treatment Effectiveness; FEOS, FEV1, Exacerbations, Oral Corticosteroids, Symptoms Score; SAQ, Severe Asthma Questionnaire. *Only external validation data was used for analysis as it was performed in a study with biologics. Risk of bias in individual studies was investigated using the COSMIN checklist for PROMs[2,3] and composite outcome measures (COSMIN RoB for non-PROMs)[4]. V= very good; A = adequate; D = doubtful; I = inadequate. ** SAQ is based on a formative model; therefore, there was no need to investigate the internal consistency. Empty cells indicate that the measurement property was not investigated.

References

1. Prinsen CAC, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res.* 2018;27(5):1147-1157.
2. Mokkink LB, de Vet HCW, Prinsen CAC, et al. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. *Qual Life Res.* 2018;27(5):1171-1179.
3. Terwee CB, Prinsen CAC, Chiarotto A, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Qual Life Res.* 2018;27(5):1159-1170.
4. Mokkink LB, Boers M, van der Vleuten CPM, et al. COSMIN Risk of Bias tool to assess the quality of studies on reliability or measurement error of outcome measurement instruments: a Delphi study. *BMC Med Res Methodol.* 2020;20(1):293.
5. Guyatt G, Oxman AD, Akl EA, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol.* 2011;64(4):383-394.
6. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol.* 2010;63(7):737-745.
7. Fitzpatrick AM, Szefler SJ, Mauger DT, et al. Development and initial validation of the Asthma Severity Scoring System (ASSESS). *Journal of Allergy and Clinical Immunology.*145(1):127-139.
8. Wildfire JJ, Gergen PJ, Sorkness CA, et al. Development and validation of the Composite Asthma Severity Index--an outcome measure for use in children and adolescents. *J Allergy Clin Immunol.* 2012;129(3):694-701.
9. Perez de Llano L, Davila I, Martinez-Moragon E, et al. Development of a Tool to Measure the Clinical Response to Biologic Therapy in Uncontrolled Severe Asthma: The FEV1, Exacerbations, Oral Corticosteroids, Symptoms Score. *J Allergy Clin Immunol Pract.* 2021.
10. Shen Q, von Maltzahn R, Nelsen L, Revicki D. Psychometric Properties of the Asthma Symptom Index in Patients with Severe Asthma. *J Allergy Clin Immunol Pract.* 2021;9(1):400-409 e401.
11. Revicki DA, Leidy NK, Brennan-Diemer F, Sorensen S, Togias A. Integrating patient preferences into health outcomes assessment: the multiattribute Asthma Symptom Utility Index. *Chest.* 1998;114(4):998-1007.
12. Globe G, Martin M, Schatz M, et al. Symptoms and markers of symptom severity in asthma--content validity of the asthma symptom diary. *Health Qual Life Outcomes.* 2015;13:21.
13. Globe G, Wiklund I, Mattera M, Zhang H, Revicki DA. Evaluating minimal important differences and responder definitions for the asthma symptom diary in patients with moderate to severe asthma. *J Patient Rep Outcomes.* 2019;3(1):22.
14. Lloyd A, Turk F, Leighton T, Walter Canonica G. Psychometric evaluation of global evaluation of treatment effectiveness: a tool to assess patients with moderate-to-severe allergic asthma. *Journal of Medical Economics.* 2007;10(3):285-296.
15. Hyland ME, Lanario JW, Pooler J, Masoli M, Jones RC. How patient participation was used to develop a questionnaire that is fit for purpose for assessing quality of life in severe asthma. *Health Qual Life Outcomes.* 2018;16(1):24.
16. Nathan RA, Sorkness CA, Kosinski M, et al. Development of the asthma control test: a survey for assessing asthma control. *J Allergy Clin Immunol.* 2004;113(1):59-65.
17. Schatz M, Kosinski M, Yarlas AS, Hanlon J, Watson ME, Jhingran P. The minimally important difference of the Asthma Control Test. *J Allergy Clin Immunol.* 2009;124(4):719-723 e711.
18. Schatz M, Sorkness CA, Li JT, et al. Asthma Control Test: reliability, validity, and responsiveness in patients not previously followed by asthma specialists. *J Allergy Clin Immunol.* 2006;117(3):549-556.
19. Pelletier T, Karagic M, Krouse R, et al. Validating the Composite Asthma Severity Index in adult asthmatics. *Journal of Allergy and Clinical Immunology.* 2019;143(2):AB11.
20. Hyland ME, Jones RC, Lanario JW, Masoli M. The construction and validation of the Severe Asthma Questionnaire. *Eur Respir J.* 2018;52(1).
21. Masoli M, Lanario JW, Hyland ME, et al. The Severe Asthma Questionnaire: sensitivity to change and minimal clinically important difference (MCID). *Eur Respir J.* 2021.
22. Lanario JW, Hyland ME, Menzies-Gow A, et al. Validation of subscales of the Severe Asthma Questionnaire (SAQ) using exploratory factor analysis (EFA). *Health Qual Life Outcomes.* 2020;18(1):336.