

Patient Length of Stay and Mortality Prediction: a survey

Aya Awad
School of Computing
University of Portsmouth
Portsmouth PO1 3HE, UK
ayallah.awad@port.ac.uk

Mohamed Bader-El-Den
School of Computing
University of Portsmouth
Portsmouth PO1 3HE, UK
Mohamed.bader@port.ac.uk

James McNicholas
Consultant in Critical Care
Queen Alexandra Hospital
Portsmouth Hospitals NHS Trust
james.mcnicholas@porthosp.nhs.uk

Abstract—Over the past few years, there has been increased interest in data mining and machine learning methods to improve hospital performance, in particular hospitals want to improve their intensive care unit (ICU) statistics by reducing the number of patients dying inside the ICU. Research has focused on prediction of measurable outcomes, including risk of complications, mortality and length of hospital stay. Length of stay (LOS) is an important metric both for healthcare providers and patients, influenced by numerous factors. In particular length of stay in critical care is of great significance, both to patient experience and the cost of care, and is influenced by factors specific to the highly complex environment of the ICU. LOS is often used as a surrogate for other outcomes, where those outcomes cannot be measured; for example as a surrogate for hospital or ICU mortality. LOS is also a parameter, which has been used to identify severity of illness and healthcare resource utilisation.

This paper examines a range of length of stay and mortality prediction applications in acute medicine and the critical care unit. It also focuses on the methods of analysing length of stay and mortality prediction. Moreover, the paper provides a classification and evaluation for the analytical methods to length of stay and mortality prediction associated with a grouping of relevant research papers published in the year 1984 till 2016 related to the domain of survival analysis. In addition, the paper highlights some of the gaps and challenges of the domain.

I. INTRODUCTION

Healthcare expenditure constitutes a significant share of the gross domestic product (GDP) of many countries. For example, in 2012 healthcare spending in the UK reached nearly a tenth (9.3%) of GDP [1], [2]. Government funding in many countries has fallen behind patient care costs, leaving healthcare institutions face the growing number of patients [3]. Accordingly, cost containment has become one of the most critical challenges in healthcare today. Hospitalization constitutes the principal cost of patient care and is therefore a main focus in healthcare management [4], [5].

Patient hospital length of stay typically refers to the number of days that an inpatient stays in a healthcare facility during a single admission [6]. It is considered one of the major indicators for the consumption of hospital resources [7], [8]. It also provides a better understanding of the flow of patients through a healthcare system which is essential for evaluating both the operational and clinical functions of such systems. Previous research have attempted to group patients

by their medical condition, assuming that each disease, illness, or procedure is associated with a recommended LOS [9]. Grubinger et al. in 2010 [10] refer to these systems as diagnosis-related-group(DRG) systems. In addition, a relative value, namely Case Mix Index (CMI) can be assigned to a DRG of patients in a medical care environment used in determining the allocation of resources to care for and/or treat the patients in the group. However, both studies assumed that all patients who fall within the same diagnosis-related-group are the same. However, length of stay is a complex metric affected by other factors including each individual's demographics, treatment complexity, complications and discharge planning which may stretch the LOS beyond the target range. A model that helps to predict a patients length of stay during a single visit - the time from hospital admission until discharge - can be an effective tool for health care providers to plan for preventive interventions and to improve the utilization of hospital resources [11]. Moreover, usually caregivers maintain an overall assessment of their patients based on important observations and trends over the first few days of admission. Some research [12], [13] demonstrate a strong correlation between LOS and mortality; however other experts/ intensivists consider LOS highly unregarded as a determinant of mortality as it is subject to influences, which may not bear upon the outcome of real interest. For example, a diagnosis which simply requires a prolonged period of hospital care, but which confers a low risk of hospital mortality, would bias the use of LOS as a surrogate for mortality; a condition which requires a series of complex treatment interventions might prolong the stay, without necessarily conferring a high mortality risk. Also, conversely presentations of high severity of illness scores, as quantified by APACHE II [14], ICNARC score [15], SOFA [16] or MPM [17], might be associated with a short LOS, because of an early decease, but also a high mortality, further undermining the correlation between LOS and mortality. In contrast, the work in [12], [13] linked LOS to mortality. Results showed significantly greater ICU, hospital and long-term mortality in patients with an ICU stay longer than than 3 days, in comparison with those who have a stay of 3 days or less.

This paper reviews LOS and mortality applications in acute

medicine and critical care unit and the correlation among them. Moreover, the paper classifies and evaluates the analytical methods available in the literature over the past three decades. In addition, the paper highlights some of the gaps and challenges of the domain.

The rest of this paper is organised as follows; section II provides a survey of the different applications of LOS in the health care domain. Section III provides a classification of the analytical methods to LOS prediction associated with a grouping of relevant research papers published in the domain. Section IV provides a survey of the different applications in mortality prediction, section V provides a classification for the methods of mortality prediction, section VI provides some examples of applications in previous literature that predicts both LOS and mortality and section VII discusses performance evaluation of the different LOS and mortality prediction models. Finally, the conclusion is drawn in section VIII.

II. APPLICATIONS OF LOS

LOS is often used as a surrogate for other outcomes in research, where those outcomes cannot be measured; for example as a surrogate for hospital mortality or ICU mortality. LOS is also a parameter, which has been used to identify severity of illness and healthcare resource utilisation [18]–[20]. As a surrogate outcome measure it is sometimes not highly regarded, as it is subject to influences, which may not bear upon the outcome of real interest. For example, a diagnosis which simply requires a prolonged period of hospital care, but which confers a low risk of hospital mortality, would bias the use of LOS as a surrogate for mortality; a condition which requires a series of complex treatment interventions might prolong the stay, without necessarily conferring a high mortality risk. Also, conversely presentations of high severity of illness scores, as quantified by APACHE II [14], ICNARC score [15], SOFA [16] or MPM [17], might be associated with a short LOS, because of an early decease, but also a high mortality, further undermining the correlation between LOS and mortality.

Moreover, LOS in hospital may be affected by factors unrelated to the disease, such as the availability of social care or community nursing support. There is an analogous effect upon discharge from critical care to the ward, in the event that there are insufficient ward beds for timely ICU discharge [21], [22]. Finally LOS may also be influenced by characteristics of the organisation including hospital management style [23], [24].

As LOS is an important determinant of both healthcare costs and patient experience, it is a high priority for it to be optimal; therefore it is also significant to identify any factors which affect it. The following two subsections will examine LOS applications in acute medicine and critical care, highlighting factors affecting LOS prediction.

A. LOS in Acute Medicine

This section presents previous studies on the different applications used in modeling LOS and its association with influencing factors with respect to patient flow. Patient flow typically refers to the progressive movement of a patient through a sequence of processes [25]. Reducing delays and making sure that the patient receives the right care at the right time will have a significant beneficial effect on the quality of service. In turn, this will improve patient outcomes and reduce the cost of care.

In 2012, Freitas et al. [26] studied variables associated with high LOS outliers, together with some hospital characteristics (administrative, economic and teaching characteristics). Results show that age, type of admission and hospital type were significantly associated with high LOS outliers. Moreover the study conducted by Caetano et al. [27] showed that the top 3 influential input attributes were the hospital episode type, the physical service where the patient is hospitalized and the associated medical speciality. However, hospital related factors on their own are not sufficient to accurately predict LOS.

An important variable associated with LOS prediction and common in several studies is the nutritional status of a patient prior to admission. Research conducted in [28]–[33], examined the effect of the variable malnutrition on patient LOS. In the study of Robinson et al. in 1987 [28], on average LOS was 15.6 days for a malnourished patient group versus 10 days for the well nourished group. However, in 1997 Chima et al. [29], showed that LOS for the two patient groups was 6 days for the at-risk for malnutrition population and 4 days for the not-at-risk for malnutrition population. The significant decrease in LOS may reflect the fact that all health institutions are under pressures of payment and reviews by government and other third-party payers. In addition, according to Correia et al. [30], length of hospital stay is shorter in the well-nourished patients; median of 6 days versus 9 days for the malnourished. Warnold and Lundhön [32], studied the clinical significance of preoperative nutritional status in 215 non-cancer patients. The variables investigated included weight loss, weight-for-height index, serum protein levels (serum albumin, transferrin, prealbumin, retinol-binding protein), delayed hypersensitivity skin testing, arm circumference, and triceps skinfold thickness. Of the markers evaluated, weight-for-height index, arm muscle circumference, serum albumin level, and weight loss correlated significantly to post-surgery outcome. In addition, Epstein et al. [33] also emphasized that underweight patients have 40% higher LOS than normal weight patients. Also according to Burritt et al. [31], low serum albumin level is the most sensitive single nutrition-related variable in the prediction of complications and length of stay.

Another important variable in a different clinical domain that was also associated with an increase in LOS was Serum Creatinine (SCr). Chertow et al. [34] evaluated the marginal effects of acute kidney injury (AKI) on mortality, LOS, and costs. Changes in serum creatinine (SCr) was used as a determinant for adverse outcomes. Results show that AKI was

consistently associated with an independent increase in LOS. Larger increases in SCr were associated with longer relative increases in hospital LOS.

B. LOS in Critical Care

There are significant potential benefits from quantification and optimisation of LOS in critical care: specifically, these relate to cost containment and clinical quality. The provision of critical care is of necessity expensive, deploying complex interventions and requiring a high intensity of clinician input to a relatively small group of patients. Greater LOS requires more critical care resource and greater cost. As critical care facilities experience increasing pressure and economic resources are more constrained, the priority given to improvements in the timeliness and efficiency of critical care, is rising [35]. Clinical quality in the critical care unit may also be affected by extended LOS. Prolonged LOS gives rise to capacity pressure; this may lead to the cancellation of elective surgery, which is both costly and harmful; it may increase the pressure to decline or delay emergency admission, which could potentially have an adverse effect upon outcome; it may dilute the attention given to the most seriously sick individuals [36].

