



Estimating Informal Caregiving Time from Patient EQ-5D Data: The Informal CARE Effect (iCARE) Tool

Maria Gheorghe¹ · Renske J. Hoefman²  · Matthijs M. Versteegh¹ · Job van Exel^{2,3} 

© The Author(s) 2018

Abstract

Background Families and friends provide a considerable proportion of care for patients and elderly people. Caregiving can have substantial effects on caregivers' lives, health, and well-being. However, because clinical trials rarely assess these effects, no information on caregiver burden is available when evaluating the cost effectiveness of treatments.

Objective This study develops an algorithm for estimating caregiver time using information that is typically available in clinical trials: the EQ-5D scores of patients and their gender.

Methods Four datasets with a total of 8012 observations of dyads of caregivers and a gamma model with a log-link estimated with the Bayesian approach were used to estimate the statistical association between patient scores on the EQ-5D-3L dimensions and the numbers of hours of care provided by caregivers during the previous week. The model predicts hours of care as mean point estimates with 95% credible intervals or entire distributions.

Results Model predictions of hours of care based on the five EQ-5D dimensions ranged from 13.06 (12.7–14.5) h/week for female patients reporting no health problems but receiving informal care to 52.82 (39.38–66.26) for male patients with the highest level of problems on all EQ-5D dimensions.

Conclusions The iCARE algorithm developed in this study allows researchers who only have patient-level EQ-5D data to estimate the mean hours of informal care received per week, including a 95% Bayesian credible interval. Caregiver time can be multiplied with a monetary value for caregiving, enabling the inclusion of informal care costs in economic evaluations. We recommend using the tool for samples that fall within the confidence intervals of the characteristics of our samples: men (age range 47.0–104.2 years), women (age range 55–103 years).

Key Points for Decision Makers

Including caregiving costs in economic evaluations is necessary in order to avoid the risk of labelling an intervention as cost effective when it actually does not maximize social welfare.

The iCARE tool (<https://www.imta.nl/tools/>) can be used to include informal care in economic evaluations based on patients' EQ-5D data in situations where it is not feasible to collect caregiver data.

The iCARE tool enables the inclusion of informal care on the cost side of economic evaluations and should be used for samples that fall within the following age ranges: men (47–104.2 years), women (55–103 years).

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s40273-018-0706-6>) contains supplementary material, which is available to authorized users.

✉ Matthijs M. Versteegh
gheorghe@imta.eur.nl

- ¹ Institute for Medical Technology Assessment, Bayle Building, Office J8-31, PO Box 1738, 3000 DR Rotterdam, The Netherlands
- ² Erasmus School of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, The Netherlands
- ³ Erasmus School of Economics, Erasmus University Rotterdam, Rotterdam, The Netherlands

1 Introduction

Economic evaluations in healthcare provide decision makers with insights about the relative efficiency of different interventions for producing health. These evaluations compare incremental costs and effects, often expressed in terms of quality-adjusted life-years, to some threshold representing either societal willingness to pay or opportunity costs within the healthcare system. An important consideration in these evaluations concerns the types of costs and effects that are to be included in the analysis [1, 2]. Advocates of a ‘societal perspective’ propose that all costs and effects occurring in society at large need to be considered [3, 4], including those associated with informal caregiving [5–10]. Moreover, excluding informal care in economic evaluations can result in non-optimal healthcare decision making [6, 11].

Family and friends often provide a substantial number of hours of care per week to address patients’ care needs [12, 13]. Therefore, healthcare interventions can affect not only patients but also the individuals in their social network [14–18]. A patient’s health can directly influence the health and well-being of the patient’s family or friends, regardless of whether they are involved in caregiving [19–22]. In addition to this family effect of ill health, family members and friends can experience positive and negative consequences from providing care, known as the caregiving effect [22]. For example, many caregivers derive fulfilment from caregiving, but caregiving is also associated with depressive feelings and physical health problems, and it can be difficult to combine caregiving with other activities, such as paid work or family life [23–26].

Outcomes for caregivers can be included on either the effect side or the cost side in economic evaluations. On the effect side, caregiver outcomes can be included as health-related quality of life next to patients’ utilities in cost-utility analysis or as a separate outcome in terms of care-related quality of life or carer experiences in multi-criteria analysis [27, 28]. Alternatively, informal care can be addressed as one of the cost components in economic evaluations in the form of the opportunity costs of the time spent on caregiving, typically by multiplying caregivers’ time investment by a value for time [6, 11]. However, due to practical considerations and methodological difficulties in identifying, measuring, and valuing informal care, these costs are usually not included in economic evaluations [6, 11, 29], mainly because most clinical and observational studies only collect data about patients and not about their caregivers [17]. Because effects of interventions on carers are not usually measured, they remain unknown [6], potentially resulting in suboptimal resource allocation recommendations.

Recent research has attempted to predict caregiver outcomes using patient health data—for example, researchers

have used data collected for measuring the family impact of meningitis to estimate changes in carers’ health based on changes in patients’ health [30], and have also focused on predicting parents’ health-related quality of life based on health information about their ill children [31]. Although these studies provide compelling examples of how to include caregivers’ health-related quality of life data in economic evaluations, they are rather restrictive in the type of data they use, and they also report poor predictive performance for their statistical models. In addition to these findings, we also investigated the relationship between care-related quality-of-life data of caregivers using the CarerQol instrument (CarerQol-7D) [27] and patient-level data, but we found this relationship to be very weak. To overcome some of these challenges, studies have proposed accounting for informal care on the cost side. For example, some studies [3, 32] have developed statistical models for predicting the number of days of care using patients’ health data. However, using the number of days of care can only provide a general indication of the real informal care costs, since the actual time commitment on each day remains unknown.

This study aimed to develop a statistical model for predicting caregivers’ time investment using patients’ health data, to serve as an input for cost calculations in economic evaluations. The study builds on previous research in several ways. First, we combined several datasets that contain health data for a large variety of patient groups in the Netherlands. This allowed more generalizable results and made it possible to develop a platform that can be used for accounting for informal care in various economic evaluations. Second, we measured informal care costs based on the hours of care during a week rather than days per week, as in previous research, which provides a more accurate indication of informal care costs. Moreover, we used caregiver-reported time investment, while previous research used patients’ reports of their caregivers’ time investment [3, 32]. Hence, we aimed to develop a prediction model for the number of hours of care received by the patient based on information that is routinely available in clinical studies: the health status as measured by the EQ-5D and the gender of the patient. A unique feature of our study is that, based on our prediction model, we have constructed a Microsoft Excel tool, the informal CARE effect (iCARE) tool (<https://www.imta.nl/tools/>), which researchers can use to estimate hours of care in the absence of data on caregivers. The prediction model and the iCARE tool enable the inclusion of informal care on the cost side of economic evaluations that only have access to patient health data as measured by the EQ-5D.

