

Empirical validity of the EQ-5D and SF-6D

Reasons for the EQ-5D and SF-6D
providing different utility estimates

Garry Barton, Tracey Sach,
Kenneth Muir, Claire Jenkinson,
Tony Avery, Micheal Doherty,
Dave Whynes



The University of
Nottingham

Outline

- Details of study undertaken
- Analyses for empirical validity of EQ-5D & SF-6D
- Analyses to examine the extent of, and reasons for, differences between EQ-5D and SF-6D scores

Lifestyle interventions for knee pain (LIKP) study

- Assessment of the cost-effectiveness of different LIKP e.g. diet, exercise, diet + exercise, leaflet
- Ascertainment of patients with knee pain
 - All patients aged ≥ 45 years in 5 UK general practices

Ascertainment questionnaire

- Personal characteristics (age, sex, ethnicity, smoking status, height+weight (BMI), occupation).
- Health conditions (back pain, hip pain, knee pain, heart disease, stroke, asthma, cancer, diabetes, rheumatoid arthritis & osteoarthritis).
- HRQL:
 - EuroQol EQ-5D questionnaire (EQ-5D)
 - SF-36 questionnaire (used to derive SF-6D)

EuroQol EQ-5D

- **Mobility**
 - I have no problems in walking about (1)
 - I have some problems in walking about (2)
 - I am confined to bed (3)
- **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**
- **Pain**
- **Anxiety / Depression**

EQ-5D utility score

- 243 potential health states (e.g. 11213)
- UK valuation based on Time Trade Off (TTO) valuation of 45 EQ-5D health states by 3395 members of the UK general public
- Regression analysis used to predict utility score for each of the 243 health states (additive)
- Minimum score -0.594

SF-6D

- 11 questions from the SF-36
- Physical functioning (3a, 3b, 3j), role limitation (4a 5b), social functioning (10), pain (7, 8), mental health (9b, 9f), vitality (9e)
- 4 to 6 levels on each dimension.
- E.g. No pain (1) → Pain that interferes with normal work extremely (6)

SF-6D

- 249 out of 18,000 potential SF-6D health states were valued by 611 members of the UK general public (5 SF-6D states valued by each person)
- Standard gamble valuation of each state.
- Various regression models (additive, inconsistencies)
- Minimum score 0.296

Participants

- 6765 patients registered with one UK general practice
- 3122 aged ≥ 45 years
- 2770 were sent an ascertainment questionnaire
- 1865 returned ascertainment questionnaire
- Missing data
 - 128 (6.9%) EQ-5D
 - 253 (13.6%) SF-6D

	%
Age: >75	19.1%
≥65 to <75	23.6%
≥55 to <65	30.7%
≥45 to <55 years	26.5%
Gender: Female	52.8%
Male	48.2%
Ethnicity: Non-white	1.9%
White	98.1%
Smoked regularly: Yes	47.0%
No	53.0%
BMI: Obese	15.6%
Overweight	36.5%
Underweight	1.3%
Normal	46.6%
Occupational Skill Level:	
Lowest skill (1)	9.2%
Skill 2	32.0%
Skill 3	23.9%
Highest skill (4)	34.9%

	%
Back pain: Yes	18.4%
No	81.6%
Hip pain: Yes	10.4%
No	89.6%
Knee pain: Yes	23.2%
No	76.8%
Heart disease: Yes	11.6%
No	88.4%
Stroke: Yes	3.0%
No	97.0%
Asthma: Yes	7.2%
No	92.8%
Cancer: Yes	5.2%
No	94.8%
Diabetes: Yes	6.1%
No	93.9%
Rheumatoid Arthritis: Yes	5.2%
No	94.8%
Osteoarthritis: Yes	12.7%
No	87.3%

	Level	%
Mobility	1	70.3
	2	29.6
	3	0.2
Self-care	1	92.3
	2	7.5
	3	0.3
Usual activities	1	65.9
	2	31.9
	3	2.2
Pain	1	43.5
	2	51.5
	3	5
Anxiety / Depression	1	71.6
	2	26.4
	3	1.9

SF-6D Health state description

	Physical	Role limitations	Pain	Mental	Vitality	Social
1	21.6	58.9	27.9	29.8	4.3	69.6
2	36.6	16.9	27.0	35.0	37.8	10.9
3	20.4	9.1	22.3	28.6	38.8	12.6
4	7.9	14.5	12.3	7.7	12.1	6.4
5	6.7	N/A	10.5	1.7	9.3	3.1
6	6.8	N/A	3.2	N/A	N/A	N/A

Empirical validity

- Do utility scores correspond with prior expectations / hypotheses?
- Reference (comparator) group – the group expected to have higher utility

