MAPPING THE SF-36 TO EQ-5D-3L IN RANDOMIZED TRIAL: THE EMOCAR STUDY(MAY 2011 - APRIL 2016)

Elise Cabout[1], MSc; Anastasiia Kabeshova[1], Ph.D.; Yann Le Teurnier[2], Md; Robert Launois[1], Ph.D.
1: Network for Evaluation in Health Economics, REES France, Paris, France; 2: CHU Nantes

Background

Cost-utility analyses need a measure to summarize the **quality of life** in a **single index**.
The *utility-preference approach* of the EQ-5D-3L offers an interval measurement instrument resulting in an overall utility score while the *psychometric approach* of the SF-36 is based on a decomposed ordinal tool allowing to explore the various dimensions of the quality of life.
Mapping technique can be used to obtain an utility score from the SF-36.
No specific regression method has been recommended for implementing such mapping.

Objective

To compare the different regression methods for mapping the SF-36 into EQ-5D-3L based on French data.

Methods

EMOCAR Study design:

- French cohort of 904 patients with carotid endarterectomy followed from May 2011 to April 2016 with 3 visits (D0, D0+30, D0+120)
- EQ-5D-3L and SF-36 questionnaires collected for each patient at each visit

Steps:

1. Split the observations into two sets using random sampling : training and validation sets
2. Estimated regression in the training set
3. Implemented in the validation set to obtain utility score

Econometric methods:

- **OLS** : ordinary least square that relies on a quantitative variable to explain → EQ-5D-3L index score
- **Logistic** : multinomial logit model that estimates a qualitative variable → each EQ-5D-3L dimension

Specification **Two approaches** concerning the explanatory variables: **(1)** *summary score-based* of the 8 dimensions of the SF-36 as quantitative variables; **(2)** *all SF-36 items* as dummy-independent categorical variables

Model performance **(i)** Predicted mean utility score; **(ii)** Mean absolute error (MAE) and mean squared errors (MSE); **(iii)** Distribution of errors

Results

Parameters	OLS		Logit	
	Specification 1 (score)	Specification 2 (all items)	Specification 1 (score)	Specification 2 (all items)
Validation set (n=652)	**0,736 (0,246) -0,377/1**			
Adjusted R^2	0,61	0,678	0,343-0,579	0,668-0,744
Mean (SD) EQ-5D-3L utility score	0,735 (0,185)	0,733 (0,20)	0,789 (0,193)	0,741 (0,241)
Min/Max EQ-5D-3L utility score	0,17/1,046	-0,078/1,075	-0,052/1	-0,199/1
MAE	0,113	0,116	0,116	0,126
MSE	0,024	0,026	0,032	0,038
P-value	0,9411	0,8193	<0,0001	0,7074

Table 1 : Perfomance of OLS and Logit models in predicting EQ-5D-3L scores

Performance

- OLS models predicted a mean EQ-5D-3L utility score quite similar to the observed value
- Logit models predicted a mean EQ-5D-3L utility score higher than the observed value
- Logit models had a higher MSE value than OLS models
- Logit model with specification 1 predicted a EQ-5D-3L score significantly different from the EQ-5D-3L observed.

Distribution of errors

- The distribution of errors are quite similar between the four models
- OLS models estimated 66% of utility values with an absolute error > 0,05 but 44% with an absolute error > 0,1
- OLS model – specification 2 predicted almost 40% of utility values that are identical to the observed values
- Logit models estimated 60% of utility values with an absolute error > 0,05 and up to 47% with an absolute error > 0,1

Comparison by age

- OLS models predicted a better EQ-5D-3L score for the younger subgroups
- Logit models predicted a better EQ-5D-3L score for the older subgroups
- Specification 1 of the logit model predicted a significantly different EQ-5D-3L score for the older subgroups

