The Impacts of Age and Gender on Mapped EQ-5D-5L Utilities and Quality-Adjusted Life Years (QALYS) in Cancer Clinical Trials

PCR119, Abstract #127212

Shaw JW[1], Bennett B[2], Russell J[2], Van Hout BA[3], Norman R[4], Frampton K[5], Moreno-Koehler A[6], Taylor F[6], Cocks K[5]

[1]Bristol-Myers Squibb, Lawrenceville, NJ, USA; [2]Bristol-Myers Squibb Pharmaceuticals Ltd, Uxbridge, UK; [3]University of Sheffield, Sheffield, UK; [4]Curtin University, Perth, Australia; [5]Adelphi Values Ltd, Bollington, UK; [6]Adelphi Values, Boston, MA, USA

Background

- For economic evaluation, treatment benefits are most commonly summarized using quality-adjusted life years (QALYs), which combine quantity and quality of survival in a single metric
- The "quality" component is often quantified through use of generic preference-based measures, among which the EQ-5D-3L is historically most representative
- BMS previously compared the sensitivity of the generic EQ-5D-3L against the condition-specific European Organisation for Research and Treatment of Cancer Quality of Life Utility Measure – Core 10 Dimensions (EORTC QLU-C10D)[1,2] and two algorithms for mapping to the EQ-5D-5L:
  - The National Institute for Health and Care Excellence (NICE) Decision Support Unit (DSU) (Hernandez-Alava)[3]
  - EuroQol Group (van Hout & Shaw)[4]

Objectives

- To explore the effects of age and gender on the derivation of mapped EQ-5D-5L utilities derived using algorithms developed by NICE DSU and EuroQol Group
- To provide further insight into the choice of mapping algorithm for use in future studies

Methods

- Analyses employed data from eight randomized controlled trials of nivolumab in which the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30 (EORTC QLQ-C30) and EQ-5D-3L were administered (Table 1)
  - Timepoints common across trials were selected for analyses (Baseline and Week 9/11-13)
- Utilities were estimated using:
  - EQ-5D-3L UK[5] value set
  - QLU-C10D UK[6] value set
  - Mapped EQ-5D-5L English value set[7] using the following algorithms:
    - DSU: Copula method[3]
    - DSU: Gaussian method[3]
    - EuroQol: including age and gender[4]
    - EuroQol: excluding age and gender[4]

Table 1: Phase 3 Trials Included in Analyses

Trial Protocol	Indication	Sample Size (All Randomized Subjects)	Treatment Groups	EQ-5D-3L	EORTC QLQ-C30
CheckMate 017	Previously Treated Advanced or Metastatic Squamous Cell Non-small Cell Lung Cancer	272	nivolumab docetaxel	X	
CheckMate 025	Advanced or Metastatic Clear-Cell Renal Cell Carcinoma Who Have Received Prior Anti-Angiogenic Therapy	821	nivolumab everolimus	X	
CheckMate 037	Advanced (Unresectable or Metastatic) Melanoma Patients Progressing Post Anti-CTLA-4 Therapy	405	nivolumab investigator's choice	X	X
CheckMate 057	Previously Treated Metastatic Non-squamous Non-small Cell Lung Cancer	582	nivolumab docetaxel	X	
CheckMate 066	Previously Untreated Unresectable or Metastatic Melanoma	418	nivolumab dacarbazine	X	X
CheckMate 067	Previously Untreated Unresectable or Metastatic Melanoma	945	nivolumab nivolumab + ipilimumab ipilimumab	X	X
CheckMate 141	Recurrent or Metastatic Platinum-refractory Squamous Cell Carcinoma of the Head and Neck	361	nivolumab investigator's choice	X	X
CheckMate 238	Completely Resected Stage IIIB/C or Stage IV Melanoma with High Risk for Recurrence	906	nivolumab ipilimumab	X	X

Statistical analyses

- Descriptive statistics were produced for the pooled trial data set as well as split by age and gender subgroups
- Resampling analysis involving a bootstrapping approach was employed to generate simulated samples with artificially skewed distributions of age and gender to assess impacts on utilities and QALYs; for example:
  - Female skew (60% female)
  - Female high skew (80% female)
  - Young skew (30% ≤45y; 25% >45y to ≤55y; 20% >55y to ≤65y; 15% >65y to ≤75y; 10% >75y)
  - Young high skew (40% ≤45y; 30% >45y to ≤55y; 20% >55y to ≤65y; 5% >65y to ≤75y; 5% >75y)
- Differential item functioning (DIF) was conducted to assess whether probabilities of EQ-5D-3L item responses differed among groups of respondents (males/females and age <65/≥65) after controlling for overall health (measured by visual analogue scale [VAS])
- Treatment-specific utilities were entered into UK cost-effectiveness models to derive QALYs for treatments

Results

Descriptive statistics

- Across EQ-5D utility indices, mean scores were higher for males than females (Figure 1) and inversely related to age (Figure 2); similar findings were observed for the QLU-C10D (data not shown)
- The EuroQol mapping algorithm yielded higher mean scores than the DSU algorithm for all age and gender subgroups
  - Figures 3 and 4 present Forest plots by gender and age, respectively, for mean difference in score between DSU (Copula) and EuroQol (including/excluding age and gender)
  - Similar findings were shown for DSU (Gaussian) compared to EuroQol (data not shown)

