



## Special issue dedicated to Odd O. Aalen

Ørnulf Borgan<sup>1</sup> · Håkon K. Gjessing<sup>2,3</sup>

Received: 17 July 2019 / Accepted: 21 August 2019 / Published online: 28 August 2019  
© Springer Science+Business Media, LLC, part of Springer Nature 2019

Odd O. Aalen celebrated his 70th birthday May 6th, 2017. In October the same year, a two-day symposium took place at the University of Oslo to celebrate Odd's many landmark contributions within survival and event history analysis, counting process theory, frailty modeling and causality. The speakers at the symposium were invited to contribute to this special issue of *Lifetime Data Analysis*, and we would like to thank them for their interesting contributions within areas of biostatistics that have been central to Odd's research. We are also grateful to Editor-in-Chief Mei-Ling Ting Lee for giving us the privilege to serve as Guest Editors for this issue.

Odd has a master's degree in statistics from the University of Oslo (1972) and a PhD degree from University of California, Berkeley (1975). After the completion of his PhD, Odd was invited by his master's thesis supervisor, Jan M. Hoem, to visit the University of Copenhagen. He spent eight months there as a visiting scholar, and his work immediately caught the attention of Niels Keiding and Søren Johansen, among others, thereby making an important contribution to the strong research tradition in survival and event history analysis at the University of Copenhagen. After completion of the compulsory military service, Odd took a position as associate professor at the newly established University of Tromsø, the northernmost university in the world. Odd spent five years in Tromsø to help establish statistics as a subject at the new university before he in 1982 came back to the University of Oslo, now as professor in statistics—and the only statistician—at the Faculty of Medicine. Odd is known as an excellent teacher with the ability to convey complex ideas in a straightforward manner, and he has written a statistics textbook in Norwegian for the specific use in courses for medical students. Over the years Odd and his colleagues have succeeded in establishing a strong research group in biostatistics at the Faculty of Medicine,

---

✉ Ørnulf Borgan  
borgan@math.uio.no

Håkon K. Gjessing  
hakon.gjessing@uib.no

<sup>1</sup> Department of Mathematics, University of Oslo, Oslo, Norway

<sup>2</sup> Norwegian Institute of Public Health, Oslo, Norway

<sup>3</sup> Department for Global Health and Primary Care, University of Bergen, Bergen, Norway

and today the Oslo Centre for Biostatistics and Epidemiology is a thriving research environment comprising more than 50 researchers.

Odd has for a period of more than 40 years contributed extensively to the development of methods for analyzing time-to-event data, and since 2003 he has been one of the editors of *Lifetime Data Analysis*. It is not possible here to give a full account of all of Odd's important contributions to biostatistics, but we will try to give an overview of his main areas of research and to highlight some of his key papers.

If there is a single topic that can be said to be crucial to almost all of Odd's scientific work, it must be a general and deep understanding of the concept of stochastic processes. Odd has always been interested in applying a general theory of stochastic processes to applied problems; by understanding concepts such as martingales, Doob-Meyer decompositions, Markov processes, local dependence, process barrier hitting times, etc., he demonstrates that these concepts can be given immediate and useful interpretations in many applied settings. A prime example of this is Odd's seminal *Annals of Statistics* paper from 1978 (Aalen 1978), based on his PhD-thesis from Berkeley, where he showed the usefulness of counting processes and martingales in lifetime data analysis and gave an in-depth study of what is now known as the Nelson-Aalen estimator. Soon after, in a joint work with Søren Johansen, Odd again used the counting process approach, combined with results on matrix valued product-integrals, to derive the Aalen-Johansen estimator for the matrix of transition probabilities of multi-state Markov models (Aalen and Johansen 1978). He then followed up by introducing the non-parametric additive hazards regression model that today bears his name (Aalen 1980, 1989). Other authors used the counting process approach to, for example, study non-parametric  $k$ -sample tests (Andersen et al. 1982) and Cox's regression model (Andersen and Gill 1982). Today, the counting process approach is an integrated part of survival and event history analysis, as presented in the books by Fleming and Harrington (1991) and Andersen et al. (1993). The events leading up to Odd's 1978 *Annals of Statistics* paper, and the subsequent development of the use of counting processes and martingale methods for analyzing lifetime data are described in Aalen et al. (2009). In the Springer series on *Breakthroughs in Statistics*, the paper was selected as one of the most influential papers in statistics in the twentieth century (McKeague 1997).

