



# INCREASING EQ-5D FROM 3 TO 5 LEVELS: IMPLICATIONS FOR USERS - DOES "NEW" MEAN "BETTER"?

*EuroQol Research Foundation  
Symposium  
ISPOR, Vienna, 2016*

*Chair: Andrew Lloyd*

[www.euroqol.org](http://www.euroqol.org)



## Introductions

**Professor Paul Kind**  
University of Leeds; HSE St Petersburg

**Dr. Joachim Marti**  
Imperial College, London

**Dr. Bas Janssen**  
EuroQol Group

**Professor Allan Wailoo**  
NICE Decision Support Unit,  
University of Sheffield

[www.euroqol.org](http://www.euroqol.org)



# Role of EQ-5D

- EQ-5D measures quality of life
- Estimation of QALYs for economic evaluation
- EQ-5D-5L developed – to address perceived limitations of EQ-5D-3L
  - new response choices
  - new preference weights

www.euroqol.org



# Moving from 3 levels to 5...

## Mobility

- I have no problems in walking about
- I have some problems in walking about
- I am confined to bed

## Self-Care

- I have no problems with self-care
- I have some problems washing or dressing myself
- I am unable to wash or dress myself

## Usual Activities (e.g. work, study, housework, family or leisure activities)

- I have no problems with performing my usual activities
- I have some problems with performing my usual activities
- I am unable to perform my usual activities

## Pain/Discomfort

- I have no pain or discomfort
- I have moderate pain or discomfort
- I have extreme pain or discomfort

## Anxiety/Depression

- I am not anxious or depressed
- I am moderately anxious or depressed
- I am extremely anxious or depressed

## Issues with 3L?

- Sensitivity
  - Insensitive to mild health problems
  - ‘Confined to bed’
- Valuation
  - UK values 25 years old
  - Worse than dead methods
  - Distribution of data - trimodal
  - Regression models

www.euroqol.org



## Developing 5L

- Large amount of work conducted to move to 5L
  - Qualitative research/ focus groups UK, France, Spain, China
- 8 international pilot studies
  - Worse than dead valuation – different formats
  - Role of DCE methods to support valuation
  - Computer administration
  - Protocol/ quality assurance/ interviewer issues
  - Modelling approaches
- Subsequent programme of work to iron out new problems!

[www.euroqol.org](http://www.euroqol.org)



## EQ-5D-5L

- Resulting EQ-5D-5L weights are now different
  - How do they differ?
  - Does this matter?
  - Would the new weights change decisions? Could it change the views of what is cost-effective?

[www.euroqol.org](http://www.euroqol.org)

# The impact of changes in EQ-5D-3L value sets

## Paul Kind

Academic Unit of Health Economics,  
Institute of Health Science,  
University of Leeds  
Leeds, UK

Centre for Health Economics,  
Management and Policy,  
HSE University  
Saint Petersburg, Russia



## Joachim Marti

Institute of Global Health Innovation,  
Imperial College  
London, UK

Imperial College  
London



Institute of Global Health Innovation

## EQ-5D-3L

### DESCRIPTION

- EQ-5D is a generic measure of health status defined in terms of 5 dimensions
- In its original format each dimension provides 3 levels of response
- Taken together these form a classification of 243 possible health states

### VALUATION

- A summated index for each EQ-5D health state can be computed
- Weights for dimensions / levels have been estimated using several methods
- For economic evaluation submitted to NICE the preferred method is Time Trade-Off (TTO)

## NICE social preferences for EQ-5D-3L

- Originate in the 1993 Measurement and Valuation of Health (MVH) study
  - 3,000+ respondents
  - Representative sample of (the then) UK
- Protocol included multiple methods
  - Ranking
  - VAS rating
  - Time Trade-Off (TTO)
- Subset of 43 EQ-5D-3L health states selected for study
- Each respondent directly evaluated 13 health states
- Estimation model constructed to interpolate values for unobserved states
- Value decrements computed for each level / dimension

## Sources of variation in value sets

### Protocol design

- Choice of method(s)
- Health state selection
- Size of choice set
- Perceptual setting
- Mode of administration

### Data analysis

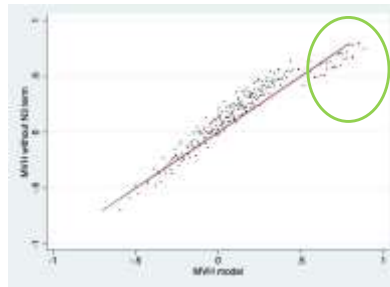
- Level of measurement
  - Form of statistical analysis
- Level of aggregation
  - Measure of central tendency
- Dummy structure
  - Interaction / main effects
  - Constant

