Article

Surviving burn injury: Drivers of length of hospital stay

Chimdimma Noelyn Onah [1,*], Richard Allmendinger [1], Julia Handl [1] and Ken Dunn [2]

[1] Alliance Manchester Business School, University of Manchester; [Chimdimma.onah@postgrad.manchester.ac.uk, Richard.allmendinger@manchester.ac.uk, Julia.handl@manchester.ac.uk]

[2] University Hospital South Manchester; ken.dunn@mft.nhs.uk

* Correspondence: Chimdimma.onah@postgrad.manchester.ac.uk; +447856976161

Abstract: With a reduction in the mortality rate of burn patients, patient length of stay (LOS) is increasingly adopted as an outcome measure. Some studies have attempted to identify factors that explain a burn patient's expected LOS. However, few have investigated the association between LOS and a patient's mental and socioeconomic status. There is anecdotal evidence for links between these factors and uncovering these will aid in better addressing the specific physical and emotional needs of burn patients, and facilitate the planning of scarce hospital resources. Here, we employ machine learning (clustering) and statistical models (regression) to investigate whether a segmentation by socioeconomic/mental status can improve the performance and interpretability of an upstream predictive model, relative to a unitary model derived for the full adult population of patients. Although we found no significant difference in the performance of the unitary model and segment-specific models, the interpretation of the segment-specific models reveals a reduced impact of burn severity in LOS prediction with increasing adverse socioeconomic and mental status. Furthermore, the models for the socioeconomic segments highlight an increased influence of living circumstances and source of injury on LOS. These findings suggest that, in addition to ensuring that the physical needs of patients are met, management of their mental status is crucial for delivering an effective care plan.

Keywords: Burn care, length of stay, mental state, socioeconomic status, clustering, predictive models, regression analysis, collaborative decision making.

---

1. Introduction

Globally, burn injury is one of the top ten causes of trauma and a leading contributor to the years of life lost from premature death, and years lived in less than full health [1]. Burn patients represent one of the most complex types of patient groups due to [2]: (i) the trauma of the injury, (ii) the prevalence of pre-existing conditions and (iii) occurrence of complications during the hospital stay. The high incidence and complex aetiology of burn injuries necessitate constant evaluation for better care management, multidisciplinary intervention approaches, effective costing, funding, and prevention programs.

To better understand and manage burn patients, it is therefore vital to understand the factors that influence the recovery of patients. Past studies have identified factors such as total burn surface area (TBSA), age, presence of inhalation injury, comorbidities, complications and, in some cases, the patient's gender as estimators of the length of stay (LOS) [3]–[8]. A well-developed prediction model of LOS on admission would aid in understanding burn patients better through the identification and management of significant predictors. The predictive model would also support patient triaging, developing personalised care pathways, resource planning, patient, family and friends counselling on the expected date of discharge [5], [9]. For the model to be fully employable, it needs collaboration with relevant stakeholders (in this case, Burn Care Specialists) to ensure the use of relevant patient cases, factors and subsequently, the production of an explainable model.

This paper focuses on the effect of adverse mental state and a patient's social environment, represented by their deprivation score – a measure of poverty -, on an adult patient's LOS. Our central hypothesis is that recovery from a burn injury is not predominantly determined by the burn injury severity but is also associated with a patient's state of mind (mental health) as well as their social

environment, as captured by socioeconomic status. We focus on the adult population, as age is an unpredictable complicator of seemingly simple burn injuries [10], [11]. Thus, throughout the burn care pathway, there has been separate development of practices suitable for paediatrics and those fit for adults.

We use LOS as the variable of interest (outcome variable) in this study as it is considered reflective of injury-induced morbidity, the cost and quality of care [12], [13]. Furthermore, with a consistent decline in mortality rates [14], [15] and thus, more people surviving very severe burn injuries, there are increased challenges in managing LOS. For this reason, LOS is considered an important performance indicator and outcome measure. To ensure the explainability and reliability of the developed model, we adopt a collaborative decision-making process, involving frequent interaction with medical practitioners.

1.1. *The effect of socioeconomic factors on health*

The British Medical Association has highlighted that poverty and social inequality adversely affect physical and mental health [16]. Their report goes on to add that these inequalities have implications such as years of life lost, costs to the economy and healthcare challenges. Before the British Medical Association, Marmot and Allen argued that socioeconomic inequalities are key determinants of a patient's health [17]. These inequalities include material circumstances, the social environment, psychosocial factors, behaviors and biological factors, and are influenced by education, occupation, income, gender, ethnicity, and race [18]. A high socioeconomic status leads to accumulated class-related individual resources that translate into higher health literacy (healthier lifestyle) and prevention practices (lower incidence) [19]. Low socioeconomic status has been documented as a risk factor for injury-related hospitalisation, severity and injury-related mortality [20].

In burn care, there is established evidence of the effect of socioeconomic status on the incidence of burn cases [2], [21]. This effect is shown in the higher burn incidence in developing countries compared to developed countries [22], [23]. In England and Wales, this is also the case: Diguiseppi and colleagues investigated residential fires in inner London and found a higher incidence rate in the most deprived areas [24]. The first exploration of burn injury spanning England and Wales over 2003 – 2011 saw a potential form of health inequalities due to socioeconomic factors reflected by geographically location [25].

Some researchers have found an association between a patient's socioeconomic characteristics and LOS for general admission and other specific diagnoses [26]–[28]. Even though socioeconomic status has been shown to influence the incidence rate of burn injuries, there is only little (although growing) evidence of the influence of income, race, educational attainment, housing status and other socioeconomic factors on the outcome of a burn injury [29]–[32]. This slow growth of evidence has been due to a lack of data, which includes variables capturing a patient's socioeconomic status.

The Index of Multiple Deprivation (IMD) is a composite measure that combines different domains of deprivation to produce an overall relative measure of deprivation [33]. The deprivation domains are income, employment, health deprivation, education, barriers to housing, crime and living environment. As a relative measure of deprivation, it can tell us if one area is more deprived than another but not by how much. Also, there are not measured on an individual level but at the level of the individual's social environment, which may limit its ability to capture fine-grained, individual-level detail on socioeconomic status. Hussain and Dunn investigated burn mortality rate and looked at the effect of socioeconomic status using IMD as a proxy [31]. They found a significant increase in the risk of death with every quintile reduction in IMD score, where lower quintile scores indicated a higher deprivation.

Understanding the effect of socioeconomic status on LOS would unveil the relationship between a patient's living environment and the speed of burn recovery. Katanien highlights that class differences in health also imply differences in lifestyle [34]. Thus, understanding the relationship between socioeconomic status and burn recovery may shed light onto the patients' desire to get back to their living environment. Alternatively, the socioeconomic status could also determine the type of burn injury. Injury type is known as a critical determinant of burn severity which in turn influences

LOS. A better understanding of these possible influences of socioeconomic factors on burn recovery would ultimately aid in an improved care pathway leading to better patient outcomes.

1.2. How adverse mental health could impede recovery

Mental health can be considered a possible marker for resilience, which motivates our interest in understanding the association of a burn patient's mental health on their LOS. Anecdotally, `motivated' patients are believed to recover faster, and to make more gains than patients that are less enthusiastic for treatment. A study which systematically reviewed literature found evidence of the adverse effect of stress in impeding wound healing [35]. The review included studies on acute and chronic clinical wounds, experimentally created punch biopsy and blister wounds, and minor damage to the skin caused by tape stripping

In burn care, there is growing evidence of the adverse effect of mental health factors on burn injury [36]–[40], both in terms of incidence and recovery. Patients with pre-existing psychiatric illness may suffer more severe burn injuries, which occurs due to several reasons including confusion, carelessness, alcohol or drug intoxication or deliberate self-harm and attempted suicide [38]. The mechanisms impeding burn recovery among mentally impaired patients may range from decreased hope or motivation to participate in one's recovery to maladaptive behaviors, including aggressive acts and agitated movements that may disrupt treatment [41]. Crucially, it was identified that stress impedes the initial inflammatory phase of wound healing, and increases susceptibility to wound infection with impairment of cellular and humoral immunity [37], [41]. [38]. Furthermore, the mental state of patients may impact their lifestyles, such as poor diet, smoking and physical fitness, all of which has an impact on recovery [38].

1.3. The interaction of mental and socioeconomic status

A patient's willingness to return to their living environment – reflective of their socioeconomic status - and to participate in their recovery – reflective of their mental health - reveals the potential interrelationship between their mental health and socioeconomic status. As summarized by a meta-analysis of 56 research papers, Lorant and colleagues find evidence of an association of high psychiatric morbidity and low socioeconomic status [42]. Those found in the lower SES groups have a higher risk of poor coping styles, stress exposure and weaker social support [43]. According to the Marmot review, there is evidence of a relationship between a patient's socioeconomic status and their psychosocial and behavioral state, which affects lifestyle, social integration and self-esteem all of which are detrimental to recovery time [17]. This evidence highlights the potential interrelationship of a patient's socioeconomic status to their mental health and the influence both may have on a patient's LOS.

To investigate the research questions raised in Section 1.1, 1.2 and 1.3, we make three hypotheses as depicted in Figure 1 and summarized below.

