Pragmatic and Observational Research

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Mapping EQ-5D-5L Score From SGRQ in Patients with Asthma and/or COPD in NOVELTY

Dan Jackson[1], Anna Quinton[2], Fanni Zhang[3], Hana Müllerová[4], Christer Janson[5], Mohsen Sadatsafavi[6] On behalf of the NOVELTY Scientific Community and NOVELTY study investigators

[1]Statistical Innovation Group, AstraZeneca, Cambridge, UK; [2]Global Health Economics and Payer Evidence, BioPharmaceuticals Business Unit, AstraZeneca, Cambridge, UK; [3]Statistical Innovation Group, AstraZeneca, Gaithersburg, MD, USA; [4]Respiratory Evidence Strategy, BioPharmaceuticals Medical, AstraZeneca, Cambridge, UK; [5]Department of Medical Sciences, Respiratory, Allergy and Sleep Research, Uppsala University, Uppsala, Sweden; [6]Respiratory Evaluation Sciences Program, Collaboration for Outcomes Research and Evaluation, Faculty of Pharmaceutical Sciences, University of British Columbia, Vancouver, BC, Canada

Correspondence: Dan Jackson, Statistical Innovation, AstraZeneca, Academy House, 136 Hills Road, Cambridge, CB2 8PA, UK, Email daniel.jackson1@astrazeneca.com

Purpose: The St George's Respiratory Questionnaire (SGRQ) measures health status in obstructive airways disease. Starkie et al proposed an algorithm for mapping the SGRQ to EQ-5D-5L, a preference-based utility measure, in chronic obstructive pulmonary disease (COPD) (Value Health 2011;14:354–60); only SGRQ total score, its squared value, and sex were included as covariates. We aimed to determine if including additional covariates could improve the performance of this algorithm type and whether amendments were required to extend this mapping to asthma or asthma+COPD.

Patients and Methods: SGRQ and EQ-5D-5L were measured from a large, global, prospective, longitudinal study in asthma and/or COPD (NOVELTY; NCT02760329). We fitted six longitudinal linear mixed models to the development sample (baseline and Year 1 data), with EQ-5D-5L as the response variable. Each model had a different combination of covariates. Mixed model repeated measures methodology was used to enable the accommodation of within-patient correlation among measurements. Restricted maximum likelihood and an unstructured covariance matrix were used to fit all models. Performance (mean square errors [MSE]) was evaluated relative to the Starkie et al algorithm in the validation sample (Year 2 and Year 3 data).

Results: A total of 6813 patients (asthma: 3546; asthma+COPD: 872; COPD: 2395) with available EQ-5D-5L and SGRQ data were included at baseline. MSEs indicated good performance, were similar across models (Year 2: 0.0302–0.0308 [45–46% variance explained]; Year 3: 0.0272–0.0277 [47–48% variance explained]), and were modestly smaller than those obtained by Starkie et al (Year 2: 0.0340; Year 3: 0.0296). Performance was similar across models in the asthma and COPD subgroups.

Conclusion: Including additional covariates and SGRQ domains resulted in similar model performance to Starkie et al, suggesting their covariates are adequate for mapping in asthma and/or COPD. NOVELTY coefficients broaden the population with chronic airways disease for whom this mapping can be applied.

Keywords: health status, utility mapping, longitudinal analysis, algorithm, economic modeling, quality of life

Introduction

Preference-based utility measures are often used in economic evaluation of health technologies as they enable the comparison of such technologies across different disease areas.[1] An example of a preference-based instrument that is widely used for utility elicitation is the EuroQol 5-Dimension 5-Level (EQ-5D-5L).[2] EQ-5D-5L is a simple, generic measure of health status covering the following five domains: mobility, self-care, usual activities, pain/discomfort and anxiety/depression, with five response levels ranging from experiencing no problems to experiencing extreme problems.[2] In contrast, disease-specific symptom or functional scores are used in clinical studies across disease areas, for example, as key endpoints in clinical trials.[3] For airways diseases, the St George's Respiratory Questionnaire (SGRQ) has been widely used in clinical studies as a comprehensive and validated disease-specific instrument for the evaluation of patient

Received: 11 December 2024
Accepted: 14 May 2025
Published: 4 June 2025

health status.[4–8] The SGRQ is a 50-item questionnaire with questions covering three domains of symptoms, activity, and impacts (psychosocial).[9,10]

Where EuroQoL 5-Dimension (EQ-5D) utility scores from validated instruments are not available, they can be predicted by mapping responses to disease-specific scores.[11] Given the extensive adoption of SGRQ to evaluate patient health status in multiple clinical studies,[4–8] algorithms for mapping SGRQ score to preference-based utility measures can have broad applicability. One such mapping algorithm has been developed and validated by Starkie et al in patients with moderate-to-very severe chronic obstructive pulmonary disease (COPD).[12] Starkie et al proposed a mapping formula that included only SGRQ total score, its squared value, and sex.[12]

The Starkie et al algorithm was derived using data from Towards a Revolution in COPD Health (TORCH), a clinical trial of patients with COPD from around two decades ago.[12] The validity of this mapping algorithm in more contemporary samples, as well as in patients with asthma, is not known, even though the algorithm has subsequently been used in patients with severe eosinophilic asthma.[13] Furthermore, more nuanced mapping, such as algorithms that use SGRQ domains rather than total score and include additional patient characteristics, has the potential to enhance the accuracy of mapping and its applicability to broader populations with obstructive airways disease.

The NOVEL observational longiTudinal studY (NOVELTY; NCT02760329) provides a unique opportunity to study a large, diverse, contemporaneous cohort of patients with physician-assigned asthma and/or COPD managed in primary and secondary care settings across 18 countries.[14] The NOVELTY study has demonstrated substantial heterogeneity within, and overlap between, populations of patients with asthma and COPD.[15] While marked heterogeneity has been observed in patients with COPD versus asthma, heterogeneity within diagnosis groups, such as for health-related-quality of life, suggests current diagnostic and severity classifications poorly differentiate between clinically-important phenotypes.[15] As such, access to this diverse cohort is of particular benefit when testing the generalizability of SGRQ to EQ-5D-5L mapping algorithms for health economic evaluations across a broad spectrum of patients with chronic airways disease.

