PharmacoEconomics - Open (2023) 7:149–161
https://doi.org/10.1007/s41669-023-00388-0

ORIGINAL RESEARCH ARTICLE

Mapping the Insomnia Severity Index Instrument to EQ-5D Health State Utilities: A United Kingdom Perspective

François-Xavier Chalet[1] · Teodora Bujaroska[2] · Evi Germeni[3] · Nizar Ghandri[2] · Emilio T. Maddalena[2] · Kushal Modi[4] · Abisola Olopoenia[4] · Jeffrey Thompson[4] · Matteo Togninalli[2] · Andrew H. Briggs[5,6]

Accepted: 12 January 2023 / Published online: 26 January 2023
© The Author(s) 2023, corrected publication 2023

Abstract

Objective This study aimed to map the Insomnia Severity Index (ISI) to the EQ-5D-3L utility values from a UK perspective.

Methods Source data were derived from the 2020 National Health and Wellness Survey (NHWS) for France, Germany, Italy, Spain, the UK and the US. Ordinary least squares regression, generalised linear model (GLM), censored least absolute deviation, and adjusted limited dependent variable mixture model (ALDVMM) were employed to explore the relationship between ISI total summary score and EQ-5D utility while accounting for adjustment covariates derived from the NHWS. Fitting performance was assessed using standard metrics, including mean-squared error (MSE) and coefficient of determination (R^2).

Results A total of 17,955 respondent observations were included, with a mean ISI score of 12.12 ± 5.32 and a mean EQ-5D-3L utility (UK tariff) of 0.71 ± 0.23. GLM gamma-log and ALDVMM were the two best performing models. The ALDVMM had better fitting performance ($R^2 = 0.320$, MSE 0.0347) than the GLM gamma-log ($R^2 = 0.303$, MSE 0.0353); in train-test split-sample validation, ALDVMM also slightly outperformed the GLM gamma-log model, with an MSE of 0.0351 versus 0.0355. Based on fitting performance, ALDVMM and GLM gamma-log were the preferred models.

Conclusions In the absence of preference-based measures, this study provides an updated mapping algorithm for estimating EQ-5D-3L utilities from the ISI summary total score. This new mapping not only draws its strengths from the use of a large international dataset but also the incorporation of adjustment variables (including sociodemographic and general health characteristics) to reduce the effects of confounders.

Key Points for Decision Makers

This study provides an algorithm to map the condition-specific Insomnia Severity Index (ISI) to the EQ-5D-3L utilities, a general health-related quality-of-life questionnaire commonly used in health technology assessment.

The mapping algorithm builds upon previous work, expanding it through the use of a large, multinational dataset and adjustment for respondent characteristics.

1 Introduction

According to the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) [1], insomnia disorder is a sleep disturbance (difficulty falling sleep, staying asleep

✉ François-Xavier Chalet
francois-xavier.chalet@idorsia.com

1 Idorsia Pharmaceuticals Ltd, Hegenheimermattweg 91, 4123 Allschwil, Switzerland

2 Visium, EPFL Innovation Park, Rte Cantonale, 1015 Lausanne, Switzerland

3 Health Economics and Health Technology Assessment (HEHTA), School of Health and Wellbeing, University of Glasgow, 1 Lilybank Gardens, Glasgow G12 8RZ, UK

4 Cerner Enviza, 2800 Rockcreek Parkway, North Kansas City, MO 64117, USA

5 Department of Health Services Research & Policy, London School of Hygiene and Tropical Medicine, 15-17 Tavistock Place, London WC1H 9SH, UK

6 Avalon Health Economics LLC, 119 Washington St, Morristown, NJ 07960, USA

or early morning awakening) marked by predominant dissatisfaction with sleep quantity or quality that is associated with substantial distress and impairments of daytime functioning. The sleep difficulty must occur at least 3 nights per week and be present for at least 3 months, while occurring despite adequate opportunity for sleep [1]. The prevalence of insomnia varies widely depending on the definition used; it ranges from an estimated 6 to 10% (Diagnostic and Statistical Manual of Mental Disorders, 4th Edition [DSM-IV] criterion) to approximately one-third of the general population (at least one symptom of insomnia) [2]. Insomnia disorder induces fatigue and mood disturbance, and impairs social, vocational, educational, and behavioural functioning [3]. In a bidirectional relation, insomnia increases the risk of psychiatric and medical comorbidities (including anxiety and depression, obesity or weight gain, obstructive sleep apnoea and hypopnea syndrome) [4, 5]. Insomnia impacts next-day functioning, health, and quality of life (QoL) and results in substantial humanistic and economic burden for the healthcare system and society [6, 7].

Several insomnia-related generic and disease-specific patient-reported outcome (PRO) instruments have been used to identify and describe the condition. These instruments include, but are not limited to, the 36-Item Short-Form Health Survey (SF-36), the Leeds Sleep Evaluation Questionnaire, the Medical Outcomes Study Sleep Scale 12, the Epworth Sleepiness Scale, the Functional Outcomes of Sleep Questionnaire, the Pittsburgh Sleep Quality Index, the Insomnia Daytime Symptoms and Impacts Questionnaire (IDSIQ)[1] and the Insomnia Severity Index (ISI). The ISI is one of the most commonly used disease-specific measures for self-perceived insomnia severity [8]. It comprises seven items assessing the severity of sleep onset and sleep maintenance difficulties (both nocturnal and early morning awakenings), satisfaction with current sleep pattern, interference with daily functioning, worry about sleep, and sleep dissatisfaction [9]. Each item is rated on a 0–4 scale and the total score ranges from 0 to 28 (with higher score denoting higher severity). The recall period is 2 or 4 weeks. To date, the ISI has been used as a PRO tool to diagnose insomnia cases, document insomnia prevalence and burden, assist clinicians in their initial evaluation of patients, determine the need for treatment, and evaluate treatment response.

The EuroQol EQ-5D and other preference-based measures are rarely reported in insomnia-related clinical trials, partially due to the lack of sensitivity of the EQ-5D for disease-specific health complaints such as fatigue and cognitive problems [10]. Therefore, as recommended in the latest guidance by the National Institute for Health and Care Excellence (NICE) [11], sleep-specific instruments need to be mapped to QoL health state utility scores to enable cost-utility analyses.

Mapping of insomnia-specific PROs to the EQ-5D has been detailed in prior literature. In August 2021 [12], NICE released a guideline that included a comparison of the different types of continuous positive airway pressure machines. The treatment effect of continuous positive airway pressure machines for obstructive sleep apnoea was measured on the Epworth Sleepiness Scale and mapped to the EQ-5D using an ordinary least squares (OLS) regression [13]. Regarding the ISI, in a 2011 study focused on insomnia, Gu et al. employed a generalised linear model (GLM) to map the ISI's seven items (Model I), summary score (Model II), and clinical categories (Model III), onto the EQ-5D-3L [14]. However, the mapping function of Gu et al. is limited by the use of a United States (US) dataset, the use of a single type of model (a GLM), and the lack of adjustment for confounding factors (e.g. comorbidities).

The primary objective of this study was to perform a new mapping between ISI and EQ-5D using a large representative dataset that includes both European (France, Germany, Italy, Spain and the UK) and US data.

2 Methods

This mapping study has been conducted following the recommendations of the Professional Society for Health Economics and Outcomes Research (ISPOR) Task Force report on 'Mapping to estimate health-state utility from non-preference-based outcome measures' [15].