Age: >75
≥65 to <75
≥55 to <65
≥45 to <55 years^R

Gender: Female
Male^R

Ethnicity: Non-white
White^R

Smoked regularly: Yes
No^R

BMI: Obese
Overweight
Underweight
Normal^R

Occupational Skill Level:
Lowest skill (1)
Skill 2
Skill 3
Highest skill (4)^R

Back pain: Yes
No^R

Hip pain: Yes
No^R

Knee pain: Yes
No^R

Heart disease: Yes
No^R

Stroke: Yes
No^R

Asthma: Yes
No^R

Cancer: Yes
No^R

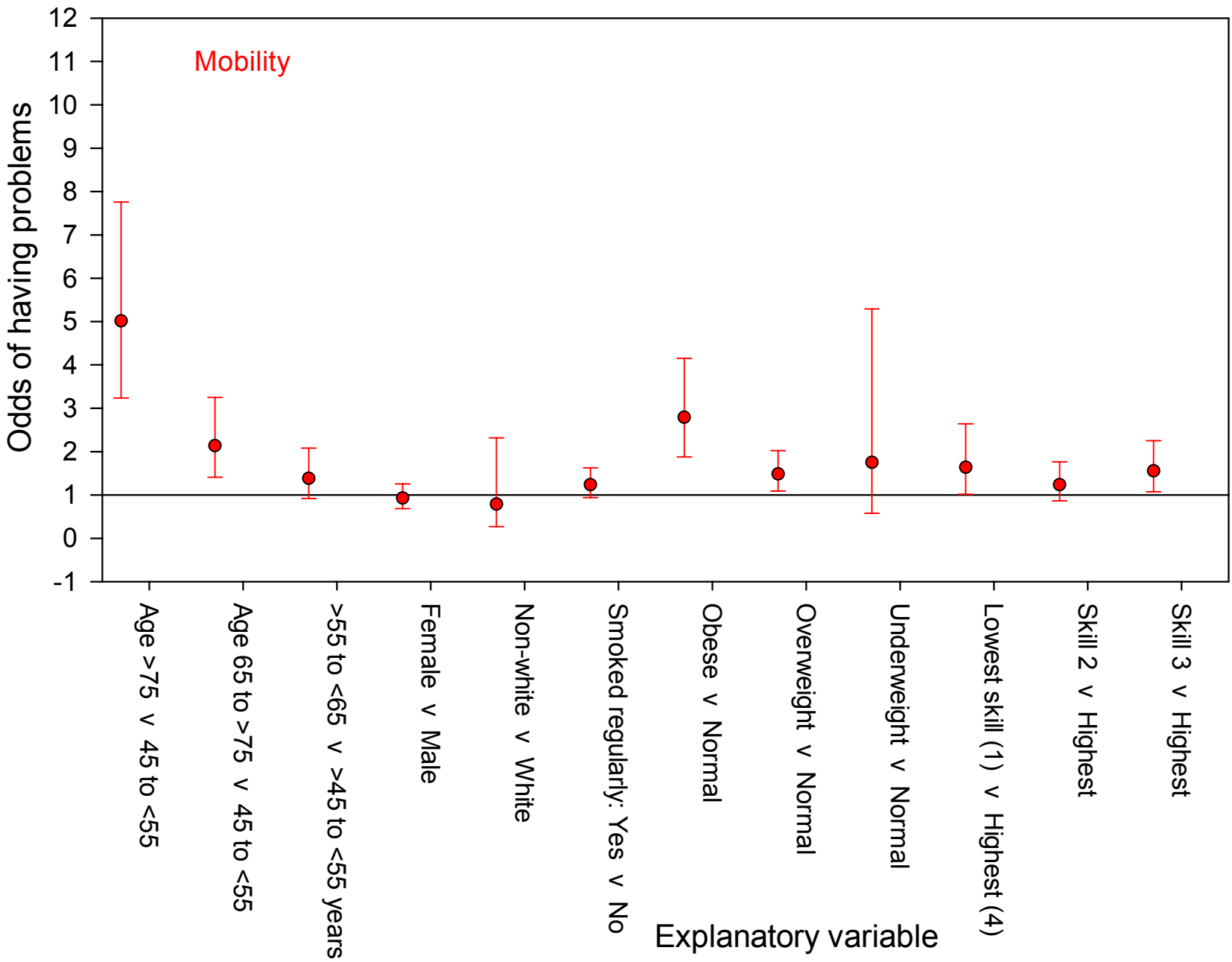
Diabetes: Yes
No^R

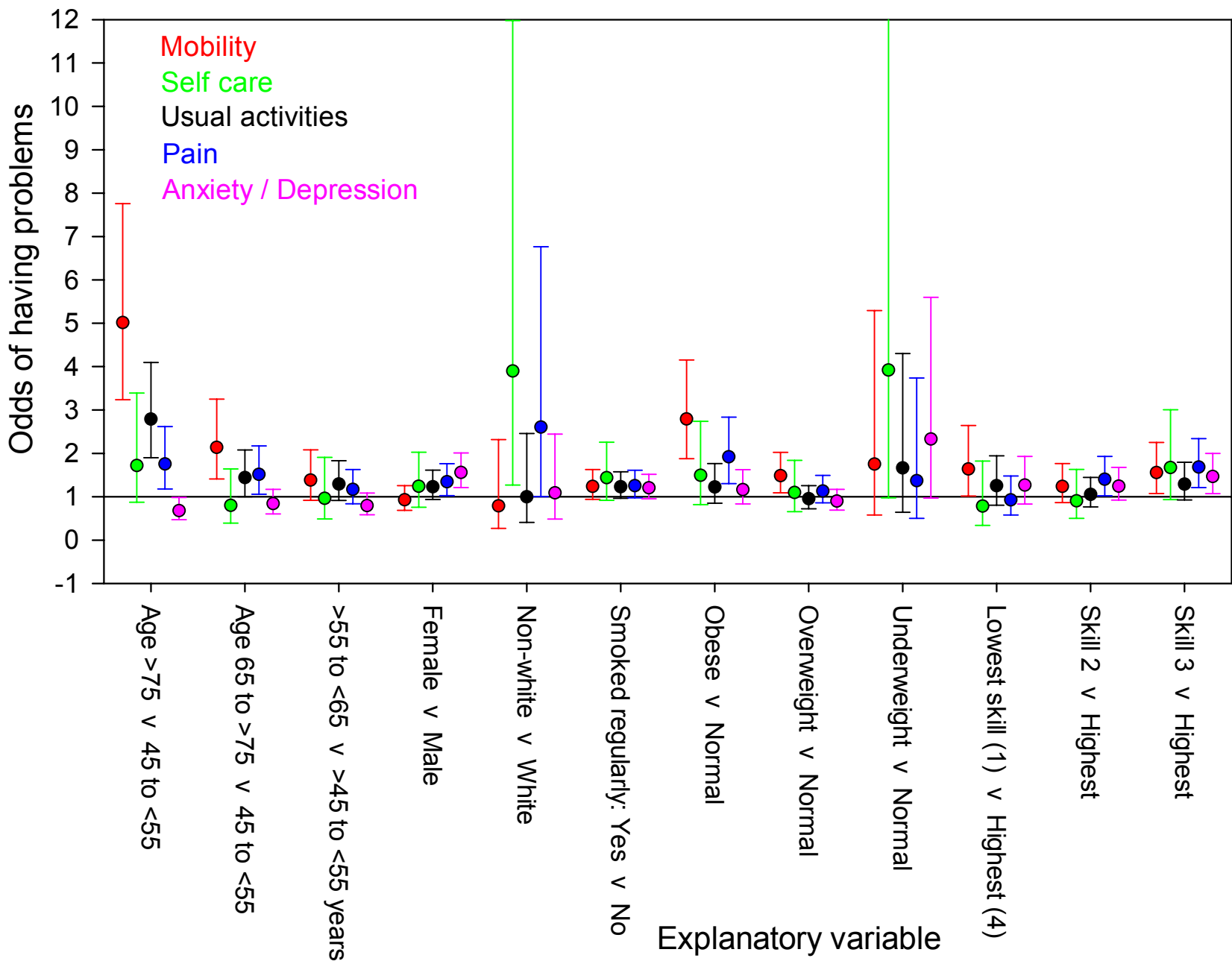
Rheumatoid Arthritis: Yes
No^R

