Health state values derived from people with multiple sclerosis for a condition-specific preference-based measure: Multiple Sclerosis Impact Scale - Eight Dimensions – Patient version (MSIS-8D-P)

Running title: A Patient Preference Based Measure for MS

Elizabeth Goodwin PhD[1], Colin Green PhD[1,2], Annie Hawton PhD[1,2]

1 Health Economics Group, Institute of Health Research, University of Exeter, Exeter, UK

2 South West Collaboration for Leadership in Applied Health Research and Care (CLAHRC), University of Exeter Medical School, University of Exeter, Exeter, UK

Corresponding author: Elizabeth Goodwin, Health Economics Group, South Cloisters, St Luke's Campus, University of Exeter, Exeter UK EX1 2LU.
e.goodwin@exeter.ac.uk
+44 1392726073

Funding Statement: Financial support for this study was provided in part by the Multiple Sclerosis Society of Great Britain and Northern Ireland. Partial funding was received from the UK NIHR Collaboration for Leadership in Applied Health Research and Care of the South West Peninsula (PenCLAHRC) to CG and AH. The funding agreements ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report. The views expressed in this publication are those of the authors and not necessarily those of the Multiple Sclerosis Society, the UK NIHR or the Department of Health.

Précis: The development of a tariff of health state values, based on the preferences of people with multiple sclerosis, for a condition-specific preference-based measure (the MSIS-8D).

Acknowledgements: Funding for this research was provided by the MS Society of Great Britain and Northern Ireland. The data for this paper was collected by the UK MS Register project (ref: 16/SW/0194) and Accent Marketing Research Ltd. We would like to acknowledge the contribution of all members of the UK MS Register team based at Swansea University Medical School, and all members of the team at Accent.

Word Count: 4000

Number of Pages: 41

Number of Figures: 1

Number of Tables: 4

Appendix:

Pages: 7

Figures: 0

Tables: 1

DRAFT

Health state values derived from people with multiple sclerosis for a condition-specific preference-based measure: Multiple Sclerosis Impact Scale - Eight Dimensions – Patient version (MSIS-8D-P)

Abstract

Objective: In economic evaluation, health outcomes are commonly quantified using quality-adjusted life-years (QALYs) derived from the preferences of a sample of the general population. It can be argued that this approach ignores the preferences of people with experience of the condition, and that patient preferences have a place in the valuation of health outcomes. Here we report the estimation of a preference-based index for an existing condition-specific preference-based measure for multiple sclerosis (MS), the MSIS-8D, based on the preferences of people with MS.

Study design: Internet time trade-off survey, eliciting preferences from people with MS.

Methods: We elicited preferences from a sample of people with MS (n=1635) across 169 MSIS-8D health states, using the time trade-off technique. We fitted ordinary least squares and random effects models to the survey data to estimate values for all health states described by the MSIS-8D.

Results: The new patient-derived index (the MSIS-8D-P) provides values ranging from 0.893 for the best possible health state to 0.138 for the worst state. The MSIS-8D-P exhibits good discriminative validity, identifying expected significant differences between groups based on presence/absence of MS, type of MS and duration since diagnosis.

Conclusions: The MSIS-8D-P index of values for MS-specific health states provides an opportunity to estimate QALYs based on patient preferences, for use in economic evaluations of treatments for MS. More broadly, it adds to the methods and data available to consider the health-related quality of life of people with MS to inform resource allocation and individual-level decisions regarding treatments for MS.

DRAFT

Highlights

What is already known about the topic?
Most commonly, the preferences that are used to estimate quality-adjusted life years (QALYs) are elicited from samples of the general population, rather than from people with the relevant health condition.

What does the paper add to existing knowledge?
This paper provides an alternative tariff of health state values for an existing preference based measure of health outcomes in multiple sclerosis (the MSIS-8D), based on the preferences of people with multiple sclerosis.

What insights does the paper provide for informing health care-related decision making?
This tariff of health state values for the MSIS-8D, based on the preferences of people with multiple sclerosis, can be considered alongside public preferences to provide a broader context for assessing the cost-effectiveness of treatments for multiple sclerosis. This paper also discusses wider issues concerning the common policy guidance on use of public preferences when undertaking economic evaluation.

DRAFT

Introduction

When considering the cost effectiveness of healthcare interventions, the effects of treatment are frequently assessed using quality-adjusted life-years (QALYs). QALYs are calculated by weighting each year of life according to its quality, on a scale from 1 (equivalent to full health) to zero (equivalent to being dead). This combines the impact of treatment on length and quality of life into one measure. QALY weights are generally estimated by eliciting preferences between health states from a sample of the general population, or from a sample of people with the condition that the intervention is designed to address (hereafter referred to as "patients"), using a preference elicitation technique [1]. These techniques enable preferences to be quantified, thereby producing the quality weights, or health state values (HSVs), required for the calculation of QALYs [2]. QALYs are commonly based on preference-based measures (PBMs) of health-related quality of life (HRQL), which use a standardised classification system for describing health states and a tariff of quality weights for all health states described by the classification system. The most commonly used PBMs, including the Euroqol EQ-5D, are designed to be generic, ie suitable for any health condition, although a growing number of condition-specific PBMs are becoming available.

