



Comparison of Caregiver and General Population Preferences for Dependency-Related Health States

Eva Rodríguez-Míguez^{1,2} · Antonio Sampayo³

Accepted: 8 August 2024 / Published online: 10 September 2024
© The Author(s) 2024

Abstract

Objective We assess whether the preferences regarding dependency-related health states as stated by informal caregivers are aligned with those expressed by the general population.

Methods The preferences of a sample of 139 Spanish informal caregivers of dependent patients are compared with those obtained via a sample of 312 persons, also from the Spanish general population. We assess 24 dependency states extracted from the DEP-6D using the time trade-off method. Descriptive statistics and regression methods are used to explore differences between the two samples.

Results Mean difference tests establish that, for all but one of the 24 states, there are no significant differences between the samples. The estimated mean values ranged from – 0.64 to 0.60 for the caregiver sample and from – 0.60 to 0.65 for the general population sample, with a correlation of 0.96. On average, the classification of states as better or worse than dead matched in both samples (except for one state). Regression models also show that sample type does not have a significant average impact. After we introduce interaction effects, only the most severe level of two dimensions, cognitive problems and housework, result in significant differences—with the caregiver sample reporting higher values for the former, and lower values for the latter.

Conclusion Caregivers and the general population exhibit quite similar preferences concerning dependency-related health states. This suggests that the results of cost-utility analyses, and the resource allocation decisions based on them, would likewise not be significantly affected by the preferences used to generate the weighting algorithm.

Key Points for Decision Makers

The caregivers' preferences for dependency-related health states are very similar, on average, to the preferences of the general population.

The analysis by dimension shows that the cognitive problems generate more disutility for the general population than for carers, and the opposite is true for the housework dimension.

Cost-utility analysis and public decision making would hardly be affected by the preferences used to generate the dependency states values, and the use of preferences from general population is recommended.

1 Introduction

The rapid ageing of the population will bring new challenges, one of which will be to delay as much as possible the loss of health-related quality of life (QoL) for older people. A frequent consequence of this deterioration in health associated with ageing is the loss of the ability to carry out basic activities of daily living, resulting in the need for help from others. This situation of dependency, not necessarily associated with ageing, translates into an enormous cost of care, which is added to the health expenditure generated by these situations. In fact, it is common for funding decisions to change when these costs are included [1, 2].

✉ Eva Rodríguez-Míguez
emiguez@uvigo.es

Antonio Sampayo
ar.sampayo@usc.es

¹ ECOBAS, Universidade de Vigo, GRIIEE, 36310 Vigo, Spain

² Instituto de Investigación Sanitaria Galicia Sur, Hospital Álvaro Cunqueiro, 36213 Vigo, Spain

³ ECOBAS, Universidade de Santiago de Compostela, 15782 Santiago de Compostela, Spain

There are multiple strategies—both pharmacologic [3, 4] and non-pharmacologic [5, 6]—to promote health for older adults, which contribute both to prevent or delay situations of dependency, and to decrease the degree of dependency when it occurs (preventing strokes, delaying both the onset and the progression of Alzheimer's disease and other dementias, rehabilitation, technical aids, education programmes to prevent falls, etc.). The benefits and costs of these strategies and others expected to emerge in the future need to be quantified to make informed decisions. Several instruments are available to assess the QoL associated with a situation of dependency—dependency-related QoL—that make it possible to estimate the effectiveness of such programmes, some of which are widely used, such as the Katz scale [7], the Barthel Index [8] or the Lawton and Brody scale [9]. However, these instruments are not preference-based and, in consequence, do not allow to measure the variation of QoL using scales recommended for economic evaluation, such as quality-adjusted life years (QALYs). To overcome these limitations, generic preference-based instruments can be used, with the EQ-5D being the most frequently mentioned in guidelines from national health technology assessment agencies [10]. However, the main limitation of these instruments is that generic measures, as opposed to specific measures of dependency, may not be sensitive enough to identify relevant changes in dependency status [11, 12]. For example, regaining the ability to eat on one's own in a Parkinson's sufferer (either through a tremor-reducing drug or an electronic device that compensates for tremors) is a major change at the individual and family level, but would rarely lead to a change in any of the generic preference-based measures, in patients with multiple comorbidities. The weak correlation between the Barthel index and the EQ-5D found in the literature supports this evidence [13–15]. In the context of dependency, there is also a growing interest in the Adult Social Care Outcomes Toolkit (ASCOT) [16, 17]. The ASCOT is another preference-based instrument that can be applied to economic evaluation. However, this tool does not measure changes in the severity of dependency but rather changes in the social care-related QoL.

The DEP-6D, is a specific instrument designed to measure changes in dependency-related QoL, which integrates the discriminatory capacity of instruments designed specifically for dependency with the desirable properties for use in economic evaluation [18]. As the authors point out “the proposed instrument provides complementary utilities to generic measures with a wider focus and should be used alongside these to provide a more accurate measurement”. This instrument provides classical QALYs anchored on a full health-dead scale as commonly used in cost-utility analysis and it has been used in different contexts, such as the

estimation of disability-adjusted life expectancy [19], or in measuring the impact of severity of dependency on caregiver burden [20]. Using the COSMIN grid in a systematic review of generic and specific instruments used to estimate QALYs, Touré et al. [21] found that 63.2% of the quality criteria assessed were rated as very good in the DEP-6D, compared to 56% of the mean for all instruments assessed.

The DEP-6D amounts to a system for classifying dependency states along six dimensions of three or four levels each, and it incorporates a preference-based scoring algorithm that assigns weights to each of the dimension levels. The weights are based on preferences of Spanish citizens obtained by way of the time trade-off (TTO) method. The use of general population preferences to estimate these weights is a strategy often recommended in the economic evaluation of health programmes [22, 23]. However, it would be desirable to analyse whether the results would be similar if based on the preferences of other involved groups. The DEP-6D describes, in general, more severe problems than those usually described by means of generic preference-based measures, with the minimum utility score predicted by the DEP-6D algorithm being -0.84 (valued on the 1-0 full health–dead QALY scale), which is considerably lower than the minimum value predicted by the Spanish algorithms of the most widely used generic instruments, such as the EQ-5D-3L [24], the EQ-5D-5L [25] and the SF-6D [26], whose minimum values are -0.65 , -0.25 and -0.36 , respectively. The authors also find that more than half (14 from 24) of the directly rated dependency states had a negative value—that is, they were considered to be worse than dead. Given the low values obtained, one wonders whether greater knowledge based on personal experience might lead to a different assessment of these states, producing an algorithm with different weights.