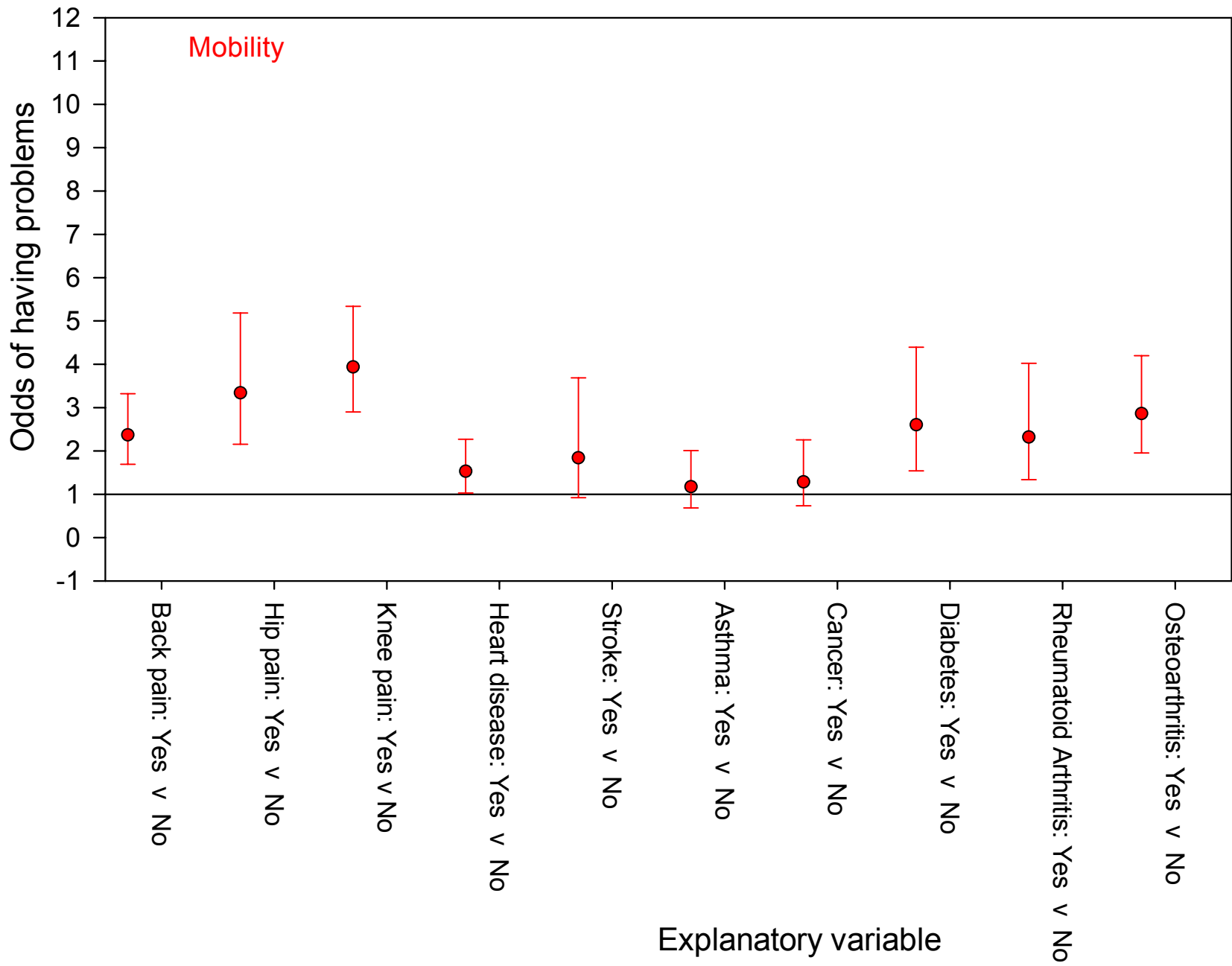
Osteoarthritis: Yes
No^R

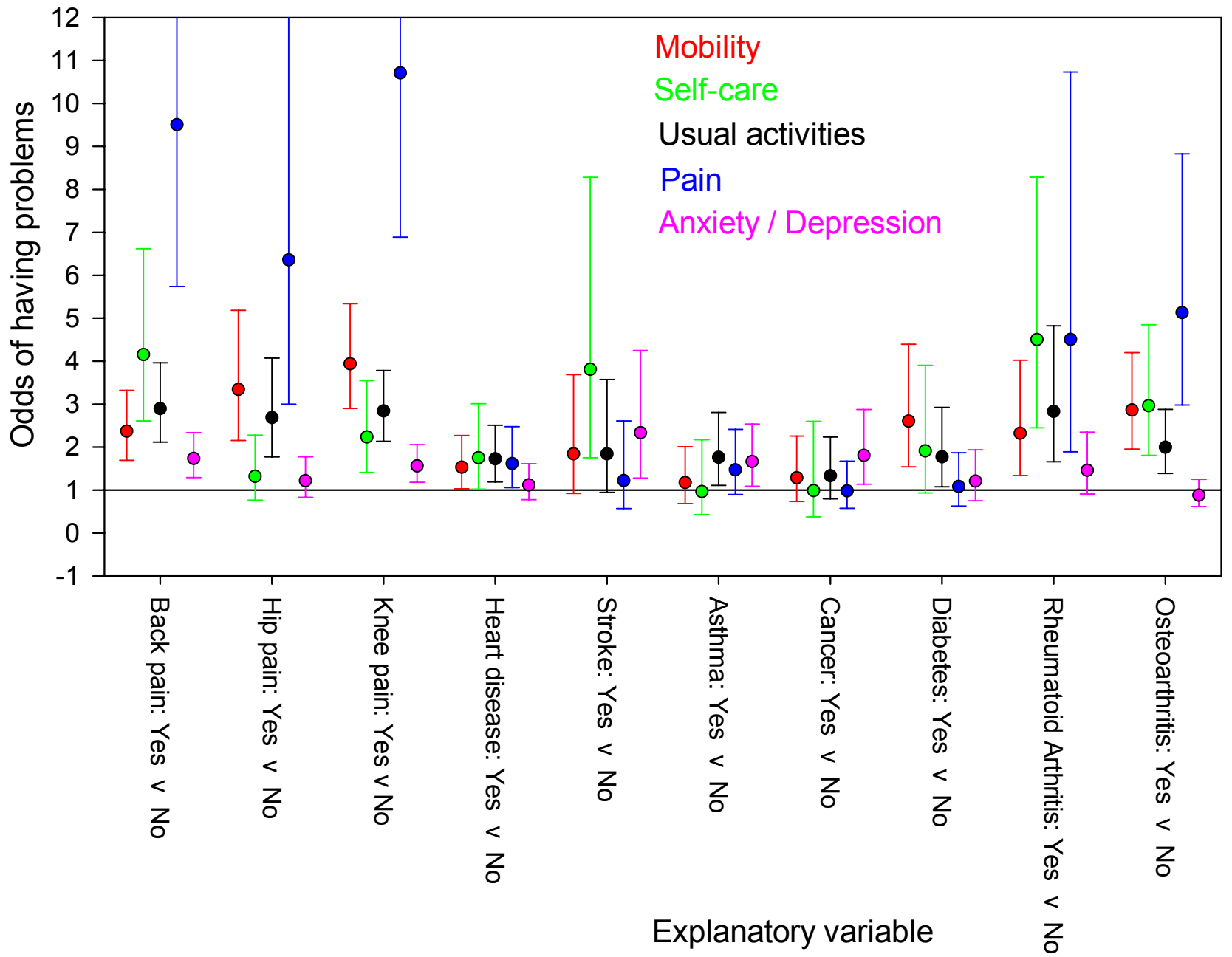
Regression analysis (1)

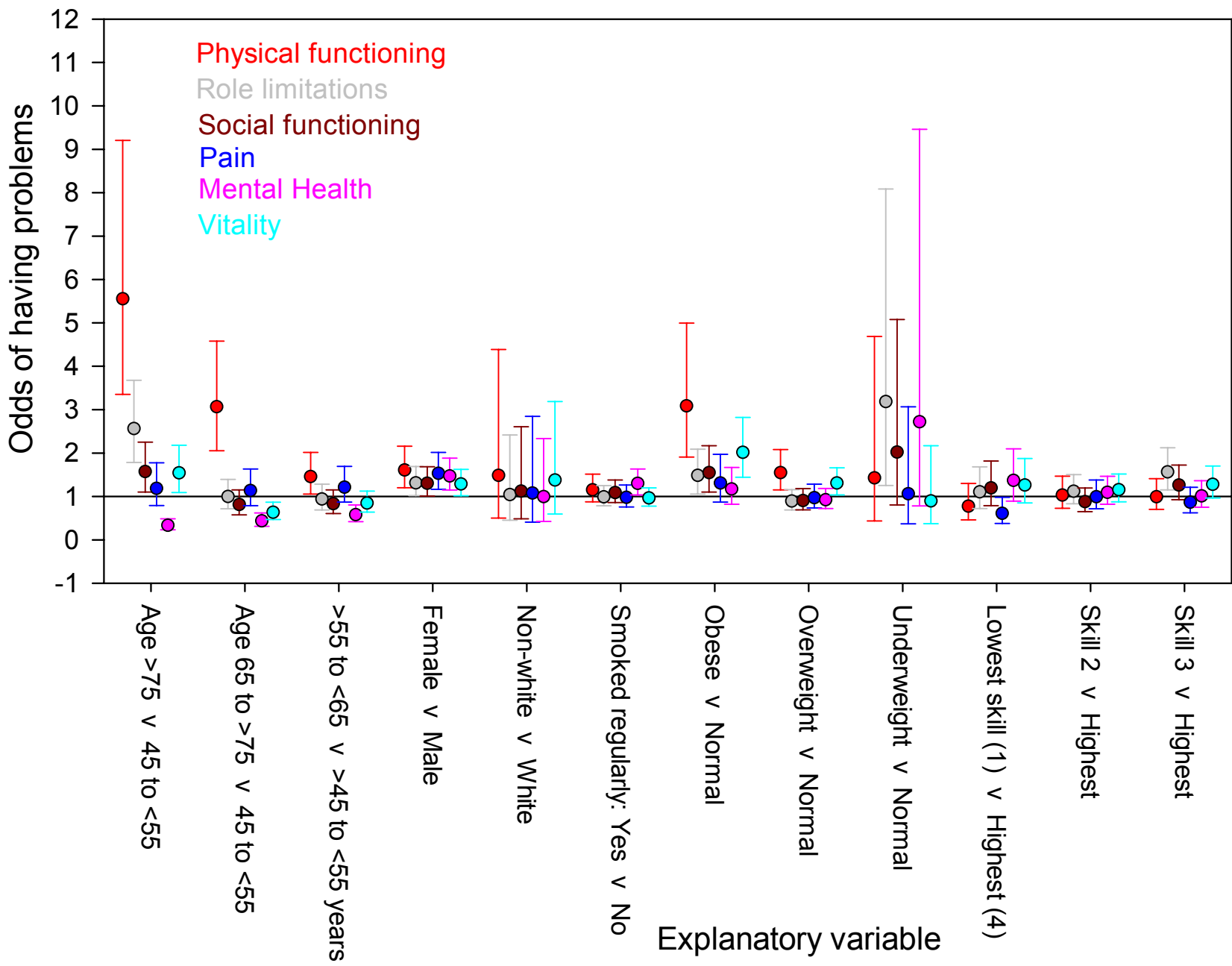
- Binary logistic regression
- Dimensions act as dependent variables
 - 0 (Level 1 i.e. No problems);
 - 1 (Level ≥ 2 i.e. some problems)
- Explanatory variables: personal characteristics & health conditions
- Odds of having some problems. e.g. back pain Vs no back pain, after controlling for confounding variables
 - OR > 1 more likely to have some problems on dimension