2.1 Data Source

Observational studies with representative patient groups, large sample sizes, and reporting of potential confounding factors (e.g. age) are the preferred mapping datasets [15]. The mapping dataset used in this study came from the 2020 National Health and Wellness Survey (NHWS) for France, Germany, Italy, Spain, the UK and the US. The NHWS is a cross-sectional, self-administered, nationwide, internet-based survey of adults (aged ≥ 18 years) that is fielded annually in select global markets. Respondents were recruited through a general purpose, web-based consumer panel via channels such as opt-in e-mails, co-registration with panel partners, and e-newsletter campaigns. A stratified random sampling procedure, with strata by sex, race/ethnicity (in the US) and age, was implemented to obtain a representative sample of the adult population in each country. The protocol

[1] The IDSIQ is the only PRO instrument assessing the impact of insomnia and its treatment on daytime impairment. This 18-item PRO tool has been developed and validated by Idorsia Pharmaceuticals Ltd with the aim to derive valid and reliable endpoints for insomnia clinical research trials and real-world studies.

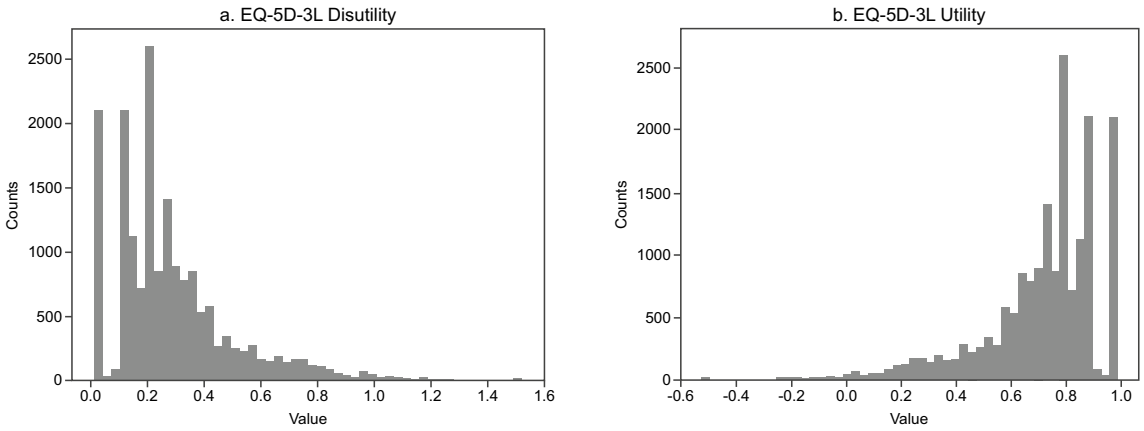


Fig. 1 EQ-5D-3L index **a** disutility and **b** utility distributions for the National Health and Wellness Survey dataset (N = 17,955). *EQ-5D-3L* EuroQol 5 dimensions 3 levels

and questionnaire for the NHWS were reviewed and granted exemption status by Pearl Institutional Review Board (IRB; Indianapolis, IN, USA; 19-KANT-204).

2.2 Study Population

Respondents were included in the modelling sample if they had reported experiencing insomnia in the past 12 months and had completed the ISI measure. Respondents excluded from the study population were those (1) reporting that they had experienced or been diagnosed with narcolepsy, sleep apnoea or sleep difficulties other than insomnia in the past 12 months; (2) reporting that they had experienced or been diagnosed with another serious condition (any type of cancer, chronic liver disease, cirrhosis, epilepsy, multiple sclerosis, muscular dystrophy or Parkinson's disease); or (3) who were pregnant at the time of data collection.

2.3 Data Inputs

NHWS data used in the mapping included sociodemographic and general health characteristics, comorbidity burden, insomnia-related measures (including ISI), current insomnia treatment, and the EQ-5D-5L health states. The respondent characteristics used as covariates for model fitting included age, sex, marital status, employment status, education level, smoking status, drinking status, comorbidities (both individual conditions and the aggregate Charlson Comorbidity Index [CCI] score [16]), body mass index (BMI), insomnia diagnosis status, treatment for insomnia status, and geography. To enable modelling of all relevant country-specific observations using a similar tariff, and to align with NICE guidelines, a UK perspective was adopted and all EQ-5D-5L health state data were converted to EQ-5D-3L utility scores using the crosswalk function of Hernández-Alava and Pudney [17].

2.4 Modelling Approach

When mapping, it is recommended that specific aspects of the EQ-5D distribution should be accounted for, i.e. the presence of large spikes, upper and lower limits, skewness, multimodality, and gaps in the range of feasible values [15]. Figure 1 illustrates the distribution of EQ-5D-3L utility data from the NHWS for the insomnia cohort. The left skew of the data can easily be converted to right skew by transforming from utility to disutility using the simple linear transformation: disutility = 1 − utility (Fig. 1).

We selected and tested multiple potentially appropriate models. First, we mapped EQ-5D-3L and ISI using OLS regression as a reference model. As well as being the most frequently used mapping model [18], OLS performs well in mean prediction [19]; however, some commentators have suggested that OLS results in systematic bias when data are not continuously distributed, as is the case for EQ-5D utility data [15]. The censored least absolute deviation (CLAD) was also tested to tackle the heteroscedasticity, non-normality, and the ceiling of the EQ-5D at 1. CLAD is a form of median regression and it is expected to perform well on the mean absolute error (MAE) metric. The question is whether or not it minimises mean square prediction error [19]. The ISI mapping work performed by Gu et al. [14] was also replicated; a gamma-log GLM was used in this analysis. To fit this model, the EQ-5D-3L was transformed into disutility as 1 − utility, so that the natural left skew of the data became a right skew (Fig. 1). A gamma-log GLM is considered an appropriate model choice because it accommodates for the skewness of EQ-5D data and prevents predictions outside of the data range. Finally, the adjusted limited dependent

variable mixture model (ALDVMM) was used to take into account both the limited nature and the multimodality trait of the EQ-5D [20]. Regarding the ALDVMMs , one should not rely on single point-estimates when handling them, but rather re-initialise the process multiple times and experiment with different optimisation methods [21]. For this reason, several methods (Broyden–Fletcher–Goldfarb–Shanno, conjugate gradient, Nelder–Mead, nlminb, Rcgmin, and Rvmmin) were employed, along with several initialisation options (zero, constant, and sann); only the best results were kept. To investigate the appropriate number of mixture models (i.e. components) for our dataset, we fitted four different ALDVMMs ranging from two to five components. Since the ALDVMM with five components did not converge, only the two to four component models are described in this study.

2.5 Covariate of Interest

ISI was the covariate of interest. We used direct mapping models by regressing the ISI total score as a continuous variable varying from 0 to 28 onto EQ-5D-3L utility.

2.6 Adjustment Covariates

Candidate covariates were derived from the information gathered in the NHWS. All models were adjusted on the entire list of available covariates that could potentially act as confounding factors. Dummy variable coding was used for all binary and categorical variables; continuous variables (including ISI) were standardised.

Binary variables encoded the respondent's sex (male, female); regular experience of pain (yes, no); having obtained an undergraduate degree or higher (yes, no); marital status (married/living with partner, single/never married/divorced/separated/widowed); current employment (employed [full time/part time/self-employed], not employed) and retirement status (retired, not retired); smoking status (current smoker, former smoker, never smoked); drinking habits (heavy drinker: 4+ times per week, low/moderate, abstains); DASD (a variable denoting self-reported depression, anxiety or post-traumatic stress disorder [yes, no]); and country (UK nationality [yes, no]). Age, BMI and the CCI were encoded as continuous variables.