# This study

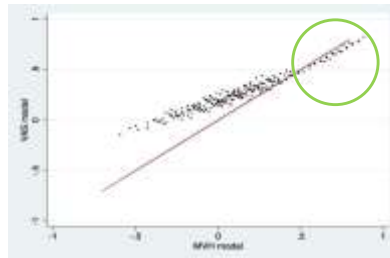
- Focuses on changes to the VALUE SET used to form index based on EQ-5D-3L health states
- Emphasis on extent of change associated with
  - structural issues in estimation model
  - deviation from preferred (TTO) method
- Takes the 1993 MVH value set as the reference comparator
- Basic question

*Does use of an alternate value set yield changes in marginal benefit that would lead to “switches” in ICER position related to given threshold?*

- Original MVH model based on 5\*2 dummy variables plus N3 and constant
- Re-estimated without N3



- Standard MVH value set based on Time Trade-Off
- Re-estimated using VAS ratings



# Marginal differences in health state values

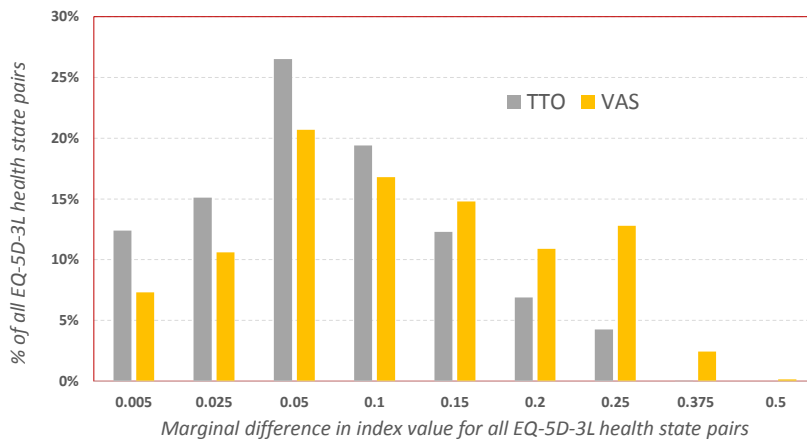
	11111	21111	12111	11211	11121	11112	.....	33333
11111	-							
21111		-						
12111			-					
11211				-				
11121					-			
11112						-		
.....							-	
33333								-

Identify 29,403 possible pairs of EQ-5D-3L health states  
 1

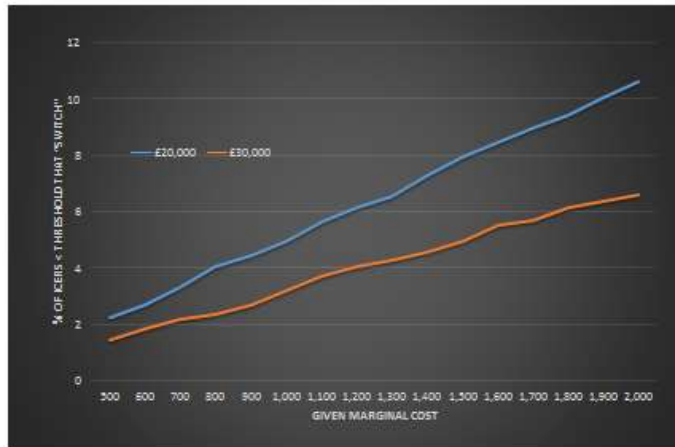
- Apply selected value set
  - Compute difference in EQ-5D<sub>index</sub> scores
- 2

- Plot the distribution of differences
  - Quantify extent of “switching” and magnitude of changes in Δ QALY
- 3

## Distribution of marginal EQ-5D<sub>index</sub> scores



## Percentage of “switching” EQ-5D health state pairs for a given marginal cost / fixed threshold



## Temporal stability ?

Do social preferences for health remain invariant with time ?  
More than 20 years on, do 1993 UK preference weights retain legitimacy for social decision-making?