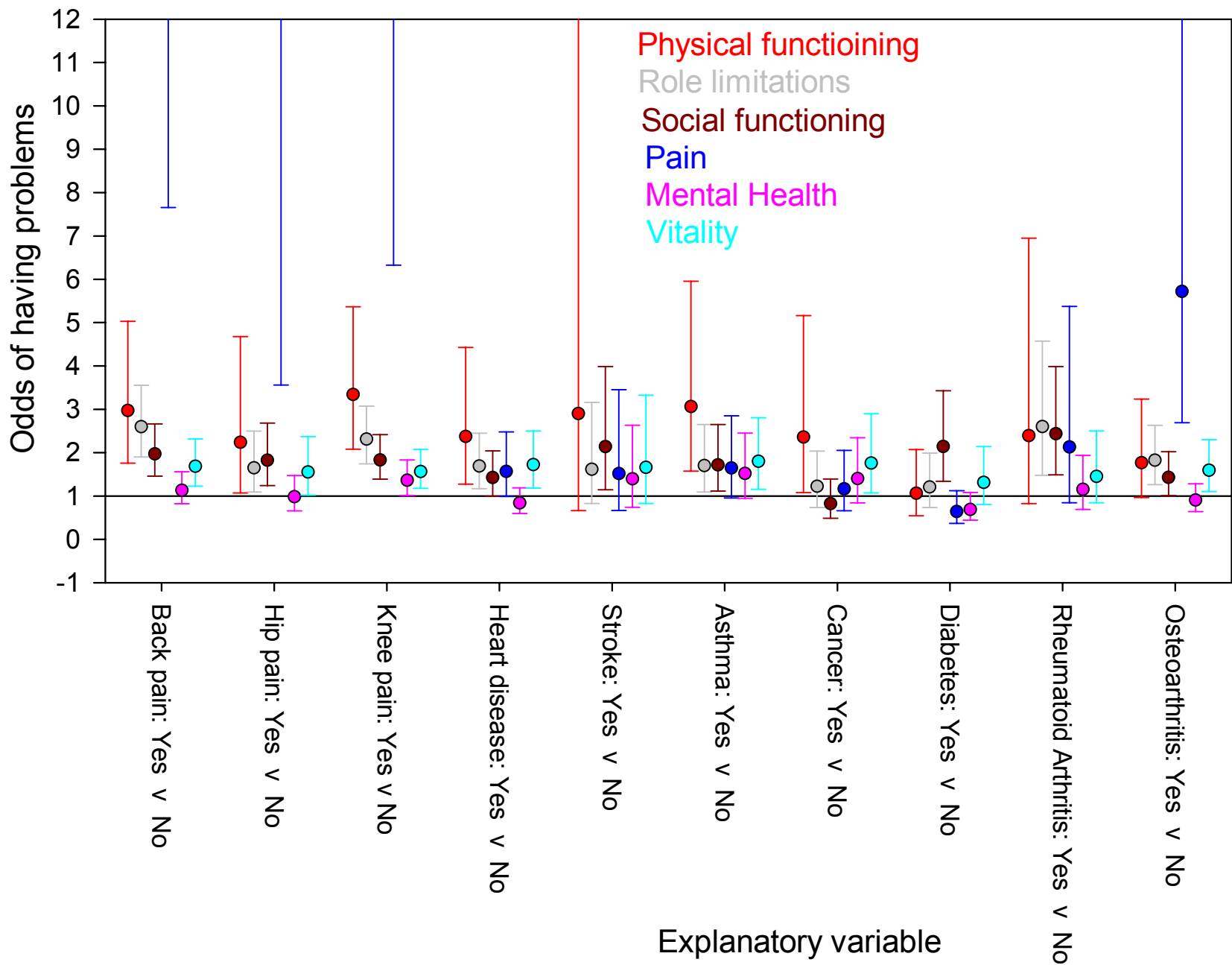












Binary logistic regression summary

	EQ-5D					SF-6D					
	Mobility	Self-care	Usual activities	Pain	Anxiety / Depression	Physical functioning	Role limitations	Social functioning	Pain	Mental Health	Vitality
Age	*		*	*	(*)	*	*	*		(*)	*
Sex				*	*	*	*	*	*	*	*
Ethnicity		*		*						*	
Smoking										*	
Obesity	*			*		*	*	*			*
Underweight	*			*		*	*	*			
SOC	*			*	*		*	*			
Back pain	*	*	*	*	*	*	*	*	*		*
Hip pain	*		*	*		*	*	*	*		*
Knee pain	*	*	*	*	*	*	*	*	*	*	*
Heart disease	*	*	*	*		*	*				*
Stroke		*			*			*			
Asthma			*		*	*	*	*			*
Cancer					*	*					*
Diabetes	*		*					*			
RA	*	*	*	*				*			
OA	*	*	*	*			*	*	*		*

Different dimensions pick up different effects,
no direct duplication / redundancy of dimensions

Utility scores for EQ-5D (SF-6D)

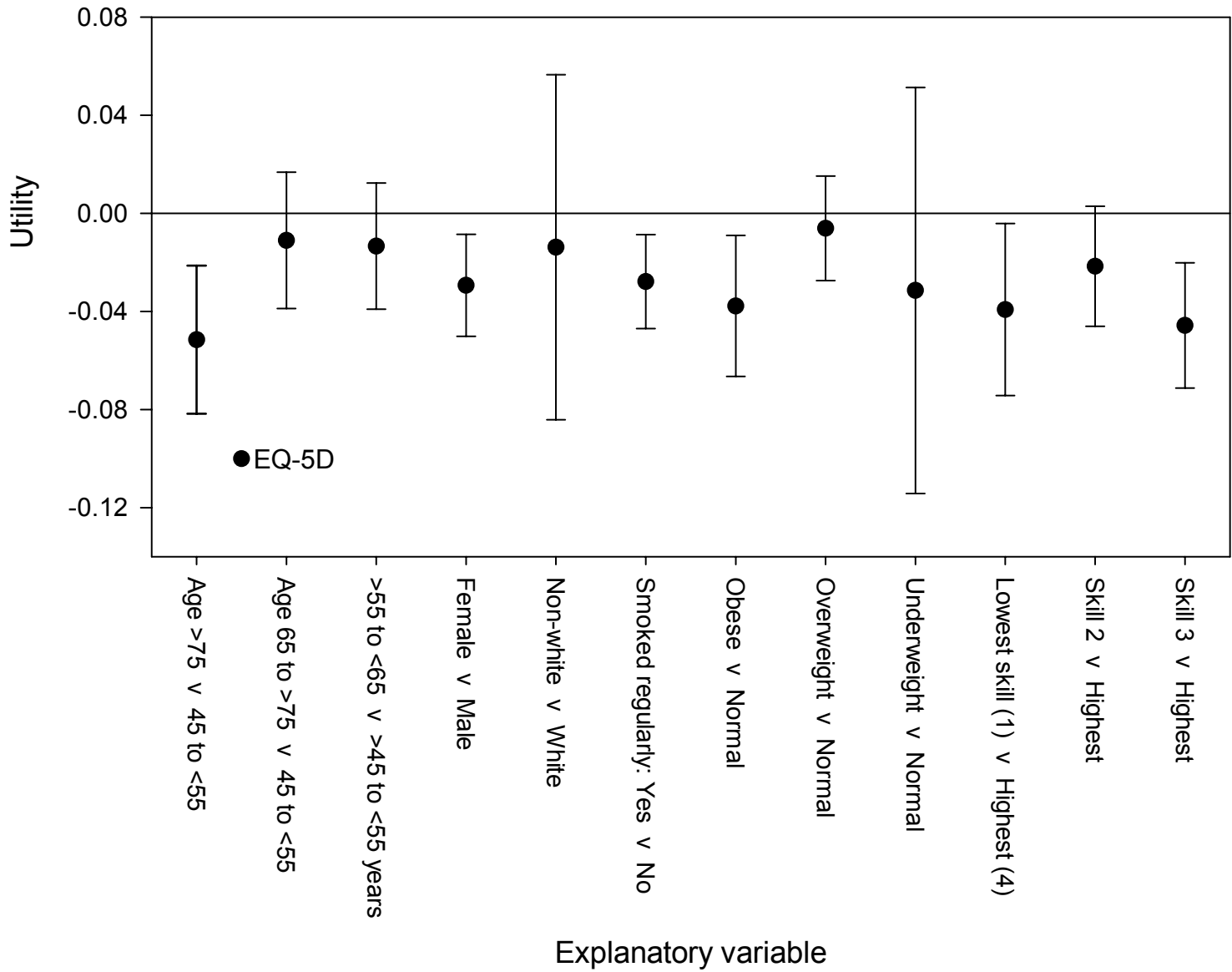
- 35.1% (4.3%) in full health (11111), utility of 1.0
- Min -0.35 (0.296)
- Mean 0.784 (0.768)
- 95% confidence interval 0.772 to 0.795
(0.761 to 0.775)

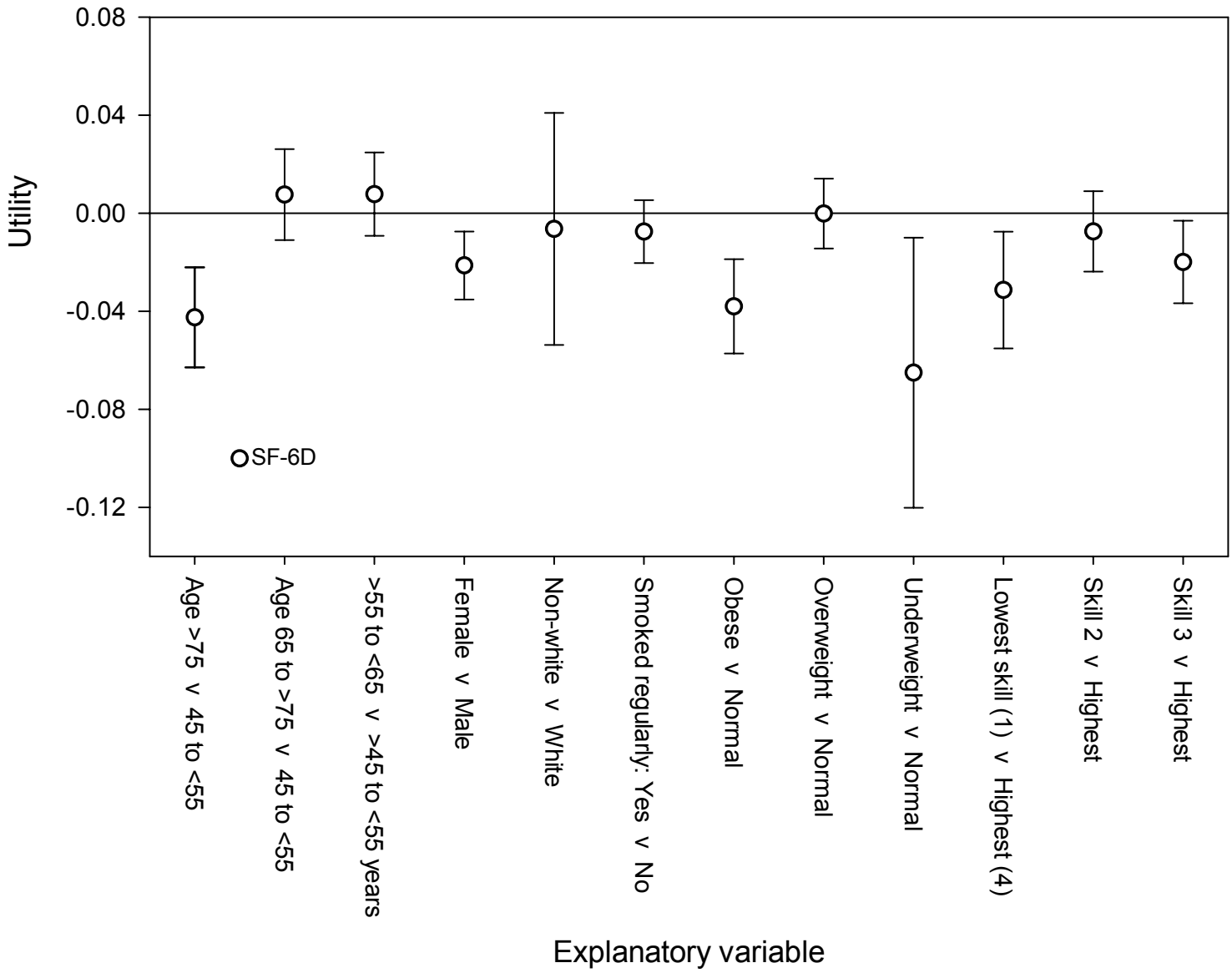
Regression analysis (2): Multiple linear regression

