Clinical decision support : distance-based, and subgroup-discovery methods in intensive care
Nannings, B.

Citation for published version (APA):
Clinical Decision Support Systems (CDSSs) are likely to play a major role in future healthcare provision. Physicians are expected to provide healthcare on the basis of the latest medical knowledge available. Moreover they have to cope with ever-increasing amounts of patient data. CDSSs can help medical professionals by providing them with targeted knowledge, relevant to the problem at hand, and may help physicians to discover important patterns or values from a mound of data that they very unlikely would discover themselves. This thesis has two parts addressing two forms of decision support: support based on discovery of “interesting” subgroups, and support embedded in a telemedicine system.

Specifically, the main focus of the first part of this thesis is the Patient Rule Induction Method (PRIM), which is a subgroup discovery algorithm, and its application in the Intensive Care Unit (ICU). PRIM can be used to discover subgroups of patients or observations that deviate markedly from the rest. The discovery of such subgroups is meant to support health care professionals and managers to improve the provision of care. For example, in the ICU the discovered subgroups can help refine blood glucose regulation guidelines, or adapt the policy for intensifying or withholding therapy.

PRIM was introduced by Friedman and Fisher and is often referred to as a “bump-hunting” algorithm. Bump hunting algorithms attempt to find regions in the input space that are associated with a high (hence the term “bump”) or low mean outcome value relative to the average value of the outcome in the whole sample. PRIM describes regions based on conjunctive conditions on input variables, e.g. “body temperature > 80 AND patient has diabetes”. An important attribute of PRIM is that it is “patient”, contrasting it with more greedy algorithms such as the widely known Classification And Regression Tree (CART) algorithm. In addition, PRIM is non-parametric, unlike logistic regression models commonly used in medicine, such as the popular Simplified Acute Physiology Score (SAPS) model in the ICU. The applicability of PRIM in medicine, and its merits relative to CART and logistic regression models like SAPS are not well understood. The first part of this thesis addresses the applicability of PRIM and the comparison of PRIM with CART and SAPS.

In Chapter 2 we apply PRIM with the aim of discovering subgroups of very elderly patients in the ICU that have a high risk of mortality. There are several reasons for seeking such subgroups. First, these subgroups may provide insight into underlying causes of mortality that may potentially be timely acted upon to increase the probability of survival. Second, high mortality subgroups are often needed in research on the efficacy and efficiency of therapeutic interventions. Third, such groups may improve case-mix adjustment to allow for comparisons of quality of care across different intensive care units. Fourth, information about probability of survival is something that patients and their families are interested in to make informed decisions about further treatments. Finally, such subgroups may influence patient admittance policy (for example, if a subgroup has an extremely high probability of death in the ICU after a specific form of surgery one may not only want to decide on whether to continue or withdraw ICU therapy but also to contemplate on the question whether to operate on such patients in the first place).
We sought subgroups on a dataset of 6617 ICU patients of at least 80 years of age that were obtained from ICUs in the Netherlands that participate in the National Intensive Care Evaluation (NICE) initiative. In addition to applying PRIM we also applied a recalibrated SAPS (version II) model. SAPS II is a commonly used method to predict mortality of intensive care patients. We compared the PRIM subgroups to those found by SAPS II. Performance of the subgroups was evaluated on a randomly selected independent test set. The performance of PRIM and SAPS II was comparable but the subgroups obtained by PRIM involved less variables and resulted in much more homogeneous groups. They are therefore likely to be more useful for decision makers.

In Chapter 3 we applied PRIM to find subgroups of ICU patients having a high blood glucose level (BGL). Despite being on Intensive Insulin Therapy (IIT), many patients suffer from hyperglycemia, which is believed to increase the risk of mortality and morbidity. In contrast to the application concerning mortality of very old patients, the input data in the hyperglycemia application is time-ordered (for example, body temperature is repeatedly measured over time) and the outcome (BGL) is continuous instead of binary. Hyperglycemia in the ICU is generally caused by a disrupted homeostasis as a result of injury or surgery. To provide treatment suggestions, most blood glucose management guidelines rely on the last measured glucose value, and sometimes on a measure describing the trend in previous glucose values and nutritional feed rates, disregarding most other available clinical data. The aim of this study was to discover subgroups of measurements having high blood glucose, and, based on these subgroups, discover potential determinants of hyperglycemia at the ICU. Further research of these potential determinants may lead to improvement of the guidelines, and in turn to a reduced mortality and morbidity.

