EEPRU

Policy Research Unit in Economic Evaluation of Health and Social Care Interventions

Research Report

Title

Examining productivity losses associated with health related quality of life using patient data

Authors

Clara Mukuria, Donna Rowen, Mónica Hernández

Simon Dixon

Correspondence to:

Clara Mukuria, Health Economics and Decision Science, University of Sheffield, Regent Court, 30 Regent Street, Sheffield, S1 4DA, UK. Email: c.mukuria@sheffield.ac.uk

Number: 019

Date: June 2014

The University Of Sheffield.

THE UNIVERSITY of York

The Policy Research Unit in Economic Evaluation of Health and Care interventions is funded by the Department of Health Policy Research Programme. It is a collaboration between researchers from then University of Sheffield and the University of York.

The Department of Health's Policy Research Unit in Economic Evaluation of Health and Care Interventions is a 5 year programme of work that started in January 2011. The unit is led by Professor John Brazier (Director, University of Sheffield) and Professor Mark Sculpher (Deputy Director, University of York) with the aim of assisting policy makers in the Department of Health to improve the allocation of resources in health and social care.

This is an independent report commissioned and funded by the Policy Research Programme in the Department of Health. The views expressed are not necessarily those of the Department. This report should be referenced as follows:

Mukuria C, Rowen D, Hernandez M, Dixon S. EEPRU research report 0002. Examining productivity losses associated with health related quality of life using patient data 2014. Available from www.eepru.org.uk

EXECUTIVE SUMMARY

This report presents the findings of a study that estimates productivity loss due to poor health, using days off paid employment and normal activities, EQ-5D and International Classification of Disease (ICD) chapter, as well as socio-demographic data. The results can be used to predict the level of productivity loss associated with EQ-5D values and specific disease diagnosis in economic evaluation.

Methods

Data was used from a prospective survey of inpatients discharged from a hospital in Wales, United Kingdom from April 2002 to January 2009. The number of days off paid employment (N=51,326) and the number of days off normal activities (N=55,726) in the last six weeks were estimated separately using the health of the patient measured by EQ-5D, ICD chapter and other health and socio-demographic data. A variety of different regression models were estimated that are appropriate for the distribution of the number of days off paid employment/normal activities dependent variable, which has large spikes at 0 (zero days off paid employment/normal activities) and 42 days (maximum number of days off paid employment/normal activities in the recall period of 6 weeks). Models were compared based on their predictive ability.

Results

The preferred model that most accurately reflected the distribution of the data was the zero-inflated negative binomial with variable inflation which performed well at approximating the spikes at 0 and 42 days. The results indicated that better health in terms of EQ-5D scores was associated with lower productivity loss in both paid employment and normal activities that did not include employment. Comorbidities were associated with greater productivity loss. ICD chapter codes had varying associations with days off paid employment and normal activities, depending on the model and dependent variable. Increasing age was associated with greater productivity loss but at a decreasing rate, while being female was associated with lower productivity loss.

Discussion

The preferred model can be used to predict changes in productivity 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 productivity data, and restrictions of the dataset to exclude some patient groups.

TABLE OF CONTENTS

List of Tables

List of Figures

1 INTRODUCTION

Economic evaluation combines information on costs and benefits to inform priority setting in health care and to inform decisions on the reimbursement of health care interventions. Benefits are typically assessed using quality adjusted life years (QALYs). QALYs combine length of life and quality of life into a single value and can therefore be used to assess interventions that affect both morbidity and mortality. Costs are typically the direct costs of providing health care but can also include indirect costs which occur as a result of having poor health.

A key question in economic evaluation is what costs and outcomes to include. Those that are directly related to health are relatively straightforward but other indirect outcomes may also be included depending on the perspective of the economic evaluation. If the focus is on the health provider, then costs and outcomes related to the patients' health have been argued to be appropriate. Where a societal perspective is taken, then outcomes beyond the patients' health and direct costs of health care should be taken into account. This includes indirect costs associated with health such as productivity losses for the patient and informal carers. Productivity refers to the economic output associated with both paid and unpaid work. The health of patients has impact upon their ability to work, often meaning that they either need to take time off work or return to work at reduced capacity which results in productivity losses. The former is referred to as absenteeism while the latter is referred to as presenteeism in the literature. [Zhang, et al. 2011] Productivity losses and the associated productivity costs can be substantial. In the UK, days off paid employment due to ill-health are estimated to be 140 million working days per year with estimated costs to employers of £9 billion, while the government spends an estimated £13 billion on health-related welfare payments. [Black and Frost, 2011]

The inclusion of productivity costs can have a substantial impact on the results of an economic evaluation. [Johannesson, et al. 2009] Some reimbursement agencies require the perspective of the health provider which may not include productivity costs. For example, National Institute for Health and Care Excellence (NICE, 2013) guidance requires health technology assessment to take a National Health Service (NHS) and Personal Social Services (PSS) perspective when assessing new technologies. Wider societal benefits such as productivity, usually from absenteeism, may be included separately, for example, as part of the sensitivity analysis, but they are usually not combined with cost utility outcomes. Although the exclusion of productivity losses has been justified on the basis of the perspective of the economic evaluation, there has been a long-standing debate on whether and how they should be included even where the perspective is a societal one. [see Zhang, et al. 2011 and Krol, et al. 2013a for reviews] Arguments against their inclusion have included ethical concerns that their inclusion would bias against those who are not working; potential problems with double-counting if respondents valuing health states take into account income losses; and problems associated with the measurement, valuation and inclusion of paid and unpaid work in economic evaluation. The focus of this report is on the measurement of productivity rather than the debate on whether or not to include these measures in economic evaluation.

Cost of illness studies also include productivity information. Such studies attempt to value all impacts of a disease in monetary terms, including morbidity and mortality effects. [Akobundu, et al. 2006] Cost of illness studies tend to use multiple sources to generate estimates, with productivity losses due to morbidity being produced from routine labour surveys, ad hoc surveys and published estimates from unrelated studies. Mortality effects are generated using life tables to identify the life expectancy of people, adjusted for any excess mortality associated with the disease under consideration. Cost of illness studies are used to highlight the relative importance of diseases, as measured by the total cost of the disease, and as such, inform health care decision making. However, their relevance to health care decision making has been called into question by some. [Shiell, et al. 1987]

Productivity losses have to be measured and valued in order for them to be taken into account. Productivity loss is typically measured using time-off paid employment, time-off unpaid employment, or time-off usual activities due to ill health. For absenteeism, this can be self-reported by patients or objective data reported by employers. [Zhang, et al. 2011] Questionnaires such as the Health Labour Questionnaire, [van Roijen, et al. 1996] have been developed to ask about time off work over a specific period. The recall period in questionnaires is typically brief, for example the last 7 or 14 days. The same approach can be applied to measuring unpaid work where respondents report the amount of time they have spent away from their unpaid work due to ill health. However, it can be difficult for respondents to separate leisure from unpaid work. There is an alternative approach to measuring time off unpaid work which involves measuring the time spent by others doing the work of the ill person. [Krol, et al. 2013a] This approach has been advocated as a way to minimise the problem of distinguishing between leisure and unpaid work time, but may result in under-estimation of productivity loss where patients stop doing the unpaid work without any replacements. Some productivity loss questionnaires also include questions to measure presenteeism. These aim to estimate the reduced productivity while at work due to poor health. However, different questionnaires have been shown to result in different measures of presenteeism in the same population, [Zhang, et al. 2011] which may reflect the difficulty of assessing presenteeism.

Generally, there are two economic theories that have been used when estimating the value of the impact of health and health care interventions on productivity. First, the human capital method focuses on the impact of health care on time off work either through morbidity or mortality. [see Weisbrod, 1961] The value is measured using the gross wage of the patient for the time spent off work. For patients who die, the present value of their lost income is generated from the time of the health care intervention to retirement, to represent the total productivity cost. However, this faces the criticism that it overestimates the impact by focusing on the potential lost production, rather than the expected loss that would be experienced where firms would be able to replace the worker after some time period known as the 'friction period'. The friction cost approach takes this into account, and for patients who die it values productivity using the gross wage in the friction period only, whilst taking into account the costs of finding and training the replacement worker. [see Koopmanschap and van Ineveld, 1992] Compensatory mechanisms for short term morbidity can also be taken into account in the friction method. This may include higher short term costs such as overtime paid to other employees or temporary workers to do the work. Alternatively, compensatory mechanisms may involve no productivity losses as individuals do the work when they return to work. Both methods have been used in economic evaluations. Various approaches have been used to value unpaid work including using the proxy approach, where paid work that closely matches the unpaid work is identified and the gross wage for this paid work is used to value it. [Sculpher, 2001 and Krol, et al. 2013a] However, this may be an overestimate of the value of unpaid work because an informal worker may not be as efficient as a paid worker. A different approach involves using the opportunity cost where the value of time is based on the net wage of the patient if they are employed as this represents the opportunity cost of working time. [Sculpher, 2001]

A question that has had little attention in the literature is the direct relationship between health related quality of life (HRQoL) and productivity losses, although this is the implicit rationale behind the inclusion of these losses in economic evaluation. Such links would allow the productivity losses associated with particular health states or utility to be predicted. Furthermore, it is of interest to know the impact of health on usual activities that exclude paid work, as this can have an impact on the ability of the patient to undertake care of other family members and may mean that the patient themselves requires formal care or informal care from family or friends. Some limited work has been done in this area. Lamers, et al. [2005] explored the relationship in Dutch patients with lower back pain and found that lower levels of HRQoL were associated with efficiency loss and absenteeism, but did not develop models to predict productivity losses due to data limitations. Krol, et al. [2013b] used a representative Dutch general population sample to collect data on HRQoL (EQ-5D) and productivity in order to develop models to predict productivity. Their results showed good external validity but were not based on actual time off paid employment but rather on hypothetical estimates made by the respondents on whether they would go to work given a particular EQ-5D state.

The aim of the research reported in this study was to explore the relationship between HRQoL and productivity using UK relevant data, which would allow decision makers to link specific utility values to expected productivity losses. The policy context of this work is the development of a new system for pricing pharmaceuticals in the UK, referred to as Value Based Pricing (VBP). [Department of Health, 2010] The proposed new mechanism will assess the cost-effectiveness of medicines by taking into account a wider scope of value, including the severity of disease and wider societal benefits such as reductions in productivity losses. [Department of Health, 2010] Given this context, it is important that the methods can be easily used within the current reimbursement process. Consequently, the regression analyses examine the impact of utility on production losses, rather than the individual responses to EQ-5D questions, as the latter are not always available. Furthermore, due to the nature of the dataset used, the focus of this report was to measure the productivity loss (paid and unpaid work) associated with absenteeism. The next sections present the methods, results and discussion.

