ISPOR 7th Asia-Pacific Conference, Suntec Convention & Exhibition Center, Singapore
Patient-Reported Outcomes & Patient Preference Research Issues
Sunday, 4 September 2016, 3:45 PM - 4:45 PM @ Room 326

IP5: ESTIMATING COUNTRY-SPECIFIC EQ-5D-5L VALUE SETS USING A HYBRID REGRESSION MODEL: IS IT A GOOD IDEA FOR ASIA?

Moderator:
Nan Luo
National University of Singapore, Singapore
Panelists:
Mark Oppe, Juan M. Ramos-Goñi
EuroQol Research Foundation, Rotterdam, The Netherlands
Kim Rand-Hendriksen
University of Oslo, Norway

Conflict of interest & disclaimer

- The moderator and panelists are members of the EuroQol Group, a not-for-profit international research organization
- The views of the moderator and panelists expressed in the workshop do not necessarily reflect the views of the EuroQol Group

EQ-5D

Background of the Issue

www.euroqol.org

EQ-5D

EQ-5D-5L

- A new version of the widely used EQ-5D instrument
- A preference-based instrument for measurement of health-related quality of life (HRQoL) consisting on:
  - A descriptive system:
    - 5 Dimensions: mobility, self-care, usual activities, pain/discomfort, anxiety/depression
    - 5 Levels on each dimension: no, slight, moderate, severe, extreme
    - Visual analogue scale (VAS)
  - National value sets:
    - Lists of values for each of the possible health state, on a cardinal scale anchored by 0 (death) and 1 (full health)

www.euroqol.org

EQ-5D

Deriving utility values using EQ-5D-5L


MOBILITY
SELF-CARE
USUAL ACTIVITIES
PAIN / DISCOMFORT
ANXIETY / DEPRESSION

"12345"

Coding:
No = 1
Slight = 2
Moderate = 3
Severe = 4
Extreme = 5

Value Set

State	Value
11111	1.00
….	
12344	0.13
12345	0.10
12351	0.05
12352	0.07
12353	0.03
…	
55555	-0.60

www.euroqol.org

EQ-5D

The EQ-5D-5L Valuation Study

- Target population – general population
- Minimum sample size – 1,000 individuals
- Data collection mode – computer-assisted personal interviewing (CAPI)
- Valuation
  - Eliciting the value of 86 EQ-5D-5L health states using the time trade-off method (10 states per participant)
  - Eliciting the preferences for 196 DCE pairs of EQ-5D-5L health states (7 pairs per participants)
- Value set estimation
  - TTO data only
  - TTO and DCE data (the 'hybrid' model)

www.euroqol.org

EQ-5D

The 'Hybrid' model

Valuation and Modeling of EQ-5D-5L Health States Using a Hybrid Approach

Juan M. Ramos-Goñi, MSc,†‡ Jose L. Pinto-Prades, PhD,§ Mark Oppe, PhD,‡ Juan M. Cabasés, PhD,‖ Pedro Serrano-Aguilar, PhD,*† and Oliver Rivero-Arias, DPhil†¶*

Background: The EQ-5D instrument is the most widely used preference-based health-related quality of life questionnaire in cost-effectiveness analysis of health care technologies. Recently, a version called EQ-5D-5L with 5 levels on each dimension was developed. This manuscript explores the performance of a hybrid approach for the modeling of EQ-5D-5L valuation data.

Methods: Two elicitation techniques, the composite time trade-off, and discrete choice experiments, were applied to a sample of the Spanish population (n=1000) using a computer-based questionnaire. The sampling process consisted of 2 stages: stratified sampling of geographic area, followed by systematic sampling in each area. A hybrid regression model combining composite time trade-off and discrete choice data was used to estimate the potential value sets using main effects as starting point. The comparison between the models was performed using the criteria of logical consistency, goodness of fit, and parsimony.