With regard to insomnia-related variables, self-reported clinician-diagnosed insomnia (yes, no) and self-reported prescribed treatment for insomnia (yes, no) were included as binary variables. Potential interactions with ISI were tested for both the self-reported diagnosis and self-reported prescribed treatment variables since these potentially relate to the severity of insomnia. Ultimately, only the Treated × ISI interaction variable was retained as there was no evidence of interaction with self-reported diagnosis.

2.7 Statistical Analysis

First, the models were fit to the full dataset, and, in the case of the ALDVMM, the optimum number of components were assessed. We then explored the predictive validity of the models by splitting the entire dataset randomly 50/50 into training and validation datasets. Models were then fit on the training dataset and the predictive ability of the models was assessed on the validation dataset. Continuous variable standardisation was performed using mean and standard deviation (SD) from the training dataset. Due to the time taken to fit the ALDVMMs , only the best fitting model from the different number of components was taken forward to this train and testing stage. For the best fitting model, the same process of optimising the model was followed as described above. This process was repeated 100 times and the results presented as averages across those 100 repetitions.

Metrics employed to measure the fitting performance were the log-likelihood, Akaike's information criterion (AIC) and the Bayesian information criterion (BIC), which provide means for comparisons of specifications within model types [15]. The AIC and BIC provide measures of model performance that account for model complexity by penalising the model for the number of parameters included. In addition, MAE, mean-squared error (MSE), and the coefficient of determination (R^2) were computed to measure the predictive performance. They are mainly used to evaluate how far observed values differ from the average of predicted values. Graphical representations of model performance depicting EQ-5D observed versus predicted values were created for each mapping model.

The predicted EQ-5D utilities from the best fitting models developed as part of the current study were compared with those obtained in a mapping algorithm between the ISI and the EQ-5D, previously published by Gu et al. [14].

The statistical analyses and modelling procedures were performed using publicly available libraries in R (aldvmm [20]), and SAS 9.4 (SAS Institute Inc., Cary, NC, USA).

3 Results

3.1 Sample Characteristics

A total of 17,955 respondent observations were included in the overall model (Table 1). Most respondents came from the US ($n = 8920$), followed by the UK ($n = 2128$), France ($n = 2069$), Italy ($n = 1853$), Germany ($n = 1792$) and Spain ($n = 1193$). Among the overall sample, the mean (SD) age was 46 (16.3) years, the majority of respondents were female (66.8%), married/living with partner (54.9%), had less than a 4-year university degree (54.9%), had some

Table 1 Sample characteristics of the NHWS 2020 for the whole sample, UK, US, France, Germany, Italy and Spain

Sample characteristics		All [N = 17,955]		UK [n = 2128]		US [n = 8920]		France [n = 2069]		Germany [n = 1792]		Italy [n = 1853]		Spain [n = 1193]	
		N/mean	%/SD	*N*/mean	%/SD	*N*/mean	%/SD	*N*/mean	%/SD	*N*/mean	%/SD	*N*/mean	%/SD	*N*/mean	%/SD
Sociodemographic															
Age	Mean, SD	46.03	16.25	47.56	16.22	45.44	16.49	48.10	16.55	44.73	16.41	47.89	15.33	43.22	14.07
Sex	Male	5963	33.2	712	33.5	2776	31.1	694	33.5	682	38.1	646	34.9	453	38.0
	Female	11,992	66.8	1416	66.5	6144	68.9	1375	66.5	1110	61.9	1207	65.1	740	62.0
Married	Single/never married/divorced/separated/widowed	8105	45.1	978	46.0	4393	49.3	838	40.5	901	50.3	589	31.8	406	34.0
	Married/living with partner	9850	54.9	1150	54.0	4527	50.8	1231	59.5	891	49.7	1264	68.2	787	66.0
Education	Less than 4-year university degree	9860	54.9	967	45.4	4788	53.7	1085	52.4	1352	75.5	1188	64.1	480	40.2
	4-year university degree or higher	8095	45.1	1161	54.6	4132	46.3	984	47.6	440	24.6	665	35.9	713	59.8
Employment status	Employed full time/employed part time/self-employed	10,534	58.7	1202	56.5	5301	59.4	1114	53.8	1109	61.9	1031	55.6	777	65.1
	Retired	3142	17.5	417	19.6	1392	15.6	563	27.2	346	19.3	315	17.0	109	9.1
	Not employed	4279	23.8	509	23.9	2227	25.0	392	19.0	337	18.8	507	27.4	307	25.7
General health															
Smoking status	Current smoker	4377	24.4	478	22.5	1639	18.4	586	28.3	625	34.9	612	33.0	437	36.6
	Former smoker	4782	26.6	642	30.2	2259	25.3	672	32.5	382	21.3	491	26.5	336	28.2
	Never smoked	8796	49.0	1008	47.4	5022	56.3	811	39.2	785	43.8	750	40.5	420	35.2
Alcohol use	Heavy drinker: 4+ times per week	2273	12.7	313	14.7	1091	12.2	278	13.4	158	8.8	296	16.0	137	11.5
	Low/moderate	10,989	61.2	1393	65.5	5216	58.5	1271	61.4	1197	66.8	1113	60.1	799	67.0
	Abstains	4693	26.1	422	19.8	2613	29.3	520	25.1	437	24.4	444	24.0	257	21.5
BMI	Mean, SD	26.67	6.51	26.85	6.46	27.73	7.05	24.88	5.20	26.56	6.39	24.57	4.86	25.11	5.02
CCI	Mean, SD	0.30	0.72	0.28	0.66	0.33	0.79	0.21	0.59	0.38	0.75	0.25	0.65	0.26	0.59
DASD	No	10,575	58.9	1055	49.6	4966	55.7	1369	66.2	1201	67.0	1275	68.8	709	59.4
	Yes	7380	41.1	1073	50.4	3954	44.3	700	33.8	591	33.0	578	31.2	484	40.6
Pain	No	10,449	58.2	1277	60.0	5295	59.4	1171	56.6	924	51.6	1107	59.7	675	56.6
	Yes	7506	41.8	851	40.0	3625	40.6	898	43.4	868	48.4	746	40.3	518	43.4
Insomnia diagnosed	No	13,097	72.9	1652	77.6	6475	72.6	1382	66.8	1260	70.3	1454	78.5	874	73.3
	Yes	4858	27.1	476	22.4	2445	27.4	687	33.2	532	29.7	399	21.5	319	26.7
Insomnia treatment	Treated	2432	13.5	182	8.6	1219	13.7	358	17.3	210	11.7	251	13.6	212	17.8
	Untreated	15,523	86.5	1946	91.5	7701	86.3	1711	82.7	1582	88.3	1602	86.5	981	82.2
ISI															
ISI score	Mean, SD	12.12	5.32	12.93	5.37	11.92	5.50	12.60	5.12	13.21	5.15	10.73	4.35	11.86	5.15
ISI1	Mean, SD	1.79	1.11	1.84	1.11	1.79	1.11	1.81	1.16	1.92	1.12	1.56	0.95	1.82	1.15
ISI2	Mean, SD	1.76	1.07	1.89	1.06	1.73	1.07	1.81	1.06	2.03	1.12	1.50	0.93	1.69	1.13
ISI3	Mean, SD	1.58	1.19	1.72	1.19	1.46	1.18	1.79	1.22	1.63	1.23	1.64	1.06	1.72	1.26
ISI4	Mean, SD	2.48	0.94	2.72	0.92	2.48	0.97	2.44	0.90	2.54	0.88	2.33	0.85	2.26	0.92