2 Methods

2.1 Data

This study used 4 survey datasets with a total of 8012 observations. Here we briefly describe the datasets; the appendix provides a more thorough description (see electronic supplementary material [ESM]). Our dataset 1 ($n = 6482$) is the Older Persons and Informal Caregivers Survey Minimum Data Set (TOPICS-MDS, <http://www.topics-mds.eu>), a database of 26 research projects on care for older persons in the Netherlands that is part of the National Care for the Elderly Programme (NPO) [33, 34]. Dataset 2 ($n = 1244$) concerns a heterogeneous group of informal caregivers selected from a representative sample in terms of age and gender of the Netherlands' general population aged 18 years and over [35]. Dataset 3 ($n = 175$) was collected with written questionnaires from a heterogeneous group of informal caregivers identified through respite centres in the Netherlands in 2003 [28]. Dataset 4 ($n = 111$) comes from the Brabant Injury Outcome Surveillance (BIOS) study, a cohort study of hip fracture patients in the Netherlands [36].

In all four datasets, health status was measured with the three-level version of the EuroQoL-5D measure (EQ-5D-3L) developed by the EuroQol Group (<http://www.euroqol.org> [37]), the most commonly used measure of patient health status in economic evaluations and the instrument of choice of the National Institute for Health Care and Excellence (NICE) for supporting reimbursement decisions [38]. Responses to the EQ-5D-3L questionnaire were provided by either the patients themselves (datasets 1 and 4) or their caregivers (datasets 2 and 3). Depending on the dataset, different questions were used to gather information on the number of informal care hours per week. For datasets 1, 2 and 3, information on the hours of informal caregiving was collected through questions on care activities from the iMTA Valuation of Informal Care Questionnaire (iVICQ [39]). For datasets 1 and 2, three questions were posed regarding the number of informal care hours during the previous week that were spent on household activities (HDL), personal care activities (ADL) and practical support (IADL). In dataset 3, these three different types of activities were split into 16 different care tasks (e.g. preparing food, personal hygiene). In dataset 4, information was collected on six care activities covering HDL, ADL and IADL activities. In all datasets, the total number of informal caregiving hours indicates the sum of the hours spent on HDL, ADL, and IADL activities during the preceding week. The maximum number of hours of care was set to 126 h/week [27] because, although some caregivers report needing to be on standby 24/7 for care needs, caregivers also need time for other activities, such as their own personal care or sleep.

Table 1 describes the four datasets and shows that they differ in terms of sample size, age of the care receivers, health of the care receivers as indicated by the EQ-5D-3L, and mean number of hours of care provided by the caregiver. Figure 1 presents histograms of the number of hours of informal care reported for the previous week in each dataset. We observe that the distribution is positively and right skewed for all datasets. All patients in the analysis reported that they received informal care during the previous week. Nonetheless, in some instances, their caregivers reported having provided zero hours of care during the preceding week.

2.2 Analyses

We developed our model, which predicts the number of hours of informal care based on the EQ-5D-3L items, in four steps. In step 1, we selected the best regression method for our model. In step 2, we pooled the four datasets and selected the best model specification for the data (including testing for interaction terms). In step 3, we estimated the model developed in steps 1 and 2 using the Bayesian approach. Finally, in step 4, we used the results from the Bayesian estimation to develop the iCARE tool, which predicts hours of care (i.e., point estimates for the mean and the entire distribution of 10,000 samples obtained from the Markov chain Monte Carlo [MCMC] simulation). We used the library `gamlss` in R [40, 41] in steps 1 and 2, and we used `Winbugs` [42] in step 3. Here we briefly describe each step; details appear in the appendix (see ESM).

In step 1, we compared the linear and gamma models using training and validation sets obtained from the four datasets. At this stage, we modelled the number of hours of care as a function of gender and the EQ-5D-3L items. To do this, we randomly trained the model on a pooled dataset formed by collapsing three datasets and then validated this model with the remaining dataset (the four datasets provided four such training and validation sets). We also developed a training and validation set by splitting the pooled dataset obtained from merging the four datasets in half, through random sampling. To assess the quality of the prediction, we used the mean absolute error (MAE), the mean square error (MSE), the root mean square error (RMSE) [43] and the original Akaike information criterion (AIC) [44] for model comparison. We developed regression models both with and without intercepts. Although models without an intercept are generally not recommended, as they can introduce bias [45], they are appealing for our setting since it seems reasonable to assume that an individual in full or perfect health as indicated by the best EQ-5D-3L state may require 0 h of informal care. However, we found that the gamma model with an intercept produced the smallest prediction error (see Appendix Table 2 in the ESM).

