Eur J Health Econ
DOI 10.1007/s10198-015-0718-5

ORIGINAL PAPER

Estimating informal care inputs associated with EQ-5D for use in economic evaluation

Donna Rowen[1] · Simon Dixon[1] · Mónica Hernández-Alava[1] · Clara Mukuria[1]

Received: 3 September 2014 / Accepted: 22 July 2015

Abstract
Objectives This paper estimates informal care need using the health of the patient. The results can be used to predict changes in informal care associated with changes in the health of the patient measured using EQ-5D.
Methods Data was used from a prospective survey of inpatients containing 59,512 complete responses across 44,494 individuals. The number of days a friend or relative has needed to provide care or help with normal activities in the last 6 weeks was estimated using the health of the patient measured by EQ-5D, ICD chapter and other health and sociodemographic data. A variety of different regression models were estimated that are appropriate for the distribution of the informal care dependent variable, which has large spikes at 0 (zero informal care) and 42 days (informal care every day).
Results The preferred model that most accurately predicts the distribution of the data is the zero-inflated negative binomial with variable inflation. The results indicate that improving the health of the patient reduces informal care need. The relationship between ICD chapter and informal care need is not as clear.
Conclusions The preferred zero-inflated negative binomial with variable inflation model can be used to predict changes in informal care associated with changes in the health of the patient measured using EQ-5D and these results can be applied to existing datasets to inform economic evaluation. Limitations include recall bias and response bias of the informal care data, and restrictions of the dataset to exclude some patient groups.

Keywords Informal care · EQ-5D · ICD chapter · Societal perspective

JEL Classification I10

Introduction

The impact of informal care is taken into account in economic evaluation if a societal perspective is used, where the costs and benefits of an intervention include those that fall outside the health system. This can help avoid an inefficient allocation of resources as it takes into account the full societal effects of an intervention rather than just those falling on the health system. In England and Wales, the National Institute for Health and Care Excellence (NICE) reference case for health technology assessment does not recommend the routine inclusion of wider societal costs or benefits. The reference case takes the perspective that benefits should be measured using direct health effects using EQ-5D (although this can include carer health) and costs should only include those incurred by the National Health Service (NHS) and personal social services (PSS) [1]. However, there is an ongoing debate about the relevance of a societal perspective in health technology assessment as a means of capturing the full economic benefits of new medicines. In the United Kingdom, this interest was crystallized in the form of a policy proposal of value-based pricing for the

Electronic supplementary material The online version of this article (doi:10.1007/s10198-015-0718-5) contains supplementary material, which is available to authorized users.

✉ Simon Dixon
s.dixon@sheffield.ac.uk

[1] Health Economics and Decision Science, School of Health and Related Research (ScHARR), University of Sheffield, Regent Court, 30 Regent Street, Sheffield S1 4DA, UK

Published online: 15 August 2015

pricing of new medicines that recognized the wider societal benefits and costs of interventions [2]. Notably, while there has been a tendency to equate wider societal benefits and costs with production losses, the value-based pricing proposals identified all non-health effects as potentially relevant, including informal care time. While the future of these proposals is uncertain, it is clear that wider societal benefits (and costs) are very much part of the health technology assessment research agenda and have the potential to change research conclusions markedly [3]. However, in order to take into account the costs of informal care for an intervention, informal care needs to be first measured and then valued using the appropriate costings for the type of care provided.

Informal care is care provided by friends, family, and volunteers to individuals who would have difficulty managing without this care. The informal care economy is substantial, with almost 6 million carers in the UK alone recorded in the 2001 census [4]. However, informal care is difficult to measure and value. In terms of measurement, the issue of joint production is important to recognize and correct for in order to avoid inflated estimates [5]. For valuation, it is important to recognize that informal care is a non-market or quasi-market of services supplied by carers who are often unpaid or who receive only nominal payments or state benefits that do not reflect the true cost or benefits of the care [6].

Typically informal care is measured using time spent caring, but it is difficult to distinguish informal care activities from usual household activities that would be undertaken regardless. A range of methods have been used to record time spent providing care, such as a time diary, recall [5, 7], experience sampling and direct observation, each with their own advantages and disadvantages. Time spent providing informal care can then be given a monetary value using a range of techniques: revealed preference using the opportunity cost, proxy good or wellbeing valuation method; or stated preference using discrete choice experiments or contingent valuation [7–9]. The monetary values that are produced by these methods can then be incorporated alongside other costs of an intervention.

The health effects of informal care that capture the effect on the carer of providing the care can also be valued using non-monetary valuation, for example by measuring their health-related quality of life using a generic preference-based measure such as EQ-5D [10] or by measuring their care-related quality of life [11] or carer experience [12]. These utility or quality of life values can then be included alongside other benefits of an intervention. Including the utility of the carer alongside the utility of the patient is straightforward if it is assumed that both should be given an equal weighting, but combining patient utility valued using EQ-5D and carer quality of life using a different measure is less straightforward.

The incorporation of costs of informal care into the overall cost of an intervention is more straightforward as they are all measured using an identical metric. However, it is important to ensure no double-counting, as the same impact from informal care cannot be measured using both costs and benefits. Willingness to accept valuations of carer time, for example, are likely to include an element of compensation for the reduction in health-related quality of life that is anticipated when caring duties are undertaken.

If informal care is to be included as a cost in economic evaluation, then the differences in informal care costs across interventions are required. However, many studies that are used to inform economic evaluation have not collected data that includes the direct measurement of informal care time across interventions. One solution is to use data that is typically collected on the health of the patient to predict the informal care required for each intervention. Given the use of EQ-5D to capture the direct health effects in economic evaluations submitted to NICE, there is interest in how informal care changes as the EQ-5D utility score of the patient changes. One exploratory study used OLS regression to estimate days of informal care used by EQ-5D utility score [13]. However, there were limitations with this study as it did not model the difference by health condition and did not use models that are most appropriate for the data.

This paper estimates informal care usage associated with the health of the patient, building upon the approach taken by Dixon et al. (2006) [13]. The analysis is based around the linkage between informal carer time and EQ-5D and International classification of diseases (ICD) chapter. Regressions are estimated that can be used to predict informal care using the EQ-5D and ICD chapter of the patient. This would enable prediction of the impact of health technologies on informal carer time via their impact on patient health. As such, the results could be applied to any cost-effectiveness analysis that uses the EQ-5D to produce its quality adjusted life years (QALYs). The informal carer time can then be given a monetary value using existing methods in the literature. These results can be applied to existing datasets to inform economic evaluation.

