A hypothetical study using fake data to map K10 psychological distress and SOFAS functioning measures toAQoL-6D health utility using data from a sample of young people presenting to primary mental health services

Alejandra R Scienceace[1,2,*] Fionn S Researchchamp[3,2]

Abstract

Background: Quality Adjusted Life Years (QALYs) are often used in economic evaluations, yet utility weights for deriving them are rarely directly measured in mental health services.

Objectives: We aimed to identify the best regression models to predict Adolescent AQoL Six Dimension (AQoL-6D) utility and evaluate the predictive ability of two candidate measures of psychological distress and functioning.

Methods: The study sample is fake data.Five ordinary least squares (OLS) and three generalised linear models (GLMs) were explored to identify the best algorithm. Predictive ability of two candidate measures of psychological distress and functioning were assessed using ten fold cross validation and forest models.

Results: K10 had the highest predictive ability in ten fold cross validation. GLM with Gaussian distribution and log link and OLS with complementary log log transformation were the best peforming models.

Conclusions: Nothing should be concluded from this study as it is purely hypothetical.

Data: Detailed results in the form of catalogues of the models produced by this study and other supporting information are available in the online repository: https://doi.org/10.7910/DVN/LYBMB0

[1] Awesome University, Shanghai
[2] August Institution, London
[3] Highly Ranked Uni, Montreal

[*] Correspondence: Alejandra R Scienceace <fake_email@fake_institute.com>

1 Introduction

This article is a scientific summary of a study that the authors implemented that has been automatically generated from study results by version 0.9.0.1 of the manuscript authoring program ttu_lng_ss [1].

Quality Adjusted Life Years (QALYs) are often used in economic evaluations, yet utility weights for deriving them are rarely directly measured in mental health services.

The objective of the study was to identify the best utility mapping models to predict Adolescent AQoL Six Dimension (AQoL-6D) utility and evaluate the predictive ability of two candidate measures of psychological distress and functioning.

2 Methods

2.1 Sample and setting

The study sample is fake data.

2.2 Measures

Data was collected on utility weights, two candidate predictors of utility weights and descriptive population characteristics.

2.2.1 Utility weights

Utility weights were assessed using the AQoL-6D multi-attribute utility instrument.

2.2.2 Candidate predictors

Two measures of psychological distress (one measure) and functioning (one measure) were used as candidate predictors to construct models.

Psychological distress was measured by Kessler Psychological Distress Scale (10 Item) (K10 - measured on a scale of 10-50). Functioning was measured by Social and Occupational Functioning Assessment Scale (SOFAS - measured on a scale of 0-100).

2.2.3 Population characteristics

Population characteristic data were age, gender, aboriginal and torres strait islander, culturally and linguistically diverse, employment type, in education, education and employment, days unable to perform usual activities, days cut back on usual activities, days out of role, primary diagnosis group, area index of relative social disadvantage and area remoteness.

2.3 Statistical analysis

We implemented the generalised form of the study analysis algorithm developed by Hamilton, Gao and colleagues [2], the key steps of which are summarised as follows.

2.3.1 Descriptive statistics

Basic descriptive statistics were used to characterise the cohort in terms of baseline population variables. Pearson's Product Moment Correlations (r) were used to determine the relationships between candidate predictors and the AQoL-6D utility score.

2.3.2 Model evaluation

We compared predictive performance of a range of models predicting AQoL-6D utility scores using the candidate predictor that had the highest Pearson correlation coefficient with utility scores. The models compared include ordinary least squares (OLS) regression models and generalised linear models (GLMs). OLS regression models used no transformation, complementary log log transformation, log transformation, logit transformation and log log transformation. GLMs used gaussian distribution and log link, beta distribution and complementary log log link and beta distribution and logit link. Ten-fold cross-validation was used to compare model fitting with training datasets and predictive ability with testing datasets using three indicators including R^2, root mean square error (RMSE) and mean absolute error (MAE) [3,4].

To evaluate whether candidate predictors could independently predict utility scores, we established multivariate prediction models using baseline data with the candidate predictor and demographic, clinical symptom and spatial covariates. Demographic covariates were aboriginal and torres strait islander, age, culturally and linguistically diverse, education and employment and gender. Clinical symptom covariates were days out of

role and primary diagnosis group. Spatial covariates were area index of relative social disadvantage and area remoteness.

2.3.3 Candidate predictor comparison

We evaluated the independent predictive ability of different candidate predictors using 10-fold cross-validation.

2.3.4 Utility mapping models

We next established models using linear mixed effect models (LMMs) and generalised linear mixed effect models (GLMMs) . Model fitting was evaluated using Bayesian R^2 [5].

2.3.5 Software

We undertook all our analyses using R version 4.3.0 [6] using the TTU package [7] (version 0.0.0.9362).

3 Results

3.1 Cohort characteristics

Participants characteristics are displayed in Table 1. This study included all 3998 participants with complete AQoL-6D data.