Most commonly, preferences are elicited from members of the general population, and this approach is specifically recommended in most policy settings, eg by the National Institute for Health and Care Excellence[3] and the Panel on Cost-Effectiveness in Health and Medicine [4,5]. However, this approach ignores the preferences of people with experience of the condition, and it has been argued that patient preferences have a place in the valuation of health [6,7]. The availability of patient preferences is all the more salient given the evidence of significant differences between public- and patient-derived HSVs [8]. Within a UK health policy context that is increasingly patient centred, with initiatives such as "No decision about me, without me" [9], we believe it is relevant and timely to review the role of patient preferences in economic evaluation. While some discussions represent a dichotomy between public or patient values, others suggest using values from both perspectives [4,6,7,10]. The common practice of reporting sensitivity analyses alongside basecase cost effectiveness results provides an opportunity for achieving this, in the context of reimbursement considerations, resource allocation decisions, and individual-level treatment decisions [8].

In previous research, we developed an MS-specific PBM, the eight dimension Multiple Sclerosis Impact Scale (MSIS-8D), with QALY weights based on the preferences of the UK general population [11]. Here, we aim to provide an alternative tariff of QALY weights for the MSIS-8D, based on the preferences of people with MS, for potential use across this spectrum of decision making.

Materials and methods

Multiple sclerosis (MS) is a chronic inflammatory condition affecting the central nervous system. In the majority of cases, people with MS experience exacerbations of symptoms (relapses) interspersed with periods of total or partial remission, before developing a progressive disease course. In around 10-15% of cases, the disease is progressive from onset [12]. Symptoms vary widely and can include physical, psychological and cognitive effects [13]. Here we describe the methods employed in conducting a valuation survey with a sample of people with MS and in undertaking analysis to estimate HSVs.

DRAFT

Health state descriptions (the MSIS-8D)

Health states for MS were described using the MSIS-8D, which was developed in response to concerns about the content validity [13-16] and sensitivity [13, 17-19] of generic PBMs in the context of MS. The MSIS-8D descriptive system was derived from the Multiple Sclerosis Impact Scale (MSIS-29), a well validated and frequently used patient-reported outcome measure for MS; this is described in detail elsewhere [20]. In summary, it represents eight dimensions of importance to the HRQL of people with MS: physical functioning, mobility, social activities, daily activities, fatigue, cognitive function, emotional well-being and depression. The original items of the MSIS-29 were primarily based on qualitative work with people with MS, alongside expert opinion and a literature review [21]. The eight dimensions covered by the MSIS-8D descriptive system were informed by a number of previous studies that used qualitative techniques to explore HRQL among people with MS [20]. Each dimension is represented by one MSIS-29 item with four response levels: not at all, a little, moderately and extremely. This constitutes a descriptive system (Figure 1) that describes 65,536 unique MS health states.

A tariff of HSVs for the MSIS-8D has been estimated previously [11]), based on preferences elicited from a representative sample of the UK general population for a sample of 169 MSIS-8D health states, which was selected using the Rasch vignette approach [2] to reflect states that are likely to be experienced by people with MS at different levels of severity. The same health states are used for the current study.
Figure 1 about here

Valuation survey

The valuation survey followed the protocol used to obtain MSIS-8D values from the general population, which was based on the Measurement and Valuation of Health (MVH) version of the time trade-off (TTO) technique. The MVH protocol was developed to generate the UK tariff of preference weights for the EQ-5D-3L [22,23]. Respondents are presented with a choice between two hypothetical scenarios: living in a suboptimal health state for a given number of years or living in perfect health for a shorter period of time. The length of time spent in perfect health is varied until the respondent is indifferent between the two scenarios. In this way, HSVs are determined by asking respondents to trade between quality and length of life [22]. The survey was administered via the internet. To ensure its suitability for the target population, people with MS were involved in developing the survey protocol. We used pre-pilot testing and an online pilot (n=55) with people with MS to finalise the valuation methods.

Prior to undertaking the TTO tasks, participants completed the MSIS-8D descriptive system for their own health to familiarise themselves with the descriptive system. As warm-up exercises, participants were asked to rank three MSIS-8D health states in order of preference and to complete a practice TTO exercise with detailed instructions. Participants were then asked to value six MSIS-8D health states. Each set of health states was stratified to include five health states covering a range of severity plus the worst possible health state (the "pits" state). Each participant was randomly assigned a set of health states. Previous MSIS-8D surveys had demonstrated that this represented an acceptable workload for participants [11,24]; this was confirmed during pilot testing with people with MS.

DRAFT

Ethical approval was obtained from the University of Exeter Medical School Ethics Committee.

Sample of people with MS

Respondents were sourced from the UK Multiple Sclerosis Register (the MS Register). The MS Register was launched in May 2011 and had 13920 members at 18th July 2016 [25]. Members are requested to complete a range of patient-reported outcome measures, including the MSIS-29, via an internet portal every three months. Other available data include socio-demographic and clinical characteristics. Initial analysis indicates that members are broadly representative of people with MS in the UK in terms of key characteristics including gender, age at onset and MS type [26,27].