Following Cubi-Molla et al. [27], different perspectives can be adopted to incorporate experience in the assessment of health states. One of them uses patients to assess their own state of health. However, it has strong limitations if we want to assess dependency situations that are very severe or linked to cognitive problems. While it is true that other studies have surveyed people with mild or moderate cognitive problems, the tasks to be performed were aimed at positioning themselves in one of the levels of the dimensions assessed (for example in the EQ-5D multi-attribute classification system). When utility weights need to be estimated, as in our study, only patients with a mild cognitive impairment can participate [28]. It should be noted that the methodology commonly used to obtain utility weights, for example the TTO method, requires very demanding tasks from a cognitive point of view, which makes it impossible for many of the health conditions described by the DEP-6D to be assessed by patients experiencing such a health condition. Another perspective that can also be considered to evaluate health states is that provided by dependent people

evaluating hypothetical situations. In this case, dependent persons without cognitive problems would be able to assess the 24 health states previously assessed by the general population. However, this option has also been rejected in our study for several reasons. On the one hand, there may be an asymmetric bias between assessments of health states that are more similar to those of the patient being interviewed, in which adaptation may play a role, and those health states that are more different [29, 30]. On the other hand, the elicitation of experience-based values would be compromised in the case of more severe hypothetical states.

Another group of interest widely used in the literature to incorporate experience-based values into the analysis is caregivers of dependent persons (vicarious experience) [31, 32]. This is the group we use in our study. The answers of informal caregivers can be a good proxy because they are familiar with the limitations, needs, and feelings of the individuals for whom they care and as the National Institute for Health and Clinical Excellence (NICE) points out, when “it is not possible to get measurements directly from patients, these should come from the person who acts as their carer” [22]. In addition, using the caregiver’s perspective allows for the incorporation of experience in a wide range of dependency situations, both in terms of current and past experience—it is common for the same caregiver to have experience in different dependency situations that the cared-for person has been through. Given that our aim is to analyse whether the preferences on dependency-related health states expressed by informal caregivers coincide with those expressed by the general population, as described in Rodríguez-Míguez et al., the assessment of the same 24 hypothetical health states also evaluated by the general population seems to us to be the most appropriate strategy [18]. Furthermore, there is evidence that when it comes to hypothetical health states, patient and caregiver preferences do not differ significantly [31].

There is, *a priori*, no clear expectation. On the one hand, care provision may induce a process of adaptation that leads to higher evaluations of these states; this phenomenon is frequently reported in the literature comparing the preferences of patients with those of the general population [33]. Although this “disability paradox” refers to patients’ preferences, a similar effect might arise when caregivers’ preferences are analysed. On the other hand, close contact with dependent persons may provide more information about the limitations and losses of well-being that such situations produce, which in turn could exacerbate the negative perception of these states [28]. Finally, it may be that the two groups express similar preferences, reinforcing the recommendation (from studies of economic evaluation) to use general population samples.

2 Method

2.1 The DEP-6D Instrument

The DEP-6D is an instrument used to characterise the level of daily dependence on others [18]. It comprises a questionnaire and a weighting algorithm that assigns a score to each of the responses provided. The DEP-6D classification system is based on six attributes or dimensions—eating, incontinence, personal care, mobility, housework, and cognition problems—each having either three or four levels (see Table 1). Correlation analysis and focus groups were used to select dimensions and levels, and the TTO method was used to estimate a preference-based scoring algorithm. A major advantage of this dependency assessment tool is that it enables the estimation of QALYs, which underpin cost-utility analysis.

2.2 Samples and Questionnaire

This study uses two cross-sectional samples from two groups: the general population and caregivers. Microdata from the general population sample were provided by Rodríguez-Míguez et al. [18]. This sample consists of 312 citizens drawn from the general population of Galicia (a region in Spain) recruited at home in 2011 by means of stratified random sampling. Caregivers’ preferences were obtained in 2016 from a convenience sample of 139 primary informal caregivers of chronically ill dependent adults who need help with activities of daily living. The caregivers were contacted in primary care centres in Pontevedra (Galicia) and interviewed by health and social workers. In both samples, the interviewers were trained to conduct the interview as follows; first, the content of the questionnaire was explained in detail and then a mock interview was conducted with one of the researchers.

Face-to-face interviews were conducted in the participant’s home in both samples. In order to reproduce as closely as possible, the interview with the general population, a pen-and-paper interview was also conducted with the sample of caregivers. Only interviews that, due to interviewer error, did not follow the pre-established TTO sequence (explained below) were eliminated—in these situations the TTO values cannot be obtained. The questionnaire begins with the presentation of the dimensions and levels of the DEP-6D. Since the questionnaire is designed so that in many cases it can be self-completed, no additional explanation is provided, unless requested by the participant. Next, participants were asked to value six DEP-6D dependency states. To increase comparability, the questionnaire and experimental design used for the

Table 1 Levels and dimensions of DEP-6D

Levels of dimensions	Dimensions
1. Does not need assistance to eat or drink 2. Needs partial aid to eat or drink (cutting, serving, etc.) 3. Needs to be given food and drink	Feeding
1. Does not have incontinence or does not need help 2. Has urinary (or faecal) incontinence and needs help for hygiene 3. Has both urinary and faecal incontinence and needs help for hygiene	Incontinence
1. Does not need help for personal care: bathing, grooming, dressing, etc. 2. Needs help only to bathe but not for the rest of personal care 3. Needs help for most personal care activities 4. Is incapable of carrying out personal care; needs someone else to perform this activity	Personal care
1. Moves independently 2. Does not need help to move within the home but does out of home 3. Needs help to move both in and out of home 4. Is incapable of changing position; bed-ridden or chair-ridden	Mobility
1. Does not need help to carry out daily housework (making food, cleaning crockery, etc.) 2. Needs daily help for daily housework 3. Is incapable of carrying out most daily housework	Housework
1. Does not need help due to cognitive/mental problems or has not these problems 2. Needs assistance to manage money, medication or to take some basic everyday decisions; generally collaborative attitude with the caregiver 3. Incapable of taking basic decisions; cannot live alone. Does not resist help 4. Incapable of taking basic decisions; cannot live alone. Does not collaborate and usually offers resistance to help	Cognition problems

caregiver sample were identical to the one used previously for the general population sample. It included a total of 24 dependency states, divided into four blocks of size six (see Table 4 in the Results section to follow); the states and blocks were obtained in the seminal work using an optimal design that captures only main effects, which means that interaction effects cannot be tested. The blocks and the states were randomly assigned to the participants. Assessment of the dependency states was carried out using the TTO method. After rating the six states in sequence, each participant was shown the six states together (on separate cards) and asked to rank them from most to least preferred. Participants' socioeconomic characteristics were also collected.

2.3 Utility Valuation Method

The traditional TTO method was used to assess the dependency states. Given that the assessment of health states may vary depending on the framing of the TTO questions [34–37], to avoid confounding effects when comparing the preference of caregivers to the general population, we used the same framework as in the original study. Participants were asked to place themselves in a hypothetical situation in which a doctor informed them that they suffer from a serious illness which, if not treated urgently, will degenerate and rapidly lead to death. The doctor also informed them that there was a treatment that, although it would not

cure the disease, would allow them to live 10 more years in a certain state of dependency (at this moment, one of the six states to be evaluated is shown). Participants were asked to choose between being treated or not. If they choose the treatment, this means that the state is regarded as better than dead (BTD); if the treatment is rejected, then the state is considered worse than dead (WTD). Depending on the answer given to this initial question, the following line of inquiry is adopted.