Figure 1. EQ-5D Score at Week 11-13 by Gender

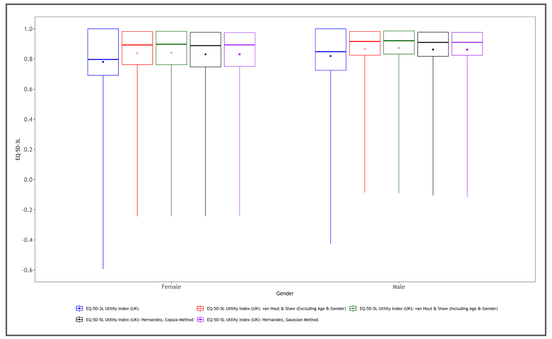


Figure 2: EQ-5D Score at Week 11-13 by Age

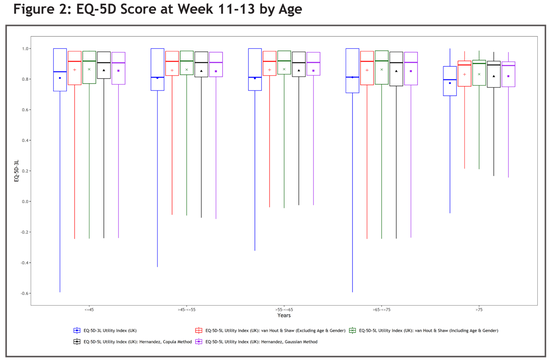


Figure 3: Forest Plot by Gender: Mean Difference in Score (95% CI) Between DSU (Copula) and EuroQol

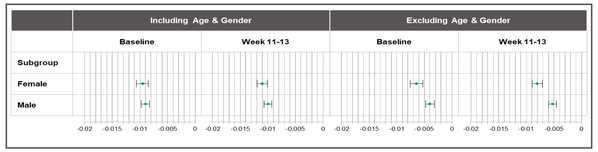
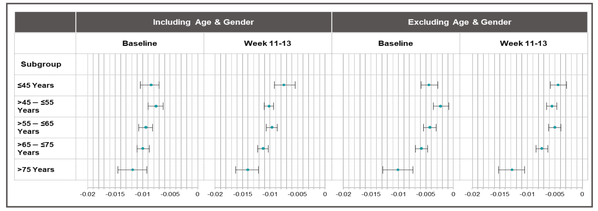


Figure 4: Forest Plot by Age: Mean Difference in Score (95% CI) Between DSU (Copula) and EuroQol



Resampling analysis

- In resampling analyses, mean scores were lower for all indices in simulated samples with a skew toward female gender (Table 2)
- At baseline, mean differences in mapped EQ-5D-5L utility index scores (original sample minus simulated sample with highest female gender skew) were:
  - 0.017 and 0.016, respectively, for the EuroQol algorithm including or excluding age and gender
  - 0.018 and 0.019, respectively, for the DSU algorithm with residuals specified as Gaussian or copula-mixture

Table 2: Utility Values with Resampled Age and Gender Distributions at Baseline

	EQ-5D-3L	EQ-5D-5L EuroQol (Excluding Age & Gender)	EQ-5D-5L EuroQol (Including Age & Gender)	EQ-5D-5L DSU (Copula Method)	EQ-5D-5L DSU (Gaussian Method)
Distribution	**Mean**	**Mean**	**Mean**	**Mean**	**Mean**
Original*	0.755	0.824	0.828	0.819	0.818
DSU†	0.756	0.824	0.828	0.818	0.817
Even‡	0.748	0.818	0.822	0.812	0.811
Young high skew	0.759	0.826	0.830	0.822	0.820
Young skew	0.753	0.822	0.826	0.817	0.816
Old skew	0.741	0.812	0.816	0.805	0.805
Old high skew	0.737	0.809	0.811	0.800	0.801
Female high skew	0.734	0.808	0.811	0.800	0.800
Female skew	0.744	0.815	0.818	0.808	0.808
Male skew	0.752	0.821	0.825	0.815	0.814
Male high skew	0.761	0.828	0.832	0.823	0.822

*Distribution in the pooled sample with ≥1 completed EQ-5D-3L
†Distribution in the dataset used to develop the DSU algorithm
‡Equal proportions of each category of age and gender subgroups

Conclusions

- Differences between the two mapped EQ-5D-5L indices were slight with the DSU index being more sensitive to sample demographic characteristics.
- Both mapping algorithms yielded higher QALY gains compared to EQ-5D-3L utilities for all skewed distribution samples, which is likely a function of the 5L descriptive system and/or valuation methodology.
- Limitations:
  - On-treatment assessment timepoints between each trial do not match exactly; an approximate on-treatment timepoint (Week 9/11-13) was employed
  - The English value set used in mapping is not currently endorsed by NICE

Differential item functioning

- Meaningful DIF was observed for the EQ-5D-3L anxiety/depression item with respect to age at Baseline (Table 3) and usual activities item with respect to gender at Week 9/11-13 (Table 4).

