ESTIMATING EQ-5D BY AGE AND SEX FOR THE UK

REPORT BY THE DECISION SUPPORT UNIT

12th January 2022

Mónica Hernández Alava, Steve Pudney, Allan Wailoo

School of Health and Related Research, University of Sheffield, UK

Decision Support Unit, ScHARR, University of Sheffield, Regent Court, 30 Regent Street
Sheffield, S1 4DA

Tel (+44) (0)114 222 0734
E-mail dsuadmin@sheffield.ac.uk
Website www.nicedsu.org.uk
Twitter @NICE_DSU

ABOUT THE DECISION SUPPORT UNIT

The Decision Support Unit (DSU) External Assessment Centre is based at the University of Sheffield with members at York, Bristol, Leicester and the London School of Hygiene and Tropical Medicine. The DSU is commissioned by The National Institute for Health and Care Excellence (NICE) to provide a research and training resource to support the Institute's Centre for Health Technology Evaluation Programme. Please see our website for further information [www.nicedsu.org.uk.](www.nicedsu.org.uk)

The production of this document was funded by the National Institute for Health and Care Excellence (NICE) through its Decision Support Unit. The views, and any errors or omissions, expressed in this document are of the authors only. NICE may take account of part or all of this document if it considers it appropriate, but it is not bound to do so.

Acknowledgements

We would like to thank OnePoll, in particular Bex Dean, for the provision of data. We received helpful comments on an earlier draft from Laura Gray (ScHARR) and Ian Watson (NICE) which have led to amendments in this final version.

This report should be referenced as follows:

Hernández Alava M., Pudney S., Wailoo A. Estimating EQ-5D by Age and Sex for the UK. NICE DSU Report. 2022.

CONTENTS

TABLES

FIGURES

ABBREVIATIONS AND DEFINITIONS

AIC	Akaike Information Criterion
ALDVMM	Adjusted Limited Dependent Variable Mixture Model
AS	Absolute Shortfall
BIC	Bayesian Information Criterion
HRQoL	Health-related quality of life
HSE	Health Survey for England
MVH	Measuring and Valuing Health Study
PS	Proportional Shortfall
QALY	Quality Adjusted Life Year

1. INTRODUCTION

Age adjustment of health-related quality of life (HRQoL) is routinely conducted in cost effectiveness models. Adjustments are made to reflect the fact that quality of life for the general population declines with age. Therefore, it can be argued that future years in the absence of the condition that may be the subject of any specific evaluation should not be assigned a value of 1, as if they would be lived in full health, but rather the age-relevant population mean. It can also be the case that adjustments need to be made to estimate the impact of specific health conditions or events if those estimates are drawn from studies using samples with different age profiles to the population of interest.

NICE specifically referred to the requirement for adjustment by age in its 2013 Methods Guide (NICE, 2013). The recently published 2022 updated methods guide (NICE, 2022) also refers to the need for age adjustment but there is an additional requirement for age adjustments when calculating the shortfall in terms of Quality Adjusted Life Years (QALY) associated with specific conditions. These are used to determine if the technology in question may attract some additional weight due to the degree of severity, as measured by absolute and proportional shortfall (AS and PS). Both measures of severity compare Quality Adjusted Life Expectancy (QALE) for the general population to QALE for those with the condition in question with current care. Health utility by age is a key input to both parts of these QALE calculations.

NICE recommends the use of the EQ-5D-3L for the calculation of health benefits. There is no favoured source for age adjustment of HRQoL in economic models in general and NICE does not recommend any specific source. Common choices for making these adjustments include values derived from various waves of the Health Survey for England (HSE, see for example Ara and Brazier, 2010 and Janssen and Szende, 2014). Others still use the original Measuring and Valuing Health (MVH) study with field work conducted in 1993 for EQ-5D-3L age norms (see Janssen and Szende, 2014). Given this variation, there is a need to review these and other potential data sources of age adjustment together with the methods used to estimate utilities from these sources. A key requirement for NICE decision-making is that there is consistency both within and across decisions. With this in mind, there may be a case for future recommendations from NICE to be more prescriptive on the approach to age weighting.

This report details the estimation of more recent health utility values, based on the EQ-5D-3L, by age and sex, for use in UK economic evaluations. Two sets of estimates are provided. The first used data from the 2014 wave of the Health Survey for England, the most recent wave including EQ-5D-3L. The second was based on data from a large scale UK survey conducted by the Economic Methods of Evaluation in Health and Social Care Policy Research Unit (EEPRU). These estimates are contrasted with existing alternatives, including an assessment of the sensitivity of calculations of absolute and proportional shortfall drawn from a sample of previous NICE Technology Appraisals.

It should be noted that we have chosen to focus on available data reporting the EQ-5D-3L and not the newer five level version (EQ-5D-5L). Currently, NICE does not recommend the use of the value set for the 5L version. Therefore, mapping is required to estimate values for 5L responses on the 3L scale. All else being equal, mapping is a sub-optimal approach compared to directly observed responses on the 3L instrument and introduces additional uncertainty. There may be other issues, including the need for consistency of methods within evaluations, and factors that are not equal across available 3L and 5L datasets (for example how current the data are) that result in a need to consider a broader set of options than those addressed in this report.

2. EXISTING ESTIMATES

2.1. HEALTH SURVEY FOR ENGLAND 2003 AND 2006 (ARA AND BRAZIER, 2010)

Ara and Brazier estimated the relationship between EQ-5D-3L and age and sex as part of a paper focussed on comparing different methods for making adjustments for utility values within cost effectiveness models. The EQ-5D-3L model formed a minor element of the overall study and, therefore, few details are reported. They made use of two waves of the Health Survey for England[1] (2003 and 2006) which comprised 26,679 observations pooled across years. These data were used to estimate the general population-based values of EQ-5D-3L by age and sex, using a linear regression. The reported model included age and age-squared as explanatory variables.