Table 1 (continued)

Sample characteristics		All [N = 17,955]		UK [n = 2128]		US [n = 8920]		France [n = 2069]		Germany [n = 1792]		Italy [n = 1853]		Spain [n = 1193]	
		N/mean	%/SD	N/mean	%/SD	N/mean	%/SD	N/mean	%/SD	N/mean	%/SD	N/mean	%/SD	N/mean	%/SD
ISI5	Mean, SD	1.71	1.03	1.79	1.03	1.71	1.07	1.77	0.98	1.89	0.98	1.47	0.92	1.61	0.97
ISI6	Mean, SD	1.29	1.09	1.37	1.10	1.27	1.13	1.43	1.04	1.45	1.05	1.04	0.94	1.13	1.02
ISI7	Mean, SD	1.51	1.08	1.59	1.07	1.48	1.11	1.55	1.04	1.75	1.08	1.20	0.93	1.62	1.02
Quality-of-life measures															
EQ-5D-3L UK tariff	Mean, SD	0.71	0.23	0.68	0.26	0.71	0.23	0.72	0.21	0.69	0.23	0.75	0.17	0.78	0.18
MCS	Mean, SD	42.14	11.30	40.88	11.45	42.52	12.45	41.60	9.87	42.39	10.09	41.41	9.00	43.31	8.62
PCS	Mean, SD	49.94	9.33	49.38	10.41	50.10	9.73	50.32	8.39	48.00	9.33	50.40	7.59	51.28	7.59
SF-6D v2	Mean, SD	0.66	0.12	0.66	0.13	0.67	0.13	0.66	0.11	0.65	0.11	0.65	0.09	0.67	0.10

Data are expressed as n and percentage, except for continuous variables, which are presented as mean and SD, as specified in each relevant row

BMI body mass index, *CCI* Charlson Comorbidity Index, *DASD* depression or anxiety or post-traumatic stress disorder, *EQ-5D-3L* EuroQol 5 dimensions 3 levels, *ISI* Insomnia Severity Index, *MCS* mental component summary, *NHWS* National Health and Wellness Survey, *PCS* physical component summary, *SF-6D v2* Short-Form 6 Dimensions version 2, *SD* standard deviation, *UK* United Kingdom, *US* United States

type of employment (58.7%), never smoked (49.0%), were low/moderate drinkers (61.2%), did not have any diagnosis of pain (58.2%) or depression/anxiety/post-traumatic stress disorder (58.9%), and had a mean (SD) BMI of 26.7 (6.5) and a mean (SD) CCI of 0.30 (0.72). The sex imbalance is corroborated by epidemiologic literature on insomnia [2]. Most respondents were experiencing insomnia but had not been diagnosed by a clinician (72.9%) and currently received no treatment for their insomnia (86.5%). Mean (SD) EQ-5D-3L scores for the overall sample, France, Germany, Italy, Spain, UK, and the US were 0.71 (0.23), 0.72 (0.21), 0.69 (0.23), 0.75 (0.17), 0.78 (0.18), 0.68 (0.26), and 0.71 (0.23), respectively. EQ-5D-3L disutility and utility (Fig. 1) distributions across the full sample combined were positively skewed and negatively skewed, respectively, which is consistent with expectations.

3.2 Model Performance

Table 2 shows the analysis aiming to optimise the number of components to select the appropriate ALDVMM. The ALDVMM with three components (hereafter, ALDVMM3) reported the best performance metrics ($R^2 = 0.320$ and MSE 0.0347). Table 3 presents the performance of the models when estimating EQ-5D utilities when fitting the entire dataset. Performance of the fits to observed data provided by each model can also be assessed visually from Fig. 2, although it is important to note that this relationship between ISI and EQ-5D-3L utility is not adjusted for covariate patterns that differ across the range of ISI. The ALDVMM3 was the best performing model on the R^2 (0.31982) and the MSE (0.034691), while the GLM gamma-log was the second-best performing model ($R^2 = 0.30309$ and MSE 0.03534).

Following the predictive modelling approach, splitting the data 50/50 into training and testing sets, Table 4 indicates the performance of the models on the test dataset. The ALDVMM3 was the best performing model based on the MSE (0.0351), and the gamma-log GLM was the second-best performing model (MSE 0.0355). Working as a median regression model, the CLAD reported the best MAE score in both the fitting and predictive approaches but performed poorly when assessed using the MSE criterion.

Model coefficients for the gamma-log GLM and ALDVMM3 are presented in Table 5 alongside the mean covariate values for the NHWS data, with the non-UK and treated variables set to zero, such that the algorithm predicts values for a UK non-treated population. Also included is the sum-product of the covariate and model coefficient columns, which provides an alternative 'intercept' value for the algorithms for those users without detailed covariate information for their own application. These algorithms from Table 5 were used to map the ISI to the EQ-5D-3L, and the predicted EQ-5D-3L utility values for each ISI score are listed

Table 2 ALDVMM component identification – model fitting performance based on the full dataset

Model	Fitting performance parameters					Model rank				
	MAE	MSE	R^2	AIC	BIC	MAE	MSE	R^2	AIC	BIC
ALDVMM2 components	0.135876	0.0349	0.3142	−15,853.7	−15,370.4	2	2	2	3	3
ALDVMM3 components	0.13591	0.034691	0.31982	−16,107.2	−15,304.2	3	1	1	2	2
ALDVMM4 components	0.1353	0.035237	0.30513	−17,804.9	−16,682.4	1	3	3	1	1

AIC Akaike information criterion, *ALDVMM* adjusted limited dependent variable mixture model, *BIC* Bayesian information criterion, *MAE* mean absolute error, *MSE* mean-squared error, R^2 coefficient of determination

Table 3 All models – model fitting performance based on the full dataset

Model	Fitting performance parameters					Model rank				
	MAE	MSE	R^2	AIC	BIC	MAE	MSE	R^2	AIC	BIC
OLS regression	0.13735	0.03535	0.30212	−9017.1	−8853.4	4	3	3	3	3
CLAD	0.13434	0.03686	0.27316	–	–	1	4	4	–	–
Gamma-log GLM	0.13573	0.03534	0.30309	−14,408.1	−14,244.4	3	2	2	2	2
ALDVMM3 components	0.13591	0.034691	0.31982	−16,107.2	−15,304.2	2	1	1	1	1

AIC Akaike information criterion, *ALDVMM* adjusted limited dependent variable mixture model, *BIC* Bayesian information criterion, *CLAD* censored least absolute deviation, *GLM* generalised linear model, *MAE* mean absolute error, *MSE* mean-squared error, *OLS* ordinary least squares, R^2 coefficient of determination

in Table 6. Figure 3 shows the EQ-5D-3L disutility associated with each one-point increase in ISI for the ALDVMM3 and gamma-log GLM compared with the Model II algorithm reported by Gu et al., and Fig. 4 shows the predicted utility scores for the same three models. In contrast to Figs. 2, 3 and 4 show the independent effect of ISI on EQ-5D-3L utility after controlling for other covariates.