Hypothesis 1. Segmentation by socioeconomic status or mental health state produces groups that are more homogenous with respect to LOS than the non-segmented population.

Hypothesis 2. We suspect an interrelationship between a patient's socioeconomic status and their mental health, and that both interactions have an impact on LOS.

Hypothesis 3. If Hypothesis 1 is supported, we expect that the relative influence of severity on burn patients reduces as the presence of adverse mental or socioeconomic state increases.

In evaluating the hypotheses, this study aims to understand, in the context of burn patients, the association of mental and socioeconomic status on length of hospital stay. We also study if there exists an interrelationship between mental health and socioeconomic status. We conduct this study by adopting a collaborative decision-making process involving the Manchester Centre for Plastic Surgery and Burns, part of the University Hospital of South Manchester. We refer to the collaborators going forward as Burn Care Specialists.
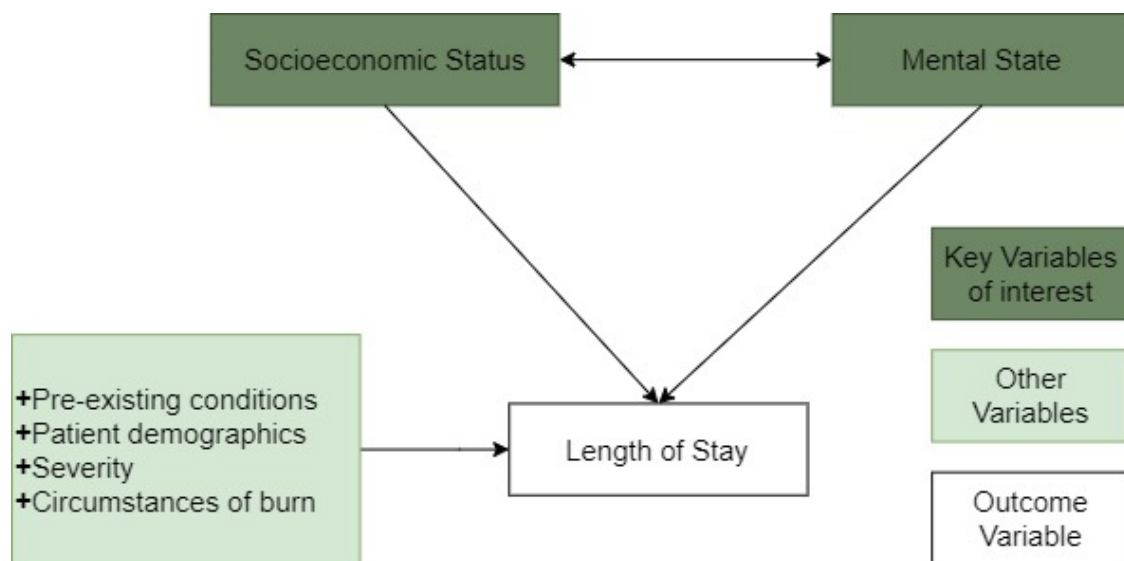
Figure 1. Relationship between socioeconomic status, psychological state, and other factors on LOS.

2.   Materials and Methods

2.1. Data

This study uses anonymised patient-level data from all burn units in England. The data covers the period, 2003 - 2019, comprising just over 50,000 admitted patients and was extracted from the international Burn Injury Database (iBID) [44]. The patient information is collected by clinicians and nurses, from the first contact with a burn service through to rehabilitation and any late reconstruction procedure [25]. The dataset includes all types of burn injury patients who attended a burn service. The dataset excludes burn injury patients that died before reaching a burn service and those with minor injuries which are cared for in the community and in hospitals that are not specialist burn services.

We focus on the adult population (age >= 16) as pediatric patients are more likely to be subjected to unpredictable complication, even with seemingly simple burn injuries, as highlighted by the 2001 National Burn Care Review Report [10]. The report argues and mandates the need for separate burn units for children and adults, due to the peculiar needs of children such as play specialist, teachers, family counsellors and intensive psychosocial support. Thus, the burn care pathway is designed to treat pediatrics separately from adults as young age is a significant complicator. In our previous study in segmentation of burn patients, we found further evidence that adult and pediatric burn patients are characterized by different variables (such as the source of injury and presence of existing disorder) and have different resource requirements [11]. Consequently, to allow for more robust predictions, in this study, we discard all cases of age less than sixteen. The cases removed also include non-survival cases as we expect them to have shorter LOS than counterparts who survive. Accurate referral of patients to burn units implies that there is an increased number of admitted minor injury cases which were previously treated in other care settings. Therefore, given our focus on predicting LOS on admission, those with no overnight stay are excluded.

2.2. Variables

As stated above, this study considers LOS as the key indicator of outcome due to its ability to reflect injury associated morbidity, cost, and efficiency. We calculate LOS as the number of days a patient is admitted for (days from admission to discharge). LOS was log-transformed to remove skew. The iBID database [42] gives us access to about 300 variables.

Our goal is to understand what variables drive the prediction of LOS at the point of admission as this will allow for improved patient triaging, resource planning, and smoother patient discharge. We exclude patient information obtained after admissions, such as complications and operation duration. Compared to a model that captures the full pathway of a patient, this exclusion will likely lower prediction accuracy, but our focus here is prediction and prediction performance at the point of admission. Other variables were removed with guidance from a Burn Care Specialist, and this

Table 1: Variables used in this study grouped by type, and an indication count for categorical variables

Type of variable (count)	Variables (number of categories in the categorical variables)
socioeconomic status (1)	IMD.
mental state (17)	self-inflicted injury, injury from suicidal attempt, number of existing psychological order, alcohol abuse, substance abuse, psychiatric disorder, depression, post-traumatic stress disorder (PTSD), anxiety, personality disorder, mania, schizophrenia, self-harm, attention deficit hyperactivity disorder (ADHD), eating disorder, learning difficulty, dementia.
pre-existing conditions (47)	non-insulin-dependent diabetes mellitus, insulin-dependent diabetes mellitus, epilepsy, asthma, bronchiolitis, upper respiratory tract infection, urinary tract infection, none specific viral illness, chronic obstructive pulmonary disease, emphysema, preadmission bacteria, perivascular disease, ischemic heart disease, past myocardial infarction, hypertension, cardiac failure, congenital heart disease, heart valve disorder, cardiac dysrhythmia, neoplasm, metastatic disease, past cerebrovascular, hemiplegia, past deep vein thrombosis, poor hearing or deafness, poor eyesight or blind, mild hepatic dysfunction, moderate to severe hepatic dysfunction, renaldys function, cerebral palsy, paraplegia, spina bifida, motor disability immobility, sensory disability, multiple sclerosis, 1st trimester – pregnancy, 2nd trimester – pregnancy, 3rd trimester – pregnancy, hepatitis b, hepatitis c, human immunodeficiency virus, anemia, clotting disorder, immune suppression, leukemia lymphoma, resus, previous complications.
demographic factors (2)	gender (2), age.
severity (31)	total burn surface area (TBSA), per cent superficial burn, per cent deep dermal burn, inhalation symptoms, body area burn occurred (face, neck, scalp, anterior chest, posterior chest, right shoulder, left shoulder, right upper arm, left upper arm, right forearm, left forearm, right-hand dorsal, right-hand palmer, left hand dorsal, left-hand palmer, abdomen, lumbar back, buttocks, perineum, right thigh, left thigh, right lower leg, left lower leg, right foot dorsal, right foot sole, left foot dorsal, left foot sole).
circumstances of burn (9)	injury type (9), source of injury (95), activity (18), category (16), living space (15), locality (24), first aid received (15), time from injury to admission (7), living circumstances (12).

includes variables with a high number of missing values, administrative variables and variables which gives information obtained from other variables. Further variable selection is done using LASSO as described in Section 2.3.

Through collaborative decision making with a Burn care Specialist, the variables were also grouped into similar types. In total, for the 107 significant variables, we identified six variable groups as summarised in Table 1. Four of these variable groups cover pre-existing conditions, demographic factors, and the severity and circumstances of the burn, respectively. These injury attributes provide important insight regarding the source of the burn, the type of burn injury sustained, the activity engaged in when the burn occurred, locality, living circumstances and space where the injury occurred, time from injury to admission and if first aid was administered. The circumstance of burn also includes an indicator of injury occurrence (for example, through assault or work-related) named injury category. Finally, the two main variable groups of interest in this study are those that relate to socioeconomic status and mental health. IMD serves as a proxy for socioeconomic status, while a group of seventeen variables serve as proxies for mental status. More information on these variables is provided below.

- **Socioeconomic Status:** Socioeconomic status must be appropriately operationalised to allow analyzing its impact. There are two approaches to studying the influence of socioeconomic status on health disparity: the compositional approach, which measures socioeconomic status by referring to characteristics of the individual, and the contextual approach, which measures socioeconomic status as the characteristics of an individual's environment [45].