Methods

Data

The Health Outcomes Data Repository, HODaR, is a dataset generated from a prospective survey of inpatients and outpatients at Cardiff and Vale NHS Hospitals Trust, which is a large University hospital in South Wales, UK [14]. The survey is linked to existing routine hospital health data to provide a dataset with socio-demographic, health-related quality of life and ICD classification data. The survey includes subjects aged 18 years or older and excludes individuals who are known to have died. The survey also excludes people with a primary diagnosis on admission of a psychological illness or learning disability. This study focuses on the inpatient sample which has 96,282 eligible observations across 66,113 individuals discharged from hospital from April 2002 to January 2009, and of these there are 59,512 complete responses across 44,494 individuals for all variables used in the analysis and this is the sample used here.

Measures

Informal care

The informal care data is self-reported by the patient and records the number of days a friend or relative has needed to provide care or help with normal activities in the last 6 weeks. Raw observations had a maximum of 45 days; here the data has been censored at 42 days, the maximum number of days in 6 weeks.

Health and socio-demographics

Health of the patient is measured using EQ-5D, a preference-based measure with five dimensions (mobility, self-care, usual activities, pain/discomfort, anxiety/depression) each with three levels of severity [15]. The preference weights used here were elicited using time trade-off with a sample of 2997 members of the UK general population and modeled to produce a utility value for every health state [10]. ICD chapter dummy variables have a value of 1 where the ICD chapter is recorded in the first ICD classification of the patient during their hospital visit, see Table 1 for a description of each of the ICD chapters. An operation dummy variable has a value of 1 if the patient had an operation during their hospital visit, and a comorbidities dummy variable has a value of 1 if multiple ICD chapters were recorded during the hospital visit. Age of the patient is measured at the time of the survey.

Analysis of data

Descriptive analysis

Descriptive statistics of the sample are generated. The distribution of EQ-5D and days the patient needed informal care are examined.

Modeling the data for the econometric models

Regression analysis is used to estimate the relationship between the health of the patient, measured using EQ-5D and ICD chapter, and days of informal care required in the last 6 weeks. Individual patient-level data is used to make best use of the variability of responses in the available dataset.

Several properties of the data need to be considered before identifying the most appropriate method for modeling days of informal care in the last 6 weeks. The variable is a count data variable, the distribution is skewed with a very large spike at zero days and another spike at 42 days, and has some repeated observations per patient. The large spike in responses at zero days is usual in this type of data and is a reflection of the large number of people who do not require informal care when in less than full health. The spike at 42 days is, in part, a consequence of a decision to censor the distribution of observations, which has a maximum of 45 days. However, there was already a peak at 42 days prior to censoring the data at 42 days. There are also smaller spikes at multiples of 5 and 7 days usually due to individuals rounding off responses. To reduce the impact of these smaller peaks in the data on the analysis, models are often estimated in the literature grouping the dependent variable. In the analysis reported here these smaller spikes are ignored since grouping the dependent variable would not allow estimation of all parameters of interest. Although some respondents have multiple observations the majority do not. There are only 10,330 out of 44,494 respondents with multiple responses and these respondents have a median of two repeated observations.

Although linear regression models based on continuous variables are often applied to count data variables, this may result in biased, inconsistent, and inefficient estimates. A linear regression model ignores that the dependent variable is limited to zero and positive integers and it will only provide appropriate estimates in datasets with a high mean of the counts. Even then, it will only be able to provide an estimate of the average effect of a covariate on the conditional mean of the process [16–18]. Alternative models that are appropriate for this type of data are estimated and compared before choosing the preferred model. For completeness some random effects models were estimated to address the issue that some respondents have multiple observations but given the small number of repeated observations robust standard errors can also be used to take into account clustering. The estimated models are Poisson; two-part model (using probit and truncated negative binomial); zero-inflated negative binomial and are described below (the random effects negative binomial model was also considered but was abandoned due to convergence difficulties). For the preferred model the exclusion of variables is explored for insignificant variables and

Table 1 Summary of sample characteristics

N		59,512
Mean EQ-5D (SD)		0.652 (0.321)
Mean age (SD)		59.00 (17.34)
Female		50.8 %
Comorbidity		73.7 %
Operation		73.7 %
Primary ICD chapter		
A	Certain infectious and parasitic diseases	199 (0.3 %)
B	Certain infectious and parasitic diseases	161 (0.3 %)
C	Neoplasms	4095 (6.9 %)
D	Neoplasms and diseases of the blood and blood-forming organs and certain disorders involving the immune system	2840 (4.8 %)
E	Endocrine, nutritional, and metabolic diseases	1043 (1.8 %)
F	Mental and behavioral disorders	100 (0.2 %)
G	Diseases of the nervous system	1530 (2.6 %)
H	Diseases of the eye and adnexa and diseases of the ear and mastoid process	3479 (5.8 %)
I	Diseases of the circulatory system	9284 (15.6 %)
J	Diseases of the respiratory system	2643 (4.4 %)
K	Diseases of the digestive system	7771 (13.1 %)
L	Diseases of the skin and subcutaneous tissue	1671 (2.8 %)
M	Diseases of the musculoskeletal system and connective tissue	6837 (11.5 %)
N	Diseases of the genitourinary system	3781 (6.4 %)
O	Pregnancy, childbirth, and the puerperium	1005 (1.7 %)
Q	Congenital malformations, deformations, and chromosomal abnormalities	148 (0.2 %)
R	Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere specified	6494 (10.9 %)
S	Injury, poisoning, and certain other consequences of external causes	2120 (3.6 %)
T	Injury, poisoning, and certain other consequences of external causes	1420 (2.4 %)
Z	Factors influencing health status and contact with health services	2891 (4.9 %)

variables that may not always be available to inform economic evaluation, e.g., the presence of comorbidities. Model fit was assessed using the distribution of the predictions of days in comparison to the observed distribution. Akaike information criterion (AIC) and Bayesian information criterion (BIC) were used to inform the selection of the model specification.

Models

This section describes the main models used in the analyses below. All of the models are capable of providing valid estimates when the dependent variable is characterized by count data. However, as each relies on different assumptions, the identification of the most appropriate model becomes an empirical issue that is also determined by the nature of the specific dataset used. The model is specified as:

$$y_i = \mathbf{x}_i^{'}\boldsymbol{\beta} + \varepsilon_i \tag{1}$$

where the dependent variable, y_i, represents days needed informal care for patient i; $\mathbf{x}_i$ is a vector of explanatory variables including EQ-5D score, dummies for ICD chapter, comorbidities and operation as well as socio-demographic characteristics of age, age-squared, and female; $\boldsymbol{\beta}$ is the corresponding vector of parameters to be estimated; and ε is the usual error term. Estimation is undertaken using Poisson, two-part model, zero-inflated negative binomial with constant inflation, and zero-inflated negative binomial with variable inflation.

