

Parametric hazard rate models for long-term sickness absence

Petra C. Koopmans · Corné A. M. Roelen ·
Johan W. Groothoff

Received: 26 November 2007 / Accepted: 29 September 2008 / Published online: 21 October 2008
© The Author(s) 2008. This article is published with open access at Springerlink.com

Abstract

Purpose In research on the time to onset of sickness absence and the duration of sickness absence episodes, Cox proportional hazard models are in common use. However, parametric models are to be preferred when time in itself is considered as independent variable. This study compares parametric hazard rate models for the onset of long-term sickness absence and return to work.

Method Prospective cohort study on sickness absence with four follow-up years of 53,830 employees working in the private sector in the Netherlands. The time to onset of long-term (>6 weeks) sickness absence and return to work were modelled by parametric hazard rate models.

Results The exponential parametric model with a constant hazard rate most accurately described the time to onset of long-term sickness absence. Gompertz–Makeham models with monotonically declining hazard rates best described return to work.

Conclusions Parametric models offer more possibilities than commonly used models for time-dependent processes as sickness absence and return to work. However, the advantages of parametric models above Cox models apply

mainly for return to work and less for onset of long-term sickness absence.

Keywords Methodology · Long-term sickness absence · Parametric models · Survival analysis · Cox proportional hazards model · Register data

Introduction

Sickness absence is an important measure for general health in the population. Long-term sickness absence is a predictor of disability and mortality (Gjesdal and Bratberg 2002; Kivimäki et al. 2003) and imposes considerable costs to both the employer and society as a whole (Henderson et al. 2005). An increase in sickness absence is associated with a higher risk of unemployment and job termination (Virtanen et al. 2006; Hesselius 2007; Koopmans et al. 2008). Revealing characteristics of employees at risk of long-term absence is important in order to reduce sickness absence, work disability and unemployment. Occupational health interventions may increase the probability of returning to work and limit economic and social deprivation associated with long-term absence. However, the impact of risk factors or interventions may vary across different stages of the sickness absence. Therefore it is important to gain insight into the time process of return to work (Joling et al. 2006).

In research on time to onset of sickness absence and the duration of sickness absence episodes, Cox proportional hazards models are widely used (Cheadle et al. 1994; Krause et al. 2001; Joling et al. 2006; Lund et al. 2006; Christensen et al. 2007; Blank et al. 2008).

However, Cox proportional hazards models do not address the shape of the baseline hazard. The hazard is the risk of an event, for example the risk of onset of long-term

P. C. Koopmans · C. A. M. Roelen
Department of Occupational Health Services,
ArboNed Noord-Oost Nederland, Groningen, The Netherlands

P. C. Koopmans · C. A. M. Roelen · J. W. Groothoff
Department of Social Medicine,
University Medical Center Groningen,
University of Groningen, Groningen, The Netherlands

P. C. Koopmans (✉)
ArboNed NV, PO Box 141,
9700 AC Groningen, The Netherlands
e-mail: petra.koopmans@arbooned.nl

sickness absence. The baseline hazard can be interpreted as the hazard function for the average individual in the sample. In Cox models, the functional form of the baseline hazard is not given, but is determined from the data. However, the course of sickness absence and reintegration cannot be understood without knowing the baseline hazard function. One way to understand the baseline hazard function is to specify it. For instance, it can be hypothesized that with increasing absence duration the probability of returning to work decreases in a certain pattern (Crook and Moldofsky 1994). Although Cox models leave the baseline hazard unspecified, duration dependence can be imposed. For instance, one may assume that the baseline hazard remains constant in time or varies exponentially with time (see e.g. Bender et al. 2005). However, parametric models are preferred when time in itself is considered a meaningful independent variable and the researcher wants to be able to describe the nature of time-dependence.