2 METHODS

2.1 Data

In order to estimate productivity losses associated with health, information on HRQoL and productivity loss in a population with wide ranging conditions and HRQoL was required. This study used a large UK patient dataset, the Health Outcomes Data Repository (HODaR). HODaR is a dataset of survey and routine hospital data collated by Cardiff Research Consortium. [Currie, et al. 2005] Survey data was collected using a prospective survey of inpatients at Cardiff and Vale NHS Hospitals Trust, which is a large University hospital in South Wales, UK. HODaR surveys were sent 6 weeks after discharge. The survey was linked to existing routine hospital health data to provide a dataset with socio-demographic, HRQoL and diagnosis data. The survey included all subjects aged 18 years or older and excludes individuals who were known to have died. The survey also excluded people with a primary diagnosis on admission of a psychological illness or learning disability. The inpatient sample had 96,282 eligible observations across 66,113 individuals discharged from hospital from April 2002 to January 2009. There were 51,326 (55,726) complete responses across 40,019 (42,442) individuals for EQ-5D, International Classification of Diseases (ICD) classification and days off paid employment (days off normal activities). 8,192 (9,318) respondents had multiple responses for days off paid employment (days off normal activities) and these respondents had a median of 2 repeated observations. Using patient data ensured that a wide range of conditions were covered which in turn ensured that the HRQoL values were more likely to cover the whole range of expected values.

2.2 Measures

2.2.1.1 Productivity loss

Survey respondents were asked two single item questions about how many days they had off from paid employment and from normal activities (excluding paid employment e.g. gardening or housework) respectively in the last 6 weeks. Respondents were also asked their employment status, but this included past employment for those who had retired i.e. there was no retired option and those who had retired were asked to complete the employment question based on past employment. As productivity losses from work only relate to those who are employed, using the whole sample in analyses is likely to overestimate productivity losses associated with poor health. A smaller sample with individuals who were aged below 66 and who did not report being unemployed, in full time education or having never worked was used in the analyses to represent those who were likely to be employed at the time they completed the survey. This resulted in three separate samples: two relating to days off employed work (the full sample and the sample focusing on those who were likely to be employed) and one relating to days off normal activities.

2.2.1.2 Health and socio-demographics

Health was measured using the EQ-5D as well as diagnosis based on the ICD. The EQ-5D is a preference-based HRQoL measure which has five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension has three levels: no problems, some problems and extreme problems; in total the measure defines 243 health states. Preferences for the scoring function were elicited using time trade-off on a random sample of 2,997 members of the adult population in UK, with scores anchored on dead (0) and perfect health (1.0). [Dolan, 1997] The resulting scores ranged from -0.594 to 1, with negative values representing states worse than dead. The EQ-5D is the recommended HRQoL measure for health technology assessment (HTA) in England and is widely used. [NICE, 2013] It was therefore useful to use it as the basis of linking health to productivity losses.

The ICD was used to record clinical diagnosis in the hospital. ICD is the standard method for classifying diseases and other health problems (see Appendix Table A1 for Chapters). The 10th version (ICD-10) [World Health Organisation, 2010] was recorded in hospital data for patients admitted to the hospital. ICD classifications were recorded as primary diagnosis (reason for admission) as well as for secondary diagnosis in the hospital data, and these were linked to survey data in HODaR. Routine data on surgical procedures was linked to survey data. HODaR also included age and gender.

2.3 Analysis

2.3.1.1 Descriptive analysis

Descriptive statistics of sample characteristics across the samples used were generated to estimate productivity losses. The distribution of days off paid employment and days off normal activities was also presented.

2.3.1.2 Regression analysis

Regression analysis was used to estimate the relationship between the HRQoL of the patient, days off paid employment due to illness in the last 6 weeks (days off paid employment), and days off normal activities using individual patient-level data, to make best use of the variability of responses in the available dataset.

2.3.1.3 Independent variables

The main independent variables included EQ-5D utility score and ICD categories. Dummy variables were used to represent the different ICD groups with 'A' (Certain infectious and parasitic diseases) excluded as the reference. A dummy variable was also included for comorbidities which were reported as secondary diagnosis. A dummy variable was created to identify those who had had a surgical procedure (operation) during their most recent hospitalisation, versus those who had not. Gender and age were also added to the models as these can be related to productivity. The best performing model in terms of predictive ability was tested for the exclusion of the comorbidity and operation variables, as these are variables which may not be available in trial data, as well as exclusion of ICD codes that were not statistically significant.

2.3.1.4 Dependent variables

The dependent variables were days off paid employment or days off normal activities, which are count data. In HODaR, the recall period in the survey was 6 weeks, giving a maximum of 42 days although some respondents reported up to 45 days which was recoded to 42. 42 days reflects the 6 week recall period but this would be an overestimate of the expected days off paid employment in a 6 week period, assuming a 5-day working week. A second dependent variable was created from the original days off paid employment variable, where any days above 30 were censored to 30 to reflect a normal 5 day working week, in order to fulfil the policy requirements which assumed that this was the average working time over a 6 week period.

There was a large proportion of respondents who did not report either days off paid employment or days off normal activities. The spike at zero is common in this type of data and is a reflection of the large number of people who do not need to take any days off after discharge. There was another spike at the maximum number of days reflecting individuals who were unable to work or undertake normal activities during the 6 week period (see Figure 1). There were also smaller spikes at multiples of 5 and 7 days usually due to individuals rounding off responses.

2.3.1.5 Models

Linear regression models based on continuous variables are often applied to discrete count data although this may result in biased, inconsistent and inefficient estimates. In addition to being count outcomes, the spikes at 0 and 42 days, as well as the smaller spikes, needed to be taken into account when selecting the appropriate model. To lessen the impact of the smaller peaks in the data on the analysis, models are often estimated in the literature grouping the dependent variable. In the present analysis, these smaller spikes were ignored, since grouping the dependent variable would not allow estimation of all parameters of interest.

Alternative models were estimated and compared before choosing the preferred model. They included Tobit; Poisson; random effects (RE) Poisson; negative binomial; two-part model (using probit and truncated negative binomial); zero-inflated negative binomial[1]. Similar methods were also described and used in Rowen, et al. [2013] to estimate the informal care needs of patients using the HRQoL of the patient.

A Tobit model allows for the spikes at both ends of the distribution to be taken into account but not the other properties of count data. The model is specified as:

$$y_i = \mathbf{x}_i'\boldsymbol{\beta} + \varepsilon_i \qquad (1a)$$

$$y = \begin{cases} 0 & \text{if } y^* \leq 0 \\ y^* & \text{if } 0 < y^* < 42 \\ 42 & \text{if } y^* \geq 42 \end{cases} \qquad (1b)$$

Where the dependent variable, y_i, represents days off paid employment, or days off normal activities for individual i; $\mathbf{x}_i$ is a vector of explanatory variables; $\boldsymbol{\beta}$ is the corresponding vector of parameters to be estimated; and ε is the usual independent and identically normally distributed zero mean error term.

The Poisson and negative binomial regression models are related count data models. The Poisson regression model assumes that the observed counts are drawn from a Poisson distribution with a mean μ_i given by:

$$\mu_i = E(y_i \mid \mathbf{x}_i) = \exp(\mathbf{x}_i'\boldsymbol{\beta}) \qquad (2)$$

The Poisson distribution has a restrictive equidispersion property, i.e. the Poisson distribution variance equals its mean. However, in practice the variance usually exceeds the mean, a feature called over dispersion. 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:

$$\eta_i = \exp(\mathbf{x}_i'\boldsymbol{\beta} + \xi_i) \qquad (3)$$

Assuming that $E[\exp(\xi)] = 1$ the expected conditional mean of the negative binomial regression can be written as:

$$E[\eta_i] = E(y_i \mid \mathbf{x}_i) = \exp(\mathbf{x}_i'\boldsymbol{\beta}) \qquad (4)$$

Overdispersion can arise not only from unobserved heterogeneity but also from different processes generating the first (in this case 0 days off normal activities or days off paid employment) and subsequent

[1]The random effects negative binomial model was also considered but was abandoned due to convergence difficulties.

events (> 0 days). This results in more zeroes in the data than what 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:

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

and

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

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 $\Pr(y_i \mid \mathbf{x}_i)$ is:

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

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 model allows the zeroes to be generated by two different processes. It supplements the Poisson or 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 non-zero 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 equation (3) together with a logit model for the probability of group membership:

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

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

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

The models were tested to identify which fitted the data best, based on the predictive performance of the models. For completeness some random effects models were also estimated to address the issue of repeated observations, but the final model just included robust standard errors to take into account clustering. STATA version 11 was used for all regression analysis and IBM SPSS version 19 was used for the descriptive statistical analysis.

3 RESULTS

3.1 Descriptive statistics

Table 1 presents descriptive statistics of the samples: the full sample for those with complete data on days off paid employment (N=51,326), the sample with complete data on days off paid employment for those who were employed (N=27,124) and the full sample for those with days off normal activities (N=55,726). The mean (SD) EQ-5D score for those with complete data on days off paid employment was 0.685 (0.305), while it was 0.673 (0.311) for those with complete data for days off normal activities. Those who were employed had higher EQ-5D scores (Mean (SD) 0.730 (0.29)). The mean (SD) number of days off paid employment was 5.02 (11.92) in the full sample, which was lower than the reported number of days off paid employment for those who were employed 8.50 (14.50) and lower than the reported number of days off normal activities 9.10 (14.50). The lower mean in the full sample was driven by the number of individuals who reported 0 days off paid employment (76.1%), most likely as they were not employed (Table 1). However, all samples had a large proportion reporting zero number of days off paid employment, with median days off paid employment and normal activities of 0. There was a spike at 42 days as well as at other time points such as multiples of 5 and 7. For example, there is a spike at 30 days. Figure 1 presents the distribution of days off paid employment, days off paid employment for employed patients and days off normal activities where these spikes are evident.

The mean (SD) age was 57.22 (17.3) and 58.15 (17.2) for the full sample of those reporting days off paid employment and days off normal activities, respectively. Those who were employed were younger, as expected based on the cut-offs applied to generate this sample. The proportions of females, patients with comorbidities and patients who had an operation during their hospital visit, was similar across the days off paid employment and days off normal activities samples. However, the employed subsample had a slightly higher proportion of females and patients who had an operation, but a lower proportion of comorbidities. The number of observations in each ICD chapter varied 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); with a few observations for A and B (infectious and parasitic diseases), F (mental and behavioural disorders), and Q (congenital malformations, deformations and chromosomal abnormalities). There were no observations for P (conditions originating in the perinatal period), or V, W, X and Y (external causes of morbidity or mortality).