The Poisson regression model assumes that the observed counts are drawn from a Poisson distribution with a mean μ_i given by:

$$\mu_i = E(y_i|\mathbf{x}_i) = \exp\left(\mathbf{x}_i^{'}\boldsymbol{\beta}\right) \tag{2}$$

The Poisson distribution rarely fits well in practice because of the well-known and restrictive equidispersion property, that is, the Poisson distribution variance equals its mean. In most applications, the variance usually exceeds the mean, a feature called overdispersion [16–18]. The negative binomial overcomes this problem by introducing unobserved heterogeneity among observations in the form of an error term, ξ_i, uncorrelated with the independent variables as follows:

$$\tilde{\mu}_i = \exp\left(\mathbf{x}_i^{'}\boldsymbol{\beta} + \xi_i\right) \tag{3}$$

where $\tilde{\mu}_i$ corresponds to Eq. (2) with unobserved heterogeneity included. Assuming that $E[\exp(\xi)] = 1$ the expected conditional mean of the negative binomial regression can be written as:

$$E[\tilde{\mu}_i] = E(y_i|\mathbf{x}_i) = \exp\left(\mathbf{x}_i^{'}\boldsymbol{\beta}\right) = \mu_i \tag{4}$$

It is clear from Eqs. (2) and (4) that both the Poisson and the negative binomial regression share the same mean structure. If the assumptions underlying the negative binomial regression are supported by the data the conditional expected rate of counts will be the same for both models but the standard errors of the Poisson regression model will be biased downwards.

Overdispersion can arise not only from unobserved heterogeneity but also from different processes generating the first and subsequent events, that is, more zeroes in the data than can be generated by the Poisson or even the negative binomial regression models. There are two leading modified count models that can be used to deal with this problem: the two-part model and the zero-inflated model.

The two-part or hurdle model relaxes the assumption that the zeroes and positive counts are generated by the same process. It combines a binary model (a probit or a logit) to predict the zeroes with a zero truncated Poisson or negative binomial regression to predict positive counts. In this model, zero is a hurdle to overcome before any positive counts can be attained and it is often interpreted as a two-stage decision-making process. Using a probit model for the binary outcome and a truncated negative binomial, the two-part model can be defined by the following two equations:

$$\Pr(y_i = 0|\mathbf{z}_i) = \Phi\left(\mathbf{z}_i^{'}\boldsymbol{\gamma}\right) \tag{5a}$$

And

$$\Pr(y_i|y_i > 0, \mathbf{x}_i) = \frac{\Pr(y_i|\mathbf{x}_i)}{1 - \left(1 + \alpha e^{\mathbf{x}_i^{'}\boldsymbol{\beta}}\right)^{-1/\alpha}} \tag{5b}$$

where $\mathbf{z}_i$ is a vector of random variables which determine the probability of a zero in the data, $\boldsymbol{\gamma}$ is the corresponding parameter vector and $\Pr(y_i|\mathbf{x}_i)$ is:

$$\Pr(y_i = m|\mathbf{x}_i) = \frac{\Gamma(y_i + \alpha^{-1})}{y_i!\Gamma(\alpha^{-1})}\left(\frac{\alpha^{-1}}{\alpha^{-1} + e^{\mathbf{x}_i^{'}\boldsymbol{\beta}}}\right)^{\alpha^{-1}}\left(\frac{e^{\mathbf{x}_i^{'}\boldsymbol{\beta}}}{\alpha^{-1} + e^{\mathbf{x}_i^{'}\boldsymbol{\beta}}}\right)^{y_i} \tag{6}$$

where $\Gamma()$ is the gamma function and α is the degree of dispersion. The vector $\mathbf{z}_i$ could in principle be identical to $\mathbf{x}_i$.

The zero-inflated negative binomial model allows the zeroes to be generated by two different processes. It supplements the negative binomial regression with a binary model to increase the zero count. It is often interpreted as heterogeneity in the population arising from two distinct unobserved groups. The outcome for one group (group 1) is zero with probability one. For the other group (group 2) the outcome might be zero but with a positive probability of a nonzero outcome. The zero-inflated negative binomial can be estimated using inflation of only the constant term or inflation of explanatory variables and the constant term. The zero-inflated negative binomial model can be defined by the negative binomial in Eq. (3) together with a logit model for the probability of group membership:

$$\Pr(\text{group } 1|\mathbf{z}_i) = \frac{e^{\mathbf{z}_i^{'}\boldsymbol{\gamma}}}{1 + e^{\mathbf{z}_i^{'}\boldsymbol{\gamma}}} \tag{7}$$

The predicted probabilities of the counts for the negative binomial model can be computed by substituting the estimated parameter values in Eq. (6). The expected counts for the model are found using the formula below:

$$E\left(y_i|\mathbf{x}_i, \mathbf{z}_i\right) = \left(1 - \frac{e^{\mathbf{z}_i^{'}\boldsymbol{\gamma}}}{1 + e^{\mathbf{z}_i^{'}\boldsymbol{\gamma}}}\right)e^{\mathbf{x}_i^{'}\boldsymbol{\beta}} \tag{8}$$

STATA version 11 was used for all regression analysis and IBM SPSS version 19 was used for the descriptive statistical analysis.

Results

Descriptive statistics

Table 1 presents descriptive statistics of the sample. Mean EQ-5D score in the sample is 0.652 (SD = 0.321) and mean age is 59.00 (SD = 17.34). Figure 1 shows that the distribution of EQ-5D score is tri-modal and there are observations across the full range of utilities (1 to −0.594). The number of observations in each ICD chapter varies greatly, with larger proportions of patients in ICD chapter I (diseases of the circulatory system), K (diseases of the digestive system), and M (diseases of the musculoskeletal system and connective tissue), and few observations for A and B (infectious and parasitic diseases), F (mental and behavioral disorders), and Q (congenital malformations, deformations and chromosomal abnormalities). There are no observations for P (conditions originating in the perinatal period), and V, W, X and Y (external causes of morbidity or mortality).