3.2 AQoL-6D and candidate predictors

Distribution of AQoL-6D total utility score and sub-domain scores are displayed in Figure 1. The mean utility score is 0.536 (SD = 0.229). The distribution of candidate predictors, K10 and SOFAS, are summarised in Table 2. K10 was found to have the highest correlation with utility score both at baseline and follow-up followed by SOFAS.

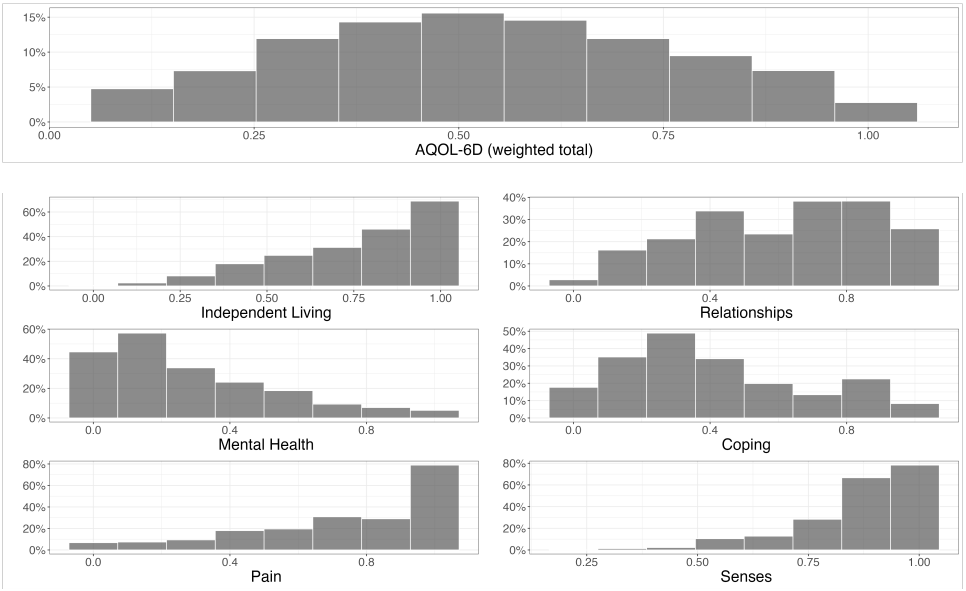

Figure 1: Distribution of AQoL-6D domains

Table 1: Participant characteristics

		(N =	3998)
Age	Mean (SD)	17.692	(3.477)
	Median (Q1 Q3)	17.000	(15.000 20.000)
	Min - Max	12.000	25.000
	Missing	0.000	
Gender	Female	2457.000	(61.456%)
	Male	1451.000	(36.293%)
	Other	90.000	(2.251%)
	Missing	0.000	
Aboriginal and torres strait islander	No	3698.000	(92.519%)
	Yes	299.000	(7.481%)
	Missing	1.000	
Culturally and linguistically diverse	No	2881.000	(72.061%)
	Yes	1117.000	(27.939%)
	Missing	0.000	
Employment type	No	2262.000	(57.926%)
	Working casually	940.000	(24.072%)
	Working full-time	309.000	(7.913%)
	Working part-time	394.000	(10.090%)
	Missing	93.000	
In education	No	1514.000	(38.771%)
	Yes	2391.000	(61.229%)
	Missing	93.000	
Education and employment	Not studying or working	782.000	(20.031%)
	Studying and working	910.000	(23.309%)
	Studying only	1480.000	(37.910%)
	Working only	732.000	(18.750%)
	Missing	94.000	
Days unable to perform usual activities	Mean (SD)	9.827	(8.854)
	Median (Q1 Q3)	8.000	(2.000 15.000)
	Min - Max	0.000	28.000
	Missing	1.000	
Days cut back on usual activities	Mean (SD)	5.989	(7.233)
	Median (Q1 Q3)	3.000	(0.000 10.000)
	Min - Max	0.000	28.000
	Missing	3.000	
	Mean (SD)	15.820	(12.465)
	Median (Q1 Q3)	15.000	(4.000 26.000)

Table 2: Candidate predictors distribution parameters and correlations with AQoL-6D utility

		(N =	3998)	*p*
Kessler Psychological Distress Scale (10 Item) (10-50)	Mean (SD)	28.569	(8.343)	
	Missing	2.000		
	Correlation with AQOL-6D	-0.655		0.000
Social and Occupational Functioning Assessment Scale (0-100)	Mean (SD)	64.821	(14.937)	
	Missing	700.000		
	Correlation with AQOL-6D	0.191		0.000

3.3 Performance of regression models

The 10-fold cross-validated model fitting index from TTU models using K10 are reported in Table A.1 in the Supplementary Material. Both GLM with Gaussian distribution and log link and OLS with complementary log log transformation were selected for further evaluation. Predictive ability of each candidate predictor using baseline data were also compared using 10-fold cross-validation.