Table 1: Summary of sample characteristics

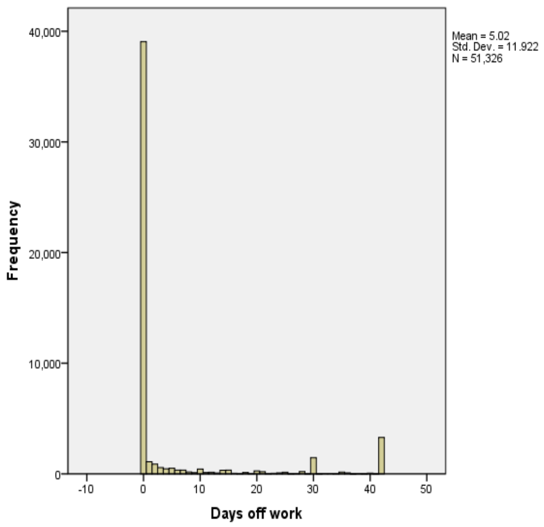
	Days off paid employment (Full sample)	Days off paid employment if employed (Aged < 66 who are likely to be employed)	Days off normal activities (Full sample)
N	51,326	27,124	55,726
Mean EQ-5D (SD)	0.685 (0.305)	0.730 (0.290)	0.673 (0.311)
Mean Days Off (SD)	5.02 (11.92)	8.50 (14.50)	9.10 (14.50)
Mean Days Off Range			
0 days	39,058 (76.1%)	16,154 (59.6%)	31,398 (56.3%)
1-7 days	4,176 (8.1%)	3,694 (13.6%)	7,493 (13.5%)
8-14 days	1,379 (2.7%)	1,239 (4.6%)	3,922 (7.0%)
15-21 days	1001 (2.0%)	914 (3.4%)	2,792 (5.0%)
22-28 days	530 (1.0%)	479 (1.8%)	1,168 (2.1%)
29-35 days	1711 (3.3%)	1,581 (5.8%)	2,191 (3.9%)
36-41 days	182 (0.4%)	163 (0.6%)	670 (1.2%)
42 days	3,289 (6.4%)	2,900 (10.7%)	6,092 (10.9%)
Mean age (SD)	57.22 (17.27)	48.58 (12.13)	58.15 (17.20)
Female	50.1%	52.6%	50.3%
Comorbidity	71.6%	67.7%	72.8%
Operation	74.5%	78.2%	74.5%
ICD category			
A	150 (0.3%)	71 (0.3%)	168 (0.3%)
B	147 (0.3%)	97 (0.4%)	157 (0.3%)
C	3388 (6.6%)	1582 (5.8%)	3832 (6.9%)
D	2574 (5.0%)	1596 (5.9%)	2720 (4.9%)
E	891 (1.7%)	508 (1.9%)	988 (1.8%)
F	81 (0.2%)	50 (0.2%)	86 (0.2%)
G	1336 (2.6%)	803 (3.0%)	1443 (2.6%)
H	2940 (5.7%)	954 (3.5%)	3203 (5.7%)
I	7824 (15.2%)	3474 (12.8%)	8612 (15.5%)
J	2205 (4.3%)	1048 (3.9%)	2362 (4.2%)

	Days off paid employment (Full sample)	Days off paid employment if employed (Aged < 66 who are likely to be employed)	Days off normal activities (Full sample)
K	6851 (13.3%)	3814 (14.1%)	7393 (13.3%)
L	1512 (2.9%)	877 (3.2%)	1564 (2.8%)
M	5634 (11.0%)	3393 (12.5%)	6337 (11.4%)
N	3454 (6.7%)	2127 (7.8%)	3629 (6.5%)
O	990 (1.9%)	767 (2.8%)	1003 (1.8%)
Q	140 (0.3%)	100 (0.4%)	141 (0.3%)
R	5589 (10.9%)	2817 (10.4%)	6024 (10.8%)
S	1882 (3.7%)	1178 (4.3%)	2000 (3.6%)
T	1140 (2.2%)	543 (2.0%)	1284 (2.3%)
Z	2598 (5.1%)	1325 (4.9%)	2780 (5.0%)
Total	51,326	27,124	55,726

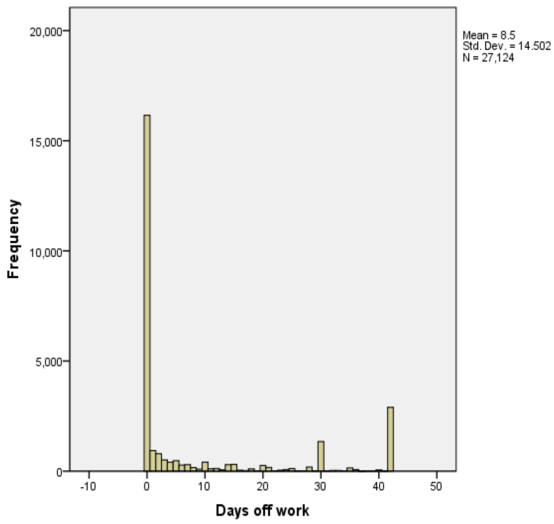
Notes: Each sample includes responses that have non-missing data for EQ-5D, ICD classification.
The sample for days off paid employment if employed has been generated by including respondents aged below 66 who did not self-report that they had never worked, were unemployed or were in full time education. All samples have full range of EQ-5D scores from -0.594 to 1.

Figure 1: Distribution of days off paid employment and days off normal activities excluding paid in the last 6 weeks

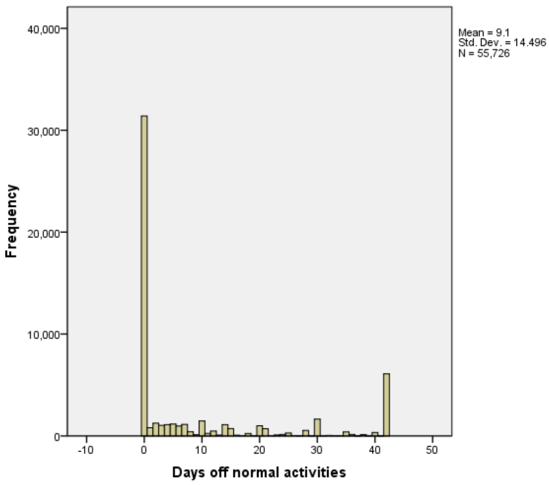
a) Days off paid employment (N=51326)



b) Days off paid employment if employed (N=27124)



c) Days off normal activities (N=55726)



The sample for days off paid employment if employed has been generated by including respondents aged below 66 who did not self-report that they had never worked, were unemployed or were in full time education.

Table 2 shows the number of days off paid employment and normal activities by EQ-5D scores. A negative relationship was expected between days off paid employment/normal activities and EQ-5D scores, i.e. the number of days off paid employment/normal activities were expected to fall as EQ-5D scores increased. In the full sample, there was no clear relationship between mean days off paid employment and EQ-5D score (*rho*=0.0122 *p*=0.006) although those with an EQ-5D score of 1 had the lowest reported mean number of days off paid employment (Table 2). This may have been due to the inclusion of those who were not employed although the relationship was still weak when the focus was on those who were employed (*rho*=-0.1222 *p*<0.001). The relationship between EQ-5D scores and days off normal activities was marginally better with those with low EQ-5D scores reporting higher mean days off normal activities compared to those with high EQ-5D scores (*rho*=-0.3668 *p*<0.001).

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

	Days off paid employment			Days off paid employment if employed			Days off normal activities			
EQ-5D range	**N**	**Mean**	**SD**	**N**	**Mean**	**SD**	**N**	**Mean**	**SD**	**Median**
-0.594 - 0	2884	5.43	13.44	1215	10.11	16.89	3433	21.5	18.25	21
0.001 - 0.099	1604	5.10	12.63	710	9.65	15.96	1884	15.98	17.05	10
0.100 - 0.199	1864	5.61	13.18	764	10.98	16.65	2150	17.71	18.10	10
0.200 - 0.299	1285	7.65	14.80	587	14.15	17.62	1439	17.95	18.28	10
0.300 - 0.499	778	9.36	15.96	362	17.70	18.11	854	19.89	18.88	15
0.500 - 0.599	3791	4.00	11.45	1320	9.85	16.32	4371	14.9	16.97	6
0.600 - 0.649	3314	5.97	13.32	1450	12.03	16.92	3739	11.56	14.99	4
0650 - 0.699	7087	6.423	13.43	3322	12.57	16.68	7753	9.88	14.24	1
0.700 - 0.749	3684	3.57	10.07	1746	6.70	12.98	3914	4.90	10.59	0
0.750 - 0.799	7474	5.54	11.98	4453	8.61	14.1	7892	6.31	11.82	0
0.800 - 0.899	4753	5.92	12.55	2739	9.43	14.76	5043	5.95	11.58	0
1	12,808	3.36	9.11	8456	4.77	10.54	13,254	2.89	7.70	0

Total	51,326	5.02	11.92	27,124	8.50	14.5	55,726	9.10	14.5	0

NB: Median days for days off paid employment in full sample as well as employed sample were 0. Minimum and maximum days off paid employment and days off normal activities were 0 and 42 respectively for all the EQ-5D groups. EQ-5D scores do not have values between 0.88 and 1.

3.2 Regression analysis

3.2.1.1 Days off paid employment- full sample

Table 3 presents predictions of days off paid employment for the estimated models. Overall the zero-inflated negative binomial with variable inflation most accurately predicted the spikes at 0 days and 42 days. The two-part model was the next best model at predicting the spike at 42 days but its predictions for less than 42 days were poor. The Poisson and random effects Poisson models were also poor at estimating the large spike at 0 days and did not capture the spike at 42 days. The distribution of days off paid employment suffers from over-dispersion, meaning the Poisson model was unlikely to be appropriate. The Tobit model also had poor predictions; it partially captured the spike at 0 days but did not predict values greater than 28 days, whereas the observed number of days off was greater than this. Akaike information criteria (AIC) and Bayesian information criteria (BIC) (not reported) both selected the zero-inflated negative binomial with variable inflation over the negative binomial model.

Table 3: Observed and predicted days off paid employment in the last 6 weeks (N=55,726)

Days	Observed		Predicted						
			Tobit	Poisson	RE Poisson	Negative binomial	Two-part model	Zero-inflated negative binomial, constant inflation	Zero-inflated negative binomial, variable inflation
	N	%	%	%	%	%	%	%	%
0	39,058	76.10	20.17	21.60	23.66	74.65	23.87	75.44	76.05
1-7	4,176	8.14	46.35	53.16	51.23	14.93	26.27	12.57	7.63
8-14	1,379	2.68	29.04	18.61	18.29	3.24	16.71	4.06	5.03
15-21	1001	1.95	4.44	5.15	4.92	1.71	10.71	2.27	3.36
22-28	530	1.03	0.00004	1.28	1.35	1.08	6.99	1.43	2.28
29-35	1711	3.34	-	0.19	0.40	0.77	4.65	0.98	1.56
36-41	182	0.36	-	0.01	0.11	0.50	2.77	0.61	0.96
42	3,289	6.41	-	0.00	0.06	3.10	8.06	2.63	3.11
Total	51,326	100	100	100	100	100	100	100	100

Regression results for the best fitting model: the zero-inflated negative binomial with variable inflation, and results for models which include truncating days above 42 to 30 and the exclusion of some of the variables, are reported in Table 4 (results for the other models are reported in Appendix Table A2). The EQ-5D score coefficient was negative and significant in all models, indicating as expected that better health (higher EQ-5D scores) was associated with less days off paid employment. As was expected, age was positive and age-squared was negative, and both were significant for all models. The female variable was significant and negative suggesting that women take fewer days off paid employment. The comorbidity coefficient was positive and statistically significant indicating, as expected, that having comorbidities was associated with more days off paid employment. The operation variable was negative but it was not statistically significant. Most of the ICD chapter variables had positive coefficients apart from Chapters L (diseases of the skin), N (diseases of the genitourinary system) and R (symptoms/signs from laboratories) which were negative. Chapters C (neoplasms), F (mental and behavioural diseases), I (diseases of circulatory system), M (diseases of musculoskeletal system), Q (congenital malformations), S (injuries) and T (burns/poisoning and other external causes) were statistically significant, indicating that compared to Chapter A (certain infectious and parasitic diseases), patients with these diagnoses had more days off paid employment. There were slight differences in the coefficient sizes of all the variables in the 30 day model. Exclusion of comorbidity, operation and ICD variables that were not statistically significant did not have a large impact on coefficients. BIC indicate that the 42 day model was preferred to the 30 day model while the models excluding comorbidity, operation and insignificant ICD codes were preferred to the ones with these variables.