- If they accept treatment, then an iterative up–down procedure was applied to find the number of years, Y_{BTD}^* , in full health (FH) at which participants were indifferent between 10 years in the state being evaluated and Y_{BTD}^* in FH. The starting value for Y_{BTD} was set equal to 5 years. Thus, participants must choose between two options: A, 10 years in the evaluated state; or B, $Y_{\text{BTD}} = 5$ years in FH. Depending on the answer provided, Y_{BTD} is adjusted up or down until the indifference point—or the interval containing it—(Y_{BTD}^*) is bounded as described next (indifference is allowed in all choices, in which case the process ends).

– If participant chooses A $\rightarrow Y_{\text{BTD}} = 8$:

- 1) if A $\rightarrow Y_{\text{BTD}} = 9$;
- 2) if B $\rightarrow Y_{\text{BTD}} = 7 \rightarrow$ if B $\rightarrow Y_{\text{BTD}} = 6$.

- If participant chooses B $\rightarrow Y_{\text{BTD}} = 2$:
 - 1) if A $\rightarrow Y_{\text{BTD}} = 4 \rightarrow$ if B $\rightarrow Y_{\text{BTD}} = 3$;
 - 2) if B $\rightarrow Y_{\text{BTD}} = 1$.
- If they refuse treatment, the next choice was between: A, Y_{WTD} years in full health followed by $10 - Y_{\text{WTD}}$ years in the state being evaluated; or B, dying. As before, the value for Y_{WTD} as initially set to 5 years and then moved up or down until the convergence process terminated.
- If participant chooses A $\rightarrow Y_{\text{WTD}} = 2$:
 - 1) if A $\rightarrow Y_{\text{WTD}} = 1$;
 - 2) if B $\rightarrow Y_{\text{WTD}} = 4 \rightarrow$ if B $\rightarrow Y_{\text{WTD}} = 3$.
- If participant chooses B $\rightarrow Y_{\text{BTD}} = 8$:
 - 1) if A $\rightarrow Y_{\text{WTD}} = 7 \rightarrow$ if B $\rightarrow Y_{\text{WTD}} = 6$;
 - 2) if B $\rightarrow Y_{\text{BTD}} = 9$.

Utility for state S , denoted $U(S)$, was obtained by using the QALY model assumptions for chronic health states – that is, the value of existing Y years in state S is equal to $U(S) \times Y$. By convention, the utility of FH is 1 and the utility of death is 0; hence $U(S)$ is obtained as follows. If a state is regarded as BTD then, under the previous assumptions, $U(S) = Y_{\text{BTD}}^*/10$. For states that are regarded as WTD, $U(S) = -Y_{\text{WTD}}^*/(10 - Y_{\text{WTD}}^*)$. However, because the negative utilities calculated in this way do not have a lower bound (and so result in distributions that are skewed strongly to the left), we bound negative values at -1 via applying the transformation suggested by Patrick et al. [38]: $U(S) = -Y_{\text{WTD}}^*/10$. Owing to our iterative up–down procedure, usually a range of values (and not an indifference value) is obtained for both Y_{BTD}^* and Y_{WTD}^* . The following example illustrates the procedure used to obtain $U(S)$. Suppose a participant, who considers a state better than dead, has chosen option A when $Y_{\text{BTD}} = 5$, option B when $Y_{\text{BTD}} = 8$, and option A when $Y_{\text{BTD}} = 7$. In that event, Y_{BTD} lies in the interval $(0.7, 0.8)$. Visual aids (identical for both samples) were used in all questions to help subjects understand the questions posed; as an example, Fig. 1 shows the visual aid presented immediately after the state has been found to be considered BTD by the participant.

2.4 Statistical Analysis

Several analyses were conducted to determine whether the preferences of caregivers are different from those of the general population. First, we checked for whether the level of inconsistency—both internal and external—and invariance

of the responses is similar in both samples. Internal consistency is analysed by identifying non-compliance with the dominance test, which is violated when a respondent places more value on a state that is logically worse in the sense that none of its dimensions presents a better situation. External consistency is measured as the extent of correlation between the ranking of states as derived from the TTO exercises and the direct ranking provided by participants. With regard to invariance, we identify those participants who assign the same value to all states, a particular case of which involves individuals who simply do not trade off ($U(S) = 1$ for all states).

Second, for each of the 24 states assessed we determined the mean utility by sample (general population vs caregivers) and checked for significant differences between those two samples. There is no consensus on how best to proceed with inconsistent participants [39, 40]; here we followed other authors in opting to exclude participants with more than one inconsistency. In this way, we seek to distinguish between systematic departures from dominance—as exhibited by the excluded participants—and random error [18, 41]. We report the states' mean values for the full sample and also for the reduced sample (hereinafter referred to as the “consistent participants”). Only the sample of consistent participants is used for all other aspects of the statistical analyses.

Third, random-effects interval-data regression models were used to obtain the weighting algorithms for each of the samples, and also to look for significant between-sample differences in preferences concerning dependency-related health states. These regression models allow us to consider both uncertainty about the exact value of indifference and the presence of left-censored value [25]. On the one hand, in most of the ratings we do not obtain a point of indifference, but an interval in which the unknown value lies. On the other hand, by construction, the values obtained had a minimum TTO value bounded at -1 , but respondents could hypothetically continue trading beyond the left bound at -1 . Therefore, we relax this lower bound assumption and consider responses at the lower bound (-1) to be censored. The model also contains a random effect to account for the dependence of repeated observations within respondents. The dependent variable for all models was $U(S)$. In model 1 and model 2 we estimate separately the weighting algorithms for each of the samples (caregivers and general population), and the independent variables are the severity levels of the dimensions. To facilitate comparisons, the relative importance of each dimension is obtained by dividing its range by the sum of the ranges of all attributes. Model 3 uses the full data set; in this model, the independent variables are the severity levels of dimensions and a dummy variable set to 1 for caregiver respondents or to 0 for respondents from the general population. Model 4 reproduces model 3 but

Fig. 1 Example of visual aid

b2

Suponga ahora que usted se encuentra en el estado que acaba de elegir (ver tarjeta) y vivirá en esa situación los próximos 10 años.
En una nueva visita al médico se le informa que hay un nuevo tratamiento que le permitiría recuperar la salud, pero vivirá menos años.

Por tanto, la situación a la que se enfrenta es la siguiente:

SIN TRATAMIENTO

10 AÑOS CON ESAS LIMITACIONES

 †

CON TRATAMIENTO

5 AÑOS EN BUENA SALUD

 †

¿Se sometería al tratamiento?