The *R* scripts to run the analysis are provided in the electronic supplementary material (ESM). In addition, a Microsoft Excel™ 'calculator', demonstrating how the ISI score can be obtained from the gamma-log GLM and ALDVMM3 models presented in the manuscript, is also available as part of the ESM. This calculator also includes the full regression coefficients, standard errors, and covariance matrices for the gamma-log GLM and ALDVMM3 models.

3.3 Validation Against the Model Developed by Gu et al.

Similar to the study by Gu et al. [14], we obtained a sufficient correlation between ISI summary total score and EQ-5D-3L health state utilities. Absolute utility values in our dataset were lower; this was expected given the use of the UK EQ-5D-3L tariff, which is known to give lower utility scores than tariffs from other countries (see, for example, Kiadaliri et al. [22]). The shape of the final mapping functions of ISI total score to EQ-5D-3L was similar between both studies (Figs. 3, 4); the current analysis resulted in a slower down curve toward the highest ISI total scores, indicating a worse QoL in those patients with more severe insomnia. Our models and that of Gu et al. [14] crossover at ISI = 12; under this threshold, any point increase in ISI results in higher disutility in our model, while the opposite is true above this threshold. As an example, a 1-point ISI increase from ISI = 22 to ISI = 23 results in a disutility of 0.020 in the study by Gu et al. [14] and 0.023 in the present study (Fig. 3). The gap between the two models widens at higher ISI scores because our model adjusts for confounding factors such as comorbidities, unlike Model II reported by Gu et al. [14]. Therefore, we believe that our analysis is more conservative. Finally, as highlighted in Figs. 3 and 4, the gamma-log GLM and the ALDVMM3 showed very similar predicted values, meaning that the difference between them is unlikely to prove important in most applications.

4 Discussion

This study expands upon prior literature by providing an updated mapping between insomnia severity (via the ISI) and a preference-based measure of health-related QoL, the EQ-5D. A prior mapping between ISI and EQ-5D performed by Gu et al. was reported in 2011 for a US population using US tariff values. The updated mapping presented here is based on this earlier work, using their same preferred functional form and supplementing it with the use of a contemporary representative dataset based on UK tariff values, consisting of validated PROs from multiple countries (France,

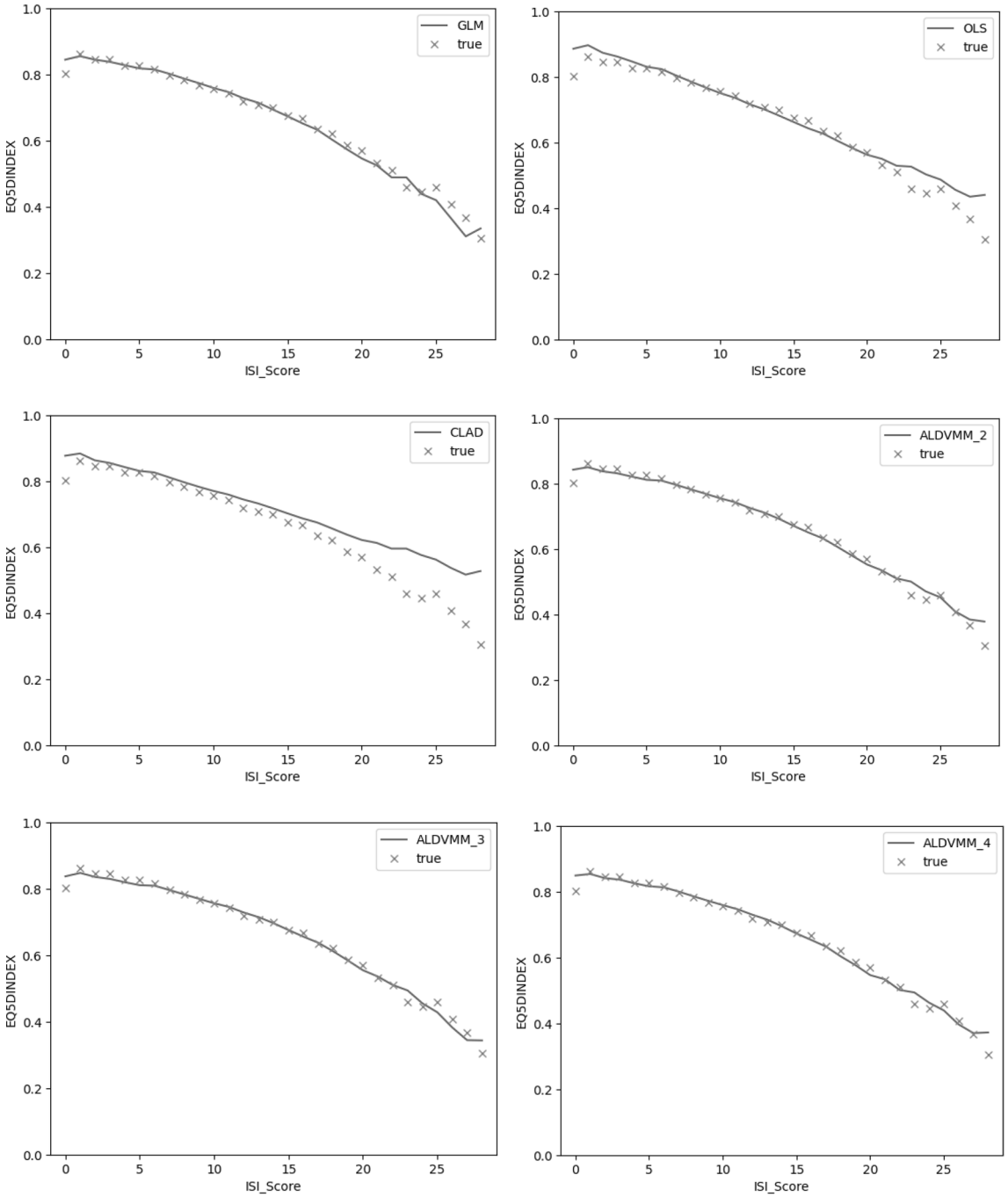


Fig. 2 Average predictions of the various models on the full dataset (solid line) and observed values (crosses): **a** GLM gamma-log; **b** OLS regression; **c** CLAD; and **d**–**f** ALDVMM with two to four components, respectively. Note: the relationship between ISI and EQ-5D in this figure does not control for covariate patterns that vary across the ISI categories. Regarding the apparently low EQ-5D-3L utility corresponding to an ISI score of 0, very few respondents scored a zero on the ISI ($n = 81$), therefore the low EQ-5D-3L utility may be considered an anomaly due to the low sample size. *ALDVMM* adjusted limited dependent variable mixture model, *CLAD* censored least absolute deviation, *EQ-5D-3L* EuroQol 5 dimensions 3 levels, *GLM* generalised linear model, *ISI* insomnia severity index, *OLS* ordinary least squares

Germany, Italy, Spain, the UK and the US), and adjusting for potential confounding factors (including patient characteristics, comorbidities and insomnia characteristics) to obtain the independent effect of ISI on EQ-5D score. The importance of adjusting for potential covariates cannot be overemphasised. Despite the similarities between our preferred algorithm and that of the earlier Gu et al. algorithm, the greater slope at higher values of ISI for the Gu et al.