Figure 1: EQ-5D-3L by age based on the estimates in Ara and Brazier (2010)

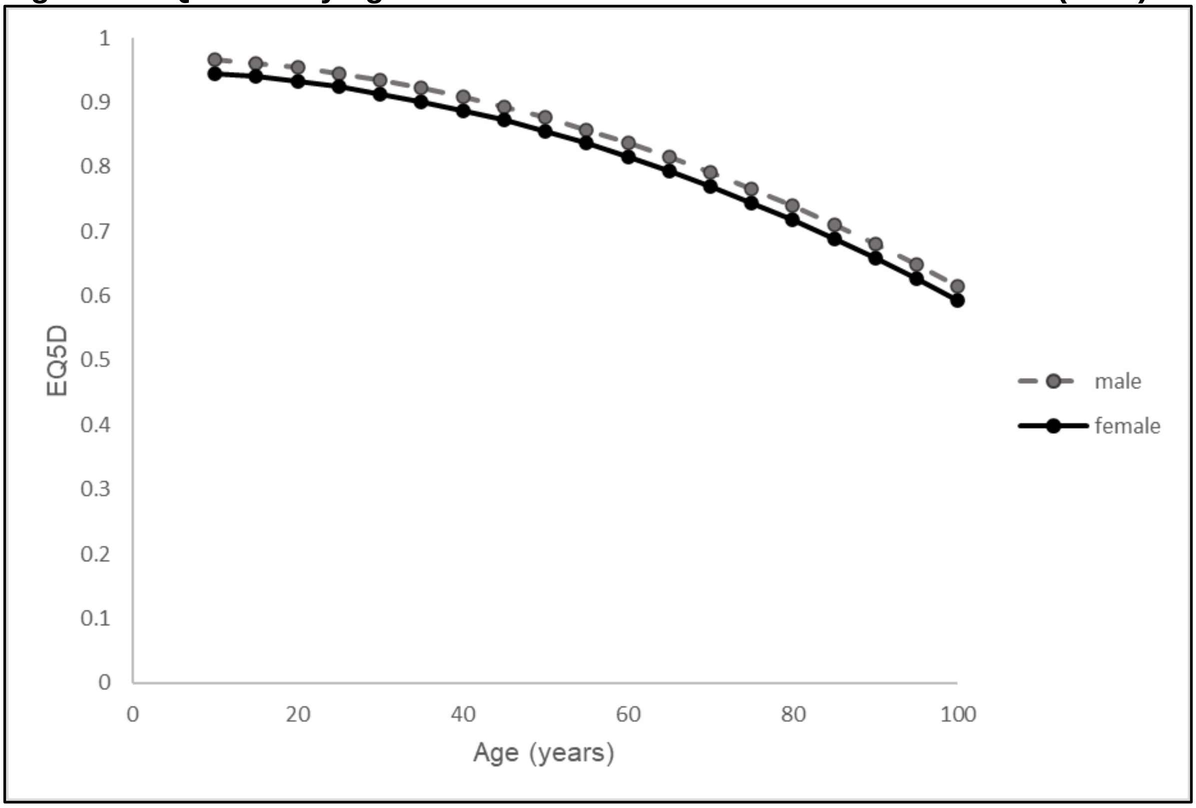


Figure 1 shows the relationship between age and sex estimated by Ara and Brazier (2010). “Sex” is an indicator variable, so the estimates for males and females are parallel with male EQ-5D-3L 0.026 higher than females at every age. The quadratic term results in EQ-5D-3L decreasing with age at an increasing rate.

We refer to these results as HSE 2003/6.

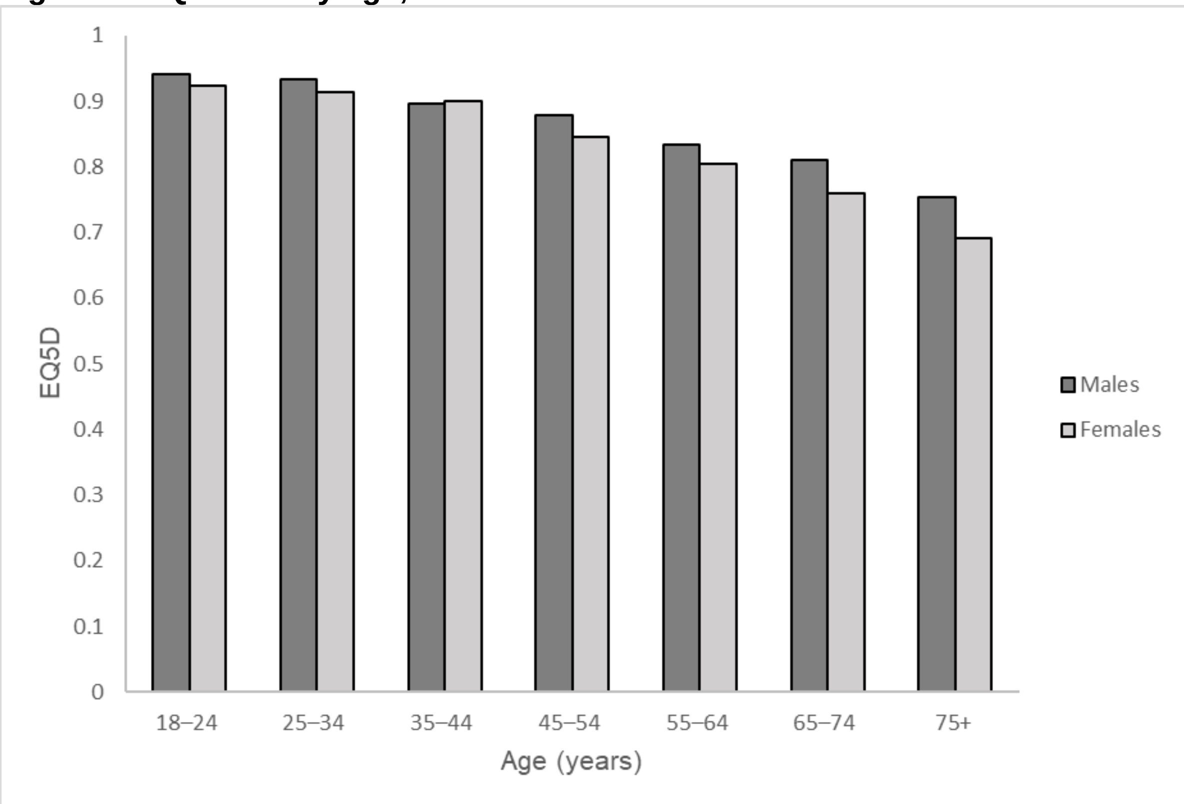
[1] See section 3 for details of the HSE.