  In this research, we have access to a contextual measure of socioeconomic status, which is the IMD index. This index combines variables of income, employment, health, education, crime, housing, access to services and living environment. As noted by Kaplan, it is paramount to study the social environment to understand the day-to-day experience of individuals and groups and how that translates into better or worse health among the richer and poorer [46]. This is also reflected by the World Health Organization report on health inequalities, which discusses the potential influence of the environment, in which people grow, live, work, age, and the systems put in place to deal with illness, on health inequalities [47]. The IMD captures the above aspects, with deprivation defined as the lack of access to opportunities and resources [48]. The iBID database holds the 2004 IMD score, which is at a local authority district level [33]. The score ranges from 0 to 100, with 100 indicating extreme deprivation. The IMD is published by the Office of National Statistics every 3-4years, however, available to us via the iBID database is the 2004 version.

- **Mental State:** Disorders of the mental state may be caused by changes of biological factors, such as changes of the brain, aberrations of the immune system, and changes in psychosocial factors, such as effects of interpersonal relations and negative body image [49]. These disorders affect a person's thinking, feeling and behaviors.

  In this study, we use some of the major indicators of an individual's mental health, as highlighted by WHOs ICD-10 Classification of Mental and Behavioral Disorders Addiction [50]. The relevant variables used in this paper indicate the presence/absence of certain mental disorders before the burn injury, i.e. pre-existing and others captured by identifying categories of burn linked to adverse mental health. We consider fifteen pre-existing mental health variables: alcohol abuse, substance abuse, psychiatric disorder, depression, post-traumatic stress disorder (PTSD), anxiety, personality disorder, mania, schizophrenia, self-harm, attention deficit hyperactivity disorder (ADHD), eating disorder, dementia and learning difficulty. The pre-existing mental conditions are captured in the iBID through patient indication and ascertained using access to available health records. Two other adverse mental state variables were identified from the category of burn occurrence: was the burn caused by self-infliction or suicide attempt.

2.3. Analytical Pipeline

To ensure the developed model is explainable and reliable, while predicting LOS accurately, we developed an analytical pipeline that entails a collaborative and iterative decision-making process. The analytical pipeline includes data pre-processing, segmenting patients by their socioeconomic and mental status, and building and testing of least absolute shrinkage and selection operator (LASSO) regression models. To ensure result validity and reliability, the analytical pipeline includes *k*-fold cross validation. This involves the training and evaluation of the model across *k* random splits of data into test (10per cent) and training (90per cent) datasets, to understand the variability of evaluation metrics and coefficients of significant predictors across these different "folds". The test dataset is a hold-out test data set used to test the reliability of results, i.e. the predicted LOS is compared to the expected LOS to identify the model's ability to predict the LOS for newly admitted patients accurately.

As indicated in Figure 2, the first step is data cleaning and pre-processing, which includes removing cases and variables that are not relevant as detailed in the Data and Variables Section. With insight from a Burn Care Specialist, cases with TBSA over 80 per cent but have a LOS under five days are flagged as outliers, as it is expected that such cases should have a higher LOS. The data pre-processing step also contains one-hot encoding, where categorical variables are converted into binary variables, excluding the first category in each variable to avoid the dummy variable trap. As a result of one-hot encoding, the number of variables available for model building increases from 107 variables to 308.
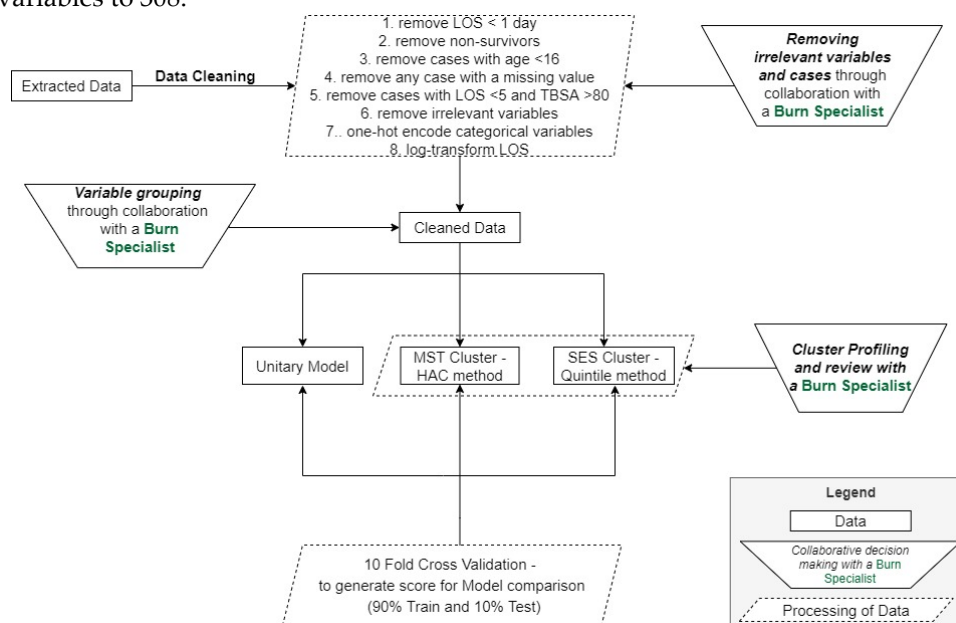

Figure 2. Analytical pipeline showing data cleaning/pre-processing to model deployment and testing.

To explore our hypothesis of increased homogeneity in terms of LOS, two segmentations were created, and predictive models developed separately for the identified segments. The first segmentation uses the socioeconomic status variable (SES-based clusters), and the second segmentation uses all seventeen variables that reflect a patient's mental health (MST-based clusters). For comparison, a predictive model with all adult patients (no segmentation) was created (Unitary model).

Due to the availability of a single segmentation variable, the quintile method -- segmentation of data into five equal parts -- is adopted for the SES-based cluster set. This method allows for segments that are explainable. It is consistent with approaches found in the literature and used by government bodies, such as the Office for National Statistics and Public Health England, where IMD is typically grouped by dividing the sample into five parts: 0-20, 20-40, 40-60, 60-80 and 80-100 [31], [51], [52]. However, as IMD is a standardized measure that reflects the relative level of deprivation of an area, the parts are only equal when segmenting population data. In line with this, we segment cases into a

group of five; the most deprived group is IMD score >=80, and the least deprived group is IMD score <20. Given that the burn care data does not reflect the general population, we would have unequal group membership across segments.

We identified the MST-based clusters using all variables that reflect mental health. The group membership was determined using a clustering algorithm as there was no prior knowledge regarding an appropriate cluster number nor an established membership criterion for creating homogenous groups in terms of mental health. Clustering is an unsupervised learning technique for partitioning sets of data items into several disjoint clusters based on a similarity measure [53], [54]. In particular, this paper uses the average link hierarchical agglomerative clustering (HAC) method [53], a well-established technique in unsupervised machine learning. Previously used in our research on segmentation of burn patients [11]. The average link linkage criterion was chosen as it is appropriate for noisy data and reduces the probability of uneven clusters. In average link, homogenous clusters are obtained by calculating the distance between clusters as the average of the pairwise distances between all observations in the two clusters.

One of the advantages of HAC for our application is its ability to work with any distance metric, to identify the similarity between data points [55]. In our case, the distance between cases (patients) is measured using the Gower metric, which was developed for datasets containing a mix of categorical, binary and numerical variables [56]. For each variable type, it uses the most appropriate distance metric. For the MST-based clustering, given that the variables are all binary, the Dice coefficient is used to calculate the distances.

Given the absence of prior knowledge of the number of clusters in the dataset, model selection is supported using internal validation variables and the interaction with Burn Care Specialists. First, we use the Silhouette and Calinski-Harabasz score to determine the most promising number of clusters. The Silhouette score of a data instance measures how similar an object is to elements in its cluster compared to elements in the nearest cluster [57].   The Calinski-Harabasz score, also known as the variance ratio criterion, is the ratio between the within-cluster dispersion and the between-cluster dispersion [58]. The partition identified at this stage were passed on for further validation to a Burn Care Specialist to ensure that segments are explainable and clinically relevant.

Due to the potential presence of multicollinearity in the data, for example, the presence of numerous variables that represent severity, LASSO [57] was selected to build the models. LASSO is a type of linear regression, differentiated by its regularisation characteristics [59]. In particular, we use the least angle regression (LARS) LASSO [60]. LARS LASSO is an efficient stepwise variable selection algorithm. As a regularisation algorithm, it uses a tuning parameter to shrink coefficients of the non-significant variables to zero, thus removing non-significant variables from the model. Where there are highly correlated variables, it tends to choose one variable only [61]. Therefore, the method is more likely to prevent overfitting, which is another motivation for using LARS LASSO. For our LARS LASSO model, to ensure low bias and variance in the model, 20-fold cross validation was used to ascertain the best value of the tuning parameter, lambda, for the regularisation process.

