Abstract.

The phenomenon of absenteeism in organizations is complex to study due to the multiple, often coincidental, multilevel and interacting causes. Although many explanatory models have already been presented, agreement is still scarce. In nearly all research so far, it has been common practice to measure absenteeism over a certain measurement period and subsequently count the number of days absent as an aggregate measure (time lost) or the number of non-adjacent absence periods, as a non-aggregate measure (absence frequency). The applicability of commonly used statistical models to analyse these data like Ordinary Least Squares (OLS) models and the Poisson regression model (PRM) can however be questioned. An assessment of the impact of the violations of their statistical assumptions sheds a light on, and may question the validity of the massive body of studies in which these previous models have been used. It can be concluded that results of the OLS model are too conservative but can therefore be ‘trusted’, while interpretations of the results of the PRM are more dangerous. As an alternative, the negative binomial regression model (NBRM) is proposed since it is less restricted by its assumptions. This could empirically be shown in a first study where various alternative statistical models have been applied to six organizational samples of real life absenteeism data.

The existing models are then further criticized because of their conceptualisation of employee absenteeism as a unitary phenomenon. A new model has therefore been advocated in which absenteeism is considered to be a compound variable of two processes: first an employee gets ill, second a decision is made on the number of days of absence. This new conceptualisation of absenteeism, proposed as a ‘dual phase’ model is tested with a zero-inflated statistical model.

The negative consequences of the usual aggregation of absenteeism data are highlighted by contrasting the results of the NBRM and the zero-inflated NBRM on aggregated absenteeism data with the technique of the multilevel logistic regression model (MLRM) on non-aggregated absenteeism data. From a second and third empirical application using absenteeism data from a sample of 930 and 11790 employees respectively, it could be argued that the MLRM has several advantages for the analysis of non-aggregated absenteeism data compared to survival analysis and frailty modeling. Furthermore, we have been able to demonstrate that the relationship of well-studied variables like age and gender with absenteeism are not as straightforward as usually thought. For example, it could be shown that this relationship is dependent on the employee’s absence status of the previous day: e.g. while researchers consistently find that women are more often absent than men, this relationship is inversed when an absence spell is ongoing. The relevance of such findings for further theorizing about employee absenteeism as well as for practical and managerial applications are numerous.

Curriculum Vitae

In 1999, Joachim Dejonckheere graduated as a Master of Science. in Clinical Psychology at the Vrije Universiteit Brussel. From September 1999 on, he taught ‘Measurement scales and descriptive statistics’, ‘Probability calculation and statistical analysis’ and ‘Multivariate analysis’ at the faculty of Psychology and Educational Sciences of the Vrije Universiteit Brussel as an assistant to professor Mark Despontin.

From May 2003 till May 2004, he worked as a statistician at De Witte & Morel, Compensations & Benefits where his responsibilities encompassed trend analysis of wages in Belgium and building statistical models for the estimation of salary components.

From May 2004 on, he is employed as a biostatistician at SGS Life Science Services where he is involved in randomisation, sample size calculation, the designing, statistical analysis and interpretation of clinical trials.

From his graduation in 1999 on, he attended several statistical courses (e.g. on Mixed models, Survival analysis) and a number of statistical congresses (e.g. Conference on complex statistical modelling in Diepenbeek, Conference on Multilevel analysis in Amsterdam).