Concordantly, the aim of this study was to validate the Starkie et al algorithm[12] in patients with asthma and/or COPD in NOVELTY, and examine whether the inclusion of additional covariates would improve the accuracy of predictions.

Materials and Methods

Study Design and Population

NOVELTY is a global, prospective observational study of patients with a physician-assigned diagnosis of asthma and/or COPD.[14] Details of the NOVELTY study design and patient population have been reported previously.[14,15] Briefly, NOVELTY included patients aged ≥18 years (or ≥12 years in some countries) from primary care and specialist centers in 18 countries in the Americas, Asia, Europe, and Oceania.

Patients completed a baseline visit and subsequent yearly visits for three years. Patient-reported measures, including the EQ-5D-5L and SGRQ, were completed by patients at baseline and at each follow-up visit. The EQ-5D-5L scale included statements across five domains and used an algorithm that converted the EQ-5D-5L questionnaire responses to EQ-5D-5L scores.[16,17] The SGRQ is a 50-item questionnaire scored on a 0–100 scale, with higher scores indicating worse health status.[9]

Baseline and Year 1 data were used as the development sample. Year 2 and Year 3 data were used as the validation samples. The intention was for all individuals to contribute to the parameter estimation to optimally use the available data.[18] To be included in the development or validation samples, a patient was required to have both EQ-5D-5L and SGRQ scores available at the yearly visit in question. Given the small fraction of missing data for physician-assessed severity, a complete case approach was used whereby patients with missing data for this variable were excluded from the development of models 3, 4, 5 and 6, and from the validation samples. This methodology is in line with Jakobsen et al 2011, which specifies that a complete case analysis may be used if the proportions of missing data are below approximately 5%.[19]

Longitudinal Linear Mixed Models

EQ-5D-5L scores were calculated using patient responses to the five questions and the R package eq5d,[20] applying the five-level crosswalk algorithm[16,17] and the UK value set.[21]

Six longitudinal linear mixed models were fitted to the development sample (baseline and Year 1), with EQ-5D-5L as the response variable. Each model had a different combination of covariates. Model 1 included the same covariates as those in the Starkie et al algorithm:[12] SGRQ total score, its squared value, and baseline sex. Model 2 included individual SGRQ domains (symptoms, activity, impact), their squared values, and sex. Model 3 included the same covariates as model 1 as well as five additional baseline covariates: age, body mass index, physician-assigned diagnosis (asthma, asthma+COPD, COPD), physician-assessed severity (mild, moderate, severe), and country (18 countries; 17 dummy-coded variables). Model 4 included the same covariates as model 2 as well as the five additional covariates included in model 3. Models 5 and 6 used backwards selection to sequentially remove the least significant baseline covariate from models 3 and 4, respectively, until all were statistically significant at the $\alpha < 0.05$ level.

Mixed model repeated measures methodology was used to enable the accommodation of within-patient correlation among measurements.[22] Restricted maximum likelihood and an unstructured covariance matrix were used to fit all models. An unstructured covariance matrix is the most general form of covariance structure.[22] The development sample included only two time points, so this structure can be easily identified. Therefore, there is no requirement to consider reduced covariance structures, such as using random intercepts and slopes.[22] Models accounted for correlations from repeated measures in the same patient over time by specifying that residual errors are correlated,[22] and ceiling effects in EQ-5D-5L scores were not accounted for. More sophisticated models that do accommodate ceiling effects are possible[23] but were not adopted here. This was because our estimation sample was longitudinal and methods of this type require extension to accommodate correlated data.[23] In addition, our primary interest lay in whether Starkie et al's linear model could be improved by including additional covariates. SGRQ scores were time varying covariates, with scores taken at either baseline or Year 1 corresponding to the visit of the EQ-5D-5L response variable.

Model performance was assessed by computing mean square error (MSE) values for predictions made in the two validation samples (Year 2 and Year 3) in the main and subgroup analyses (below); this methodological approach is supported by Brazier et al.[24] Smaller MSEs indicate better performing models. Validation sample predictions and MSEs were computed in subsets of patients who had data for all covariates included in the models, to ensure that all six models were validated in the same set of patients. In addition to MSE, models were also compared in terms of the proportion of variation in EQ-5D-5L explained.

Subgroup Analyses

Subgroup analyses were performed according to the physician-assigned diagnosis (asthma, asthma+COPD, or COPD), whereby models were developed in the subgroups and were subsequently validated. Patients with asthma+COPD were included in both the asthma and COPD subgroups.

Sensitivity Analyses

To assess the impact of missing EQ-5D-5L and SGRQ data, we performed a sensitivity analysis where these data were imputed. To assess the impact of alternative disease severity classification, we performed a sensitivity analysis where disease severity was determined based on the Global Initiative for Asthma (GINA) treatment step classification, as opposed to physician-assessed severity in the main analysis. In this analysis, the GINA 2017 treatment steps (1–5) were used as a categorical baseline covariate for severity instead of physician-assessed severity in patients with asthma and asthma+COPD.[25] These treatment steps represented the GINA guideline recommendations that applied at the time the data were collected. Patients for whom GINA treatment step was unknown were excluded from models 3 and 4, the stepwise procedure giving rise to models 5 and 6, and from the validation samples to ensure that all six models were validated using the same validation datasets.

In their original report, Starkie et al rounded coefficients of intercept, SGRQ, $SGRQ^2$, and sex to four decimal places. An additional comparison was performed in which the coefficients from the model that was subsequently selected for further analysis were rounded to the same number of decimal places,[12] to evaluate the potential loss of prediction accuracy due to rounding.