2.2. HEALTH SURVEY FOR ENGLAND 2010 (JANSSEN AND SZENDE 2014)

Janssen and Szende report the results from the HSE 2010 (fieldwork completed during 2008), as simple mean EQ-5D-3L values for sex and age categories: age groups: 18–24, 25–34, 35–44, 45–54, 55–64, 65–74, 75+ years. The study had a sample size of 14,763. Figure 2 is drawn using the mid-point of each age category and an age of 85yrs for the upper unbounded category. EQ-5D-3L decreases with age. EQ-5D-3L for males is higher than that for females at all ages except the category 35-44 where females score marginally higher (0.900 vs 0.897).

We refer to these results as HSE 2010.

Figure 2: EQ-5D-3L by age, HSE 2010



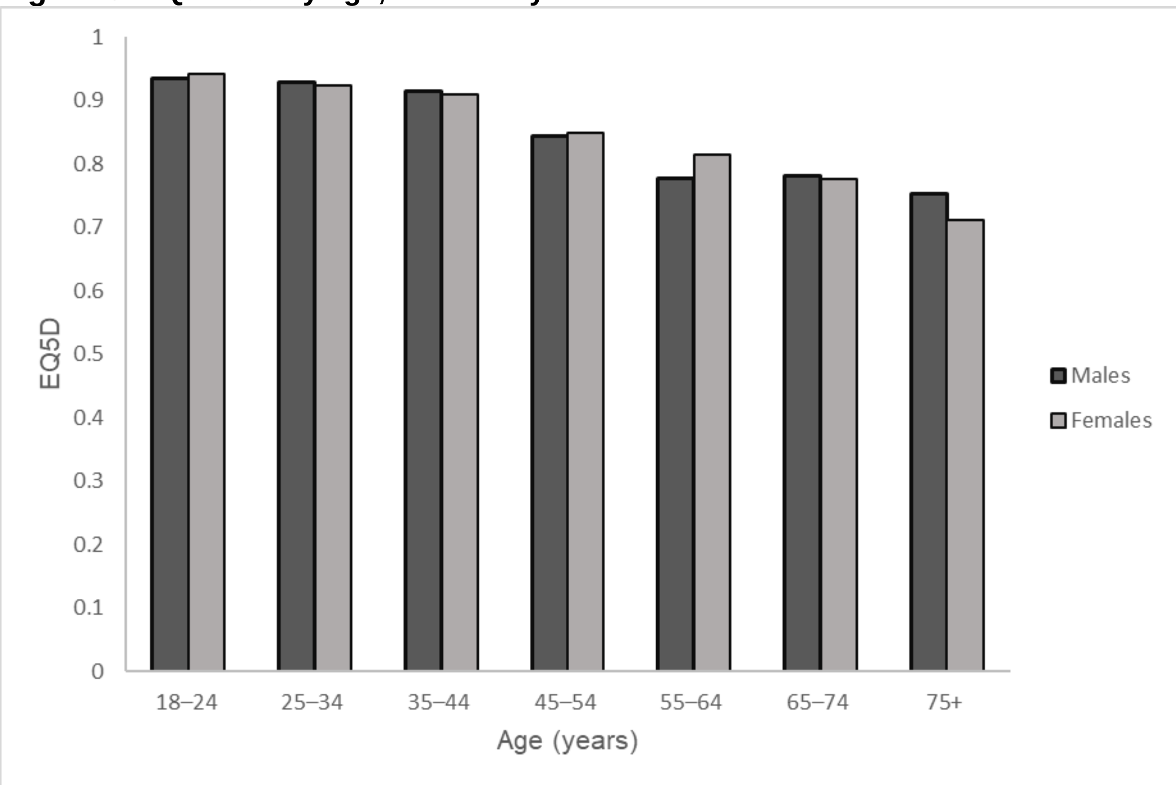
2.3. MEASURING AND VALUING HEALTH STUDY 1998 (REPORTED IN JANSSEN AND SZENDE 2014)

The MVH Study (Kind et al 1998, reported in Janssen and Szende 2014) was the original EQ-5D-3L valuation study conducted in 1993, interviewing 3,395 adults in the UK. Results are reported as the mean by sex and age categories, corresponding to those above for the HSE 2010.

Figure 3 plots the mean EQ-5D-3L by age category, separately for males and females. There is no consistent pattern between sexes. Values are similar up to the age of 55-64yrs at which point females attain substantially higher EQ-5D-3L values (0.815 vs 0.777). Mean EQ-5D-3L amongst males is higher than females in the age category of 75 years and above (0.753 vs 0.712). EQ-5D-3L decreases with age in all categories except for males aged 65-74yrs who have a higher mean EQ-5D-3L than males aged 55-64yrs (0.781 vs 0.777). This could be due to the relatively low numbers of sample responses in these categories.

We refer to these results as MVH 1993.