Table 4 Model performance based on the predictive test dataset across 100 repetitions

Model	Fitting performance parameters		Model rank	
	MAE	MSE	MAE	MSE
OLS regression	0.137606	0.035486	4	3
CLAD	0.134726	0.037013	1	4
Gamma-log GLM	0.135979	0.035475	3	2
ALDVMM3 components	0.135785	0.035091	2	1

ALDVMM adjusted limited dependent variable mixture model, *CLAD* censored least absolute deviation, *GLM* generalised linear model, *MAE* mean absolute error, *MSE* mean-squared error, *OLS* ordinary least squares

algorithm can likely be attributed to an association with greater comorbidities. We see this effect in our own data when comparing the fitted values in Fig. 2, where the algorithm is predicted back onto the observed data, with Fig. 4, where the ISI relationship is shown at the mean of all other covariates. The lower slope after adjustment shows the independent effect of ISI on utility, which represents the most appropriate approach to linking a clinical effect on ISI to estimated effect on utility.

In addition to the work by Gu et al. [14] and in agreement with the ISPOR Task Force report on 'Mapping to estimate health-state utility from non-preference-based outcome measures' [15], other appropriate models (i.e. OLS regression, CLAD and ALDVMM) were tested against the gamma-log GLM model. The goal of our study was to develop a mapping algorithm from a UK perspective, which was accomplished by applying the latest NICE-approved UK tariff to all observations based on a crosswalk algorithm derived from the original EQ-5D-3L tariff. The use of a UK tariff for all five non-UK countries in the NHWS dataset means that remaining country-specific differences relate to potential differences in how subjects complete the instruments in different jurisdictions, not the tariff weights. This

Table 5 Regression coefficients for the gamma-log GLM and ALDVMM models

Covariates	Gamma-log GLM	ALDVMM (continued)				
		comp1	comp2	comp3	delta1	delta2
Intercept	−1.337	0.814	0.421	0.307	3.783	2.522
Female	−0.054	−0.001	−0.026	0.096	0.628	0.296
Married	−0.029	0.001	−0.011	−0.004	−0.014	−0.386
Degree educated	−0.059	0.011	0.02	0.012	−0.134	−0.301
Employed	−0.118	0.011	0.089	−0.091	0.805	0.583
Retired	−0.052	0.007	0.081	0	0.317	0.361
Current smoker[a]	0.118	−0.014	0.007	−0.039	−0.664	−0.279
Former smoker[a]	0.049	−0.009	−0.007	−0.02	0.556	0.764
Heavy drinker[a]	−0.043	0.007	0.078	−0.058	−0.307	−0.233
Low–moderate drinker[a]	−0.08	0.007	0.037	0.039	0.076	−0.232
DASD	0.292	−0.044	−0.049	0.179	−1.131	−0.245
Pain	0.25	−0.052	−0.083	−0.09	−0.236	−0.042
Insomnia treated	0.04	0.012	0.039	−0.302	0.299	0.938
Insomnia diagnosed	−0.008	0.001	−0.021	−0.079	0.301	0.242
Non-UK	−0.06	0	0.089	0.122	−0.37	−0.431
ISI score (std)	0.25	−0.029	−0.011	−0.174	−1.279	−0.578
Age (std)	0.005	−0.014	−0.021	−0.005	0.099	−0.259
BMI (std)	0.067	−0.014	−0.014	−0.03	−0.167	−0.058
CCI (std)	0.074	−0.019	0.012	0.006	−0.6	−0.193
ISI x treatment	−0.055	0.004	−0.021	0.136	−0.167	−0.43
Sigma		−2.291	−1.596	−1.379		
Sum product[b]	−1.284	0.786	0.446	0.337	4.173	2.689

ALDVMM adjusted linear dependent variable mixture model, *BMI* body mass index, *CCI* Charlson Comorbidity Index, *DASD* depression or anxiety or post-traumatic stress disorder, *GLM* generalised linear model (fitted on the disutility scale with a gamma distribution family and a log link),*ISI* Insomnia Severity Index, *UK* United Kingdom

[a]Smoker and alcohol user status are encoded as three-level categorical variables

[b]SumProduct cross multiplies the covariate column with the coefficient column and gives an 'alternative intercept' that can be used where detailed information on covariates are not available

Table 6 EQ-5D utility scores predicted for gamma-log GLM - and ALDVMM3 models

ISI	Gamma-log GLM	ALDVMM
0	0.843	0.832
1	0.836	0.825
2	0.828	0.818
3	0.819	0.81
4	0.811	0.802
5	0.802	0.793
6	0.792	0.784
7	0.782	0.775
8	0.772	0.765
9	0.761	0.754
10	0.749	0.743
11	0.737	0.731
12	0.725	0.719
13	0.711	0.706
14	0.697	0.692
15	0.683	0.677
16	0.668	0.662
17	0.652	0.645
18	0.635	0.627
19	0.617	0.608
20	0.599	0.588
21	0.58	0.567
22	0.56	0.544
23	0.538	0.52
24	0.516	0.495
25	0.493	0.468
26	0.469	0.439
27	0.443	0.409
28	0.416	0.377

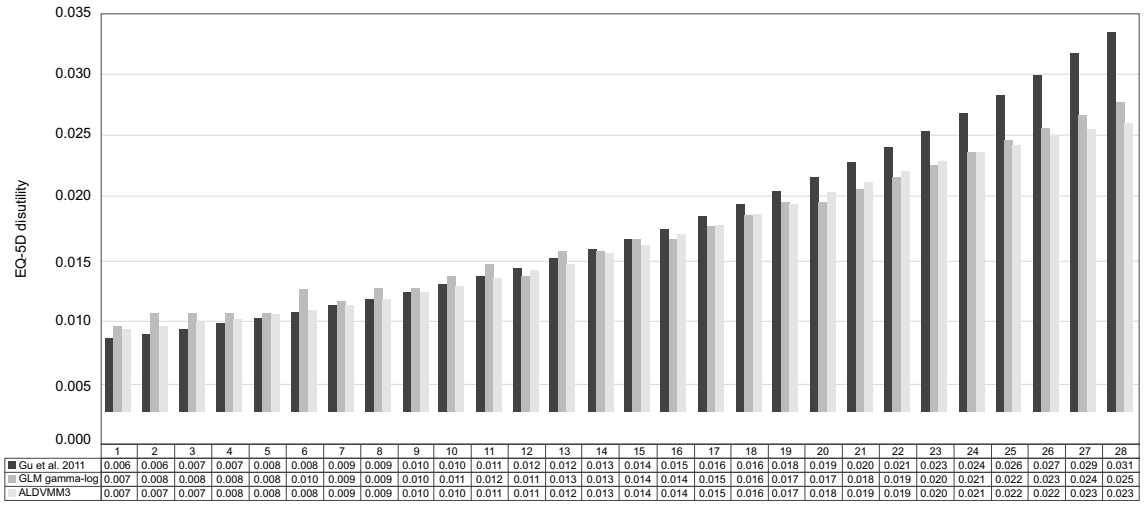
ALDVMM adjusted linear dependent variable mixture model, *GLM* generalised linear model, *ISI* Insomnia Severity Index

was handled in our analysis by using a fixed-effect covariate adjusting for subjects not in the UK, leaving the UK as the reference category.