Figure 2 presents the distribution of days the patient needed informal care, showing a huge spike at 0, smaller

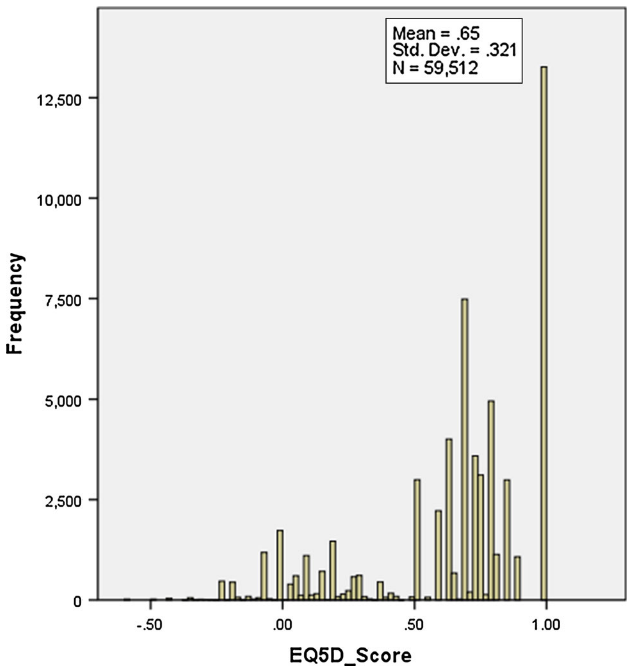

Fig. 1 Distribution of EQ-5D

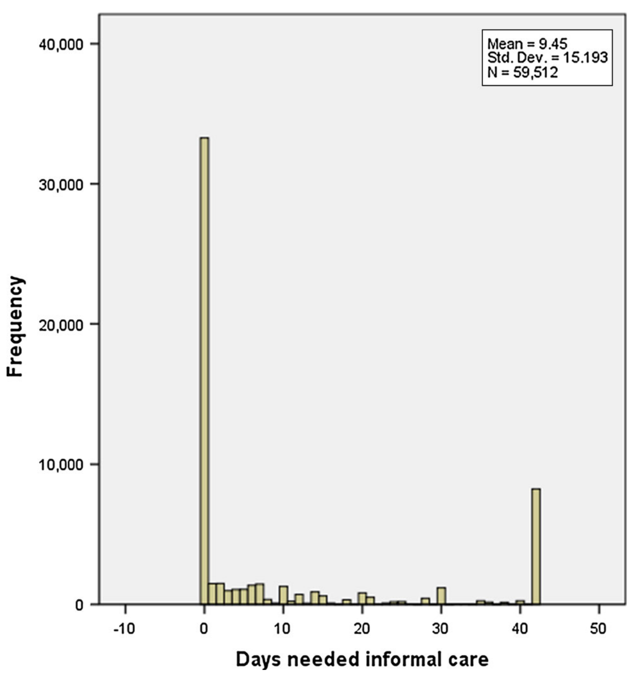

Fig. 2 Distribution of days needed care from a friend or relatives in the last 6 weeks

spikes at multiples of 5 and 7 and a large spike at 42 days (maximum number of days in the last 6 weeks).

Table 2 presents the mean days the patient needed informal care by EQ-5D score. As expected the number of days increase as EQ-5D score decreases and health

Table 2 Summary of days needed informal care in the last 6 weeks by EQ-5D score

EQ-5D score	Mean days (SD)	*N*
1	1.46 (5.39)	13,268
$0.75 \leq$ EQ-5D < 1	3.84 (9.43)	13,128
$0.5 \leq$ EQ-5D < 0.75	10.87 (15.64)	21,515
$0.25 \leq$ EQ-5D < 0.5	19.15 (18.46)	2374
$0 \leq$ EQ-5D < 0.25	19.95 (18.45)	4940
$-0.25 \leq$ EQ-5D < 0	27.65 (17.05)	4106
EQ-5D < -0.25	32.90 (16.54)	181

decreases, meaning that patients in more severe health needed more informal care. Patients with an EQ-5D score of 1 have a mean of 1.46 days (SD = 5.39) where they needed informal care. This may be expected given that EQ-5D measures health today whereas patients were asked to record their informal care need over the previous 6 weeks.

Regression analysis

Table 3 presents regression analysis estimating informal care effects using days the patient needed informal care in the last 6 weeks. The sign and significance of the coefficients for the dummies for ICD chapters varies across the models. Coefficients have a consistent sign and are always significant for ICD chapters H, K, N, O, and R. EQ-5D score is negative and significant in all models as expected, meaning that as EQ-5D score increases the number of days the patient needed informal care decreases. The comorbidities dummy variable is positive and significant in all models, suggesting that patients with comorbidities need more days of informal care. The operation dummy is negative and significant in the Probit model of the two-part model, but is not significant in the other models. The coefficients for age/100 and (age/100)-squared are almost insignificant in the Poisson model but become highly significant when using alternative models. The signs of the coefficients of the age variables as well as the gender are in general consistent across models, with any remaining inconsistency explained by the model misspecifications described later in this Section. Table 4 presents the marginal effects for all models. There are substantial differences in the marginal effects across the different models. For example, changes in the primary ICD chapter H reduce the informal care needed by three and a half days in the Poisson model, the zero-inflated negative binomial with variable inflation predicts the reduction to be only two days. An increase in the EQ-5D score of 0.1, reduces the number of days of care needed by just under two days, less than a day and just over two days for the Poisson, zero-inflated negative binomial with constant inflation and the