Different types of parametric models can be distinguished, depending on the type of time dependence of the hazard rate (Blossfeld and Rohwer 2002), as shown in Fig. 1. In exponential models, the hazard rate is assumed to be constant. Weibull models assume a hazard function that is a power function of duration. Log-logistic models permit non-monotonic hazard functions in which hazard rates can increase and then decrease or vice versa. Log-normal models are quite similar to log-logistic models, though the distribution of the error term is specified to be standard normal. Gompertz–Makeham models assume the hazard rate to be an exponential function of duration times.

The impact of risk factors or interventions may vary in different stages of sickness absence (Krause et al. 2001). When a researcher wants to investigate the effect of covariates on sickness absence and assumes that the effect of these covariates is different depending on the duration of the absence episode, it is better to use parametric models. Some parametric models have a level parameter and a shape parameter, which is allowed to depend on covariates

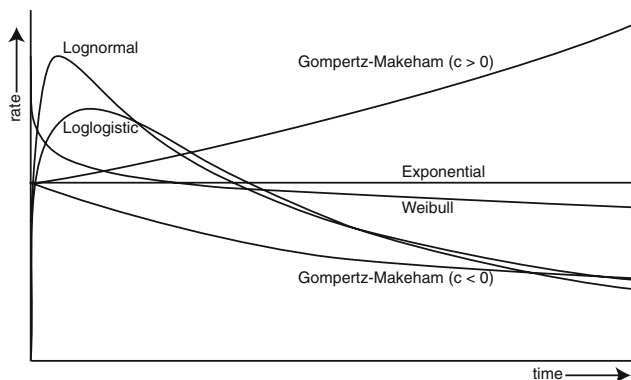


Fig. 1 Different parametric models for time-dependency of the hazard rate

and to vary between groups. The Cox model may include time-dependent covariates. However, the change in covariate value does not affect the shape of the hazard but shifts the hazard to a different level. Also Cox models consume more degrees of freedom than models with parametric duration dependence. One degree of freedom is calculated for every category used in the analysis. For example, when 10 age categories are defined, 10 degrees of freedom are used, one for every baseline hazard. Parametric models only use a limited number of parameters and a corresponding lower number of degrees of freedom. Therefore parametric models are more parsimonious and have more power as compared to Cox models.

The aim of this study was to investigate the time to onset of long-term sickness absence and return to work after long-term sickness absence by means of parametric hazard rate models, in order to identify which model fitted the data best. Instead of modelling total sickness absence (e.g. Joling et al. 2006), we choose to focus on long-term (i.e. more than six consecutive weeks) sickness absence because it has been reported that short term sickness absence is a different construct affected by different factors (Allebeck and Mastekaasa 2004).

Methods

Study design and population

The study population consisted of 53,830 employees of three large and nationally spread Dutch companies in the postal and telecommunications sector. Functions in these companies included sorting and delivery of mail, (parcel) transportation, call center and post office tasks, telecommunication (e.g. mechanics, sales, IT), back-office work, and executive functions. The study design is described elsewhere (Koopmans et al. 2008). Employees aged 55 years or older in the base year were excluded because of possible bias due to senior regulations or early retirement. The study population consisted of 37,955 men (mean age 41 years, SD = 8) and 15,875 women (mean age 39 years, SD = 8). Sickness absence data were retrieved from the occupational health department registry. Long-term sickness absence was defined as absence due to sickness for more than six consecutive weeks. Sickness absence episodes between 1998 and 2001 were recorded. Overlapping and duplicated absence episodes were corrected for. We investigated the time to onset of the first long-term sickness absence and the duration of all long-term sickness absence episodes. In case an employee had not suffered a long-term absence before 31 December 2001 or before the end of the employment period, the period was right censored. For the return to work models, data of employees ($N = 16,433$)

who had at least one long-term absence episode between 1998 and 2001 were used. Return to work was defined as resumption of the contracted work hours/week in one's job. Long-term sickness absence episodes which did not end at 31 December 2001, or which could not be recorded because the employee left employment, were right censored.

Statistics

Survival data were plotted using SPSS life tables. The rates of onset of long-term sickness absence and return to work were parameterized using Transition Data Analysis (TDA, version 6.4f). The time to onset of long-term absence was recorded from days into weeks. The duration of long-term sickness absence was counted in days, but to make the calculations possible, 42 days were subtracted from the absence duration, in order to obtain 1 as the lowest value.