Translation: Now suppose that you are in the state you have chosen (see card) and you will live in that situation for the next 10 years. On a new visit to the doctor, you are informed that there is a new treatment that would allow you to regain your health, but you will live fewer years. Therefore, the situation you face is the following:

WITHOUT TREATMENT: 10 YEARS WITH THOSE LIMITATIONS

WITH TREATMENT: 5 YEARS IN GOOD HEALTH

Would you undergo the treatment?

controls for the most relevant sociodemographic variables (gender, age, and educational level). Finally, the independent variables in model 5 are the interactions between severity levels and sample type; thus, we explore whether (and, if so, how) the sample's effects differ among the dimensions.

Written informed consent was obtained from each participant, and the Committee on Ethics of Clinical Research in Galicia approved the study.

3 Results

In the sample of caregivers, all questions have correctly followed the established itinerary to obtain TTO values—in the general population sample only 2 observations had been eliminated. Table 2 presents a descriptive analysis of the sample. As expected, the sample of caregivers differs markedly from that of the general population. In particular, the former consists mostly of women and has a higher mean age than the general population. The caregiver sample is also characterized by a lower educational level and

a lower frequency of people working, with a quarter of the sample being retired people. The distribution of family income (excluding missing data) is similar across the samples.

Table 3 reports the level of consistency and invariance by sample. The two samples yield similar results, with no significant differences being observed. There is a high percentage of participants who meet all the consistency tests, and if we include those who commit no more than one inconsistency then this percentage rises to 94% in both samples. The correlation between the direct ranking and the resulting TTO method is only moderate yet is similar in both samples. The number of participants who always provide the same answer for all states (invariant utility) is higher in the general population than in caregivers, although the difference is not statistically significant ($p = 0.228$).

Table 4 shows the mean values of the 24 states that were directly evaluated. The labelling of the states refers to the levels in which the state is found in each of the six dimensions that make up the DEP-6D (e.g., a dependent in state 122222 does not need assistance to eat or drink but has a level-2 severity in all the other dimensions). Figure 2 plots the mean values, by sample, ordered according to the sum of the (ordinal)

Table 2 Characteristics of the two samples (%)

	Caregivers (<i>n</i> = 139)	General population (<i>n</i> = 312)
Female	91.4	47.4
Age (mean)	55.0	41.5
Education		
Primary studies or less	59.7	37.5
Secondary	30.2	39.4
University	10.1	23.1
Lives with partner	33.8	39.9
Labor status		
Employed	20.1	59.6
Pensioner/retired	25.9	10.9
Unemployed	41.0	15.7
Domestic tasks	10.1	8.3
Student/others	2.9	5.5
Home income (€ monthly)		
< 500	1.4	5.9
500–999	15.8	13.2
1000–1499	29.5	30.5
1500–1999	25.9	25.7
2000–2999	19.4	16.9
≥ 3000	7.9	7.7
# Household members (mean)	3.1	3.5
CD		
No CD	0.0	53.2
CD (not living together)	16.55	40.1
CD (living together)	83.45	6.7

CD close dependent

Table 3 Consistency and invariance by sample

	Caregivers (<i>n</i> = 139)	General population (<i>n</i> = 312)
Meet all dominance tests	80.6%	79.2%
Fail only one dominance test	13.7%	14.7%
Individual Spearman's rho between TTO and direct ranking (mean)	0.65	0.66
$U(S) = 1$ for all states (no trade-off)	0%	0%
Same utility for all states	3.6%	6.4%

TTO time trade-off, $U(S)$ utility for state

values of the levels that make up each state. According to the mean difference tests whose results are reported in Table 4, there are no significant differences between the samples (caregivers vs general population)—except for the state 112132 ($p = 0.008$ in the full samples and $p = 0.013$ in the samples

of consistent participants). Pearson's coefficient of correlation between the mean values of the two samples is 0.96. Mean values of the consistent samples ranged from -0.69 to 0.64 for the caregiver's sample and from -0.62 to 0.66 for the general population sample. On average, the classification of states as better or worse than dead coincides in both of those samples (except for state 212223).

Table 5 (models 1 and 2) displays the impact of dimension levels on the utility weights (weighting algorithm) by sample. For the estimation, the two most severe levels of the dimension “personal care” (level 3 and level 4 in Table 1) have been combined into a single level. The reason is that, in the general population sample, level 4 of personal care had less of an effect on QoL than level 3 of personal care (although the two parameters are not significantly different, $p = 0.568$); hence, we followed the approach of Rodríguez-Míguez et al. [18] in grouping them together. We also grouped them in the caregiver sample to facilitate comparisons between samples; in this case, the parameters reflected the hypothesized effects—level 3 of personal care had less of an effect on QoL than level 4 of personal care—($p = 0.058$). The remaining parameters have the expected sign and direction: as the situation in a given dimension worsens, its utility declines. Comparing models 1 and 2 reveals that the largest differences are observed in the most severe level of the “housework” dimension (caregivers report greater disutility than the general population) and in the severest level of “cognitive problems” (for which caregivers report less disutility). Table 6 shows the relative importance of the 6 dimensions, as well as their confidence intervals, for each of the samples. The relative importance is quite similar between samples for all dimensions except for the housework and cognitive problems dimensions, both with a difference of around 9 percentage points. Models 1 and 2 allow the estimation of the utility score associated with any dependency situation described by the DEP-6D by adding to the constant the estimated value of each of its levels.

We can see from model 3 that, at the aggregate level, the average score in the sample of caregivers is 0.063 points less than the general population sample—although this difference is not significant ($p = 0.36$). Model 4 establishes that these results do not change substantially when we control for sociodemographic variables. And in model 5, where we allowed for differences by dimension, again there are no significant differences between samples except for the most severe level of cognitive problems (-0.607 for the general population and -0.444 for caregivers; $p = 0.032$) and the most severe level of housework (-0.104 for the general population and -0.240 for caregivers; $p = 0.031$). These results do not change substantially when we control for sociodemographic variables (results not displayed).