Results

Study Population

In total, 11,192 patients were included in the NOVELTY baseline population (Table 1 and Figure 1). Of these, 6813 patients with available EQ-5D-5L and SGRQ data were included in the baseline sample (mean age, standard deviation [SD] 60.3 (14.5) years; 47.7% male); 3546 had asthma, 872 had asthma+COPD, and 2395 had COPD. The Year 1 sample included 5607 patients with both EQ-5D-5L and SGRQ data (Tables 1 and 2) and together with the baseline sample comprised the development sample. The Year 2 validation sample and the Year 3 validation sample included 4593 and 4643 patients with both EQ-5D-5L and SGRQ data available, respectively (Tables 1 and 2). A summary of the EQ-5D-5L outcome data is shown

Table 1 Baseline Demographics and Clinical Characteristics for Patients Included in the Baseline, Year 1, Year 2, and Year 3 Samples

Parameter	Development Sample			Validation Samples	
	Overall NOVELTY Baseline Population (N=11,192)	Baseline Population Included in Main Analysis (N=6813)	Year 1 Population Included in Main Analysis (N=5607)	Year 2 Population Included in Main Analysis (N=4593)	Year 3 Population Included in Main Analysis (N=4643)
Age (years), mean (SD)	58.7 (15.8)	60.3 (14.5)	61.2 (13.5)	61.7 (13.0)	61.2 (13.2)
Male, n (%)	5345 (47.8)	3250 (47.7)	2747 (49.0)	2302 (50.1)	2253 (48.5)
BMI (kg/m^2), mean (SD)	28.1 (6.7)	28.1 (6.6)	28.0 (6.4)	28.0 (6.3)	27.9 (6.3)
Physician-assigned diagnosis, n (%)					
Asthma	5908 (52.8)	3546 (52.0)	2845 (50.7)	2272 (49.5)	2450 (52.8)
Asthma+COPD	1401 (12.5)	872 (12.8)	733 (13.1)	628 (13.7)	625 (13.5)
COPD	3883 (34.7)	2395 (35.2)	2029 (36.2)	1693 (36.9)	1568 (33.8)
Physician-assessed severity, n (%)					
Mild	3532 (31.6)	2116 (31.1)	1672 (29.8)	1363 (29.7)	1445 (31.1)
Moderate	3917 (35.0)	2384 (35.0)	2005 (35.8)	1638 (35.7)	1668 (35.9)
Severe	3732 (33.3)	2308 (33.9)	1927 (34.4)	1591 (34.6)	1529 (32.9)
Missing[a]	11 (0.1)	5 (0.1)	3 (0.1)	1 (0.0)	1 (0.0)
Country, n (%)					
Argentina	521 (4.7)	292 (4.3)	249 (4.4)	204 (4.4)	216 (4.7)
Australia	818 (7.3)	454 (6.7)	358 (6.4)	305 (6.6)	293 (6.3)
Brazil	202 (1.8)	140 (2.1)	84 (1.5)	59 (1.3)	53 (1.1)
Canada	1165 (10.4)	742 (10.9)	549 (9.8)	415 (9.0)	464 (10.0)
Columbia	252 (2.3)	193 (2.8)	71 (1.3)	68 (1.5)	69 (1.5)
Denmark	97 (0.9)	85 (1.2)	63 (1.1)	49 (1.1)	44 (0.9)
France	747 (6.7)	404 (5.9)	316 (5.6)	268 (5.8)	258 (5.6)
Germany	769 (6.9)	508 (7.5)	438 (7.8)	336 (7.3)	357 (7.7)
Italy	590 (5.3)	418 (6.1)	347 (6.2)	288 (6.3)	276 (5.9)
Japan	820 (7.3)	613 (9.0)	572 (10.2)	516 (11.2)	565 (12.2)
South Korea	606 (5.4)	396 (5.8)	306 (5.5)	241 (5.2)	232 (5.0)
Mexico	126 (1.1)	109 (1.6)	92 (1.6)	73 (1.6)	61 (1.3)
The Netherlands	318 (2.8)	-[b]	194 (3.5)	158 (3.4)	153 (3.3)
Norway	52 (0.5)	36 (0.5)	29 (0.5)	20 (0.4)	28 (0.6)
Spain	975 (8.7)	602 (8.8)	563 (10.0)	464 (10.1)	456 (9.8)
Sweden	335 (3.0)	261 (3.8)	219 (3.9)	183 (4.0)	211 (4.5)
Great Britain	890 (8.0)	555 (8.1)	454 (8.1)	422 (9.2)	443 (9.5)
USA	1909 (17.1)	1005 (14.8)	703 (12.5)	524 (11.4)	464 (10.0)
SGRQ total, mean (SD)	35.2 (22.1)	35.4 (22.2)	32.3 (22.8)	31.6 (23.0)	30.4 (22.5)

Note: [a]Patients with missing physician-assessed severity data were excluded from the development of models 3, 4, 5 and 6, and from the validation samples. [b]EQ-5D-5L data were not available for The Netherlands at baseline.

Abbreviations: BMI, body mass index; COPD, chronic obstructive pulmonary disease; EQ-5D-5L, EuroQoL 5 Dimensions 5 Levels Health Questionnaire; N, total number of patients; n, number of patients with data; SD, standard deviation.

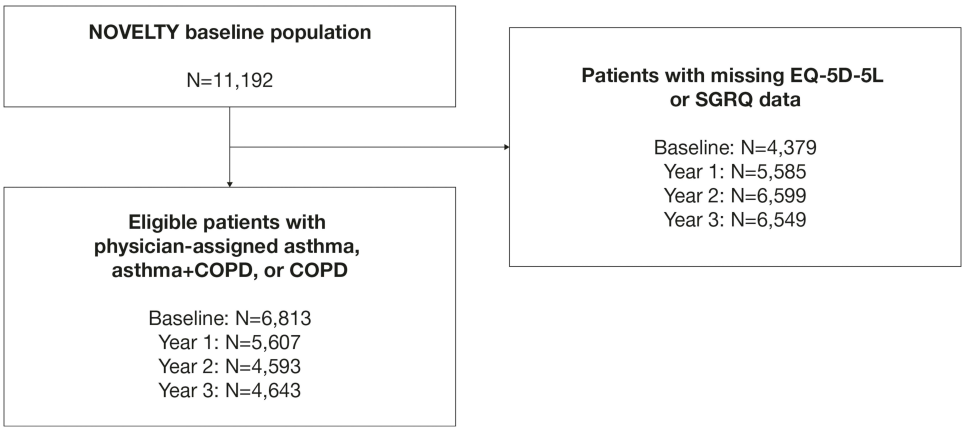

Figure 1 Number of patients included in this analysis.
Abbreviations: COPD, chronic obstructive pulmonary disease; EQ-5D-5L, EuroQoL 5 Dimensions 5 Levels Health Questionnaire; SGRQ, St George's Respiratory Questionnaire.

in Table 2; the median (interquartile range) EQ-5D-5L score in the baseline sample was 0.795 (0.648–1), at Year 1 was 0.827 (0.664–1), at Year 2 was 0.836 (0.668–1), and at Year 3 was 0.819 (0.671–1). The mean (SD) SGRQ score in the baseline sample was 35.4 (22.2), at Year 1 was 32.3 (22.8), at Year 2 was 31.6 (23.0), and at Year 3 was 30.4 (22.5).