Invitation emails were circulated to all current members of the MS Register during March 2016. Our target sample size was 1500 respondents, in order to generate approximately 40 observations per health state, based on a review of the literature describing the valuation of health states for condition-specific PBMs [28] with an adjustment to allow for the increased variance that may result from internet administration [29]. In previous valuation surveys, this was sufficient for the estimation of a regression model to predict HSVs for the MSIS-8D [11].

Data cleaning and descriptive analysis

When estimating a tariff for a PBM, it is common practice to exclude data from respondents who provide responses that are internally inconsistent or illogical. This study adopted the exclusion criteria developed for previous MSIS-8D surveys, which were based on the condition-specific PBM development literature [11]. During the practice TTO exercise, respondents were screened out of the survey if they considered ten years in full health to be worse than or equivalent to ten years with health problems, or considered ten years in full health to be worse than or equivalent to being dead. Following data collection, respondents were excluded from the analysis if they:

- gave the same value to all health states (unless they valued all health states as equivalent to full health),
- gave all states a value less than or equal to zero,
- valued the pits state at least as highly as all other states,
- gave the least severe state a lower value than all other states, or
- provided three or more inconsistent responses with a difference in HSV of at least 0.1 ie they valued a dominated health state as better than a logically better alternative by the equivalent of one year in the TTO exercise.

Negative HSVs (ie health states considered to be worse than being dead) were transformed onto a scale from 0 to -1 following the method used to estimate the UK tariff for the EQ-5D-3L [23].

Modelling to obtain health state values

HSVs for all MSIS-8D health states were estimated using the standard regression model [2]:

$$h_{ij} = f(\beta' X_{\lambda\partial}) + \varepsilon_{ij}$$

where h_{ij} represents the TTO value; i represents individual health states; j represents individual respondents; f represents the functional form; X represents a vector of dummy

explanatory variables for each level λ of dimension ∂ of the classification system, where level $\lambda = 1$ acts as a baseline; and ε_{ij} represents the error term.

We estimated individual-level and mean-level ordinary least squares (OLS) models, fixed or random effects (RE) models to account for the clustering of data by respondent, and RE Tobit models to allow for censoring of HSV data between -0.975 and 1 [30]. Any inconsistent coefficients, where a less severe item-level resulted in a greater HRQL decrement than a more severe item-level, were merged and the analysis was re-run to produce a consistent model. Additional versions of these models were generated by merging adjacent item-levels represented by coefficients that were non-significant at the 95% level [31]. We did not assume that the best MSIS-8D health state represents perfect health, therefore the constant was not constrained to unity [28].

Models were compared in terms of the proportion of coefficients that were significant, mean absolute error (MAE) of predicted HSVs and the number of health states with absolute errors greater than 5% and 10% (equivalent to six months and one year in the TTO exercise respectively) [2]. This enabled selection of a preferred model to generate an index of HSVs for the MSIS-8D based on the preferences of people with MS. Here we do not compare the preferences of people with MS and those of the general public. A detailed examination of this comparison is reported in a companion paper [32].

Discriminative validity

The sensitivity of the MSIS-8D-P index was assessed by exploring its discriminative validity [2]. Empirical data suggest that progressive types of MS have a greater impact on HRQL than relapsing-remitting MS and that the HRQL of people with MS decreases over time, therefore we would expect these differences to be reflected in HSVs [33,34]. Independent t-tests were used to determine the ability of the MSIS-8D-P index to distinguish between sub-groups of respondents to the valuation survey, based on type of MS and duration since diagnosis. Additional analysis was undertaken using MSIS-8D responses from previous surveys of the general population [11,24] to assess the ability of the MSIS-8D-P index to distinguish between people with and without MS.

Analyses were undertaken using Microsoft Excel and Stata.

DRAFT

Results

Valuation survey

In total, 3565 members of the MS Register entered the website, of whom 1635 (46%) completed the survey. Of these, 39 (2.39%) provided inconsistent or illogical responses and were excluded from the analysis. No differences were apparent in the characteristics of excluded and included respondents.
Table 1 presents the socio-demographic and clinical characteristics of the 1596 respondents who were included in the analysis compared to those of all MS Register members (for whom data was available) at the time the survey was closed. The characteristics of the survey respondents reflected those of the MS Register members overall in terms of age, gender, employment status, type and duration of MS and self-reported health status, although a higher proportion of respondents had a university education. Table 1 also reports respondents' views of task comprehension and difficulty. The majority (84%) reported that they found the TTO questions easy or very easy to understand. Nearly half of respondents (47%) found it difficult to choose between the scenarios, but only 7% said they found this very difficult.

The mean observed HSVs for the 169 health states included in the survey, presented in the Appendix, ranged from 0.15 for the pits state to 0.94 for the best state. The mean number of observations per health state was 48. The distribution of individual observed HSVs over the full possible range of values reflected the left skew and clustering at 0 and 1 that are typical of TTO data [30].

The 169 health states were assigned to 24 severity groups, based on the sum of response-levels across all dimensions [35]. The average HSV for each severity group is shown in Table 2. The pattern is consistent with the expected direction of preferences, ie mean observed HSVs decrease as severity increases, with the exception of three discrepancies (group 2, group 4 and group 12).