Table A.2 illustrates that K10 had the highest predictive ability followed by SOFAS.

The confounding effect of other participant characteristics when using the candidate predictors in predicting utility score were also evaluated. Days out of role, age, gender and education and employment were found to independently predict utility scores in models for both candidate predictors *($p<0.01$)*.

3.4 Utility mapping models

Regression coefficients of the baseline score and score changes (from baseline to follow-up) estimated in individual GLMM and LMM models are summarised in Table 3. Bayesian R^2 and modelled residual standard deviations (SDs) from each model are reported. In both GLMM and LMM models, the prediction models using K10 had the highest R^2 (0.438 and 0.609). R^2 was between 0.082 and 0.438 for all GLMMs and between 0.428 and 0.609 for all LMMs.

Table 3: Estimated coefficients from utility mapping models

	GLMM - Gaussian distribution and log link					LMM - complementary log log transformation				
Parameter	Estimate	SE	CI (95%)	R2	Sigma	Estimate	SE	CI (95%)	R2	Sigma
K10 model				**0.438**	**0.171**				**0.609**	**0.480**
SD (Intercept)	0.034	0.022	0.001, 0.079			0.304	0.164	0.016, 0.570		
Intercept	0.280	0.015	0.249, 0.310			1.437	0.034	1.370, 1.504		
K10 scaled	-3.297	0.061	-3.419, -3.177			-6.114	0.114	-6.339, -5.890		
SOFAS model				**0.082**	**0.220**				**0.428**	**0.570**
SD (Intercept)	0.067	0.044	0.003, 0.157			0.443	0.220	0.044, 0.740		
Intercept	-1.048	0.040	-1.129, -0.972			-0.931	0.061	-1.053, -0.810		
SOFAS scaled	0.653	0.060	0.539, 0.771			0.980	0.091	0.802, 1.163		

Note: The K10 and SOFAS parameters were first multiplied by 0.01.

Distribution of observed and predicted utility scores and their association from GLMM (Gaussian distribution and log link) and LMM (complementary log log transformation) using K10 are plotted in Figure 2.

We also evaluated models with cdaysoor, dage, dgender and dstudyingworking at baseline and cdaysoor, dage, dgender and dstudyingworking change from baseline added to psychological distress, functioning, clinical symptom and demographic predictors (see Tables A.3 and A.4).

Detailed summaries of all models are available in the online data repository (see "Availability of data and materials").

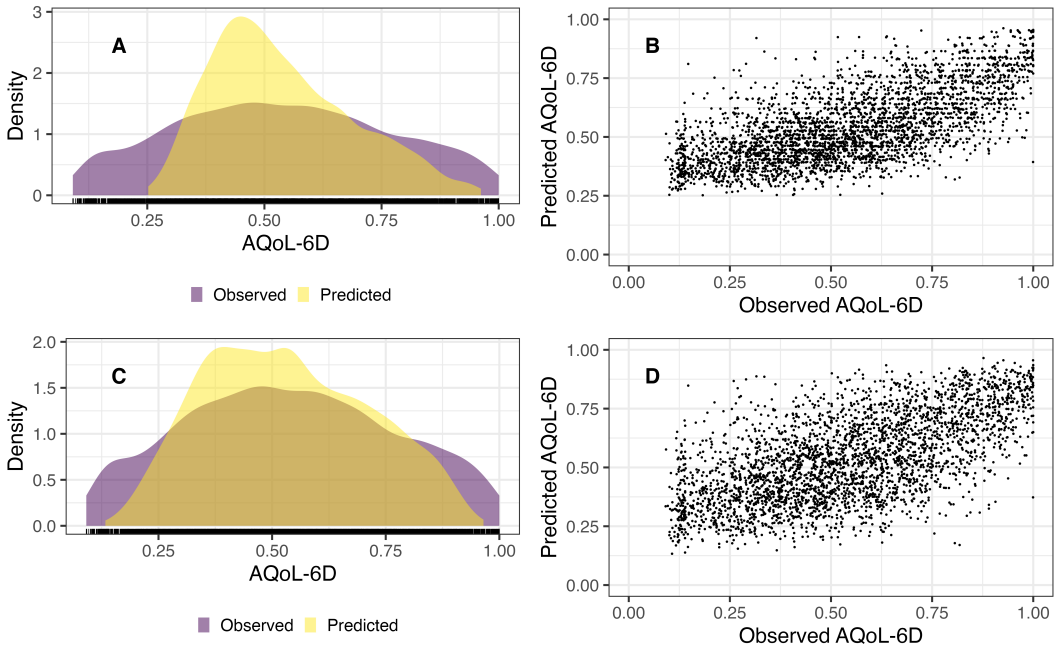

Figure 2: Comparison of observed and predicted AQoL-6D score from longitudinal model using K10 (A) Density plots of observed and predicted utility scores (GLMM (Gaussian distribution and log link)) (B) Scatter plots of observed and predicted utility scores by timepoint (GLMM (Gaussian distribution and log link)) (C) Density plots of observed and predicted results (LMM (complementary log log transformation)) (D) Scatter plots of observed and predicted results by timepoint (LMM (complementary log log transformation))

4 Conclusions

Nothing should be concluded from this study as it is purely hypothetical.