Table 4 Descriptive statistics of observed TTO values by sample

	State ^a	All participants							Consistent participants						
		Caregivers			General population			<i>p</i> value	Caregivers			General population			<i>p</i> value
		<i>n</i>	Mean	SD	<i>n</i>	Mean	SD		<i>n</i>	Mean	SD	<i>n</i>	Mean	SD	
Block 1	122222	34	0.09	0.65	82	0.17	0.54	0.465	32	0.12	0.66	77	0.18	0.53	0.607
	133334	34	-0.37	0.56	82	-0.45	0.47	0.403	32	0.39	0.56	77	-0.49	0.45	0.335
	211121	34	0.60	0.49	82	0.58	0.47	0.888	32	0.64	0.45	77	0.60	0.46	0.696
	214232	34	0.00	0.64	82	0.06	0.59	0.605	32	0.02	0.65	77	0.04	0.60	0.871
	313331	34	0.11	0.71	82	0.17	0.60	0.614	32	0.15	0.71	77	0.17	0.61	0.873
	323433	34	-0.42	0.54	82	-0.49	0.51	0.480	32	-0.41	0.55	77	-0.53	0.48	0.252
Block 2	111221	36	0.60	0.48	78	0.65	0.42	0.518	35	0.62	0.46	77	0.66	0.42	0.664
	112132	36	0.02	0.68	78	0.35	0.57	0.008	35	0.04	0.68	77	0.35	0.57	0.013
	112211	36	0.51	0.58	78	0.60	0.45	0.355	35	0.50	0.58	77	0.60	0.45	0.322
	223234	36	-0.44	0.52	78	-0.47	0.54	0.822	35	-0.43	0.52	77	-0.48	0.53	0.610
	234333	36	-0.54	0.43	78	-0.42	0.51	0.235	35	-0.57	0.40	77	-0.44	0.50	0.185
	333122	36	-0.37	0.55	78	-0.23	0.62	0.259	35	-0.37	0.55	77	-0.24	0.61	0.304
Block 3	111112	33	0.38	0.49	75	0.40	0.57	0.905	30	0.44	0.46	66	0.50	0.49	0.594
	113233	33	-0.27	0.60	75	-0.12	0.66	0.263	30	-0.31	0.60	66	-0.12	0.66	0.202
	213322	33	-0.26	0.60	75	-0.07	0.64	0.144	30	-0.30	0.59	66	-0.07	0.64	0.098
	222131	33	0.08	0.58	75	0.24	0.60	0.191	30	0.10	0.58	66	0.26	0.60	0.235
	234431	33	-0.57	0.49	75	-0.37	0.62	0.102	30	-0.60	0.48	66	-0.37	0.62	0.075
	334234	33	-0.64	0.40	75	-0.54	0.51	0.293	30	-0.69	0.35	66	-0.55	0.51	0.179
Block 4	123121	36	0.29	0.57	77	0.30	0.57	0.984	34	0.31	0.59	73	0.32	0.58	0.917
	212223	36	0.02	0.66	77	-0.16	0.67	0.198	34	0.00	0.67	73	-0.15	0.68	0.301
	233432	36	-0.46	0.53	77	-0.45	0.58	0.946	34	-0.52	0.48	73	-0.48	0.57	0.743
	314434	36	-0.55	0.45	77	-0.60	0.51	0.563	34	-0.58	0.42	73	-0.62	0.50	0.686
	324332	36	-0.45	0.49	77	-0.32	0.61	0.234	34	-0.46	0.50	73	-0.32	0.61	0.244
	333231	36	-0.21	0.65	77	-0.19	0.64	0.841	34	-0.26	0.64	73	-0.21	0.65	0.709

SD standard deviation, *TTO* time trade-off

^aThe first number of the state refers to the level of the first dimension, and so on

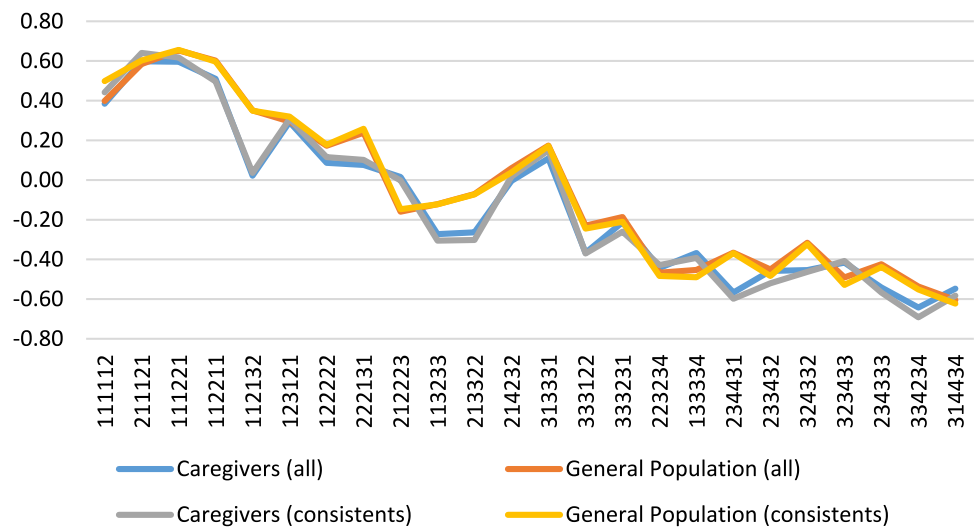
Fig. 2 Values of DEP-6D states by sample

Table 5 DEP-6D models: Comparing caregivers and the general population (GP)

	MODEL 1 (carers)			MODEL 2 (GP)			MODEL 3 (full data)			MODEL 4 (full data)			MODEL 5 (full data)			
	Coef.	SE	p	Coef.	SE	p	Coef.	SE	p	Coef.	SE	p	Coef.	SE	p	
Cons	0.839	0.068	<0.001	0.776	0.047	<0.001	0.875	0.097	<0.001	0.837	0.152	<0.001	Cons	0.794	0.040	<0.001
Feeding_2	-0.095	0.031	0.002	-0.159	0.031	<0.001	-0.139	0.023	<0.001	-0.139	0.026	<0.001	Feeding_2 × GP	-0.158	0.031	<0.001
Feeding_3	-0.182	0.038	<0.001	-0.220	0.037	<0.001	-0.207	0.028	<0.001	-0.207	0.029	<0.001	Feeding_2 × carer	-0.095	0.036	0.008
Incontinence_2	-0.113	0.046	0.014	-0.134	0.031	<0.001	-0.129	0.023	<0.001	-0.129	0.027	<0.001	Feeding_3 × GP	-0.219	0.039	<0.001
Incontinence_3	-0.300	0.040	<0.001	-0.286	0.034	<0.001	-0.293	0.026	<0.001	-0.293	0.027	<0.001	Feeding_3 × carer	-0.179	0.046	<0.001
Personal_2	-0.230	0.067	<0.001	-0.122	0.044	0.006	-0.153	0.031	<0.001	-0.153	0.034	<0.001	Incontinence_2 × GP	-0.133	0.029	<0.001
Personal_3+4	-0.327	0.067	<0.001	-0.249	0.057	<0.001	-0.271	0.042	<0.001	-0.271	0.043	<0.001	Incontinence_2 × carer	-0.113	0.041	0.006
Mobility_2	-0.071	0.049	0.147	-0.099	0.037	0.008	-0.090	0.031	0.003	-0.090	0.023	<0.001	Incontinence_3 × GP	-0.285	0.030	<0.001
Mobility_3	-0.163	0.058	0.005	-0.144	0.043	0.001	-0.150	0.035	<0.001	-0.150	0.030	<0.001	Incontinence_3 × carer	-0.303	0.036	<0.001
Mobility_4	-0.289	0.068	<0.001	-0.363	0.048	<0.001	-0.337	0.042	<0.001	-0.336	0.036	<0.001	Personal_2 × GP	-0.126	0.039	<0.001
Housework_2	-0.169	0.059	0.004	-0.067	0.043	0.117	-0.096	0.030	<0.001	-0.096	0.030	0.001	Personal_2 × carer	-0.221	0.054	<0.001
Housework_3	-0.260	0.061	<0.001	-0.096	0.045	0.035	-0.146	0.030	<0.001	-0.145	0.033	<0.001	Personal_3+4 × GP	-0.250	0.057	<0.001
Cognition_2	-0.278	0.040	<0.001	-0.249	0.028	<0.001	-0.257	0.020	<0.001	-0.257	0.027	<0.001	Personal_3+4 × carer	-0.322	0.065	<0.001
Cognition_3	-0.349	0.046	<0.001	-0.437	0.035	<0.001	-0.410	0.026	<0.001	-0.410	0.027	<0.001	Mobility_2 × GP	-0.101	0.031	0.001
Cognition_4	-0.442	0.046	<0.001	-0.609	0.038	<0.001	-0.554	0.033	<0.001	-0.554	0.034	<0.001	Mobility_2 × carer	-0.063	0.042	0.133
Carer							-0.063	0.069	0.360	-0.074	0.061	0.228	Mobility_3 × GP	-0.147	0.038	<0.001
Age										0.000	0.002	0.856	Mobility_3 × carer	-0.156	0.043	<0.001
Female										0.043	0.049	0.377	Mobility_4 × GP	-0.364	0.050	<0.001
Secondary										-0.015	0.068	0.829	Mobility_4 × carer	-0.280	0.069	<0.001
University										0.030	0.068	0.663	Housework_2 × GP	-0.077	0.033	0.021
													Housework_2 × carer	-0.144	0.044	0.001
													Housework_3 × GP	-0.104	0.036	0.004
													Housework_3 × carer	-0.240	0.063	<0.001
													Cognition_2 × GP	-0.250	0.023	<0.001
													Cognition_2 × carer	-0.275	0.036	<0.001
													Cognition_3 × GP	-0.435	0.038	<0.001
													Cognition_3 × carer	-0.352	0.055	<0.001
													Cognition_4 × GP	-0.607	0.038	<0.001
													Cognition_4 × carer	-0.444	0.057	<0.001
# Participants	131			293			424			424			424			
# Observations	786			1758			2544			2544			2544			
# Left-cens. Obs.	112			316			428			428			428			
Wald χ^2	1095.23			906.39			1797.28			2239.40			3786.02			
Rho	0.640	0.037		0.509	0.030		0.547	0.024		0.546	0.032		0.550	0.021		