Table 1 Summary statistics for the four datasets used in this study

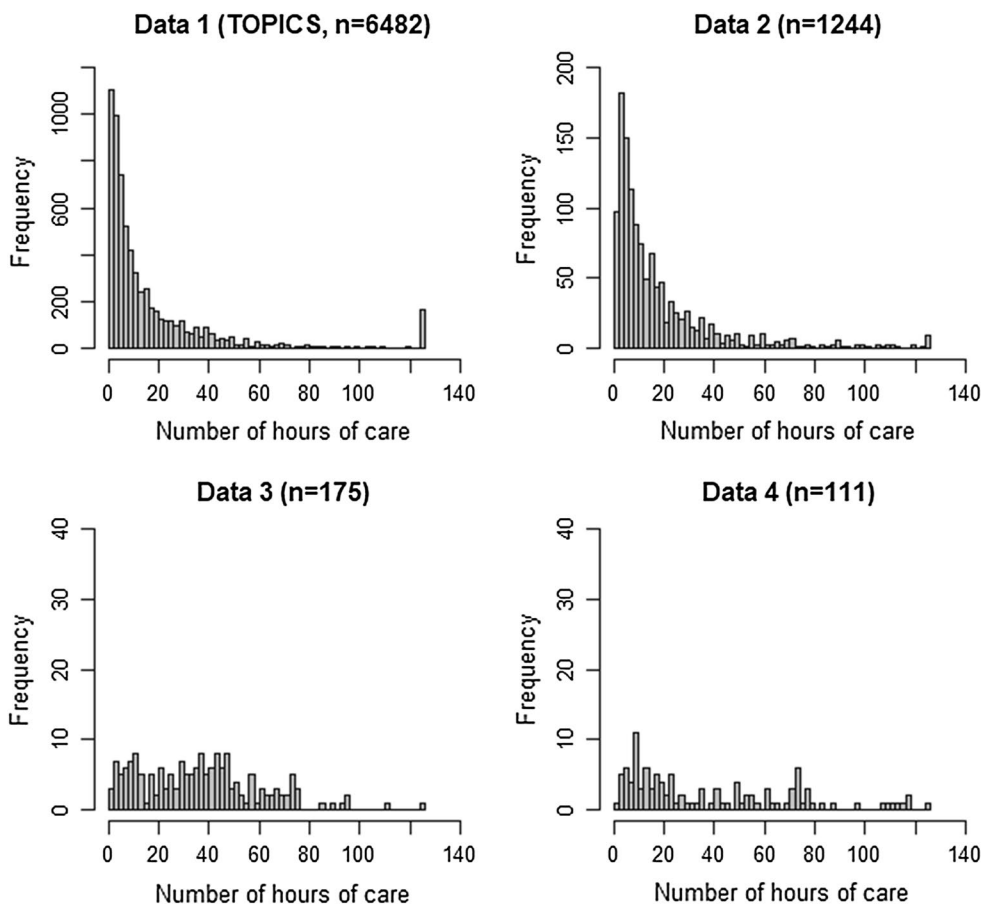
	Dataset 1 (<i>n</i> = 6482)		Dataset 2 (<i>n</i> = 1244)		Dataset 3 (<i>n</i> = 175)		Dataset 4 (<i>n</i> = 111)		All four datasets (<i>n</i> = 8012) Men and women
	Men	Women	Men	Women	Men	Women	Men	Women	
Number of observations	2310	4172	421	823	88	87	28	83	8012
Minimum number of hours of care	0	0	1	1	0	0	3	0.5	0
Mean number of hours of care	19.8	14.1	21.6	16.7	41.9	30.9	42	37.9	17.3
Maximum number of hours of care	126	126	126	126	111.8	126	108	126	126
Minimum age (year) of care receiver	54	55	2	1	6	8	65.1	66.4	1
Mean age (year) of care receiver	79.3	81.4	56.9	67.3	66	72.1	77.4	80.2	77.8
Maximum age (year) of care receiver	98	103	99	99	100	95	95.1	97.1	103
% with 0 h of care	1	0.86	0	0	1.14	1.15	0	0	0.76
% with 126 h of care	1.56	0.6	0.95	0.61	0	1.15	0	1.2	0.9
EQ-5D items									
% Mobility 1 (no problems)	28	24	33	19	17	16	43	44	26
% Mobility 2 (some problems)	68	71	59	73	66	63	57	54	69
% Mobility 3 (a lot of problems)	4	5	8	8	17	21	0	2	5
% Anxiety 1 (no problems)	67	63	44	47	27	23	71	75	61
% Anxiety 2 (some problems)	29	32	42	43	58	56	29	19	33
% Anxiety 3 (a lot of problems)	4	5	14	10	15	21	0	6	6
% Self care 1 (no problems)	57	58	41	40	23	19	78	63	54
% Self care 2 (some problems)	31	29	42	45	38	41	18	27	32
% Self care 3 (a lot of problems)	12	13	17	15	39	40	4	10	14
% Daily activities 1 (no problems)	43	38	20	15	7	1	57	48	35
% Daily activities 2 (some problems)	39	45	54	60	35	37	36	40	45
% Daily activities 3 (a lot of problems)	18	17	26	25	58	62	7	12	20
% Pain 1 (no pain)	46	33	35	21	22	24	64	54	35
% Pain 2 (some pain)	45	53	48	56	61	62	36	39	51
% Pain 3 (a lot of pain)	9	14	17	23	17	14	0	7	14

In step 2, we pooled the four datasets and used the gamma method to select the best model specification by comparing models, including interaction terms for both the mean and the scale parameters of the gamma models. We minimized the generalized AIC (GAIC) for different penalties: $GAIC = -2 * l(\theta) + df * \#$, where $-2 * l(\theta)$ is the fitted deviance, df denotes the total degrees of freedom, and $\#$ is the penalty for each degree of freedom used in the model [41, 46]. We selected the model specification based on the GAIC with penalties $\# = 2, 5, 9$; $\# = 2$ produces the original AIC [44], $\# = 9$ gives the Schwarz Bayesian information criterion (SBC), and we also considered the GAIC with a penalty of 5, a value falling between the AIC and the SBC. We found that the best model specification as indicated by the GAIC with various penalties is one in which the shape parameter is modelled as a log-link function of the EQ-5D-3L dummy variables and the gender of the patient (see Appendix Table 3 in the ESM). In fact, because gender as a main effect as well as in interaction with other variables produced large benefits in terms of model fit, we opted to fit separate models for men and women. We found that no

other interactions were significant and that adding model specifications for the scale parameter did not significantly improve the model fit.

To develop a model that provides predictions while offering a suitable measure of uncertainty associated with unknown future events, we used a Bayesian approach in step 3 to estimate probability distributions over future events. We used the MCMC algorithm, which yielded a Monte Carlo sample from the posterior predictive distribution over future quantities. We used the Bayesian estimates in step 4 to develop the iCARE tool, which predicts hours of informal care for a caregiver using the patient's EQ-5D-3L data. The tool includes estimations of hours of informal care based on both EQ-5D-3L and EQ-5D-5L data, where the 5L data are transformed to 3L data using the probability matrix from the 5L to 3L cross-walk [47]. Note that the predictions of caregivers' hours of informal care are based on a sample of individuals who all receive care. Since, in many populations, not all individuals receive care, the estimate produced by the toolkit must be corrected for the proportion of patients that receive care before using the estimate in cost-effectiveness

Fig. 1 Histograms of the four datasets



analyses. To further exemplify its use, below we apply the toolkit to a previously published cost-effectiveness model for multiple sclerosis and demonstrate the impact of including caregiver burden on the incremental cost-effectiveness ratio.