Availability of data and materials

Detailed results in the form of catalogues of the models produced by this study and other supporting information are available in the online repository: https://doi.org/10.7910/DVN/LYBMB0

Ethics approval

The study was reviewed and granted approval by no-one.

Funding

The study was funded by no-one.

Conflict of Interest

None declared

References

1. Hamilton M, Gao C. ttu_lng_ss: Create a Draft Scientific Manuscript For A Utility Mapping Study [Internet]. Zenodo; 2023. doi:10.5281/zenodo.5976987
2. Hamilton MP, Gao CX, Filia KM, Menssink JM, Sharmin S, Telford N, et al. Predicting quality adjusted life years in young people attending primary mental health services. medRxiv. Cold Spring Harbor Laboratory Press; 2021; doi:10.1101/2021.07.07.21260129
3. Hastie T, Tibshirani R, Friedman J. The elements of statistical learning: Data mining, inference, and prediction. Springer Science & Business Media; 2009.
4. Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. Ijcai. Montreal, Canada; pp. 1137–1145.
5. Gelman A, Goodrich B, Gabry J, Vehtari A. R-squared for bayesian regression models. The American Statistician. 2019;73: 307–309. doi:10.1080/00031305.2018.1549100
6. R Core Team. R: A language and environment for statistical computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2020. Available: https://www.R-project.org/
7. Gao C, Hamilton M. TTU: Implement Transfer to Utility Mapping Algorithms with Ready4 [Internet]. Zenodo; 2023. doi:10.5281/zenodo.5646593

A Supplementary Material

A.1 Additional tables

Table A.1: 10-fold cross-validated model fitting index for different GLM and OLS models using K10 as predictor with the baseline data

	Training model fit			Testing model fit		
Model	R2	RMSE	MAE	R2	RMSE	MAE
OLS						
No transformation	0.429	0.173	0.141	0.428	0.173	0.141
Complementary log log transformation	0.426	0.174	0.140	0.425	0.174	0.140
Log transformation	0.408	0.176	0.143	0.407	0.176	0.143
Logit transformation	0.392	0.179	0.142	0.390	0.179	0.142
Log log transformation	0.344	0.186	0.146	0.342	0.186	0.146
GLM						
Gaussian distribution and log link	0.433	0.173	0.140	0.432	0.173	0.140
Beta distribution and complementary log log link	0.431	0.173	0.140	0.430	0.173	0.140
Beta distribution and logit link	0.425	0.174	0.141	0.423	0.174	0.141

Results are averaged over ten folds. RMSE: Root Mean Squared Error; MAE: Mean Absolute Error

Table A.2: 10-fold cross-validated model fitting index for different candidate predictors estimated using GLM with Gaussian distribution and log link with the baseline data

	Training model fit			Testing model fit		
Model	R2	RMSE	MAE	R2	RMSE	MAE
K10	0.433	0.173	0.140	0.432	0.173	0.140
SOFAS	0.041	0.225	0.187	0.040	0.225	0.187

Results are averaged over ten folds. RMSE: Root Mean Squared Error; MAE: Mean Absolute Error

Table A.3: Estimated coefficients from utility mapping models based on individual candidate predictors with days out of role, age and gender and education and employment using GLMM (Gaussian distribution and log link)

Parameter	Estimate	SE	CI (95%)	R2	Sigma
K10 cdaysoor model				**0.451**	**1.050**
SD (Intercept)	0.322	0.525	0.001, 1.499		
Intercept	0.106	0.264	-0.383, 0.303		
K10 scaled	-2.531	0.725	-3.129, -1.380		
cdaysoor	-0.004	0.003	-0.008, 0.003		
K10 dage model				**0.442**	**0.171**
SD (Intercept)	0.038	0.023	0.002, 0.085		
Intercept	0.361	0.027	0.309, 0.414		
K10 scaled	-3.268	0.060	-3.385, -3.152		
dage	-0.005	0.001	-0.008, -0.002		
K10 dgender model				**0.440**	**0.171**
SD (Intercept)	0.035	0.022	0.002, 0.080		
Intercept	0.258	0.017	0.225, 0.290		
K10 scaled	-3.262	0.061	-3.379, -3.140		
dgenderMale	0.031	0.010	0.011, 0.051		
dgenderOther	-0.015	0.038	-0.089, 0.058		
K10 dstudyingworking model				**0.440**	**0.171**
SD (Intercept)	0.035	0.022	0.001, 0.081		
Intercept	0.237	0.019	0.199, 0.275		
K10 scaled	-3.287	0.062	-3.409, -3.166		
dstudyingworkingBoth	0.065	0.015	0.035, 0.095		
dstudyingworkingStudy	0.044	0.014	0.017, 0.072		
dstudyingworkingWork	0.041	0.017	0.008, 0.073		
K10 cdaysoor 2 GLM GSN LOG				**0.477**	**1.297**
SD (Intercept)	0.471	0.577	0.001, 1.585		
Intercept	5.271	6.931	0.335, 17.309		
K10 scaled	-2.002	1.009	-3.087, -0.676		
cdaysoor	0.002	0.018	-0.006, 0.071		
dage	-0.249	0.375	-0.891, -0.004		
K10 cdaysoor 3 GLM GSN LOG				**0.464**	**0.716**
SD (Intercept)	1.542	2.537	0.004, 5.936		
Intercept	0.500	0.643	-0.015, 1.825		
K10 scaled	-2.008	1.012	-3.095, -0.675		
cdaysoor	0.003	0.019	-0.006, 0.068		