The critical care unit is also an environment which is well suited to exploiting data for mathematical modelling and prediction, both because of analytical experience and data availability. There are well developed methodologies for performance benchmarking. The increasing use of electronic Clinical Information Systems, means that computer analysis can now be performed directly on the patient record, rather than after specific-to-purpose hand data extraction; the physiological and laboratory data sets are relatively large by comparison with other patient groups.

In England, Wales and Northern Ireland, the benchmarking of critical care unit performance is conducted by the Intensive Care National Audit and Research Centre (ICNARC), by means of its Case Mix Programme (CMP) [15]. The CMP uses rigorous methods to ensure data are complete, valid and reliable [37], [38]; admissions are scored for severity using an in house scoring system and also the APACHE II model, and then a predicted hospital mortality for admissions is calculated. Comparison is made with actual mortality and a Standardised Mortality Ratio is generated quarterly [37], [38]. Another example of a non-commercial database of this kind is that held by the Australia and New Zealand Intensive Care Society, which contains data on over 900,000 ICU stays [39].

Several research groups have investigated LOS in the ICU as it has been felt to be a suitable target for improvement [12], [13]. LOS has been linked to mortality; research in 2006 showed significantly greater ICU, hospital and long-term mortality in patients with an ICU stay longer than than 3 days, in comparison with those who have a stay of 3 days or less. Others have sought to develop models which predict LOS; Buchman et al [38] predicted chronicity in a surgical intensive care unit by classifying patients LOS in accordance with a seven day norm. Levin et al [40] developed a model to produce real-time, updated forecasts of patients intensive care

LOS using naturally generated provider orders. The model was designed to be integrated within a computerised decision support system to improve patient flow management. The study compared the predicted LOS to the actual LOS based on fixed variables, such as age, source of admission and readmission status; temporal variables, such as current LOS, day of the week, time of the day; and order-based predictor variables grouped by medication, ventilation, laboratory, diet, activity, foreign body and extra-corporeal membrane oxygenation.

LOS prediction would help with capacity planning. At present, LOS prediction tools are not used in mainstream critical care practice. Surges in demand are managed reactively, requiring considerable staffing flexibility and variability in the balance between demand and capacity. It is possible that accurate prediction of LOS would help to align these quantities in critical care, and improve resource allocation, in particular staffing resource. According to Celi et al [36], healthcare delivery has worked as well as it has to date because clinicians are bright, hard-working, and well-intentioned, not because systems are well designed nor data systematically harnessed.

It follows that the presence of complete, highly detailed critical care databases is essential, if the potential benefits of modelling and prediction are to be fully realised [12], [41]. Several commercial ICU databases have been developed, archiving patient demographics and aggregating information such as underlying disease, severity of illness, and hospital-specific information such as LOS, mortality and readmission. For example, among the commercial ICU databases is APACHE Outcomes, created at Cerner by merging APACHE [14] with Project IMPACT [42], and includes data from about 150,000 ICU stays since 2010. The commercial Philips eICU, a telemedicine intensive care support provider, archives data from participating ICUs; Philips eICU is estimated to maintain a database of over 1.5 million ICU stays, and is adding 400,000 patient records per year from over 180 subscribing hospitals in the US. More ambitious still is the Multiparameter Intelligent Monitoring in Intensive Care (MIMIC) II database established in October 2003. Developed by an interdisciplinary team from academia (MIT), industry (Philips Medical Systems) and clinical medicine (Beth Israel Deaconess Medical Center), the database incorporates two different types of medical data: clinical data is stored in a relational database and bedside monitoring is stored in flat binary files. There are over 25,000 patients in the MIMIC II relational database, which permits the systematic capture, analysis and integration of information contained within the massive quantity of data generated by each critical care admission. Clearly these kinds of datasets and the several listed earlier could be used to investigate LOS as they provide well-structured high quality data. They could also be exploited for broader research activity.

III. ANALYTICAL METHODS TO LOS PREDICTION

This section explores the methods used in the field of calculating and predicting patient LOS. After surveying the previous literature, LOS prediction methods were categorized into 4 subgroups as shown in Figure I. A classification of the

reviewed papers based on LOS prediction methods is shown in table I.

A. Arithmetic and Statistical Approaches to LOS Prediction

Despite the complex nature of the metric LOS, simple arithmetic methods still exist for the calculation of LOS [87]. Arithmetic methods usually compute the average length of stay or the median as shown. However, this is a very simple way to measure length of stay as it assumes that LOS is normally distributed, typically LOS has an exponential distribution. Also Vasilakis et al. 2003 [48] illustrates how average LOS can be a misleading measure; the research proposes alternative statistical techniques survival analysis, on stroke patients, aging 65 years and over. Survival Analysis [88] is a branch of statistics that typically uses LOS data to study the effect of different patient attributes on survival time.

In addition, Figure 1 highlights a special type of statistical method which includes the analysis of covariates Regression Analysis. Covariates are defined in the context of LOS as the patient's characteristics and external factors which possibly predict LOS. Within this type are found linear regression and logistic regression, which is a special case of survival models [87]. The models developed often include the patient's diagnoses, procedures, gender and age [26], [44]. Moreover, Freitas et al. [26] used regression models to examine the association of some administrative variables from inpatient episodes in public acute care hospitals in the Portuguese National Health Service with high LOS outliers. The variables include year of discharge, comorbidities, age, adjacent DRG complexity (ADRG), readmission, admission and DRG type, discharge status, distance from residence to hospital and hospital type. Results show that age, type of admission and hospital type were significantly associated with high LOS outliers.

A hospital is a complex stochastic system, therefore simple deterministic approaches for planning and managing such a system is considered inadequate to provide a complete and accurate analysis [4], [89], [90]; also the resulting models which are mostly based on simple rules modelled with regression trees, are usually further adjusted manually according to medical knowledge, decreasing the predictive accuracy of successive models [50] [91] [92] [93]. Grubinger et al. [10] argue that any minor change in the data of such simple models can lead to a completely different tree, although all of these trees can be statistically accurate. As a result, the work presented in their research used the bootstrap-based model method bumping [94] to build diverse regression tree models through systematic re-sampling (uniform randomness) of the data. Bootstrap methods are most commonly based on the idea of combining and averaging models to reduce prediction error. Examples of such methods include Bagging [94], Boosting [95] and Random Forests [96].

A data-driven approach [97]–[100], which will be discussed thoroughly in subsection B, can be used to predict which patients seem likely to experience an extended LOS by analyzing survival data using decision trees (also called survival trees), artificial neural networks, ensemble methods...etc. Usu-

ally these approaches are used to predict categorical survival outcomes (dead or alive) for a given set of patient attributes, or used to measure patient length of stay above or below a certain threshold.

In contrast with these data-driven approaches listed above [97]–[100], Caetano et al. [27] do not perform a classification task to LOS, instead a more information pure regression approach is adopted which predicts the actual number of LOS days and not classes. The study describes 14 input covariates to the LOS target variable. Six regression techniques were tested and compared: Average Prediction (AP), Multiple Regression (MP), Decision Trees (DT), Artificial Neural Network (ANN) ensemble, Support Vector Machines (SVM) and Random Forests (RF). The best results were obtained by the Random Forest model to reveal high impact of inpatient clinical process attributes, instead of the patient's characteristics. Effective predictions can aid healthcare institutions and clinicians to improve their decisions about patient managements and resource allocations [46], [47], [51], [101].

Despite such attempts, Marshall et al. and Garg et al. [5], [7] argue that data-driven methods among the other statistical models fail to address the inherent uncertainty, complexity and heterogeneity in health processes. To address such issues, a more reliable way is to model patient flow as it presents the temporal dimension as well as the structural dimension of the system [7]. Numerous probabilistic models have been proposed to address the issue of LOS, namely Markov models, phase-type distributions, conditional phase-type distributions, compartmental and simulation modeling [5], [44], [50], [51], [102]. Such models may be used for planning health services for both acute and chronic patients. These models will be discussed thoroughly in subsections C, D and E.

B. A Data-Driven Approach to LOS Prediction

Whereas most previous research examines LOS numerically [103]–[105], several studies take a data-driven approach to LOS prediction. A data-driven approach refers to a predictive model that is based on data-mining techniques, such as classification, clustering...etc. Such techniques are used to discover useful patterns in large datasets by showing novel and interesting relationships among data variables. Data Mining techniques facilitate the creation of knowledge and support clinical decision making, in what is known as medical data mining [106], [107].