3 Results

Figure 2 shows posterior distributions of the 10,000 MCMC simulation samples for the regression coefficients of the mean hours of care (μ), modelled separately for male and female patients. The regression coefficients consist of an intercept and EQ-5D-3L responses at levels 2 (some problems) and 3 (a lot of problems) relative to the baseline level 1 (no problems) for each EQ-5D-3L question. The distributions of the regression coefficients estimated for level 3 relative to level 1 were wider (illustrating more variation or higher uncertainty) and shifted to the right compared with those estimated for level 2 relative to level 1. For some EQ-5D-3L items, the two distributions largely overlapped. In the model for men, the EQ-5D-3L estimates for levels 2 and 3 overlapped for anxiety, mobility, and pain, while they overlapped for anxiety and self-care in the model for women. In general, the relationship between the hours of informal

care and the EQ-5D-3L score was stronger for men than for women.

Table 2 illustrates the summary regression coefficient estimates (mean with 95% credible intervals, median, MC error) based on the 10,000 MCMC sample distributions. Table 2 shows that, for both men and women, the coefficients at levels 2 and 3 for most EQ-5D-3L questions had a positive sign relative to the item responses indicated at level 1, suggesting that the number of informal care hours increased as the health state indicated by the EQ-5D-3L score worsened. There are exceptions for a few coefficients at level 2, which had a negative sign relative to level 1 (usual activities for men and pain for women); however, these coefficients were not statistically significant and, as indicated by Fig. 2, their posterior distributions overlapped with the coefficient distributions at level 3. Moreover, in both models (men and women), the coefficient for mobility at level 2 was negative, which means that our models cannot distinguish between mobility levels 1 and 2. For these situations, we considered levels 1 and 2 as interchangeable in the iCARE tool and included only changes of level 1 or level 2 relative to level 3.

Figure 3 compares posterior hours-of-care predictions (distribution and mean) for individuals in the best health state (11111) and in the worst health state (33333), as

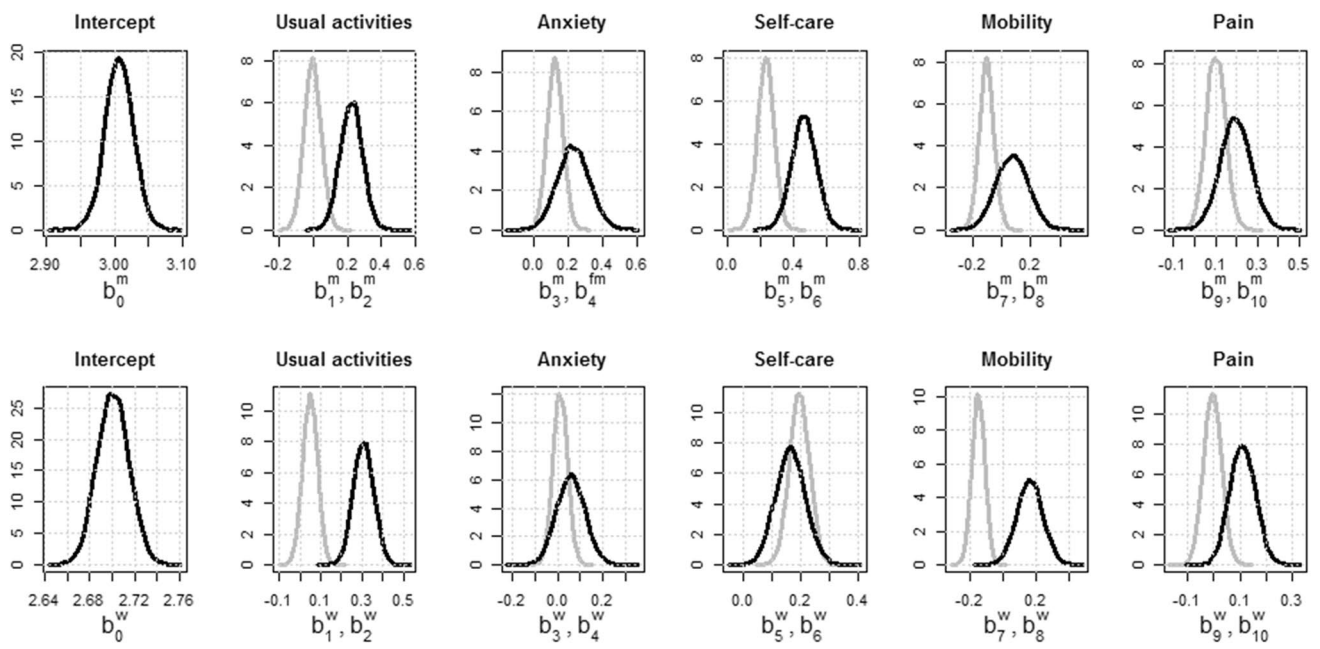


Fig. 2 Posterior distributions (10,000 MCMC samples) of the regression coefficients of the mean μ for the models for male patients (top) and for female patients (bottom). Coefficients for the EQ-5D item levels 2 (grey lines) and 3 (black lines) are relative to the baseline category level 1

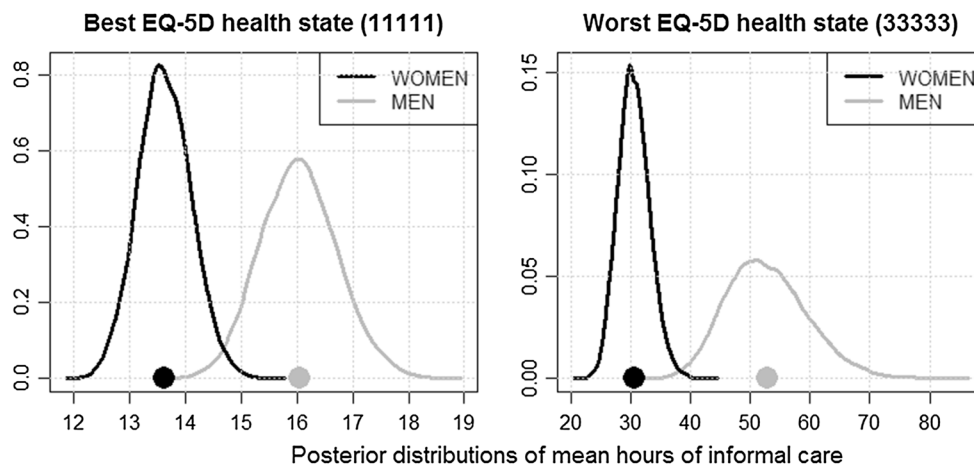
Table 2 Posterior summary estimates of the regression coefficients for the mean μ on log-link scale

