Personalised medicine for multiple outcomes : methods and application

Rüten-Budde, A.J.

CHAPTER 9

Discussion and future perspectives

This thesis aimed at developing clinically relevant survival models, in particular the development and validation of prediction models for use in clinical practice. In the medical field statistical models need to address the research questions asked by clinicians and take into account the particular data structure of the problem at hand. Models used for prediction need to additionally be validated before they can be used in clinical practice. This final chapter, summarizes and discusses previous chapters. They are put in broader perspective and future research directions are suggested.

§9.1 Custom-made models

Statistical research is often motivated by a practical problem. A statistician is assigned to answer an important research question with an often already collected data set. Many times standard techniques can be applied to analize the data. However, frequently the data does not perfectly meet the criteria for standard techniques or the research question suggests to use more advanced techniques. In some cases, existing methods cannot properly address the problem at hand and new methodology needs to be developed.

Survival methodology for example, was developed to analyze time to event data. This type of data is incomplete because not for all subjects the event of interest can be observed, referred to as right-censoring. This is the most common type of censoring found in survival data, however there are many more types of missing information that occur in practice. Methods have been developed to deal with different kinds of missing information.

Some survival methods such as the Kaplan-Meier estimator, the log-rank test and the Cox model are well established in clinical literature [60]. These methods are often adequate to answer clinically relevant research questions and clinicians can understand the output of these methods. Sometimes however, when the research question or the data are complex other statistical methods need to be used. Even though a variety of statistical models have been developed by statisticians there seems to be a disconnect between available methods and methods used for clinical research [118, 159, 132]. Several reasons for this exist. First, clinical researchers may not be aware of the best method to answer the research question. Second, a lack of understanding of complex methods and the output leads to them being undesirable. Third, a lack of available software makes methods difficult to apply.

Regardless the reason for the discrepancy between available statistical methods and methods applied in clinical research a close collaboration between statistician and clinician can contribute to the solution. Additionally, statisticians can benefit from such a collaboration twofold. On the one hand, statisticians can demonstrate the application of their methods and popularize them while at the same time benefiting clinical research. On the other hand, inspiration for new statistical methods may come from interesting clinical research questions. The interdisciplinary collaboration between statistician and clinician is of tremendous importance and may deliver great contributions to both fields [60, 132].

An example of statistical models that are under used in clinical practice are competing risks models [93, 156, 24]. In a frail patient population in which a non terminal event of interest is studied the competing risk of death may prevent the event of interest to be observed. Ignoring the risk of death will lead to wrong conclusions about the incidence of the event of interest and the effect of risk factors on the event of interest. A correct analysis of competing risks is therefore of great importance. In Chapter 2 we proposed a new competing risks model for two competing events that is able to quantify hospital heterogeneity using correlated frailties. The cause-specific proportional hazards model is used to model the risk of the two competing events and additionally frailties on the hospital level are modelled with gamma distribution. The model was developed for multi-center data which is data that was collected from multiple treatment centers. This is often necessary to obtain sufficient data for rare disease. The model adjusts for competing risks while at the same time quantifying the difference in risk between hospitals using Empirical Bayes methodology. It is a delicate task to compare the performance of hospitals and Empirical Bayes methodology is used to take the larger variances of smaller hospitals into account. In Chapter 5 Fine and Gray's model for competing risks was used to predict the cumulative incidence of local recurrence in soft tissue sarcoma patients. Local recurrence is an adverse event for patients who were surgically treated to remove the primary tumor. It means that tumor growth was found at site of surgery and its occurrence may mean additional surgery for some patients. It is therefore an important event for patients and clinicians and the predicted probability of local recurrence is an important information that can be used in patient care.

Multi-state models naturally extend competing risks models. The evolution of disease can sometimes be described by a series of events such as relapse, recovery and death that a patient may experience. These can be directly modelled with a multi-state model. Even though multi-state models can describe disease progression close to reality they are seldom applied in clinical practice. Apart from being more difficult to apply they are also still not well known in the clinical world. If applied they need to be carefully explained so that parameters are interpreted properly. Multi-state models are however an important tool that can give a deeper insight into the association of risk factors and the different states of disease which should be exploited in clinical research. In Chapter 4 and 8 multi-state models were used to understand the disease process for soft tissue sarcoma patients and Ewing sarcoma patients, respectively. In Chapter 4 a multi-state model to investigate the association of risk factors and adverse disease events for soft tissue sarcoma patients was proposed. Several prognostic factors such