Table 4: Regression output for days off paid employment in the last 6 weeks – Zero inflated negative binomial variable inflation

	42 days comorbidity and operation	30 days comorbidity and operation	30 days excluding comorbidity and operation	30 days excluding comorbidity and operation significant ICD codes only
	$\hat{\beta}$	$\hat{\beta}$	$\hat{\beta}$	$\hat{\beta}$
Primary ICD chapter				
B	0.001	0.087	0.037	-
	(0.998)	(0.701)	(0.872)	
C	0.495**	0.456***	0.479***	0.467***
	(0.012)	(0.008)	(0.007)	(0.000)
D	0.092	0.092	0.112	-
	(0.645)	(0.598)	(0.533)	
E	0.132	0.133	0.078	-
	(0.527)	(0.468)	(0.680)	
F	0.487*	0.393	0.373	-
	(0.074)	(0.104)	(0.132)	
G	0.252	0.241	0.185	-
	(0.208)	(0.170)	(0.308)	
H	0.107	0.123	0.060	-
	(0.597)	(0.489)	(0.744)	
I	0.406**	0.370**	0.363**	0.352***
	(0.038)	(0.031)	(0.041)	(0.000)
J	0.042	0.049	0.016	-
	(0.833)	(0.779)	(0.930)	
K	0.070	0.077	0.014	-
	(0.721)	(0.657)	(0.938)	
L	-0.051	-0.049	-0.120	-
	(0.805)	(0.786)	(0.521)	
M	0.363*	0.336*	0.289	0.277***
	(0.064)	(0.050)	(0.104)	(0.000)

	42 days comorbidity and operation	30 days comorbidity and operation	30 days excluding comorbidity and operation	30 days excluding comorbidity and operation significant ICD codes only
N	-0.001	-0.002	-0.038	-
	(0.996)	(0.990)	(0.833)	
O	0.119	0.082	0.037	-
	(0.565)	(0.653)	(0.844)	
Q	0.375*	0.355*	0.333*	0.320***
	(0.088)	(0.067)	(0.094)	(0.000)
R	-0.056	-0.051	-0.082	-
	(0.777)	(0.770)	(0.646)	
S	0.387**	0.366**	0.386**	0.375***
	(0.049)	(0.034)	(0.031)	(0.000)
T	0.338*	0.319*	0.323*	0.311***
	(0.090)	(0.068)	(0.074)	(0.000)
Z	0.090	0.071	0.025	-
	(0.655)	(0.689)	(0.892)	
EQ-5D score	-0.733***	-0.579***	-0.602***	-0.596***
	(0.000)	(0.000)	(0.000)	(0.000)
Age	0.055***	0.050***	0.051***	0.052***
	(0.000)	(0.000)	(0.000)	(0.000)
Age-squared	-0.001***	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Female	-0.063***	-0.054***	-0.061***	-0.055***
	(0.001)	(0.001)	(0.000)	(0.000)
Comorbidity	0.219***	0.189***	-	-
	(0.000)	(0.000)		
Operation	-0.008	-0.001	-	-
	(0.710)	(0.961)		
Constant	1.891***	1.756***		
	(0.000)	(0.000)		
Inflated variables	$\hat{\gamma}$	$\hat{\gamma}$	$\hat{\gamma}$	$\hat{\gamma}$
Constant	4.367***	4.383***	4.222***	4.220***
	(0.000)	(0.000)	(0.000)	(0.000)
EQ-5D score	-0.078**	-0.057	-0.030	-0.030
	(0.043)	(0.180)	(0.474)	(0.476)
Age	-0.226***	-0.226***	-0.226***	-0.226***
	(0.000)	(0.000)	(0.000)	(0.000)
Age-squared	0.003***	0.003***	0.003***	0.003***
	(0.000)	(0.000)	(0.000)	(0.000)
Female	0.172***	0.172***	0.176***	0.177***
	(0.000)	(0.000)	(0.000)	(0.000)
Comorbidity	-0.239***	-0.242***		
	(0.000)	(0.000)		
Dispersion (α)	0.985	0.774	0.785	0.789
	(0.021)	(0.017)	(0.017)	(0.017)
Observations	51326	51326	51326	51326
Number of patients	40019	40019	40019	40019
Log likelihood	-71321	-69342	-69438	-69461
AIC	2.788	2.703	2.707	2.707
BIC	-413279	-417636	-417478	-417572

Notes: p values in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%

Table 5 presents the predictions of days off paid employment, and there is little variation in days predicted using 42 days compared to 30 days models. Most obviously, censoring forced the model to better predict the presence of a spike of observations at 30 days, which was also present in the uncensored dataset, although it still under-predicted the size of the spike by over 50% (4.02 predicted vs. 10.08 censored observed). Other, less noticeable effects were a slight increase in the overestimation for groups between 8 and 29 days. There was a very small variation in the predictions of days off paid employment when comorbidity, operation and ICD codes were excluded from the censored models.

Table 5: Predictions for days off paid employment in the last 6 weeks using Zero inflated negative binomial variable inflation models

Days	Observed		Predicted	Observed		Predicted		
	Frequency	Percent	variable inflation 42 days comorbidity and operation			variable inflation 30 days comorbidity and operation	variable inflation 30 days excluding comorbidity and operation	variable inflation 30 days significant ICD codes only
	N	%	%	N	%	%	%	%
0	39058	76.1	76.05	39058	76.1	76.07	76.06	76.06
1-7	4176	8.14	7.63	4176	8.14	7.82	7.87	7.87
8-14	1379	2.68	5.03	1379	2.68	5.76	5.75	5.75
15-21	1001	1.95	3.36	1001	1.95	3.71	3.71	3.71
22-29	539	1.05	2.54	539	1.05	2.61	2.58	2.61
30	1459	2.84	0.25	5173	10.08	4.02	3.99	4.00
31-41	425	0.84	2.01					
42	3,289	6.41	3					
Total	51326	100	100	51326	100	100	100	100

3.2.1.2 Days off paid employment for employed patients

Table 6 presents the predictions of days off paid employment for the estimated models for those who were employed. As with the results based on the full sample, the best fitting models were the zero inflated negative binomial with constant and variable inflation models, with the latter performing slightly better. However, the two-part model performed better at predicting the 42 day spike, although it under-predicted the 0 days spike and over-predicted the number of days off paid employment for all the other days. The Tobit and Poisson models were poor predictors.

Table 6: Predictions for days off paid employment in the last 6 weeks, if employed

Days	Observed		Predicted						
	Frequency	Percent	Tobit	Poisson	RE Poisson	Two-part model	Negative binomial	Zero-inflated negative binomial, constant inflation	Zero-inflated negative binomial, variable inflation
	N	%	%	%	%	%	%	%	%
0	16154	59.6	-	0.81	0.99	40.41	56.92	59.52	59.49
1-7	3694	13.5	38.91	50.99	51.37	18.57	22.87	12.69	12.73
8-14	1239	4.5	47.6	34.74	35	12.7	6.27	8.59	8.55
15-21	914	3.4	12.73	10.27	9.49	8.55	3.47	5.75	5.74
22-28	479	1.9	0.76	2.58	2.35	5.8	2.26	3.91	3.89
29-35	1581	5.8	-	0.51	0.62	3.98	1.61	2.68	2.68
36-41	231	0.6	-	0.08	0.15	2.41	1.02	1.64	1.64
42	2900	10.7	-	0.01	0.05	7.56	5.59	5.22	5.29
Total	27,124	100		100	100.00	100	100	100	100

Regression results for the zero-inflated negative binomial with variable inflation model, as well as the 30 day truncation and exclusion of various variables, are reported in Table 7 (regression results for the other models are reported in Appendix Table A3). Results were similar to the full sample with some variations on coefficient size and statistical significance. The EQ-5D score coefficient was negative, age was positive and age squared was negative, comorbidity was positive and all these coefficients were statistically significant across all the models. The female coefficient was negative across all models but was only statistically significant in the 30 day model, excluding comorbidity and operation. The operation coefficient was negative and not statistically significant. ICD Chapter variables were mainly positive apart from Chapters L and R and similar chapters to those in the previous section were statistically significant (Chapters C, F, I, M, Q, S, T). Coefficient sizes as well as significance varied across the 42 day and 30 day models. Exclusion of comorbidity, operation and insignificant ICD codes changed the coefficient size of the remaining variables slightly. BIC indicate that the 42 day model was preferred to the 30 day model and the model excluding comorbidity, operation and insignificant ICD codes was preferred to the model with these variables.

Table 7: Regression output for days off paid employment in the last 6 weeks, if employed - Zero inflated negative binomial variable inflation models

	42 days comorbidity and operation	30 days comorbidity and operation	30 days excluding comorbidity and operation	30 days excluding comorbidity and operation, significant ICD codes only
	$\hat{\beta}$	$\hat{\beta}$	$\hat{\beta}$	$\hat{\beta}$
Primary ICD chapter				
B	0.014	0.067	0.010	
	(0.957)	(0.775)	(0.968)	
C	0.513**	0.455**	0.475**	0.458***
	(0.010)	(0.011)	(0.011)	(0.000)
D	0.142	0.121	0.139	
	(0.484)	(0.503)	(0.463)	
E	0.138	0.119	0.066	
	(0.517)	(0.531)	(0.739)	
F	0.570**	0.458**	0.438*	0.423***
	(0.026)	(0.043)	(0.064)	(0.004)
G	0.296	0.255	0.197	
	(0.145)	(0.159)	(0.300)	
H	0.149	0.146	0.080	
	(0.468)	(0.425)	(0.678)	
I	0.439**	0.383**	0.372**	0.356***
	(0.027)	(0.031)	(0.046)	(0.000)
J	0.067	0.058	0.022	
	(0.742)	(0.752)	(0.907)	
K	0.104	0.092	0.023	
	(0.602)	(0.605)	(0.902)	
L	-0.060	-0.075	-0.151	
	(0.778)	(0.693)	(0.445)	
M	0.397**	0.351**	0.301	0.285***
	(0.046)	(0.048)	(0.106)	(0.000)
N	0.026	0.002	-0.041	
	(0.897)	(0.990)	(0.828)	
O	0.135	0.083	0.036	
	(0.521)	(0.658)	(0.856)	
Q	0.392*	0.343*	0.307	
	(0.084)	(0.089)	(0.141)	
R	-0.050	-0.057	-0.090	
	(0.804)	(0.750)	(0.631)	
S	0.417**	0.376**	0.391**	0.376***
	(0.037)	(0.035)	(0.037)	(0.000)
T	0.367*	0.329*	0.328*	0.312***
	(0.071)	(0.070)	(0.084)	(0.000)
Z	0.106	0.066	0.016	
	(0.606)	(0.717)	(0.932)	
EQ-5D score	-0.732***	-0.585***	-0.608***	-0.599***
	(0.000)	(0.000)	(0.000)	(0.000)
Age	0.031***	0.028***	0.030***	0.032***
	(0.000)	(0.000)	(0.000)	(0.000)
Age-squared	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)