Figure 3: EQ-5D-3L by age, MVH study 1993



3. HSE 2014 DATASET

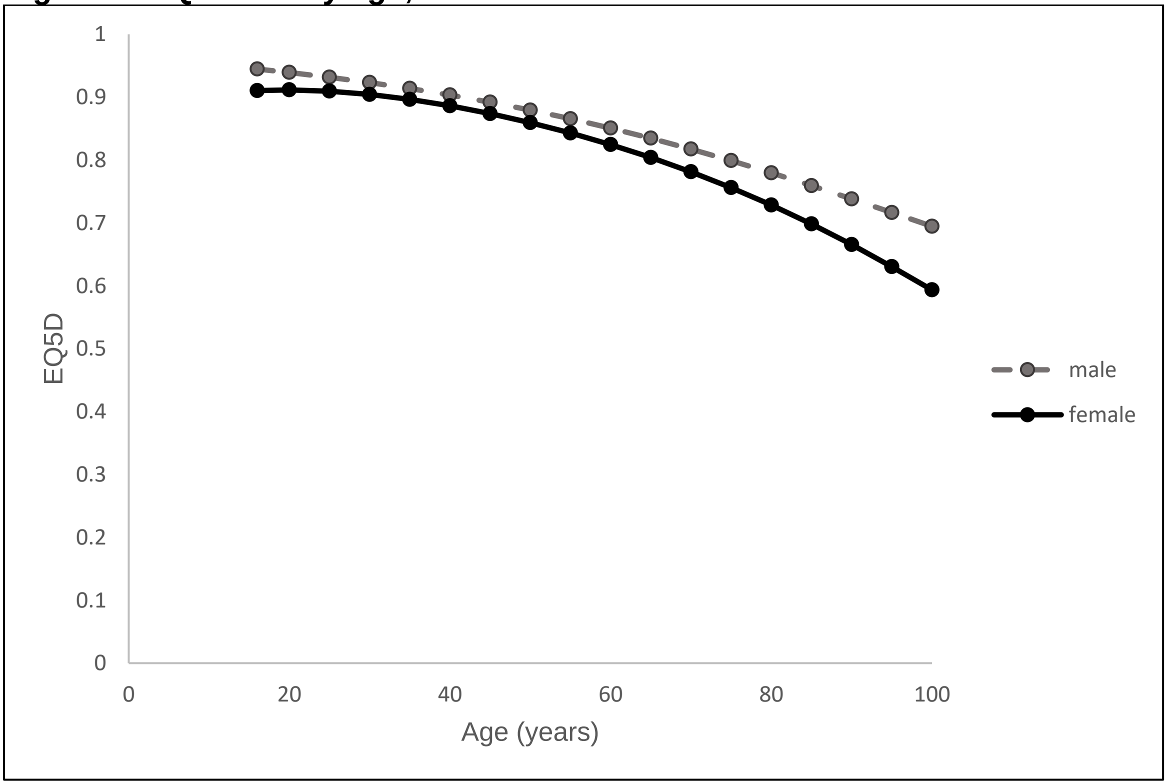
HSE is an annual survey of the English population, focussing on core questions about an individual's health and with modules that focus on specific topics and vary from year to year. There are separate surveys for adults aged 16 and over (the focus of the work presented here) and children. The survey is conducted by face-to-face interview with an additional nurse visit for those that agree. Importantly, sampling is intended to ensure the results are representative of the English population living in private households. The survey implements a clustered, multi-stage stratified sample design. In addition to the sample design variables, weighting variables are included in the dataset to adjust for selection, non-response and the population profile. Estimates provided in this report using this survey take the clustering, stratification and weighting into account.

The latest available wave of the HSE that includes the EQ-5D-3L was published in 2014. 8,077 individuals age 16 and over were included; 992 respondents had one or more EQ-5D-3L dimensions missing reducing the sample size to 7,085 (3,111 males and 3,974 females). To maintain anonymity within small subgroups in the dataset, the HSE has adopted the policy of top-coding age at 90 since the 2014 wave. For most respondents, exact age is provided. However, for those aged 90 years and above, there is a single 90+ category. For this small group (0.57% of the sample) we have assigned an age value of 92, which is the mean age of those aged 90+ in the 2013 HSE sample (which was not subject to top-coding). It is important to deal with this issue correctly, since assumptions made for this extreme group will affect the estimates for the entire population and have implications for any model-based analysis. It is possible to show theoretically that regression-type models give unbiased estimates when grouped variables are imputed with the appropriate conditional mean as we have done here (Manski and Tamer 2002).

We estimate Adjusted Limited Dependent Variable Mixture Models (ALDVMM) separately for males and females to allow for different EQ-5D-3L age profiles. Standard model selection criteria is not available when using complex survey design variables. As a first step we estimated all models assuming random sampling. Then the selected models were re-estimated taking into account the clustering, stratification and weighting. We estimate models with up to four components. Age and age-squared terms were included in each of the components. Models including age as a predictor of component membership were preferred by both Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). Models with four components were difficult to identify and resulted in a large proportion of insignificant coefficients (10 out of 22) clearly indicating model overfitting. Therefore, we select three component models and refit them including clustering, stratification and weighting variables. The final parameter estimates are presented in the appendix.

Figure 4 plots the mean EQ-5D-3L by age implied by these models. Males show higher mean EQ-5D-3L throughout but the difference is smaller in the mid-thirties to mid-forties.

Figure 4: EQ-5D-3L by age, HSE 2014



4. EEPRU DATASET

The EEPRU survey, conducted online by OnePoll in April 2020, was designed as a basis for mapping between the 3L and 5L versions of EQ-5D and included adults aged 18 and over. The survey was not intended to sample from a representative group of the UK population. Responses came from existing UK polling panels which are themselves made up of respondents considered representative of the UK population in measured aspects. For the intended mapping purpose, it was important to obtain observations across the EQ5D severity range. However, for the current purposes, the EEPRU study provides a very large sample (almost 50,000 cases), with data on age, EQ-5D-3L and a number of other contextual variables. It offers another very up-to-date basis for estimation of an age adjustment model. Again, there is a complication in age measurement, since age was collected as a 7-group categorical variable rather than an exact figure. However, for a large majority of the sample (36,445 individuals) it has been possible to recover exact ages.