Modelling health state values
The Hausman test yielded a non-significant result ($chi^2(24) = 14.05$; $p = 0.95$), indicating that a fixed effects specification would produce a similar result with reduced efficiency, therefore a random effects specification was used [36]. The individual-level OLS and RE models each had one coefficient (corresponding to level 2 of the Fatigue dimension) with an inconsistent sign, and the aggregate-level OLS model produced three coefficients with an inconsistent sign (for Mobility levels 2 and 3 and for Fatigue level 2). Consistent versions of these models were created by merging the affected levels.
Table 3 summarises the consistent individual OLS, aggregate-level OLS and RE models and the original Tobit model. All models had coefficients that were consistent with expected preferences, ie for each dimension of the MSIS-8D, coefficient values decreased as the level of severity increased. The consistent individual-level OLS performed the least well, with the highest MAE (0.0469), number of errors over 0.1 (17) and number of errors over 0.05 (63). Just over one third (35%) of coefficients were significant.

The consistent mean-level OLS and RE models performed similarly well. The Tobit model performed better in terms of significant coefficients (58.33%), however it had a slightly higher MAE (0.0391) and number of errors over 0.1 (7).

DRAFT

In order to avoid the loss of information that occurs when data is aggregated, the mean-level OLS model was not considered further. The remaining RE and Tobit models, however, had a relatively low proportion of significant coefficients. Therefore, various options for merging affected item-levels were explored. The best-performing of these parsimonious models (RE Version 2 and Tobit Version 2) are presented in Table 3. Results for all other models are available from the authors on request.

Selection of preferred model

Overall, RE Version 2 had superior predictive ability, with fewer errors >0.1 and a lower MAE than Tobit Version 2. On this basis, the parsimonious RE Version 2 is the recommended model for estimation of a single index for the MSIS-8D based on the preferences of people with MS. The index ranges from 0.893 for the best MSIS-8D health state to 0.138 for the pits state. This indicates that respondents considered the best state to be less than perfect health, ie they assumed decrements in HRQL beyond the dimensions included in the classification system. This is not unusual for a condition-specific PBM [28].

This preferred model, as presented in Table 3, enables a HSV to be calculated for any MSIS-8D health state, by summing the constant and the coefficient for each item depending upon its level. For example, the predicted value for the MSIS-8D health state (3,3,2,3,4,2,2,1) is calculated as:
Constant + Physical(3) + Social(3) + Mobility(2) + Daily activities(3) + Fatigue(4) + Emotion(2) + Cognition(2) + Depression(1)
= 0.893 - 0.065 - 0.032 - 0.003 - 0.020 - 0.063 - 0.015 - 0.027 + 0 = 0.668

Discriminative validity

Table 4 presents data describing the discriminative ability of the MSIS-8D-P. The results of the t-tests provided strong evidence that the index is capable of discriminating between sub-groups of people with MS that would be expected to differ in terms of their HRQL ($p<0.0001$). Significantly lower HSVs were observed for those with progressive rather than relapsing-remitting MS, and for those with a disease duration of ten or more years since diagnosis. A large, significant difference was also observed between survey respondents with and without MS.

DRAFT

Discussion

We have elicited preferences from people with MS to derive an alternative tariff of HSVs for an existing health state classification system. We refer to this as the MSIS-8D-P (Multiple Sclerosis Impact Scale – 8 Dimensions – People with MS). This provides an additional source of information about the impact of MS and its treatment on HRQL, alongside the original tariff of MSIS-8D values, which is based on the preferences of the general population. Both tariffs are suitable for use across all types of MS, to assess HRQL and to estimate QALYs, and can be derived directly from patient-reported responses to the MSIS-29, a well validated and frequently used patient-reported outcome measure for MS. The methods employed in this study are well-accepted for the generation of HSVs for condition-specific descriptive systems [2]. In terms of the mean absolute error and the proportion of health states with a prediction error greater than 0.05 or 0.1, the preferred model for the estimation of MSIS-8D-P values compares favourably with models that have been developed to estimate tariffs for other condition-specific PBMs [28]. The MSIS-8D-P exhibits good discriminative validity, suggesting that it is sensitive to differences and changes in the HRQL of people with MS.

There are two potential uses for the MSIS-8D-P: providing QALY weights from the perspective of people with MS to inform economic evaluations and providing a source of data on the HRQL of people with MS to inform condition-specific resource allocation and individual-level treatment decisions. In addition, the availability of tariffs for the MSIS-8D classification system from people with MS and from the general population enables a full comparison to be drawn between public and patient values for MS health states. This analysis is reported in a companion paper.

Strengths and limitations of study

The results of the questions regarding self-reported task comprehension, and the nature of the preference data gathered, indicate that it is possible to administer a complex technique such as the TTO via the Internet to people with a chronic condition such as MS, which can, in some cases, affect cognitive functioning. The direct involvement of people with MS in reviewing the survey protocol, the careful construction of instructions and warm-up tasks, and the pre-pilot and pilot tests were instrumental in ensuring that the survey was appropriate for the target population.