Model Development

In models 1 and 2, sex, which was the only baseline covariate included, was statistically significant (Table 3). Backwards stepwise selection removed baseline age from models 3 and 4, resulting in models 5 and 6.

Model Validation

Validation MSE values indicated good performance and were similar across models 1–6 (Year 2: 0.0302–0.0308; Year 3: 0.0272–0.0277) and were smaller than those obtained using the Starkie et al algorithm[12] (Year 2: 0.0340; Year 3: 0.0296), indicating better performing models (Table 3). EQ-5D-5L score sample variance was 0.0556 for Year 2 and 0.0524 for Year 3, indicating that just under 50% (Year 2: 45–46%; Year 3: 47–48%) of the total variation in the EQ-5D-5L scores is explained by the models.

Since validation MSEs were similar across all models, model 1 was chosen for further analysis given that this was the most parsimonious model. Model 1 performed well with estimated coefficients for all patients differing from those of the Starkie et al algorithm[12] (Table 4). Predictions were similar for model 1 and the Starkie et al algorithm when rounding to four decimal places, which is consistent with the rounding used by Starkie et al.[12] Furthermore, predictions lay in a feasible utility range (Figure 2).

Table 2 EQ-5D-5L Outcome Data and Missing Data in the Development and Validation Samples

Visit	Sample Type	1st Quartile EQ-5D-5L Score	Median EQ-5D-5L Score	3rd Quartile EQ-5D-5L Score	Patients with EQ-5D-5L Scores Equal to 1, n	Patients with Missing EQ-5D-5L Score Data, n (%)	Patients with Missing EQ-5D-5L Score or SGRQ Data, n (%)	Eligible Patients with Physician-Assigned Asthma, Asthma+COPD or COPD, n
Baseline	Development	0.648	0.795	1	1830	4217 (37.7)	4379 (39.1)	6813
Year 1	Development	0.664	0.827	1	1792	5461 (48.8)	5585 (49.9)	5607
Year 2	Validation	0.668	0.836	1	1513	6507 (58.1)	6599 (59.0)	4593
Year 3	Validation	0.671	0.819	1	1474	6442 (57.6)	6549 (58.5)	4643

Abbreviations: COPD, chronic obstructive pulmonary disease; EQ-5D-5L, EuroQoL 5 Dimensions 5 Levels Health Questionnaire; n, number of patients with data; SGRQ, St George's Respiratory Questionnaire.

https://doi.org/10.2147/POR.S508814

Model Performance Within Disease Groups

For the asthma subgroup, 4,418, 3,578, 2,900, and 3,075 patients with asthma (and asthma+COPD) were included in the baseline, Year 1, Year 2, and Year 3 samples, respectively. For the COPD subgroup, 3,267, 2,762, 2,321, and 2,193 patients with COPD (and asthma+COPD) were included in the baseline, Year 1, Year 2, and Year 3 samples, respectively. As with the main analysis, sex was a statistically significant baseline covariate in models 1 and 2 in both the asthma and COPD subgroup analyses. Backwards stepwise selection removed different baseline covariates from models 5 and 6 in each subgroup analysis (Table 3).

Model 1 performed well in the patient subgroups, with performance similar in all models in the asthma and COPD subgroups (Table 3). As with the main analysis, estimated coefficients for model 1 predictions in patients with asthma (and asthma+COPD) and patients with COPD (and asthma+COPD) differed from those of the Starkie et al algorithm[12] (Table 4).

Estimated coefficients were different for the COPD subgroup compared with the main and asthma subgroup analyses; however, model 1 predictions were similar across the main and subgroup analyses (Figure 3), although some noticeable

Table 3 Model Performance in All Patients, Patients with Asthma (and asthma+COPD), and Patients with COPD (and asthma+COPD)

Analysis	Model	SGRQ Total Score or Domains Used	Baseline Covariates Included	MSE (Year 2)	MSE (Year 3)
Main analysis					
All patients	1	Total	Sex	0.0308	0.0273
	2	Domains	Sex	0.0308	0.0272
	3	Total	Sex, age, BMI, physician-assessed severity, physician-assigned diagnosis, country	0.0302	0.0277
	4	Domains	Sex, age, BMI, physician-assessed severity, physician-assigned diagnosis, country	0.0303	0.0275
	5	Total	Sex, BMI, physician-assessed severity, physician-assigned diagnosis, country	0.0302	0.0277
	6	Domains	Sex, BMI, physician-assessed severity, physician-assigned diagnosis, country	0.0303	0.0275
	Starkie et al[12]	Total	Sex	0.0340	0.0296
Subgroup analyses					
Asthma subgroup[a]	1	Total	Sex	0.0307	0.0257
	2	Domains	Sex	0.0308	0.0254
	3	Total	Sex, age, BMI, physician-assessed severity, physician-assigned diagnosis, country	0.0302	0.0257
	4	Domains	Sex, age, BMI, physician-assessed severity, physician-assigned diagnosis, country	0.0303	0.0254
	5	Total	Sex, BMI, physician-assigned diagnosis, country	0.0302	0.0258
	6	Domains	BMI, country	0.0303	0.0255
COPD subgroup[b]	1	Total	Sex	0.0341	0.0309
	2	Domains	Sex	0.0342	0.0307
	3	Total	Sex, age, BMI, physician-assessed severity, physician-assigned diagnosis, country	0.0334	0.0315
	4	Domains	Sex, age, BMI, physician-assessed severity, physician-assigned diagnosis, country	0.0335	0.0314
	5	Total	BMI, physician-assessed severity, country	0.0335	0.0316
	6	Domains	BMI, physician-assessed severity, country	0.0336	0.0314

Note: [a]A subset of patients with asthma (possibly comorbid with COPD). [b]A subset of patients with COPD (possibly comorbid with asthma).
Abbreviations: BMI, body mass index; COPD, chronic obstructive pulmonary disease; MSE, mean square error; SGRQ, St George's Respiratory Questionnaire.