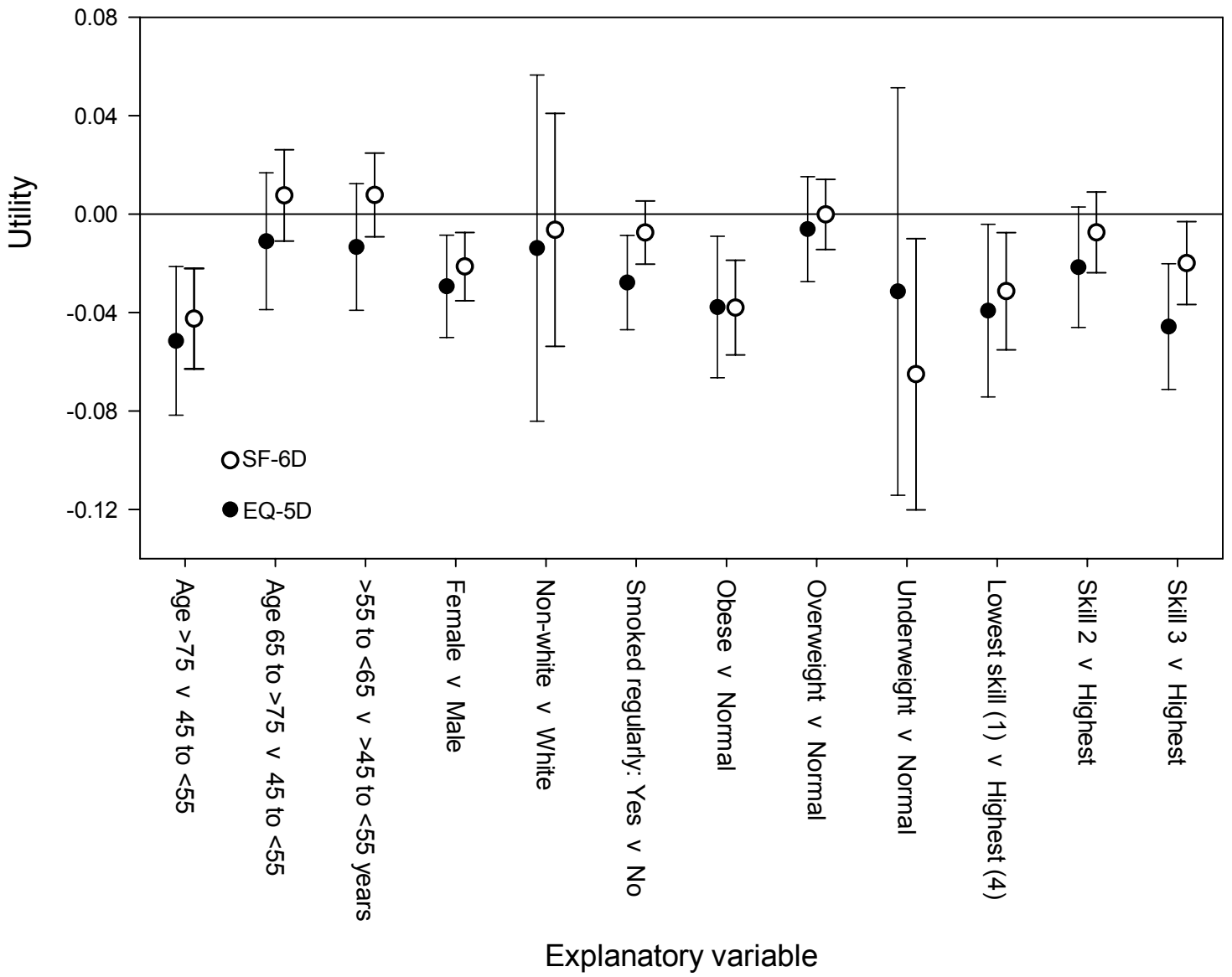
- Utility scores act as dependent variables
- Explanatory variables: personal characteristics & health conditions
- Parameter estimate – Difference in utility

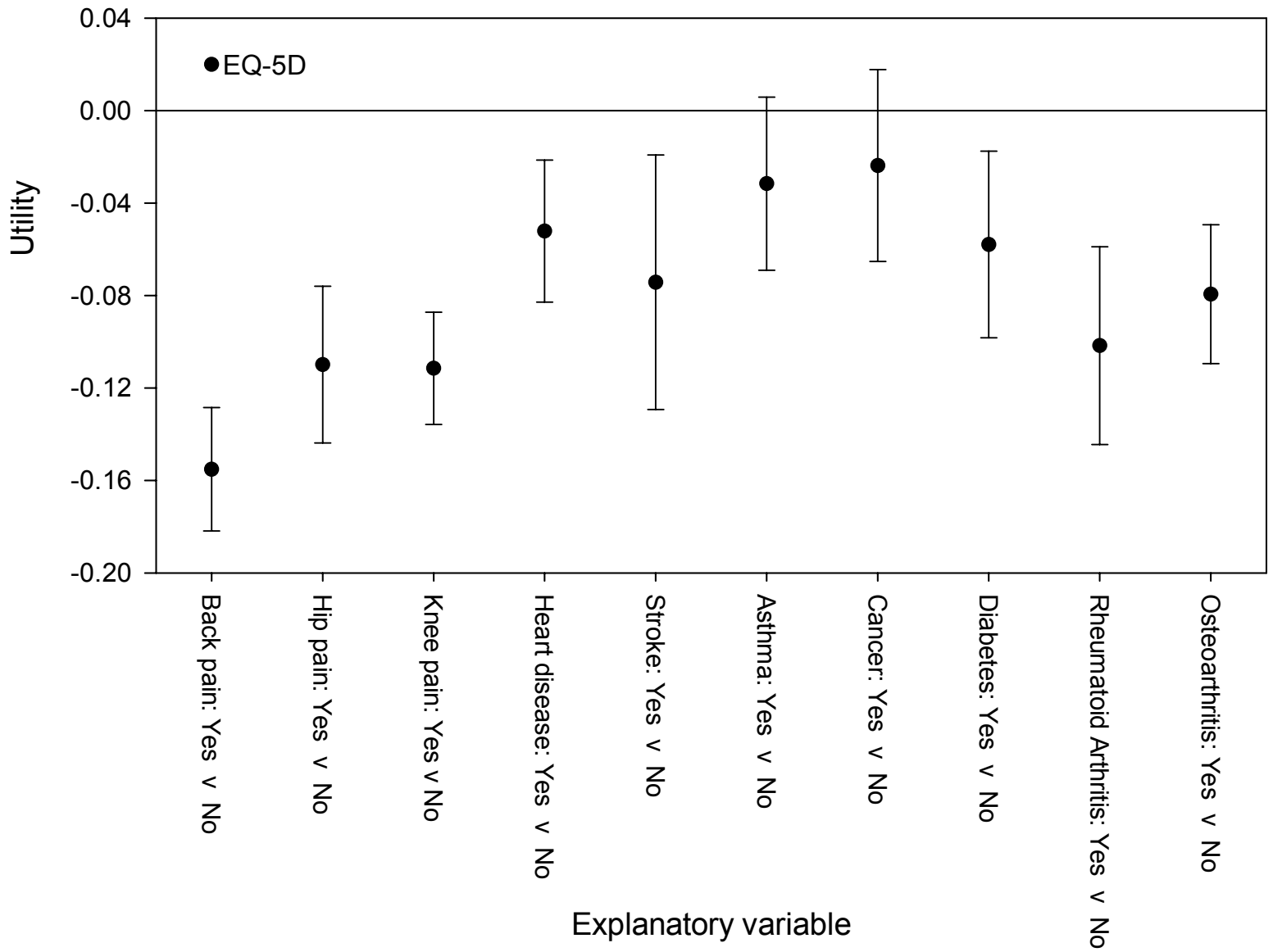
back pain V no back pain, whilst controlling for confounding variables