	42 days comorbidity and operation	30 days comorbidity and operation	30 days excluding comorbidity and operation	30 days excluding comorbidity and operation, significant ICD codes only
Female	-0.030	-0.023	-0.030*	-0.026
	(0.121)	(0.181)	(0.075)	(0.114)
Comorbidity	0.223***	0.192***		
	(0.000)	(0.000)		
Operation	-0.015	-0.006		
	(0.507)	(0.761)		
Constant	2.276***	2.139***	2.269***	2.237***
	(0.000)	(0.000)	(0.000)	(0.000)
Inflated variables	$\hat{\gamma}$	$\hat{\gamma}$	$\hat{\gamma}$	$\hat{\gamma}$
Constant	2.435***	2.443***	2.232***	2.228***
	(0.000)	(0.000)	(0.000)	(0.000)
EQ-5D score	0.382***	0.403***	0.434***	0.434***
	(0.000)	(0.000)	(0.000)	(0.000)
Age	-0.164***	-0.163***	-0.162***	-0.162***
	(0.000)	(0.000)	(0.000)	(0.000)
Age-squared	0.002***	0.002***	0.002***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
Female	0.109***	0.108***	0.113***	0.113***
	(0.000)	(0.000)	(0.000)	(0.000)
Comorbidity	-0.289***	-0.291***		
	(0.000)	(0.000)		
Dispersion (α)	0.961	0.757	0.768	0.773
	(0.021)	(0.017)	(0.017)	(0.017)
Observations	27124	27124	27124	27124
Number of patients	22367	22367	22367	22367
Log likelihood	-60848	-58911	-59012	-59041
AIC	4.489	4.346	4.353	4.355
BIC	-154854	-158726	-158556	-158632

Notes: p values in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%

Table 8 presents associated predictions of days off paid employment for those who are employed for the best fitting models. There were no significant differences across the different specifications or when comparing these results to those using 42 days, except from those already noted in the previous section.

Table 8: Predictions for days off paid employment in the last 6 weeks, if employed using Zero inflated negative binomial models

Days	Observed		Predicted	Observed		Predicted		
			variable inflation 42 days comorbidity and operation			variable inflation 30 days comorbidity and operation	variable inflation 30 days	variable inflation 30 days significant ICD codes only
	N	%	%	N	%	%	%	%
0	16154	59.56	59.49	16154	59.56	59.5	59.5	59.5
1-7	3694	13.62	12.73	3694	13.62	13.03	13.14	13.09
8-14	1239	4.56	8.55	1239	4.56	9.78	9.78	9.76
15-21	914	3.37	5.74	914	3.37	6.34	6.34	6.34
22-29	488	1.79	4.34	488	1.79	4.46	4.46	4.46
30	1346	5	0.42	4635	17.09	6.87	6.82	6.83
31-41	389	1.4	3.45					
42	2900	10.7	5.29					
Total	27,124	100	100	27,124	100	100	100	100

3.2.1.3 Days off normal activities

The predicted results for the regressions of days off normal activities are reported in Table 9. Similar to the normal days off paid employment predictions, the Poisson and Tobit models were poor at predicting the spikes at 0 and 42 days; the two-part model did better and the negative binomial models performed the best. The zero-inflated negative binomial model with variable inflation performed best in terms of predicting the spikes at 0 and 42 days. The Tobit, Poisson, RE Poisson and the two part models over-predicted the days off normal activities from 1 to 28 days, and with the exception of the two part model, under-predicted days off normal activities from 29 to 41 days. The negative binomial over-predicted days off normal activities for 1 to 7 days and generally under-predicted for the rest of the distribution; whereas the zero-inflated models under-predicted for 1 to 7, and 29 to 35 days, and over-predicted for the rest of the distribution.

Table 9: Predictions for days off normal activities excluding paid employment in the last 6 weeks

	Observed					Predicted			
Days			Tobit	Poisson	RE Poisson	Negative binomial	Two-part model	Zero-inflated negative binomial, constant inflation	Zero-inflated negative binomial, variable inflation
	N	%	%	%	%	%	%	%	%
0	31398	56.34	-	0.79	0.51	52.16	43.54	56.23	56.32
1 -7	7493	13.46	46.69	52.06	54.4	25.97	16.94	12.59	11.28
8-15	3922	7.04	35.15	31.59	34.18	7.1	14.26	10.51	10.06
16-21	2792	5.01	10.71	8.91	7.98	3.83	9.43	7.11	7.14
22-28	1168	2.08	6.31	3.74	2.31	2.42	5.92	4.63	4.84
29-35	2191	3.91	1.12	1.78	0.51	1.66	3.66	2.98	3.25
36-41	670	1.21	0.02	0.66	0.08	1.06	2	1.7	1.93
42	6092	10.93	-	0.46	0.03	5.79	4.23	4.24	5.19
Total	55,726	100	100	100	100	100	100	100	100

Table 10 presents the regression results for days off normal activities, using the zero-inflated negative binomial with variable inflation model results, models which excluded comorbidity and operation, and non-significant ICD codes. (NB: a 30 day truncation was not assessed as normal activities are not restricted to the working week) The coefficient for the EQ-5D score was negative, age was positive, age squared was negative, female was negative, comorbidity was positive, and all were statistically significant, which mirrored the results for days off paid employment. However, the operation coefficient was positive and statistically significant which indicated that having an operation was associated with more days off normal activities. The ICD chapter variables were also different with 9 of the included chapters having negative coefficients, though most were not statistically significant. Similar to days off paid employment, Chapters C (neoplasms), M (diseases of musculoskeletal system), Q (congenital malformations), S (injuries) and T (burns/poisoning and other external causes) were positive and statistically significant. Additionally, Chapters H (diseases of the eye, ear and mastoid process) and O (pregnancy) were statistically significant but they were negative, which indicated that patients in these groups were less likely to report days off normal activities compared to those in Chapter A (certain infectious and parasitic diseases). The exclusion of the comorbidity and operation variables as well as ICD chapters that were not significant changed the size of some of the coefficients for the remaining variables. Table 11 shows the predicted days off normal activities and there is little difference across the models, with the full model performing slightly better at the 42 day spike.

Table 10: Regression output for days off normal activities excluding paid employment in the last 6 weeks using Zero inflated negative binomial models with variable inflation

	Days off normal activities All variables	Days off normal activities Excluding operation and comorbidity	4 Days off normal activities Statistically significant ICD chapters excluding operation, comorbidity
	$\widehat{\boldsymbol{\beta}}$		
Primary ICD			
B	0.009	0.014	
	(0.950)	(0.925)	
C	0.211**	0.257**	0.298***
	(0.043)	(0.015)	(0.000)
D	0.019	0.054	
	(0.861)	(0.617)	
E	-0.018	-0.041	
	(0.871)	(0.720)	
F	-0.244	-0.223	
	(0.246)	(0.294)	
G	0.068	0.053	
	(0.528)	(0.627)	
H	-0.190*	-0.230**	-0.188***
	(0.080)	(0.035)	(0.000)
I	0.167	0.187*	0.229***
	(0.108)	(0.075)	(0.000)
J	0.056	0.055	
	(0.601)	(0.605)	
K	-0.033	-0.045	
	(0.749)	(0.669)	
L	-0.088	-0.121	
	(0.434)	(0.284)	
M	0.223**	0.216**	0.258***
	(0.031)	(0.039)	(0.000)
N	-0.039	-0.044	
	(0.711)	(0.681)	
O	-0.283**	-0.304**	-0.264***
	(0.020)	(0.013)	(0.000)
Q	0.270*	0.285**	0.327***
	(0.053)	(0.043)	(0.001)
R	-0.125	-0.127	
	(0.232)	(0.229)	
S	0.390***	0.435***	0.477***
	(0.000)	(0.000)	(0.000)
T	0.258**	0.285***	0.327***
	(0.015)	(0.008)	(0.000)
Z	-0.067	-0.068	
	(0.530)	(0.531)	
EQ-5D score	-0.778***	-0.787***	-0.782***
	(0.000)	(0.000)	(0.000)
Age	0.012***	0.013***	0.014***
	(0.000)	(0.000)	(0.000)
Age-squared	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)
Female	-0.090***	-0.092***	-0.090***
	(0.000)	(0.000)	(0.000)
Comorbidity	0.172***		
	(0.000)		
Operation	0.023*		
	(0.083)		

	Days off normal activities All variables	Days off normal activities Excluding operation and comorbidity	4 Days off normal activities Statistically significant ICD chapters excluding operation, comorbidity
Constant	2.824***	2.923***	2.865***
	(0.000)	(0.000)	(0.000)
Inflated variables	$\hat{y}$	$\hat{y}$	$\hat{y}$
Constant	-0.769***	-0.921***	-0.923***
	(0.000)	(0.000)	(0.000)
EQ-5D score	2.085***	2.111***	2.112***
	(0.000)	(0.000)	(0.000)
Age	-0.027***	-0.029***	-0.029***
	(0.000)	(0.000)	(0.000)
Age-squared	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)
Female	-0.183***	-0.175***	-0.175***
	(0.000)	(0.000)	(0.000)
Comorbidity	-0.274***		
	(0.000)		
Dispersion (α)	0.657	0.664	0.666
Observations	55726	55726	55726
No. of patients	42442	42442	42442
Log likelihood	-130826	-130994	-131027
AIC	4.697	4.702	4.703
BIC	-346970	-346668	-346723

Notes: p values in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%

Table 11: Predictions for days off normal activities excluding paid employment in the last 6 weeks using zero-inflated negative binomial with variable inflation models

Days	Observed		Predicted		
			Days off normal activities All variables	Days off normal activities Excluding operation and comorbidity	Days off normal activities Statistically significant ICD chapters excluding operation, comorbidity
	N	%	%	%	%
0	31398	56.34	56.32	56.32	56.32
1 -7	7493	13.46	11.28	11.34	11.32
8-15	3922	7.04	10.06	10.05	10.05
16-21	2792	5.01	7.14	7.14	7.14
22-28	1168	2.08	4.84	4.84	4.84
29-35	2191	3.91	3.25	3.23	3.25
36-41	670	1.21	1.93	1.92	1.93
42	6092	10.93	5.19	5.17	5.17
Total	55,726	100	100	100	100

4 DISCUSSION

This report presents regression analyses that enable the estimation of productivity losses associated with the health of the patient. These analyses enable the prediction of days off paid employment and days off normal activities using the health and socio-demographic characteristics of the patient. The linkage between EQ-5D and ICD chapter with productivity loss enables the measurement of the impact of interventions on productivity lost, via their impact on health. The estimate number of days can then be converted into a cost value using available methods in the existing literature.

4.1 HRQoL and productivity

The relationship between EQ-5D and productivity losses was consistent and significant, where lower EQ-5D score meant higher productivity losses. This result is broadly similar to that reported by Krol, et al. [2013b] though the models were different. Patients with comorbidities also consistently had significantly higher productivity losses. The relationship between whether the patient had an operation and productivity losses was not as clear, but generally indicated that having an operation reduced the number of days off paid employment, but increased the number of days off normal activities. However, only the latter was statistically significant. There may be several explanations for these differences: the type of operation may have differed across those who were in work and the rest of the sample; the flexibility of choice in doing normal activities may have allowed respondents to take more time off from this while recovering, compared to time off work which may be dictated by other considerations such as limitations on how long they could have off, or the level of income they would receive while off sick.