The ALDVMM3 model proved to be the best fitting/predictive model, albeit by a very small margin, compared with the GLM gamma-log model, the second-best fitting/predictive model. The OLS regression and the CLAD models were the poorest performing both on fitting and prediction criteria. Despite its better performance, the ALDVMM3 was a time-consuming model to fit as it was solved by iteration and required numerous starting values to ensure that the fit was to a global rather than local maxima (this took more than 2 h to fit on a standard laptop). It is not surprising that the ALDVMM3 results showed a better fit to the entire dataset, given that this model uses five times as many parameters as the gamma-log GLM. Although the advantage of ALDVMM3 over the gamma-log GLM is reduced in the predictive task, it still performs better overall and therefore may be preferred by some analysts who are concerned only with getting an accurate algorithm for mapping purposes. For others, who also value parsimony and interpretability, the gamma-log GLM algorithm may be preferred to the ALDVMM3. Nevertheless, the practical differences between them are small (Figs. 3, 4) and readers who want to implement an ALDVMM algorithm can easily do so using the simple Microsoft Excel calculator available in the ESM.

In the absence of preference-based measures, the algorithms presented herein can be used to predict utility values from the ISI total summary score. Mapping studies between disease-specific instruments and generic preference-based instruments are common, driven largely by the need for quality-adjusted life-year (QALY) metrics in reimbursement decisions. Indeed, the latest NICE guidance specifically refers to the need for mapping studies to obtain EQ-5D estimates, where direct measurement of the EQ-5D is lacking. However, the legitimacy of any mapping exercise must be grounded in the conceptual overlap between the two instruments [24].

Our analysis is limited in that it may potentially underestimate the burden of insomnia due to the condition being potentially insufficiently captured by the EQ-5D. There is no empirical test for this, but some conceptual understanding of the problem can be inferred from the literature. For example, Perneger and Courvoisier [25] identify 'Sleep' and 'Fatigue/energy' as two of five possible dimensions missing from the EQ-5D. Fatigue has been the subject of so-called 'bolt-on' developments of the EQ-5D [10], most notably because fatigue is a common adverse effect of many health conditions and treatments for them, but it is clear that fatigue will also be a major consequence of next-day functioning for individuals suffering from insomnia. If the EQ-5D itself fails to capture important dimensions of QoL related to sleep deprivation, then no mapping, however statistically accurate, will be able to account for this deficiency and the QALY burden of insomnia will inevitably be underestimated. Another limitation of this study is that adjusting the results for other countries would require re-running the analyses with both a country-specific tariff and a country-specific fixed-effect covariate. However, we believe our approach is optimal since it utilises the full dataset of values and provides the most robust results. Moreover, EQ-5D-5L value sets exist for several countries [23]. Even though full health is reported less frequently on the EQ-5D-5L, as compared with the EQ-5D-3L, the average severity of reported problems is also less (i.e. slight problems instead of moderate problems). The comparability between scores obtained from the 3L and 5L versions of the EQ-5D depends on the values attached to health



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
Gu et al. 2011	0.006	0.006	0.007	0.007	0.008	0.008	0.009	0.009	0.010	0.010	0.011	0.012	0.012	0.013	0.014	0.015	0.016	0.016	0.018	0.019	0.020	0.021	0.023	0.024	0.026	0.027	0.029	0.031
GLM gamma-log	0.007	0.008	0.008	0.008	0.008	0.010	0.009	0.009	0.010	0.011	0.012	0.011	0.013	0.013	0.014	0.014	0.015	0.016	0.017	0.017	0.018	0.019	0.020	0.021	0.022	0.023	0.024	0.025
ALDVMM3	0.007	0.007	0.007	0.008	0.008	0.008	0.009	0.009	0.010	0.010	0.011	0.011	0.012	0.013	0.014	0.014	0.015	0.016	0.017	0.018	0.019	0.019	0.020	0.021	0.022	0.022	0.023	0.023

Fig. 3 Mapping functions on the full dataset: EQ-5D health state disutilities per ISI summary total point increase in the current study compared with the study by Gu et al. [14]. In this figure, the independent effect of ISI on utility is shown after controlling for other covariates. *ALDVMM* adjusted limited dependent variable mixture model, *EQ-5D* EuroQol 5 dimensions, *GLM* generalised linear model, *ISI* Insomnia Severity Index

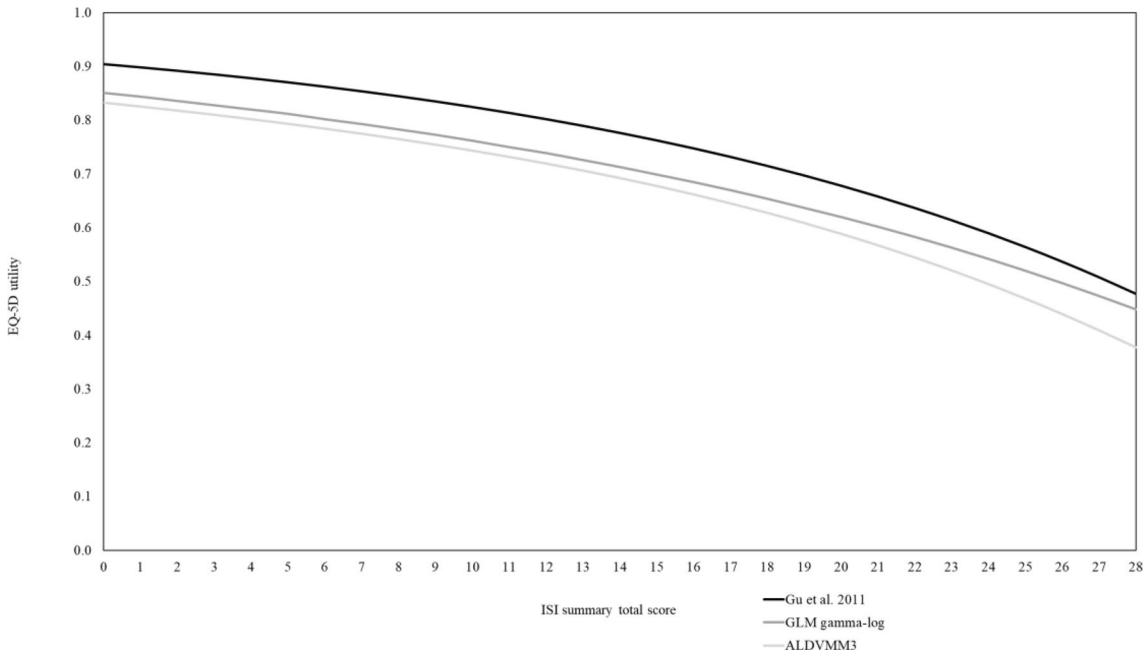


Fig. 4 Mapping functions on the full dataset: EQ-5D health state utilities per ISI summary total score (0–28) in the current study compared with the study by Gu et al. [14]. In this figure, the independent effect of ISI on utility is shown after controlling for other covariates. *ALDVMM* adjusted limited dependent variable mixture model, *EQ-5D* EuroQol 5 dimensions, *GLM* generalised linear model, *ISI* Insomnia Severity Index

states, which vary across countries. In some countries, the results obtained from the 3L and 5L versions of the EQ-5D will be similar, while in other countries, the two EQ-5D versions may produce different results [26]. Therefore, using country-specific value sets instead of the UK EQ-5D-3L in our analysis could have resulted in different results and provided a more international perspective.