as histology, grade, depth and size [83, 117, 169, 53, 73, 146, 140, 98, 139, 29] were already identified. An increase in risk for local recurrence following an intralesional margin resection was recognised [73, 139, 91], the association between margin status and survival and between local recurrence and survival however was still unclear [140, 106, 164, 29, 23, 74, 110, 102, 111]. After surgery a patient is alive with no evidence of disease and may then have local recurrence, distant metastasis or die. We proposed a three state multi-state model to obtain more insight into how certain prognostic factors affect phases of the evolution of disease with particular interest in surgical margin and radiotherapy. In our analysis distant metastasis and death were modelled as a single state, leading to three states (starting state, local recurrence, distant metastasis/death). The multi-state model showed that wide surgical margins and the use of (neo)adjuvant radiotherapy decreased the risk of local recurrence but had little effect on the risk of distant metastasis/death without local recurrence. This study contributed to a better understanding of the effect of risk factors on the different states of disease progression. In Chapter 8 we developed a multi-state model for Ewing sarcoma patients that were treated surgically to gain insight into the effect of surgical margin, histological response and radiotherapy on disease progression. Other studies had conducted single end point analyses of risk factors for adverse events in Ewing sarcoma. Multi-state models can estimate the effect of risk factors on different disease states simultaneously. Five states of disease progression were considered in the multi-state model. After surgery a patient enters the starting state, from here he may develop local recurrence, distant metastasis in the lungs, distant metastasis at a different site and he may die. The data only contained information of the first adverse event a patient experienced, so transitions from adverse events to other non terminal events were not possible. It was found that disease extent at start of treatment and histological response had a strong association for the transitions to distant metastasis and death. For the transition to local recurrence the location of the tumor and surgical margin were important risk factors.

Competing risks and multi-state models are natural model choices when the evolution of disease can be represented by multiple events. Another reason to apply more complex models lies in the data collection process. Methods for survival analysis were developed to deal with the common issue of right-censoring and other forms of missing information that can be present in survival data. Sometimes the event of interest cannot be observed exactly because it may only be diagnosed at pre-specified follow-up visits by for example a blood test. If the test is positive then it is only known that the event had occurred prior to the test. If regular tests are conducted then it is known that the event occurred between the time of the last negative and the first positive test. This type of censoring is referred to as interval-censoring. Methods for interval-censored data have been developed but they are not frequently used in clinical practice. We studied the effect that ignoring interval-censoring has on the predictive accuracy of a binary time-dependent marker in Chapter 3. The binary time-dependent marker may represent the presence or absence of disease that can be acquired over time. Prediction models for survival outcomes can inform patient and clinician and give an indication of a patients prognosis. Dynamic prediction models are able to incorporate updated information from time-dependent variables into the predictions

as they become available. The predictive accuracy of such time-dependent marker is therefore a parameter of interest. We conducted a simulation study to investigate the predictive accuracy of a binary time-dependent marker for the outcome death in the presence of interval-censoring. The marker may represent the presence or absence of disease that a patient can acquire over time. We studied several data scenarios and compared four different models, one of them ignoring interval-censoring. Different interval lengths between the observation of the time-dependent marker were studied. We used Area Under the Curve (AUC) measures that were adapted to survival outcomes to quantify the predictive accuracy of the time-dependent marker and found that the spacing between observation times did have a large effect on the AUC. The results of this study suggest to take interval-censoring into account when evaluating the predictive accuracy of a time-dependent marker.

The choice of statistical method can be motivated from the research question and the collected data. There are however choices to be made like for example regarding the selection of variables to be included in the model. Throughout this thesis clinical researchers chose the variables that were included in the models. In Chapter 4 and 8 multi-state models were estimated and not all covariate effects were estimated for all transitions. The reason for this was the limited number of patients that made some transitions. The selection of covariates for these transitions was done by clinical experts. Categorization of covariates and dichotomization of continuous covariates have been motivated clinically as well and these have changed throughout the time of this research. In Chapter 4 the variable age has been modelled in three categories while in Chapter 5, 6, and 7 age has been modelled continuously. The choice to include a quadratic effect of age into the models of Chapter 6, and 7 have been based on the significant nonlinearity of this variable. For the dynamic prediction model of Chapter 6, and 7 the covariates have been chosen by clinical experts, the interactions with time have been chosen using a backward selection procedure described in Chapter 6.