Although the approach taken in this study was informed by best practice guidance on the development of condition-specific PBMs [2], it has some limitations. In keeping with national guidelines [3], HSVs were elicited using the MVH version of the TTO. This asks respondents to imagine remaining in a specified health state for ten years, with no changes in that health state during that time. However, MS is usually characterised by alternating periods of relapse and remission, or by ongoing progression [12]. This may have caused confusion for respondents with MS and may have affected the values they attributed to health states.

The proportion of coefficients that were significant in the initial RE model was relatively small compared to other models that have been estimated to predict HSVs for condition-specific PBMs. As a result, we merged four pairs of adjacent dimension-levels to produce the preferred model (along with an additional pair that was merged to address one coefficient with an unexpected sign, which is not unusual for models of this type) [28]. This increased

DRAFT

both the number and the proportion of coefficients that were significant. The levels that were merged and the coefficients that remained non-significant in the preferred model indicate that the preferences of people with MS were not sensitive to shifts from level 1 (not at all) to level 2 (a little) for the Social, Mobility, Daily Activities, Emotion and Depression dimensions, or to shifts from level 2 (a little) to level 3 (moderately) for the Mobility, Daily Activities and Fatigue dimensions.

Patient or public values?

Two main arguments have been put forward to support the use of patient preferences in economic evaluation. The first rests on the theory of welfare economics, which posits that the well-being of a society equals the sum of the utilities of its individual members. This implies that it is more appropriate to base decisions regarding public programmes on the preferences of those set to gain or lose directly from the decision, rather than a wider sample, many of whom will be unaffected. The second is more prosaic: patients are likely to have more experience of poor health and are hence better placed to value how this affects quality of life. Conversely, it is argued that societal preferences should guide resource allocation in order to reflect the views of those who are funding the service [37], while some have expressed concerns that strategic bias may be introduced into HSVs if patients attempt to maximise the possibility of treatments being considered cost-effective [38]. Furthermore, there is debate over whether the differences between public and patient values represent a better understanding of the impact of health states by those who have experienced them, or less desirable effects of ill health on how people assess their situation such as distortions in HSVs or negative forms of adaptation (eg people failing to recognise how poor their current health is, what full health feels like and what it would allow them to do, or lowering their expectations) [10].

There are arguments for and against the use of patient values to inform resource allocation decisions, however neither approach can claim superior theoretical or empirical validity, and the choice between the two is likely to affect which types of intervention are considered cost-effective [7]. It has therefore been suggested that cost-effectiveness results based on patient preferences should be used in conjunction with results based on the preferences of the general population [4,6,7,10]. The MSIS-8D-P provides the information required to apply this approach in the context of MS.
The use of condition-specific PBMs is not limited to informing the allocation of resources across whole healthcare systems. It has been suggested that, while public preferences are better suited to system-wide decision-making, patient values are more appropriate for informing condition-specific resource allocation and individual-level treatment decisions [4]. A recent systematic review of the literature describing the development of condition-specific PBMs [28] identified 21 instruments with tariffs based on patient preferences. Of these, 18 were specifically designed to inform individual or clinical-level decision-making, rather than system-wide resource allocation. Such measures provide useful information about the factors that influence patients' experiences of living with disease and that inform their decisions between treatment alternatives, and the relative importance of these factors.

Conclusion

This new tariff of HSVs for the MSIS-8D, based on the preferences of people with MS, provides an additional source of information to assess the impact of MS and treatments for people with MS, providing a broader context for resource allocation decision-making. We

DRAFT

recommend that such information should be used to determine the impact of using QALYs based on public preferences versus QALYs derived from preferences of those with experience of the condition, and to consider more broadly the common policy guidance on use of public preferences when undertaking economic evaluation. The MSIS-8D-P may also help to inform resource allocation within ring-fenced budgets for MS and individual-level treatment decisions.

DRAFT

Figure 1: The MSIS-8D classification system

In the *past two weeks*, how much has your MS limited your ability to ...	Not at all	A little	Moderately	Extremely
Do physically demanding tasks?	1	2	3	4

In the *past two weeks*, how much have you been bothered by ...	Not at all	A little	Moderately	Extremely
Limitations in your social and leisure activities at home?	1	2	3	4
Being stuck at home more than you would like to be?	1	2	3	4
Having to cut down the amount of time you spent on work or other daily activities?	1	2	3	4
Feeling mentally fatigued?	1	2	3	4
Feeling irritable, impatient or short-tempered?	1	2	3	4
Problems concentrating?	1	2	3	4
Feeling depressed?	1	2	3	4