Results: Twenty-seven participants from the 1000 were removed following the exclusion criteria. The best-fitted model included 2 significant interaction terms but resulted in marginal improvements in model fit compared to the main effects model. We therefore selected the model results with main effects as a potential value set for this methodological study, based on the parsimony criteria. The results showed that the main effects hybrid model was consistent, with a range of utility values between 1 and −0.224.

Conclusion: This paper shows the feasibility of using a hybrid approach to estimate a value set for EQ-5D-5L valuation data.

Key Words: utility theory, quality of life, maximum likelihood estimation, time trade-off, discrete choice experiment

(*Med Care* 2014;00: 000–000)

EQ-5D

Published country-specific EQ-5D-5L value sets

Country	TTO data only	TTO + DCE (the 'hybrid' model)
England		√*
Spain		√**
The Netherlands	√	
Uruguay	√	
Japan	√	
South Korea	√	

*The working OHE paper is online but the Journal paper is still under review

**Spanish team published a methodological test of the hybrid approach while the value set has not been published yet

www.euroqol.org

EQ-5D

The issue

- Shall we adopt the 'hybrid' model to estimate EQ-5D-5L value sets in Asia?
  - What is better as valuation technique: TTO or DCE?
  - How can TTO and DCE data be combined to predict EQ-5D-5L health states?
  - Is 'hybrid' a better approach than the TTO only approach?

www.euroqol.org

EQ-5D

The panelists

Mark Oppe

- **Development of the EQ-VT and the hybrid model**

Juan M. Ramos-Goñi

- **Hybrid models frameworks: the use of the "hyreg" command (Stata) and (R)**

Kim Rand-Hendriksen

- **A critique of hybrid models**

www.euroqol.org

Development of the EQ-VT and the hybrid model

Mark Oppe, PhD
EuroQol Research Foundation

Singapore, September 2016

www.euroqol.org

Past approaches to valuation

- Early valuation research on EQ-5D used VAS
- The UK MVH study first to use the TTO
- Became the 'default' protocol used in other countries
- Somewhat inconsistent approaches between countries limited comparability

www.euroqol.org

EQ-5D

Development of the EQ-VT

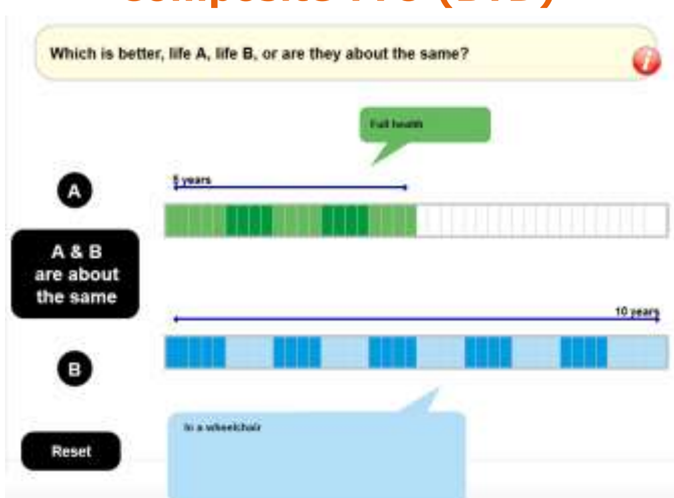
- Why a new valuation protocol?
  - Develop better valuation methods for valuing EQ-5D-5L
  - Take advantage of advances in computer-based methods
  - Provide a fully documented, evidence-based protocol to be used in all countries – ensure consistency