§9.2 Prediction models in clinical practice

Statistical models are able to make predictions as well as provide interpretable parameters, which contribute to the understanding of the underlying event process. The main goal of prediction models is to make good predictions which can be verified by validating predictions using a data set that has not been used in the model building process. In Chapter 5, 6 and 7 we developed and validated prediction models that predict the probability of local recurrence and survival for soft tissue sarcoma patients. In the first two chapters the models were validated internally using cross validation and in the last chapter the prediction model was validated externally using an external data set.

The parameters of the prediction models can be interpreted clinically and contribute to the understanding of the underlying disease process. The variables used in the models were chosen from clinical experts. The baseline covariates included in the models were of two different kinds: (1) they included patient- and disease-specific covariates, such as age, histology subtype, and tumor size. (2) they included treatment related covariates, such as surgical margin and radiotherapy treatment. The

first type of covariates are given and cannot be influenced. The treatment related covariates however, can be influenced by the treating clinician. In Chapter 4 and 5 we found that a wider surgical margin and radiotherapy treatment were associated with a decreased probability of local recurrence compared to an intralesional margin and no radiotherapy treatment. In Chapter 6, 7 we found that a wider surgical margin was associated to longer survival for a prediction time point just after surgery, but as the patient survived a period of time after surgery, no association with survival was found. Because of these associations the predictions of survival probability for a wider margin and radiotherapy treatment are higher than for an intralesional margin and no radiotherapy treatment. However these predictions do not make a fair comparison between treatment options because the prediction models developed in this thesis depend on current treatment practice. The data used for model development are not randomized controlled trials. For this reason they cannot be used to base treatment decisions on.

§9.3 Future perspectives

In this thesis clinically relevant survival models have been developed. In Chapter 3, 4 and 8 we used multi-state methodology to investigate risk factors for soft tissue sarcoma and Ewing sarcoma. The hazard ratios estimated with these models illustrate the association between covariates and disease related events and are very informative for clinicians. Because of the large number of parameters and insufficient transitions between states not all possible transitions between disease related events could be considered in the multi-state models. In Chapter 4 the event distant metastasis and death were combined because of the small number of patients transitioning from local recurrence to these states. Additionally, because of the small number of patients making some transitions it was not possible to estimate the effect of all relevant covariates for each transition. The soft tissue sarcoma data set has grown since the research conducted in Chapter 4 and now would allow for a more sophisticated multi-state model, one with more states and transitions and more covariates per transition. At the moment a project on the construction of a more complex multi-state model for soft tissue sarcoma patients is ongoing. In Chapter 8 a multi-state model for Ewing sarcoma was estimated based on data from the GPOH registry (Gesellschaft für Pädiatrische Onkologie und Hämatologie) treated in or according to the EURO-E.W.I.N.G 99 (EE99) protocol [8]. This data comprises only half of the total available data. At the moment researchers at Leiden University are working on receiving the second half of this data to conduct a more sophisticated analysis.

In Chapter 5, 6 and 7 we developed prediction models for patients with soft tissue sarcoma. We started with two models to be used at baseline just after surgery which were implemented in the PERSARC mobile application. They predict the probability of overall survival and local recurrence at 3, 5, and 10 years post surgery. In Chapter 6 a dynamic prediction model for soft tissue sarcoma patients was developed to predict the probability of surviving an additional 5 years from a prediction time point during follow-up. This dynamic model was updated and externally validated in Chapter 7 and is in the process of being implemented in the PERSARC mobile application.

Implementation of prediction tools into clinical practice remains challenging despite their utility. The adoption of prediction models to support shared decision making in clinical practice is a current subject of interest. A group of researchers from the Leiden University Medical Center were granted funding to implement shared decision making in treatment decisions in high-grade soft tissue sarcoma of the extremities in the Netherlands. The goal is to ensure that soft tissue sarcoma patients receive personalised care, in which risks and benefits of treatment options and patient preferences are balanced. Part of the implementation strategy is the introduction of the PERSARC mobile application to clinical practice through educational outreach. This type of research may lead to improved prediction tools and facilitate their introduction to clinical practice.

Finally, we stress the importance of future interdisciplinary collaboration between statisticians and clinicians. Multi-state models and dynamic prediction models are important tools that are underused in the clinical field. Other complex available methods should be introduced and explained carefully by statisticians to the clinical community so that they are able to benefit from their advantages.