DRA

Table 1: Characteristics of respondents to the preference elicitation survey

Characteristic	MSIS-8D survey sample Number	Percentage	MS Register members
Gender			
Female	1,145	73%	72%
Male	424	27%	28%
Age group			
25 and under	7	0%	1%
26 to 35	85	5%	7%
36 to 45	304	19%	20%
46 to 55	504	32%	30%
56 to 65	463	30%	27%
Over 65	205	13%	15%
Employment status			
Economically active	633	41%	39%
Economically inactive	912	59%	61%
Highest level of education			
University	658	43%	33%
Occupational	464	30%	34%
Compulsory	298	19%	26%
Other	125	8%	7%
Type of MS			
RRMS	745	49%	51%
SPMS	394	26%	25%
PPMS	241	16%	14%
Benign	79	5%	5%
Unknown	55	4%	5%
Time since diagnosis			
Under 10 years	602	41%	40%
10 to 19 years	530	36%	36%
20 years and over	330	23%	24%
Respondents' self-reported raw scores on the MSIS-8D			
MSIS-8D total score			
Mild (score 8-16)	441	28%	
Moderate (score 17-24)	686	43%	
Severe (score 25-32)	468	29%	
MSIS-8D physical score			
Mild (score 4-8)	443	28%	
Moderate (score 9-12)	514	32%	
Severe (score 13-16)	638	40%	
MSIS-8D psychological score			
Mild (score 4-8)	652	41%	
Moderate (score 9-12)	616	39%	
Severe (score 13-16)	327	21%	
Respondents' self-reported task comprehension			
What were the questions like to understand?			
Very easy	391	24.50%	
Easy	946	59.27%	

DRAFT

Difficult	239	14.97%
Very difficult	20	1.25%
How easy or difficult was it to make choices between the options you were asked to think about?		
Very easy	135	8.46%
Easy	588	36.84%
Difficult	755	47.31%
Very difficult	118	7.39%

DRAFT

Table 2: Mean health state values by severity group

Severity group	Total score	Mean	SD	Min	Max	Obs	Number of health states
0	8	0.943	0.150	0	1	54	1
1	9	0.882	0.203	0	1	487	8
2	10	0.843	0.214	0.025	1	118	3
3	11	0.853	0.216	-0.500	1	203	4
4	12	0.794	0.262	-0.975	1	247	5
5	13	0.819	0.233	-0.2	1	280	6
6	14	0.804	0.246	-0.275	1	333	7
7	15	0.801	0.237	-0.975	1	391	8
8	16	0.730	0.290	-0.975	1	463	9
9	17	0.713	0.285	-0.725	1	421	10
10	18	0.675	0.342	-0.900	1	434	10
11	19	0.648	0.335	-0.825	1	448	10
12	20	0.618	0.362	-0.925	1	520	11
13	21	0.627	0.368	-0.975	1	491	10
14	22	0.587	0.363	-0.975	1	434	9
15	23	0.545	0.395	-0.825	1	391	8
16	24	0.490	0.420	-0.975	1	345	8
17	25	0.451	0.419	-0.825	1	377	8
18	26	0.415	0.443	-0.975	1	322	7
19	27	0.405	0.430	-0.975	1	235	6
20	28	0.348	0.461	-0.975	1	253	5
21	29	0.339	0.478	-0.900	1	196	4
22	30	0.287	0.488	-0.975	1	131	3
23	31	0.157	0.486	-0.975	1	406	8
24	32	0.146	0.480	-0.975	1	1596	1

SD = standard deviation; min = minimum observed value; max = maximum observed value; obs = observations

DRAFT

Table 3: Regression models for the estimation of health state values

	Consistent individual OLS		Consistent mean OLS		Consistent RE model		Tobit model		Preferred model: RE Version 2		Tobit Version 2	
	Coeff	p	Coeff	p	Coeff	p	Coeff	p	Coeff	p	Coeff	p
Physical												
A little	-0.034	0.006	-0.037	0.045	-0.040	0.008	-0.072	0.000	-0.047	0.000	-0.080	0.000
Moderately	-0.036	0.082	-0.040	0.098	-0.042	0.052	-0.088	0.001	-0.065	0.000	-0.092	0.000
Extremely	-0.147	0.000	-0.151	0.000	-0.151	0.000	-0.225	0.000	-0.175	0.000	-0.230	0.000
Social												
A little	-0.022	0.173	-0.022	0.228	-0.025	0.108	-0.032	0.089			-0.037	0.030
Moderately	-0.047	0.077	-0.044	0.069	-0.051	0.027	-0.070	0.009	-0.032	0.019	-0.083	0.000
Extremely	-0.071	0.039	-0.077	0.009	-0.086	0.001	-0.115	0.000	-0.067	0.001	-0.128	0.000
Mobility												
A little	-0.001	0.935			-0.003	0.820	-0.006	0.752	-0.003	0.856	-0.006	0.716
Moderately	-0.001	0.960			-0.017	0.449	-0.019	0.462				
Extremely	-0.084	0.018	-0.079	0.000	-0.092	0.001	-0.097	0.001	-0.077	0.000	-0.084	0.000
Daily activities												
A little	-0.012	0.385	-0.010	0.586	0.000	0.996	-0.009	0.629				
Moderately	-0.035	0.172	-0.032	0.190	-0.013	0.568	-0.029	0.267	-0.020	0.132	-0.027	0.077
Extremely	-0.064	0.063	-0.065	0.029	-0.039	0.135	-0.053	0.084	-0.048	0.015	-0.051	0.022
Fatigue												
A little							-0.003	0.859			-0.003	0.842
Moderately	-0.034	0.068	-0.037	0.040	-0.020	0.206	-0.033	0.184	-0.021	0.137		
Extremely	-0.077	0.005	-0.086	0.001	-0.062	0.004	-0.089	0.003	-0.063	0.002	-0.060	0.010
Emotion												
A little	-0.017	0.173	-0.016	0.260	-0.016	0.187	-0.033	0.048	-0.015	0.203	-0.034	0.041
Moderately	-0.031	0.174	-0.030	0.197	-0.042	0.034	-0.060	0.014	-0.042	0.035	-0.077	0.001
Extremely	-0.049	0.165	-0.052	0.090	-0.070	0.008	-0.089	0.005	-0.069	0.009	-0.106	0.000