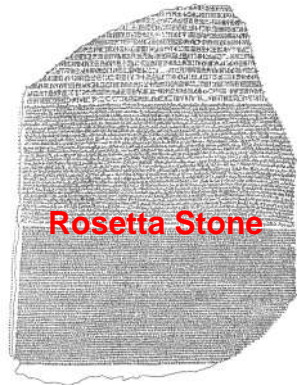


## Summary

- Potential value set differences can be linked with multiple causal factors
- Parameter uncertainty is less of an issue than structural model design
- These are “trumped” by the choice of valuation method
- These results are based on hypothetical changes to the value of incremental benefits – NOT observed



## From EQ-5D-3L to EQ-5D-5L



*Speaker:* **M.F. Bas Janssen, PhD**  
Senior Researcher, EuroQol Research Foundation,  
Rotterdam, The Netherlands

[www.euroqol.org](http://www.euroqol.org)



# EQ-5D-3L versus EQ-5D-5L

**Mobility**

I have no problems in walking about

I have some problems in walking about

I am confined to bed

**Self-Care**

I have no problems with self-care

I have some problems washing or dressing myself

I am unable to wash or dress myself

**Usual Activities** (e.g. work, study, housework, family or leisure activities)

I have no problems with performing my usual activities

I have some problems with performing my usual activities

I am unable to perform my usual activities

**Pain/Discomfort**

I have no pain or discomfort

I have moderate pain or discomfort

I have extreme pain or discomfort

**Anxiety/Depression**

I am not anxious or depressed

I am moderately anxious or depressed

I am extremely anxious or depressed

**MOBILITY**

I have no problems in walking about

I have slight problems in walking about

I have moderate problems in walking about

I have severe problems in walking about

I am unable to walk about

**SELF-CARE**

I have no problems washing or dressing myself

I have slight problems washing or dressing myself

I have moderate problems washing or dressing myself

I have severe problems washing or dressing myself

I am unable to wash or dress myself

**USUAL ACTIVITIES** (e.g. work, study, housework, family or leisure activities)

I have no problems doing my usual activities

I have slight problems doing my usual activities

I have moderate problems doing my usual activities

I have severe problems doing my usual activities

I am unable to do my usual activities

**PAIN / DISCOMFORT**

I have no pain or discomfort

I have slight pain or discomfort

I have moderate pain or discomfort

I have severe pain or discomfort

I have extreme pain or discomfort

**ANXIETY / DEPRESSION**

I am not anxious or depressed

I am slightly anxious or depressed

I am moderately anxious or depressed

I am severely anxious or depressed

I am extremely anxious or depressed

The EuroQol Group. EuroQol-a new facility for the measurement of health-related quality of life. Health Policy, 1990, 16(3):199-208.

Herdman et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Quality of Life Research, 2011, 20(10):1727-1736.

www.euroqol.org



## Methods

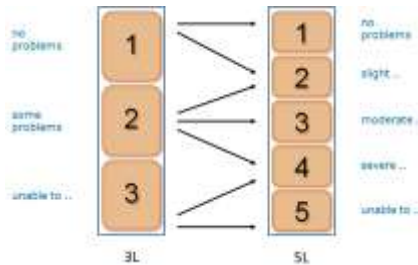
- I. Equivalence of descriptive distributions of 3L vs 5L in various population groups
- II. Equivalence of 3L vs 5L value sets
- III. Combined distributional equivalence of description+valuation in 3L vs 5L
- IV. Final test of 3L vs 5L:  
Discriminative power, using empirical values (utilities) in known groups comparison**

www.euroqol.org

# I. Projecting 3L on 5L

*description*

## Redistribution from 3L to 5L



Resulting in: reduced ceiling effects;  
 increased discriminatory power; improved  
 convergent validity; similar construct  
 validity