Coefficient	Mean	SD	MC error	2.50%	Median	97.50%
Model for male patients						
b_0^m (intercept)	3.0070*	0.0205	0.0002	2.9660	3.0070	3.0470
b_1^m (ACT2)	-0.0048	0.0500	0.0005	-0.1019	-0.0051	0.0940
b_2^m (ACT3)	0.2228*	0.0655	0.0008	0.0935	0.2233	0.3513
b_3^m (ANX2)	0.1205*	0.0464	0.0005	0.0312	0.1202	0.2127
b_4^m (ANX3)	0.2298*	0.0929	0.0009	0.0515	0.2285	0.4147
b_5^m (SC2)	0.2307*	0.0498	0.0005	0.1329	0.2310	0.3266
b_6^m (SC3)	0.4622*	0.0748	0.0008	0.3144	0.4621	0.6078
b_7^m (MO2)	-0.1030*	0.0504	0.0006	-0.2022	-0.1032	-0.0022
b_8^m (MO3)	0.0728	0.1094	0.0013	-0.1386	0.0738	0.2857
b_9^m (PAIN2)	0.1013*	0.0463	0.0005	0.0109	0.1008	0.1925
b_{10}^m (PAIN3)	0.1969*	0.0741	0.0008	0.0523	0.1961	0.3444
Model for female patients						
b_0^w (intercept)	2.7010*	0.0143	0.0002	26.730	27.010	27.290
b_1^w (ACT2)	0.0494	0.0359	0.0005	-0.0214	0.0497	0.1184
b_2^w (ACT3)	0.3071*	0.0493	0.0006	0.2124	0.3070	0.4032
b_3^w (ANX2)	0.0119	0.0324	0.0003	-0.0516	0.0115	0.0763
b_4^w (ANX3)	0.0584	0.0628	0.0007	-0.0619	0.0585	0.1831
b_5^w (SC2)	0.1955*	0.0355	0.0005	0.1260	0.1953	0.2638
b_6^w (SC3)	0.1643*	0.0524	0.0007	0.0640	0.1640	0.2682
b_7^w (MO2)	-0.1527*	0.0392	0.0005	-0.2303	-0.1526	-0.0778
b_8^w (MO3)	0.1611*	0.0789	0.0008	0.0071	0.1610	0.3141
b_9^w (PAIN2)	-0.0052	0.0345	0.0004	-0.0721	-0.0051	0.0616
b_{10}^w (PAIN3)	0.1102*	0.0491	0.0005	0.0164	0.1095	0.2076

MC Monte Carlo, SD standard deviation

*Significant; credible interval does not include zero

Fig. 3 Posterior predictions of mean hours of informal care for the best and the worst health states as indicated by the EQ-5D. The large dots represent the posterior mean distribution estimates



indicated by the EQ-5D-3L instrument. For male patients, the predictions of the mean hours of informal care per week varied between 16.0 (95% CI 14.7–17.4) and 52.8 (95% CI 39.4–66.3) for the best (11111) and the worst (33333) health state. For female patients, the variation was between 13.6 (95% CI 12.7–14.5) for the best (11111) and 32.0 (95% CI 26.2–37.8) for the worst (33333) health state. Hence, in line with the raw dataset, the numbers of hours of informal care predicted for men were considerably greater than for women.

We imported the estimated probability distributions using the Bayesian approach from Winbugs to Microsoft Excel and developed the iCARE tool to provide users with entire estimated probability distribution predictions and corresponding point estimates of the mean hours of informal care based on the EQ-5D-3L items, with 95% credible intervals.

We used the iCARE tool in a case study to estimate caregiver burden for multiple sclerosis (MS) patients, using data from a previously published cost-effectiveness model [48]. The model had four health states based on the Expanded Disability Status Scale (EDSS): EDSS1 (EDSS 0–2.5),

EDSS2 (EDSS 3–5.5), EDSS3 (EDSS 6–7.5), and EDSS4 (EDSS 8–9.5), and two relapse states. Observed EQ-5D-3L data for patients in each of the four EDSS-based health states ($n = 1295$) were available, and we used these to calculate caregiver hours per week for each health state. Since not all of the MS patients received informal care (especially those in the better EDSS states), and the iCARE algorithm is based on a sample of patients who all receive care, the estimated values had to be corrected for the proportion of patients who received care. For this, we used values for the proportion of patients in each EDSS class who receive care as provided in [49]. We translated the resulting estimate into costs per month using a 14 Euro (€) per hour cost price for caregiver time, as recommended in the Dutch costing manual [50]. The reference price of 1 h of informal care here is the market value of household activities as stated by the Central Administration Office (CAK) in the Netherlands. Costs and effects were discounted by 3% annually. Table 3 shows that including caregiver burden in this example reduces the lifetime incremental costs of glatiramer acetate versus symptom

Table 3 Example application incorporating estimated caregiver burden (monthly caregiver costs)

Health state	1 (EDSS 0–2.5)	2 (EDSS 3–5.5)	3 (EDSS 6–7.5)	4 (EDSS 8–9.5)	Relapse 1	Relapse 2
Mean EQ-5D utility (Dutch tariff)	0.74	0.63	0.52	0.33	0.65	0.54
N	371	557	278	19		
Hours per week	15.03	15.68	16.96	23.59		
Hours per month	60.12	62.73	67.84	94.355		
Proportion receiving informal care ^a	22%	58%	58%	75%		
Costs per month	€186.00	€508	€549	€990.73	Equal to EDSS2	Equal to EDSS3
SD costs per month	€37.10	€190.27	€212.54	€321.33		
ICER without estimated informal care						€161,319
ICER with estimated informal care						€145,265

EDSS Expanded Disability Status Scale, ICER Incremental cost effectiveness ratio, SD standard deviation

^aBased on Uitdehaag et al. [49]

management from €22,369 to €20,774; consequently, the incremental cost-effective ratio (ICER) was reduced by approximately 10% (from €161,319 to €145,265).