The data-driven approach classification is used to generate early alerts with respect to a target LOS range for a specific diagnosis related group (DRG). For example, Buchman et al. [108] predict chronicity in a surgical intensive care unit by classifying patients LOS in accordance with a recommended seven-day norm. In response to the need for effective resource planning and cost containment, Mobley et al. [104] predict the LOS of patients receiving post-coronary care over the range of 120 days. Frye et al. [109] use a technique to predict whether the LOS of patients suffering from burns will fall within a one-week period. Cheng et al. in 2009 [43] introduce a study that examines the LOS management of appendectomy patients

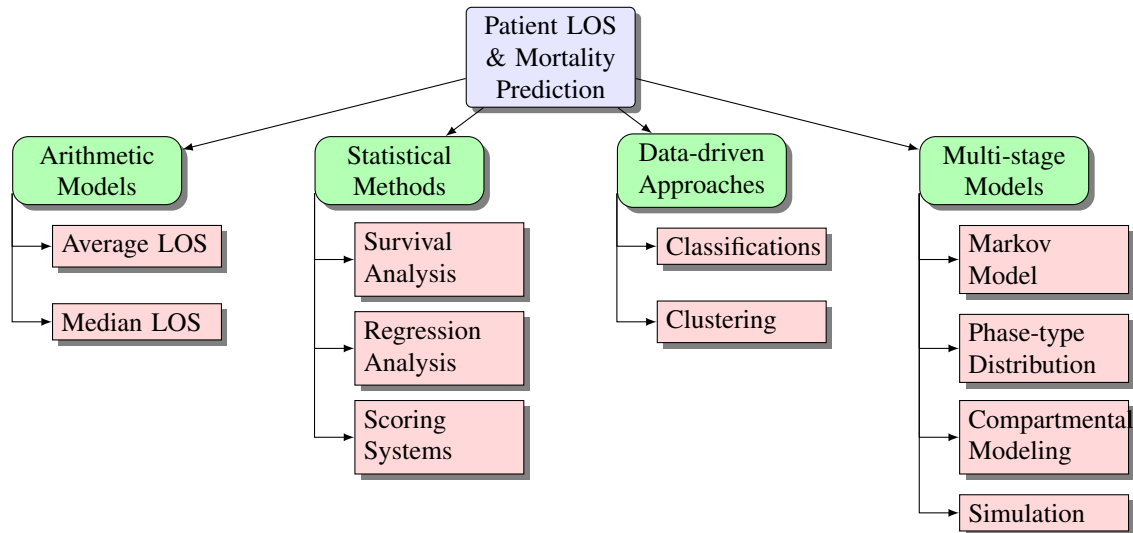


Fig. 1. Classification of LOS & Mortality Prediction Methods.

by building and empirically evaluating an automatic prediction system to identify those patients whose LOS will likely exceed the recommended five-day period. Hachesu et al. [46] apply three classification algorithms namely, decision tree, support vector machines (SVM), and artificial neural network (ANN) to draw an accurate model to predict the LOS of heart patients. Thirty six input variables were used to predict the target variable, LOS. The findings demonstrated that the SVM was the best fit. There was a significant tendency for LOS to be longer in patients with lung or respiratory disorders and high blood pressure. One of the interesting findings was that most single patients (64.3%) had a LOS less than or equal to 5 days, whereas 41.2 % of married patients had a LOS greater than 10 days. The most significant variables affecting LOS were drug categories, such as nitrates and anticoagulants as well as coronary artery disease (CAD) diagnosis. Comorbidity is also a strong predictor of prolonged LOS. Comorbidity is the presence of one or more additional diseases or disorders co-occurring with a primary disease or disorder. There was a significant tendency for LOS to be longer in patients with lung or respiratory disorders and high blood pressure. Gender was significant in predicting LOS since men had longer LOS than women. Age played a notable role as well since analysis revealed that patients aged less than 50 and greater than or equal 80 statistically had increased mean LOS.

Moreover, Rowan et al. [110] implemented a software package demonstrating that artificial neural networks (ANNs) could be used as an effective LOS stratification instrument in postoperative cardiac patients. In [11], Azari et al. propose an approach for predicting hospital length of stay using a multi-tiered data mining approach. They form training sets, using groups of similar claims identified by k-means clustering and perform classification using ten different classifiers. They consistently found that using clustering as a precursor to form the training set gives better prediction results as compared to non-clustering based training sets. Binning the LOS to three

groups of short, medium and long stays, their method identifies patients who need aggressive or moderate early interventions to prevent prolonged stays.

Liu et al. [47] applied two classifiers: decision tree C4.5 & its successor R-C4.5s, Naive Bayesian classifier (NBC) and its successor NBCs to a geriatric hospital dataset, called Clinics Dataset, containing 4722 patient records including patient demographic details, admission reasons, discharge details, outcome and LOS, to predict inpatient LOS for long stay patients. According to [111], C4.5 is one of the classifiers, which has the best combinations in terms of error rate and speed. Also, R-C4.5s combines branches with little classification contribution and thus resulted in building more robust and smaller trees [112]. In addition, NBC is robust and insensitive to missing data as stated in the work of Liu et al. [47].

In addition, phase-type survival trees and mixed distribution survival trees are used to cluster stroke-related patients into clinically meaningful groups with respect to length of stay where partitioning is based on covariates, such as gender, age at time of admission, primary diagnosis code, treatment outcome and discharge destination [5], [44]. Moreover, Kudyba et al. in 2010 [45] utilize the method of neural networks to analyze data describing inpatient cases to examine the effect of the independent variables of patient demographics, primary payer, admission and discharge dates, physician specialty, and detailed radiology procedural variables (including the sum of radiology hours) on the dependent variable of length of stay excess per patient case for a major New Jersey based healthcare provider. Also, artificial neural networks, decision trees and ensemble methods are used in developing an intelligent decision support system- INTCare for intensive medicine in the ICU of the Hospital Santo Antonio (HAS) in Porto, Portugal [35]. In addition, the bootstrap-based method bumping is used by [10] to build diverse and more accurate regression tree models for DRG systems in Austria. Eight datasets are used consisting of patient's main diagnosis, sec-

TABLE I
A SUMMARY OF RESEARCH PAPERS GROUPED BY ANALYTICAL METHODS TO LOS AND MORTALITY PREDICTION

	Method	references	applications	Common evaluation methods
Length of stay	Data-driven & Data Mining	[43], [10], [44], [27], [5], [35], [45], [46], [47]	Stroke unit, Intensive Medicine, Appendectomy, General	Sensitivity, Specificity, Harmonic measure (F-measure, precision and recall average) and RBF kernel
	Statistical methods	[26], [40], [48], [49], [34], [30], [29]	Acute care, intensive care, stroke patients, ICU cardiac surgery, nutrition and acute kidney	Mean, median, std. deviation, skewness, kurtosis, minmax, 25th percentile, 75th percentile, confidence interval and area under the receiver-operating characteristic curve (AUROC).
	Multi-stage methods	[50], [51], [52]	healthcare of elderly patients, Obstetric Unit (pregnancy and childbirth) and Emergency Departments	Sensitivity analysis, simulation models, generalized Erlang, hyperexponential and Coxian.
Mortality prediction	Scoring Systems	[53], [54], [55], [56], [57], [58], [59], [12], [60], [61], [62], [63], [64], [65], [66], [67], [68], [69], [70], [71]	Critical Care	confidence interval and area under the receiver-operating characteristic curve (AUROC)
	Data Mining	[72], [73], [74], [75], [76], [77], [78], [79], [80], [81], [62], [82]	Critical Care	confidence interval and area under the receiveroperating characteristic curve (AUROC)
LOS + Mort.	Arithmetic, Statistical (regression & scoring systems) & Data Mining Methods	[83], [84], [85], [86]	Acute Medicine, Critical Care (trauma & surgical patients)	Mean, confidence interval, AUROC, C-statistic

ondary diagnoses, procedures, number of diagnoses, number of procedures, gender and age as well as patients' LOS.

C. Markov Model and Phase-type Distributions

Markov and semi-Markov chain models are models that assume sub-groups of patients are homogeneous and events occur at equally spaced intervals of time; queueing models and deterministic models of the transition of patients between states. These techniques are useful for examining patient flow in large population groups where Markov assumptions can be made [25]. Phase-type distributions describe the time to absorption of a finite Markov chain in continuous time when there is a single absorbing state and the stochastic process starts in a transient state [87].

The first probabilistic approach describes a special type of Markov model known as the Coxian phase-type distribution and its further development into the conditional phase-type distribution. The Coxian phase-type distribution, allows the representation of the continuous duration of stay of patients in hospital as a series of sequential phases which the patients progress through until they leave the hospital completely [113].

It is possible to expand the theory of Coxian phase-type distributions to include a network of additional interrelated variables (such as patient characteristics) that may interact to influence patient LOS Conditional phase-type distribution. This approach allows the incorporation of discrete and continuous variables representing causality. Marshall et al. [114] [115]

[7] [116] uses conditional phase-type distribution to model the LOS of elderly patients in hospital. The approach illustrates data on hospital processes for a number of geriatric patients along with personal details, admissions reasons, dependency levels and destination (the causal network). The final model represents patient LOS in terms of five of the most significant patient variables in the dataset, namely patient age, gender, admission method into hospital, Barthel grade (dependency score) and destination on departure from hospital.

D. Compartmental modeling

The second general approach described is the compartmental model. Compartmental modeling of patient flow is a type of mathematical model used for describing the way patients are transmitted among the compartments of a healthcare system. Each compartment is assumed to be a homogeneous entity within which the entities being modelled are equivalent. For instance, in a pharmaceutical model, the compartments may represent different sections of a body within which the concentration of a drug is assumed to be uniformly equal. Another example, in a healthcare facility, the compartments may represent the different stages that patient goes through-acute, long-stay and death.

Haigeng Xie et al. [117] present a model-based approach to extract from an administrative social care dataset, high-level length of stay patterns of residents in long-term care (LTC). A continuous-time Markov model, a residents stay in both residence care (RC) and nursing care (NC) is modeled

as consisting of a short-stay and a long-stay phase, was used to show the flow of residents within and between RC and NC, as well as discharge from RC and NC. The model has been extended to incorporate residents features, such as gender. The final model showed that gender has a significant influence on transition rates.