We investigated the following models (Blossfeld and Rohwer 2002):

- (1) Exponential model: the hazard rate can vary with different sets of covariates, but is assumed to be time constant; the hazard function and survivor function are $r(t) = a$, respectively $G(t) = \exp(-at)$, with $t =$ time and $a =$ constant.
- (2) Gompertz–Makeham model: the hazard rate increases or decreases monotonically with time. The hazard function is given by the expression $r(t) = a + b \exp(ct)$, in which a , b and c are constants and $t =$ time. For long durations the hazard rate declines towards the value of parameter a (the Makeham term). If $b = 0$ the model reduces to an exponential model $r(t) = a$, which states the hazard rate is constant over time. The parameter c is the shape parameter. If the parameter c is negative, we conclude that increasing duration of the process leads to a declining hazard rate. If the parameter c is positive, increasing duration leads to an acceleration of the hazard rate.
- (3) Weibull model: the hazard rate increases or decreases exponentially with time: $r(t) = ba^b t^{b-1}$, but like the Gompertz model, it can also be used to model monotonically decreasing ($0 < b < 1$) or increasing rates ($b > 1$). An exponential model is obtained in the special case of $b = 1$.
- (4) Log-logistic model: this model is even more flexible than the Gompertz and Weibull distributions. The hazard rate function is:

$$r(t) = \frac{ba^b t^{b-1}}{1 + (at)^b}$$

For $b \leq 1$ the hazard rate monotonically declines (Gompertz–Makeham) and for $b > 1$ the hazard rate rises monotonically to a maximum and then

decreases monotonically. Thus this model can be used to test a monotonically declining time-dependence against a non-monotonic pattern. This is the most commonly recommended model if the hazard rate is bell-shaped.

- (5) Log-normal model: this model implies a non-monotonic relationship between the hazard rate and the duration; the hazard rate increases to a maximum and then decreases.
- (6) Generalized gamma models can be used to discriminate between exponential, Weibull and log-normal models. It has three parameters: a , b and k of which a can take all values, but b and k must be positive. Special cases are the exponential model, if $b = 1$ and $k = 1$, the Weibull model if $k = 1$, and a log-normal model is reached if $k \rightarrow \infty$.

Nested models are compared using the likelihood ratio (LR) test. Under the null hypothesis that the models do not differ the likelihood test statistic approximately follows a χ^2 distribution with m degrees of freedom where m is the number of additionally included covariates. The LR-test statistic is computed as two times the difference between the log likelihoods (LL): $LR = 2 [LL(\text{present model}) - LL(\text{reference model})]$.

The use of likelihood ratio tests is limited to nested models. In order to compare non-nested models we used the graphical methods described by Blossfeld and Rohwer (2002). We performed a non-parametric estimation of a survivor function using the product limit estimation (Kaplan and Meier 1958). Then, given a parametric assumption, the survivor function is transformed so that the results become a linear function that can be plotted. If the model is appropriate, the resulting plot should be linear and the accuracy of the fit can be evaluated with the R^2 measure. The graphical check, however, is not possible for the Gompertz–Makeham model (unless $a = 0$ or $c = 0$). Pseudoresiduals were also computed to check the statistical fit of the parametric models (Cox and Snell 1968). If the model is appropriate, the pseudoresiduals should follow approximately a standard exponential distribution. A plot of the logarithm of the survivor function against the residuals should be a straight line that passes through the origin (Blossfeld and Rohwer 2002).

Ethical approval

Ethical approval was sought from the Medical Ethics Committee of the University Medical Center Groningen, who advised that according to Dutch law ethical clearance was not required for this secondary study on sickness absence data.

Results

Between 1998 and 2001, 16,433 employees (30%) had a total of 22,159 long-term sickness absence episodes. The majority of workers (73%; 11,923) who were long-term absent had one episode; 21% ($N = 3,495$) had two episodes and 6% ($N = 1,015$) had three or more long-term absence episodes.