DRAFT

Cognition												
A little	-0.027	0.054	-0.029	0.088	-0.027	0.058	-0.028	0.106	-0.027	0.030	-0.028	0.104
Moderately	-0.055	0.022	-0.053	0.033	-0.052	0.013	-0.058	0.018	-0.052	0.008	-0.057	0.019
Extremely	-0.107	0.002	-0.102	0.001	-0.115	0.000	-0.121	0.000	-0.116	0.000	-0.120	0.000
Depression												
A little	-0.006	0.706	-0.001	0.968	0.000	0.974	-0.016	0.341			-0.030	0.047
Moderately	-0.044	0.102	-0.040	0.106	-0.041	0.050	-0.065	0.006	-0.040	0.008	-0.079	0.000
Extremely	-0.166	0.000	-0.170	0.000	-0.141	0.000	-0.156	0.000	-0.140	0.000	-0.168	0.000
Constant	0.902	0.000	0.902	0.000	0.894	0.000	1.089	0.000	0.893	0.000	1.092	0.000
Model performance												
Coefficients		23		22		23		24		19		21
Sig coefficients		8 (34.78%)		10 (47.62%)		11 (47.83%)		14 (58.33%)		15 (78.95%)		17 (80.95%)
Mean absolute error		0.0469		0.0349		0.0361		0.0391		0.0364		0.0399
No of errors > 0.1		17		2		2		7		3		8
No of errors > 0.05		63		46		50		49		52		51
Obs (respondents)		9576 (1596)		169 (1596)		9576 (1596)		9576 (1596)		9576 (1596)		9576 (1596)
Wald chi2		NA		NA		8897.14 (23)		9305.22 (24)		8893.62 (19)		9300.19 (21)
Prob > chi2		NA		NA		<0.001	<0.001	<0.001		<0.001		<0.001
Overall R-sq		NA		NA		0.3014		NA		0.3013		NA
Log likelihood		NA		NA		NA		-3846.31		NA		NA
F		185.91 (24, 9551)		133.94 (21, 147)		NA		NA		NA		NA
Prob>F		<0.001		<0.001		NA		NA		NA		NA
R-sq		0.3017		0.9503		NA		NA		NA		NA
Adj R-sq		NA		0.9432		NA		NA		NA		NA
RMSE		0.3767		0.0469		NA		NA		NA		NA

DRAFT

Table 4: Discriminative validity of the MSIS-8D-P

		Mean	SD	Frequency	t-statistic	p-value
Disease status	No MS	0.766	0.172	3490	28.931	<0.0001
	MS	0.613	0.186	1635		
Duration of MS	Under 10 yrs	0.645	0.187	612	4.943	<0.0001
	10 yrs or over	0.597	0.182	882		
MS type	Relapsing	0.666	0.177	760	-12.651	<0.0001
	Progressive	0.547	0.175	652		
SD = standard deviation						

DRAFT

References

1. Weinstein MC, Torrance G, McGuire A. QALYs: The basics. Value Health. 2009;12:S5-S9.
2. Brazier JE, Rowen D, Mavranezouli I, et al. Developing and testing methods for deriving preference-based measures of health from condition-specific measures (and other patient-based measures of outcome). Health Technol Assess. 2012;16:1-114.
3. NICE. Guide to the methods of technology appraisal 2013. National Institute for Health and Care Excellence (NICE), 2013. http://www.nice.org.uk/article/pmg9/chapter/foreword. Accessed 23rd October 2017.
4. Gold MR, Siegel JE, Russell LB, Weinstein, MC. Cost-effectiveness in health and medicine. New York: Oxford University Press, 1996.
5. Sanders GD, Neumann PJ, Basu A, et al. Recommendations for Conduct, Methodological Practices, and Reporting of Cost-effectiveness Analyses Second Panel on Cost-Effectiveness in Health and Medicine. JAMA. 2016;316:1093-103.
6. Nord E, Pinto JL, Richardson J, Menzel P, Ubel P. Incorporating societal concerns for fairness in numerical valuations of health programs. Health Econ. 1999;8:25–39.
7. Versteegh MM, Brouwer WBF. Patient and general public preferences for health states: A call to reconsider current guidelines. Soc Sci Med. 2016;165:66-74.
8. Stiggelbout AM, de Vogel-Voort E. Health state utilities: a framework for studying the gap between the imagined and the real. Value Health. 2008;11:76-87.
9. Department of Health. Liberating the NHS: No decision about me, without me – Government response to the consultation. London: Department of Health, 2012. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/216980/Liberating-the-NHS-No-decision-about-me-without-me-Government-response.pdf. Accessed 23rd October 2017.
10. Menzel P, Dolan P, Richardson J, Olsen JA. The role of adaptation to disability and disease in health state valuation: a preliminary normative analysis. Soc Sci Med. 2002;55:2149-58.
11. Goodwin E, Green C, Spencer A. Estimating a preference-based index for an eight dimensional health state classification system derived from the Multiple Sclerosis Impact Scale (MSIS-29). Value Health. 2015;18:1025-36.
12. Zajicek J, Freeman J, Porter B. Mulitple Sclerosis Care: A Practical Manual. Oxford: Oxford University Press, 2007.
13. Hemmett L, Holmes J, Barnes M, Russell N. What drives quality of life in multiple sclerosis? QJM. 2004;97:671-6.
14. Fisk JD, Brown MG, Sketris IS, Metz LM, Murray TJ, Stadnyk KJ. A comparison of health utility measures for the evaluation of multiple sclerosis treatments. J Neurol Neurosurg Psychiatry. 2005;76:58-63.
15. Kuspinar A, Mayo NE. Do generic utility measures capture what is important to the quality of life of people with multiple sclerosis? Health and Qual Life Outcomes. 2013;11:1-10.
16. Orme M, Kerrigan J, Tyas D, Russell N, Nixon R. The effect of disease, functional status, and relapses on the utility of people with multiple sclerosis in the UK. Value Health. 2007;10:54-60.
17. Benito-León J, Morales JM, Rivera-Navarro J, Mitchell A. A review about the impact of multiple sclerosis on health-related quality of life. Disabil Rehabil. 2003;25:1291-303.
18. Gruenewald DA, Higginson IJ, Vivat B, Edmonds P, Burman RE. Quality of life measures for the palliative care of people severely affected by multiple sclerosis: a systematic review. Mult Scler. 2004;10:690-704.