**Equivalent and enhanced across**

# II. Comparing 3L vs 5L

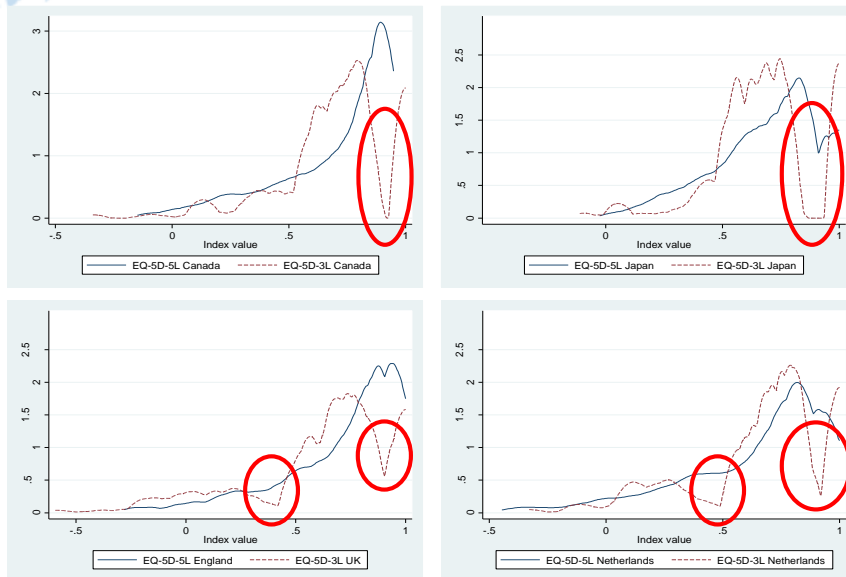
*valuation*

Models characteristics	Canada		China		England/UK		Japan		Netherlands		South Korea		Spain	
	3L	5L	3L	5L	3L	5L	3L	5L	3L	5L	3L	5L	3L	5L
Intercept	0.111	0.051	0.039	-	0.081	-	0.152	0.061	0.071	0.047	0.050	0.096	0.024	-
Interaction parameters														
N3			0.022		0.269				0.234		0.050		0.291	
Num45sq		0.0085												
Slope					0.9675									
C4											0.078			
Highest value	1	0.949	1	1	1	1	1	1	1	1	1	1	1	1
Second highest value	0.844	0.929	0.887	0.955	0.883	0.951	0.804	0.895	0.897	0.918	0.913	0.883	0.914	0.956
Lowest value	-0.340	-0.148	-0.149	-0.386	-0.594	-0.281	-0.111	-0.025	-0.329	-0.446	-0.171	-0.066	-0.654	-0.416
Upper gap	0.16	0.02	0.11	0.05	0.12	0.05	0.20	0.11	0.10	0.08	0.09	0.12	0.09	0.04
Range	1.34	1.10	1.15	1.39	1.59	1.28	1.11	1.03	1.33	1.45	1.17	1.07	1.65	1.42

**Different weighting structure, country variance** ↑



### III. Combined equivalence 3L vs 5L



5L enhanced, fewer discontinuities

[www.euroqol.org](http://www.euroqol.org)



### IV. 3L vs 5L value sets *discriminative power*

- Which instrument differentiates best between healthy / sick, and between mildly / severely affected patients?
- Using a dataset containing 3L *and* 5L responses from 3919 people, ...
- ... and using value sets from countries where both a 3L *and* 5L value set is available (N=7) ...
- ... provides seven end-to-end comparisons of 3L vs 5L utilities
- Primary endpoint: the F-Statistic (Variation Between Sample Means / Variation Within the Samples)

[www.euroqol.org](http://www.euroqol.org)

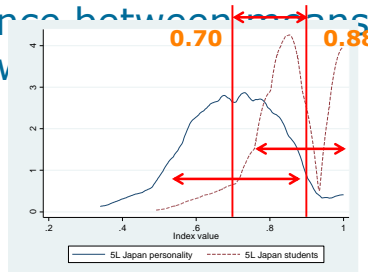


## Discriminative power explained..

- F statistic expresses ability to discriminate between

		Canada	
		F stat.	F ratio
Student vs COPD	3L	397	0.89
	5L	352	

- Based on difference between means and SDs of the two



www.euroqol.org



## Healthy students vs patients

		Canada		China		England/UK		Japan		Netherlands		South Korea		Spain	
		F stat.	F ratio	F stat.	F ratio	F stat.	F ratio	F stat.	F ratio	F stat.	F ratio	F stat.	F ratio	F stat.	F ratio
COPD	3L	397	0.89	519	0.88	358	0.91	449	1.00	261	1.11	521	0.77	437	0.84
	5L	352		458		324		448		289		400		365	
Diabetes	3L	48	0.94	71	0.87	54	0.69	45	0.89	29	0.91	82	0.51	82	0.52
	5L	45		61		37		40		27		42			
Liver disease	3L	36	0.53	47	0.64	33	0.40	29	0.60	22	0.39	57	0.21	53	0.30
	5L	19		30		13		17		9		12		16	
RA	3L	398	0.82	528	0.83	351	0.87	492	0.96	234	1.18	514	0.82	418	0.87
	5L	328		437		304		470		276		419		366	
CVD	3L	305	1.10	416	1.03	301	1.02	311	1.31	223	1.20	424	0.85	382	0.88
	5L	334		427		307		409		268		361		336	
Stroke	3L	565	1.00	806	0.84	561	0.96	651	1.14	527	0.94	555	1.04	575	1.00
	5L	563		680		536		741		497		577		573	
Depression	3L	223	0.74	234	0.82	181	0.96	202	1.11	185	0.95	231	0.84	198	0.98
	5L	165		192		175		224		176		194		194	
Personality disorder	3L	400	0.76	381	0.99	311	1.06	314	1.31	319	1.08	388	0.91	342	0.95
	5L	302		377		328		410		344		355		325	