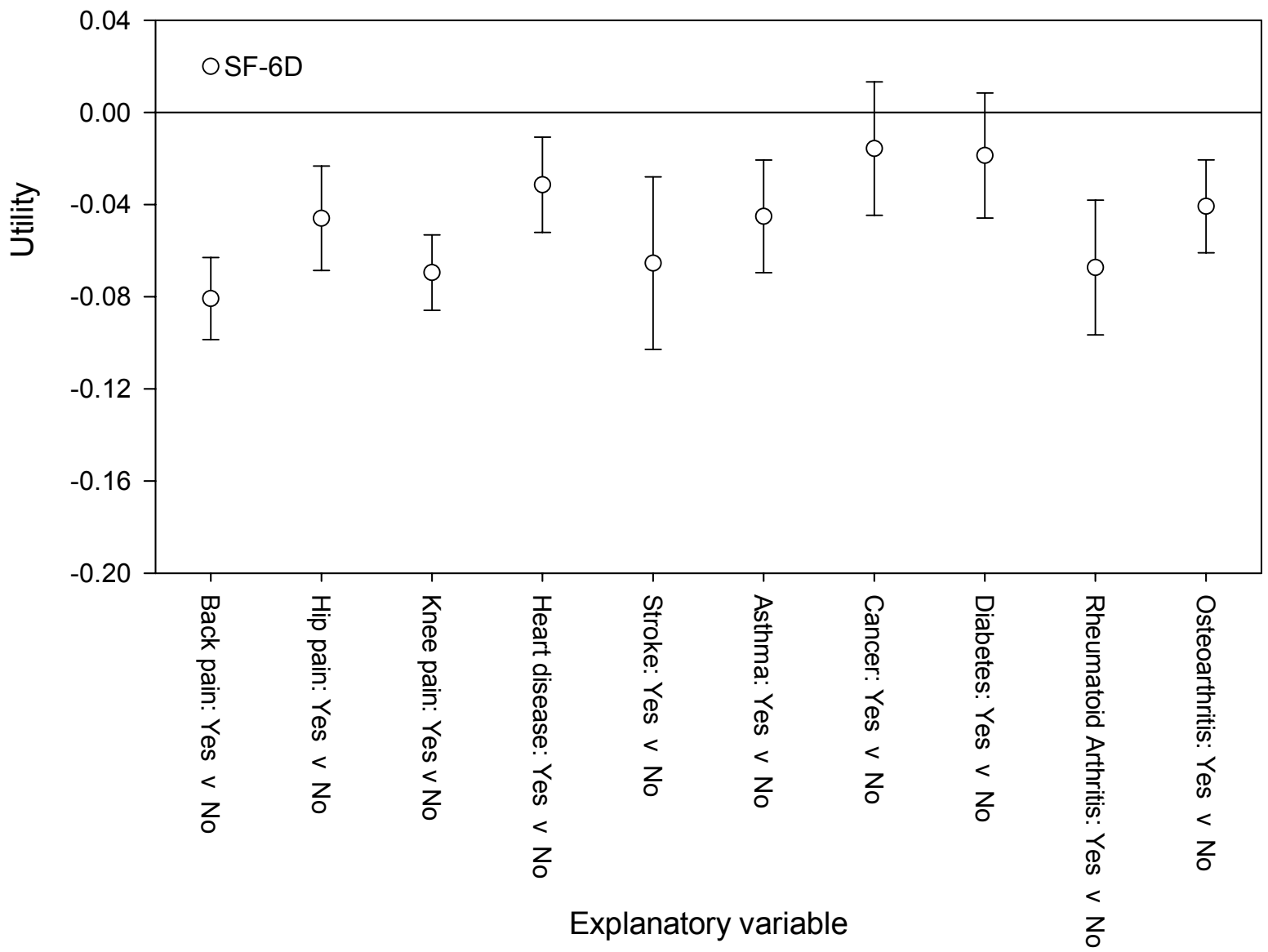
- Estimate <0 predicted to have lower utility