The LARS LASSO algorithm was fitted on the two cluster sets and the unitary model:

- SES-based clusters: the socioeconomic variable used for clustering was excluded as an independent variable in the LASSO model. Thus, all SES-based models included 307 independent variables
- MST-based clusters: In this cluster set, all mental health variables were excluded. Thus, the MST-based models were fitted using 291 independent variables.
- Unitary model: The LASSO model fitted for the unitary model contained all 308 independent variables, plus three interaction terms. The interaction terms were included to enable the evaluation of how the presence of key aspects (mental health, socioeconomic status and burn severity) is associated with LOS. The variables chosen to represent each of those aspects were selected by drawing on literature. For socioeconomic status, IMD was used while TBSA was used for severity. To accurately represent mental status, a new variable - *MST_count* - was created that represents the count of the adverse mental health including pre-existing mental disorder and the other two mental health variables identified from the category of burn. The three interaction terms are *TBSA*IMD, TBSA* MST_count* and *IMD* MST_count*.

Given the adoption of 10-fold cross validation models, the generated coefficients from each model were averaged to identify the expected population coefficient. Subsequently, the effect size (which expresses variable importance) was obtained for each variable by standardising the coefficients from the LASSO model. The coefficients from multiple linear models represent the relationship between the given variable and the target variable, assuming all other variables are constant. With the inclusion of variables with different units, there is a need to standardise the generated coefficients. This is done by multiplying each variable's coefficient by the standard deviation of the variable. This means that the greater the variance of a variable then the larger the standardized coefficient (and therefore the impact of the variable on the output), all else being equal.

Model fit and performance were ascertained using R-squared (R^2), and weighted root mean square error (*wRMSE*) on the training and test data. To achieve objective comparison across segmentations, an adapted root mean square error measure was devised. We calculated the weighted sum of RMSE across all clusters, with weights reflecting the proportion of the data covered by each cluster (Equation 1).

$$wRMSE = \sum_{k=0}^{K} \frac{n_k}{n} \sqrt{\frac{1}{n_k} \sum_{i=1}^{n_k} (Y_i - \hat{Y}_i)^2},$$ (1)

here n_k is the sample size in cluster k, n is the total sample size of all clusters in a cluster set and $\hat{Y}_i$ is the predicted value and Y_i the observed value.

To understand the proportion of the variance in the dependent variable that is predictable from the independent variable, we measure R^2 (Equation 2).

$$R^2 = 1 - \left[\frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y}_i)^2}\right],$$ (2)

where $\bar{Y}_i$ is the mean of observed values in the training data.

3. Results

In total, 19,474 admitted adult patients with at least an overnight stay were analysed. There is a consistent decline in average LOS over time (Figure 3). A reduction in the number of admitted patients is seen from 2011.
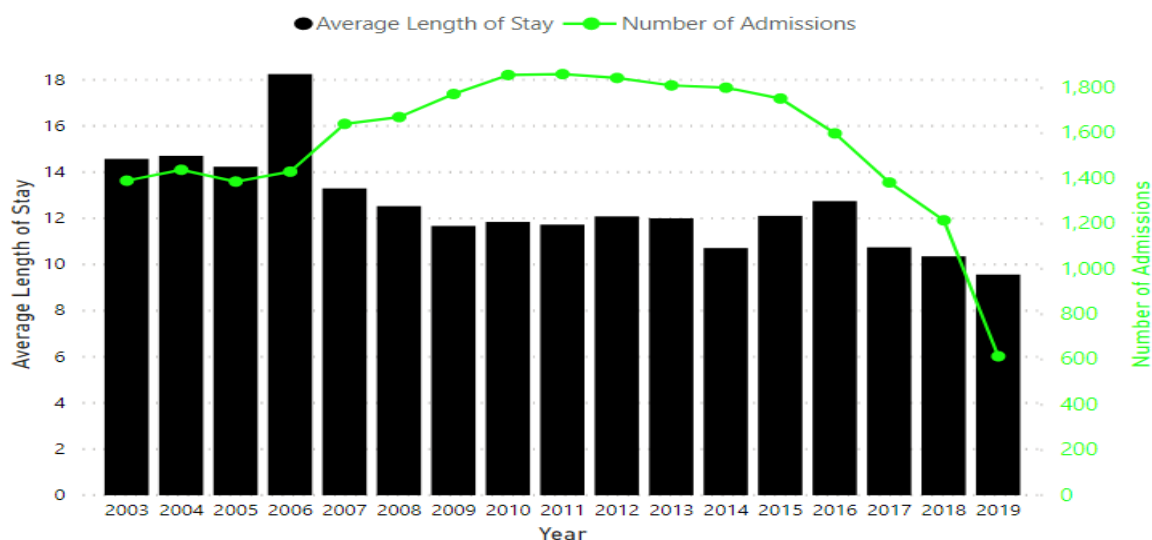

Figure 3: Number of admissions and change in the average length of stay from 2003 - 2019

Our initial data exploration provides some first insight into the variability in LOS and key drivers behind it. The LOS of adult patients ranges between a day and 455 days with a median of five days and an interquartile range (IQR) of eleven days. This high variability highlights the potential benefits of constructing patient segments. We find that there are more males admitted (68 per cent) than females (32 per cent), with the median LOS for women at seven days higher than that of males at five days. The age of admitted adult cases ranges from 16 to 113 years old, and LOS varies with age: the older the patients, the higher the median LOS.

Of 19,474 adult patients, 20 per cent have one or more pre-existing mental condition. Patients with a mental health related pre-existing condition have a median LOS of 10 days – the lowest median

was found for patients with PTSD at five days and the highest amongst those with dementia at 17 days. The higher the number of pre-existing conditions, the higher the median LOS except for those with 7 or 9 disorders with a median LOS of seven days and five days, respectively. Patients with one to four pre-existing mental health disorders have a median LOS of 10 days, while those with no pre-existing mental health conditions have a median LOS of 5 days (see Figure 4a).
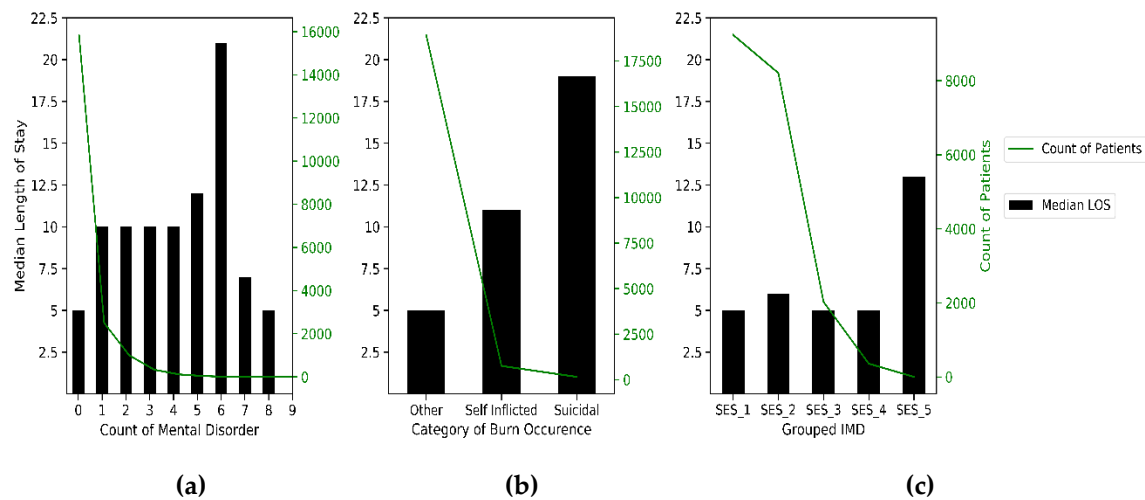

Figure 4. Patient profile by key variables of interest: (**a**) This shows varying median LOS by the number of pre-existing mental disorder; (**b**) This shows varying median LOS by category of burn grouped into two mental state and other (non-mental state-related); (**c**) With deprivation increasing from left to right, this shows varying median LOS by socioeconomic status (SES).

Exploring how the burn injury was inflicted, 95 per cent of cases are caused accidentally or through assault and have a median LOS of five days (Figure 4b). Patients with injury infliction from suicide attempt accounted for 1 per cent and had a median LOS of 20 days, while those due to self-infliction accounted for 4 per cent, with a median LOS of 11 days. Exploring IMD, where 0 IMD indicates relatively no deprivation and 100 extreme deprivation, there is a slight increase in LOS as IMD score increases and a significant correlation with LOS. From Figure 4c, median LOS is the same at five days for the first, third and fourth SES quintile groups, and six in the second group (20 - 39 IMD score). As expected, the median LOS is at the highest for the most deprived group (SES_5), with a value of 13 days.

These striking differences suggest support for our hypothesis that segmentation by socioeconomic status or mental health produces groups that are more homogenous with respect to LOS than the non-segmented population. We now proceed to consider whether a segmentation by socioeconomic status and another by the mental health of patients can reduce LOS variability and result in better prediction compared to the unitary model. For the SES-based cluster set, group membership was identified by quintile classification of IMD. Given the IMD range of 0 – 100, this creates five groups. The first being the least deprived group (IMD<20), the second group has IMD of 20 to 39, the third is 40 to 59, fourth is 60 to 79 and the fifth group has the most deprived cases with IMD >=80. Though it is expected when analysing the population in England that these quintile groups should have equal membership, due to low admission into burn care units from the two most deprived group, we see low case membership in SES_4 and SES_5. Thus, to avoid creating a heterogeneous group in terms of deprivation indicators, cases assigned to these group were removed from the analysis altogether.