4 Discussion

This study shows that the relationship between the number of hours of care indicated by a caregiver and the EQ-5D-3L score of the care receiver can successfully be used to develop a prediction model that estimates informal care effects for economic evaluations in clinical studies when these data are unavailable. We developed the iCARE tool to derive predictions (mean, 95% credible intervals and entire distributions based on 10,000 MCMC samples) of hours of informal care using patient-level EQ-5D-3L data.

Our results show that informal care effects are higher for patients with lower health scores, in line with findings of previous studies [3, 32]. Moreover, our results indicate that men required more hours of informal care than women. The predicted mean hours of informal care per week for male patients varied between 16.0 (95% CI 14.7–17.4) for the best (11111) and 52.8 (95% CI 39.4–66.3) for the worst (33333) EQ-5D-3L health state. For female patients, the variation was between 13.6 (95% CI 12.7–14.5) for the best (11111) and 32.0 (95% CI 26.2–37.8) for the worst (33333) health state.

Our study has several noteworthy strengths. First, it contributes to the economic evaluation field by enabling researchers to estimate the effects of interventions on informal carers. Considering the increasing appeal to the public in many countries to contribute to the care of their loved ones, it is important to account for the effects thereof, irrespective of the perspective adopted in the evaluation of healthcare interventions. It is evident that these effects should be included in evaluations from the societal perspective. But considering the mounting evidence of the effects of providing care on the health and wellbeing of informal caregivers, evaluations from a healthcare perspective should also consider examining and presenting these effects alongside the health effects of interventions on patients. After obtaining information on the hours of informal care using the iCARE tool and correcting this estimate for the proportion of the sample that is receiving care, the next step for researchers is to multiply the hours of informal care by a value for time spent on informal care. This value can be derived from health economic manuals or the literature; for example, by using the value of a close market substitute who could be hired to perform the specific caregiving tasks [51] or values from stated preference methods, such as contingent valuation studies or discrete choice experiments [27, 51–53]. Second, from a statistical perspective, we have addressed some important challenges associated with the development

of the prediction model. We used a variety of datasets in this study that include information on hours of care and the EQ-5D-3L items, allowing us to test the robustness of the prediction models for various populations. Furthermore, we adopted a Bayesian estimation procedure that enabled us to obtain suitable measures of uncertainty through use of the MCMC algorithms. This is important, especially for using the prediction model results in economic evaluations, as it will allow proper propagation of uncertainty in these models.

Our study also has several limitations. First, some of the datasets did not provide self-completed information on the care receiver's EQ-5D-3L: for two of the four datasets, accounting for about 18% of the pooled dataset used to estimate the prediction model, the EQ-5D-3L questionnaire was completed by the caregiver instead of the care receiver, which may have introduced bias in our analyses. Second, we only had data on the hours provided by the primary caregiver, while patients could have had more than just one caregiver and hence may have received more hours of care in total. The model we developed is not conditional on the number of caregivers that a patient might have had but only on patient gender and EQ-5D-3L items. This means that the estimated number of hours of care from the model is independent of the number of caregivers. Third, the four datasets are rather unbalanced in terms of sample size, as they represented 81%, 15.5%, 2.1%, and 1.4% of the pooled dataset. Obviously, the pooled estimate was dominated by the larger dataset (TOPICS). This might have cancelled some of the diversity introduced by the different datasets, although the largest dataset comprised 26 different studies in a variety of patient populations. However, the pooled databases, while adding more heterogeneity, result in a model with similar errors and increased variability (or range of model predictions) as the one using TOPICS data only, and this is potentially beneficial for developing a toolkit that could be used in diverse health technology assessments. Fourth, the model is based on data from the Netherlands only. The validity of the model's predictions for other countries with different healthcare systems and perhaps different informal caregiving needs and traditions is therefore unclear. Hence, we recommend investigating country-specific tools using local data. Fifth, we purposefully restricted the prediction model to information that is routinely available in clinical studies. Including only patient EQ-5D and gender was an a priori decision made considering that the main purpose of the tool was to be used in studies in which no information about the caregiving is available. More often than not, information about the caregiver is not available in clinical trials and observational studies; hence, we focused on developing a tool that would be useful in studies where only patient data is available. Besides, we expect that in

studies where information about the caregiver is collected, this will likely include information about hours of care and there would therefore be no use for this tool. We also developed models including the relationship between the number of hours of care and the age of the patient; however, we found that this was difficult to explain and did not improve model fit. In addition, we developed models including other explanatory variables such as relationship of the care receiver with the caregiver (partner, son/daughter, other), age of the caregiver, if the care receiver is institutionalized or not, co-habitation of care receiver with caregiver. We only found a significant relationship for the relationship between the caregiver and the care receiver and the number of hours of care. However, as mentioned, we decided not to include more variables in the model as such specific information is unlikely to be generally available in the clinical trials and this would highly restrict the usability of the tool. Lastly, since the iCARE tool predicts only the mean hours of care through the health status and gender of the patient, the tool is not an appropriate instrument for interventions where other aspects beyond the scope of the health of patients can influence care needs; for example, interventions that target the location of services. For instance, an intervention that reduces the length of a hospital stay will have a major impact on the hours of informal care but may not directly affect the health of the patient, and hence the effect would not be captured by our model. Without disregarding these important considerations, the iCARE tool can be used to include informal care in economic evaluations based on patients' EQ-5D data in research situations where it is not feasible to collect caregiver data. Furthermore, although the estimates are based on relatively weak associations, it is important to note that this may pose a problem for predictions at an individual level but not so much a problem for estimating the mean hours of care within a population of carers, as we do here. We recommend using the tool for samples that fall within the confidence intervals of the characteristics of our samples: men (age range: 47–104.2), women (age range: 55–103).

It is worth noting that there are equity implications from using different models for estimating caregiver burden for male and female patients. However, these are not a consequence of the model developed here, but a consequence of including caregiving costs in health economic models in general, as these will differ between men and women. For example, the Dutch costing manual for productivity losses specifies a different hourly wage for men and women, resulting in equity implications upon inclusion of productivity losses. However not all countries follow such recommendations. Contrary to the Dutch guidelines, the US guidelines do not recommend the use of gender-specific wage rates because of potential bias. As such, we acknowledge that the

developed model here takes into consideration countries with similar recommendations to the Netherlands. We recommend taking this aspect into consideration when using the tool.