Most of the ICD chapters were positive for days off paid employment, though they were not statistically significant. Conditions associated with ICD chapters C (neoplasms), F (mental and behavioural disorders), I (circulatory system), M (musculoskeletal system), Q (congenital malformations), S (injury) and T (burns/poisoning/external causes) were associated with higher number of days off paid employment than certain infectious and parasitic diseases (A). ICD chapters C, M, Q, S and T were also associated with a higher number of days off normal activities. However, there were more ICD chapters with negative coefficients for days off normal activities, with statistically significant effects for H (diseases of the eye, ear and mastoid process) and O (pregnancy) indicating they were associated with a lower number of days off normal activities compared to Chapter A. Diagnosis was therefore associated in different ways to days off paid employment compared to days off normal activities for some conditions, which is not unexpected.

Exclusion of comorbidity, operation and insignificant ICD codes offered some improvements for the models as indicated by the BIC statistic, with slight changes in coefficient sizes. The change in coefficient size was expected as comorbidity was initially statistically significant and its effect was subsequently captured by the other variables when it was omitted. However, omission of the comorbidity variable had little impact on days predicted.

4.2 Socio-demographics

The sign and significance of age and age-squared were consistent for days off paid employment and days off normal activities, where age was positive and age-squared was negative. This was expected as older people are more likely to have poorer health and this would have an impact on productivity, but at a decreasing rate. Being female was associated with a significantly lower number of days off paid employment and normal activities.

4.3 Model selection

The distribution of days off paid employment and days off normal activities in the last 6 weeks 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 spike at 42 days. A variety of different regression models were estimated that are appropriate for the type and distribution of the data. 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 the 42 day spike, but still under-predicted the very large spike at 0 days. The negative binomial and 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. Unsurprisingly, none of the models were able to accommodate the spikes at multiples of 5 and 7 days.

The choice of preferred model between the negative binomial, zero-inflated negative binomial model with constant inflation and the zero-inflated negative binomial model with variable inflation, needed 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 productivity cost as these individuals have had to take every day off work in the last 6 weeks, and the very large spike at 0 days, which represents zero productivity cost. However, between around 5 and 41 days, all models over-predict or under-predict the number of days.

Model selection criteria indicated that for days off paid employment, the 42 day models were preferred to the 30 day models. The loss of information that resulted when the days off were censored, led to slight changes in the coefficients and some differences in the predicted days, though these were small. The primary motivation behind censoring productivity losses at 30 days is that it represents the maximum productivity loss among fully employed people for a six week recall period. Based on this argument, any reported productivity loss above 30 days must necessarily include non-working days too (e.g. weekends). As such, including these additional days would have overestimated the predicted productivity losses. However, censoring at 30 days ignored all the observations under 30 days that appeared to include weekends; there are small spikes at 7, 14, 21 and 28 days suggestive of whole weeks rather than working weeks being included in patient responses. An alternative approach would have been to interpret 42 days of lost productivity as reflecting a 100% loss in productivity. As such, a pro rata adjustment of all data would be more sensible. However, this would adjust data that was considered 'legitimate'; there were spikes at multiples of 5 days suggestive of only working weeks being reported by some respondents. As such, there does not appear to be a single interpretation of values in excess of 30 days; therefore, there is not a clearly superior data adjustment. Consequently, the data adjustment undertaken, whilst solving the need for a maximum productivity loss of 5 days per week, assumes a particular problem with reporting by respondents that is not fully supported by the data. However, it is accepted that alternative adjustments are also prone to error. The reason for favouring one adjustment over another is not clear, nor is the sensitivity of predictions to alternative adjustment methods.

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. Across the different samples, the majority of observations related to zero and 42 days; 76% for paid employment, 60% for paid employment if employed and 56% for normal activities. This would suggest that for this patient population, accurate predictions of these parts of the distribution are very important. However, it should also be considered that the HODaR patient population is relatively sick, with most NHS activity being related to healthier patients who would be expected to have a greater proportion of zero days off paid employment. 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 is considered to be the most relevant.

4.4 Limitations

These analyses will not be appropriate for predicting productivity losses when: 1) EQ-5D is inappropriate for the patient population; 2) the relationship between EQ-5D, days off paid employment and days off normal activities is inappropriate for the patient population. Furthermore, results based on HODaR may not be generalizable to some populations. Although HODaR data provided a large sample of patients with different conditions, some conditions based on ICD codes had very small groups, which may have limited the applicability of results in these groups. In particular, the exclusion of those with a primary diagnosis of a psychological illness in HODaR means that the results may not apply to those who have these conditions. The analysis using HODaR data was based on patients who had been recently hospitalised, and they may be significantly different from other patient populations where these results would be applied. A further related limitation of using the HODaR data is the fact that respondents reported past employment for those who were retired. This was addressed by excluding those who were above retirement age and who were therefore more likely to be retired, but this may not have captured everyone and so estimates may have been biased.

There was evidence of a lack of a relationship between days off paid employment and EQ-5D scores. This was in part due to the large number of respondents who report 0 days off paid employment (76% full sample and 60% if employed) indicating they did not take days off paid employment following discharge. Furthermore, the mismatch between the recall periods between the days off paid employment and the EQ-5D may have also affected the relationship for those who reported positive count days. The recall period for days off paid employment was the past 6 weeks, while EQ-5D is based on health 'today'; this mismatch may have resulted in a mismatch between health and productivity loss. Respondents may have had time off work in the 6 week period but their health on the day they completed the questionnaire i.e. today in EQ-5D, may have been good, for example, 2,779/8,456 respondents who had an EQ-5D score of 1 reported days off paid employment in the employed sample with 332 of these reporting 42 days off paid employment. Furthermore, some of those in full health who reported a large number of days off paid employment may have conditions that are not captured by the EQ-5D because it is inappropriate for the population.

A final concern with regards to HODaR relates to potential recall bias in the 6 week recall period. The spikes in the data where full weeks have been counted e.g. 7, 14 etc. indicated that respondents were rounding off days which may indicate some form of recall bias. Although recommended recall periods of 3 months for productivity have been suggested, [Zhang, et al. 2011] the evidence in HODaR suggests that there is some bias over a shorter period.

While this study provides a means to estimate the productivity losses associated with HRQoL, the valuation of these losses can be challenging where there is no information on the type of work or usual activities which have not been undertaken.

4.5 Conclusion

Models to predict days of work and days off normal activities were estimated using a large patient dataset. These allow productivity losses associated with HRQoL to be estimated for inclusion in economic evaluation. The use of a large patient dataset meant that the models were based on HRQoL from a number of conditions. There were a number of limitations and further research is recommended to address the generalizability of the models.

5 REFERENCES

Akobundu E, Ju J, Blatt L, Mullins C D. Pharmacoeconomics. *Cost-of-Illness Studies*. 2006; 24(9): 869-890.

Black C, Frost D. The Stationery Office. *Health at work – an independent review of sickness absence*. 2011; https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/181060/health-at-work.pdf.

Brooks R, EuroQol Group. Health Policy. *EuroQol: the current state of play.* 1996; 37: 53-72.

Currie C J, McEwan P, Peters J R, Patel T C, Dixon S. Value in Health. *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.* 2005; 8: 581-590.

Department of Health. *Value base pricing: impact assessment*. London. Department of Health, 2010; (available at: http://www.dh.gov.uk/en/Consultations/Liveconsultations/DH_122760).

Dixon S et al. Pharmacoeconomics. *Incorporating carer effects into economic evaluation.* 2006; 24(1): 43-53.

Dolan P. Medical Care. *Modeling valuations for EuroQol health states.* 1997; 35: 1,095-108.

Johannesson M, Jönsson B, Jönsson L, Kobelt G, Zethraeus N. Office of Health Economics Briefing. *Why Should Economic Evaluations of Medical Innovations Have a Societal Perspective?* 2009.

Koopmanschap M A, van Ineveld B M. Social Science and Medicine. *Towards a new approach for estimating indirect costs of disease.* 1992; 34: 1,005-1,010.

Krol M, Brouwer W, Rutten F. PharmacoEconomics. *Productivity Costs in Economic Evaluations: Past, Present, Future.* 2013a; 1-13.

Krol M, Stolk E, Brouwer W. The European Journal of Health Economics. *Predicting productivity based on EQ-5D: an explorative study.* 2013b; 1-11.

Lamers L M, Meerding W J, Severens J L, Brouwer W B. Quality of life research. *The relationship between productivity and health-related quality of life: an empirical exploration in persons with low back pain.* 2005; *14*(3), 805-813.

National Institute of Health and Care Excellence (NICE). *Guide to the methods of technology appraisal 2013*. 2013; London: NICE.

Rowen D, Dixon S, Mukuria C, Hernandez M. EEPRU Research Report. *Examining informal care associated with health related quality of life using patient data.* 2013; http://www.eepru.org.uk/Informal%20Care%20Research%20Report.pdf

Sculpher M. The role and estimation of productivity costs in economic evaluation. In Drummond M, McGuire A, (Eds) *Economic evaluation in health care: merging theory with practice*, *2*. 2001; Oxford University Press.

Shiell A, Gerard K, Donaldson C. Health policy. *Cost of illness studies: an aid to decision-making?* 1987; 8(3), 317-323.

van Roijen L, Essink-Bot M L, Koopmanschap M A, Bonsel G, Rutten F F. International Journal of Technology Assessment in Health Care. *Labor and health status in economic evaluation of health care. The health and labor questionnaire.* 1996; 12(3), 405-415.

Weisbrod B A. The valuation of human capital. *Journal of Political Economy.* 1961; 69: 425-436.

World Health Organisation (WHO) *International Statistical Classification of Diseases and Related Health Problems, 10th Revision (ICD-10) Volume 2 Instruction Manual*. 2010; http://www.who.int/classifications/icd/ICD10Volume2_en_2010.pdf

Zhang W, Bansback N, Anis A H. Social science & medicine*. Measuring and valuing productivity loss due to poor health: a critical review.* 2011; *72*(2), 185-192.