As discussed in the Analytical Pipeline Section, the MST-based cluster set was identified by adopting an approach which allows for identification of an optimal number of groups and cluster sets in terms of the seventeen mental state variables. The most suitable number of groups is identified from examining Figure 5, which shows that at $k = 10$, group separation on mental state variables is at its maximum.
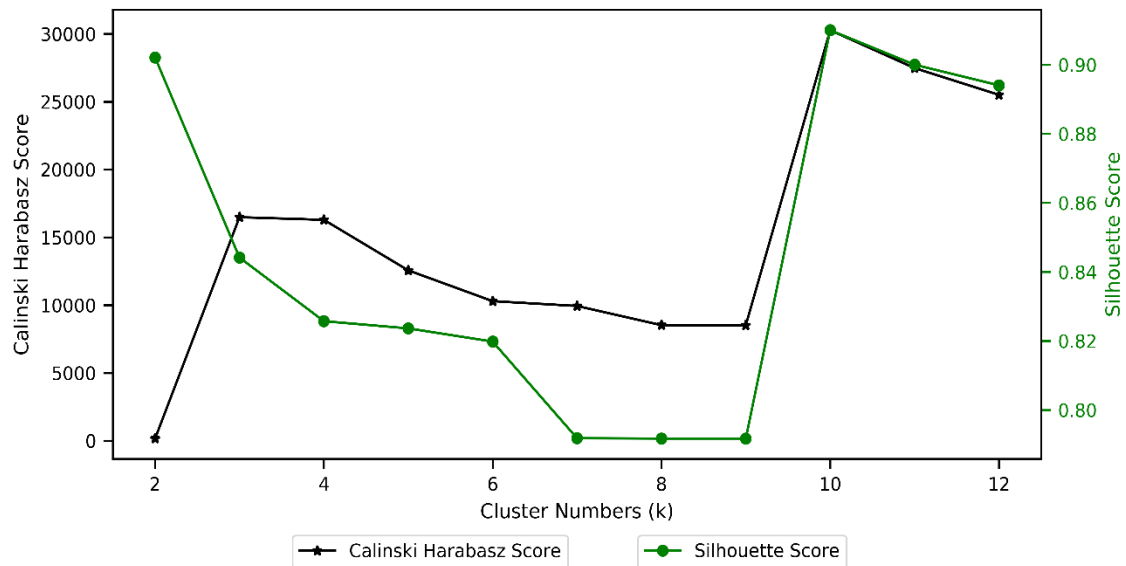
Figure 5: *k*-selection using Calinski-Harabasz and Silhouette Score.

Cluster characterisation further showcases a clear separation and easy to describe grouping. The first group comprises all cases with no mental health-based disorder, and the second group comprises those with one mental health-based disorder, which is either a pre-existing disorder or derives from the category of injury infliction. The third group comprises those patients with two mental health disorders - either as two pre-existing mental conditions or as one pre-existing mental condition and injury infliction by a suicide attempt or self-infliction. The remaining seven groups comprise patients having three or more mental health disorders, with each group distinct by the increasing number of mental health conditions present. However, these groups are small, which means that LASSO modelling is not feasible as the number of variables exceeds the group membership.

A Burn Specialist was, therefore consulted to review the group membership. A collaborative decision-making process led to the adoption of a less granular two-group structure, representing the presence or absence of mental state disorder, respectively. This segmentation was judged to be easy to interpret and clinically relevant by the burn specialist. Thus, for this paper, the small clusters (2,3,4,5,6,7,8,9,10) are merged, to produce a single cluster of adverse mental state.

In summary, the segmentation steps produced $k = 10$ (with $k >= 2$ classes merged) for the MST-based segmentation and $k = 5$ (with the fourth and fifth cluster excluded from LASSO model building due to low membership) for the SES-based clustering. The final groups and their characteristics are

Table 2: Profile of Identified Clusters by average LOS, TBSA and Age

Cluster (count)	Characteristics	Average LOS (SD)	Average TBSA (SD)	Average Age
MST_1 (15,418)	No mental based condition	10 (19.42)	6 (7.73)	45 (19.74)
MST_2 (4056)	Presence of mental based condition	20 (32.19)	9 (13.65)	47 (15.01)
SES_1 (9238)	Least deprived (IMD <20)	11 (19.82)	6 (8.48)	47 (19.83)
SES_2 (8,209)	IMD between 20 and 39	12 (22.14)	6 (8.53)	44 (19.00)
SES_3 (2,027)	IMD between 40 and 59	12 (22.28)	6 (8.57)	42 (18.3)

shown in Table 2. With regards to LOS and TBSA, the two MST-based clusters are different as its

mean, and standard deviation reflects different LOS, TBSA and age. We see an increased average LOS as the number of the adverse mental state increased. This clear distinction across clusters is not found in SES-based clusters, most likely due to IMD not being an individual-level measure of deprivation. Having identified the MST-based and SES-based patient groups, these and the unitary adult group were used to build our LASSO models. The interpretation of the model coefficients will reveal if the impact of severity reduces with an increased presence of adverse socioeconomic/mental status, depicted by decreased severity effect size.   Similarly, the significance of the interaction terms will provide insight regarding the presence of an interrelationship between severity, mental state, and socioeconomic status.
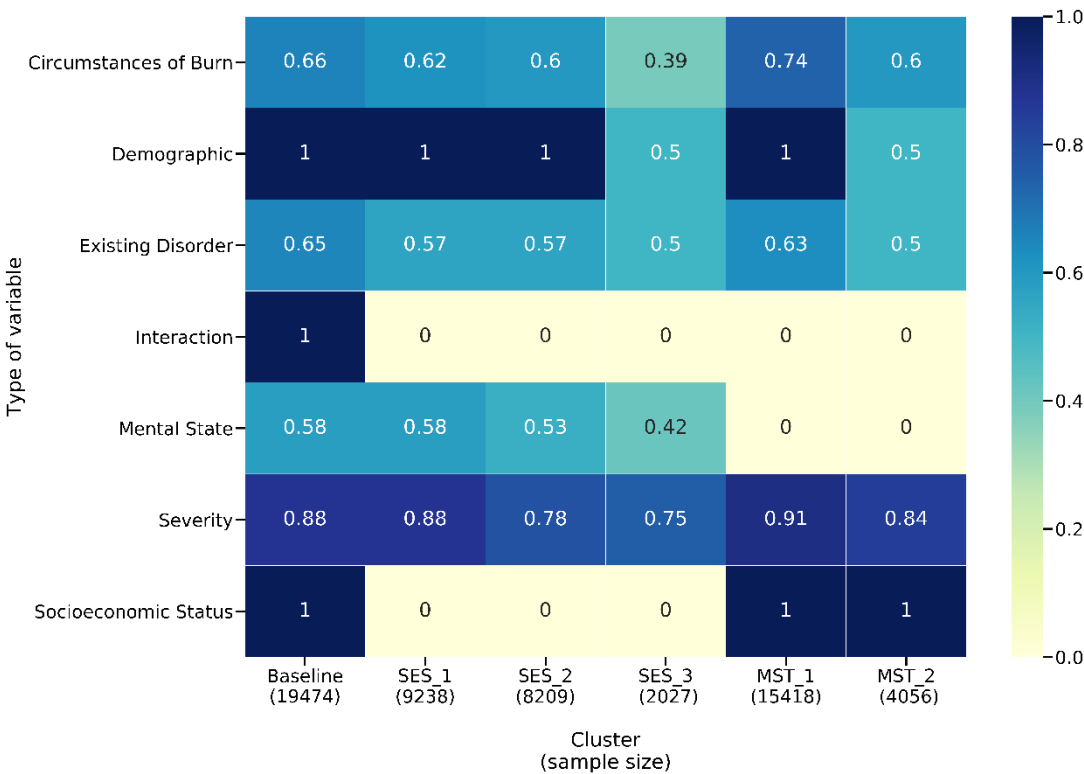

Figure 6: Heatmap showing the proportion of variables that have non-zero coefficient in each type for MST-based, SES-based and unitary models.

The results of the unitary model indicate 212 significant variables. Evaluating the summed effect of the severity variable group, as expected, it had the highest sum compared to all other variable groups at 0.92. The majority of variables describing severity (except for indicators of burn occurrence in the face or to the right forearm) have a positive effect size, indicating higher severity is associated with a longer LOS. In total, 88 per cent (Figure 6) of severity variables have a non-zero coefficient. We found that eight of the top fifteen variables which predict longer LOS in the unitary model are severity variables (see Figure 7). As might be expected, TBSA is the most important predictor of burn. Existing disorder variable type has 65 per cent of it as significant and two variables included in the top fifteen contributors to longer LOS. As expected, increased age leads to an elevated risk of increased LOS and is the second most important predictor of LOS.

Focusing on the aspects of interest in this paper, the results show that IMD, and therefore 100 per cent of variables related to socioeconomic status, and 58 per cent of the mental state variables (depression, personality disorder, schizophrenia, PTSD, substance abuse, psychiatric disorder, anxiety, dementia, learning difficulty, alcohol abuse and count of pre-existing mental disorder) are determined to be significant predictors of LOS. IMD's effect size is 0.02. In comparison, mental state variables had a summed effect of 0.14. Many of the mental health variables have a positive effect size, and two of them are amongst the top fifteen predictors of longer LOS. In the unitary model, all variables capturing a patient's demographic (age and gender), are also significant predictors of LOS.