- 10 multinational pilot studies
  - Different modes of administration
  - Different types of TTO
  - Different secondary tasks (VAS, DCE, BWS)

www.euroqol.org

EQ-5D

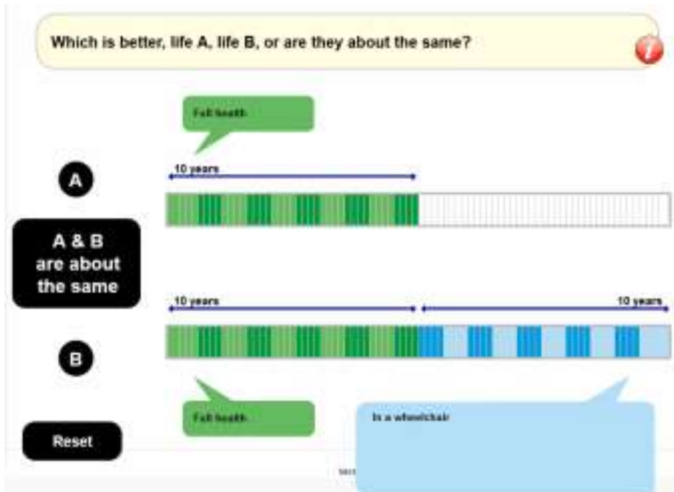
Composite TTO (BTD)



www.euroqol.org

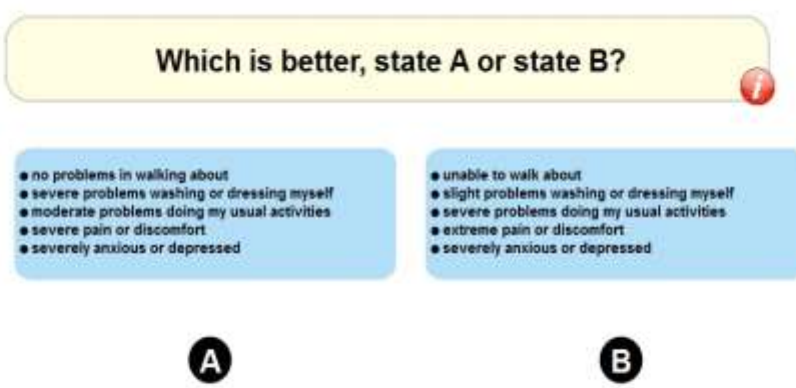
EQ-5D

Composite TTO (WTD)



EQ-5D

DCE paired comparisons



www.euroqol.org

Experimental Design

DESIGN SPECIFICATIONS	cTTO		DCE	
N respondents	1000		1000	
N blocks	10		28	
N states/pairs	80	+ 6 fixed	186	+ 10 fixed very mild
N states/pairs per resp	10		7	
N obs per state/pair	100	(for the set of 80 states)	36	
Optimisation Algorithm	Monte Carlo simulation		Bayesian efficient design	

www.euroqol.org

Tasks included in EQ-VT version 2.0

Introduction

- Self reported health on the EQ-5D-5L descriptive system
- Self reported health on the EQ-VAS
- Background questions

Composite Time Trade-Off

- Instructions and example of TTO task, 3 practice states
- TTO valuation of 10 EQ-5D-5L states
- TTO debriefing/structured feedback
- TTO feedback module

Discrete Choice

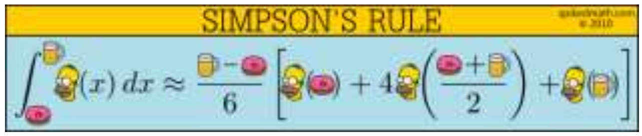
- Instructions of DC task
- DC valuation of 7 pairs of EQ-5D-5L states
- DC debriefing/structured feedback

Cyclic quality control process

www.euroqol.org

Modelling TTO and DCE

- Individuals have a utility function which determines their preferences over health states



- TTO & DCE methods both try to measure the same utility function
- TTO & DCE each have their own weaknesses
  - e.g. scale compatibility (BTD vs WTD) for C-TTO
  - e.g. no anchors for use in QALY calculations for DCE
- Which method should we choose?

www.euroqol.org

TTO, DCE or both?