Table 3 Regression output for days needed informal care in the last 6 weeks

	(1) Poisson $\hat{\beta}$	(2.1) Two-part model: Probit $\hat{\beta}$	(2.2) Two-part model: Truncated negative binomial $\hat{\beta}$	(3) Zero-inflated negative binomial, constant inflation $\hat{\beta}$	(4) Zero-inflated negative binomial, variable inflation $\hat{\beta}$
Primary ICD chapter					
B	−0.173 (0.319)	0.191 (0.290)	−0.143 (0.340)	−0.165 (0.286)	−0.152 (0.308)
C	0.077 (0.536)	0.006 (0.963)	0.037 (0.689)	0.039 (0.681)	0.040 (0.657)
D	−0.254 (0.046)**	0.313 (0.013)**	−0.079 (0.405)	−0.109 (0.272)	−0.094 (0.323)
E	−0.079 (0.549)	0.250 (0.059)*	−0.013 (0.895)	−0.038 (0.715)	−0.021 (0.834)
F	−0.467 (0.036)**	0.400 (0.054)*	−0.248 (0.196)	−0.278 (0.162)	−0.275 (0.159)
G	0.012 (0.923)	0.063 (0.624)	−0.001 (0.991)	−0.003 (0.979)	0.002 (0.986)
H	−0.373 (0.003)***	0.370 (0.003)***	−0.185 (0.049)**	−0.229 (0.020)**	−0.207 (0.028)**
I	−0.011 (0.932)	0.062 (0.617)	−0.025 (0.783)	−0.024 (0.800)	−0.026 (0.773)
J	−0.103 (0.408)	0.128 (0.314)	−0.105 (0.255)	−0.125 (0.194)	−0.108 (0.240)
K	−0.357 (0.004)***	0.312 (0.012)**	−0.203 (0.026)**	−0.237 (0.013)**	−0.220 (0.016)**
L	−0.267 (0.039)**	0.345 (0.007)***	−0.114 (0.241)	−0.164 (0.109)	−0.128 (0.186)
M	0.022 (0.858)	0.012 (0.922)	−0.003 (0.974)	0.002 (0.982)	−0.001 (0.991)
N	−0.322 (0.010)**	0.225 (0.073)*	−0.226 (0.016)**	−0.245 (0.012)**	−0.236 (0.012)**
O	−0.764 (0.000)***	0.323 (0.014)**	−0.537 (0.000)***	−0.617 (0.000)***	−0.559 (0.000)***
Q	0.087 (0.635)	−0.054 (0.756)	−0.009 (0.952)	0.003 (0.985)	−0.004 (0.978)
R	−0.347 (0.005)***	0.386 (0.002)***	−0.189 (0.039)**	−0.232 (0.015)**	−0.208 (0.023)**
S	0.297 (0.017)**	−0.251 (0.048)**	0.137 (0.135)	0.161 (0.095)*	0.148 (0.106)
T	0.079 (0.527)	0.044 (0.733)	0.069 (0.456)	0.068 (0.486)	0.069 (0.457)
Z	−0.329 (0.010)***	0.445 (0.000)***	−0.092 (0.327)	−0.154 (0.119)	−0.113 (0.228)
EQ-5D score	−1.891 (0.000)***	2.205 (0.000)***	−0.870 (0.000)***	−1.028 (0.000)***	−0.858 (0.000)***
Comorbidity	0.279 (0.000)***	−0.169 (0.000)***	0.152 (0.000)***	0.186 (0.000)***	0.148 (0.000)***
Operation	−0.014 (0.351)	0.074 (0.000)***	0.013 (0.304)	0.004 (0.782)	0.009 (0.504)
Age/100	0.049 (0.846)	2.460 (0.000)***	1.887 (0.000)***	1.697 (0.000)***	1.948 (0.000)***
(Age/100)-squared	0.410 (0.060)*	−2.123 (0.000)***	−1.157 (0.000)***	−0.965 (0.000)***	−1.203 (0.000)***
Female	0.170 (0.000)***	−0.334 (0.000)***	−0.022 (0.070)*	0.011 (0.374)	−0.023 (0.061)*
Constant	2.882 (0.000)***	−1.887 (0.000)***	2.671 (0.000)***	2.737 (0.000)***	2.654 (0.000)***
Inflated variables				$\hat{\gamma}$	$\hat{\gamma}$
Constant				0.125 (0.000)***	−3.343 (0.000)***
EQ-5D score					4.123 (0.000)***
Comorbidity					−0.393 (0.000)***
Age					4.828 (0.000)***
Age-squared					−4.012 (0.000)***
Female					−0.563 (0.000)***
Dispersion (α)			0.796 (0.010)	0.875 (0.014)	0.821 (0.012)
Observations	59,512	59,512	26,240	59,512	59,512

Figures in *parentheses* are *p* values calculated with robust standard errors

* significant at 10 %; ** significant at 5 %; *** significant at 1 %

zero-inflated negative binomial with variable inflation, respectively. These differences are all statistically significant at 5 %.

Table 5 and Fig. 3 present predictions of days needed care for the estimated models. The generated predictions are the mean proportions of the population with a predicted given number of days, estimated using the mean probability across all individuals of the predicted probability for each given number of days. These predictions are presented to enable comparison of how well each model predicts the observed distribution of days needed informal care. The two-part model and the zero-inflated negative binomial

Table 4 Marginal effects—days needed informal care in the last 6 weeks

	(1) Poisson	(2.1) Two-part model: probit	(2.2) Two-part model: truncated negative binomial	(3) Zero-inflated negative binomial, constant inflation	(4) Zero-inflated negative binomial, variable inflation
Primary ICD chapter					
B	−1.636 (1.643)	0.059 (0.056)	−2.995 (3.141)	−1.315 (1.234)	−1.440 (1.413)
C	0.727 (1.174)	0.002 (0.039)	0.767 (1.916)	0.312 (0.759)	0.383 (0.863)
D	−2.403 (1.205)***	0.097 (0.039)***	−1.652 (1.983)	−0.864 (0.787)	−0.885 (0.895)
E	−0.750 (1.252)	0.077 (0.041)	−0.275 (2.089)	−0.304 (0.832)	−0.198 (0.943)
F	−4.416 (2.102)***	0.124 (0.064)	−5.202 (4.025)	−2.211 (1.581)	−2.599 (1.845)
G	0.116 (1.195)	0.019 (0.040)	−0.022 (1.985)	−0.021 (0.786)	0.016 (0.894)
H	−3.522 (1.197)**	0.114 (0.039)**	−3.880 (1.974)***	−1.821 (0.785)***	−1.955 (0.891)***
I	−0.099 (1.162)	0.019 (0.038)	−0.521 (1.894)	−0.190 (0.750)	−0.246 (0.853)
J	−0.977 (1.181)	0.039 (0.039)	−2.210 (1.941)	−0.998 (0.769)	−1.026 (0.874)
K	−3.373 (1.170)**	0.096 (0.038)***	−4.268 (1.915)***	−1.884 (0.760)***	−2.085 (0.863)***
L	−2.522 (1.220)***	0.107 (0.040)**	−2.386 (2.037)	−1.301 (0.812)	−1.216 (0.920)
M	0.207 (1.160)	0.004 (0.038)	−0.063 (1.895)	0.017 (0.751)	−0.010 (0.854)
N	−3.040 (1.188)***	0.069 (0.039)	−4.743 (1.960)***	−1.950 (0.777)***	−2.231 (0.883)***
O	−7.217 (1.393)**	0.100 (0.040)***	−11.277 (2.467)**	−4.913 (0.994)**	−5.293 (1.112)**
Q	0.824 (1.736)	−0.017 (0.053)	−0.191 (3.135)	0.022 (1.215)	−0.039 (1.411)
R	−3.279 (1.170)**	0.119 (0.038)**	−3.963 (1.916)***	−1.845 (0.760)***	−1.965 (0.864)***
S	2.803 (1.179)***	−0.078 (0.039)***	2.882 (1.930)	1.277 (0.764)	1.404 (0.869)
T	0.751 (1.187)	0.014 (0.040)	1.456 (1.952)	0.539 (0.774)	0.654 (0.880)
Z	−3.106 (1.201)**	0.137 (0.039)**	−1.930 (1.968)	−1.224 (0.784)	−1.072 (0.889)
EQ-5D score	−17.866 (0.189)**	0.681 (0.006)**	−18.261 (0.351)**	−8.179 (0.139)**	−20.589 (0.193)**
Comorbidity	2.635 (0.166)**	−0.052 (0.005)**	3.201 (0.317)**	1.476 (0.128)**	2.589 (0.159)**
Operation	−0.135 (0.145)	0.023 (0.005)**	0.283 (0.275)	0.030 (0.109)	0.083 (0.125)
Age	0.459 (2.365)	0.759 (0.069)**	39.607 (4.530)**	13.507 (1.807)**	3.841 (2.300)
Age-squared	3.877 (2.063)	−0.656 (0.062)**	−24.285 (3.841)**	−7.682 (1.527)**	0.744 (1.992)
Female	1.603 (0.137)**	−0.103 (0.004)**	−0.456 (0.252)	0.091 (0.102)	1.487 (0.131)**
Observations	59,512	59,512	26,240	59,512	59,512