The EQ-5D-3L by age profile in the EEPRU data is quite different from the profile in HSE data. Figures 4 and 5 show the two profiles, where we have used nonparametric local regression smoothing to ensure that any local details are captured. The HSE age profiles fall fairly steadily with age, while the EEPRU age profile is highly non-monotonic and suggests a slight overall improvement in EQ-5D-3L with age.

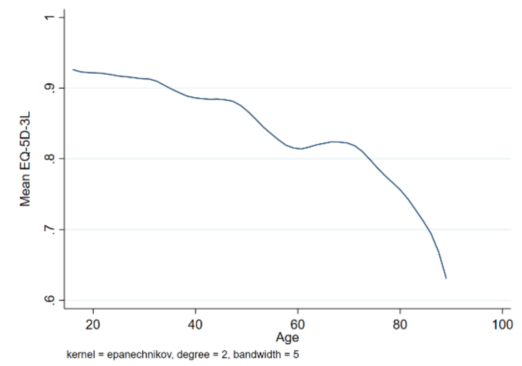


Figure 5 HSE 2014 data: local quadratic regression (5-year smoothing bandwidth)

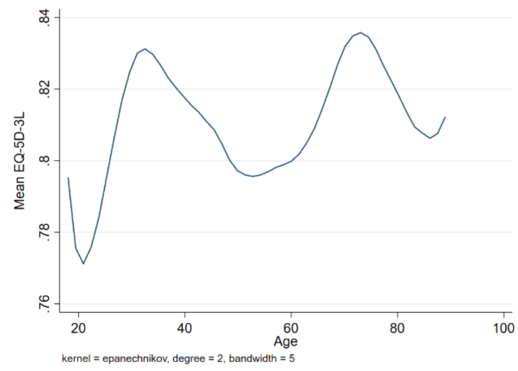


Figure 6 EEPRU data: local quadratic regression (5-year smoothing bandwidth)

There is clearly something interesting at work here. Further analysis confirms that the differences cannot be attributed to differences in sample characteristics between the HSE and EEPRU surveys. The most important difference between the two surveys is timing: HSE fieldwork was conducted in the fairly stable period of 2014, while EEPRU data collection took place during lockdown for Covid-19. Specifically, the fieldwork for the EEPRU data collection was conducted between 1st April 2020 and 1st May 2020. From the 23rd March 2020, individuals were required to stay at home and these restrictions were not lifted until after the data collection period.

The source of the difference in EQ-5D-3L age profiles becomes apparent if we look at the age profiles of responses to the five separate domain questions comprising the EQ-5D-3L descriptive system. Figure 77 shows these for the HSE and they look much as we might expect: the anxiety/depression domain shows no systematic change until a deterioration late in life, whereas all other domains show a fairly steady deterioration through the lifecourse.

Figure 8 shows a broadly similar pattern for mobility, usual activities and pain, but the profiles for the other two domains are quite different – the prevalence of difficulties with self-care and experience of anxiety or depression is particularly high for younger people, declining strongly with age (apart from a deterioration in self-care in old age). Lockdown caused particular difficulties for young adults, many in confined accommodation, and there is compelling evidence that the mental health of young people suffered particularly badly (see for example Fancourt et al 2020) – despite the fact that Covid-19 risks were more serious for older people. The effect of this reversal in the age gradient for self-care and anxiety/depression, when fed through the EQ-5D valuation formula, leads to an unclear relationship with age.