Inconsistent observations have been excluded in all models. Bootstrap standard errors in all models

Coef. coefficient, SE standard error

Table 6 Relative importance of dimensions by sample

	Caregivers			General population		
	Relative importance	95% CI		Relative importance	95% CI	
Feeding	0.101	0.062	0.140	0.121	0.084	0.157
Incontinence	0.167	0.117	0.217	0.157	0.121	0.193
Personal care	0.182	0.110	0.253	0.137	0.075	0.198
Mobility	0.160	0.097	0.223	0.199	0.154	0.245
Housework	0.145	0.085	0.204	0.052	0.004	0.101
Cognition problems	0.245	0.190	0.301	0.334	0.303	0.365

CI confidence interval

4 Discussion

In this study we compared the preferences of the general population and those of caregivers when assessing 24 dependency-related (hypothetical) health states from the DEP-6D instrument. With regard to the validity of the information obtained, no internal and external consistency tests exhibited significant differences between samples, and the invariant responses were low in both samples. As for the utility scores assigned to the states by respondents, significant differences were observed in only one of the 24 states assessed. The estimated regression models confirm that sample type does not have a significant average impact, even when we control for the participants' socioeconomic characteristics.

However, as mentioned below, one of the limitations of the caregiver sample is its small size, which decreases its statistical power. Therefore, it seems appropriate to comment on the mean differences between the two samples despite their non-significance. The general population sample assigns a mean value that is 0.063 points higher than the caregiver sample's mean—0.074 when we control for socioeconomic characteristics. In addition to the statistical non-significance of these differences, they are also not large in the sense of a “minimally important difference” (difference perceived by patients as beneficial)—a difference of 0.074 in the EQ-5D, which also uses the TTO method when assessing health states, has been estimated to constitute a minimally important difference [42].

The similarity in ratings that we find is in line with the results of other studies comparing the preferences of the general population and patients; see the meta-analyses by Dolders et al. [43] and Peeters and Stiggelbout [44]. Although Peeters and Stiggelbout find at the aggregate level that patients give higher ratings than the general population, the ratings of the two groups were similar when both rated hypothetical health states. In the field of dependency, Netten et al. [16] similarly found that, when using the ASCOT instrument to assess outcomes of social care for adults, there is insufficient difference between the

general population and service users to justify a different weighting algorithm.

However, although there is no significant average impact, the analysis by dimension provides additional information that is worth commenting on. In particular, the most severe level of the cognitive problems dimension generates more disutility for the general population than for carers, and the opposite is true for the most severe level of the dimension housework; there are no significant differences in the other comparisons. Regarding the differences found between carers and the general population in the cognitive problems dimension, the study most similar to ours is that of Rowen et al. [28]. These authors also focus on the differences between caregivers and the general population and find, using the TTO method, that caregivers of patients with mild dementia gave systematically lower values for hypothetical dementia health states than did members of the general population—an outcome at odds with our study's results, where for the mild and moderate levels of the cognitive problems dimension no significant differences were found and for severe states caregivers provide higher values. Yet, as those authors point out, the wording and labelling of the classification system may have affected these results. Although both samples rated the same description of hypothetical health states, caregiver participants knew that the study was about dementia whereas respondents from the general population were not informed that the health states evaluated were dementia specific. That difference may account for the latter sample's higher valuation.

Several hypotheses can be put forward to help explain the results obtained in our study, both at the aggregate level and by dimension. First, it should be noted that, although the most severe levels of the cognitive problems and housework dimensions are significantly different between the samples, their impact goes in the opposite direction; the cognitive problems dimension generates more disutility for the general population than for caregivers, and the opposite occurs with housework. Therefore, both impacts can exert a countervailing effect, which may help to explain, at least partially, the low aggregate

impact obtained. Second, there is evidence that the perspective in which we place the participant may play some role [27]. For example, in the study by Lobchuk et al. [45], lower values of health states are obtained when the caregiver imagines him/herself in that situation than when the patient's perspective is adopted. This would help to explain the lower (although not significant) rating provided by the caregivers in our sample. Third, given that we tried to incorporate experience in the assessment of health states, the fact that 47% of the general population sample has a close dependent relative—although only 6.7% live with him or her—may make their preferences closer to those of the caregiver sample. To try to isolate this possible effect, model 3 was re-estimated but considering 3 values for the variable sample, depending on whether the observation belongs to the general population without a close dependent (less experience), to the general population with a close dependent (more experience) or to the sample of caregivers. Although no significant differences were found between any of the 3 groups, the results show the expected direction, the greater the experience the smaller the difference with the sample of caregivers. Thus, with respect to the general population without a close dependent, the general population with a close dependent reduces the mean valuation of the states by 0.69 ($p = 0.198$) and the sample of caregivers by 0.095 ($p = 0.135$). Fourth, the higher values obtained for the most severe level of the cognitive problems dimension could be explained by an adaptation process leading to a higher assessment of these states. Although this phenomenon usually applies to assessments provided by patients [46], as mentioned above, a similar effect could occur when caregiver preferences are analyzed. Finally, the sample of caregivers mostly comprises women, and one might wonder whether this may have influenced the greater impact of the housework dimension in the sample of caregivers. Given that women tend to dedicate more time to housework than men, the limitations in performing it may have a greater impact, which could explain the observed results. To test this hypothesis, model 5 was replicated considering only the subsamples of women. If this were the main cause of the differences found in the most severe level of the housework dimension, the differences should be considerably reduced. The results obtained do not support this hypothesis: the parameter *housework_3* is -0.16 for the general population and -0.29 for female caregivers; $p = 0.10$).