- TTO: trade-off between quality of life and length of life
  - How many years are you willing to give up to avoid being in impaired health?
- DCE: trade-off between quality of life and quality of life
  - Which health state is better?
- Both questions provide relevant information
- View TTO and DCE as complementary sources of information instead of competing

Include both types of information in a single hybrid model

www.euroqol.org

EQ-5D

Log likelihood of basic hybrid model (OLS & clogit)

$$lnL = -\frac{1}{2} * \sum_{j\in C}\left\{\ln(2\pi\sigma^2) + \left(\frac{y_j - x\beta}{\sigma}\right)^2\right\}$$

$$+ \sum_{j\in D}\left\{ln\left(\frac{1}{1+e^{(-x\beta')}}\right) * y_j + ln\left(\frac{e^{(-x\beta')}}{1+e^{(-x\beta')}}\right) * (1-y_j)\right\}$$

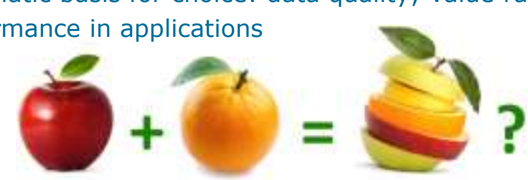
proportional rescaling parameter θ, such that
$\beta' = \beta * \theta$

www.euroqol.org

EQ-5D

Apples & Oranges or a Fruit Salad?

- Hybrid:
  - Uses all available information
  - Hybrid estimates are typically between estimates of TTO alone and estimates of DCE
  - DCE can help mitigate issues present in TTO and v.v.
- Since the “true” utilities are not known, ultimately the choice is a normative one:
  - Which (imperfect) utility theory?
  - Which (imperfect) data collection technique?
- Pragmatic basis for choice: data quality; value range; performance in applications



www.euroqol.org

Hybrid models frameworks The use the “hyreg” command (Stata)

Juan M. Ramos-Goñi, MSc
EuroQol Research Foundation

Singapore, September 2016

www.euroqol.org

History of modelling approaches

- EQ-5D-5L valuation studies were first launched in 2012, with Spain, UK, The Netherlands, Canada and China being the first countries to test EQ-VT
- The first test of the hybrid model using 5L valuation data was done using Spanish data
- The test indicates that the approach is feasible, but having some limitations

www.euroqol.org

Programing the hybrid model

- First implementation was made in R by Ben van Hout (not user friendly code)
- In parallel a Stata implementation was made by Juan M. Ramos-Goñi (not user friendly code)
- Improvements were started in parallel:
  - Random coefficients for TTO data (Ben van Hout)
  - Inclusion of interval data for TTO data (Benjamin Craig and Juan M. Ramos-Goñi)
  - Censoring TTO observations (Ben van Hout)
  - Including the mix with a conditional probit instead of logit (Benjamin Craig and Juan M. Ramos-Goñi)
- At the end it was decided to integrate as much features as possible in "user friendly commands" for Stata and R.

The Stata "hyreg" command

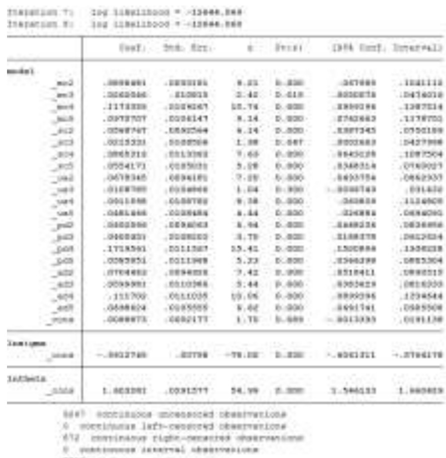
- **Syntax**

hyreg *depvar1* [*depvar2*] [*indepvars*] [if] [in] ,
datatype(*varname*)
[interval
contdist(*normal* | *logistic*)
dichdist(*normal* | *logistic*)
ll(#) ul(#)
hetcont(*varlist*) hetdich(*varlist*)
noconstant
vce(oim | opg | robust | cluster *varname*) *maximize options*]