Figures in *parentheses* are robust standard errors

* significant at 10 %; ** significant at 5 %; *** significant at 1 %

models are the best at predicting. Overall the zero-inflated negative binomial with variable inflation and the two-part model most accurately predict the spikes at 0 days and 42 days but the two-part model seems to understate the frequency of the number of days between 1 and 7. The Poisson model is poor at estimating the large spike at 0 days, with the predictions instead forming a lower tail from 1 to 14 days, which does not reflect the observed distribution and there is no spike at 42 days. The Poisson model is appropriate when the mean and variance are similar, yet the distribution of days needed informal care suffers from overdispersion, where the variance is much larger than the mean. A test of the Poisson model versus the negative binomial model emphatically rejects the restrictions of the Poisson model at all standard significance levels. The Vuong test of the zero-inflated negative binomial against the standard negative binomial also rejects the standard negative binomial in favor of the zero-inflated versions at all standard significance levels.

The specification of the model with the most accurate predictions, the zero-inflated negative binomial with variable inflation, is explored by excluding the comorbidity and operation variables and insignificant ICD chapters (Table TA1 in the Appendix). The comorbidity and operation variables are excluded, as these may not always be available in datasets used in economic evaluation, although it should be recognized that this lack of data may lead to biased estimates. An additional set of analysis is also undertaken by excluding statistically insignificant ICD chapters to obtain a more parsimonious model, although AIC and BIC indicate that the original specification of the model is preferred (Table TA1 in the Appendix).

Table 5 Predictions of days needed informal care in the last 6 weeks

Days	Observed frequency	Observed percent	Poisson	Two-part model	Zero-inflated negative binomial, constant inflation	Zero-inflated negative binomial, variable inflation
0	33,272	55.91	1.50	55.99	55.55	55.78
1–7	8887	14.94	56.39	10.37	15.64	12.53
8–14	3615	6.07	24.50	9.46	10.14	9.49
15–21	2388	4.01	7.44	6.64	6.32	6.60
22–28	971	1.64	4.63	4.60	4.00	4.54
29–35	1542	2.60	2.77	3.19	2.60	3.14
36–41	592	1.01	1.29	1.96	1.51	1.93
42	8245	13.85	1.47	6.07	4.26	6.00

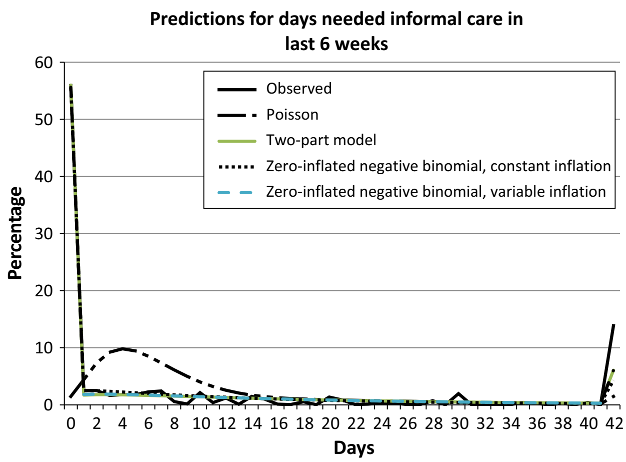

Fig. 3 Predictions of days needed informal care in the last 6 weeks (all models)

Discussion

This paper has presented regression analyses that enable the estimation of informal care associated with the health of the patient. These analyses enable the prediction of days the patient needed informal care in the last 6 weeks using the health and sociodemographic characteristics of the patient. The linkage between EQ-5D and ICD chapter with informal care time enables the measurement of the impact of medicines on carer time via their impact on health. The estimate of days can then be converted into a cost value using available methods in the existing literature and used in economic evaluation of health interventions.

The approach is much simpler than the alternative approaches that have been used in evaluations to date, which have collected care data for the relevant patient-carer population [7]. It also removes the need to collect additional data on informal care to accompany the results of a randomized controlled trial for use in cost-effectiveness models. This method has the advantage that it can be used for all economic evaluations in submissions to agencies such as NICE where EQ-5D scores are available, providing consistency across all evaluations. The method can be used to evaluate both new treatments and displaced treatments. This is advantageous for agencies assessing new interventions where there is a need to consider displacement occurring in a health care system with limited resources. However, it must be recognized that direct measurement of informal care use will be more accurate than the informal care need predicted here, and for this reason the prediction approach presented here should be considered a second-best approach to direct measurement.

The distribution of days the patient needed informal care provided a variety of challenges for the regression analysis, with a very large spike at 0 days, spikes at multiples of 5 and 7 days and a large spike at 42 days. A variety of different regression models were estimated that are appropriate for the type and distribution of the data, and recommendations made on the preferred model. This can be used to inform future studies with similar data type and distribution. The Poisson model was poor at predicting the number of days, and taking into account repeated observations at the individual level using the random effects Poisson model did not improve these predictions. The two-part model was much better at predicting, but still underpredicted the very large spike at 0 days. The zero-inflated negative binomial models were consistently the best models at predicting the distribution of the number of days and were capable of describing the spikes at zero and 42 days. Not surprisingly, none of the models were able to accommodate the multiple spikes at multiples of 5 and 7 days.

The choice of preferred model between the zero-inflated negative binomial model with constant inflation and the zero-inflated negative binomial model with variable inflation needs to take into account their econometric performance. The zero-inflated negative binomial model with variable inflation was best overall at predicting both the spike at 42 days, which represents a large informal care cost as these individuals needed informal care every day in the last 6 weeks, and the very large spike at 0 days, which represents zero informal care. In contrast, the zero-inflated negative binomial model with constant inflation predicted the spike at 0 days but underpredicted the spike at 42 days. However, between around 8 days and 41 days all models overpredicted the number of days. This overprediction is possibly a consequence of the spike at 42 days which none of the models can fully capture, and as a result, this distorts the estimated distribution in the tail. Given that the spike appears to be partly an artefact of imprecise survey responses (as discussed below), this overprediction/distortion may represent a bias in estimated values in this range.