The traditional methods in survival and event history analysis ignore the heterogeneity between individuals that is not captured by observed covariates. Around 1980, the effect of unobserved heterogeneity for lifetime data was discussed in the demographic literature under the name of "frailty" (Vaupel et al. 1979), and frailty models soon found their way into the statistical literature. In particular, Philip Hougaard derived important results on possible distributions for the frailty (Hougaard 1984, 1986). Inspired by this work, Odd started to study frailty models, focusing on the impact that frailty may have on the shape of the population hazard (Aalen 1987b, 1988), and he used frailty models, e.g., to help the understanding of mechanisms that may cause testicular cancer (Aalen and Tretli 1999; Moger et al. 2004). Again with a stochastic process perspective, Odd explored how frailty itself could be modeled using an underlying "stochastic disease process." By assuming, for instance, that an event was triggered at the boundary hitting time of a diffusion process with random drift, he demonstrated how properties of the underlying process are reflected in the shape of

the observed population hazard (Aalen 1995; Aalen and Gjessing 2001), and that similar effects appear when the hazard itself is modeled as a process (Gjessing et al. 2003; Aalen and Gjessing 2004). A general observation was that many such population hazard functions tend to stabilize over time, as a consequence of the underlying frailty process converging to its quasi-stationary distribution.

During a sabbatical at the MRC Biostatistics Unit in Cambridge in the 1990s, Odd became involved in work that monitored the development of the AIDS epidemic in England and Wales. Backcalculation is widely used to estimate HIV incidence rates, and in collaboration with researchers at the Biostatistics Unit, Odd showed how the method could be improved through the use of a Markov model that describes the progress of an HIV infected person through various stages and by incorporating available knowledge of the times of first positive HIV test (Aalen et al. 1997). The method has been put into practical use, and has later been extended and improved (e.g. Sweating et al. 2005; Brizzi et al. 2019, in this special issue).

From the start of his research career, Odd has been interested in understanding how statistics may help to throw light on the causes and course of a disease. An early work uses the concept of local independence for Markov chains (Schweder 1970), which is closely related to Granger causality (Granger 1969), to investigate if the risk of obtaining a certain skin disease for women increases after menopause (Aalen et al. 1980). This is a simple example of dynamic modeling where the occurrence of one event (menopause) changes the risk that another event (skin disease) may occur. Dynamic modeling and local independence was further investigated by Odd for general stochastic processes using the Doob-Meyer decomposition (Aalen 1987a); see Didelez (2008) for an in-depth study of this approach. Dynamic modeling is also crucial for dynamic path analysis (Fosen et al. 2006), where the aim is to understand the effect of an internal (endogenous) time-dependent covariate. An extensive discussion of the dynamic viewpoint is given in Aalen et al. (2012). Local independence can be formulated causally (Didelez 2019, in this special issue), and dynamic path analysis has now been formulated in a causal setting using Pearl's do-operator (Aalen et al. 2019).

Many of Odd's research interests are reflected in his book co-authored by us (Aalen et al. 2008); in keeping with his focus, its subtitle "a process point of view" indeed seemed unavoidable.

This special issue contains contributions on a number of topics within survival and event history analysis that have been central to Odd's research. With a focus on causality, the paper by Vanessa Didelez: *Defining causal mediation with a longitudinal mediator and a survival outcome*, asks causality questions in the context of mediation processes with a time-to-event outcome. The analytical approach is based on graphical models that allow decomposition of treatment into different pathways, representing different mechanisms for disease, for instance. Pål Christie Ryalen, Mats Julius Stensrud and Kjetil Røysland: *The additive hazard estimator is consistent for continuous time marginal structural models*, study continuous-time marginal structural modeling, where the application of treatment weights is used to obtain causal interpretations of results. They demonstrate the important role of the additive hazard estimator for estimating treatment weights in continuous time, and prove consistency of the model under the right conditions. Ditte Nørbo Sørensen, Torben Martinussen and

Eric Tchetgen Tchetgen: *A causal proportional hazards estimator under homogeneous or heterogeneous selection in an IV setting*, propose a novel estimator for a causal proportional hazards. The method uses an instrumental variable (IV) and describes unobserved confounding using a selection bias function. Similarities between the observed hazard function and a Cox-Aalen model is exploited to develop the estimation procedure and explore its large sample properties.