Table A.3: Estimated coefficients from utility mapping models based on individual candidate predictors with days out of role, age and gender and education and employment using GLMM (Gaussian distribution and log link) *(continued)*

Parameter	Estimate	SE	CI (95%)	R2	Sigma
dgenderMale	-0.367	0.429	-0.989, 0.047		
dgenderOther	0.681	0.709	-0.088, 1.453		
K10 cdaysoor 4 GLM GSN LOG				**0.455**	**0.171**
SD (Intercept)	0.841	1.416	0.001, 3.294		
Intercept	-0.158	0.688	-1.945, 0.274		
K10 scaled	-2.362	1.006	-3.123, -0.674		
cdaysoor	0.003	0.020	-0.008, 0.069		
dstudyingworkingBoth	-0.210	0.452	-0.991, 0.082		
dstudyingworkingStudy	0.373	0.556	0.016, 1.325		
dstudyingworkingWork	-0.104	0.262	-0.526, 0.067		
K10 dage 2 GLM GSN LOG				**0.441**	**0.171**
SD (Intercept)	0.035	0.022	0.002, 0.081		
Intercept	0.340	0.028	0.285, 0.395		
K10 scaled	-3.230	0.062	-3.351, -3.109		
dage	-0.005	0.001	-0.008, -0.002		
dgenderMale	0.032	0.010	0.012, 0.053		
dgenderOther	-0.005	0.039	-0.084, 0.069		
K10 dage 3 GLM GSN LOG				**0.443**	**0.170**
SD (Intercept)	0.036	0.022	0.002, 0.082		
Intercept	0.354	0.033	0.290, 0.420		
K10 scaled	-3.253	0.062	-3.370, -3.131		
dage	-0.007	0.002	-0.010, -0.004		
dstudyingworkingBoth	0.069	0.016	0.039, 0.100		
dstudyingworkingStudy	0.028	0.015	0.000, 0.057		
dstudyingworkingWork	0.057	0.017	0.024, 0.091		
K10 dgender 2 GLM GSN LOG				**0.443**	**0.170**
SD (Intercept)	0.037	0.023	0.002, 0.083		
Intercept	0.206	0.021	0.165, 0.247		
K10 scaled	-3.241	0.062	-3.365, -3.119		
dgenderMale	0.038	0.010	0.018, 0.058		
dgenderOther	-0.015	0.038	-0.089, 0.056		
dstudyingworkingBoth	0.071	0.016	0.041, 0.102		
dstudyingworkingStudy	0.049	0.014	0.021, 0.077		
dstudyingworkingWork	0.045	0.017	0.012, 0.078		

Table A.3: Estimated coefficients from utility mapping models based on individual candidate predictors with days out of role, age and gender and education and employment using GLMM (Gaussian distribution and log link) *(continued)*