Irvine et al. [118] describes the development of a two-stage continuous-time Markov model that describes the movement of patients through geriatric hospitals. Patients are initially admitted to the acute state from which they transfer to the long-stay state or leave the hospital completely through discharge or death state. McClean et al. [119] extends the stochastic Markov model presented in [118] to a three-stage one and attaches different costs to each stage thus taking cost into account. Taylor et al. [120] uses a continuous time Markov model and applies it to the case of a four compartmental model, where the four stages are acute, long-stay, community, and dead. The model estimates the expected number of patients at any time t in each stage. Taylor et al. [121] extends these models to contain six stages. Garg et al. 2012 [5] proposed a novel distribution, multi-absorbing state phase-type distribution, as a generalization of the single absorbing state Coxian phase-type distribution for representing a Markov process having more than one absorbing states. The approach effectively forecasts the bed requirements in a care unit considering the effect of several factors, such as patient demography age and gender, as well as treatment outcome based on diagnosis and patients expected destination after discharge, which may also affect a patients LOS in hospital.

E. Simulation modeling

Simulation-based models simulate scenarios which replicate real life in an attempt to understand the complex health processes and their interactions [5]. Vasilakis et al. 2003 [48] illustrate how average LOS can be a misleading measure. The research proposes alternative statistical techniques, such as survival analysis, the application of mixed exponential and phase-type distributions demonstrated in two dynamic models of patient flow compartmental model (small, medium and long stay) and discrete event simulation model, introducing capacity constraints in the various stages of the model, such as bed blockage and refuse-admission rates.

In addition, Griffin et al. 2011 [51], developed a simulation model using a path-based approach for an obstetric unit to study tradeoffs in blocking and system efficiency. The model focuses on patient flow, considering patient classification, blocking effects, time dependent arrival and departure patterns, and statistically supported distributions for LOS. Moreover, the study conducted by Wang et al. [52] in the Emergency department at a community hospital, Entral Baptist Hospital in Lexington, KY, uses a discrete-event simulation model to evaluate patient outcome, identify the impact of critical resources and procedures, conduct "What if" analysis for various staffing and operational scenarios, and provide recommendations for hospital management.

Discrete event simulation models allow patients to have individual attributes and to interact with resource provision but they are more time consuming to test and run. They are particularly suitable for models of systems of patient care where the constraints on resource availability are important. They may also be used on unconstrained population models with several thousands of patients. A significant development in simulation is the facility to model entities so that they can participate in more than one activity simultaneously and interrupt each other. The credibility of any model is dependent on reliable data which are not always readily available in the British health service [25].

IV. APPLICATIONS IN MORTALITY PREDICTION

The primary concern of any healthcare system is to relieve the patient symptoms, prevent complications and prolong the patient's life. In order to achieve these goals, it is crucial in the ICU to provide the correct treatment and to predict clinical deterioration early enough so preventive or curative actions can be taken in time. Extensive bedside monitoring in hospital ICUs has resulted in complex data-intensive environment regarding patient physiology, which presents a rich context for clinical data analysis. The majority of mortality prediction research has focused on severity of illness scoring systems designed for risk estimation at 24 hours after ICU admission or data-mining algorithms that help predict mortality. The following two subsections will illustrate the use of different scoring systems and data-mining algorithms in predicting mortality in critical care.

A. Scoring Systems in Critical Care Mortality Prediction

A number of researchers have explored using daily severity of illness scores. In 1993, Le Gall et al. suggested that despite likely being too time-consuming for most ICUs, daily scores would be the most efficient way to evaluate the progression of risk of death [54]. Rue et al. found that the mortality prediction on the current-day was the most informative in fact, the mortality probability at admission and on previous days did not improve performance from the current days score [70]. The importance of the current-day mortality prediction that Rue et al. observed confirms Lemeshow et al.s finding that the most important features change between the admission MPM model and the 24, 48 and 72 hour MPM models. The logistic regression equation also changes between 24-hour intervals to reflect an increasing probability of mortality [122]. From their observations, Lemeshow et al. make the general observation that a patient in the ICU with a steady clinical profile is actually getting worse. In addition, others have confirmed the usefulness of daily severity scores; Wagner et al. showed strong results looking at daily risk predictions based on the APACHE III score and several additional variables such as the primary reason for ICU admission and treatment before ICU admission. Wagner et al.s study relied on over 17,440 patients from 40 U.S. hospitals [123].

B. Data Mining in Critical Care Mortality Prediction

Till today a standard statistical method such as logistic regression used by the scoring systems has been well received by critical care professionals to predict the risk of mortality or adverse events for patients with critical illnesses or injuries admitted to ICU. Despite warnings from many of the original researchers and several studies [124], many caregivers have come to expect the availability of a severity score to assist them in treating individual patients. However, these predictions are not accurate enough for individual patients and no tools exist to reliably predict an individual patient's progress on a critical care condition in a timely manner [62], [78]. As a result, local customized mortality prediction models could perform better as compared to the corresponding current standard severity scoring system. The study conducted in [125] revealed better results for all three subsets of patients: patients with acute kidney injury (AUC = 0.875 for ANN, vs. SAPS, AUC = 0.642), patients with subarachnoid hemorrhage (AUC = 0.958 for BN, vs. SAPS, AUC = 0.84), and elderly patients undergoing open heart surgery (AUC = 0.94 for ANN, vs. EuroSCORE, AUC = 0.648). Moreover, studies performed in [62], [72], [74], [77], [78], [126]–[132] concluded that more flexible nonparametric approaches based on data mining techniques, such as artificial neural networks (ANN), Support Vector Machines (SVM), and decision trees (DTs) might perform at least as well, if not better, than standard logistic regression in ICU mortality prediction. Also the use of untransformed explanatory variables resulted in better results than those transformed using scores/ weights [62].

In 1996, Dybowski et al. [72], reported a significantly improved AUROC using artificial neural networks as compared to standard logistic regression. Also research in [73] reported better performance of ANN over APACHE III. However, research in [74]–[76] found that logistic regression and neural networks performed similarly for ICU mortality prediction. Such conflicting results on the performance of different prediction tools reveal that no single algorithm invariably outperforms all others; it depends on the underlying population being tested, the set of explanatory variables available and the outcome of interest. Contradicting results were reported for other techniques as well. For instance, in 2011 Ribas et al. [77] showed that the use of SVMs resulted in increased prediction accuracy as compared to the APACHE II score. Likewise, the study conducted in [78] compared the predictive accuracy of ANN, SVM and DT derived from the University of Kentucky Hospital's ICU patients' data with the APACHE III scoring system. Results showed that the best performing model is the Clementine's C5.0 algorithm (DT) followed by SVM, APACHE III and ANN.

As mentioned earlier, there is no single algorithm that outperforms others; it depends on the population of interest, the variables measured and the outcome being tested. However, some models reveal strengths over others in certain aspects. For example, the major advantage for the use of DTs over other models lies in its descriptive modeling as it explains hidden

clinical implications unlike ANNs which lacks logic between input and output nodes. From another perspective, DT, RF, ANN, Bayesian networks and kernel methods such as SVM can handle large size data samples and integrate background knowledge into analysis [82].

V. METHODS TO MORTALITY PREDICTION

This section explores the methods used in predicting patient mortality in critical care. After surveying the previous literature, critical care mortality prediction methods were categorized into 2 subgroups as shown in Figure I. A classification of the reviewed papers based on mortality prediction methods is shown in table I.

A. Scoring Systems for Mortality Prediction

Scoring systems can be divided into two categories, those that assess disease severity on admission and use it to predict outcome, for example, Acute Physiology and Chronic Health Evaluation (APACHE) [53], Simplified Acute Physiology Score (SAPS) [54], Mortality Probability Model (MPM) [55], and those scores that assess the presence and severity of organ dysfunction, for example, Sequential Organ Failure Assessment (SOFA) [56]. The SOFA score is limited to 6 organs by looking at respiration, coagulation, liver, cardiovascular, central nervous system, and renal measurements. For each organ, the score provides an assessment of derangement between 0 (normal) and 4 (highly deranged).

Several work in literature have discussed and compared mortality prediction models in intensive care that rely on a panel of experts or statistical models, namely logistic regression [12], [53]–[55], [57]–[60]. Scoring systems, such as Acute Physiology and Chronic Health Evaluation (APACHE) [53] and Simplified Acute Physiology Score (SAPS) [54] assess disease severity on admission and use it to predict outcome. The objective of these models is to compare groups of patients and characterize disease severity from patient demographics and physiological variables obtained within the first 24 hours after ICU admission. However, despite their simplicity it is claimed that these models are not reliable enough for prediction of individual patients since they provide a value that can be averaged for a group of patients [61]. To this day, SAPS II (Simplified Acute Physiology Score) [54] and APACHE II (Acute Physiology and Chronic Health Evaluation) [53] remain the most widely used scores in clinical practice despite attempts for their modification [12], [60], [62], specifically tailored for other populations, such as France, Southern Europe and Mediterranean countries, and to Central and Western Europe [63]–[66].

In 1985, the original model of the APACHE scoring system (1981) was revised and simplified to create APACHE II [67], now the worlds most widely used severity of illness score [12]. The score relies on a panel of experts for variable selection and weights. In APACHE II, there are just 12 physiological variables, compared to 34 in the original score. The effects of age and chronic health status are incorporated directly into the model, weighted according to their relative impact, to

give a single score with a maximum of 71. The worst value recorded during the first 24 hours of a patient's admission to the ICU is used for each physiological variable. The score is not recalculated during the stay; it is by definition an admission score. If a patient is discharged from the ICU and readmitted, a new APACHE II score is calculated. However, many researchers have validated the use of severity of illness scores in settings that deviate from their original design. Alternative settings have included populations such as coronary care patients or subarachnoid hemorrhage patients or days subsequent to the initial 24 hours after admission [68]–[71]. APACHE III [63] was developed in 1991 and in 2002/2003 APACHE IV was developed, which provides length of stay prediction equations [12]. A more detailed comparison of the current scoring systems is available in [12].