\*Green cells indicate a significant F ratio in favor of 5L, red cells in favor of 3L (95% CI, 3000 bootstrap samples)

\*All comparisons were significant at P < 0.05

Overall 3L discriminates even better compared to 5L. This is true for most countries

www.euroqol.org



## Mildly vs moderately or severely diseased patients

		Canada		China		England/UK		Japan		Netherlands		South Korea		Spain	
		F stat.	F ratio	F stat.	F ratio	F stat.	F ratio	F stat.	F ratio	F stat.	F ratio	F stat.	F ratio	F stat.	F ratio
COPD vs diabetes	3L	105		114	1.03	86	1.03	116	1.08	78	1.12	114	0.94	92	1.06
	5L	94	0.90	117		89		126		87		106		98	
RA vs diabetes	3L	93	0.84	103	0.97	73	1.04	121	0.99	59	1.31	95	1.08	72	1.25
	5L	79		100		76		121		78		103		90	
CVD vs diabetes	3L	73	1.14	86	1.17	67	1.19	75	1.44	62	1.25	88	1.04	76	1.11
	5L	83		100		79		109		77		92		85	
Stroke vs diabetes	3L	211	0.99	267	0.91	195	1.05	238	1.16	203	0.96	203	1.04	201	1.07
	5L	209		242		205		276		195		211		215	
Depression vs diabetes	3L	43	0.58	28	0.82	27	1.13	33	1.18	46	0.88	26	1.13	19	1.75
	5L	25		23		30		39		40		29		33	
Personality vs diabetes	3L	89	0.54	45	0.91	58	1.00	50	1.38	95	0.92	50	1.15	41	1.23
	5L	49		41		58		69		88		57		50	
COPD vs liver disease	3L	195	1.27	237	1.20	194	1.23	229	1.24	148	1.48	238	1.20	217	1.18
	5L	247		286		240		284		220		286		256	
RA vs liver	3L	172	1.25	217	1.16	168	1.27	235	1.17	115	1.77	202	1.39	177	1.36
	5L	216		251		213		274		203		281		240	
CVD vs liver	3L	132	1.61	174	1.35	146	1.42	149	1.57	112	1.68	181	1.32	175	1.23
	5L	213		235		207		235		188		239		215	
Stroke vs liver	3L	444	1.18	583	1.00	446	1.16	521	1.21	410	1.16	462	1.18	473	1.13
	5L	521		582		516		632		476		543		534	
Depression vs liver	3L	82	1.00	69	1.07	69	1.41	76	1.33	85	1.30	63	1.58	57	1.80
	5L	82		73		97		101		111		101		102	
Personality vs liver	3L	163	0.92	107	1.16	139	1.27	113	1.50	178	1.25	117	1.56	112	1.38
	5L	150		125		176		169		223		182		155	

\*Green cells indicate a significant F ratio in favor of 5L, red cells in favor of 3L (95% CI, 3000 bootstrap samples)

**Overall 5L discriminates much better compared to 3L among different severity groups. This is true across countries.**

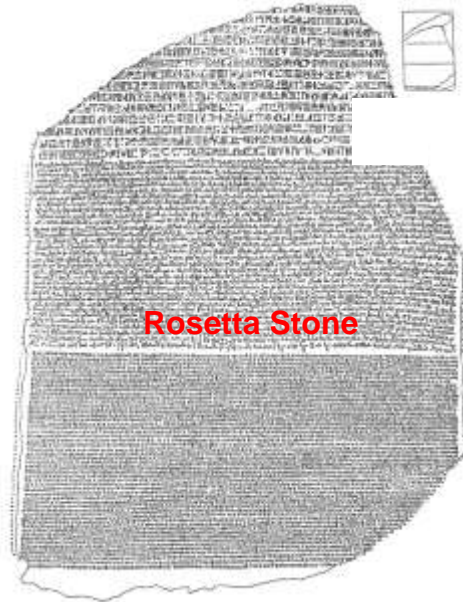
pl.org



## Conclusions

- 5L shows enhanced psychometrics on descriptive data
- 5L shows smoother more 'natural' distributions, this affects discriminative power in general
- 3L seems discriminates better than 5L when comparing healthy students vs patients; 5L discriminates better among severity classes. This is true for most countries although various language versions plus matching country value sets show different overall values  
*Explanation: in healthy students the winner (11111) 'takes all' enhancing the contrast with the remaining states*
- Caveats: Limitations of data (some groups from a single country); student cohort (Poland) as proxy for a healthy general population sample

www.euroqol.org



[www.euroqol.org](http://www.euroqol.org)



## Comparing 3L and 5L EQ-5D

### Acknowledgements

#### 3L to 5L mapping:

Dr Monica Hernandez, Prof Steve Pudney –  
SchARR, University of Sheffield

#### Case studies:

Dr Monica Hernandez, Dr Sabine Grimm –  
SchARR, University of Sheffield

Dr Manuel Gomes, Dr Zia Sadique – LSHTM,  
London.