Parameter	Estimate	SE	CI (95%)	R2	Sigma
SOFAS cdaysoor model				**0.264**	**1.181**
SD (Intercept)	0.298	0.415	0.002, 1.218		
Intercept	-0.414	0.470	-0.815, 0.429		
SOFAS scaled	-0.086	0.788	-1.381, 0.578		
cdaysoor	-0.011	0.004	-0.015, -0.001		
SOFAS dage model				**0.091**	**0.219**
SD (Intercept)	0.064	0.044	0.003, 0.164		
Intercept	-0.806	0.055	-0.914, -0.700		
SOFAS scaled	0.641	0.058	0.530, 0.757		
dage	-0.013	0.002	-0.018, -0.009		
SOFAS dgender model				**0.104**	**0.218**
SD (Intercept)	0.073	0.045	0.003, 0.161		
Intercept	-1.076	0.039	-1.153, -1.001		
SOFAS scaled	0.634	0.058	0.525, 0.747		
dgenderMale	0.109	0.015	0.080, 0.137		
dgenderOther	-0.116	0.059	-0.239, -0.003		
SOFAS dstudyingworking model				**0.077**	**0.221**
SD (Intercept)	0.055	0.040	0.002, 0.146		
Intercept	-1.083	0.043	-1.167, -1.000		
SOFAS scaled	0.644	0.061	0.526, 0.763		
dstudyingworkingBoth	0.058	0.023	0.015, 0.102		
dstudyingworkingStudy	0.066	0.021	0.026, 0.107		
dstudyingworkingWork	0.016	0.025	-0.032, 0.065		
SOFAS cdaysoor 2 GLM GSN LOG				**0.344**	**1.207**
SD (Intercept)	0.597	0.777	0.002, 1.924		
Intercept	6.578	7.505	-0.605, 17.516		
SOFAS scaled	-0.274	0.796	-1.381, 0.564		
cdaysoor	0.000	0.024	-0.014, 0.076		
dage	-0.343	0.366	-0.898, -0.010		
SOFAS cdaysoor 3 GLM GSN LOG				**0.329**	**0.745**
SD (Intercept)	0.713	1.158	0.003, 3.439		
Intercept	0.301	1.356	-0.849, 2.535		
SOFAS scaled	-0.273	0.792	-1.375, 0.564		
cdaysoor	-0.009	0.012	-0.013, -0.002		

Table A.3: Estimated coefficients from utility mapping models based on individual candidate predictors with days out of role, age and gender and education and employment using GLMM (Gaussian distribution and log link) *(continued)*

Parameter	Estimate	SE	CI (95%)	R2	Sigma
dgenderMale	-0.338	0.460	-0.989, 0.117		
dgenderOther	0.638	0.758	-0.212, 1.463		
SOFAS cdaysoor 4 GLM GSN LOG				**0.262**	**0.198**
SD (Intercept)	0.860	1.427	0.002, 3.332		
Intercept	-0.935	0.377	-1.961, -0.643		
SOFAS scaled	0.187	0.516	-0.673, 0.589		
cdaysoor	0.004	0.032	-0.014, 0.085		
dstudyingworkingBoth	-0.237	0.437	-0.991, 0.059		
dstudyingworkingStudy	0.365	0.556	0.004, 1.324		
dstudyingworkingWork	-0.154	0.224	-0.527, 0.016		
SOFAS dage 2 GLM GSN LOG				**0.114**	**0.217**
SD (Intercept)	0.070	0.044	0.003, 0.155		
Intercept	-0.825	0.054	-0.934, -0.720		
SOFAS scaled	0.619	0.058	0.507, 0.734		
dage	-0.014	0.002	-0.018, -0.010		
dgenderMale	0.114	0.014	0.085, 0.142		
dgenderOther	-0.089	0.058	-0.205, 0.021		
SOFAS dage 3 GLM GSN LOG				**0.092**	**0.219**
SD (Intercept)	0.063	0.041	0.003, 0.149		
Intercept	-0.813	0.063	-0.939, -0.686		
SOFAS scaled	0.622	0.060	0.506, 0.740		
dage	-0.014	0.002	-0.019, -0.009		
dstudyingworkingBoth	0.061	0.022	0.018, 0.106		
dstudyingworkingStudy	0.034	0.021	-0.006, 0.076		
dstudyingworkingWork	0.046	0.025	-0.003, 0.095		
SOFAS dgender 2 GLM GSN LOG				**0.105**	**0.218**
SD (Intercept)	0.066	0.044	0.003, 0.158		
Intercept	-1.121	0.042	-1.204, -1.040		
SOFAS scaled	0.617	0.058	0.503, 0.731		
dgenderMale	0.116	0.015	0.087, 0.146		
dgenderOther	-0.114	0.060	-0.236, -0.001		
dstudyingworkingBoth	0.078	0.023	0.035, 0.123		
dstudyingworkingStudy	0.081	0.021	0.041, 0.121		
dstudyingworkingWork	0.030	0.024	-0.018, 0.077		

Table A.3: Estimated coefficients from utility mapping models based on individual candidate predictors with days out of role, age and gender and education and employment using GLMM (Gaussian distribution and log link) *(continued)*

Parameter	Estimate	SE	CI (95%)	R2	Sigma

Note: The K10 and SOFAS parameters were first multiplied by 0.01.