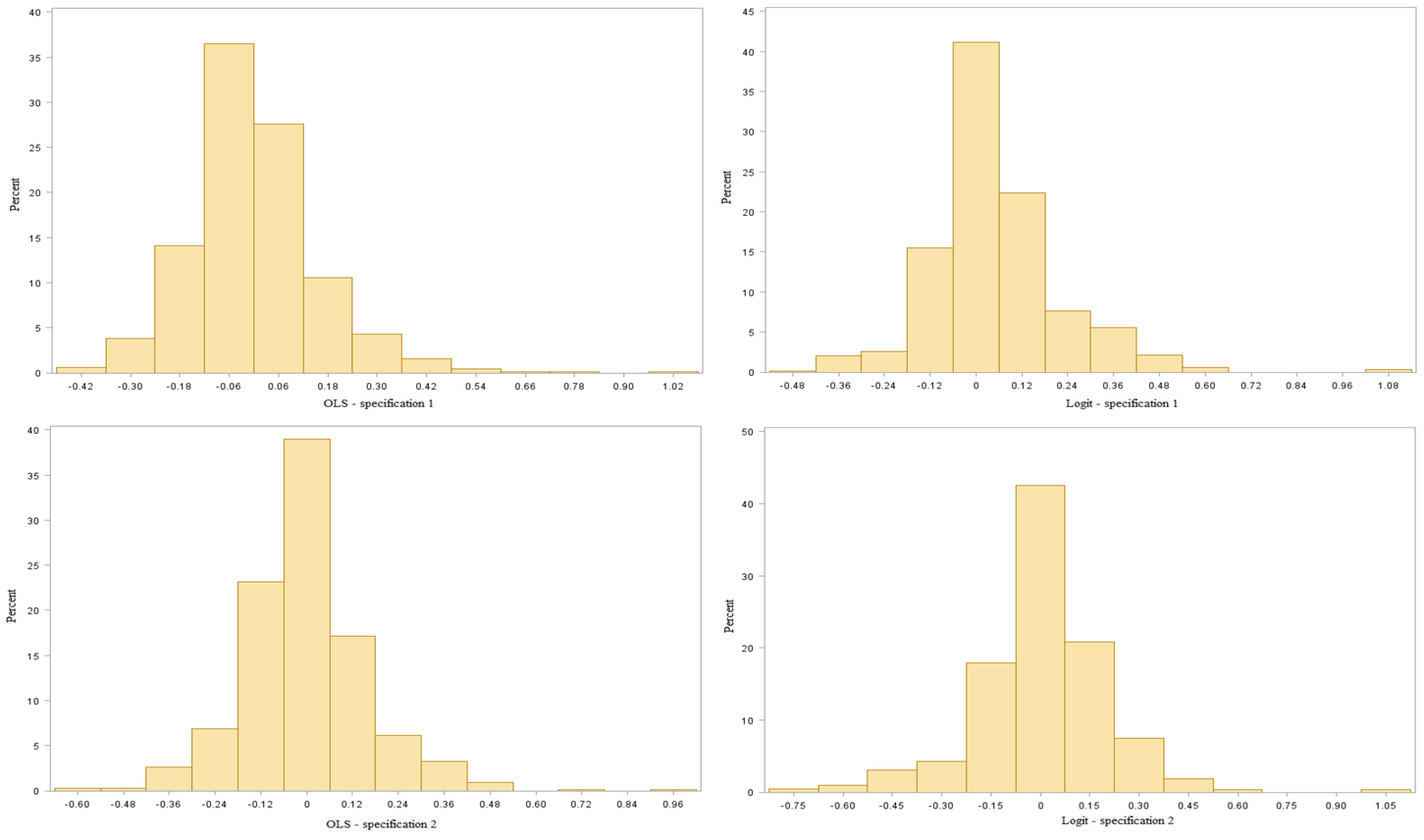

Figure 1 : Distribution of errors in EQ-5D-3L predicted scores

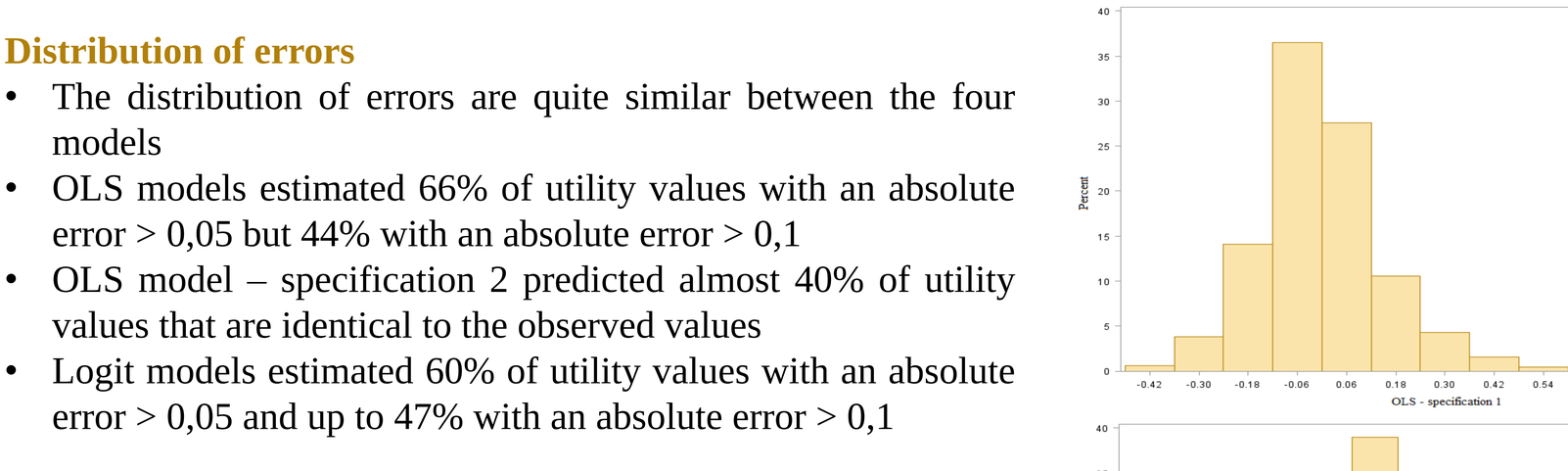

Figure 2 : Observed and predicted EQ-5D-3L scores : comparison to OLS and Logit models

Observed vs. predicted

OLS models predicted EQ-5D-3L score higher than 1 while logit models are limited to 1 (calculation of the utility score after regression with the French tariffs)

- OLS models predicted EQ-5D-3L scores closer to the observed score than logit models
- Low EQ-5D-3L scores are underestimated with all models

Conclusion

Our study suggests that models using OLS method produce the best results for mapping SF-36 into EQ-5D-3L utility scores. Specification 2, using all items of SF-36, has also better performance. Logit model with specification 1 give poor conclusions.
Low EQ-5D-3L utility scores are poorly predicted with all models but this phenomenon is always observed in the existing literature.

References

1. Chevalier J. Mesure de l'utilité attachée aux états de santé. Valorisation de l'index d'utilité EQ-5D et évolution de l'échelle actuelle en France [Thèse pour le Doctorat en Sciences Economiques]. Université Paris IX Dauphine; 2010.
2. Chuang L-H, Kind P. Converting the SF-12 into the EQ-5D: an empirical comparison of methodologies. Pharmacoeconomics. 2009;27(6):491-505.
3. Gray AM, Rivero-Arias O, Clarke PM. Estimating the Association between SF-12 Responses and EQ-5D Utility Values by Response Mapping. Medical Decision Making. 1 janv 2006;26(1):18-29.
4. Rowen D, Brazier J, Roberts J. Mapping SF-36 onto the EQ-5D index: how reliable is the relationship? Health Qual Life Outcomes. 31 mars 2009;7:27.

28, rue d'Assas – 75006 Paris France; Tel: +33 1 44 39 16 90; E-mail: **launois.reesfrance@wanadoo.fr**; Web: **www.rees-france.com**