Onset of long-term sickness absence

From the generalized gamma distributions with $k = 1$ it can be seen that the exponential model and the Weibull model give the best fit (see Table 1). The Weibull model does not have a better fit than the exponential model ($LR(1) = 2$, $p = 0.157$). The Gompertz–Makeham model does have a better fit than the exponential model:

$LR(2) = 10$ ($p = 0.007$). The negative C-parameter of the Gompertz–Makeham model indicates a declining rate of long-term absence with increasing duration. In Fig. 2 the graphical checks are plotted. The plots of the exponential and the Gompertz–Makeham models show a straight line suggesting good fits. However, the exponential model is the simplest of the parametric alternatives, and seems a good choice because of that simplicity.

In Fig. 3 the actual and estimated long-term absence onset rates are presented.

Return to work

According to the likelihood tests, the Gompertz–Makeham model ($LR(2) = 7,636$, $p < 0.001$) or the Weibull model ($LR(1) = 5,288$, $p < 0.001$) give a better fit for return to work than the exponential model (Table 1). In the

Table 1 Different parametric models for long-term absence onset and return to work, expressed as parameters with standard errors (SE) and log likelihoods (LL)

| Variables | Onset of long-term absence | | | Return to work | | |
|-----------------------------------|----------------------------|--------|---------|----------------|--------|---------|
| | Parameter | SE | LL | Parameter | SE | LL |
| Exponential model | | | | | | |
| A | -6.2371 | 0.0078 | -118348 | -5.3212 | 0.0075 | -113744 |
| Gompertz–Makeham model | | | | | | |
| A | -7.4575 | 0.9907 | -118343 | -6.9978 | 0.0560 | -109926 |
| B | -6.5326 | 0.3942 | | -4.6678 | 0.0123 | |
| C | -0.0006 | 0.0003 | | -0.0057 | 0.0002 | |
| Weibull model | | | | | | |
| A | -6.2497 | 0.0111 | -118347 | -5.1555 | 0.0110 | -111100 |
| B | -0.0118 | 0.0073 | | -0.3753 | 0.0050 | |
| Log-logistic model | | | | | | |
| A | -5.9845 | 0.0108 | -118350 | -4.4048 | 0.0114 | -109874 |
| B | 0.0800 | 0.0071 | | 0.0593 | 0.0061 | |
| Log-normal model | | | | | | |
| A | 6.2706 | 0.0145 | -119466 | 4.4031 | 0.0118 | -109783 |
| B | 0.6969 | 0.0062 | | 0.5060 | 0.0062 | |
| C | -0.0161 | 0.0007 | | -1.0990 | 0.1575 | |
| Generalized gamma ($k = 0.5$) | | | | | | |
| A | 6.2555 | 0.0106 | -118379 | 5.4536 | 0.0108 | -112045 |
| B | -0.2572 | 0.0075 | | 0.2969 | 0.0059 | |
| Generalized gamma ($k = 10$) | | | | | | |
| A | 6.2183 | 0.0126 | -118489 | 4.6523 | 0.0113 | -109993 |
| B | 0.4375 | 0.0066 | | 0.4634 | 0.0055 | |
| Generalized gamma ($k = 1,000$) | | | | | | |
| A | 6.1744 | 0.0132 | -118676 | 4.4396 | 0.0114 | -109807 |
| B | 0.5830 | 0.0063 | | 0.4863 | 0.0054 | |

Fig. 2 Graphical checks of different parametric models for the long-term absence onset rate with **a** graphical check of distributional assumptions, and **b** graphical checks of the pseudoresiduals