Like the APACHE scores, SAPS was calculated from the worst values obtained during the first 24 hours of ICU admission. In 1993, Le Gall et al. [54] used logistic regression analysis to develop SAPS II, which includes 17 variables: 12 physiological variables, age, type of admission and 3 variables related to underlying disease. The SAPS II score was validated using data from consecutive admissions to 137 ICUs in 12 countries [54]. Research in [64] introduced an expanded SAPS II by adding six admission variables: age, gender, length of pre-ICU hospital stay, patient location before ICU, clinical category and whether drug overdose was present. Results show that the expanded SAPS II performed better than the original and a customized SAPS II, with an AUROC (area under the receiver operating characteristic curve) of 0.879. A study conducted by Gilani et al. [60], showed that the prognostic accuracy of APACHE II was excellent for (AUC: 0.828) score and acceptable for APACHE III (AUC: 0.782) and SAPS II (AUC: 0.778) scores. According to the clinical review conducted by Vincent et al. [12], the different types of scores should be seen as complementary, rather than competitive and mutually exclusive.

B. Data Mining Techniques for Mortality Prediction

Various authors have advocated the use of machine learning techniques for predicting ICU mortality over the use of logistic regression methods. Research in [72] and [73] have reported better performance of Artificial Neural Networks (ANN) over logistic regression. However, research in [74]–[76] found that logistic regression and neural networks performed similarly for ICU mortality prediction. Others [77]–[81] found that Decision trees and Support Vector machines performed better. In 2011, Ribas et al. [77] showed that the use of SVMs resulted in increased prediction accuracy as compared to the APACHE II score. Likewise, the study conducted in [78] compared the predictive accuracy of ANN, SVM and DT derived from the University of Kentucky Hospital's ICU patients' data with the APACHE III scoring system. Results showed that the best performing model is the Clementine's C5.0 algorithm (DT) followed by SVM, APACHE III and ANN. These results confirm with earlier findings in Delen et al. [80] which also reported that C5.0 was the best predictor with the highest

accuracy of 93.6% in predicting breast cancer survivability. In addition, [81] concluded that a decision tree used in their study provided a clinically acceptable mining result in predicting susceptibility of prostate carcinoma patients at low risk for lymph node spread. On the other hand, Ramon et al. [130] reported that the AUCs of decision tree based algorithms (decision tree learning, 65%; first order random forests, 81%) yielded smaller areas compared to those of naive Bayesian networks (AUC, 85%) and tree-augmented naive Bayesian networks (AUC, 82%) in their study on a small dataset containing 1,548 mechanically ventilated ICU patients. Similarly Pirracchio et al. [62] reported that Bayesian Additive Regression Trees (BART) is the best candidate when using transformed variables, while Random Forests outperformed all other candidates when using untransformed variables. Other authors achieved improved mortality prediction using a method based on SVMs [79].

Such conflicting results on the performance of different prediction tools reveal that no single algorithm invariably outperforms all others; it depends on the population of interest, the variables measured and the outcome being tested. However, some models reveal strengths over others in certain aspects. For example, the major advantage for the use of DTs over other models lies in its descriptive modeling as it explains hidden clinical implications unlike ANNs which lacks logic between input and output nodes. From another perspective, DT, RF, ANN, Bayesian networks and kernel methods such as SVM can handle large size data samples and integrate background knowledge into analysis [82].

VI. APPLICATIONS IN CONCURRENT PREDICTION OF LOS AND MORTALITY

There are several works in literature that handle both LOS and mortality prediction concurrently. The prediction models include arithmetic models, such as the mean and median; statistical models, such as regression analysis and data-driven models, such as Bayesian Network. The following are some examples of applications that attempt to concurrently predict patient LOS and mortality.

The work in [83] aims to assess the impact of the introduction of an early warning scoring system (SEWS) on physiological observations and patient outcomes in acute admissions at point of entry to care. The admission of the early warning score correlated both with in-hospital mortality ($P < 0.001$) and length of stay ($P = 0.001$). Moreover, Clark et al. [84] developed a method for predicting concurrently both hospital survival and LOS for seriously ill patients from three trauma centers in Maine, with particular attention to the competing risks of death or discharge alive as determinants of LOS. Poisson regression was used to develop a model for each type of terminal event, with risk factors on admission contributing proportionately to the subsequent rates for each outcome in each interval. Mean LOS and cumulative survival were calculated from a combination of the resulting piecewise exponential models [84]. Similarly, Risk stratification indices (RSIs) for length of stay and mortality endpoints were derived from aggregate

risk associated with individual diagnostic and procedure codes. Results showed that RSI is a broadly applicable and robust system for assessing hospital length of stay and mortality for groups of surgical patients based solely on administrative data [85]. Cai et al. [86] built a Bayesian Network model to estimate the probability of a hospitalized patient being at home, in the hospital, or dead for each of the next 7 days. Electronic health records from 32,634 patients admitted to a Sydney metropolitan hospital via the emergency department from July 2008 through December 2011 were used. The model achieved an average daily accuracy of 80% and AUROC of 0.82. The models predictive ability was highest within 24 hours from prediction (AUROC=0.83) and decreased slightly with time. Death was the most predictable outcome with a daily average accuracy of 93% and AUROC of 0.84 [86].

VII. MEASURING THE PERFORMANCE OF MORTALITY AND LOS PREDICTION MODELS

Mortality prediction is considered a binary classification problem where a classifier attempts to identify whether a patient will live or die. Evaluating classifier performance shows how well a method improves classification. Traditionally classification accuracy is used to measure classifier performance. The classification accuracy gives a good idea of classifier performance when the dataset is balanced, however when the dataset suffers from class imbalance (i.e the number of instances belonging to one class outnumbers that of any other class(es)) some problems emerge. For example, in a binary classification problem, if the majority class outnumbers the minority class 9:1, and all instances were classified as the majority class, the classifier would have an accuracy of 90%, despite 0% of the minority class being classified correctly. In a binary dataset, the classification accuracy shows the number of correctly classified minority instances (true positives), incorrectly classified minority instances (false positives), correctly classified majority instances (true negatives) and incorrectly classified majority instances (false negatives) as follows:

$$ACC = \frac{TP + TN}{P + N} \quad (1)$$

For this reason, better evaluative measures that are independent of the class imbalance ratio and sufficiently recognizes the minority class are preferred. One of such measures used in this paper is the Geometric Mean which calculates the geometric mean between the sensitivity or recall (true positive rate) and specificity (true negative rate).

$$Sensitivity = \frac{TP}{TP + FN} \quad (2)$$

$$Specificity = \frac{TN}{TN + FP} \quad (3)$$

The G-mean is an effective evaluative criterion as it is not dependent on the data distribution. The Gmean is defined as [133]:

$$Gmean = \sqrt{Sensitivity * Specificity} \quad (4)$$

It accentuates the balancing between the specificity and sensitivity while maximizing the recognition between the minority and the majority class. Other evaluative measures used in the paper are Precision (also called positive predictive value) which is the fraction of retrieved instances that are relevant. Sensitivity and precision can be combined using a metric known as the FMeasure [134].

$$Precision = \frac{TP}{TP + FN} \quad (5)$$

$$FMeasure = \frac{2.Recall.Precision}{Recall + Precision} \quad (6)$$

In addition, area under the receiver-operating characteristic curve (AUROC) is also a widely used measure of classification performance in mortality prediction. It is a graphical plot that illustrates the performance of a binary classifier system. The curve is created by plotting the true positive rate (TPR) against the false positive rate (FPR) at various threshold settings.

There are several quantitative methods for measuring the performance of the different LOS prediction models. Measuring the performance of the statistical methods to LOS prediction include several approaches, such as the mean, median, standard deviation, kurtosis, min-max, confidence and AUROC are the most commonly used in literature [26], [40], [48], [49], [34], [30], [29]. As for the data-driven and data mining methods to LOS prediction, there are several evaluation approaches in literature, such as sensitivity, specificity, F-measure or the harmonic mean of precision and recall [43], [10], [44], [27], [5], [35], [45], [46], [47]. On the other hand, sensitivity analysis, simulation models, generalized Erlang, hyper-exponential and Coxian models are the most widely used in measuring the performance of multi-stage models to LOS prediction [50], [51] and [52].

VIII. CONCLUSION

This paper has presented a comprehensive review of the methods and applications of LOS and mortality prediction in acute medicine and critical care.

An introduction to LOS theory and mortality prediction and the main drivers behind the interest in such research was given in the opening section. In addition, several applications in LOS and mortality prediction were demonstrated both in acute medicine and in the intensive care unit environment in particular, highlighting the challenges facing both physicians and information engineers today. However, only a small sample of such applications considered the prediction of both LOS and mortality concurrently. We see this as a limitation in our paper that could be enhanced in future work. An analysis of the various LOS and mortality prediction methods were presented and compared. Moreover, the paper provides a classification for some state-of-the-art literature based on each paper's analytical method utilised to predict LOS and mortality. The four main categories include: (1) Arithmetic methods, (2) Statistical methods, (3) data-driven methods and (4) multi-stage methods. The classification presents a brief

summary of the analytical method, dataset and the evaluation method utilised. Given that LOS and mortality are relatively complex matrices as they are influenced by various external uncontrollable factors, there is no one good-for-all technique that serves its prediction. At present, in most cases several algorithms are tested, tweaked based on some domain knowledge or some performance criteria to enhance the accuracy of prediction. This is considered another limitation as there is no one reliable technique for prediction; it all depends on the situation at hand. However, it is clear that much research remains to be done, especially that the physiological and laboratory datasets in both Acute Medicine and the Critical Care environment are relatively large and well-structured in commercial and non-commercial databases.