Multi-state models provide a flexible extension of the standard survival analysis model, by allowing different transition intensities between the various states. For Markovian multi-state models, the Aalen-Johansen estimator (extended to allow for fixed covariates) may be used to estimate (covariate dependent) transition probabilities, but the estimator is not valid for non-Markov models. However, Rune Hoff, Hein Putter, Ingrid Sivesind Mehlum and Jon Michael Gran: *Landmark estimation of transition probabilities in non-Markov multi-state models with covariates*, show how one may use a landmark version of the Aalen-Johansen estimator to estimate transition probabilities also for non-Markov models. The idea is to base the inference on a subsample of the full data, consisting of the individuals who are in a specific state at a given landmark time-point. Per Kragh Andersen, Jules Angst and Henrik Ravn: *Modeling marginal features in studies of recurrent events in the presence of a terminal event*, combine recurrent states with “not-at-risk” states as well as a terminal endpoint state. Instead of explicit modeling of all transition intensities, they study Nelson-Aalen-like estimators for the marginal means of the number of events. Vern Farewell, Li Su and Christopher Jackson: *Partially hidden multi-state modelling of a prolonged disease state defined by a composite outcome*, use a partially hidden multi-state Markov model with latent states to describe the development of a rheumatic disease. The approach makes it possible to use data from all clinical visits, also those with missing data on some variables.

In the setting of unobserved heterogeneity, Niels Keiding, Katrine Lykke Albertsen, Helene Rytgaard and Anne Lyngholm Sørensen: *Prevalent cohort studies and unobserved heterogeneity*, explore selection effects. In ordinary prospective studies, unobserved covariates are often accounted for by the inclusion of a frailty variable in hazard rate models. In the case of prevalent cohort studies, however, there is an additional effect of unobserved heterogeneity in that different individuals have different probabilities of being selected into the study. The paper surveys available methods to counter this issue. Robin Henderson, Ralitsa Mihaylova and Paul Oman: *A dual frailty model for lifetime analysis in maritime transportation*, propose a model that may be used to study the change of ownership of commercial shipping vessels. Unexplained heterogeneity in the data is captured by a dual frailty model with one frailty term for the seller and one for the buyer, and estimation is performed using Markov chain Monte Carlo.

The final paper of this special issue concerns modeling of the AIDS epidemic. Based on the work of Aalen and others, Francesco Brizzi, Paul J. Birrell, Martyn T. Plummer, Peter Kirwan, Alison E. Brown, Valerie C. Delpech, Owen Noel Gill and Daniela De Angelis: *Extending Bayesian back-calculation to estimate age and time specific HIV incidence*, extend multistate CD4-count models to enable estimation of age- and time-dependent HIV incidence. Using bivariate splines and tensor product splines, the authors obtain smoothing over both age and time.