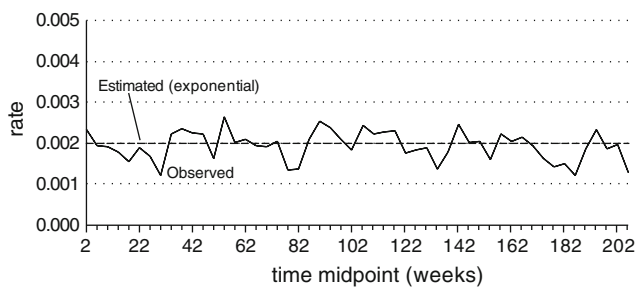
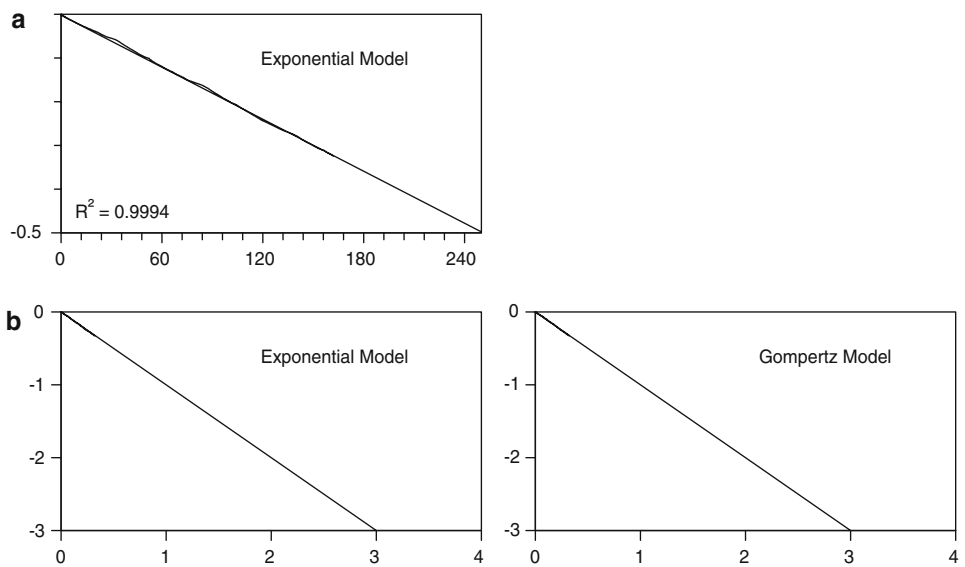


Fig. 3 Observed and estimated long-term absence onset rates according to the exponential model

generalized gamma distribution the fit increased with increasing k . Therefore the log-normal model seems to be a better choice to describe the data than Weibull model. Subsequently, we compared the log-logistic, the log-normal and the Gompertz–Makeham model.

When plotting the transformed survivor function (*a*) and the pseudoresiduals (*b*) of these functions, the best fit was found for the Gompertz–Makeham model (Fig. 4). The pseudoresiduals in the log-logistic and the log-normal

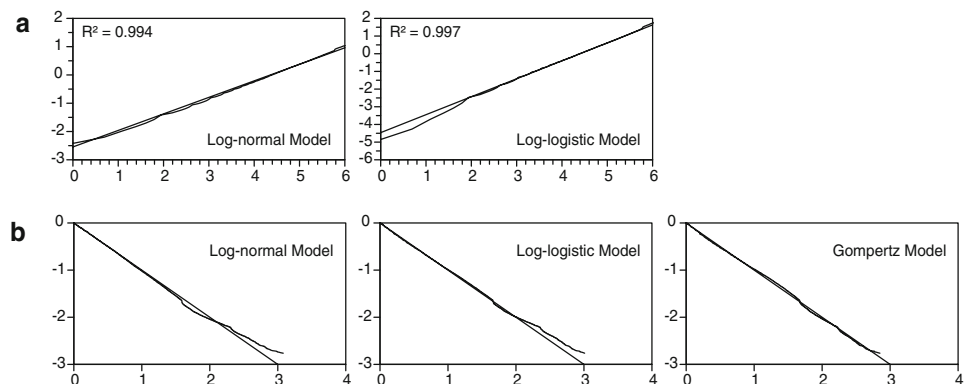
model distribution depart from linearity in the highest values of the residuals.