Table 4 Estimated Coefficients for Model 1 Predictions in All Patients, Patients with Asthma (and asthma+COPD), and Patients with COPD (and asthma+COPD), Compared with Starkie et al

Covariate	Coefficient			
	Starkie et al[12]	All patients (SE)	Asthma subgroup (SE)	COPD subgroup (SE)
Intercept	0.9617	0.9434 (0.0045)	0.9486 (0.0049)	0.9259 (0.0085)
SGRQ total score/100	–0.13[a]	–0.2942 (0.0233)	–0.3178 (0.0273)	–0.2352 (0.0389)
(SGRQ total score/100)^2	–1[b]	–0.5153 (0.0281)	–0.4765 (0.0341)	–0.5909 (0.0435)
Male	0.0231	0.0105 (0.0033)	0.0102 (0.0040)	0.0160 (0.0051)

Note: Models presented using SGRQ total score/100 as a covariate to help ensure sufficient significant figures are captured. [a]The coefficient for SGRQ total score is −0.0013. [b]The coefficient for SGRQ total score squared is −0.0001.
Abbreviations: COPD, chronic obstructive pulmonary disease; SE, standard error; SGRQ, St George's Respiratory Questionnaire.

differences were observed for low (<20) and high (>60) SGRQ scores, due to the SGRQ quadratic coefficients having the most impact at the data extremities. To explore if patients with COPD require different coefficients, a final model was fitted to all patients as in model 1 but also including an indicator for the patient having COPD at baseline as a covariate, as well as interactions between this indicator and the other model coefficients. The main effect of the COPD indicator was estimated to be –0.024 and was statistically significant (p = 0.014). This results in lower predictions for patients with COPD versus other diagnostic groups (Figure 3) and suggests that use of the lower intercept for patients with COPD may be important. The strength of the evidence for interactions of the COPD indicator with other covariates varied (sex p = 0.499; SGRQ total score p = 0.037; SGRQ total score squared p = 0.005).

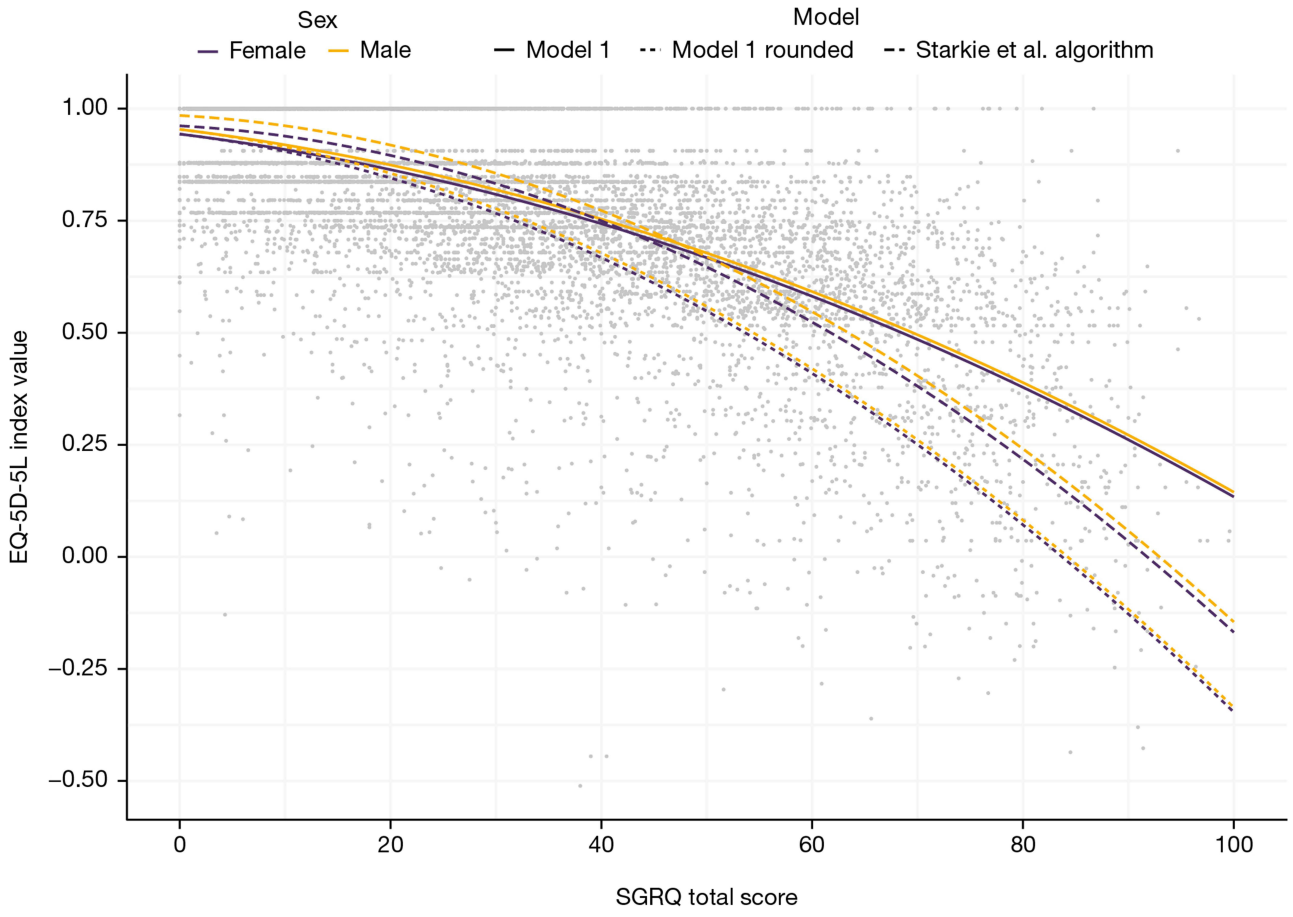

Figure 2 Comparison of predictions from model 1 with those from the Starkie et al algorithm.
Note: Starkie et al algorithm from: Starkie et al. Value Health 2011;14:354–60[12]. Grey dots depict all available data points from both validation samples (Year 2 and Year 3) to support in interpreting the predictive power of the model.
Abbreviations: EQ-5D-5L, EuroQoL 5 Dimensions 5 Levels Health Questionnaire; SGRQ, St George's Respiratory Questionnaire.