Dr David Meads, John O'Dwyer – University of  
Leeds

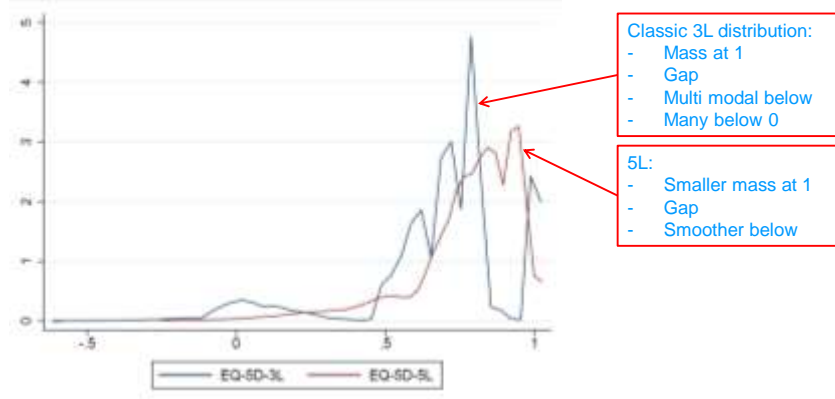
### Allan Wailoo

Professor of Health Economics  
Director, NICE Decision Support Unit  
Health Economics and Decision Science  
School of Health and Related Research (SchARR)  
University of Sheffield, UK

Datasource	National Databank for Rheumatic Diseases	EuroQoL
N	5,192	3,691
Patient characteristics	Rheumatoid Arthritis	Range of disorders (and students)
Setting	United States and Canada	Denmark, England, Italy, the Netherlands, Poland, and Scotland
Method	Postal and web. 5L first then 3L. Massive separation.	5L first then 3L, little separation
Year	January 2011	
Range of disease	3L: -0.594 to 1 5L: -0.226 to 1	3L: -0.594 to 1 5L: -0.281 to 1

31/10/2016 © The University of Sheffield

## Empirical distribution functions of 3L and 5L (NDB Jan 2011)



31/10/2016 © The University of Sheffield





## Modelling method

33

- Need to map from 3L to 5L, and the other way round
- Joint model of the two descriptive systems, conditional on age and sex
- Copula-based model:
  - 10 equation model (5 items of the descriptive system x 2 instruments) allowing for the correlation between each
  - Differences in utility scores (UK tariffs) then follow from the relationship in descriptive systems
  - Overall difference made up of two parts:
    - Responses to descriptive systems
    - Tariffs for health states

31/10/2016 © The University of Sheffield



## Modelling Headlines

34

- The relationship between 3L and 5L is different between datasets
  - Is this because the distribution of disease severity is different?
  - Is it because of disease specific aspects?
  - Different design issues?
- The models work very well in predicting 5L utility from 3L
- Stata code will be available to translate any 3L state into a predicted 5L, and vice versa.

31/10/2016 © The University of Sheffield

Title	ICER (inc QALYs)		
	3L	5L EuroQoL	5L NDB
CARDERA	13,666 (0.084)	15,252 (0.075)	14,846 (0.077)
CACTUS	3,058 (0.15)	9,481 (0.05)	23,022 (0.02)
RAIN a)	184,700 (0.02)	738,800 (0.005)	1,231,333 (0.003)
RAIN b)	294,137 (0.051)	714,333 (0.021)	714,333 (0.021)
IMPROVE	-44,617 (0.052)	-48,113 (0.046)	-54,742 (0.042)
COUGAR II	27,180 (0.115)	26,434 (0.119)	26,484 (0.118)
ARCTIC	112,193 (0.058)	162,744 (0.043)	152,130 (0.046)

CARDERA – Early Rheumatoid Arthritis. Methotrexate vs Methotrexate plus steroid, CACTUS – Computer Assisted therapy for Asphasia, RAIN - traumatic brain injury pathways a) Dedicated neurocritical care units vs combined neurocritical and general critical care units b) Early transfer vs no or late transfer IMPROVE - suspected ruptured aortic aneurysms. Comparison of endovascular repair vs open repair. COUGAR2 - Docetaxel and Active Symptom Control versus Active Symptom Control Alone for Refractory Oesophagogastric Adenocarcinoma. ARCTIC - Attenuated dose Rituximab with ChemoTherapy In CLL  
31/10/2016 © The University of Sheffield