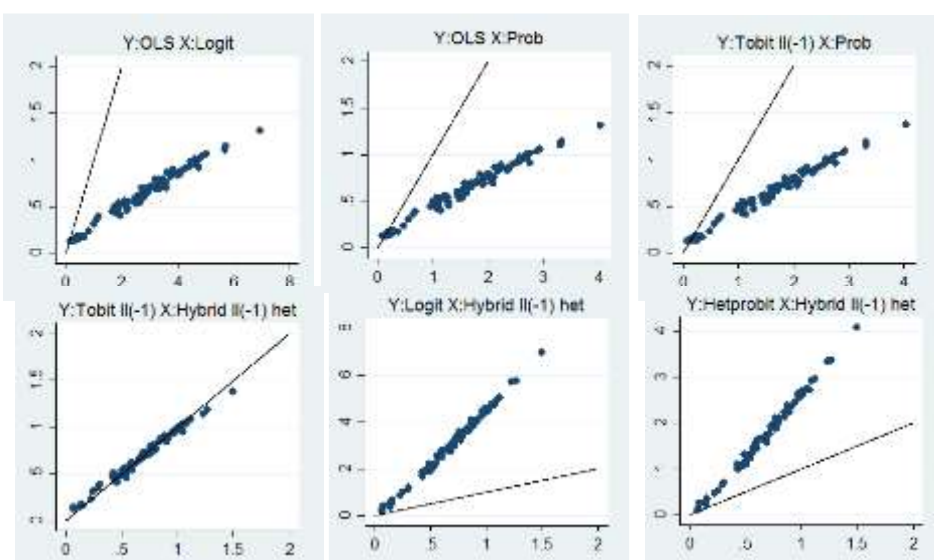
"hyreg output"

$$lnL = -\frac{1}{2} * \sum_{j \in C} \left\{ \ln(2\pi\sigma^2) + \left(\frac{y_j - x\beta}{\sigma}\right)^2 \right\}$$

$$+ \sum_{j \in D} \left\{ ln \left(\frac{1}{1 + e^{(-x\beta * \theta)}} \right) * y_j + ln \left(\frac{e^{(-x\beta * \theta)}}{1 + e^{(-x\beta * \theta)}} \right) * (1 - y_j) \right\}$$



Are the hybrid assumptions sensible?



EQ-5D

Summary

- The hybrid approach is feasible
- DCE predictions and TTO predictions are highly correlated
- High concordance between TTO models and hybrid models
- High correlation between DCE models and hybrid models
- The estimated coefficient from hybrid model are more precise than (<S.t error) than the ones from DCE or TTO models