https://doi.org/10.2147/POR.S508814

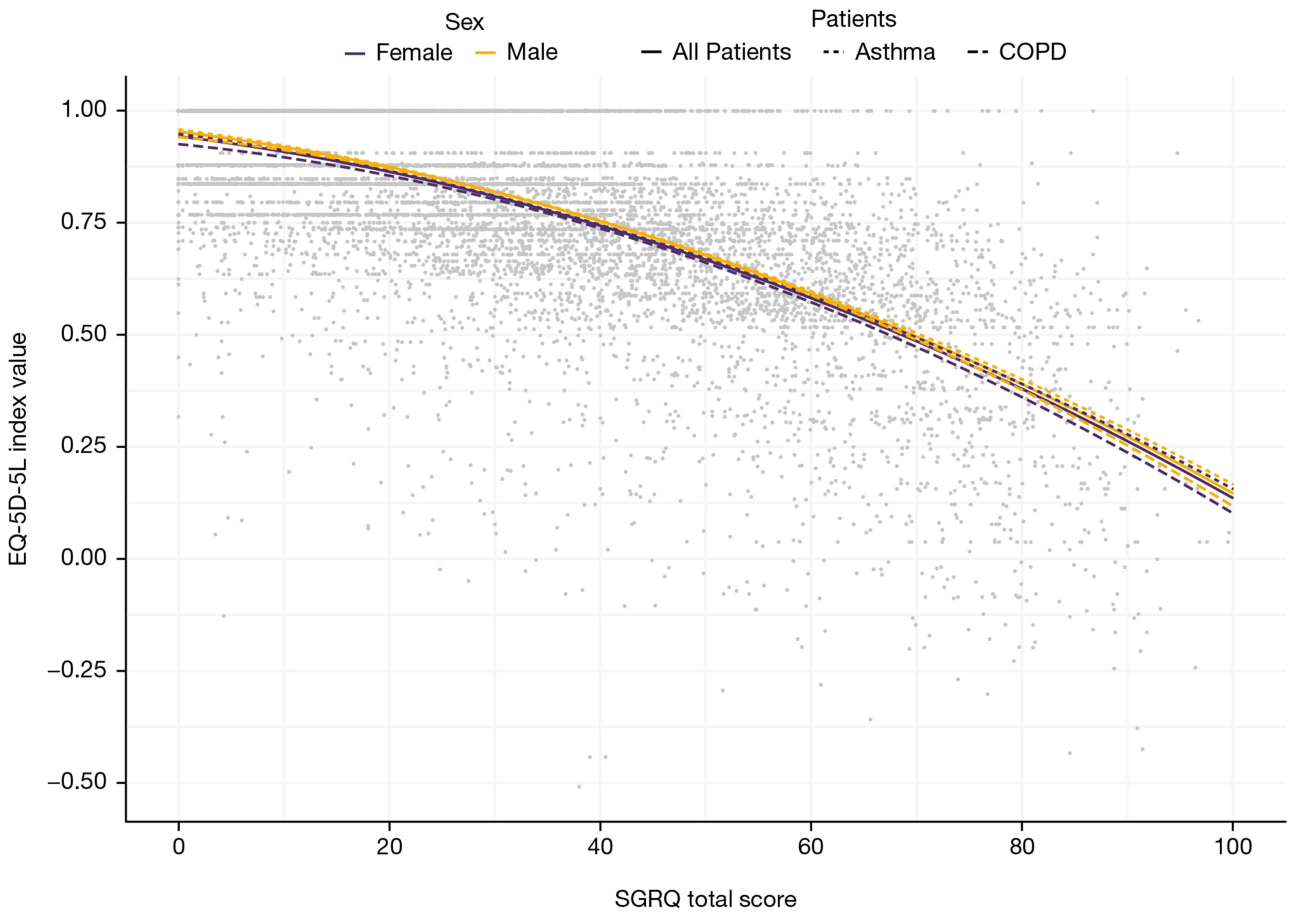

Figure 3 Comparison of subgroup analyses of model 1 predictions.
Note: Grey dots depict all available data points from both validation samples (Year 2 and Year 3) to support in interpreting the predictive power of the model.
Abbreviations: COPD, chronic obstructive pulmonary disease; EQ-5D-5L, EuroQoL 5 Dimensions 5 Levels Health Questionnaire; SGRQ, St George's Respiratory Questionnaire.

For example, using model 1 (all patients) for prediction, envisage a male patient with a SGRQ total score of 40. From the coefficients reported in Table 4, the predicted EQ-5D-5L is given by $0.9434 - 0.2942 \times \left(\frac{40}{100}\right) - 0.5153 \times \left(\frac{40}{100}\right)^2 + 0.0105 = 0.7538$ (to four decimal places). Supplementary Table 1 provides additional information about the model fits in Table 4 to allow calculation of covariances and standard errors for predictions and simulation of data from these models.

Sensitivity Analyses

While there were missing data for both EQ-5D-5L and SGRQ outcomes in the NOVELTY population (Figure 1), the sensitivity analysis that involved imputation of missing data showed that the key findings are robust when this is taken into account (see Supplementary Material — "Imputation of missing EQ-5D-5L and SGRQ values" and Supplementary Tables 2 and 3). Furthermore, the imputed EQ-5D-5L scores are not at the ceiling, suggesting that the lack of accommodation for ceiling effects in our modelling is not a paramount limitation.

In the asthma subgroup sensitivity analysis, the GINA treatment step was unknown for 494 patients with asthma and 59 patients with asthma+COPD, and these patients were excluded from models 3 and 4, the stepwise procedure giving rise to models 5 and 6, and from the validation samples as described in the sensitivity analysis section of the Methods. The conclusions from the asthma subgroup analysis were found to be robust when using GINA 2017 treatment step instead of physician-assessed severity as a covariate. Backwards stepwise selection removed the baseline covariate GINA 2017 treatment step from models 3 and 4 (Supplementary Table 4).