Data for this study were physiological measurements collected in an 18 bed mixed general-surgical intensive care unit of a teaching hospital. For each patient multiple measurements over time for various variables were available. We included only measurements within the first 24 hours, as normoglycemia (normal glucose level) should be achieved within this period while hyperglycemia was found to be still prevalent.

Prior to applying PRIM we investigated the literature for known determinants of hyperglycemia. PRIM was able to find several subgroups of high glucose measurements which were validated with the independent test set. Aside from well known determinants (e.g. the previous glucose value obtained from the previous measurement) we also found additional candidate determinants of which their relation to blood glucose is less clear. More research is needed to determine whether these potential determinants may help to improve blood glucose management guidelines.

In Chapter 4 we compared PRIM to the Classification And Regression Trees (CART) algorithm using a large high dimensional real-world clinical dataset and searched for circumstances in which the PRIM algorithm is at a disadvantage. We used a multi-center dataset consisting of 41183 records of intensive care patients with 86 input variables and mortality (survival or non-survival) as the outcome variable.
Because there are factors that hinder the direct comparison of PRIM and CART we followed an extensive analysis strategy consisting of 10 different comparison scenarios. The algorithms were compared using the performance measures odds ratios and coverage. We used bootstrapping (with Laplace smoothing) to obtain estimates and confidence intervals.

In most cases CART significantly outperformed PRIM. Further analysis revealed that PRIM’s inferiority could be attributed to its failure to find a large contiguous subgroup that was found by CART at once. More specifically PRIM has trouble “peeling” observations of a discrete ordinal variable which had a mode (in its distribution) located at its highest value. Since such variables are ubiquitous in clinical medicine we recommend to incorporate a backtracking mechanism (such as beam-search) in PRIM and let it make use of global information in assessing the utility of peeling a variable.

In the second part of this thesis we investigated CDSSs in a telemedicine context for which we coined the term Decision Support Telemedicine Systems (DSTSs). These systems are likely to become more common in the near future to cater for the need of having medical information available any time and place, and to support medical professionals in keeping up to date with the latest medical knowledge and coping with the large amounts of data that are available to them.

Although much research dedicated to telemedicine and CDSSs separately exist, this is not the case in the area where these two fields intersect. Based on a systematic literature search with a focus on keywords pertaining to telemedicine and CDSSs, we aimed to create a useful conceptualization of DSTSs focusing on those areas that are important for DSTSs.

While studying the literature in search of DSTSs, it became clear that the descriptions of such systems were often incomplete and/or vague, as important properties were not described (e.g. not reporting on the reasoning method pertaining to the CDSS component). In Chapter 5 we proposed a characterizing property set for DSTSs and applied this set to describe a number of DSTSs. The set consists of 14 properties that can be used to describe and cluster DSTSs. The properties are grouped in three categories that we refer to as the problem dimension (medical problem and the environment where the DSTS is used), process dimension (behavior and dynamic aspects) and system dimension (physical system aspects). Properties of the problem dimension are related to e.g. the purpose of a DSTS, what kind of human agents are involved, and what kind of medical task is supported. Properties of the process dimension are related to e.g. whether the process is synchronous or asynchronous. Properties of the system dimension are related to e.g. what type of reasoning method the system uses to support a decision and what type of data it processes. Unexpectedly the literature did not reveal emerging properties that are unique to DSTSs.

In Chapter 6 we proposed a definition for DSTSs. This definition is a combination and harmonization of definitions for telemedicine and CDSSs that we found in the literature.
Additionally, we proposed a general conceptual model of a DSTS and a number of template models for different typical DSTSs. Such models can help stakeholders of a DSTS such as medical professionals, CDSS developers and telemedicine experts to quickly gain insights specific to DSTSs that could be used during the system’s requirements analysis or further development. The models were created using the Unified Modeling Language (UML).

In Chapter 7 we provide a summary of the principle findings of this thesis. The main contribution of this thesis is to provide a better understanding of CDSSs from an application and comparison perspective (based on the PRIM algorithm), and by formulating a conceptual framework for understanding CDSSs in a telemedicine context from a bird’s eye view.