The three interaction terms, MST_count*TBSA, MST_count*SES and SES*TBSA, are all significant predictors, providing evidence for our second hypothesis. As the significant interactions show evidence that the effect of a patient's mental health state on LOS is different for different values of socioeconomic status and vice versa. Similarly, the effect of TBSA on LOS varies with varying values of socioeconomic and mental health state. The MST_count*TBSA coefficient reveals a negative relationship with LOS, which reveals that the effect of TBSA (severity) on LOS will decrease with the increased presence of mental health disorders.
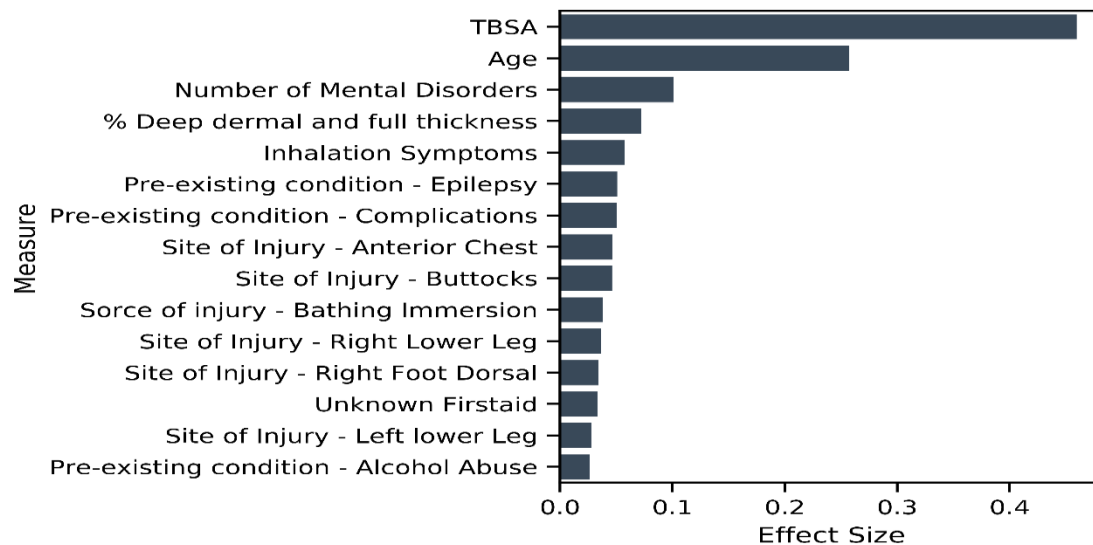

Figure 7: Variables of importance – top 15 contributors from 212 significant variables in the unitary model.

To investigate our third hypothesis, we compare the cluster-specific models, exploring the proportion of variable types that are significant predictors of LOS. Across all models – unitary (Figure 7) and cluster wise (Figures 8 and 9) – TBSA and age are the two most important predictors of LOS. Inhalation symptoms, percentage of full dermal burn and pre-existing complications also appeared as a top fifteen predictor in all the models though in varying order. All clusters, including the unitary model, have a proportion of severity, pre-existing disorders, demographic factor and circumstances of burn as a significant predictor.

For SES-based clusters, the proportion of significant predictors across the variable types varies across the cluster set, indicating that deprivation quintiles differ in terms of the impact of specific variables on LOS. More specifically, the proportion of variables significant in each variable type reduces as deprivation increases. Most notably is severity, mental state and existing disorder. For the most deprived group – SES_3 – seven of the top fifteen variables of importance are not found in the unitary model (Figure 7 and Figure 6c). These seven variables are related to living circumstances and the source of injury. SES_2 has all variables found in the unitary model included within the top 15 predictors, except for the variables that indicate the living space in which the burn occurred (Figure 7 and Figure 8b). SES_1 was mostly similar to the unitary model; the difference in both models is that SES_1 has a variable indicating living circumstance as significant (Figure 6 and Figure 7a).

In the MST-based clusters, we observe a reduction in the proportion of significant severity variables as the extent of mental disorders increases (91 per cent for no mental disorder and 84 per cent for mental disorder group – see Figure 6). This pattern is found across all variable types except for IMD, which is a significant predictor in both MST models but has a higher effect size in MST_1
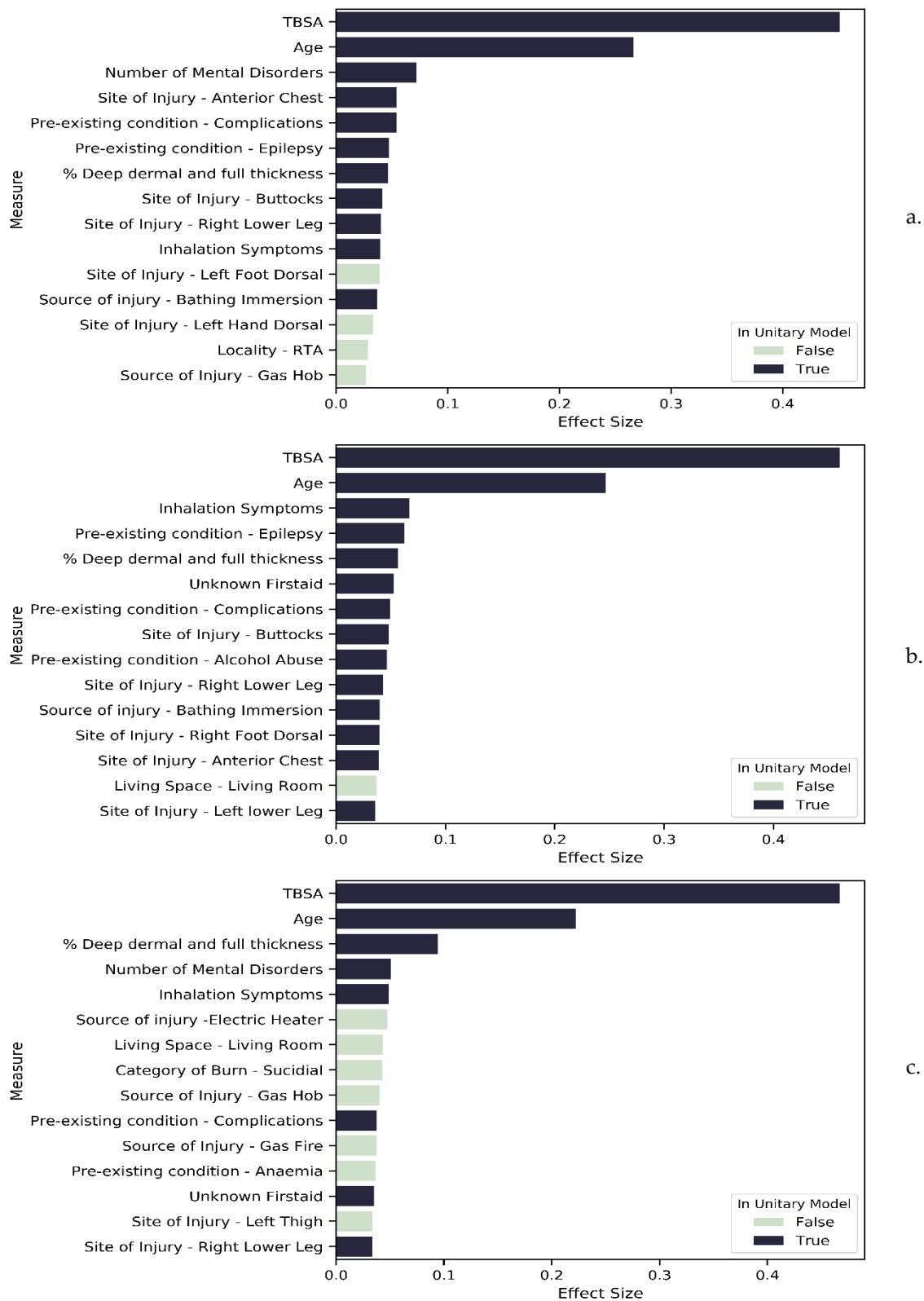

Figure 8. Variables of importance: (**a**) top 15 contributors from 196 significant variables in SES1; (**b**) top 15 contributors from 188 significant variables in SES2; (**c**) top 15 contributors from 137 significant variables in SES3.