Discussion

In this work, we developed an algorithm to map the SGRQ to EQ-5D-5L and evaluated model performance against a previously published algorithm across a spectrum of patients with chronic airways disease. Validation MSEs indicated good performance and were similar across models. The models had reasonable predictive power, explaining approximately half of the variation in EQ-5D-5L scores. The covariates used by Starkie et al[12] (SGRQ total score, its squared value, and baseline sex) were adequate for mapping in this patient population. Including additional baseline covariates and SGRQ domains resulted in a similar model performance to the Starkie et al algorithm, although the coefficients used in our analysis may be more suitable for situations where the patient population is more similar to that included in the NOVELTY study, versus those used in the Starkie et al algorithm.[12]

Our results have important implications for research. We confirm that the Starkie et al mapping algorithm, developed based on data collected >15 years ago,[12] performs well using contemporary, longitudinal, multi-country data. The observation that the adaptations made in this analysis only slightly improved the performance of the Starkie et al algorithm,[12] suggests that it is adequate for mapping in patients with asthma and/or COPD, and since the SGRQ is widely used in clinical research in airways diseases,[4–8] the mapping algorithm can have wide applicability. The Starkie et al algorithm[12] was developed in patients with moderate-to-very severe COPD from the TORCH clinical trial, whereas our analysis included a much broader population of patients from NOVELTY with physician-assigned asthma, asthma+COPD, or COPD, with a range of physician-assessed severities, from real-world primary care and specialist settings in 18 countries. Although the findings for the subgroup analyses are slightly unclear in terms of whether or not patients with COPD require their own set of coefficients, data suggest that the algorithm derived from all patients, regardless of their diagnosis, can also be applied to patients with COPD. Conclusions were robust when using GINA 2017 treatment steps instead of physician-assessed severity as a covariate but the use of GINA treatment steps resulted in more missing covariate data versus physician-assessed severity and, as GINA treatment step was not selected in models 5 and 6, it does not appear to be a useful predictor.

A noteworthy observation is that Starkie et al's[12] decision to only report their coefficients to four decimal places, and so the coefficient of $SGRQ^2$ to the first significant figure, may have resulted in a loss of model performance. The Starkie et al[12] predictions lay in between our model, where we report all decimal places for the coefficients provided by the statistical software, and our rounded model, where we only report coefficients to four decimal places; predictions for our model were more similar to those of Starkie et al when rounding to four decimal places. This highlights that numerical precision of reporting coefficients is critical when making predictions, as SGRQ total score squared can become large. When making comparisons, one methodological advantage of our study is the use of mixed model repeated measures and using their ability to predict future patient outcomes to validate them; when making predictions in the validation samples, no use of the estimated correlation structure is made, resulting in a more equitable comparison between the Starkie et al model[12] and our model.

Some limitations of this work should be acknowledged. As NOVELTY is a longitudinal study, the development and validation samples were generated using data for the same patients across different yearly timepoints. The drawback of this approach is that the development and validation samples were not independent of each other, resulting in data correlation between the samples, making comparisons with other algorithms derived from entirely different datasets difficult. However, splitting the development and validation samples in this way can improve the model's predictive power by taking advantage of the longitudinal nature of the data.[18] Future mapping analyses should use samples independent from those reported here and by Starkie et al[12] to validate the mapping algorithms and coefficients should be reported to a greater numerical accuracy than four decimal places. Furthermore, accounting for the ceiling effect of EQ-5D-5L in the modeling might improve the accuracy of the mapping, albeit at the cost of added complexity. Additionally, although validation across 18 countries means that our algorithm is likely applicable to diverse settings, inclusion of country as a covariate with the UK tariff for mapping may also generate conceptual challenges. However, an advantage of using the most parsimonious model (model 1) is that the resulting model predictions are widely applicable as they do not depend on variables such as country that may make their application more challenging in other datasets. While our findings show that EQ-5D-5L utility scores can be predicted through mapping from SGRQ total score, as previously described by Starkie et al,[12] it should be noted that EQ-5D-5L utility scores derived directly from validated instruments should be used, where possible.

In conclusion, our analysis validates the use of Starkie-type mappings in patients with asthma and/or COPD. The inclusion of additional baseline covariates and SGRQ domains resulted in a similar model performance. Our updated coefficients

derived from NOVELTY have potential for use in health economics models due to the broader recruited population versus Starkie et al.[12] In addition to marginal improvement in the accuracy of mapping (in terms of lower MSEs), a particular advantage of NOVELTY-derived coefficients in health economic modeling is their applicability to patients with asthma and asthma+COPD. Our mapping algorithm could also be considered for use in other respiratory diseases in the future.

Data Sharing Statement

Data underlying the findings described in this manuscript may be obtained in accordance with AstraZeneca's data sharing policy described at https://astrazenecagrouptrials.pharmacm.com/ST/Submission/Disclosure. Data for studies directly listed on Vivli can be requested through Vivli at www.vivli.org. Data for studies not listed on Vivli could be requested through Vivli at https://vivli.org/members/enquiries-about-studies-not-listed-on-the-vivli-platform/. AstraZeneca Vivli member page is also available outlining further details: https://vivli.org/ourmember/astrazeneca/. The NOVELTY protocol is available at https://astrazenecagrouptrials.pharmacm.com.

Ethics Approval

The NOVELTY study was performed in accordance with ethical principles consistent with the Declaration of Helsinki, International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, and Good Clinical Practice. The NOVELTY protocol was approved in each participating country by the relevant independent ethics committees and institutional review boards, who are listed in full in Supplementary Table 7, and all patients provided written informed consent (with legal guardian consent for adolescent patients).

Acknowledgments

The authors would like to thank the patients who participated in this study and the NOVELTY Scientific Community and the NOVELTY study investigators who are listed in full in Supplementary Tables 5 and 6. Medical writing support, under the direction of the authors, was provided by Niall Tyrer, MBiolSci, and Jilly Hope, PhD, CMC Connect, a division of IPG Health Medical Communications, funded by AstraZeneca, Cambridge, UK, in accordance with Good Publication Practice (GPP 2022) guidelines (Ann Intern Med. 2022;175(9):1298-1304).

The abstract of this paper was presented at the ISPOR EU 2023 congress as a poster presentation with interim findings. The poster's abstract was published the Supplement of Value in Health: https://doi.org/10.1016/j.jval.2023.09.2528.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

The NOVELTY study is funded by AstraZeneca.

Disclosure

D Jackson, A Quinton, H Müllerová and F Zhang are employees and shareholders of AstraZeneca. C Janson has received honoraria from AstraZeneca, Boehringer Ingelheim, Chiesi, GlaxoSmithKline, Novartis, and Orion for lectures. M Sadatsafavi has received honoraria from AstraZeneca for participation in the NOVELTY study. The authors report no other conflicts of interest in this work.