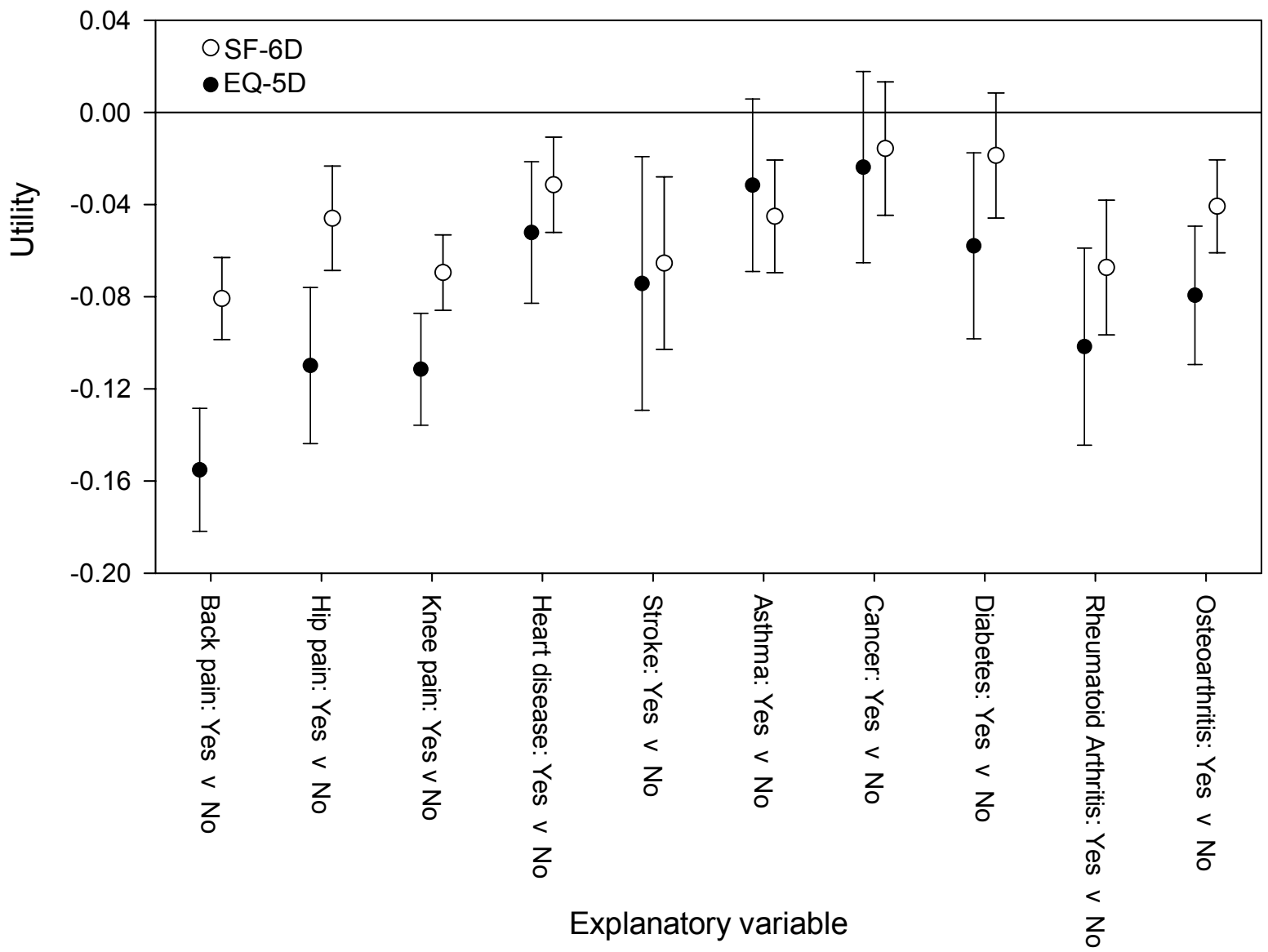












Regression analysis (2): summary

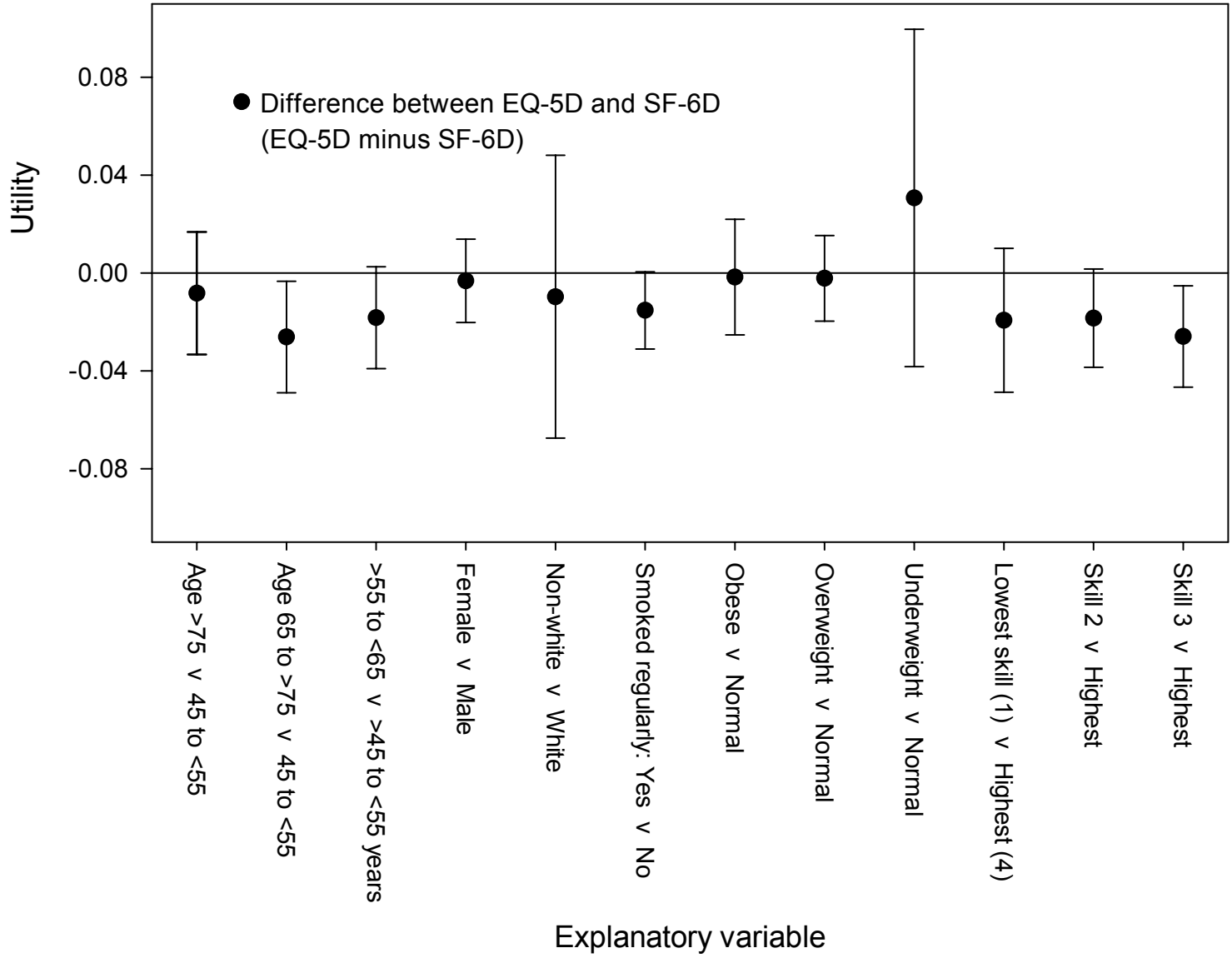
- EQ-5D utility scores, significant independent loss: older age, female, smoked regularly, obese, lower occupational skill, back pain, hip pain, knee pain, heart disease, stroke, diabetes, RA, OA
- SF-6D utility scores, significant independent loss: older age, female, obese, underweight, lower occupational skill, back pain, hip pain, knee pain, heart disease, stroke, asthma, RA, OA
- Utility scores in line with expectations / hypotheses
 - Evidence that EQ-5D & SF-6D are empirically valid

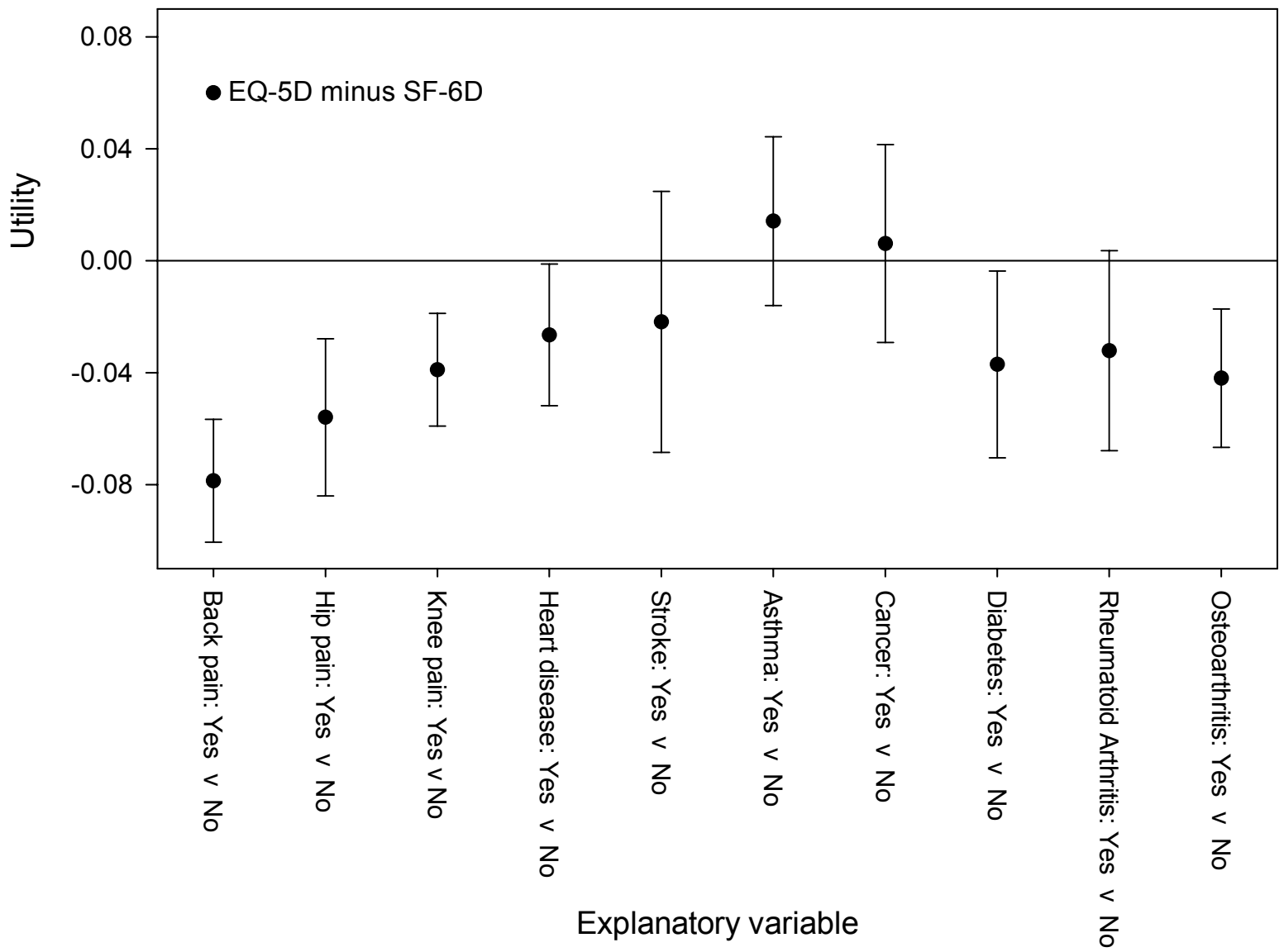
Differences between EQ-5D and SF-6D scores

- Quantifying the extent of the differences
- Identifying for whom scores differ
- Offering reasons for different utility estimates

Regression analysis (3): Multiple linear regression

- Difference in utility between EQ-5D and SF-6D (EQ-5D score minus SF-6D score)
- Explanatory variables: personal characteristics & health conditions
- Parameter estimate:
 - Difference between EQ-5D and SF-6D
 - Estimate <0 predicted to have greater utility loss on EQ-5D





Explanatory variable	Parameter estimate
Constant:	0.096***
Age: >75 v ≥ 45 to <55 years ^R	-0.008
≥ 65 to <75 v ≥ 45 to <55 years ^R	-0.026*
≥ 55 to <65 v ≥ 45 to <55 years ^R	-0.018
Gender: Female v Male ^R	-0.003
Ethnicity: Non-white v White ^R	-0.010
Smoked regularly: Yes v No ^R	-0.015
BMI: Obese v Normal ^R	-0.002
Overweight v Normal ^R	-0.002
Underweight v Normal ^R	0.031
Occupational Skill Level: Lowest (1) v	-0.019
2 v Highest ^R	-0.018
3 v Highest ^R	-0.026***
Back pain: Yes v No ^R	-0.079***
Hip pain: Yes v No ^R	-0.056***
Knee pain: Yes v No ^R	-0.039***
Heart disease: Yes v No ^R	-0.026*
Stroke: Yes v No ^R	-0.022
Asthma: Yes v No ^R	0.014
Cancer: Yes v No ^R	0.006
Diabetes: Yes v No ^R	-0.037*
Rheumatoid Arthritis: Yes v No ^R	-0.032
Osteoarthritis: Yes v No ^R	-0.042**

* p<0.05, ** p<0.01, *** p<0.001

Regression analysis (3): summary

- Patients in worse health states tend to have lower utility scores on the EQ-5D, up to 0.274 lower overall
- Healthier patients tend to have higher EQ-5D scores, 0.096 higher for best health state
- Explanations:
 - On EQ-5D 35.1% in full health, SF-6D 4.3%
 - Minimum score -0.35 (EQ-5D), 0.296 (SF-6D)