Table A.4: Estimated coefficients from utility mapping models based on individual candidate predictors with days out of role, age and gender and education and employment using LMM (complementary log log transformation)

Parameter	Estimate	SE	CI (95%)	R2	Sigma
K10 cdaysoor model				**0.625**	**0.470**
SD (Intercept)	0.307	0.156	0.024, 0.543		
Intercept	1.437	0.033	1.373, 1.502		
K10 scaled	-5.619	0.120	-5.854, -5.381		
cdaysoor	-0.009	0.001	-0.011, -0.007		
K10 dage model				**0.588**	**0.498**
SD (Intercept)	0.277	0.164	0.013, 0.517		
Intercept	1.682	0.056	1.571, 1.795		
K10 scaled	-6.063	0.114	-6.285, -5.840		
dage	-0.015	0.003	-0.020, -0.009		
K10 dgender model				**0.611**	**0.482**
SD (Intercept)	0.310	0.151	0.020, 0.552		
Intercept	1.391	0.037	1.318, 1.461		
K10 scaled	-6.042	0.117	-6.269, -5.807		
dgenderMale	0.071	0.020	0.032, 0.112		
dgenderOther	-0.032	0.064	-0.158, 0.095		
K10 dstudyingworking model				**0.754**	**0.365**
SD (Intercept)	0.428	0.156	0.025, 0.586		
Intercept	1.350	0.039	1.273, 1.429		
K10 scaled	-6.111	0.114	-6.337, -5.888		
dstudyingworkingBoth	0.143	0.029	0.087, 0.198		
dstudyingworkingStudy	0.092	0.026	0.042, 0.143		
dstudyingworkingWork	0.088	0.030	0.030, 0.149		
K10 cdaysoor 2 OLS CLL				**0.561**	**0.518**
SD (Intercept)	0.236	0.140	0.013, 0.490		
Intercept	1.708	0.056	1.598, 1.818		
K10 scaled	-5.547	0.122	-5.786, -5.306		
cdaysoor	-0.009	0.001	-0.011, -0.008		
dage	-0.016	0.003	-0.022, -0.011		
K10 cdaysoor 3 OLS CLL				**0.630**	**0.470**
SD (Intercept)	0.317	0.142	0.032, 0.519		
Intercept	1.390	0.036	1.321, 1.461		
K10 scaled	-5.544	0.123	-5.783, -5.302		
cdaysoor	-0.009	0.001	-0.011, -0.007		

Table A.4: Estimated coefficients from utility mapping models based on individual candidate predictors with days out of role, age and gender and education and employment using LMM (complementary log log transformation) *(continued)*

Parameter	Estimate	SE	CI (95%)	R2	Sigma
dgenderMale	0.074	0.020	0.034, 0.114		
dgenderOther	-0.042	0.064	-0.168, 0.084		
K10 cdaysoor 4 OLS CLL				**0.605**	**0.488**
SD (Intercept)	0.287	0.149	0.019, 0.516		
Intercept	1.369	0.039	1.293, 1.445		
K10 scaled	-5.632	0.123	-5.872, -5.393		
cdaysoor	-0.009	0.001	-0.010, -0.007		
dstudyingworkingBoth	0.109	0.029	0.052, 0.163		
dstudyingworkingStudy	0.081	0.026	0.031, 0.132		
dstudyingworkingWork	0.052	0.030	-0.007, 0.111		
K10 dage 2 OLS CLL				**0.715**	**0.400**
SD (Intercept)	0.395	0.161	0.042, 0.573		
Intercept	1.641	0.058	1.528, 1.755		
K10 scaled	-5.994	0.116	-6.222, -5.766		
dage	-0.015	0.003	-0.020, -0.010		
dgenderMale	0.077	0.020	0.037, 0.114		
dgenderOther	-0.006	0.065	-0.129, 0.119		
K10 dage 3 OLS CLL				**0.640**	**0.457**
SD (Intercept)	0.314	0.186	0.012, 0.539		
Intercept	1.687	0.067	1.555, 1.818		
K10 scaled	-6.053	0.116	-6.282, -5.824		
dage	-0.019	0.003	-0.026, -0.013		
dstudyingworkingBoth	0.146	0.030	0.087, 0.203		
dstudyingworkingStudy	0.045	0.028	-0.010, 0.099		
dstudyingworkingWork	0.128	0.032	0.066, 0.191		
K10 dgender 2 OLS CLL				**0.568**	**0.512**
SD (Intercept)	0.264	0.142	0.016, 0.509		
Intercept	1.287	0.042	1.203, 1.369		
K10 scaled	-6.026	0.117	-6.253, -5.797		
dgenderMale	0.085	0.020	0.046, 0.125		
dgenderOther	-0.029	0.064	-0.151, 0.097		
dstudyingworkingBoth	0.156	0.030	0.099, 0.214		
dstudyingworkingStudy	0.103	0.027	0.050, 0.156		
dstudyingworkingWork	0.098	0.031	0.037, 0.160		

Table A.4: Estimated coefficients from utility mapping models based on individual candidate predictors with days out of role, age and gender and education and employment using LMM (complementary log log transformation) *(continued)*