The hazard rates of the Gompertz–Makeham model and the observed rates are plotted in Fig. 5. Figure 5 shows a remarkable increase in the observed return to work rate at 365 days.

Discussion

Sickness absence is an important outcome measure in epidemiologic research on public health and occupational health intervention studies (Kivimäki et al. 2003; Ruotsalainen et al. 2006). The time concept is an important aspect in sickness absence research. Studies can focus on how long employees are absent from work, how long it takes them to return to work when sick listed, or how long an individual works between different sick leave spells (Hensing 2004). Despite its importance, the time concept has not been investigated in detail. It is known that the probability of return to work decreases as a function of

Fig. 4 Graphical checks of different parametric models for the return to work rate with **a** graphical check of distributional assumptions and **b** graphical checks of the pseudoresiduals



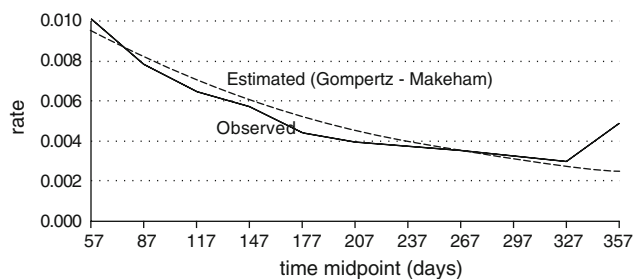


Fig. 5 Observed and estimated return to work rates according to the Gompertz–Makeham model

time, but the actual pattern of this duration dependence has hardly been investigated (Joling et al. 2006).

Researchers often do not specify a parametric form of the baseline hazard function, because they are not interested in it or have no reference as what it might look like. The Cox regression offers a neat way to avoid this issue. The advantage of Cox regression is that the data determine the shape of the hazard function that best fits them. The disadvantage is that data are, as a rule, rather irregular. Parametric models are more useful when a researcher wants to have information what the baseline hazard function might look like.

The advantage of parametric models is that they give a succinct summary of a large amount of data. From our study it appeared that parametric models—in which the hazard function is specified—were accurate in describing the time-dependence of long-term sickness absence: the exponential model for the time to onset of long-term absence and the Gompertz–Makeham model for return to work. The exponential model assumes that the hazard rate from work to long-term sickness absence is constant over time. In our population, the onset of long-term sickness absence can be described by only one parameter. The Gompertz–Makeham model assumes that the hazard rate from long-term sickness absence to work declines monotonically with time, meaning that most employees resume work at an early stage and with increasing absence duration the return to work rate decreases.

However, the models selected do have some shortcomings. The exponential model does not help to overcome some of the disadvantages of the Cox model: (1) the exponential model has a constant hazard, and therefore cannot accommodate duration dependence; (2) the exponential model is a form of proportional hazards model—hazard rate ratios from this model will be independent of time. Also regarding the irregular shape of the observed hazard rate in Fig. 3, it could be argued that Cox models are as adequate for analyzing time to onset to long-term absence as are parametric models.

The return to work rate showed an increase at 365 days of absence. This may be an artefact, because, up to 2004,

disability pension was granted in the Netherlands after 1 year of incapacity to work. Part of the employees may be granted a disability pension and therefore the absence episode will be ended, and others will prefer to return to work instead of receiving a disability pension. The Gompertz–Makeham model does not provide in this increase in the return to work rate. Since 2004 employers pay their employees on sick leave for 2 years and the disability pension date is moved accordingly. It is recommended to study whether the return to work rate of long-term sickness absence since 2004 will be different from before.

Time can be interpreted as a proxy for time-varying causal factors of long-term sickness absence, such as the commitment to the organization, psychosocial factors, medical follow-up and sickness benefits. Given the difficulty of measuring these theoretically important concepts over time, time-dependent parametric models are useful for modelling the changes in the hazard rate over time. Based on our results, we recommend that future sickness absence studies address the issue of time-dependence of return to work using parametric models.