6 APPENDIX

Table A1: ICD-10 Chapters

A	Certain infectious and parasitic diseases
B	Certain infectious and parasitic diseases
C	Neoplasms
D	Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism
E	Endocrine, nutritional and metabolic diseases
F	Mental and behavioural disorders
G	Diseases of the nervous system
H	Diseases of the eye and adnexa and diseases of the ear and mastoid process
I	Diseases of the circulatory system
J	Diseases of the respiratory system
K	Diseases of the digestive system
L	Diseases of the skin and subcutaneous tissue
M	Diseases of the musculoskeletal system and connective tissue
N	Diseases of the genitourinary system
O	Pregnancy, childbirth and the puerperium
P	Certain conditions originating in the perinatal period
Q	Congenital malformations, deformations and chromosomal abnormalities
R	Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified
S	Injuries
T	Burns, poisoning and certain other consequences of external causes
V-Y	External causes of morbidity and mortality
Z	Factors influencing health status and contact with health services
U	Codes for special purposes

Table A2: Regression output for days off paid employment in the last 6 weeks

	(1)		(2)	(3)	(4)	(5.1)	(5.2)	(6)	(7)
	Tobit	Tobit marginal effects	Poisson	RE Poisson	Negative binomial	Two-part model: Probit	Two-part model: Truncated negative binomial	Zero-inflated negative binomial, constant inflation	Zero-inflated negative binomial, variable inflation
Primary ICD chapter	$\hat{\beta}$	$\hat{\beta}$	$\hat{\beta}$	$\hat{\beta}$	$\hat{\beta}$	$\hat{\beta}$	$\hat{\beta}$	$\hat{\beta}$	$\hat{\beta}$
B	-17.516**	-1.327***	-0.527	0.475***	-1.549*	-0.495**	0.047	-0.725	0.001
	(0.034)		(0.108)	(0.002)	(0.070)	(0.027)	(0.852)	(0.307)	(0.998)
C	6.362	0.504	0.473*	0.770***	-0.775	-0.001	0.495**	-0.227	0.495**
	(0.372)		(0.064)	(0.000)	(0.349)	(0.997)	(0.014)	(0.742)	(0.012)
D	-13.907**	-1.065***	-0.365	0.336***	-1.515*	-0.424**	0.127	-0.799	0.092
	(0.050)		(0.155)	(0.000)	(0.068)	(0.022)	(0.533)	(0.250)	(0.645)
E	-16.117**	-1.226***	-0.360	0.341***	-1.642**	-0.498***	0.171	-0.790	0.132
	(0.026)		(0.178)	(0.001)	(0.049)	(0.009)	(0.419)	(0.258)	(0.527)
F	-10.344	-0.796	-0.152	0.339**	-0.834	-0.439	0.524*	-0.126	0.487*
	(0.357)		(0.709)	(0.025)	(0.355)	(0.110)	(0.054)	(0.863)	(0.074)
G	-6.583	-0.510	-0.060	0.175*	-1.058	-0.258	0.273	-0.511	0.252
	(0.359)		(0.818)	(0.068)	(0.209)	(0.168)	(0.182)	(0.463)	(0.208)
H	-14.539**	-1.112***	-0.418	0.171*	-1.710**	-0.445**	0.143	-0.957	0.107
	(0.043)		(0.111)	(0.086)	(0.040)	(0.018)	(0.490)	(0.170)	(0.597)
I	-0.646	-0.051	0.218	0.784***	-0.833	-0.156	0.414**	-0.253	0.406**
	(0.927)		(0.391)	(0.000)	(0.315)	(0.398)	(0.039)	(0.714)	(0.038)
J	-14.157**	-1.083***	-0.413	0.397***	-1.106	-0.427**	0.080	-0.570	0.042
	(0.046)		(0.115)	(0.000)	(0.198)	(0.021)	(0.698)	(0.418)	(0.833)
K	-13.408*	-1.032***	-0.381	0.196**	-1.470*	-0.393**	0.100	-0.774	0.070
	(0.057)		(0.135)	(0.035)	(0.078)	(0.032)	(0.621)	(0.266)	(0.721)
L	-18.165**	-1.378***	-0.580**	0.001	-1.545*	-0.525***	0.004	-0.817	-0.051
	(0.012)		(0.029)	(0.991)	(0.076)	(0.005)	(0.984)	(0.249)	(0.805)
M	-0.223	-0.017	0.174	0.635***	-0.808	-0.116	0.372*	-0.345	0.363*
	(0.975)		(0.492)	(0.000)	(0.334)	(0.529)	(0.064)	(0.619)	(0.064)
N	-12.757*	-0.980***	-0.399	0.362***	-1.351	-0.360*	0.026	-0.775	-0.001
	(0.070)		(0.118)	(0.000)	(0.106)	(0.050)	(0.898)	(0.264)	(0.996)
O	-19.242***	-1.455***	-0.512*	0.237**	-1.673**	-0.607***	0.180	-0.751	0.119
	(0.007)		(0.051)	(0.034)	(0.045)	(0.001)	(0.393)	(0.281)	(0.565)
Q	-3.998	-0.311	0.092	0.478***	-1.102	-0.229	0.387*	-0.498	0.375*
	(0.621)		(0.747)	(0.001)	(0.193)	(0.288)	(0.085)	(0.478)	(0.088)
R	-14.730**	-1.130***	-0.460*	0.034	-1.561*	-0.413**	-0.020	-0.868	-0.056
	(0.036)		(0.071)	(0.717)	(0.060)	(0.024)	(0.920)	(0.210)	(0.777)
S	6.177	0.490	0.354	0.839***	-0.800	0.088	0.385*	-0.389	0.387**
	(0.385)		(0.165)	(0.000)	(0.334)	(0.635)	(0.056)	(0.573)	(0.049)
T	-9.840	-0.758**	-0.102	0.526***	-1.320	-0.366*	0.360*	-0.530	0.338*
	(0.172)		(0.694)	(0.000)	(0.111)	(0.051)	(0.077)	(0.445)	(0.090)
Z	-18.79***	-1.426***	-0.514**	0.099	-1.719**	-0.579***	0.150	-0.891	0.090

	(1)		(2)	(3)	(4)	(5.1)	(5.2)	(6)	(7)
	Tobit	Tobit marginal effects	Poisson	RE Poisson	Negative binomial	Two-part model: Probit	Two-part model: Truncated negative binomial	Zero-inflated negative binomial, constant inflation	Zero-inflated negative binomial, variable inflation
	(0.008)		(0.048)	(0.295)	(0.040)	(0.002)	(0.466)	(0.203)	(0.655)
EQ-5D score	-6.725***	-0.528***	-0.394***	-0.874***	-0.547***	0.060**	-0.718***	-0.696***	-0.733***
	(0.000)		(0.000)	(0.000)	(0.000)	(0.021)	(0.000)	(0.000)	(0.000)
Comorbidity	5.988***	0.468***	0.333***	0.139***	0.294***	0.091***	0.228***	0.255***	0.219***
	(0.000)		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Operation	-1.161*	-0.091*	-0.002	-0.077***	-0.063	-0.041**	-0.000	-0.019	-0.008
	(0.093)		(0.927)	(0.000)	(0.407)	(0.029)	(0.991)	(0.666)	(0.710)
Age	4.822***	0.378***	0.233***	0.186***	0.220***	0.113***	0.056***	0.199***	0.055***
	(0.000)		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Age-squared	-0.062***	-0.005***	-0.003***	-0.002***	-0.003***	-0.001***	-0.001***	-0.002***	-0.001***
	(0.000)		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Female	-3.385***	-0.266***	-0.166***	-0.232***	-0.202***	-0.050***	-0.073***	-0.159***	-0.063***
	(0.000)		(0.000)	(0.000)	(0.000)	(0.002)	(0.000)	(0.000)	(0.001)
Constant	-84.621***		-2.083***	-0.710***	-0.261	-1.897***	1.827***	0.127	1.891***
	(0.000)		(0.000)	(0.000)	(0.760)	(0.000)	(0.000)	(0.858)	(0.000)
Inflated variables								$\hat{\gamma}$	$\hat{\gamma}$
Constant								0.689***	4.367***
								(0.000)	(0.000)
EQ-5D score									-0.078**
									(0.043)
Comorbidity									-0.239***
									(0.000)
Age									-0.226***
									(0.000)
Age-squared									0.003***
									(0.000)
Female									0.172***
									(0.000)
Dispersion (α)					10.931		0.966	1.743	0.985
					(0.208)		(0.020)	(0.094)	(0.021)
Observations	51326		51326	51326	51326	51326	12268	51326	51326
Number of patients				40019					

Notes: p values in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%

Table A3: Regression output for days off paid employment in the last 6 weeks, if employed

	(1)		(2)	(3)	(4)	(5.1)	(5.2)	(6)	(7)
	Tobit	Tobit marginal effects	Poisson	RE Poisson	Negative binomial	Two-part model: Probit	Two-part model: Truncated negative binomial	Zero-inflated negative binomial, constant inflation	Zero-inflated negative binomial, variable inflation
	$\hat{\boldsymbol{\beta}}$	$\hat{\boldsymbol{\beta}}$	$\hat{\boldsymbol{\beta}}$	$\hat{\boldsymbol{\beta}}$	$\hat{\boldsymbol{\beta}}$	$\hat{\boldsymbol{\beta}}$	$\hat{\boldsymbol{\beta}}$	$\hat{\boldsymbol{\beta}}$	$\hat{\boldsymbol{\beta}}$
Primary ICD chapter									
B	-12.203*	-1.222*	-0.458	0.763***	-0.171	-0.452*	0.066	0.042	0.014
	(0.076)		(0.142)	(0.000)	(0.603)	(0.050)	(0.796)	(0.869)	(0.957)
C	14.658***	1.521***	0.553**	1.254***	0.809***	0.200	0.517**	0.530***	0.513**
	(0.006)		(0.013)	(0.000)	(0.001)	(0.263)	(0.011)	(0.008)	(0.010)
D	-8.736*	-0.881*	-0.265	0.825***	-0.068	-0.378**	0.183	0.155	0.142
	(0.098)		(0.241)	(0.000)	(0.787)	(0.034)	(0.374)	(0.442)	(0.484)
E	-8.840	-0.890*	-0.252	0.767***	-0.133	-0.373**	0.173	0.138	0.138
	(0.109)		(0.293)	(0.000)	(0.619)	(0.044)	(0.424)	(0.517)	(0.517)
F	-2.797	-0.284	0.037	0.878***	0.621	-0.373	0.616**	0.615**	0.570**
	(0.793)		(0.922)	(0.000)	(0.124)	(0.207)	(0.016)	(0.016)	(0.026)
G	-1.370	-0.140	0.034	0.558***	0.206	-0.194	0.322	0.305	0.296
	(0.799)		(0.883)	(0.000)	(0.416)	(0.285)	(0.121)	(0.133)	(0.145)
H	-7.775	-0.785	-0.268	0.684***	-0.022	-0.333*	0.184	0.154	0.149
	(0.148)		(0.249)	(0.000)	(0.932)	(0.066)	(0.379)	(0.453)	(0.468)
I	6.726	0.691	0.321	1.131***	0.591**	-0.013	0.450**	0.458**	0.439**
	(0.199)		(0.150)	(0.000)	(0.017)	(0.942)	(0.027)	(0.021)	(0.027)
J	-9.800*	-0.986*	-0.329	0.752***	-0.142	-0.401**	0.112	0.082	0.067
	(0.068)		(0.156)	(0.000)	(0.578)	(0.026)	(0.591)	(0.687)	(0.742)
K	-7.278	-0.737	-0.269	0.603***	-0.055	-0.298*	0.136	0.114	0.104
	(0.162)		(0.228)	(0.000)	(0.826)	(0.092)	(0.506)	(0.569)	(0.602)
L	-12.319**	-1.235**	-0.491**	0.337***	-0.343	-0.452**	-0.001	-0.048	-0.060
	(0.022)		(0.039)	(0.005)	(0.193)	(0.013)	(0.996)	(0.819)	(0.778)
M	4.160	0.427	0.231	1.101***	0.484*	-0.050	0.412**	0.410**	0.397**
	(0.425)		(0.298)	(0.000)	(0.050)	(0.779)	(0.043)	(0.039)	(0.046)
N	-6.883	-0.696	-0.283	0.742***	-0.055	-0.274	0.056	0.041	0.026
	(0.188)		(0.208)	(0.000)	(0.825)	(0.122)	(0.786)	(0.840)	(0.897)
O	-12.477**	-1.250**	-0.420*	0.666***	-0.210	-0.538***	0.199	0.159	0.135
	(0.019)		(0.071)	(0.000)	(0.415)	(0.003)	(0.352)	(0.449)	(0.521)
Q	4.450	0.457	0.259	0.859***	0.375	-0.054	0.405*	0.394*	0.392*
	(0.496)		(0.314)	(0.000)	(0.195)	(0.804)	(0.080)	(0.082)	(0.084)
R	-9.927*	-1.001**	-0.392*	0.339***	-0.245	-0.352**	-0.009	-0.039	-0.050
	(0.057)		(0.080)	(0.002)	(0.325)	(0.047)	(0.966)	(0.846)	(0.804)
S	11.098**	1.148**	0.405*	1.305***	0.648***	0.204	0.420**	0.421**	0.417**
	(0.036)		(0.070)	(0.000)	(0.009)	(0.256)	(0.040)	(0.035)	(0.037)
T	-0.537	-0.055	0.067	0.862***	0.310	-0.160	0.385*	0.373*	0.367*
	(0.922)		(0.769)	(0.000)	(0.223)	(0.384)	(0.064)	(0.067)	(0.071)