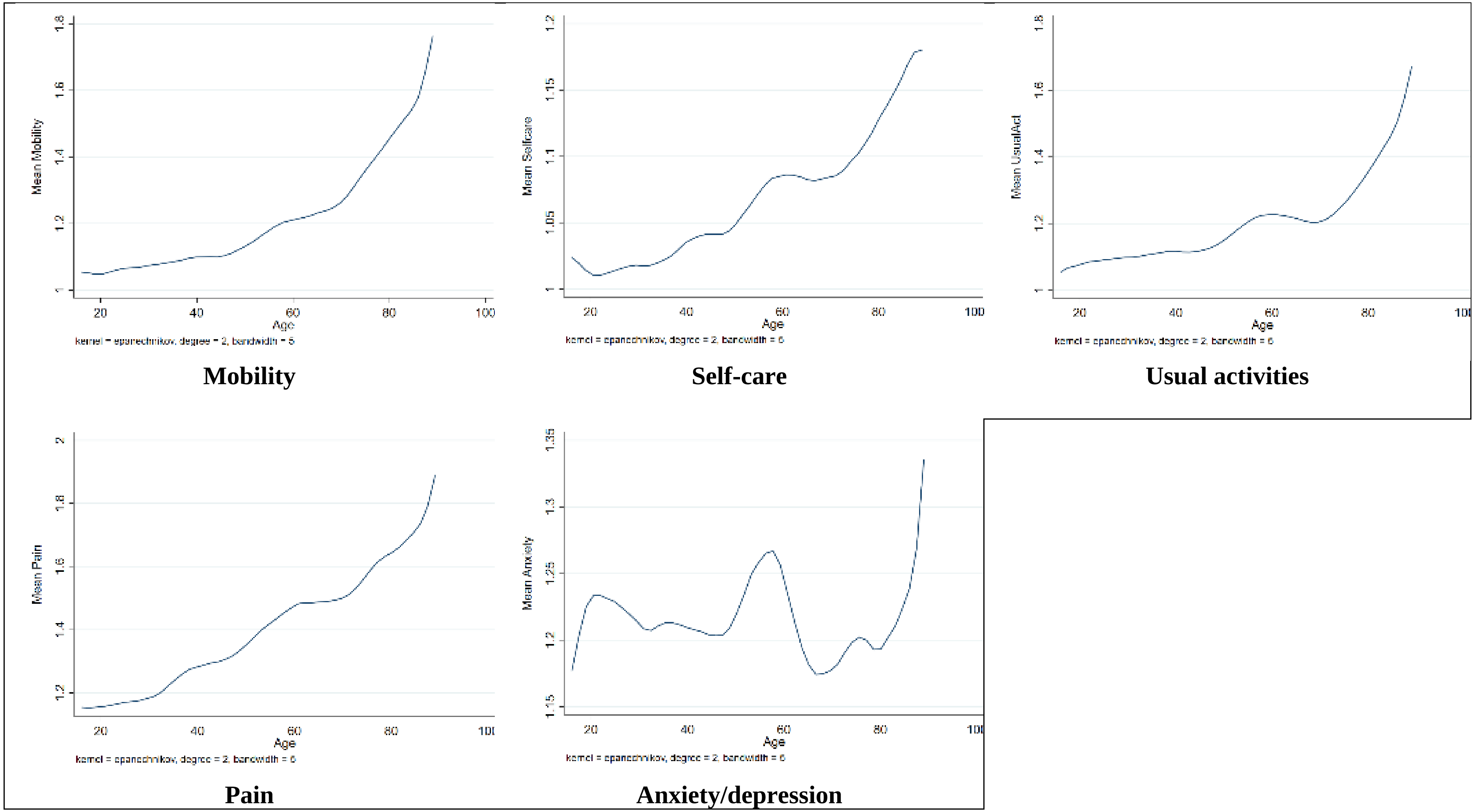


Figure 7 HSE2014 data: smoothed age profiles for EQ-5D-3L domain items

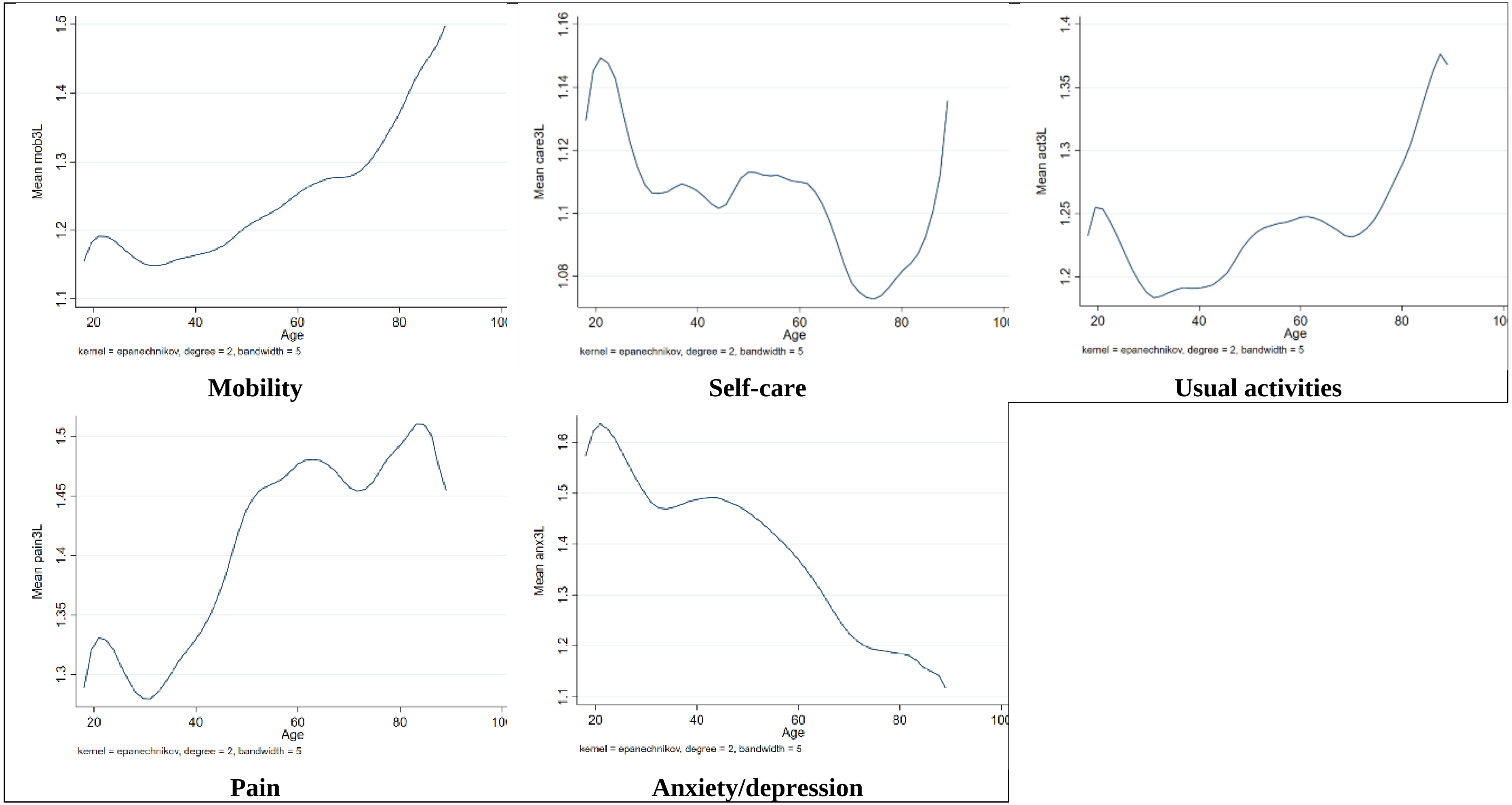


Figure 8 EEPRU data: smoothed age profiles for EQ-5D domain items

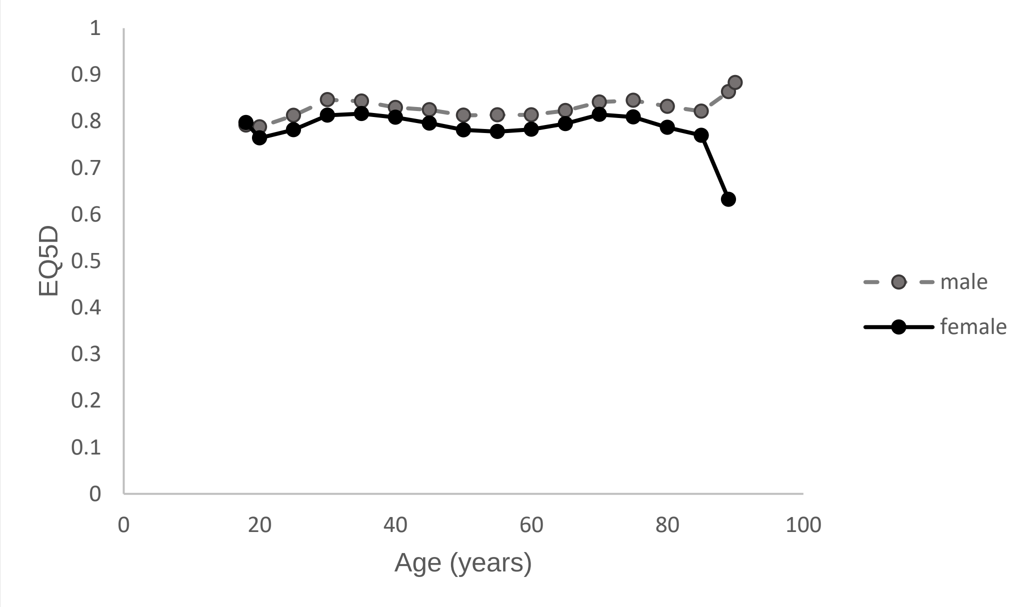
This interesting finding raises some important questions. Can this result be confirmed in other datasets? How long will the Covid-19 effect last? Will the age profile ever return to the form that we observe pre-covid? What recommendations should be made for age adjustment in the meantime?

In our view, the best advice at present would be to use the updated age profile based on HSE 2014, but NICE should be aware that there may be long-term implications of the Covid-19 epidemic and (particularly) the lockdown response to it. There are also potential implications for the interpretation of clinical studies conducted since 2020 where quality of life has been assessed.

With these caveats in mind, we analysed the EEPRU data. The shape of the relationship between EQ-5D-3L and age makes the use of quadratic terms inappropriate. AIC and BIC suggest a minimum of 6^{th} order polynomials on age based on simple regressions to be able to capture the shape of the relationship. Therefore, we instead estimated nonparametric local regression smoothing separately for males and females. The smoothing is calculated at each of the observed ages in the sample (18 to 89 for females and 18 to 90 for males) and we use a local quadratic polynomial in the smoothing with a bandwidth of 5 years.

Figure 9 shows the EQ-5D-3L age profile by sex based on the EEPRU data. Similar to the other datasets, males show a higher mean EQ-5D-3L by age. The shape of the relationships is similar for both sexes but the drop in EQ-5D-3L in the younger group is larger for females and the lines strongly diverge towards the end after age 85 although this might be partly due to the small number of observations in that age group.

Figure 9: EQ-5D-3L by age, EEPRU 2020



5. COMPARISON OF ESTIMATES

5.1. AGE, SEX SCENARIOS

We provide comparisons between different methods in terms of the mean EQ-5D-3L predicted scores for a range of ages and for both males and females in Table 1. The most significant differences occur between the HSE 2003/6 and the EEPRU studies for males. At age 30 the difference is 0.088 and at age 90 rises to 0.203. The HSE 2003/6 study does estimate the lowest scores for males aged 90 and over of all the datasets.