## References

- Aalen OO (1978) Nonparametric inference for a family of counting processes. *Ann Stat* 6:701–726
- Aalen OO (1980) A model for non-parametric regression analysis of life times. In: Klonecki W, Kozek A, Rosinski J (eds) *Mathematical statistics and probability theory, Lecture Notes in Statistics*, vol 2. Springer, New York, pp 1–25
- Aalen OO (1987a) Dynamic modelling and causality. *Scand Actuar J* 1987:177–190
- Aalen OO (1987b) Two examples of modelling heterogeneity in survival analysis. *Scand J Stat* 14:19–25
- Aalen OO (1988) Heterogeneity in survival analysis. *Stat Med* 7:1121–1137
- Aalen OO (1989) A linear regression model for the analysis of life times. *Stat Med* 8:907–925
- Aalen OO (1995) Phase type distributions in survival analysis. *Scand J Stat* 22:447–463
- Aalen OO, Gjessing HK (2001) Understanding the shape of the hazard rate: a process point of view (with discussion). *Stat Sci* 16:1–22
- Aalen OO, Gjessing HK (2004) Survival models based on the Ornstein–Uhlenbeck process. *Lifetime Data Anal* 10:407–423
- Aalen OO, Johansen S (1978) An empirical transition matrix for nonhomogeneous Markov chains based on censored observations. *Scand J Stat* 5:141–150
- Aalen OO, Tretli S (1999) Analyzing incidence of testis cancer by means of a frailty model. *Cancer Causes Control* 10:285–292
- Aalen OO, Borgan Ø, Keiding N, Thormann J (1980) Interaction between life history events: nonparametric analysis of prospective and retrospective data in the presence of censoring. *Scand J Stat* 7:161–171
- Aalen OO, Farewell VT, De Angelis D, Day NE, Gill ON (1997) A Markov model for HIV disease progression including the effect of HIV diagnosis and treatment: application to AIDS prediction in England and Wales. *Stat Med* 16:2191–2210
- Aalen OO, Borgan Ø, Gjessing HK (2008) *Survival and event history analysis. A process point of view*. Springer, New York
- Aalen OO, Andersen PK, Borgan Ø, Gill RD, Keiding N (2009) History of applications of martingales in survival analysis. *Journal Électronique d’Histoire des Probabilités et de la Statistique* 5:1–28
- Aalen OO, Røysland K, Gran JM, Ledergerber B (2012) Causality, mediation and time: a dynamic viewpoint. *J R Stat Soc Ser A* 175:831–861
- Aalen OO, Stensrud MJ, Didelez V, Daniel R, Røysland K, Strohmaier S (2019) Time-dependent mediators in survival analysis: modeling direct and indirect effects with the additive hazards model. *Biometrical J*. <https://doi.org/10.1002/bimj.201800263>
- Andersen PK, Gill RD (1982) Cox’s regression model for counting processes: a large sample study. *Ann Stat* 10:1100–1120
- Andersen PK, Borgan Ø, Gill RD, Keiding N (1982) Linear non-parametric tests for comparison of counting processes, with application to censored survival data (with discussion). *Int Stat Rev* 50:219–258 Amendment 52:225
- Andersen PK, Borgan Ø, Gill RD, Keiding N (1993) *Statistical models based on counting processes*. Springer, New York
- Brizzi F, Birrell PJ, Plummer MT, Kirwan P, Brown AE, Delphech VC, Gill ON, De Angelis D (2019) Extending Bayesian back-calculation to estimate age and time specific HIV incidence. *Lifetime Data Anal*. <https://doi.org/10.1007/s10985-019-09465-1>
- Didelez V (2008) Graphical models for marked point processes based on local independence. *J R Stat Soc Ser B* 70:245–264
- Didelez V (2019) Defining causal mediation with a longitudinal mediator and a survival outcome. *Lifetime Data Anal*. <https://doi.org/10.1007/s10985-018-9449-0>
- Fleming TR, Harrington DP (1991) *Counting processes and survival analysis*. Wiley, New York
- Fosen J, Ferkingstad E, Borgan Ø, Aalen OO (2006) Dynamic path analysis—a new approach to analyzing time-dependent covariates. *Lifetime Data Anal* 12:143–167
- Gjessing HK, Aalen OO, Hjort NL (2003) Frailty models based on Lévy processes. *Adv Appl Probab* 35:532–550
- Granger CWJ (1969) Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 37:424–438
- Hougaard P (1984) Life table methods for heterogeneous populations: distributions describing the heterogeneity. *Biometrika* 71:75–83

- Hougaard P (1986) Survival models for heterogeneous populations derived from stable distributions. *Biometrika* 73:387–396 (Correction 5:395)
- McKeague IW (1997) Introduction to Aalen (1978): Nonparametric inference for a family of counting processes. In: Kotz S, Johnson NL (eds) *Breakthroughs in statistics*, vol III. Springer, New York, pp 347–353
- Moger TA, Aalen OO, Heimdal K, Gjessing HK (2004) Analysis of testicular cancer data using a frailty model with familial dependence. *Stat Med* 23:617–632
- Schweder T (1970) Composable Markov processes. *J Appl Probab* 7:400–410
- Sweeting MJ, De Angelis D, Aalen OO (2005) Bayesian back-calculation using a multi-state model with application to HIV. *Stat Med* 24:3991–4007
- Vaupel JW, Manton KG, Stallard E (1979) The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography* 16:439–454

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.