Title	ICER (inc QALYs)		
	3L	5L EuroQoL	5L NDB
CARDERA	13,666 (0.084)	15,252 (0.075)	14,846 (0.077)
CACTUS	3,058 (0.15)	9,481 (0.05)	23,022 (0.02)
RAIN a)	184,700 (0.02)	738,800 (0.005)	1,231,333 (0.003)
RAIN b)	294,137 (0.051)	714,333 (0.021)	714,333 (0.021)
IMPROVE	-44,617 (0.052)	-48,113 (0.046)	-54,742 (0.042)
COUGAR II	27,180 (0.115)	26,434 (0.119)	26,484 (0.118)
ARCTIC	112,193 (0.058)	162,744 (0.043)	152,130 (0.046)

Marginal health gain lower with 5L ICERs ↑

Except COUGAR II (advanced cancer trial): Mortality gains important!

CARDERA – Early Rheumatoid Arthritis. Methotrexate vs Methotrexate plus steroid, CACTUS – Computer Assisted therapy for Asphasia, RAIN - traumatic brain injury pathways a) Dedicated neurocritical care units vs combined neurocritical and general critical care units b) Early transfer vs no or late transfer IMPROVE - suspected ruptured aortic aneurysms. Comparison of endovascular repair vs open repair. COUGAR2 - Docetaxel and Active Symptom Control versus Active Symptom Control Alone for Refractory Oesophagogastric Adenocarcinoma ARCTIC - Attenuated dose Rituximab with ChemoTherapy In CLL  
31/10/2016 © The University of Sheffield

Title	ICER (inc QALYs)		
	3L	5L EuroQoL	5L NDB
CARDERA	13,666 (0.084)	15,252 (0.075)	14,846 (0.077)
CACTUS	3,058 (0.15)	9,481 (0.05)	23,022 (0.02)
RAIN a)	184,700 (0.02)	738,800 (0.005)	1,231,333 (0.003)
RAIN b)	294,137 (0.051)	714,333 (0.021)	714,333 (0.021)
IMPROVE	-44,617 (0.052)	-48,113 (0.046)	-54,742 (0.042)
COUGAR II	27,180 (0.115)	26,434 (0.119)	26,484 (0.118)
ARCTIC	112,193 (0.058)	162,744 (0.043)	152,130 (0.046)

Marginal health gain is usually lower when using NDB mapping compared to EuroQoL dataset

CARDERA – Early Rheumatoid Arthritis. Methotrexate vs Methotrexate plus steroid, CACTUS – Computer Assisted therapy for Asphasia, RAIN - traumatic brain injury pathways a) Dedicated neurocritical care units vs combined neurocritical and general critical care units b) Early transfer vs no or late transfer IMPROVE - suspected ruptured aortic aneurysms. Comparison of endovascular repair vs open repair. ARCTIC - Attenuated dose Rituximab with ChemoTherapy In CLL

Title	ICER (inc QALYs)		
	3L	5L EuroQoL	5L NDB
CARDERA	13,666 (0.084)	15,252 (0.075) 89%	14,846 (0.077) 92%
CACTUS	3,058 (0.15)	9,481 (0.05) 33%	23,022 (0.02) 13%
RAIN a)	184,700 (0.02)	738,800 (0.005) 25%	1,231,333 (0.003) 15%
RAIN b)	294,137 (0.051)	714,333 (0.021) 41%	714,333 (0.021) 41%
IMPROVE	-44,617 (0.052)	-48,113 (0.046) 89%	-54,742 (0.042) 81%
COUGAR II	27,180 (0.115)	26,434 (0.119) 103%	26,484 (0.118) 103%
ARCTIC	112,193 (0.058)	162,744 (0.043) 74%	152,130 (0.046) 79%

Impact is particularly pronounced in CACTUS (aphasia in stroke) and RAIN (traumatic brain injury) studies

Severity of patients?  
- RAIN approx 0.3 at baseline  
- But CARDERA only 0.4

CARDERA – Early Rheumatoid Arthritis. Methotrexate vs Methotrexate plus steroid, CACTUS – Computer Assisted therapy for Asphasia, RAIN - traumatic brain injury pathways a) Dedicated neurocritical care units vs combined neurocritical and general critical care units b) Early transfer vs no or late transfer IMPROVE - suspected ruptured aortic aneurysms. Comparison of endovascular repair vs open repair. ARCTIC - Attenuated dose Rituximab with ChemoTherapy In CLL