- **Why shouldn't it be done?**

EQ-5D

A critique of hybrid models

Kim Rand-Hendriksen, PhD
University of Oslo

Singapore, September 2016

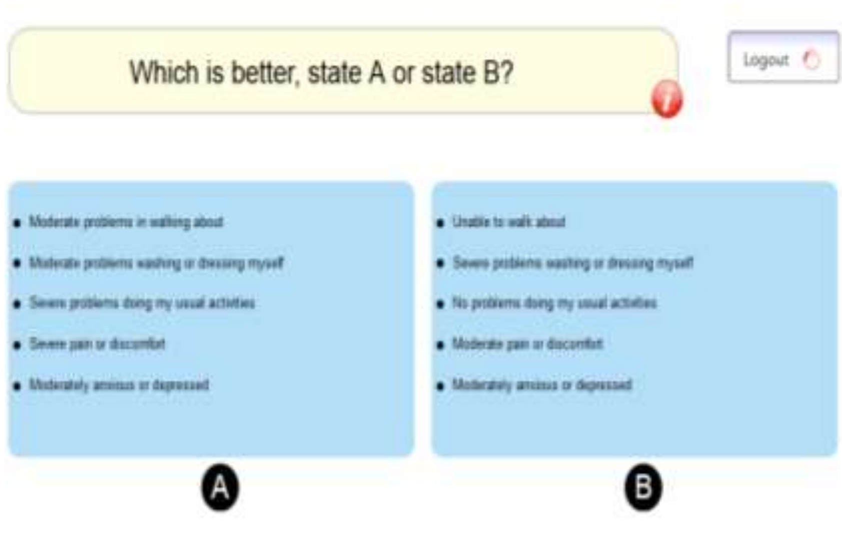
Battle plan

1. Conceptual issues
   - The relationship between utilities and DCE
   - What we know, and what we don't know
   - Lack of obvious counterfactual
2. Practical issues
   - Shared constant term/intercept between DCE and TTO
   - "Flat" areas when combining two data types with different maxima
   - Weights
   - Problems with the handling of differences between the TTO and DCE functions
3. Conclusion

The relationship between utilities and DCE (RUT)

- DCE/TTO hybrid models rest on the assumption that DCE and TTO are equally valid, or that it is unknown which is more valid
- TTO measures strength of preference directly, on and individual level. Population aggregates of this will therefore take into account variation in strength of preference.
- DCE, when applied to a population, as opposed to repeated measures of an individual, does not (necessarily) take into account variation in individual strength of preference.
- Choices for health states could reflect differences in "taste" for health
  - Consider a choice between chocolate and caramel ice cream. If it is observed that 60% prefer chocolate, we cannot directly infer that chocolate has a higher value than caramel, since the minority preferring caramel could display a substantially greater willingness to pay than the proponents of chocolate. TTO catches this difference (at least in theory), while DCE does not.
  - This is a general critique of DCEs for health state valuation, and does not apply only to hybrids.

www.euroqol.org



Lack of obvious counterfactual

- With mean-based modeling of TTO, predictions can be directly compared to observed means. This allows leave-out cross-validation with a true counterfactual for comparison.
- For "pure" DCE models, predictions can be compared to observed choice probabilities.
- With hybrid models, performance cannot be easily measured by these kinds of comparison.
- For more complex hybrids (i.e. predicting intervals, handling censoring, heteroscedastic standard deviations, models for the link between TTO and DCE…), determining model validity becomes very tricky.
- Likelihood-based comparison remains possible, but are uninformative as to the validity of the assumptions behind the likelihood function.

Practical issues
Shared intercept/constant term

1. hyreg value _mo2-_ad5, datatype(_method) nocons

2. hyreg value _mo2-_ad5, datatype(_method)

- Code 1 fits a model with 20 parameters to both TTO and DCE data, with no constant term/intercept
- Code 2 fits the same model, adding a constant term. All parameters, including the constant term shared, meaning that they are fitted to both TTO and DCE observations.
- Unfortunately, the constant term does not mean the same for the two kinds of data, and the sign of the constant for DCE is arbitrary. I will illustrate with an example.

Constant problems

- Data from the DHS (demographic and health surveys) run by USAID
- Age, sex, height, and weight for approx. 3000 children aged 0-5 years
- Linear regression model to predict height based on age (dummies for 1, 2, 3, and 4 years) and sex (dummy for girl)

h = INTERCEPT + S + A1 + A2 + A3 + A4

Parameters:

	Estimate	Std..Error
SIGMA	10.154619	0.1839315
INTERCEPT	64.086948	0.6976458
S	-1.389705	0.5211695
A1	14.798526	0.9016843
A2	8.354346	0.8481761
A3	8.964731	0.8017649
A4	7.344159	0.7644284

Intercept is interpretable as estimated average height for boys at <1 years.

Constant problems cont'd

- Generate 10 000 "DCE"s, by random sampling (with replacement)
- Target variable 1 if left child is tallest. Ties removed. Conditional logit model:

Parameters:

	Estimate	Std..Error
INTERCEPT	-0.003116987	0.01568021
S	-0.097730655	0.02234994
A1	1.004000093	0.04862737
A2	0.629414872	0.03525029
A3	0.573683424	0.03238180
A4	0.528308318	0.03269077

- Here, the intercept is the average right/left bias, which is negligible due to the random sampling.