5 Conclusions

In the absence of preference-based measures, this study provides updated mapping algorithms for estimating EQ-5D utilities from the ISI summary total score. This new mapping draws its strengths from the use of a large international dataset and also the incorporation of adjustment variables (including sociodemographic and general health characteristics) to reduce the effects of confounders. Nevertheless, users should consider whether the limitations of the EQ-5D instrument itself in capturing all relevant domains of insomnia could still lead to underestimation of utility in application of the algorithms reported here.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s41669-023-00388-0.

Acknowledgements The authors would like to thank Dr Charles Morin for his review and comments on this manuscript.

Declarations

Funding Idorsia Pharmaceuticals Ltd funded this research study, including access to the NHWS data and the analytical work by Visium.

Conflicts of interest/competing interest Teodora Bujaroska, Nizar Ghandri, Emilio Tanowe Maddalena, and Matteo Togninalli are employees of Visium Ltd, which received funding from Idorsia Pharmaceuticals Ltd to conduct this study. Kushal Modi, Abisola Olopoenia and Jeffrey Thompson are employees of Cerner Enviza, which received funding from Idorsia Pharmaceuticals Ltd to conduct this study. Andrew Briggs and Evi Germeni were contracted through Avalon Health Economics LLC, which received fees from Idorsia Pharmaceuticals Ltd. François-Xavier Chalet is an employee of, and holds shares in, Idorsia Pharmaceuticals Ltd.

Ethics approval The protocol and questionnaire for the NHWS were reviewed and granted exemption status by Pearl Institutional Review Board (Indianapolis, IN, USA; 19-KANT-204).

Data availability The NHWS is a commercial dataset. Access to the dataset used in this present paper must be requested to Cerner Enviza via the following link: https://www.cernerenviza.com/contact-us. Data access conditions will be detailed by the responsible team.

Consent to participate All respondents have provided informed consent prior to their participation.

Consent for publication (from patients/participants) All respondents have provided informed consent prior to their participation, with the potential for publication.

Code availability R scripts to reproduce the analyses are available in the ESM. Code needs to be run against the original dataset.

Author contributions Conceptualisation: F-XC, AB. Methodology: F-XC, JT, MT, AB. Formal analysis and investigation: TB, NG, ETM, KM. Writing—original draft preparation: F-XC, NG, ETM, KM, JT, MT, AB. Writing—review and editing: F-XC, TB, EG, AO, JT, MT, AB. Approval of the final version: All authors.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc/4.0/.

References

1. American Psychiatric Association. Diagnostic and statistical manual of mental disorders: DSM-5. Washington, DC: American Psychiatric Association; 2013.
2. Ohayon MM. Epidemiology of insomnia: what we know and what we still need to learn. Sleep Med Rev. 2002;6(2):97–111.
3. Riemann D, Baglioni C, Bassetti C, Bjorvatn B, Dolenc Groselj L, Ellis JG, et al. European guideline for the diagnosis and treatment of insomnia. J Sleep Res. 2017;26(6):675–700.
4. Rosekind MR, Gregory KB. Insomnia risks and costs: health, safety, and quality of life. Am J Manag Care. 2010;16(8):617–26.
5. Vgontzas AN, Liao D, Pejovic S, Calhoun S, Karataraki M, Basta M, et al. Insomnia with short sleep duration and mortality: the Penn State cohort. Sleep. 2010;33(9):1159–64.
6. Wickwire EM, Shaya FT, Scharf SM. Health economics of insomnia treatments: the return on investment for a good night's sleep. Sleep Med Rev. 2016;30:72–82.
7. Hafner M, Stepanek M, Taylor J, Troxel WM, Van Stolk C. Why sleep matters—the economic costs of insufficient sleep: a cross-country comparative analysis. Rand Health Q. 2017;6(4):11.
8. Morin CM. Insomnia: psychological assessment and management. New York: Guilford press; 1993.
9. Bastien CH, Vallières A, Morin CM. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. Sleep Med. 2001;2(4):297–307.
10. Geraerds A, Polinder S, Spronk I, Olde Loohuis AG, de Groot A, Bronner MB, et al. Sensitivity of the EQ-5D-5L for fatigue and cognitive problems and their added value in Q-fever patients. Qual Life Res. 2022;31(7):2083–92.
11. National Institute for Health and Care Excellence (NICE). Health technology evaluations: the manual. NICE; 2022.
12. NICE guideline no. 202. Obstructive sleep apnoea/hypopnoea syndrome and obesity hypoventilation syndrome in over 16s. Methods. London: National Institute for Health and Care Excellence; 2021.
13. McDaid C, Griffin S, Weatherly H, Duree K, Van der Burgt M, Van Hout S, et al. Continuous positive airway pressure devices for the treatment of obstructive sleep apnoea-hypopnoea syndrome: a systematic review and economic analysis. Health Technol Assess. 2009;13(4):iii–iv (**xi**).
14. Gu NY, Botteman MF, Ji X, Bell CF, Carter JA, van Hout B. Mapping of the Insomnia Severity Index and other sleep measures to EuroQol EQ-5D health state utilities. Health Qual Life Outcomes. 2011;9(1):1–13.
15. Wailoo AJ, Hernandez-Alava M, Manca A, Mejia A, Ray J, Crawford B, et al. Mapping to estimate health-state utility from non-preference-based outcome measures: an ISPOR good practices for outcomes research task force report. Value Health. 2017;20(1):18–27.

16. Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi J-C, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care. 2005;43:1130–9.
17. Hernández-Alava M, Pudney S. Mapping between EQ-5D-3L and EQ-5D-5L: a survey experiment on the validity of multi-instrument data. Health Econ. 2022;31(6):923–39.
18. 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):1–9.
19. Pullenayegum EM, Tarride JE, Xie F, Goeree R, Gerstein HC, O'Reilly D. Analysis of health utility data when some subjects attain the upper bound of 1: are Tobit and CLAD models appropriate? Value Health. 2010;13(4):487–94.
20. Alava MH, Wailoo A. Fitting adjusted limited dependent variable mixture models to EQ-5D. Stata J. 2015;15(3):737–50.
21. Pletscher M. Adjusted Limited Dependent Variable Mixture Models of Health State Utilities in R. 2022.
22. Kiadaliri AA, Eliasson B, Gerdtham U-G. Does the choice of EQ-5D tariff matter? A comparison of the Swedish EQ-5D-3L index score with UK, US, Germany and Denmark among type 2 diabetes patients. Health Qual Life Outcomes. 2015;13(1):1–10.
23. EuroQol. Comparing EQ-5D-3L and EQ-5D-5L value sets. 2022. https://euroqol.org/eq-5d-instruments/3l-vs-5l/comparing-eq-5d-3l-and-eq-5d-5l-value-sets/.
24. Dzingina MD, McCrone P, Higginson IJ. Does the EQ-5D capture the concerns measured by the Palliative care Outcome Scale? Mapping the Palliative care Outcome Scale onto the EQ-5D using statistical methods. Palliat Med. 2017;31(8):716–25.
25. Perneger TV, Courvoisier DS. Exploration of health dimensions to be included in multi-attribute health-utility assessment. Int J Qual Health Care. 2011;23(1):52–9.
26. Janssen MF, Buchholz I, Golicki D, Bonsel GJ. Is EQ-5D-5L better than EQ-5D-3L over time? A head-to-head comparison of responsiveness of descriptive systems and value sets from nine countries. Pharmacoeconomics. 2022;40(11):1081–93.