We must point out the limitations of our study. First, given that our aim was to compare the preferences of a sample of caregivers with the general population sample obtained by Rodríguez-Míguez et al. [18], the questionnaire design was the same and therefore shared the limitations acknowledged in the original study. Thus, the design did not allow us to test for interaction effects between dimensions—that is, the

weights we estimate can be represented only by an additive main-effects model without interaction term—and utilities can only be obtained with an accuracy of 0.05 points. In addition, there is evidence that the assessment of health states can vary depending on the framing of the TTO questions—for example the duration of the state to be valued or the lead (or lag) time in FH attached [35, 37], the starting point [34], or the iteration procedure [36], among others. Although for comparability purposes our study follows the same TTO framework as the original study, the impact of this choice may affect comparability with other Spanish studies that allow for the estimation of QALYs. Thus, while our design is similar to that used for the estimation of the Spanish EQ-5D-3L and therefore QALYs are highly comparable, the comparison with the Spanish EQ-5D-5L [25] is more questionable because there is evidence that the composite-TTO values tend to be higher than the traditional TTO values for health states WTD [47].

Second, there are limitations related to the characteristics of the sample of caregivers. The small sample size and a possible selection bias may reduce the generality of our results, potentially compromising the external validity of the results. Although we have a representative sample of the general population, the same cannot be said about our sample of caregivers—whose members were recruited among those attending an appointment with the social worker in different primary health centres. This fact will naturally bias the sample of caregivers toward more severe dependents, which in turn may have influenced our results. In any case, we argue that if there are scant differences between the general population and caregivers of people with severe dependency then one could not reasonably expect there to be differences when caregivers of people with milder dependency are considered. That being said, informal caregivers of more severely dependent patients may well have endured different levels of dependency of the persons they care for, and this experience could endow them with a broad knowledge of different states of dependency. With respect to this sample's small size, it could easily have affected the statistical significance (or non-significance) of some results. Still, it is worth noting that small samples are common in this field; in the previously cited study of Rowen et al. [28], for example, only 71 patient caregivers participated in the TTO exercise.

In conclusion, our study provides a DEP-6D weighting algorithm that is based on a sample of caregivers and thus complements the previously estimated weighting algorithm based on general population preferences. We find that caregivers' preferences for hypothetical health states derived from the DEP-6D scarcely differ from preferences derived from the general population. However, some significant differences by dimension have been found. Since the results of caregivers have been extracted from a convenience sample our recommendation—pending further studies that yield new information on this topic—is to use the algorithm

derived from the general population because it is based on a larger sample and also aligns with the approach recommended for health economic evaluation.

Acknowledgements The authors thank the caregivers who participated in this study and the health and social care workers who recruited them and conducted the interviews. The contribution of Ana Rodríguez-González in the compilation of the information is particularly appreciated.

Declarations

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. This study was financed by the Spanish Ministry of Science, Innovation and Universities and FEDER (RTI2018-099403-B-I00 and PID2020-118119GB-I00) and Xunta de Galicia (ED431C2023/05). Funding for open access charge: Universidade de Vigo/ CRUE-CISUG.

Conflict of interest The authors have no competing/conflicting interests to declare.

Ethical approval for human subject research All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. This study was reviewed and approved by the Committee on Ethics of Clinical Research in Galicia (ref. 2012/381). The anonymity and confidentiality of the participants were preserved.

Informed consent Written informed consent was obtained from all participants. Participants were informed about their freedom for refusal.

Availability of data Data will be available upon reasonable request by contacting the corresponding author.

Consent for publication (from patients/participants) Not applicable.

Code availability Not applicable.

Authors' contributions All co-authors played a role in designing and implementing the research, analysing the results and writing the manuscript.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

References

1. Krol M, Papenburg J, van Exel J. Does including informal care in economic evaluations matter? A systematic review of inclusion and impact of informal care in cost-effectiveness studies. *Pharmacoeconomics*. 2015;33:123–35.
2. Peña-Longobardo LM, Rodríguez-Sánchez B, Oliva-Moreno J, Aranda-Reneo I, López-Bastida J. How relevant are social costs in economic evaluations? The case of Alzheimer's disease. *Eur J Health Econ*. 2019;20:1207–36.
3. Rountree SD, Atri A, Lopez OL, Doody RS. Effectiveness of antimentia drugs in delaying Alzheimer's disease progression. *Alzheimers Dement*. 2013;9(3):338–45.
4. Meretoja A, Keshtkaran M, Tatlisumak T, Donnan GA, Churilov L. Endovascular therapy for ischemic stroke: save a minute—save a week. *Neurology*. 2017;88(22):2123–7.
5. McLaren AN, LaMantia MA, Callahan CM. Systematic review of non-pharmacologic interventions to delay functional decline in community-dwelling patients with dementia. *Aging Ment Health*. 2013;17(6):655–66.
6. Santos FH, Zurek J, Janicki MP. Efficacy of healthy aging interventions for adults with intellectual and developmental disabilities: a systematic review. *Gerontologist*. 2022;62(4):e235–52.
7. Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. Studies of illness in the aged: the index of ADL: a standardized measure of biological and psychosocial function. *JAMA*. 1963;185(12):914–9.
8. Mahoney FI, Barthel D. Functional evaluation: the Barthel index. *Md State Med J*. 1965;14:56–61.
9. Lawton MP, Brody EM. Assessment of older people: Self-maintaining and instrumental activities of daily living. *Gerontologist*. 1969;9(3):179–86.
10. Kennedy-Martin M, Slaap B, Herdman M, van Reenen M, Kennedy-Martin T, Greiner W, Boye KS. Which multi-attribute utility instruments are recommended for use in cost-utility analysis? A review of national health technology assessment (HTA) guidelines. *Eur J Health Econ*. 2020;21:1245–57.
11. Donaldson C, Atkinson A, Bond J, Wright K. Should QALYs be programme-specific? *J Health Econ*. 1988;7(3):239–57.
12. Chisholm D, Healey A, Knapp M. QALYs and mental health care. *Soc Psychiatry Psychiatr Epidemiol*. 1997;32:68–75.
13. Hickson M, Frost G. An investigation into the relationships between quality of life, nutritional status and physical function. *Clin Nutr*. 2004;23(2):213–21.
14. Van Exel NJA, Scholte op Reimer WJM, Koopmanschap MA. Assessment of post-stroke quality of life in cost-effectiveness studies: the usefulness of the Barthel Index and the EuroQoL-5D. *Qual Life Res*. 2004;13:427–33.
15. Kaambwa B, Billingham L, Bryan S. Mapping utility scores from the Barthel index. *Eur J Health Econ*. 2013;14:231–41.
16. Netten A, Burge P, Malley J, Potoglou D, Towers AM, Brazier J, Forder J. Outcomes of social care for adults: developing a preference-weighted measure. *Health Technol Assess*. 2012;16(16):1–166.
17. Shiomiwa T, Moriyama Y, Nakamura-Thomas H, Morikawa M, Fukuda T, Batchelder L, Malley J. Development of Japanese utility weights for the Adult Social Care Outcomes Toolkit (ASCOT) SCT4. *Qual Life Res*. 2020;29:253–63.
18. Rodríguez-Míguez E, Abellán-Perpiñán JM, Alvarez XC, González XM, Sampayo AR. The DEP-6D, a new preference-based measure to assess health states of dependency. *Soc Sci Med*. 2016;153:210–9.
19. Llopart-Carles N, García-López S, Rejas-Gutierrez J. Disability-adjusted life expectancy lost due to pain severity and usual