	(1)		(2)	(3)	(4)	(5.1)	(5.2)	(6)	(7)
	Tobit	Tobit marginal effects	Poisson	RE Poisson	Negative binomial	Two-part model: Probit	Two-part model: Truncated negative binomial	Zero-inflated negative binomial, constant inflation	Zero-inflated negative binomial, variable inflation
Z	-12.490**	-1.252**	-0.415*	0.446***	-0.275	-0.501***	0.167	0.116	0.106
	(0.019)		(0.071)	(0.000)	(0.279)	(0.005)	(0.425)	(0.571)	(0.606)
EQ-5D score	-15.54***	-1.586***	-0.621***	-0.815***	-0.799***	-0.193***	-0.711***	-0.732***	-0.732***
	(0.000)		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Comorbidity	6.546***	0.666***	0.345***	0.109***	0.285***	0.114***	0.232***	0.237***	0.223***
	(0.000)		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Operation	-2.710***	-0.277***	-0.063**	-0.063***	-0.117***	-0.097***	-0.002	-0.015	-0.015
	(0.000)		(0.020)	(0.000)	(0.000)	(0.000)	(0.915)	(0.513)	(0.507)
Age	3.855***	0.394***	0.143***	0.142***	0.160***	0.103***	0.033***	0.044***	0.031***
	(0.000)		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Age-squared	-0.051***	-0.005***	-0.002***	-0.002***	-0.002***	-0.001***	-0.000***	-0.000***	-0.000***
	(0.000)		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Female	-1.980***	-0.202***	-0.088***	-0.141***	-0.095***	-0.016	-0.041**	-0.041**	-0.030
	(0.001)		(0.000)	(0.000)	(0.001)	(0.376)	(0.030)	(0.030)	(0.121)
Constant	-57.04***		-0.049	-0.400*	-0.313	-1.307***	2.193***	2.070***	2.276***
	(0.000)		(0.857)	(0.079)	(0.305)	(0.000)	(0.000)	(0.000)	(0.000)
Inflated variables								$\hat{\gamma}$	$\hat{\gamma}$
Constant								0.297***	2.435***
								(0.000)	(0.000)
EQ-5D score									0.382***
									(0.000)
Comorbidity									-0.289***
									(0.000)
Age									-0.164***
									(0.000)
Age-squared									0.002***
									(0.000)
Female									0.109***
									(0.000)
Dispersion (α)					6.967		0.944	0.958	0.961
					(0.089)		(0.020)	(0.021)	(0.021)
Observations	27124		27124	27124	27124	27124	10970	27124	27124
Number of patients				22367					

Notes: p values in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%

Table A4: Regression output for days off normal activities excluding paid employment in the last 6 weeks

	(1) Tobit	Tobit marginal effects	(2) Poisson	(3) RE Poisson	(4) Negative binomial	(5.1) Two-part model: Probit	(5.2) Two-part model: Truncated negative binomial	(6) Zero-inflated negative binomial, constant inflation	(7) Zero-inflated negative binomial, variable inflation
	$\hat{\beta}$	$\hat{\beta}$	$\hat{\beta}$	$\hat{\beta}$	$\hat{\beta}$	$\hat{\beta}$	$\hat{\beta}$	$\hat{\beta}$	$\hat{\beta}$
Primary ICD chapter									
B	-7.240	-0.933*	-0.180	0.131	-0.279	-0.305*	0.021	0.003	0.009
	(0.132)		(0.344)	(0.133)	(0.233)	(0.073)	(0.884)	(0.983)	(0.950)
C	1.810	0.237	0.134	0.636***	0.141	-0.044	0.212**	0.210**	0.211**
	(0.621)		(0.335)	(0.000)	(0.410)	(0.728)	(0.045)	(0.047)	(0.043)
D	-8.852**	-1.139***	-0.289**	0.336***	-0.263	-0.341***	0.031	0.016	0.019
	(0.016)		(0.042)	(0.000)	(0.131)	(0.007)	(0.773)	(0.880)	(0.861)
E	-9.404**	-1.207***	-0.288*	0.276***	-0.355*	-0.351***	-0.005	-0.025	-0.018
	(0.014)		(0.055)	(0.000)	(0.054)	(0.008)	(0.963)	(0.830)	(0.871)
F	-19.516***	-2.436***	-0.721***	0.564***	-0.729**	-0.613***	-0.209	-0.240	-0.244
	(0.001)		(0.007)	(0.000)	(0.014)	(0.002)	(0.311)	(0.254)	(0.246)
G	-4.044	-0.525	-0.081	0.311***	-0.126	-0.177	0.073	0.065	0.068
	(0.278)		(0.570)	(0.000)	(0.471)	(0.167)	(0.500)	(0.548)	(0.528)
H	-14.116***	-1.795***	-0.613***	-0.033	-0.666***	-0.464***	-0.167	-0.211*	-0.190*
	(0.000)		(0.000)	(0.538)	(0.000)	(0.000)	(0.127)	(0.056)	(0.080)
I	0.259	0.034	0.072	0.520***	0.162	-0.069	0.170	0.170	0.167
	(0.943)		(0.605)	(0.000)	(0.344)	(0.580)	(0.105)	(0.105)	(0.108)
J	-6.834*	-0.883**	-0.177	0.222***	-0.221	-0.301**	0.066	0.052	0.056
	(0.064)		(0.213)	(0.000)	(0.205)	(0.017)	(0.540)	(0.628)	(0.601)
K	-8.332**	-1.077***	-0.296**	0.243***	-0.262	-0.288**	-0.024	-0.036	-0.033
	(0.021)		(0.033)	(0.000)	(0.127)	(0.020)	(0.822)	(0.736)	(0.749)
L	-11.163***	-1.427***	-0.400***	0.110**	-0.487***	-0.404***	-0.070	-0.097	-0.088
	(0.003)		(0.007)	(0.045)	(0.007)	(0.002)	(0.535)	(0.394)	(0.434)
M	3.931	0.518	0.190	0.655***	0.271	0.074	0.224**	0.224**	0.223**
	(0.277)		(0.169)	(0.000)	(0.111)	(0.551)	(0.033)	(0.033)	(0.031)
N	-7.439**	-0.961***	-0.290**	0.388***	-0.224	-0.256**	-0.031	-0.040	-0.039
	(0.041)		(0.039)	(0.000)	(0.195)	(0.040)	(0.770)	(0.707)	(0.711)
O	-12.428***	-1.583***	-0.709***	-0.144*	-0.682***	-0.488***	-0.250**	-0.293**	-0.283**
	(0.001)		(0.000)	(0.076)	(0.000)	(0.000)	(0.040)	(0.017)	(0.020)
Q	0.689	0.090	0.150	0.940***	0.196	-0.110	0.274*	0.271*	0.270*
	(0.889)		(0.436)	(0.000)	(0.367)	(0.509)	(0.051)	(0.054)	(0.053)
R	-10.192***	-1.311***	-0.368***	0.203***	-0.442***	-0.326***	-0.113	-0.131	-0.125
	(0.005)		(0.008)	(0.000)	(0.010)	(0.009)	(0.286)	(0.215)	(0.232)
S	10.453***	1.394***	0.425***	0.865***	0.490***	0.157	0.389***	0.391***	0.390***
	(0.005)		(0.002)	(0.000)	(0.004)	(0.215)	(0.000)	(0.000)	(0.000)
T	0.469	0.061	0.108	0.594***	0.187	-0.139	0.264**	0.258**	0.258**
	(0.901)		(0.444)	(0.000)	(0.286)	(0.283)	(0.014)	(0.016)	(0.015)

	(1)		(2)	(3)	(4)	(5.1)	(5.2)	(6)	(7)
	Tobit	Tobit marginal effects	Poisson	RE Poisson	Negative binomial	Two-part model: Probit	Two-part model: Truncated negative binomial	Zero-inflated negative binomial, constant inflation	Zero-inflated negative binomial, variable inflation
Z	-12.791***	-1.631***	-0.456***	0.071	-0.565***	-0.468***	-0.048	-0.079	-0.067
	(0.001)		(0.001)	(0.175)	(0.001)	(0.000)	(0.657)	(0.466)	(0.530)
EQ-5D score	-42.76***	-5.603***	-1.418***	-1.138***	-1.942***	-1.198***	-0.777***	-0.823***	-0.778***
	(0.000)		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Comorbidity	5.278***	0.688***	0.281***	0.158***	0.281***	0.118***	0.175***	0.184***	0.172***
	(0.000)		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Operation	0.070	0.009	0.036**	-0.003	-0.020	-0.015	0.025*	0.021	0.023*
	(0.869)		(0.033)	(0.674)	(0.329)	(0.307)	(0.060)	(0.114)	(0.083)
Age	0.457***	0.060***	0.022***	0.001	0.022***	0.013***	0.012***	0.013***	0.012***
	(0.000)		(0.000)	(0.796)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Age-squared	-0.006***	-0.001***	-0.000***	-0.000	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)		(0.000)	(0.501)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Female	1.644***	0.215***	0.010	0.032	0.059***	0.134***	-0.091***	-0.084***	-0.090***
	(0.000)		(0.494)	(0.168)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	17.809***		2.467***	2.276***	2.837***	0.706***	2.809***	2.831***	2.824***
	(0.000)		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Inflated variables								$\hat{\gamma}$	$\hat{\gamma}$
Constant								0.203***	-0.769***
								(0.000)	(0.000)
EQ-5D score									2.085***
									(0.000)
Comorbidity									-0.274***
									(0.000)
Age									-0.027***
									(0.000)
Age-squared									0.000***
									(0.000)
Female									-0.183***
									(0.000)
Dispersion (α)					5.663		0.652	0.673	0.657
					(0.054)		(0.008)	(0.009)	(0.008)
Observations	55726		55726	55726	55726	55726	24328	55726	55726
Number of patients				42442					

Notes: p values in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%