The shape of the baseline hazard may give clues for the ideal moment of intervention programmes aimed at reducing long-term sickness absence. According to the Gompertz–Makeham model of return to work, the probability of success of an intervention to stimulate return to work decreases with the duration of sickness absence. Joling et al. (2006) tested several types of Weibull models of duration dependence for sickness absence. They found positive duration dependence: the return to work rate increased over time. We found negative duration dependence: the return to work rate decreased monotonically over time. The difference is probably due to the fact that Joling et al. analyzed both short term absences and long-term absences, while we focused on sickness absence lasting longer than 6 weeks.

Using the appropriate model, it is possible to estimate how many employees are still absent any point in time after their sickness notice. By adding predictors to the model, it is possible to investigate the presence of variable duration dependence across workers. Early interventions could be targeted to the type of workers most likely to be subject to negative duration dependence (Joling et al. 2006). The Gompertz–Makeham model of return to work has three parameters (A, B and C) to which covariates can be linked. Covariates in the B-term have an impact on the return to work rate. Covariates in the C-term test whether these effects increase or decrease with absence duration. The importance and direction of the influence of covariates on return to work “in the long run” is assessed by linking covariates to the A-term.

About 27% of the long-term absentees had two or more long-term absence episodes. The units of analysis in

survival analysis are episodes and this lowers the standard error of covariate estimates, as compared to an analysis based on independent observations, increasing the possibility of finding significant effects of covariates. There are techniques to deal with this dependence. For example, a model accommodating multiple spells can be applied. It is also possible to add a time-invariant unobserved hazard rate constant specific for each individual ('frailty models'). It summarizes the impact of 'omitted' variables on the hazard rate and can be regarded as person characteristics, for example someone's health status. Christensen et al. (2007) and Joling et al. (2006) applied frailty models to sickness absence data. Christensen et al. demonstrated that frailty models had higher statistical power than standard methods. Combining parametric models with frailty models may be a powerful tool in sickness absence research.

Alternatively, multi-state models may be a useful application to sickness absence research. In multi-state models it is possible to model individuals moving among a finite number of stages, for example from work to sickness absence to work disability or back to work again. Stages can be transient or absorbing (or definite), with death being an example of an absorbing state. To each of the possible transitions covariates can be linked. In multi-state models assumptions can be made about the dependence of hazard rates on time (Putter et al. 2007; Meira-Machado et al. 2008; Lie et al. 2008).

Our results are relevant for further absence research in which the application of parametric hazard rate models should be encouraged. It is important to visualize the baseline hazard and detect risk factors which are associated with certain stages in the sickness absence process. Using these models, groups at risk of long-term absence can be detected and interventions can be timed in order to reduce long-term sickness absence. The choice of a parametric model should be theory-driven instead of data-driven. The current study gives a promising impulse to the development of such a theory.