REFERENCES

- [1] Wang YC, McPherson K, Marsh T, Gortmaker SL, Brown M. Health and economic burden of the projected obesity trends in the USA and the UK. *The Lancet*. 2011;378(9793):815–825.
- [2] Comas-Herrera A, Wittenberg R, Pickard L. European study of long-term care expenditure. *Long-Term Care Expenditure in an Ageing Society*. 2003;p. 111.
- [3] Roberts A, Marshall L, Charlesworth A. *A decade of austerity? The funding pressures facing the NHS from 2010/11 to 2021/22*. London: Nuffield Trust. 2012;.
- [4] Bain CA, Taylor PG, McDonnell G, Georgiou A. Myths of ideal hospital occupancy. *Medical Journal of Australia*. 2010;193(5):311.
- [5] Garg L, McClean SI, Barton M, Meenan BJ, Fullerton K. Intelligent patient management and resource planning for complex, heterogeneous, and stochastic healthcare systems. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*. 2012;42(6):1332–1345.
- [6] Huntley DA, Cho DW, Christman J, Csernansky JG. Predicting length of stay in an acute psychiatric hospital. *Psychiatric Services*. 2014;.
- [7] Marshall, A H , McClean, S I , Shapcott, C M , Hastie, I R , Millard PH. Developing a Bayesian belief network for the management of geriatric hospital care. *Healthcare Management Science*. 2001;p. 25–30.
- [8] Marshall A, Vasilakis C, El-Darzi E. Length of stay-based patient flow models: recent developments and future directions. *Health Care Management Science*. 2005;8(3):213–220.
- [9] S Shea, R V Sideli, W DuMouchel, G Pulver, R R Arons PDC. Computer-generated informational messages directed to physicians: Effect on length of hospital stay. *J Amer Med Info Assoc*. 1995;2(1):58–64.
- [10] Grubinger T, Kobel C, Pfeiffer KP. Regression tree construction by bootstrap: Model search for drug-systems applied to austrian health-data. *BMC Medical Informatics and Decision Making*. 2010;10(1):9.
- [11] Azari A, Janeja VP, Mohseni A. Healthcare Data Mining: Predicting Hospital Length of Stay (PHLOS). *International Journal of Knowledge Discovery in Bioinformatics (IJKDB)*. 2012;3(3):44–66.
- [12] Vincent JL, Singer M. Critical care: advances and future perspectives. *The Lancet*. 2010;376(9749):1354–1361.
- [13] Vincent JL. Is the current management of severe sepsis and septic shock really evidence based? *PLoS medicine*. 2006;3(9):e346.
- [14] Abbott R, Setter M, Chan S, Choi K. APACHE II: prediction of outcome of 451 ICU oncology admissions in a community hospital. *Annals of Oncology*. 1991;2(8):571–574.
- [15] Harrison DA, Parry GJ, Carpenter JR, Short A, Rowan K. A new risk prediction model for critical care: The Intensive Care National Audit & Research Centre (ICNARC) model*. *Critical care medicine*. 2007;35(4):1091–1098.
- [16] Vincent JL, Moreno R, Takala J, Willatts S, De Mendonça A, Bruining H, et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. *Intensive care medicine*. 1996;22(7):707–710.
- [17] Lemeshow S, Teres D, Klar J, Avrunin JS, Gehlbach SH, Rapoport J. Mortality Probability Models (MPM II) based on an international cohort of intensive care unit patients. *Jama*. 1993;270(20):2478–2486.
- [18] Lim A, Tongkumchum P. Methods for analyzing hospital length of stay with application to inpatients dying in Southern Thailand. *Global Journal of Health Science*. 2009;1(1):p27.
- [19] Chang KC, Tseng MC, Weng HH, Lin YH, Liou CW, Tan TY. Prediction of length of stay of first-ever ischemic stroke. *Stroke*. 2002;33(11):2670–2674.
- [20] Jiang X, Qu X, Davis L. Using Data Mining to Analyze Patient Discharge Data for an Urban Hospital. In: *DMIN*; 2010. p. 139–144.
- [21] Appelros P. Prediction of length of stay for stroke patients. *Acta Neurologica Scandinavica*. 2007;116(1):15–19.
- [22] Robinson GH, Davis LE, Leifer RP. Prediction of hospital length of stay. *Health services research*. 1966;1(3):287.
- [23] McMullan R, Silke B, Bennett K, Callachand S. Resource utilisation, length of hospital stay, and pattern of investigation during acute medical hospital admission. *Postgraduate medical journal*. 2004;80(939):23–26.
- [24] Vahidi R, Kushavar H, Khodayari R. Factors affecting coronary artery patients hospital length of stay of Tabriz Madani hospital 2005-2006. *Journal of Health Administration*. 2006;9(25):63–68.
- [25] Davies R, Davies H. Modelling patient flows and resource provision in health systems. *Omega*. 1994;22(2):123–131.
- [26] Freitas A, Silva-Costa T, Lopes F, Garcia-Lema I, Teixeira-Pinto A, Brazdil P, et al. Factors influencing hospital high length of stay outliers. *BMC health services research*. 2012 Jan;12(1):265.
- [27] Caetano N, Laureano R, Cortez P. A data-driven approach to predict hospital length of stay: a Portuguese case study. 2014;.
- [28] Robinson G, Goldstein M LG. Impact of nutritional status on DRG length of stay. In: *JPEN*; 1987. p. 49–51.
- [29] Chima CS, Barco K, LA DEWITT M, Maeda M, Teran JC, Mullen KD. Relationship of nutritional status to length of stay, hospital costs, and discharge status of patients hospitalized in the medicine service. *Journal of the American Dietetic Association*. 1997;97(9):975–978.
- [30] Isabel T D Correia M. The impact of malnutrition on morbidity, mortality, length of hospital stay and costs evaluated through a multivariate model analysis. *Clinical Nutrition*. 2003 Jun;22(3):235–239.
- [31] Anderson CG, Moxness K, Meister, J BM. The sensitivity and specificity of nutrition-related variables in relationship to the duration of hospital stay and the rate of complications. In: *Mayo Clin Proc*.; 1984. p. 477–485.
- [32] Warnold I LK. Clinical significance of preoperative nutritional status in 215 non cancer patients. In: *Ann Surg*.; 1984. p. 299–305.
- [33] Epstein AM , Red JL HM. The relation of body weight to length of stay and charges for patients undergoing elective surgery: as study of two procedures. *Am J Public Health*. 1987;p. 993–997.
- [34] Chertow GM, Burdick E, Honour M, Bonventre JV, Bates DW. Acute kidney injury, mortality, length of stay, and costs in hospitalized patients. *Journal of the American Society of Nephrology : JASN*. 2005 Nov;16(11):3365–70.
- [35] Portela F, Santos MF, Ivaro Silva, Rua F, Abelha A, Machado J. Adoption of Pervasive Intelligent Information Systems in Intensive Medicine. *Procedia Technology*. 2013;9:1022 – 1032. {CENTERIS} 2013 - Conference on {ENTERprise} Information Systems / ProjMAN 2013 - International Conference on Project Management/ {HCIST} 2013 - International Conference on Health and Social Care Information Systems and Technologies.
- [36] Anthony Celi L, Mark RG, Stone DJ, Montgomery RA. Big data in the intensive care unit. Closing the data loop. *American journal of respiratory and critical care medicine*. 2013;187(11):1157–1160.
- [37] Harrison DA, Brady AR, Rowan K. Case mix, outcome and length of stay for admissions to adult, general critical care units in England, Wales and Northern Ireland: the Intensive Care National Audit & Research Centre Case Mix Programme Database. *Critical Care*. 2004;9(Suppl 3):S1.
- [38] Nolan J, Laver S, Welch C, Harrison D, Gupta V, Rowan K. Outcome following admission to UK intensive care unit after cardiac arrest: a secondary analysis of the ICNARC Case Mix Programme Database*. *Anaesthesia*. 2007;62(12):1207–1216.
- [39] Stow PJ, Hart GK, Hignett T, George C, Herkes R, McWilliam D, et al. Development and implementation of a high-quality clinical database: the Australian and New Zealand Intensive Care Society Adult Patient Database. *Journal of critical care*. 2006;21(2):133–141.
- [40] Levin SR, Harley ET, Fackler JC, Lehmann CU, Custer JW, France D, et al. Real-time forecasting of pediatric intensive care unit length of stay using computerized provider orders. *Critical care medicine*. 2012 Nov;40(11):3058–64.