5 Conclusions

Informal caregivers make an important contribution to societal welfare. Including caregiving costs in economic evaluations is necessary in order to avoid the risk of labelling an intervention as cost effective when it actually does not maximize social welfare. This is supported by many national guidelines for cost-effectiveness studies advising researchers to adopt a societal perspective when identifying the relevant costs and effects of an intervention. This study is important because it facilitates the inclusion of informal caregiver effects in economic evaluations that lack this information. However, the algorithm we present is only a 'second-best' alternative, as actual data on caregivers will obviously be more accurate than predictions, and therefore preferable. Finally, although one may question the validity of the predictions of a model based on data from the Netherlands when used for other countries, the ranges of caregiver time seem reasonable, and we would argue that it is better to have a reasonable estimate than to completely neglect the effects of interventions on informal carers.

Author contributions MG performed the analysis and wrote the manuscript. RJH, MMV and JvE provided feedback on the analysis and contributed to writing the paper.

Compliance with Ethical Standards

Data Availability Statement The data that support the findings of this study are available but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are, however, available from the authors upon reasonable request and with permission from ZonMw.

Funding This study was funded by by ZonMw (Grant number 60-63300-98-506).

Conflict of interest MG has no conflicts of interest. RJH has no conflicts of interest. MMV has no conflicts of interest. JvE has no conflicts of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Drummond MF, Sculpher MJ, Torrance GW. *Methods for the economic evaluation of health care programmes*. 3rd ed. Oxford: Oxford University Press; 2005.
2. Brouwer WB, van Exel NJ, Baltussen RM, Rutten FF. A dollar is a dollar is a dollar—or is it? *Value Health*. 2006;9(5):341–7.
3. Dixon S, Walker M, Salek S. Incorporating carer effects into economic evaluation. *Pharmacoeconomics*. 2006;24(1):43–53.
4. Jacobs P, Ohinmaa A, Brady B. Providing systematic guidance in pharmacoeconomic guidelines for analysing costs. *Pharmacoeconomics*. 2005;23(2):143–53.
5. Brouwer WB, Koopmanschap MA. On the economic foundations of CEA: ladies and gentlemen, take your positions! *J Health Econ*. 2000;19(4):439–59.
6. 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(2):123–35.
7. Jones C, Edwards RT, Hounsborne B. Health economics research into supporting carers of people with dementia: a systematic review of outcome measures. *Health Qual Life Outcomes*. 2012;10:142-7525-10-142.
8. Wiseman V. Caring: the neglected health outcome? Or input? *Health Policy*. 1997;39(1):43–53.
9. Davidson T, Levin LA. Is the societal approach wide enough to include relatives? Incorporating relatives' costs and effects in a cost-effectiveness analysis. *Appl Health Econ Health Policy*. 2010;8(1):25–35.
10. Al-Janabi H, Coast J, Flynn TN. What do people value when they provide unpaid care for an older person? A meta-ethnography with interview follow-up. *Soc Sci Med*. 2008;67(1):111–21.
11. Goodrich K, Kaambwa B, Al-Janabi H. The inclusion of informal care in applied economic evaluation: a review. *Value Health*. 2012;15(6):975–81.
12. Verbeek-Oudijk D, Woittiec I, Eggink E, Putnam L. Who cares in Europe? *Geron*. 2015;17(1):62–5.
13. Colombo F, Llana-Nozal A, Mercier J, Tjadens F. *Help wanted? Providing and paying for long-term care*. Paris: OECD Publishing; 2011.
14. Brouwer WB. Too important to ignore: informal caregivers and other significant others. *Pharmacoeconomics*. 2006;24(1):39–41.
15. Culyer AJ. The nature of the commodity 'health care' and its efficient allocation. *Oxf Econ Pap*. 1971;23(2):189–211.
16. Basu A, Meltzer D. Implications of spillover effects within the family for medical cost-effectiveness analysis. *J Health Econ*. 2005;24(4):751–73.
17. Al-Janabi H, Nicholls J, Oyeboode J. The need to “carer proof” healthcare decisions. *BMJ*. 2016;352:i1651.
18. Wittenberg E, Prosser LA. Health as a family affair. *N Engl J Med*. 2016;374(19):1804–6.
19. Al-Janabi H, van Exel J, Brouwer W, Trotter C, Glennie L, Hannigan L, Coast J. Measuring health spillovers for economic evaluation: a case study in meningitis. *Health Econ*. 2016;25(12):1529–44.
20. Bobinac A, van Exel NJ, Rutten FF, Brouwer WB. Health effects in significant others: separating family and care-giving effects. *Med Decis Mak*. 2011;31(2):292–8.
21. Wittenberg E, Ritter GA, Prosser LA. Evidence of spillover of illness among household members: EQ-5D scores from a US sample. *Med Decis Mak*. 2013;33(2):235–43.
22. Bobinac A, van Exel NJ, Rutten FF, Brouwer WB. Caring for and caring about: disentangling the caregiver effect and the family effect. *J Health Econ*. 2010;29(4):549–56.
23. Brouwer WB, van Exel NJ, van den Berg B, van den Bos GA, Koopmanschap MA. Process utility from providing informal care: the benefit of caring. *Health Policy*. 2005;74(1):85–99.
24. van der Lee J, Bakker TJ, Duivenvoorden HJ, Dröes RM. Multivariate models of subjective caregiver burden in dementia: a systematic review. *Ageing Res Rev*. 2014;15:76–93.
25. Beach SR, Schulz R, Yee JL, Jackson S. Negative and positive health effects of caring for a disabled spouse: longitudinal findings from the caregiver health effects study. *Psychol Aging*. 2000;15(2):259–71.
26. Pinquart M, Sorensen S. Differences between caregivers and non-caregivers in psychological health and physical health: a meta-analysis. *Psychol Aging*. 2003;18(2):250–67.
27. Hoefman RJ, van Exel J, Brouwer W. How to include informal care in economic evaluations. *Pharmacoeconomics*. 2013;31(12):1105–19.
28. Brouwer WB, van Exel NJ, van Gorp B, Redekop WK. The Carer-Qol instrument: a new instrument to measure care-related quality of life of informal caregivers for use in economic evaluations. *Qual Life Res*. 2006;15(6):1005–21.
29. Stone PW, Chapman RH, Sandberg EA, Liljas B, Neumann PJ. Measuring costs in cost-utility analyses: variations in the literature. *Int J Technol Assess Health Care*. 2000;16(1):111–24.
30. Al-Janabi H, Manca A, Coast J. Predicting carer health effects for use in economic evaluation. *PLoS One*. 2017;12(9):e0184886.
31. Defenderfer EK, Rybak TM, Davies WH, Berlin KS. Predicting parent health-related quality of life: evaluating conceptual models. *Qual Life Res*. 2017;26(6):1405–15.
32. Rowen D, Dixon S, Hernández-Alava M, Mukuria C. Estimating informal care inputs associated with EQ-5D for use in economic evaluation. *Eur J Health Econ*. 2016;17(6):733–44.
33. Lutomski JE, Baars MA, Schalk BW, Boter H, Buurman BM, den Elzen WP, Jansen AP, Kempen GI, Steunenberg B, Steyerberg EW, et al. The development of the Older Persons and Informal Caregivers Survey Minimum DataSet (TOPICS-MDS): a large-scale data sharing initiative. *PLoS One*. 2013;8(12):e81673.
34. Lutomski JE, van Exel NJ, Kempen GI, Moll van Charante EP, den Elzen WP, Jansen AP, Krabbe PF, Steunenberg B, Steyerberg EW, Olde Rikkert MG, Melis RJ. Validation of the care-related quality of life instrument in different study settings: findings from The Older Persons and Informal Caregivers Survey Minimum DataSet (TOPICS-MDS). *Qual Life Res*. 2015;24(5):1281–93.
35. Hoefman RJ, van Exel J, Brouwer WB. Measuring the impact of caregiving on informal carers: a construct validation study of the CarerQol instrument. *Health Qual Life Outcomes*. 2013;11:173.
36. de Jongh MA, Kruithof N, Gosens T, van de Ree CL, de Munter L, Brouwers L, Polinder S, Lansink KW. Prevalence, recovery patterns and predictors of quality of life and costs after non-fatal injury: the Brabant Injury Outcome Surveillance (BIOS) study. *Inj Prev*. 2017;23(1):59. <https://www.ncbi.nlm.nih.gov/pubmed/27154507>. Accessed 24 Feb 2018.
37. Kind P, Brooks R, Rabin R, editors. *EQ-5D concepts and methods: a developmental history*. Dordrecht: Springer; 2005.
38. NICE. *Guide to the methods of technology appraisal 2013*. London: National Institute of Health and Care Excellence; 2013.
39. Hoefman R, Van Exel N, Brouwer WBF. iVICQ: iMTA Valuation of Informal Care questionnaire. Updated 2011. 2011. <https://www.imta.nl/carerqol/>. Accessed 02 Dec 2013.
40. Rigby RA, Stasinopoulos DM. Generalized additive models for location, scale and shape. *J R Stat Soc Ser C (Appl Stat)*. 2005;54(3):507–54.
41. Rigby B, Stasinopoulos M. A flexible regression approach using GAMLSS in R; 2009. <https://pdfs.semanticscholar.org/ed3d/f96aa477d6454d5737403a5aab751ccd9315.pdf>