DRAFT

19. Opara JA, Jaracz K, Brola W. Quality of life in multiple sclerosis. J Med Life. 2010;3:352-8.
20. Goodwin E, Green C. A QALY Measure for Multiple Sclerosis: Developing a Patient-Reported Health State Classification System for an MS-Specific Preference-Based Measure. Value Health. 2015;18:1016-24.
21. Hobart JC, Riazi A, Lamping DL, Fitzpatrick R, Thompson AJ. Improving the evaluation of therapeutic interventions in multiple sclerosis: development of a patient-based measure of outcome. Health Technol Assess. 2004;8:1-60.
22. Gudex C. Time trade-off user manual: Props and self-completion methods. York: The MVH Group, Centre for Health Economics, University of York, 1994. http://www.york.ac.uk/che/pdf/op20.pdf. Accessed 23rd October 2017.
23. Dolan P. Modeling Valuations for EuroQol Health States. Med Care. 1997;35:1095-108.
24. Green C, Goodwin E, Hawton A. "Naming and Framing": The Impact of Labeling on Health State Values for Multiple Sclerosis. Med Decis Making. 2017;37:703-14.
25. MS Register website. https://www.ukmsregister.org/Portal/Home#about. Accessed 18th July 2016.
26. Ford DV, Jones KH, Middleton RM, et al. The feasibility of collecting information from people with multiple sclerosis or the UK MS Register via a web portal: characterising a cohort of people with MS. BMC Med Inform Decis Mak. 2012;12:1-8.
27. Mackenzie IS, Morant SV, Bloomfield GA, MacDonald TM, O'Riordan J. Incidence and prevalence of multiple sclerosis in the UK 1990–2010: a descriptive study in the General Practice Research Database. J Neurol Neurosurg Psychiatry. 2014;85:76-84.
28. Goodwin E, Green C. A Systematic Review of the Literature on the Development of Condition-Specific Preference-Based Measures of Health. Appl Health Econ Health Policy. 2016;14:161-83.
29. Norman R, King MT, Clarke D, Viney R, Cronin P, Street D. Does mode of administration matter? Comparison of online and face-to-face administration of a time trade-off task. Qual Life Res. 2010;19:499–508.
30. Brazier J, Ratcliffe J,, Salomon JA, Tsuchiya A. Measuring and valuing health for economic evaluation. Oxford: Oxford University Press, 2007.
31. Versteegh MM, Leunis A, Uyl-de Groot CA, Stolk EA. Condition-specific preference-based measures: benefit or burden? Value Health. 2012;15:504-13.
32. Goodwin E, Green C, Hawton A. What difference does it make? A comparison of health state preferences elicited from the general population and from people with multiple sclerosis. Value Health. in submission.
33. Hawton A, Green C. Health utilities for multiple sclerosis. Value Health. 2016;19:460-8.
34. Benito-León J, Morales JM, Rivera-Navarro J. Health-related quality of life and its relationship to cognitive and emotional functioning in multiple sclerosis patients. Eur J Neurol. 2002;9:497–502.
35. Kind P. A revised protocol for the valuation of health states defined by the EQ-5D-3L classification system: Learning the lessons from the MVH study. York: Centre for Health Economics, University of York, 2009.
36. Greene WH. Econometric Analysis. 7th ed. Boston: Prentice Hall, 2012.
37. Brazier J, Akehurst R, Brennan A, et al. Should patients have a greater role in valuing health states? Appl Health Econ Health Policy. 2005;4:201–8.
38. Dolan P. Whose preferences count? Med Decis Making. 1999;19:482-86.

DRAFT