- [41] Gardner RM, Hawley WL, East TD, Oniki TA, Young HFW. Real time data acquisition: recommendations for the Medical Information Bus (MIB). *International journal of clinical monitoring and computing*. 1991;8(4):251–258.
- [42] Cook SF, Visscher WA, Hobbs CL, Williams RL, Committee PICI, et al. Project IMPACT: results from a pilot validity study of a new observational database. *Critical care medicine*. 2002;30(12):2765–2770.
- [43] Hu PJH. A Data-Driven Approach to Manage the Length of Stay for Appendectomy Patients. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*. 2009 Nov;39(6):1339–1347.
- [44] Garg L, Mcclean S, Meenan BJ, Millard P. Phase-Type Survival Trees and Mixed Distribution Survival Trees for Clustering Patients Hospital Length of Stay. 2011;22(1):57–72.
- [45] Kudyba S, Gregorio T. Identifying factors that impact patient length of stay metrics for healthcare providers with advanced analytics. *Health informatics journal*. 2010 Dec;16(4):235–45.
- [46] Hachisu PR, Ahmadi M, Alizadeh S, Sadoughi F. Use of data mining techniques to determine and predict length of stay of cardiac patients. *Healthcare informatics research*. 2013;19(2):121–129.
- [47] Liu P, Lei L, Yin J, Zhang W, Najjun W, El-Darzi E. Healthcare data mining: predicting inpatient length of stay. 2006;.
- [48] Marshall A, Vasilakis C, El-Darzi E. Length of stay-based patient flow models: recent developments and future directions. *Health Care Management Science*. 2005;8(3):213–220.
- [49] Hein OV, Birnbaum J, Wernecke K, England M, Konertz W, Spies C. Prolonged intensive care unit stay in cardiac surgery: risk factors and long-term-survival. *The Annals of thoracic surgery*. 2006 Mar;81(3):880–5.
- [50] Fackrell M. Modelling healthcare systems with phase-type distributions. *Health Care Management Science*. 2008 Jun;12(1):11–26.
- [51] Griffin J, Xia S, Peng S, Keskinocak P. Improving patient flow in an obstetric unit. *Health care management science*. 2012 Mar;15(1):1–14.
- [52] Wang J, Li J, Tussey K, Ross K. Reducing length of stay in emergency department: A simulation study at a community hospital. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*. 2012;42(6):1314–1322.
- [53] Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. *Critical care medicine*. 1985;13(10):818–829.
- [54] Le Gall JR, Lemeshow S, Saulnier F. A new simplified acute physiology score (SAPS II) based on a European/North American multicenter study. *Jama*. 1993;270(24):2957–2963.
- [55] Lemeshow S, Teres D, Klar J, Avrunin JS, Gehlbach SH, Rapoport J. Mortality Probability Models (MPM II) based on an international cohort of intensive care unit patients. *Jama*. 1993;270(20):2478–2486.
- [56] Vincent JL, De Mendonça A, Cantraine F, Moreno R, Takala J, Suter PM, et al. Use of the SOFA score to assess the incidence of organ dysfunction/failure in intensive care units: results of a multicenter, prospective study. *Critical care medicine*. 1998;26(11):1793–1800.
- [57] Le Gall JR, Loirat P, Alperovitch A, Glaser P, Granthil C, Mathieu D, et al. A simplified acute physiology score for ICU patients. *Critical care medicine*. 1984;12(11):975–977.
- [58] Poole D, Rossi C, Latronico N, Rossi G, Finazzi S, Bertolini G, et al. Comparison between SAPS II and SAPS 3 in predicting hospital mortality in a cohort of 103 Italian ICUs. Is new always better? *Intensive care medicine*. 2012;38(8):1280–1288.
- [59] Rosenberg AL. Recent innovations in intensive care unit risk-prediction models. *Current opinion in critical care*. 2002;8(4):321–330.
- [60] Gilani MT, Razavi M, Azad AM, et al. A comparison of Simplified Acute Physiology Score II, Acute Physiology and Chronic Health Evaluation II and Acute Physiology and Chronic Health Evaluation III scoring system in predicting mortality and length of stay at surgical intensive care unit. *Nigerian Medical Journal*. 2014;55(2):144.
- [61] Hug C. Detecting hazardous intensive care patient episodes using real-time mortality models; 2009.
- [62] Pirracchio R, Petersen ML, Carone M, Rigon MR, Chevret S, van der Laan MJ. Mortality prediction in intensive care units with the Super ICU Learner Algorithm (SICULA): a population-based study. *The Lancet Respiratory Medicine*. 2015;3(1):42–52.
- [63] Knaus WA, Wagner DP, Draper EA, Zimmerman JE, Bergner M, Bastos PG, et al. The APACHE III prognostic system. Risk prediction of hospital mortality for critically ill hospitalized adults. *Chest Journal*. 1991;100(6):1619–1636.
- [64] Le Gall JR, Neumann A, Hemery F, Bleriot JP, Fulgencio JP, Garrigues B, et al. Mortality prediction using SAPS II: an update for French intensive care units. *Critical Care*. 2005;9(6):R645.
- [65] Metnitz B, Schaden E, Moreno R, Le Gall JR, Bauer P, Metnitz PG, et al. Austrian validation and customization of the SAPS 3 Admission Score. *Intensive care medicine*. 2009;35(4):616–622.
- [66] Moreno RP, Metnitz PG, Almeida E, Jordan B, Bauer P, Campos RA, et al. SAPS 3From evaluation of the patient to evaluation of the intensive care unit. Part 2: Development of a prognostic model for hospital mortality at ICU admission. *Intensive care medicine*. 2005;31(10):1345–1355.
- [67] Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. *Critical care medicine*. 1985;13(10):818–829.
- [68] Schuster H, Schuster F, Ritschel P, Wilts S, Bodmann K. The ability of the Simplified Acute Physiology Score (SAPS II) to predict outcome in coronary care patients. *Intensive care medicine*. 1997;23(10):1056–1061.
- [69] Hekmat K, Kroener A, Stuetzer H, Schwinger RH, Kampe S, Bennink GB, et al. Daily assessment of organ dysfunction and survival in intensive care unit cardiac surgical patients. *The Annals of thoracic surgery*. 2005;79(5):1555–1562.
- [70] Rué M, Quintana S, Álvarez M, Artigas A. Daily assessment of severity of illness and mortality prediction for individual patients. *Critical care medicine*. 2001;29(1):45–50.
- [71] Schuiling WJ, de Weerd AW, Dennesen PJ, Algra A, Rinkel GJ. The simplified acute physiology score to predict outcome in patients with subarachnoid hemorrhage. *Neurosurgery*. 2005;57(2):230–236.
- [72] Dybowski R, Gant V, Weller P, Chang R. Prediction of outcome in critically ill patients using artificial neural network synthesised by genetic algorithm. *The Lancet*. 1996;347(9009):1146–1150.
- [73] Nimgaonkar A, Sudarshan S. Predicting hospital mortality for patients in the intensive care unit: a comparison of artificial neural networks with logistic regression models. *Intensive Care Med*. 2004;30:248–253.
- [74] Clermont G, Angus DC, DiRusso SM, Griffin M, Linde-Zwirble WT. Predicting hospital mortality for patients in the intensive care unit: a comparison of artificial neural networks with logistic regression models. *Critical care medicine*. 2001;29(2):291–296.
- [75] Wong L, Young J. A comparison of ICU mortality prediction using the APACHE II scoring system and artificial neural networks. *Anaesthesia*. 1999;54(11):1048–1054.
- [76] Doig G, Inman K, Sibbald W, Martin C, Robertson J. Modeling mortality in the intensive care unit: comparing the performance of a back-propagation, associative-learning neural network with multivariate logistic regression. In: *Proceedings of the Annual Symposium on Computer Application in Medical Care*. American Medical Informatics Association; 1993. p. 361.
- [77] Ribas VJ, López JC, Ruiz-Sanmartín A, Ruiz-Rodríguez JC, Rello J, Wojdel A, et al. Severe sepsis mortality prediction with relevance vector machines. In: *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*. IEEE; 2011. p. 100–103.
- [78] Kim S, Kim W, Park RW. A comparison of intensive care unit mortality prediction models through the use of data mining techniques. *Healthcare informatics research*. 2011;17(4):232–243.
- [79] Citi L, Barbieri R. PhysioNet 2012 challenge: predicting mortality of ICU patients using a cascaded SVM-GLM paradigm. In: *Computing in Cardiology (CinC), 2012*. IEEE; 2012. p. 257–260.
- [80] Delen D, Walker G, Kadam A. Predicting breast cancer survivability: a comparison of three data mining methods. *Artificial intelligence in medicine*. 2005;34(2):113–127.
- [81] Crawford ED, Batuello JT, Snow P, Gamito EJ, McLeod DG, Partin AW, et al. The use of artificial intelligence technology to predict lymph node spread in men with clinically localized prostate carcinoma. *Cancer*. 2000;88(9):2105–2109.
- [82] Meyfroidt G, Guiza F, Ramon J, Bruynooghe M. Machine learning techniques to examine large patient databases. *Best Practice & Research Clinical Anaesthesiology*. 2009;23(1):127–143.
- [83] Paterson R, MacLeod D, Thetford D, Beattie A, Graham C, Lam S, et al. Prediction of in-hospital mortality and length of stay using an early warning scoring system: clinical audit. *Clinical Medicine*. 2006;6(3):281–284.