42. Spiegelhalter D, Thomas A, Nicky Best N, Lunn D. WinBUGS user manual: version 1.4. Cambridge: MRC Biostatistics Unit; 2003.
43. Longworth L, Rowen D. The use of mapping methods to estimate health state utility values. Support document 10. London: National Institute for Clinical Excellence; 2011.
44. Akaike H. Information theory and an extension of the maximum likelihood principle. In: Parzen E, Tanabe K, Kitagawa G, editors. Selected papers of Hirotugu Akaike. Springer series in statistics (perspectives in statistics). New York: Springer; 1988.
45. Gelman A, Hill J, editors. Data analysis using regression and multilevel/hierarchical models. Cambridge: Cambridge University Press; 2007.
46. Stasinopoulos DM, Rigby RA. Generalized additive models for location scale and shape (GAMLSS) in R. *J Stat Softw*. 2007;23(7):1–46.
47. van Hout B, Janssen MF, Feng YS, Kohlmann T, Busschbach J, Golicki D, Lloyd A, Scalone L, Kind P, Pickard AS. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. *Value Health*. 2012;15(5):708–15.
48. Versteegh M. Impact on the incremental cost-effectiveness ratio of using alternatives to EQ-5D in a Markov model for multiple sclerosis. *Pharmacoeconomics*. 2016;34(11):1133–44.
49. Uitdehaag B, Kobelt G, Berg J, Capsa D, Dalén J. New insights into the burden and costs of multiple sclerosis in Europe: results for the Netherlands. *Mult Scler*. 2017;23(2_suppl):117–29.
50. Kanters TA, Clazien AM, Bouwmans N, van der Linden N, Swan Tan S, Hakkaart-van Roijen L. Update of the Dutch manual for costing studies in health care. *PLoS One*. 2017;12(11):e0187477.
51. van den Berg B, Brouwer W, van Exel J, Koopmanschap M, van den Bos GA, Rutten F. Economic valuation of informal care: lessons from the application of the opportunity costs and proxy good methods. *Soc Sci Med*. 2006;62(4):835–45.
52. de Meijer C, Majer IM, Koopmanschap MA, van Baal PH. Forecasting lifetime and aggregate long-term care spending accounting for changing disability patterns. *Med Care*. 2012;50(8):722–9.
53. Mentzakis E, Ryan M, McNamee P. Using discrete choice experiments to value informal care tasks: exploring preference heterogeneity. *Health Econ*. 2011;20(8):930–44.
54. Mihaylova B, Briggs A, O'Hagan A, Thompson SG. Review of statistical methods for analysing healthcare resources and costs. *Health Econ*. 2011;20:897–916.
55. Cooper NJ, Lambert PC, Abrams KR, Sutton AJ. Predicting costs over time using Bayesian Markov chain Monte Carlo methods: an application to early inflammatory polyarthritis. *Health Econ*. 2007;16:37–56.
56. Mullahy J. Specification and testing of some modified count data models. *J Econom*. 1986;33(3):341–65.
57. Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences. *Stat Sci*. 1992;7(4):457–72.