- analgesic treatment among older adults with osteoarthritis in Spain. *Aging Clin Exp Res.* 2021;33:1285–95.
20. Rodríguez-González AM, Rodríguez-Míguez E, Claveria A. Determinants of caregiving burden among informal caregivers of adult care recipients with chronic illness. *J Clin Nurs.* 2021;30(9–10):1335–46.
21. Touré M, Kouakou CR, Poder TG. Dimensions used in instruments for QALY calculation: a systematic review. *Int J Environ Res Public Health.* 2021;18(9):4428.
22. National Institute for Health and Care Excellence. NICE health technology evaluations: the manual. Process and methods [PMG36]. 2022.
23. European Network for Health Technology Assessment (EUnetHTA). Methods for health economic evaluations—a guideline based on current practices in Europe. 2015.
24. Badia X, Roset M, Herdman M, Kind P. A comparison of United Kingdom and Spanish general population time trade-off values for EQ-5D health states. *Med Decis Mak.* 2001;21(1):7–16.
25. Ramos-Goñi JM, Craig BM, Oppe M, Ramallo-Fariña Y, Pinto-Prades JL, Luo N, Rivero-Arias O. Handling data quality issues to estimate the Spanish EQ-5D-5L value set using a hybrid interval regression approach. *Value Health.* 2018;21(5):596–604.
26. Abellán Perpiñán JM, Sánchez Martínez FI, Martínez Pérez JE, Méndez I. Lowering the ‘Floor’ Of The Sf-6d scoring algorithm using a lottery equivalent method. *Health Econ.* 2012;21(11):1271–85.
27. Cubi-Molla P, Shah K, Burström K. Experience-based values: a framework for classifying different types of experience in health valuation research. *Patient Patient Cent Outcomes Res.* 2018;11(3):253–70.
28. Rowen D, Mulhern B, Banerjee S, Tait R, Watchurst C, Smith SC, Brazier JE. Comparison of general population, patient, and carer utility values for dementia health states. *Med Decis Mak.* 2015;35(1):68–80 (**Neurology**, 88(22), 2123–2127).
29. 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(12):2149–58.
30. 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.
31. Hofman CS, Makai P, Boter H, Buurman BM, de Craen AJ, Olde Rikkert MG, Melis RJ. Establishing a composite endpoint for measuring the effectiveness of geriatric interventions based on older persons’ and informal caregivers’ preference weights: a vignette study. *BMC Geriatr.* 2014;14:1–8.
32. Michalowsky B, Hoffmann W, Mohr W, Rädke A, Xie F. Comparing the psychometric properties of EQ-5D-3L and EQ-5D-5L proxy ratings by informal caregivers and a health professional for people with dementia. *Health Qual Life Outcomes.* 2022;20(1):140.
33. Loewenstein G, Ubel PA. Hedonic adaptation and the role of decision and experience utility in public policy. *J Public Econ.* 2008;92(8–9):1795–810.
34. Augestad LA, Stavem K, Kristiansen IS, Samuelsen CH, Rand-Hendriksen K. Influenced from the start: anchoring bias in time trade-off valuations. *Qual Life Res.* 2016;25:2179–91.
35. Devlin N, Buckingham K, Shah K, Tsuchiya A, Tilling C, Wilkinson G, Van Hout B. A comparison of alternative variants of the lead and lag time TTO. *Health Econ.* 2013;22(5):517–32.
36. Lenert LA, Cher DJ, Goldstein MK, Bergen MR, Garber A. The effect of search procedures on utility elicitation. *Med Decis Mak.* 1998;18(1):76–83.
37. Pinto-Prades JL, Rodríguez-Míguez E. The lead time trade-off: the case of health states better than dead. *Med Decis Mak.* 2015;35(3):305–15.
38. Patrick DL, Starks HE, Cain KC, Uhlmann RF, Pearlman RA. Measuring preferences for health states worse than death. *Med Decis Mak.* 1994;14(1):9–18.
39. Devlin NJ, Hansen P, Kind P, Williams A. Logical inconsistencies in survey respondents’ health state valuations—a methodological challenge for estimating social tariffs. *Health Econ.* 2003;12(7):529–44.
40. Bravata DM, Nelson LM, Garber AM, Goldstein MK. Invariance and inconsistency in utility ratings. *Med Decis Mak.* 2005;25(2):158–67.
41. Ryan M, Netten A, Skåtun D, Smith P. Using discrete choice experiments to estimate a preference-based measure of outcome—an application to social care for older people. *J Health Econ.* 2006;25(5):927–44.
42. Walters SJ, Brazier JE. Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. *Qual Life Res.* 2005;14:1523–32.
43. Dolders MG, Zeegers MP, Groot W, Ament A. A meta-analysis demonstrates no significant differences between patient and population preferences. *J Clin Epidemiol.* 2006;59(7):653–64.
44. Peeters Y, Stiggelbout AM. Health state valuations of patients and the general public analytically compared: a meta-analytical comparison of patient and population health state utilities. *Value Health.* 2010;13(2):306–9.
45. Lobchuk MM, McClement SE, Daeninck PJ, Shay C, Elands H. Asking the right question of informal caregivers about patient symptom experiences: multiple proxy perspectives and reducing interrater gap. *J Pain Symptom Manag.* 2007;33(2):130–45.
46. Ubel PA, Loewenstein G, Schwarz N, Smith D. Misimagining the unimaginable: the disability paradox and health care decision making. *Health Psychol.* 2005;24(4S):S57.
47. Xie F, Pullenayegum E, Gaebel K, Bansback N, Bryan S, Ohinmaa A, Johnson JA. How different are composite and traditional TTO valuations of severe EQ-5D-5L states? *Qual Life Res.* 2016;25:2101–8.