Further explanations

- Differing health state descriptions (Anal 1)
- Different valuation techniques:
 - TTO (EQ-5D) v SG (SF-6D)
 - Health state duration v probabilities
- Compare scores on common scale (0.296 to 1)
- Attach different values to EQ-5D health state descriptions, re-scale EQ-5D scores upwards

Figure 1

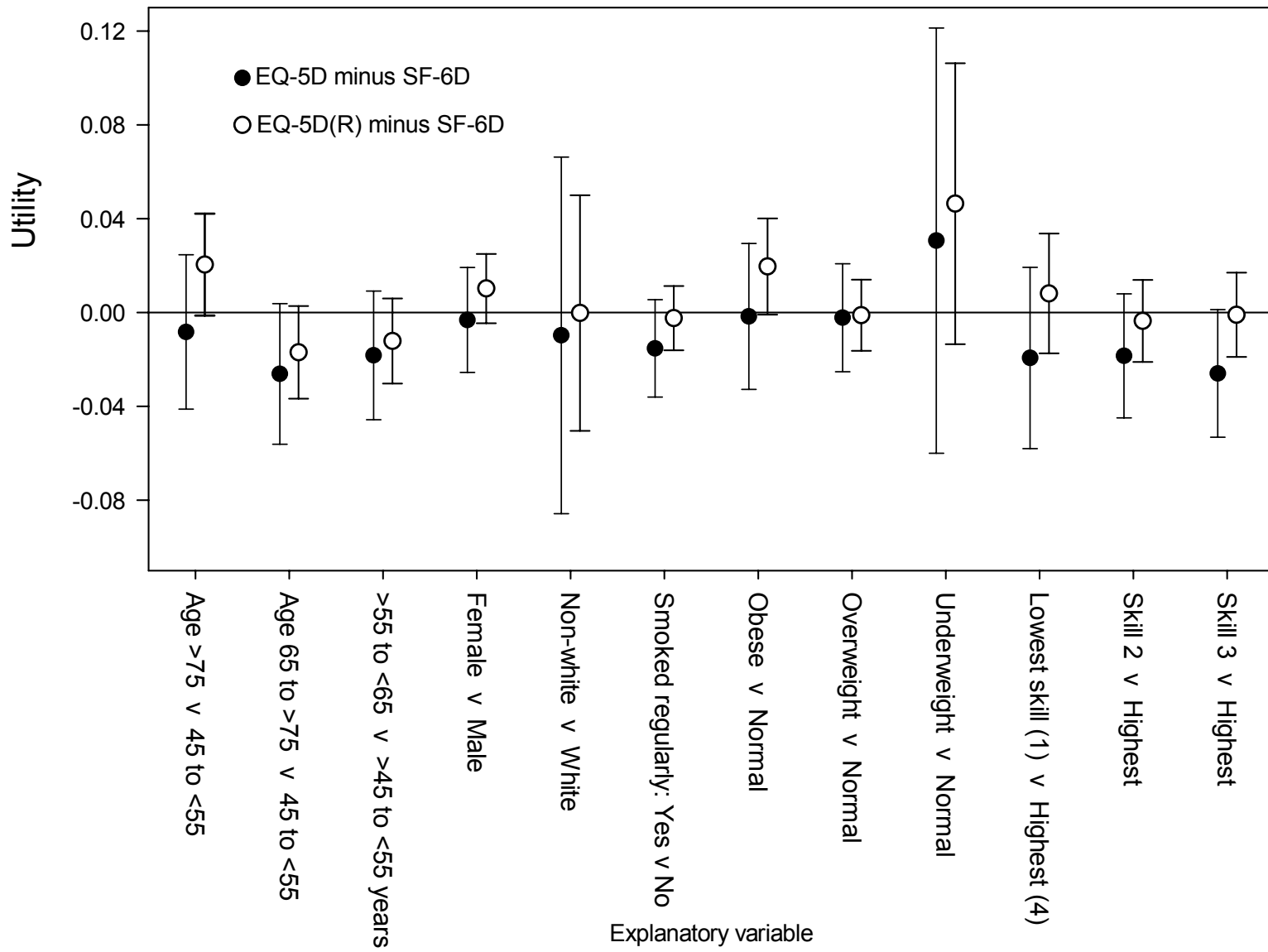
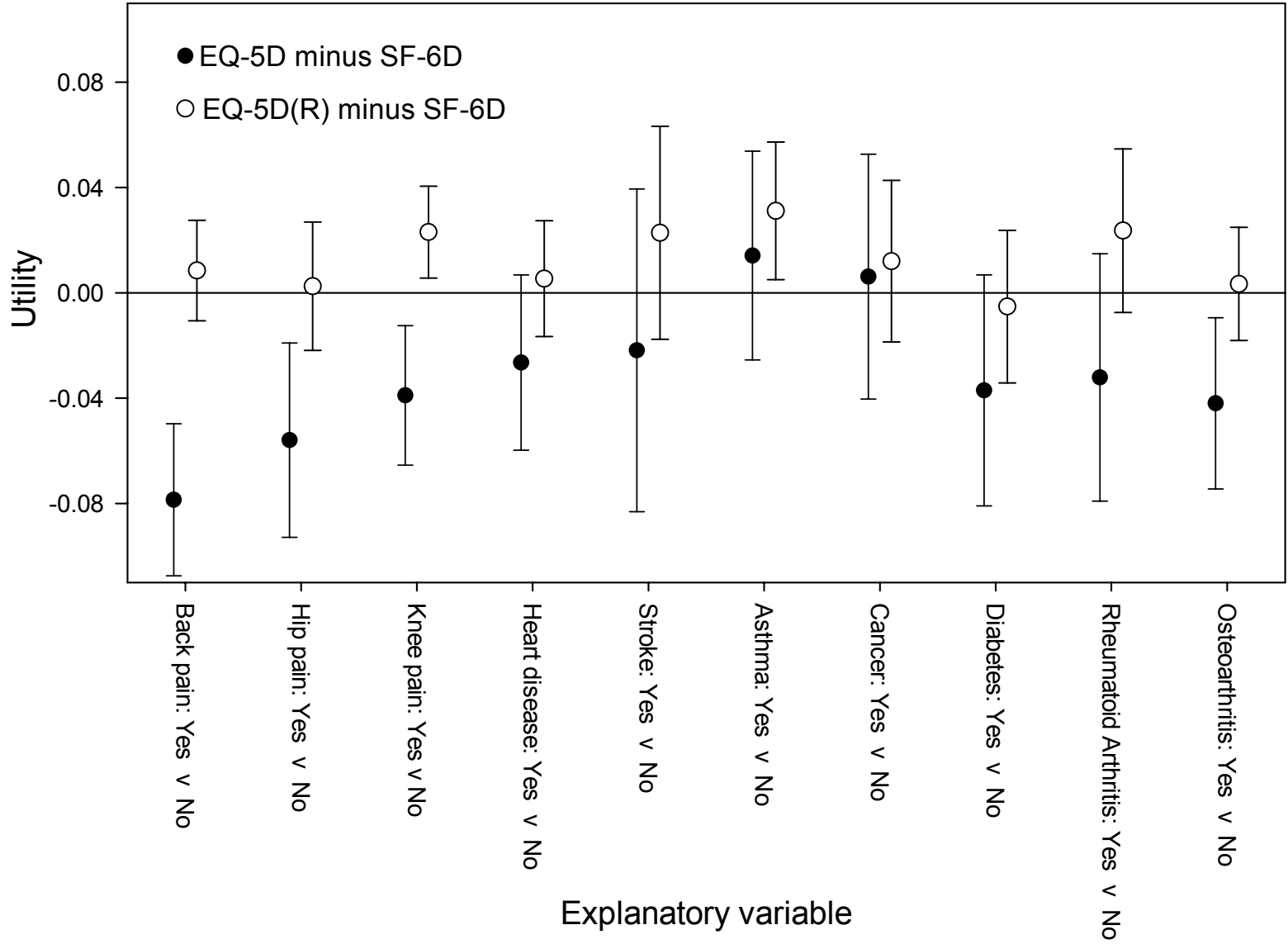


Figure 2



Explanations (continued)

- Dependent variable:
 - EQ-5D minus SF-6D Ad $R^2 = 13.6\%$
 - EQ-5D (R) minus SF-6D Ad $R^2 = 5.0\%$
- Fewer systematic differences when EQ-5D scores are re-scaled (re-valued)
- Evidence that differences between EQ-5D and SF-6D utility scores arise (at least in part) because of different valuation techniques.

Overall summary

- Evidence that different dimensions (of EQ-5D and SF-6D) detect problems associated with different conditions
- Evidence of empirical validity on EQ-5D and SF-6D
- Healthier patients tend to have higher utility on EQ-5D
- Less healthy patients tend to have lower utility on EQ-5D
 - Implies benefits of alleviating conditions higher on EQ-5D
- Utility differences (at least in part) due to different valuation techniques