For females, the largest difference in EQ-5D-3L at age 30 years occurs between the MVH and EEPRU studies, a difference of 0.112. At other ages the differences between study results are smaller.

5.2. BURDEN OF ILLNESS

NICE intends to use “severity weights” in its latest Methods Guide (NICE, 2022). Weights in excess of 1 will be assigned to the health benefits generated by health technologies that satisfy particular criteria defined in terms of AS and PS. Both AS and PS are measures of the difference between the number of Quality Adjusted Life Years (QALYs) generated for an individual of the average age of the patient cohort with and without the condition in question. The calculation of both AS and PS relies, in part, on the quality-of-life score assigned to otherwise healthy adults by age. We were able to change this aspect of the calculation. Other parts of the calculation are the source of data for life expectancy by age, the source of utility weights used for current treatment, and the analytical techniques used for these data. We did not change the decision models where the number of QALYs associated with the condition (treated with standard care) was calculated so there is a slight inconsistency introduced. For the sake of demonstrating sensitivity of the estimates to the source used, this is not considered a significant issue.

We calculated both AS and PS for a subset of 364 NICE Technology Appraisal decisions. The focus of this report is not the accuracy of those calculations per se but to demonstrate sensitivity of estimates to the source used for health utility by age. In addition, there is a requirement to safeguard commercially confidential information. We therefore report only summary statistics here.

Table 1: Expected EQ-5D-3L by age and sex

	Males					Females				
	HSE 2003 /6	HSE 2010	MVH	HSE 2014	EEPRU	HSE 2003 /6	HSE 2010	MVH	HSE 2014	EEPRU
Age (yrs)										
30	0.934	0.934	0.930	0.923	0.846	0.913	0.913	0.925	0.904	0.813
50	0.876	0.879	0.845	0.879	0.813	0.855	0.846	0.849	0.859	0.781
70	0.791	0.810	0.781	0.817	0.841	0.770	0.760	0.777	0.781	0.815
90	0.680	0.753	0.753	0.738	0.883	0.659	0.692	0.712	0.672	0.632

HSE = Health Survey for England, MVH = Measuring and Valuing Health, EEPRU – Economic Evaluation Policy Research Unit.