Acknowledgments The authors wish to thank Prof. Dr. ir. F.J.C. Willekens (Professor of Demography at the Population Research Center, University of Groningen) for his valuable suggestions on the transition rate analysis and his comments on earlier drafts of this paper.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Allebeck P, Mastekaasa A (2004) Chapter 5. Risk factors for sick leave: general studies. *Scand J Public Health* 32:49–108. doi:10.1080/14034950410021853
- Bender R, Augustin T, Blettner M (2005) Generating survival times to simulate Cox proportional hazard models. *Stat Med* 24:1713–1723. doi:10.1002/sim.2059
- Blank L, Peters J, Pickvance S, Wilford J, MacDonald E (2008) A systematic review of the factors which predict return to work for people suffering episodes of poor mental health. *J Occup Rehabil* 18:27–34. doi:10.1007/s10926-008-9121-8
- Blossfeld HP, Rohwer G (2002) Techniques of event history modeling. New approaches to causal analysis, 2nd edn. Lawrence Erlbaum, Mahwah
- Cheadle A, Franklin G, Wolfhagen C, Savarino J, Liu PY, Salley C et al (1994) Factors influencing the duration of work-related disability: a population-based study of Washington state workers' compensation. *Am J Public Health* 84:190–196
- Christensen KB, Andersen PK, Smith-Hansen L, Nielsen ML, Kristensen TS (2007) Analyzing sickness absence with statistical models for survival data. *Scand J Work Environ Health* 33:233–239
- Cox DR, Snell EJ (1968) A general definition of residuals. *J R Stat Soc Ser B Methodol* 30:248–275
- Crook J, Moldofsky H (1994) The probability of recovery and return to work from work disability as a function of time. *Qual Life Res* 3(suppl 1):97–109. doi:10.1007/BF00433383
- Gjesdal S, Bratberg E (2002) The role of gender in long-term sickness absence and transition to permanent disability benefits. *Eur J Public Health* 12:180–186. doi:10.1093/eurpub/12.3.180
- Henderson M, Glozier N, Elliot KH (2005) Long term sickness absence. *BMJ* 330:802–803. doi:10.1136/bmj.330.7495.802
- Hensing G (2004) Chapter 4. Methodological aspects in sickness-absence research. *Scand J Public Health* 32:44–48. doi:10.1080/14034950410021844
- Hesseliuss P (2007) Does sickness absence increase the risk of unemployment? *J Socio-Econ* 36:288–310. doi:10.1016/j.socec.2005.11.037
- Joling C, Groot W, Janssen PPM (2006) Duration dependence in sickness absence: how can we optimize disability management intervention strategies? *J Occup Environ Med* 48:803–814. doi:10.1097/01.jom.0000222583.70927.3e
- Kaplan EL, Meier P (1958) Nonparametric estimation from incomplete observations. *J Am Stat Assoc* 53:457–481. doi:10.2307/2281868
- Kivimäki M, Head J, Ferrie JE, Shipley MJ, Vahtera J, Marmot MG (2003) Sickness absence as a global measure of health: evidence from mortality in the Whitehall II prospective cohort study. *BMJ* 327:364–368. doi:10.1136/bmj.327.7411.364
- Koopmans PC, Roelen CAM, Groothoff JW (2008) Frequent and long-term absence as a risk factor for work disability and job termination among employees in the private sector. *Occup Environ Med* 65:494–499
- Krause N, Frank JW, Dasinger LK, Sullivan TJ, Sinclair SJ (2001) Determinants of duration of disability and return-to-work after work-related injury and illness: challenges for future research. *Am J Ind Med* 40:464–484. doi:10.1002/ajim.1116
- Lie SA, Eriksen HR, Ursin H, Hagen EM (2008) A multi-state model for sick-leave data applied to a randomized control trial study of low back pain. *Scand J Public Health* 36:279–283. doi:10.1177/1403494807086979
- Lund T, Labriola M, Christensen KB, Bültmann U, Villadsen E (2006) Return to work among sickness-absent Danish employees: prospective results from the Danish Work Environment Cohort Study/National Register on Social Transfer Payments. *Int J Rehabil Res* 29:229–235. doi:10.1097/01.mrr.0000210056.24915.c2
- Meira-Machado LF, Una-Alvarez JD, Cadarso-Suarez C, Andersen P (2008) Multi-state models for the analysis of time-to-event data. *Stat Methods Med Res*. doi:10.1177/0962280208092301

- Putter H, Fiocco M, Geskus RB (2007) Tutorial in biostatistics: competing risks and multi-state models. *Stat Med* 26:2389–2430. doi:[10.1002/sim.2712](https://doi.org/10.1002/sim.2712)
- Ruotsalainen JH, Verbeek JH, Salmi JA, Jauhiainen M, Laamanen I, Pasternack I et al (2006) Evidence on the effectiveness of occupational health interventions. *Am J Ind Med* 49:865–872. doi:[10.1002/ajim.20371](https://doi.org/10.1002/ajim.20371)
- Virtanen M, Kivimäki M, Vahtera J, Elovainio M, Sund R, Virtanen P et al (2006) Sickness absence as a risk factor for job termination, unemployment, and disability pension among temporary and permanent employees. *Occup Environ Med* 63:212–217. doi:[10.1136/oem.2005.020297](https://doi.org/10.1136/oem.2005.020297)