Now as a hybrid

- OLS for continuous, conditional logit for generated "DCE"

Parameters:

	Estimate	Std..Error
SIGMA	25.976202	0.5459067
THETA	38.619536	0.8219985
INTERCEPT	7.743260	0.5916663
S	2.191158	0.7061509
A1	51.815555	1.1331342
A2	19.021508	1.1407489
A3	18.434260	1.0677370
A4	15.905160	1.0498862

- Now, the intercept has no direct interpretation.
- The model also fits quite badly:

	Alone	Hybrid
Loglik for continuous:	-5695,9	-7126,4
Loglik for "DCE" :	-4176,7	-4340,6

EQ-5D

Hybrid with separate intercepts

- INTERCEPT for continuous, INTERCEPT_DCE for “DCE”

Parameters:

	Estimate	Std..Error
SIGMA	10.15649782	0.1839824
INTERCEPT_TTO	64.11278521	0.5847958
INTERCEPT_DCE	-0.04476864	0.2262171
THETA	14.42715917	0.3858576
S	-1.40353266	0.2748719
A1	14.57344660	0.5649097
A2	8.89915133	0.4536258
A3	8.44354327	0.4271501
A4	7.54903476	0.4180064

	Alone	Hybrid1	Hybrid2
Loglik for continuous:	-5695,9	-7126,4	-5695,3
Loglik for “DCE” :	-4176,7	-4340,6	-4176,8

EQ-5D

Right/left bias

"Original"	Estimate	Std. Error	Reversed	Estimate	Std. Error
(Intercept)	0.21164586	0.0266645	(Intercept)	-0.21164586	0.0266645
incr_mo2_diff	0.44230775	0.0526517	incr_mo2_diff	0.44230775	0.0526517
incr_mo3_diff	-0.00271160	0.0602873	incr_mo3_diff	-0.00271160	0.0602873
incr_mo4_diff	0.77762394	0.0592446	incr_mo4_diff	0.77762394	0.0592446
incr_mo5_diff	1.16562366	0.0653388	incr_mo5_diff	1.16562366	0.0653388
incr_sc2_diff	0.35921265	0.0576461	incr_sc2_diff	0.35921265	0.0576461
incr_sc3_diff	-0.00847601	0.0625683	incr_sc3_diff	-0.00847601	0.0625683
incr_sc4_diff	0.67629896	0.0661428	incr_sc4_diff	0.67629896	0.0661428
incr_sc5_diff	0.39924359	0.0615881	incr_sc5_diff	0.39924359	0.0615881
incr_ua2_diff	0.27154495	0.0549798	incr_ua2_diff	0.27154495	0.0549798
incr_ua3_diff	0.04449129	0.0615065	incr_ua3_diff	0.04449129	0.0615065
incr_ua4_diff	0.53745959	0.0619102	incr_ua4_diff	0.53745959	0.0619102
incr_ua5_diff	0.57164557	0.0636063	incr_ua5_diff	0.57164557	0.0636063
incr_pd2_diff	0.25449076	0.0580883	incr_pd2_diff	0.25449076	0.0580883
incr_pd3_diff	0.03962980	0.0633029	incr_pd3_diff	0.03962980	0.0633029
incr_pd4_diff	0.80788425	0.0642941	incr_pd4_diff	0.80788425	0.0642941
incr_pd5_diff	0.30395110	0.0638620	incr_pd5_diff	0.30395110	0.0638620
incr_ad2_diff	0.27174584	0.0603230	incr_ad2_diff	0.27174584	0.0603230
incr_ad3_diff	0.19483877	0.0615435	incr_ad3_diff	0.19483877	0.0615435
incr_ad4_diff	0.88367616	0.0655376	incr_ad4_diff	0.88367616	0.0655376
incr_ad5_diff	0.52791784	0.0654440	incr_ad5_diff	0.52791784	0.0654440

www.euroqol.org

EQ-5D

Right/left bias with hybrids and shared intercept

- Sign of DCE influences joint intercept, and model fit

A - B

	Estimate	Std..Error
THETA	47.51251	1.0218045
INTERCEPT	19.70110	0.6816674
S	2.68336	0.7147274
A1	43.69928	1.1060360
A2	17.45344	1.1748502
A3	17.21160	1.1099331
A4	13.35570	1.0740309
SIGMA	21.36575	0.4746664