Table 3: Differential Item Functioning Models with VAS and Age as Covariates (Baseline)

EQ-5D-3L Item	No Problems (n)	Some Problems (n)	Extreme Problems (n)	Proportional Odds Assumption p-value*	VAS p-value	Age p-value	VAS×Age p-value
Mobility	2954	1131	23	-- †	--	--	--
Self-care	3713	373	20	-- †	--	--	--
Usual activities	2596	1327	182	0.02	<0.001	0.41	0.17
Pain/Discomfort	1948	1929	228	0.08	<0.001	0.20	0.33
Anxiety/Depression	2424	1583	99	0.11	<0.001	0.27	**0.01**

*<0.05 indicates assumption does not hold
†Insufficient data (n<50 with "Extreme problems" response) for model

Table 4: Differential Item Functioning Models with VAS and Gender as Covariates (Week 9/11-13)

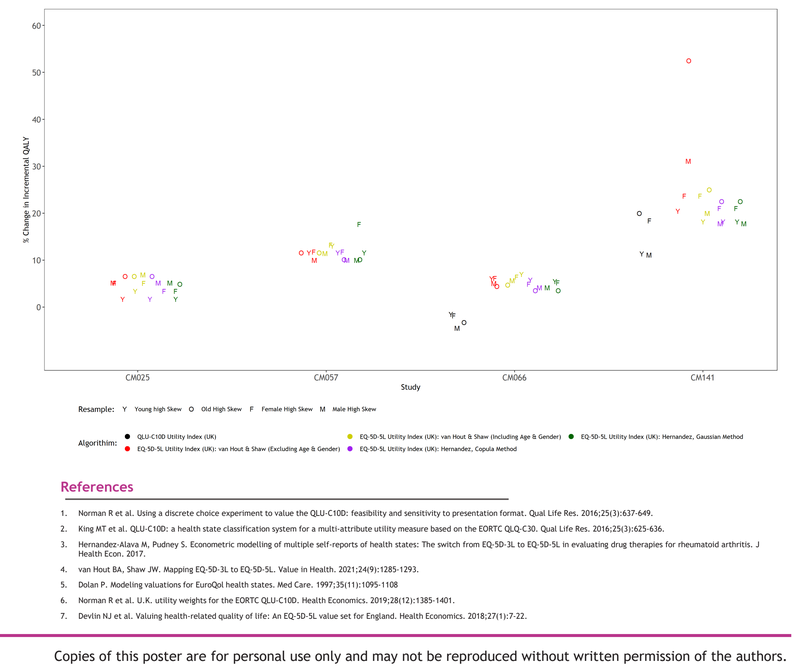
EQ-5D-3L Item	No Problems (n)	Some Problems (n)	Extreme Problems (n)	Proportional Odds Assumption p-value*	VAS p-value	Gender p-value	VAS×Gender p-value
Mobility	1903	600	7	-- †	--	--	--
Self-care	2331	163	15	-- †	--	--	--
Usual activities	1671	779	59	0.88	<0.001	0.12	**0.02**
Pain/Discomfort	1337	1104	69	<0.001	<0.001	0.60	0.16
Anxiety/Depression	1714	762	34	-- †	--	--	--

*<0.05 indicates assumption does not hold
†Insufficient data (n<50 with "Extreme problems" response) for model

Incremental QALYs

- EQ-5D-5L mapping algorithms yielded higher QALY gains compared to EQ-5D-3L utilities for all skewed distribution samples, as exemplified in Figure 5 (data presented for one trial per indication)
- QLU-C10D UK utilities are closer to EQ-5D-3L utilities (presented for studies in which EORTC QLQ-C30 was administered)

Figure 5. Incremental QALY Percentage Changes Compared to EQ-5D-3L (UK)



References

1. Norman R et al. Using a discrete choice experiment to value the QLU-C10D: feasibility and sensitivity to presentation format. Qual Life Res. 2016;25(3):637-649.
2. King MT et al. QLU-C10D: a health state classification system for a multi-attribute utility measure based on the EORTC QLQ-C30. Qual Life Res. 2016;25(3):625-636.
3. Hernandez-Alava M, Pudney S. Econometric modelling of multiple self-reports of health states: The switch from EQ-5D-3L to EQ-5D-5L in evaluating drug therapies for rheumatoid arthritis. J Health Econ. 2017.
4. van Hout BA, Shaw JW. Mapping EQ-5D-3L to EQ-5D-5L. Value in Health. 2021;24(9):1285-1293.
5. Dolan P. Modeling valuations for EuroQol health states. Med Care. 1997;35(11):1095-1108
6. Norman R et al. U.K. utility weights for the EORTC QLU-C10D. Health Economics. 2019;28(12):1385-1401.
7. Devlin NJ et al. Valuing health-related quality of life: An EQ-5D-5L value set for England. Health Economics. 2018;27(1):7-22.

Email: James.Shaw@bms.com