compared to MST_2. For MST_1 and MST_2, the ten most important predictors are predominantly variables that reflect severity, age inhalation symptom, previous complications, except for MST_2 where locality of burn is the tenth most important predictor (Figure 9a and b).
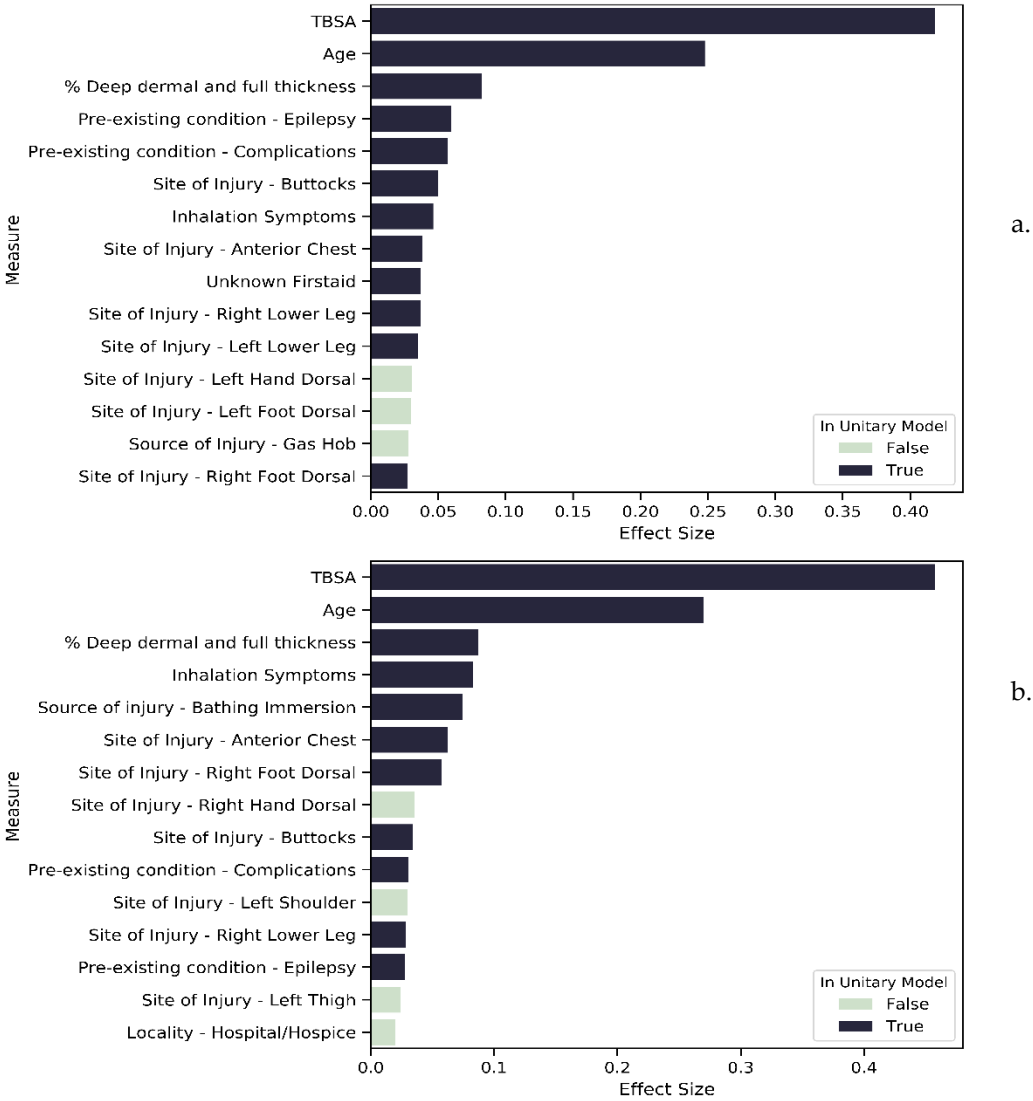

Figure 9. Variables of importance: (**a**) top 15 contributors from 215 significant variables in MST1; **(b)** top 15 contributors from 177 significant variables in MST2.

Comparing the models using wRMSE, we find no significant difference in the performance of the cluster wise compared to the unitary model (Figure 10a). This is an important finding, as it suggests that the cluster wise models have captured the most important influence of mental health and socioeconomic status. With mental health and socioeconomic status significant in the unitary model, prediction accuracy of the cluster wise models were not impacted when these key variables were removed. The cluster wise models bring clear benefits in terms of the complexity and explainability of the resulting regression models, and the potential to exploit the underlying segmentation in a personalised care pathway and creation of reimbursement models. Except for MST_1, the unitary model has the highest number of significant variables, whereas the cluster wise models give rise to simpler models, highlighting the improved homogeneity of these patient groups. This is also evident from Figure 10b, which illustrates that, except for MST_1, the cluster-specific regression models have a higher R^2 score than the unitary model, indicating a higher ability to explain the variability in LOS through a regression function.
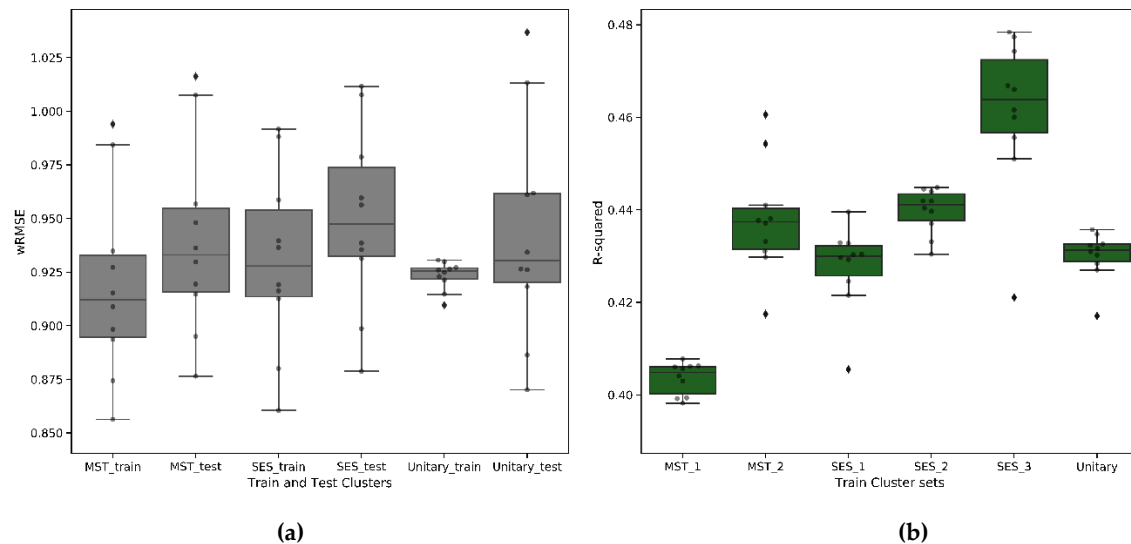
Figure 10. Evaluation metric: **(a)** wRMSE from each k-fold models (train and test) averaged across a cluster (b) R-squared on train models on each cluster set.

4. Discussion

In this study, IMD and a set of mental health variables (seventeen in total) were examined to ascertain the association between these factors and LOS. These variables include pre-existing mental health conditions in patients and cause of injury due to self-infliction or suicide. IMD was used as a measure of socioeconomic status, capturing contextual variables of deprivation. The results of our baseline model support previous research, highlighting that TBSA, age, inhalation injury and depth of burn are the most important predictors of burn [62]–[66]. We further provide evidence for the association of mental state variables with LOS, showing that the number of mental state conditions is a significant, top three predictor of LOS.

For mental state, its influence on LOS was observable in the exploratory analysis as depicted in Figure 4, in the form of a trend of increased median LOS with an increased presence of adverse mental health in a patient. Median LOS for injury caused by a suicide attempt or self-infliction was found to be higher for these patients compared to those caused by accident or assault. This exploratory analysis reveals that patients with adverse mental health are more likely to spend a long time admitted to hospital than patients without these conditions. This finding is in line with anecdotal evidence and with previous research on the impact of adverse mental/psychological state on the LOS of burn patients [37]–[39], [41], [67]. Our research builds on this previous work in two ways. Firstly, we use a larger sample size compared to previous research. Secondly, with the adoption of cluster wise models, we explored how the impact of variables on LOS changes with different levels of adverse mental health status.

The MST-based cluster result gives better insight into the relationship between LOS and mental disorder and its interrelationship with socioeconomic status. A clear trend is shown: with an increasing presence of mental disorder, there is a reduced number of burn severity variables that significantly predict LOS. The cumulative effect size reveals a higher association of LOS and severity variables in the no mental disorder group compared to the groups with any form of mental disorder. This begins to answer Hypothesis 1 and 3, indicating greater homogeneity in groups of patients with similar mental state compared the general adult burn patient population. The resultant higher cumulative effect size of severity variable in the no mental disorder group compared to those groups with any level of mental disorder further evidence the need for grouping to obtain better homogeneity. Therefore, severity, as well as mental disorders, need to be triaged effectively to enable better injury management and outcomes. Our results corroborate Tarrier's and colleagues who found that patients with pre-existing mental conditions and self-inflicted injuries are associated with difficult clinical management and higher LOS [38]. The analysis of the cluster profile (Table 2)

revealed an association of increased mental disorder to increased TBSA; this is further evidenced by the significance of the interaction term that accounts for these two variable types. Thus, revealing a potential interrelation with these two variables and the need for a personalised triaging process which takes into account a patient's mental state.