Parameter	Estimate	SE	CI (95%)	R2	Sigma
SOFAS cdaysoor model				**0.547**	**0.514**
SD (Intercept)	0.465	0.181	0.044, 0.673		
Intercept	-0.464	0.060	-0.583, -0.344		
SOFAS scaled	0.777	0.085	0.610, 0.944		
cdaysoor	-0.021	0.001	-0.023, -0.019		
SOFAS dage model				**0.523**	**0.506**
SD (Intercept)	0.488	0.242	0.019, 0.759		
Intercept	-0.462	0.091	-0.643, -0.286		
SOFAS scaled	0.962	0.089	0.788, 1.137		
dage	-0.026	0.004	-0.033, -0.019		
SOFAS dgender model				**0.290**	**0.655**
SD (Intercept)	0.328	0.200	0.013, 0.668		
Intercept	-0.995	0.061	-1.115, -0.876		
SOFAS scaled	0.967	0.090	0.789, 1.145		
dgenderMale	0.210	0.028	0.155, 0.264		
dgenderOther	-0.204	0.094	-0.388, -0.019		
SOFAS dstudyingworking model				**0.616**	**0.463**
SD (Intercept)	0.576	0.170	0.118, 0.745		
Intercept	-0.999	0.065	-1.127, -0.871		
SOFAS scaled	0.961	0.091	0.784, 1.140		
dstudyingworkingBoth	0.123	0.042	0.042, 0.206		
dstudyingworkingStudy	0.120	0.037	0.048, 0.195		
dstudyingworkingWork	0.042	0.044	-0.042, 0.130		
SOFAS cdaysoor 2 OLS CLL				**0.541**	**0.499**
SD (Intercept)	0.435	0.218	0.032, 0.710		
Intercept	0.029	0.092	-0.149, 0.210		
SOFAS scaled	0.751	0.087	0.577, 0.922		
cdaysoor	-0.021	0.001	-0.023, -0.019		
dage	-0.027	0.004	-0.034, -0.020		
SOFAS cdaysoor 3 OLS CLL				**0.564**	**0.458**
SD (Intercept)	0.427	0.261	0.007, 0.724		
Intercept	-0.524	0.062	-0.642, -0.403		
SOFAS scaled	0.760	0.087	0.587, 0.930		
cdaysoor	-0.021	0.001	-0.023, -0.019		

Table A.4: Estimated coefficients from utility mapping models based on individual candidate predictors with days out of role, age and gender and education and employment using LMM (complementary log log transformation) *(continued)*

Parameter	Estimate	SE	CI (95%)	R2	Sigma
dgenderMale	0.193	0.026	0.142, 0.243		
dgenderOther	-0.198	0.091	-0.373, -0.020		
SOFAS cdaysoor 4 OLS CLL				**0.622**	**0.455**
SD (Intercept)	0.513	0.179	0.075, 0.719		
Intercept	-0.490	0.066	-0.617, -0.363		
SOFAS scaled	0.771	0.085	0.605, 0.939		
cdaysoor	-0.021	0.001	-0.023, -0.019		
dstudyingworkingBoth	0.039	0.039	-0.039, 0.115		
dstudyingworkingStudy	0.084	0.035	0.014, 0.153		
dstudyingworkingWork	-0.053	0.042	-0.136, 0.030		
SOFAS dage 2 OLS CLL				**0.515**	**0.513**
SD (Intercept)	0.483	0.215	0.035, 0.756		
Intercept	-0.518	0.091	-0.695, -0.339		
SOFAS scaled	0.945	0.089	0.772, 1.118		
dage	-0.026	0.004	-0.034, -0.019		
dgenderMale	0.215	0.028	0.162, 0.270		
dgenderOther	-0.157	0.094	-0.341, 0.022		
SOFAS dage 3 OLS CLL				**0.442**	**0.557**
SD (Intercept)	0.443	0.218	0.038, 0.758		
Intercept	-0.466	0.108	-0.667, -0.254		
SOFAS scaled	0.933	0.089	0.762, 1.114		
dage	-0.028	0.005	-0.038, -0.019		
dstudyingworkingBoth	0.125	0.041	0.047, 0.207		
dstudyingworkingStudy	0.050	0.039	-0.021, 0.127		
dstudyingworkingWork	0.098	0.047	0.007, 0.186		
SOFAS dgender 2 OLS CLL				**0.500**	**0.537**
SD (Intercept)	0.481	0.209	0.031, 0.718		
Intercept	-1.087	0.065	-1.215, -0.960		
SOFAS scaled	0.940	0.091	0.763, 1.118		
dgenderMale	0.224	0.029	0.167, 0.278		
dgenderOther	-0.196	0.096	-0.385, -0.009		
dstudyingworkingBoth	0.159	0.041	0.078, 0.240		
dstudyingworkingStudy	0.148	0.037	0.076, 0.218		
dstudyingworkingWork	0.070	0.043	-0.013, 0.156		

Table A.4: Estimated coefficients from utility mapping models based on individual candidate predictors with days out of role, age and gender and education and employment using LMM (complementary log log transformation) *(continued)*

Parameter	Estimate	SE	CI (95%)	R2	Sigma

Note: The K10 and SOFAS parameters were first multiplied by 0.01.