Model choice should also consider the decision-making context within which the predictions will be made. If a model fits well overall, but is poor at describing an important part of the distribution, then this should be taken into consideration. For this population, which represents patients recently undergoing treatment, accurate predictions of the observations at zero and 42 days is very important as these account for almost 70 % of the observations. Having considered both the fit of the models and the decision-making context within which the models will be applied, the zero-inflated negative binomial model with variable inflation was considered to be the most relevant. By way of example, the estimated equation needed to calculate the number of days of informal care based on model (4) and a 60-year-old female patient within ICD Chapter B, with no comorbidities or recent operation is given in the Appendix.

The relationship between EQ-5D and informal care was consistent and significant, where lower EQ-5D score meant higher informal care. Patients with comorbidities also consistently and significantly had higher informal care. The relationship between whether the patient had an operation and informal care was not clear. The sign and significance of many of the dummies for ICD chapter varied across models, but were consistent and significant for 5 ICD chapters. Exclusion of comorbidity and operation variables and insignificant ICD chapters from the full specification produced a deterioration in model performance (as indicated using AIC and BIC statistics), but had little impact on days predicted. The sign and significance of age, age-squared and a female dummy variable varied.

The analyses built upon the approach taken by Dixon et al. (2006) [13] using the same dataset but with additional observations (additional years of data) and regression models that are more appropriate for the data. Different model specifications are estimated here as Dixon et al. focused upon the use of dummy variables for the severity levels of each EQ-5D dimension rather than the utility score and did not incorporate ICD chapter or other health variables into their models. Dixon et al. found that using EQ-5D severity levels as independent variables, rather than the EQ-5D score, produced better models. However, such models were not explored here, as their applicability is limited because analysts rarely have access to patient level data. The results of these alternative formulations are consistent, with improvements in the health of the patient meaning reduced informal care.

Several other studies have identified relationships between informal care and patient health-related quality of life. Brouwer and colleagues [19] undertook a simple linear regression of patient health-related quality of life and caregiver employment status on amount of informal care, relating to a sample of one hundred and fifty-three patients with rheumatoid arthritis. Both explanatory variables were negatively related to amount for informal care provided.

Wimo et al. [20] undertook more complex linear regressions to explain the amount and type of formal care for ninety-two patients with dementia. In addition to patient characteristics, carer characteristics and amount of formal care received were also used as explanatory variables, although they were not consistently significant across all types of informal care. Neubauer and colleagues [21] extended this preceding research further by looking at the amount of informal care provided across multiple caregivers. This showed different relationships between patient health-related quality of life and amount of informal care for the primary caregiver and all caregivers.

While these studies have advantages over our study in terms of data quality, and have also extended the analysis to explore other important factors, they lack the breadth of disease coverage that we have been able to provide. Additionally, the ability of other researchers to use the results of these studies is limited by the need for data on all the explanatory variables. Our work is based primarily on sociodemographic variables that are commonly available in all clinical studies.

Krol et al. [22] adopt a different approach to predicting predict productivity loss (in terms of both absenteeism and presenteeism) using patient quality of life. In their study, a sample of the Dutch general population was asked to imagine the impact on their productivity associated with hypothetical EQ-5D states. This has the advantage that the EQ-5D health states were selected using an orthogonal design, meaning that the health states reflected a range of severity and there were a large and almost identical number of observations for each health state, meaning that the data was not dependent upon the incidence of observed health states in the sample. However, the use of stated preference where respondents imagine the impact on their productivity is likely to be less accurate than observational data where respondents report the actual impact on their productivity with their current EQ-5D health state. The advantage of the data used in this study is that there is a direct linkage between EQ-5D and informal care rather than relying upon hypothetical informal care values.

Limitations

One limitation of the analyses undertaken here is that they will not be appropriate for predicting informal care when either the EQ-5D is inappropriate for the patient population, such as for visual impairment in macular degeneration problems where the EQ-5D has been found to perform poorly [23]. The analyses may also be less reliable for the ICD chapters where there are fewer observations as the relationship is based on a smaller sample.

A further limitation of this study is that the dependent variable is generated from a question asking patients to recall how many days a friend or relative has needed to provide care or help them with normal activities in the last six weeks. It is well documented that questions of recall may be prone to bias and potentially systematic error. This may be exacerbated by the 6-week recall period as recalling care or help needed over the last 6 weeks is a difficult task. Potential problems with the data are illustrated by the fact that raw observations had a maximum of 45 days and spikes at multiples of 5 and 7 days, suggesting that respondents were approximating the number of days using weeks or months (for example, 6 weeks may be converted into one and a half months, which is approximately $30 + 15 = 45$ days). Whereas the spikes at multiples of 5 and 7 days were not altered as it may reflect a real propensity for carers to plan their input in terms of whole weeks, responses of greater than 42 days were censored at 42 days. Alternative approaches are possible, for example, excluding or adjusting responses of over 42 days instead of censoring. While the recall period of the informal care question in HODaR may be an issue, the framing of the question in terms of ‘friends and relatives’ overcomes one problem with previous studies which have often focused on the primary caregiver. Neubauer and colleagues [18] have shown that estimating the time of the primary care giver can produce a substantial underestimate of the costs of informal care.

Another issue relates to the underlying patient population within HODaR. The vast majority of patients were recently discharged from hospital, and as such, are not in a stable health state. In addition, the EQ-5D captures the health of the patient at a single point at the end of the 6-week recall period. Consequently, a patient with a low EQ-5D on discharge (with informal care needs) could recover to full health (with no informal care needs) by the time they complete the EQ-5D at 6 weeks. This could lead to a systematic overestimation of the informal care needs for relatively mild health states.

A further limitation is that the dataset excludes people with a primary diagnosis on admission of a psychological illness or learning disability and therefore is not recommended for use in health interventions for these patients. Furthermore, the sample may suffer from response bias as only respondents with complete data on the variables used in the models were analyzed.

The most straightforward and inexpensive method of providing a monetary valuation of the days of informal care is to use existing values in the literature for specified tasks or hours. However, the dependent variable creates challenges for the monetary valuation of the number of days of informal care as the HODaR dataset does not include data on the number of hours in each day that the patient required informal care or the tasks that were undertaken during that time. We therefore recommend that future surveys of this nature should also collect data on the number of hours and tasks undertaken. Without this information the accuracy and precision of the monetary valuation of informal care will be compromised.