- [84] Clark DE, Ryan LM. Concurrent prediction of hospital mortality and length of stay from risk factors on admission. *Health services research*. 2002;37(3):631–645.
- [85] Sessler DI, Sigl JC, Manberg PJ, Kelley SD, Schubert A, Chamoun NG. Broadly applicable risk stratification system for predicting duration of hospitalization and mortality. *The Journal of the American Society of Anesthesiologists*. 2010;113(5):1026–1037.
- [86] Cai X, Perez-Concha O, Coiera E, Martin-Sanchez F, Day R, Roffe D, et al. Real-time prediction of mortality, readmission, and length of stay using electronic health record data. *Journal of the American Medical Informatics Association*. 2016;23(3):553–561.
- [87] Guzman Castillo M. Modelling patient length of stay in public hospitals in Mexico. University of Southampton; 2012.
- [88] D R Cox DO. *Analysis of Survival Data*. U.K.: Chapman & Hall; 1984.
- [89] McCarthy SM. Hospital capacity: what is the measure and what is the goal? *The Medical Journal of Australia*. 2010;193(5):252–253.
- [90] Jones RP. Myths of ideal hospital size. *Medical Journal of Australia*. 2010;193(5):298–300.
- [91] L Garg, S I McClean, B J Meenan PHM. A non-homogeneous discrete time Markov model for admission scheduling and resource planning in a care system. *Health Care Management Science*. 2010;13(2):155–169.
- [92] L Garg, S I McClean, B J Meenan PHM. Non-homogeneous Markov models for sequential pattern mining of healthcare data. *IMA J Management Math*. 2009;20(4):327–344.
- [93] L Garg, S I McClean, M Barton, B J Meenan KF. Forecasting hospital bed requirements and cost of care using phase type survival trees. In: *IEEE Conference Intell*; 2010. p. 185–190.
- [94] Breiman L. Bagging predictors. *Machine learning*. 1996;24(2):123–140.
- [95] Friedman JH. Greedy function approximation: a gradient boosting machine. *Annals of statistics*. 2001;p. 1189–1232.
- [96] Breiman L. Random forests. *Machine learning*. 2001;45(1):5–32.
- [97] Vapnik V. *The Nature of Statistical Learning Theory*; 1995.
- [98] W O Kim, H K Kil, J W Kang HRP. Prediction on lengths of stay in the postanesthesia care unit following general anesthesia: Preliminary study of the neural network and logistic regression modeling. *J Korean Med Sci*. 2000;15(1):25–30.
- [99] W E Pofahl, S M Walczak, E Rhone SDI. Use of an artificial neural network to predict length of stay in acute pancreatitis. *Mer Surgeon*. 1998;64(9):868–872.
- [100] Azari A, Janeja VP, Mohseni A. Predicting hospital length of stay (PHLOS): A multi-tiered data mining approach. In: *Data Mining Workshops (ICDMW)*, 2012 IEEE 12th International Conference on. IEEE; 2012. p. 17–24.
- [101] Rowan M, Ryan T, Hegarty F, OHare N. The use of artificial neural networks to stratify the length of stay of cardiac patients based on preoperative and initial postoperative factors. *Artificial Intelligence in Medicine*. 2007;40(3):211–221.
- [102] Vasilakis C, Marshall AH. Modelling nationwide hospital length of stay: opening the black box**. *Journal of the Operational Research Society*. 2005;56(7):862–869.
- [103] J Grigsby, R Kookan JH. Simulated neural networks to predict outcomes, costs, and length of stay among orthopedic rehabilitation patients. *ARCh Phys Med Rehabil*. 1994;75(10):1077–1081.
- [104] B A Mobley, R Leasure LD. Artificial neural network predictions of lengths of stay in a post-coronary care unit. *Heart Lung*. 1995;24(3):251–256.
- [105] B Zernikow, K Holtmannspotter, E Michel, F Hornschuh, K Groote KHH. Predicting length of stay in preterm neonates. *Eur J Pediatr*. 1999;158(1):59–62.
- [106] Silva Á, Cortez P, Santos MF, Gomes L, Neves J. Mortality assessment in intensive care units via adverse events using artificial neural networks. *Artificial Intelligence in Medicine*. 2006;36(3):223–234.
- [107] Silva Á, Cortez P, Santos MF, Gomes L, Neves J. Rating organ failure via adverse events using data mining in the intensive care unit. *Artificial intelligence in medicine*. 2008;43(3):179–193.
- [108] T G Buchman, K L Kubos, A J Seidler MJS. A comparison of statistical and connectionist models for the prediction of chronicity in a surgical intensive care unit. *Crit Care Med*. 1994;22(5):750–762.
- [109] K E Frye, S D Izenberg, M D Williams AL. Simulated biologic intelligence used to predict length of stay and survival of bums. *J Burn Care Rehabil*. 1996;17:540–546.
- [110] Rowan M, Ryan T, Hegarty F, OHare N. The use of artificial neural networks to stratify the length of stay of cardiac patients based on preoperative and initial postoperative factors. *Artificial Intelligence in Medicine*. 2007;40(3):211 – 221.
- [111] Lim TS, Loh WY, Shih YS. A comparison of prediction accuracy, complexity, and training time of thirty-three old and new classification algorithms. *Machine learning*. 2000;40(3):203–228.
- [112] Yao Z, Liu P, Lei L, Yin J. R-C4. 5 Decision tree model and its applications to health care dataset. In: *Services Systems and Services Management*, 2005. Proceedings of ICSSSM'05. 2005 International Conference on. vol. 2. IEEE; 2005. p. 1099–1103.
- [113] Faddy M, McClean S. Analysing data on lengths of stay of hospital patients using phase-type distributions. *Applied Stochastic Models in Business and Industry*. 1999;15(4):311–317.
- [114] Marshal, A H , McClean, S I , Shapcott, C M , Millard PH. Modeling patient duration of stay to facilitate resource management of geriatric hospitals. *Health care management science*. 2002;p. 313–319.
- [115] Marshal, A H MSI. Conditional phase-type distributions for modelling patient length of stay in hospital. *International Transactions in Operational Research*. 2003;p. 567–576.
- [116] Golúke N, Huibers C, Stalpers S, Taekema D, Vermeer S, Jansen P. An observational, retrospective study of the length of stay, and its influencing factors, among elderly patients at the Emergency Department. *European Geriatric Medicine*. 2015;.
- [117] Xie H, Chausalet TJ, Millard PH. A Model-Based Approach to the Analysis of Patterns of Length of Stay in Institutional Long-Term Care. *IEEE Transactions on Information Technology in Biomedicine*. 2006 Jul;10(3):512–518.
- [118] Irvine, V , McClean, S I , Millar PH. Stochastic models for geriatric inpatient behavior. *IMA Journal of Mathematics Applied in Medicine and Biology*. 1994;p. 207–216.
- [119] McClean, S I , McAlea, B , Millard PH. Using a Markov reward model to estimate spend-down costs for a geriatric department. *Journal of Operational Research Society*. 1998;p. 1021–1025.
- [120] Taylor, G J , McClean, S I , Millar PH. Continuous-time Markov model for geriatric patient behaviour. *Applied Stochastic Models and Data Analysis*. 1998;p. 315–323.
- [121] Taylor, G J , McClean, S I , Millard PH. Stochastic Models of Geriatric Patient Bed Occupancy Behaviour. *Journal of the royal statistical society series*. 2000;p. 1–10.
- [122] Lemeshow S, Klar J, Teres D, Avrunin JS, Gehlbach SH, Rapoport J, et al. Mortality probability models for patients in the intensive care unit for 48 or 72 hours: a prospective, multicenter study. *Critical care medicine*. 1994;22(9):1351–1358.
- [123] Wagner DP, Knaus WA, Harrell FE, Zimmerman JE, WATIS C. Daily prognostic estimates for critically ill adults in intensive care units: results from a prospective, multicenter, inception cohort analysis. *Critical care medicine*. 1994;22(9):1359–1372.
- [124] SCHAUFER JH, Maurer A, Jochimsen F, Emde C, WEGSCHEIDER K, ARNTZ HR, et al. Outcome prediction models on admission in a medical intensive care unit: Do they predict individual outcome?. *Critical care medicine*. 1990;18(10):1111–1118.
- [125] Celi LA, Galvin S, Davidzon G, Lee J, Scott D, Mark R. A database-driven decision support system: Customized mortality prediction. *Journal of personalized medicine*. 2012;2(4):138–148.
- [126] Foltran F, Berchiolla P, Giunta F, Malacarne P, Merletti F, Gregori D. Using VLAD scores to have a look insight ICU performance: towards a modelling of the errors. *Journal of evaluation in clinical practice*. 2010;16(5):968–975.
- [127] Gortzis LG, Sakellaropoulos F, Ilias I, Stamoulis K, Dimopoulou I. Predicting ICU survival: a meta-level approach. *BMC health services research*. 2008;8(1):1.
- [128] Lucas P. Bayesian analysis, pattern analysis, and data mining in health care. *Current opinion in critical care*. 2004;10(5):399–403.
- [129] Sierra B, Serrano N, Larrañaga P, Plasencia EJ, Inza I, Jiménez JJ, et al. Using Bayesian networks in the construction of a bi-level multi-classifier. A case study using intensive care unit patients data. *Artificial Intelligence in Medicine*. 2001;22(3):233–248.
- [130] Ramon J, Fierens D, Güiza F, Meyfroidt G, Blockeel H, Bruynooghe M, et al. Mining data from intensive care patients. *Advanced Engineering Informatics*. 2007;21(3):243–256.
- [131] Silva I, Moody G, Scott DJ, Celi LA, Mark RG. Predicting in-hospital mortality of icu patients: The physionet/computing in cardiology chal-

- lenge 2012. In: Computing in Cardiology (CinC), 2012. IEEE; 2012. p. 245–248.
- [132] Silva Á, Cortez P, Santos MF, Gomes L, Neves J. Mortality assessment in intensive care units via adverse events using artificial neural networks. *Artificial Intelligence in Medicine*. 2006;36(3):223–234.
- [133] Kubat M, Matwin S, et al. Addressing the curse of imbalanced training sets: one-sided selection. In: *ICML*. vol. 97. Nashville, USA; 1997. p. 179–186.
- [134] Bader-El-Den M, Teitei E, Adda M. Hierarchical classification for dealing with the class imbalance problem. 2016;.