DCE	CONTINUOUS
-4946.764	-6828.629

B - A

	Estimate	Std..Error
THETA	52.772017	1.1346890
INTERCEPT	2.155600	0.7820390
S	4.221512	0.8438539
A1	54.971610	1.3074188
A2	19.586300	1.3842987
A3	19.516540	1.3060186
A4	14.812707	1.2673559
SIGMA	27.772534	0.6207201

DCE	CONTINUOUS
-4985.903	-7228.311

www.euroqol.org

EQ-5D

Removing the intercept

- No impact from arbitrary choice of A and B, but (in some cases) bad model fit

Parameters:

	Estimate	Std..Error
THETA	52.669965	1.1332921
S	4.333265	0.8536047
A1	56.244804	1.2443716
A2	19.891977	1.3961106
A3	19.795430	1.3172189
A4	15.023894	1.2801413
SIGMA	28.611836	0.5587462

DCE	CONTINUOUS
-4944.324	-7273.685

www.euroqol.org

EQ-5D

Separate intercepts

- With separate intercepts, model fit is improved

A - B

	Estimate	Std..Error
THETA	18.309480	0.4926709
INTERCEPT_DCE	3.849551	0.2821264
S	-1.471070	0.3105478
A1	13.487405	0.5971305
A2	9.250974	0.5088136
A3	8.999900	0.4824521
A4	7.280235	0.4662980
SIGMA	10.162608	0.1841837
INTERCEPT	64.634485	0.5934544

DCE	CONTINUOUS	sum
-4735.784	-5696.184	-10431.968

B - A

	Estimate	Std..Error
THETA	18.309480	0.4926709
INTERCEPT_DCE	-3.849551	0.2821264
S	-1.471070	0.3105478
A1	13.487405	0.5971305
A2	9.250974	0.5088136
A3	8.999900	0.4824521
A4	7.280235	0.4662980
SIGMA	10.162608	0.1841837
INTERCEPT	64.634485	0.5934544

DCE	CONTINUOUS	sum
-4735.784	-5696.184	-10431.968

www.euroqol.org

EQ-5D

"Flat" areas when combining two data types with different maxima

- Since the hybrid maximizes the sum of TTO and DCE log-likelihood, and the two are often different, parameter changes that improve TTO fit can often reduce DCE fit, and vice versa
- This results in ranges of parameter values for which the sum of log-likelihoods changes very little – "flat" areas. Such flat areas make the model unstable, in that quite small changes can result in relatively large changes in the resulting fitted model.

www.euroqol.org

Weights

- Maximum likelihood is a sum of likelihoods for each prediction over each observation. Increasing the number of observations increases the maximum likelihood.
- When maximizing the sum of two different sums of likelihoods, the relative weight of one type of data over the other will be a function of how many observations are present of each.
- The fitting function is not sensitive to the absolute magnitude of likelihoods, but to the magnitude of the change from small changes in the parameters.
- If a change of one unit for a parameter results in a positive change of 1.1 for the sum likelihood for TTO, and -1 for DCE, a >10% increase in the number of DCE observations will reverse the direction of change to the fitted model.

Conclusions

- TTO and DCE are different
- We know more about the problems with TTO than with DCE
- We might not be combining two measures of the same, but two measures of different things
- Various practical issues that have not been adequately addressed yet
- Are hybrids interesting?
  Yes.
- Are we at the point where we should replace TTO-only models with TTO/DCE-hybrids?
  My personal opinion is that this is premature.

EQ-5D

Brief Responses from Juan M. Ramos-Goñi & Mark Oppe

www.euroqol.org

EQ-5D

Open Discussion

www.euroqol.org