Preface

This book is about statistical inference using multi-state survival models. The aim is to introduce and explain methods to describe stochastic processes that consist of transitions between states over time. The book is targeted at applications in medical statistics, epidemiology, demography, and social statistics. An example of an application is a three-state process for dementia and survival in the older population. Such a process can be described by an illness-death model, where state 1 is the dementia-free state, state 2 is the dementia state, and state 3 is the dead state. Statistical analysis can investigate potential associations between the risk of moving to the next state and variables such as age, gender, or education. Statistical analysis can also be used to predict the multi-state survival process. A prediction of specific interest is residual life expectancy; that is, prediction of the expected number of years of life remaining at a given age. When the model describes an illness-death process in the older population, total residual life expectancy at a given age can be subdivided into healthy life expectancy and life expectancy in ill-health.

Applications in this book concern longitudinal data. Typically the data are subject to interval censoring in the sense that some of the transition times are not observed but are known to lie within a given time interval. For example, the time of onset of dementia is latent but when longitudinal data are available the onset may be known to lie in the time interval defined by two successive observations.

Methodologically, multi-state modelling is an elegant combination of statistical inference and the theory of stochastic processes. With this book, I aim to show that the statistical modelling is versatile and allows for a wide range of applications. The computation that is involved in fitting the models can be considerable, but in many cases existing software can be utilised.

I hope that the book will be of interest to diverse groups of readers. Firstly, the book introduces multi-state survival modelling for researchers who are new to the subject. After the introductory first chapter, the book discusses the multi-state survival model as an extension of the (two-state) standard survival model. Further topics are discrete-time models versus continuous-time models, theory on continuous-time Markov chains, parametric models for transition-specific baseline hazards, maximum likelihood inference, model validation, and prediction. The appendix includes code for the R software.

Secondly, for readers with subject-matter knowledge, the book can serve as a reference and it also offers advanced topics such as the modelling of time-dependent hazards, a general scoring algorithm for maximum likeli-

hood inference, methods for Bayesian inference, estimation of state-specific life expectancies, semi-parametric models, and specification and estimation of frailty models.

For the software used in the data analysis, some details are provided in Appendix C. Software is also available on the author's website, `www.ucl.ac.uk/~ucakadl/Book`.

The book assumes knowledge of mathematical statistics at third-year undergraduate or at MSc level. This knowledge can be based on courses in statistics or applied mathematics, but also on courses in other disciplines with a strong statistical programme.