Table 2: Absolute and Proportional Shortfall for a sample of 364 NICE decisions using different methods for calculating utility by age

	HSE 2003 /6		HSE 2010		MVH		HSE 2014		EEPRU	
	AS	PS	AS	PS	AS	PS	AS	PS	AS	PS
Mean	9.39	0.62	9.32	0.62	9.46	0.62	9.64	0.64	9.47	0.64
Median	9.77	0.64	9.75	0.64	9.89	0.64	10.09	0.65	10.08	0.65
SD	4.14	0.27	4.13	0.27	4.13	0.27	4.11	0.27	4.04	0.29
5th percentile	2.57	0.18	2.62	0.18	2.75	0.19	2.88	0.20	2.90	0.18
95th percentile	15.85	0.96	15.74	0.96	15.87	0.96	16.11	0.98	15.43	1.02

HSE = Health Survey for England, MVH = Measuring and Valuing Health, EEPRU – Economic Evaluation Policy Research Unit.

Table 2 shows that neither AS nor PS are sensitive to the source of age adjustment of EQ-5D-3L. The mean AS varies between 9.32 (using the HSE 2010) and 9.64 (using HSE 2014). Mean PS varies between 0.62 and 0.64. All other summary statistics also demonstrate only trivial differences.

6. CONCLUSIONS

Age adjustment of health utility values is commonplace in decision models submitted to NICE. This requirement is retained in the 2022 Methods Guide and the introduction of disease severity weights will provide a further need for such analyses. The calculation of both proportional and absolute shortfalls require an estimate of QALE for the general population of equivalent age to the population being considered in any given appraisal. This in turn requires estimates of health utility by age. Currently there is no recommended source for age weights. Some widely used approaches are based on data that is approaching 30 years old.

This report provides estimates of EQ-5D-3L by age and sex from two sources. The first is the most recent HSE containing EQ-5D-3L which was 2014. The second was a large-scale population survey undertaken by EEPRU.

The HSE showed that EQ-5D-3L declines with age and this is consistent with published studies. The EEPRU dataset shows a less clear relationship with ageing. Further inspection identifies that the “anxiety/depression” and, to a lesser extent, “self-care” domains of the EQ-5D were responsible for this difference. We consider that this is likely to be due to the timing of the fieldwork which coincided with the early stages of the Covid-19 lockdown in the UK. At the present time, there is insufficient information to judge if the impact of Covid-19 has had a permanent impact on the health, particularly mental health, of the general population. Research into this issue would be welcome.

We found that the differences between data sources do not make a significant difference to estimates of AS or PS. Therefore, for the sake of consistency our recommendation to NICE is to use the most up to date information available that has direct observation of EQ-5D-3L from the HSE (2014).

7. REFERENCES

Ara R, Brazier JE. Populating an economic model with health state utility values: Moving toward better practice. Value Health 2010; 13: 509–18.

Fancourt D, Steptoe A, Bu F. (submitted 2020) “Trajectories of anxiety and depressive symptoms during enforced isolation due to Covid-19: longitudinal analyses of 36,520 adults in England”, non peer reviewed pre-print available at: https://www.medrxiv.org/content/10.1101/2020.06.03.20120923v3

Janssen B & Szende A, (2014) “Population norms for the EQ5D”, in Szende A, Janssen B, Cabases J, editors. Self-reported population health: An international perspective based on the EQ5D, 2014, Dordrecht : Springer.

Kind P, Dolan P, Gudex C, Williams A (1998) Variations in population health status: results from a United Kingdom national questionnaire survey. BMJ 316(7133):736–741

Manski, C. F. and Tamer, E. (2002). “Inference on Regressions with Interval Data on a Regressor or Outcome” Econometrica 70: 519-546.

NICE (2013) Guide to the Methods of Technology Appraisal, available at:

https://www.nice.org.uk/process/pmg9/resources/guide-to-the-methods-of-technology-appraisal-2013-pdf-2007975843781

NICE (2022) Guide to the Methods of Technology Appraisal, available at: https://www.nice.org.uk/process/pmg36/chapter/introduction-to-health-technology-evaluation

NICE (2021) Review of methods for health technology evaluation programmes: proposals for change Available at: https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/chte-methods-and-processes-consultation

8. APPENDIX

Estimated EQ-5D-3L models using the HSE 2014 dataset

		Males		Females	
		Coefficient	Std. error	Coefficient	Std. error
Comp. 1	Age/10	-0.1609	0.0425	-0.0774	0.0426
	$(Age/10)^2$	0.0129	0.0037	0.0064	0.0034
	intercept	0.5660	0.1137	0.2990	0.1339
Comp. 2	Age/10	-0.0080	0.0124	-0.0147	0.0114
	$(Age/10)^2$	-0.0007	0.0011	-0.0003	0/0011
	intercept	0.8502	0.0319	0.8708	0.0279
Comp. 3	Age/10	-0.0316	0.1067	0.2043	0.0898
	$(Age/10)^2$	-0.0051	0.0089	-0.0241	0.0087
	intercept	1.8194	0.4700	1.1659	0.3039
Probability	Age/10	0.2804	0.0613	0.4028	0.0612
Comp. 1	intercept	-4.3366	0.3678	-4.4767	0.4012
Probability	Age/10	0.1445	0.0359	0.1937	0.0345
Comp.2	intercept	-1.5277	0.2332	-1.3549	0.2245
	Sigma 1	0.1140	0.0131	0.1282	0.0108
	Sigma 2	0.0694	0.0091	0.0831	0.0055
	Sigma 3	0.5402	0.1305	0.5230	0.1137