References

1. Brazier J, Ara R, Rowen D, Chevrou-Severac H. A review of generic preference-based measures for use in cost-effectiveness models. *Pharmacoeconomics*. 2017;35(Suppl 1):21–31. doi:10.1007/s40273-017-0545-x

2. EUROQOL. EQ-5D-5L user guide; 2019. Available from: https://euroqol.org/wp-content/uploads/2023/11/EQ-5D-5LUserguide-23-07.pdf. Accessed August 14, 2024.
3. Weldring T, Smith SM. Patient-reported outcomes (PROs) and patient-reported outcome measures (PROMs). *Health Serv Insights*. 2013;6:61–68. doi:10.4137/HSI.S11093
4. Afroz N, Gutzwiller FS, Mackay AJ, Naujoks C, Patalano F, Kostikas K. Patient-reported outcomes (PROs) in COPD clinical trials: trends and gaps. *Int J Chron Obstruct Pulmon Dis*. 2020;15:1789–1800. doi:10.2147/COPD.S235845
5. Mahler DA, Kerwin E, Ayers T, et al. FLIGHT1 and FLIGHT2: efficacy and safety of QVA149 (indacaterol/glycopyrrolate) versus its monocomponents and placebo in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2015;192(9):1068–1079. doi:10.1164/rccm.201505-1048OC
6. Zhong N, Wang C, Zhou X, et al. LANTERN: a randomized study of QVA149 versus salmeterol/fluticasone combination in patients with COPD. *Int J Chron Obstruct Pulmon Dis*. 2015;10:1015–1026. doi:10.2147/COPD.S84436
7. Hashimoto S, Ikeuchi H, Murata S, Kitawaki T, Ikeda K, Banerji D. Efficacy and safety of indacaterol/glycopyrronium in Japanese patients with COPD: a subgroup analysis from the SHINE study. *Int J Chron Obstruct Pulmon Dis*. 2016;11:2543–2551. doi:10.2147/COPD.S111408
8. Arpinelli F, Carone M, Riccardo G, Bertolotti G. Health-related quality of life measurement in asthma and chronic obstructive pulmonary disease: review of the 2009-2014 literature. *Multidiscip Respir Med*. 2015;11:5. doi:10.1186/s40248-016-0040-9
9. Jones PW, Quirk FH, Baveystock CM, Littlejohns P. A self-complete measure of health status for chronic airflow limitation. The St. George's respiratory questionnaire. *Am Rev Respir Dis*. 1992;145(6):1321–1327. doi:10.1164/ajrccm/145.6.1321
10. US Food and Drug Administration. Chronic obstructive pulmonary disease: use of the St. George's respiratory questionnaire as a PRO assessment tool. *Guidance for industry*; 2018. Available from: https://www.fda.gov/files/drugs/published/Chronic-Obstructive-Pulmonary-Disease–Use-of-the-St.-George%E2%80%99s-Respiratory-Questionnaire-as-a-PRO-Assessment-Tool-Guidance-for-Industry.pdf. Accessed April 22, 2022.
11. Dakin H, Abel L, Burns R, Yang Y. Review and critical appraisal of studies mapping from quality of life or clinical measures to EQ-5D: an online database and application of the MAPS statement. *Health Qual Life Outcomes*. 2018;16(1):31. doi:10.1186/s12955-018-0857-3
12. Starkie HJ, Briggs AH, Chambers MG, Jones P. Predicting EQ-5D values using the SGRQ. *Value Health*. 2011;14(2):354–360. doi:10.1016/j.jval.2010.09.011
13. Whittington MD, McQueen RB, Ollendorf DA, et al. Assessing the value of mepolizumab for severe eosinophilic asthma: a cost-effectiveness analysis. *Ann Allergy Asthma Immunol*. 2017;118(2):220–225. doi:10.1016/j.anai.2016.10.028
14. Reddel HK, Gerhardsson de Verdier M, Agustí A, et al. Prospective observational study in patients with obstructive lung disease: NOVELTY design. *ERJ Open Res*. 2019;5(1):00036–2018. doi:10.1183/23120541.00036-2018
15. Reddel HK, Vestbo J, Agustí A, et al. Heterogeneity within and between physician-diagnosed asthma and/or COPD: NOVELTY cohort. *Eur Respir J*. 2021;58(3):2003927. doi:10.1183/13993003.03927-2020
16. van Hout B, Janssen MF, Feng YS, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. *Value Health*. 2012;15(5):708–715. doi:10.1016/j.jval.2012.02.008
17. Omelyanovskiy V, Musina N, Ratushnyak S, et al. Valuation of the EQ-5D-3L in Russia. *Qual Life Res*. 2021;30(7):1997–2007. doi:10.1007/s11136-021-02804-6
18. Collins GS, Dhiman P, Ma J, et al. Evaluation of clinical prediction models (part 1): from development to external validation. *BMJ*. 2024;384:e074819.
19. Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials - a practical guide with flowcharts. *BMC Med Res Methodol*. 2017;17(1):162. doi:10.1186/s12874-017-0442-1
20. Package 'eq5d'. EQ-5D: methods for analysing 'EQ-5D' data and calculating 'EQ-5D' index scores; 2023. Available from: https://cran.r-project.org/web/packages/eq5d/eq5d.pdf. Accessed February 6, 2024.
21. Devlin NJ, Shah KK, Feng Y, Mulhern B, van Hout B. Valuing health-related quality of life: an EQ-5D-5L value set for England. *Health Econ*. 2018;27(1):7–22. doi:10.1002/hec.3564
22. The Stats Geek. Mixed model repeated measures (MMRM) in Stata, SAS and R; 2020. Available from: https://thestatsgeek.com/2020/12/30/mixed-model-repeated-measures-mmrm-in-stata-sas-and-r/. Accessed February 6, 2024.
23. Alava MH, Wailoo A. Fitting adjusted limited dependent variable mixture models to EQ-5D. *Stata J*. 2015;15(3):737–750. doi:10.1177/1536867X1501500307
24. Brazier JE, Yang Y, Tsuchiya A, Rowen DL. A review of studies mapping (or cross walking) non-preference based measures of health to generic preference-based measures. *Eur J Health Econ*. 2010;11(2):215–225. doi:10.1007/s10198-009-0168-z
25. Global Initiative for Asthma. Global strategy for asthma management and prevention: 2017 report. Available from: https://ginasthma.org/wp-content/uploads/2019/01/2017-GINA.pdf. Accessed February 6, 2024.