Title	ICER (inc QALYs)			5L NDB *
	3L	5L EuroQoL	5L NDB	
CARDERA	13,666 (0.084)	15,252 (0.075)	14,846 (0.077)	18,100 (0.065)
CACTUS	3,058 (0.15)	9,481 (0.05)	23,022 (0.02)	
RAIN a)	184,700 (0.02)	738,800 (0.005)	1,231,333 (0.003)	Better mapping model uses HAQ and pain as covariates Lowers marginal QALY still further
RAIN b)	294,137 (0.051)	714,333 (0.021)	714,333 (0.021)	
IMPROVE	-44,617 (0.052)	-48,113 (0.046)	-54,742 (0.042)	
COUGAR II	27,180 (0.115)	26,434 (0.119)	26,484 (0.118)	
ARCTIC	112,193 (0.058)	162,744 (0.043)	152,130 (0.046)	

CARDERA – Early Rheumatoid Arthritis. Methotrexate vs Methotrexate plus steroid, CACTUS – Computer Assisted therapy for Asphasia, RAIN - traumatic brain injury pathways a) Dedicated neurocritical care units vs combined neurocritical and general critical care units b) Early transfer vs no or late transfer IMPROVE - suspected ruptured aortic aneurysms. Comparison of endovascular repair vs open repair. COUGAR2 - Docetaxel and Active Symptom Control versus Active Symptom Control Alone for Refractory Oesophagogastric Adenocarcinoma. ARCTIC - Attenuated dose Rituximab with ChemoTherapy In CLL  
31/10/2016 © The University of Sheffield

- Impact on CACTUS model states (Computer Assisted therapy for Asphasia):



Difference:  
0.07 (3L) vs 0.02 (5L)



## Discussion

41

### Limitations

- Investigate impact of missing data imputation for CEAs
- More case studies needed
- Need to better understand the impact by severity/disease area
  - Stata program to allow anyone to do this

### Mapping between 3L and 5L

- The relationship between 3L and 5L instruments is well represented by the copula model approach
- The relationship is different depending on the dataset
- The relationship is better modelled including HAQ and pain in the RA dataset
  - Do we need disease specific 3L/5L data and disease specific explanatory models for future HTA?

31/10/2016 © The University of Sheffield



## Discussion

42

### Impact on ICERs

- Large differences in incremental QALYs and ICERs
  - Movement up the severity scale and compression within smaller range
- Difference may be larger for more severe patients
  - More compression at this end of distribution
- Where mortality is a big driver the worsening in the ICER can be offset
- Impact also depends on the dataset used for mapping model
  - 3L and 5L results cannot be interpreted in the same way
  - Simple proportional adjustment not appropriate. Changes differ across the distribution

31/10/2016 © The University of Sheffield



## Conclusions – Kind & Marti

- Show how cost effectiveness estimates may change with relatively small changes in HRQL methods
- Small changes to methods can change interpretation of cost effectiveness
  - In EQ-5D-5L there were many small changes
- Also underlines sensitivity to valuation method – broader lessons for HTA bodies that accept many different approaches to utility estimation

[www.euroqol.org](http://www.euroqol.org)



## Conclusions - Janssen

- Examines the effect of changing the descriptive system from 3L to 5L
  - Evidence suggests 5L is an improvement
- Valuation data
  - National value sets show 3L-5L differences
  - 3L better in some analyses; 5L better in other analyses
  - Is the 3L advantage actually pseudo-sensitivity – arising from very large number of students in full health?
  - More work needed to tease apart these effects?

[www.euroqol.org](http://www.euroqol.org)



## Conclusions - Wailoo

- Explores relationship between 3L and 5L through mapping
- Identifies important effects on ICER
- 5L data is inflated and squeezed into smaller range
- Raises ICER estimates?
- But is this just a UK phenomenon?
  - Other 5L national value sets are more similar to 3L sets in their distribution

[www.euroqol.org](http://www.euroqol.org)



## Wrap up

- 5 items has produced a powerful tool
  - 5L is different to 3L and produces different data
  - Testing shows some advantages for 5L, but not universally
- Significant ongoing role for 3L
  - Much ongoing research, new value sets, updating old value sets
- But also ongoing shift to 5L
  - Implications for decision makers need to be understood
- EQ Group also revisiting the fundamental questions of description and valuation that underpin our measures

[www.euroqol.org](http://www.euroqol.org)