Our approach also assumes that health is exogenous to the amount of informal care received by an individual. If health is also determined by informal care our estimates will suffer from simultaneity bias. Unfortunately, the dataset does not provide appropriate instrumental variables with which to undertake two-stage least squares estimation. However, we consider that the amount of informal care would be expected to have a small effect on EQ-5D scores, given the descriptive system that underpins is heavily based around functioning. Impacts on mental health and wellbeing would be more likely, and only the former is included (crudely) within the EQ-5D.

The issues of measuring joint production and identifying care activities as opposed to incidental household activities were identified earlier. The dataset used here does not distinguish between the different types of input and output associated with the quantities of time recorded. However, the primary purpose for the analyses conducted here is for use in cost-effectiveness analyses, which are based on incremental differences between interventions. As such, any bias introduced by these issues should cancel each other out, as long as the nature of the interventions does not fundamentally change the level of joint production or mix of care/household activities across the interventions. Finally, the predictions of the models have not been validated in a separate dataset, and their predictive accuracy has not been tested in the context of incremental differences in informal care usage across different interventions.

This paper has shown that patients with lower EQ-5D scores have higher informal care needs than patients with higher EQ-5D scores. The preferred regression model estimated here can be used to predict changes in informal care associated with changes in the health of the patient measured using EQ-5D. This information can be used to inform economic evaluation of health interventions. The main limitations of this study relate to the underlying dataset, in particular, exclusions of some patient groups from the patient population and the imprecision of measurement. If a dataset can be generated that overcomes these issues, arguably the analysis presented here has shown that a robust relationship between patient EQ-5D utility scores and informal care time can be estimated. Such an approach has many advantages over the direct measurement of informal care in each individual patient population. However, the degree of imprecision introduced by estimation rather than observation is currently unknown and therefore this method is a second-best alternative to direct measurement. It is recommended that future research should specify and generate a more appropriate dataset for re-estimating the relationship between EQ-5D utility scores and amount of informal care. This relationship should then be used in studies that have collected informal care time data, so that the differences between the two methods can be calculated.

Acknowledgments We are grateful to Gavin Roberts and Danny Palnoch from the Department of Health for their comments throughout this project. This is an independent study commissioned and funded by the Policy Research Programme in the Department of Health. The study was undertaken by the Policy Research Unit in Economic Evaluation of Health and Care interventions (EEPRU) funded by the Department of Health Policy Research Programme. The views expressed are not necessarily those of the Department.

References

1. National Institute of Health and Care Excellence (NICE): Guide to the methods of technology appraisal 2013. NICE, London (2013)
2. Department of Health. Value based pricing: impact assessment (2010). http://www.dhsspsni.gov.uk/value-based_pricing_impact_assessment.pdf. Accessed 12 Aug 2015
3. 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 **33**, 123–135 (2015)
4. Buckner, L., Yeandle, S.: Valuing carers: calculating the value of carers support. Carers UK, UK (2011)
5. van den Berg, B., Spauwen, P.: Measurement of informal care: an empirical study into the valid measurement of time spent on informal caregiving. Health Econ. **15**, 447–460 (2006)
6. Weatherly, H., Faria, R., van den Berg, B.: Valuing informal care for economic evaluation. In: Culyer, A. (ed.) Encyclopedia of Health Economics. Elsevier, UK (2014)
7. Goodrich, K., Kaambwa, B., Al-Janabi, H.: The inclusion of informal care in applied economic evaluation: a review. Value Health **15**, 975–981 (2012)
8. Koopmanschap, M.A., van Exel, J.N., van den Berg, B., Brouwer, W.: An overview of methods and applications to value informal care in economic evaluations of healthcare. PharmacoEconomics **26**, 269–280 (2008)
9. van den Berg, B., Brouwer, W., Koopmanschap, M.A.: Economic valuation of informal care: an overview of methods and applications. Eur J Health Econ **5**, 36–45 (2004)
10. Dolan, P.: Modeling valuations for EuroQol health states. Med. Care **35**, 1095–1108 (1997)
11. Brouwer, W., van Excel, N., can Gorp, B., Redekop, W.: The CarerQol instrument: a new instrument to measure care-related quality of life of informal caregivers for use in economic evaluations. Qual. Life Res. **15**, 1005–1021 (2006)
12. Al-Janabi, H., Flynn, T., Coast, J.: Estimation of a preference-based carer experience scale. Med. Decis. Making **31**, 458–468 (2011)
13. Dixon, S., Walker, M., Salek, S.: Incorporating carer effects into economic evaluation. PharmacoEconomics **24**, 43–53 (2006)
14. Currie, C.J.M., McEwan, P., Peters, J., Patel, T., Dixon, S.: The routine collation of health outcomes data from hospital treated subjects in the Health Outcomes Data Repository (HODaR): descriptive analysis from the first 20,000 subjects. Value Health **8**, 581–590 (2005)
15. Brooks, R.: EuroQol: the current state of play. Health Policy **37**, 53–72 (1996)
16. Cameron, A.C., Trivedi, P.K.: Regression Analysis for Count Data. Econometric Society Monograph, vol. 30. Cambridge University Press, Cambridge (1998)
17. Long, J.S., Freese, J.: Regression models for Categorical Dependent Variables Using Stata, 2nd edn. Stata Press, StataCorp LP, Texas (2006)
18. Wooldridge, J.M.: Econometric Analysis of Cross Section and Panel Data, 2nd edn. MIT Press, Cambridge (2010)
19. Brouwer, W.B., van Exel, N.J., van den Berg, B., Dinant, H.J., Koopmanschap, M.A., van den Bos, G.A.: Burden of caregiving: evidence of objective burden, subjective burden, and quality of life impacts on informal caregivers of patients with rheumatoid arthritis. Arthritis Rheum. **51**, 570–577 (2004)
20. Wimo, A., von Strauss, E., Nordberg, G., Sassi, F., Johansson, L.: Time spent on informal and formal care giving for persons with dementia in Sweden. Health Policy **61**, 255–268 (2014)
21. Neubauer, S., Holle, R., Menn, P., Grossfeld-Schmitz, M., Graesel, E.: Measurement of informal care time in a study of patients with dementia. Int. Psychogeriatr. **20**, 1160–1176 (2008)
22. Krol, M., Stolk, E., Brouwer, W.: Predicting productivity based on EQ-5D: an explorative study. Eur. J. Health Econ. **14**, 465–475 (2014)
23. Espallargues, M., Czoski-Murray, C.J., Bansback, N.J., Carlton, J., Lewis, G.M., Hughes, L.A., Brand, C.S., Brazier, J.E.: The impact of age-related macular degeneration on health status utility values. Invest. Ophthalmol. Vis. Sci. **46**, 4016–4023 (2005)