The SES-cluster based results show a similar trend. The least deprived group have the highest proportion of severity variables as significant, and the most deprived have the lowest number of severity variables as significant. We also found that in the SES-based cluster, variables that reflect the mental health of patients were found as significant predictors. This may be linked to a potential interrelationship between mental state and socioeconomic status. Further evidenced in the unitary model where the interaction term, of mental state and socioeconomic variables, has a significant positive effect, revealing that with an increased presence of adverse mental health condition, the effect of deprivation increases. Supporting the 2010 Marmot review, which found evidence of an influence of the socioeconomic state on an individual's psychosocial and behavioral state [17]. The difference in the top 15 predictors in the most deprived group compared to the unitary model, shows the impact of the living circumstance of patients on burn outcome, as these have a higher effect in this group. Thus, revealing the influence of an individual's living environment on an increased probability of a more severe burn injury. Researchers have found that living environment and other socioeconomic facts increases burn incidence and its severity [29]–[32]. Our research corroborates this and for the first time identifies that certain living circumstances – which indirectly reflects a patient's socioeconomic status - as top predictors of LOS.

5. Conclusions

Our study suggests segmentation of burn patients in terms of their socioeconomic status or mental state. The groups we identified are easy to explain and facilitate a better understanding of patients by clinicians. In particular, the usefulness of these groups was evaluated with the development and comparison of a set of cluster specific regression models. Evaluating significant predictors of LOS, we found that with adverse mental state or socioeconomic status of a patient, the influence of severity (established predictors of LOS) reduces. This finding highlights the importance of factors such as socioeconomic status and mental health on a patient's recovery.

The adoption of a collaborative decision-making process in our research ensured the production of models that allows easy understanding by clinicians and thus the ability to understand the reasoning of model outcomes. With explainable results, this study will aid burn units in better triaging patients by personalising their care pathway, such as taking their mental state and socioeconomic status into account.

References

[1]   World Health Organization, "The Global Burden of Disease. 2004 Update," 2008. Accessed: Dec. 11, 2018. [Online]. Available: https://www.who.int/healthinfo/global_burden_disease/GBD_report_2004update_full.pdf?ua=1.

[2]   World Health Organisation, "Burns," 2018. https://www.who.int/news-room/fact-sheets/detail/burns (accessed Jan. 23, 2020).

[3]   B. Moores, M. M. Rahman, J. A. D. Settle, and F. S. C. Browning, "On the predictability of the length of patient stay in a burns unit," 1975. Accessed: Dec. 13, 2018. [Online]. Available: https://ac.els-cdn.com/0305417975900030/1-s2.0-0305417975900030-main.pdf?_tid=85b68d8a-58a6-4fcd-98f6-7a48c9453bf2&acdnat=1544708655_5dc1557dca972a86dceb9e047e5a2406.

[4]   L. M. Sanderson, P. A. Buffier, R. R. Perry, and S. J. Blackwell, "A Multivariate Evaluation of Determinants of Length of Stay in a Hospital Burn Unit," *J. Burn Care Rehabil.*, vol. 2, no. 3, pp. 142–149, 1981, doi: https://doi.org/10.1097/00004630-198105000-00004.

[5]   A. Hussain and K. W. Dunn, "Predicting length of stay in thermal burns: A systematic review of prognostic factors," *Burns*, vol. 39, no. 7, pp. 1331–1340, Nov. 2013, doi: 10.1016/j.burns.2013.04.026.

[6]   M. D. Peck, L. Mantelle, and C. G. Ward, "Comparison of Length of Hospital Stay to Mortality Rate in a Regional Burn Center," *J. Burn Care Rehabil.*, vol. 17, no. 1, pp. 39–44, Jan. 1996, doi: 10.1097/00004630-199601000-00010.

[7]   D. Andel, L.-P. Kamolz, M. Niedermayr, K. Hoerauf, W. Schramm, and H. Andel, "Which of the Abbreviated Burn Severity Index Variables Are Having Impact on the Hospital Length of Stay?," 2007, doi: 10.1097/BCR.0B013E31802C9E8F.

[8]   S. J. Chong, C. Song, T. W. Tan, G. Kusumawijaja, and K. Y. Chew, "Multi-variate analysis of burns patients in the Singapore General Hospital Burns Centre (2003–2005)," *Burns*, vol. 35, no. 2, pp. 215–220, Mar. 2009, doi: 10.1016/J.BURNS.2008.06.014.

[9]   S. Barnes, E. Hamrock, M. Toerper, S. Siddiqui, and S. Levin, "Real-time prediction of inpatient length of stay for discharge prioritization," *J. Am. Med. Informatics Assoc.*, vol. 23, no. e1, pp. e2–e10, 2016, doi: 10.1093/jamia/ocv106.

[10]   National Burn Care Review Committee, "National Burn Care Review Committee Report Standards and Strategy for Burn Care A REVIEW OF BURN CARE IN THE BRITISH ISLES Contents," 2001. Accessed: Jul. 28, 2019. [Online]. Available: http://79.170.40.160/britishburnassociation.org/wp-content/uploads/2017/07/NBCR2001.pdf.

[11]   C. N. Onah, R. Allmendinger, J. Handl, P. Yiapanis, and K. W. Dunn, "A Clustering-Based Patient Grouper for Burn Care," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11872 LNCS, 2019, pp.

123–131.

[12]   K. J. Brasel, H. J. Lim, R. Nirula, and J. A. Weigelt, "Length of Stay: An Appropriate Quality Measure?," *Arch. Surg.*, vol. 142, no. 5, pp. 461–466, May 2007, doi: 10.1001/archsurg.142.5.461.

[13]   British Burn Association, "Outcome Measures for Adult and Paediatric Services Second Edition 2," 2016. Accessed: Feb. 12, 2019. [Online]. Available: https://www.britishburnassociation.org/wp-content/uploads/2018/05/BBA-Outcome-Measures-for-Adult-Paediatric-Services-2nd-Edition.pdf.

[14]   B. H. Bowser, F. T. Caldwell, J. A. Baker, and R. C. Walls, "Statistical methods to predict morbidity and mortality: self assessment techniques for burn units," *Burns*, vol. 9, no. 5, pp. 318–326, May 1983, doi: 10.1016/0305-4179(83)90077-3.

[15]   N. Caetano, P. Cortez, and R. M. S. Laureano, "Using data mining for prediction of hospital length of stay: An application of the CRISP-DM methodology," in *Lecture Notes in Business Information Processing*, 2015, vol. 227, pp. 149–166, doi: 10.1007/978-3-319-22348-3_9.

[16]   British Medical Association, "Health at a Price: Reducing the Impact of Poverty. A Briefing from the Board of Science.," 2017. [Online]. Available: https://www.bma.org.uk/media/2084/health-at-a-price-2017.pdf.

[17]   M. Marmot and R. Bell, "Fair society, healthy lives," *Public Health*, vol. 126, no. SUPPL.1, p. S4, 2012, doi: 10.1016/j.puhe.2012.05.014.

[18]   M. G. Marmot, R. Fuhrer, S. L. Ettner, N. F. Marks, L. L. Bumpass, and C. D. Ryff, "Contribution of Psychosocial Factors to Socioeconomic Differences in Health," *Milbank Q.*, vol. 76, no. 3, pp. 403–448, 1998, doi: 10.1111/1468-0009.00097.

[19]   I. Paccoud, J. Nazroo, and A. Leist, "A Bourdieusian approach to class-related inequalities: the role of capitals and capital structure in the utilisation of healthcare services in later life," *Sociol. Health Illn.*, pp. 1467-9566.13028, Nov. 2019, doi: 10.1111/1467-9566.13028.

[20]   L. Moore, A. F. Turgeon, M. J. Sirois, V. Murat, and A. Lavoie, "Influence of socioeconomic status on trauma center performance evaluations in a Canadian trauma system," *J. Am. Coll. Surg.*, vol. 213, no. 3, pp. 402–409, Sep. 2011, doi: 10.1016/j.jamcollsurg.2011.05.007.

[21]   J. A. Locke, A. MacKay Rossignol, and J. F. Burke, "Socioeconomic factors and the incidence of hospitalized burn injuries in New England Counties, USA," *Burns*, vol. 16, no. 4, pp. 273–277, 1990, doi: 10.1016/0305-4179(90)90138-M.

[22]   M. D. Peck, "Epidemiology of burns throughout the world. Part I: Distribution and risk factors," *Burns*, vol. 37, no. 7, pp. 1087–1100, Nov. 2011, doi: 10.1016/j.burns.2011.06.005.

[23]   C. Smolle., J. Cambiaso-Daniel, A. A. Forbes, P. Wurzer, G. Hundeshagen, L. K. Branski, F. Huss, and L. Kamolz, "Recent trends in burn epidemiology worldwide: A systematic review," *Burns*, vol. 43, no. 2, pp. 249–257, Mar. 2017, doi: 10.1016/j.burns.2016.08.013.

[24]   C. Diguiseppi, P. Edwards, C. Godward, I. Roberts, and A. Wade, "Urban residential fire and flame injuries: a population based study," *Inj. Prev.*, vol. 6, no. 4, p. 250, 2000, doi: 10.1136/ip.6.4.250.

[25]   N. Stylianou, I. Buchan, and K. W. Dunn, "A review of the international Burn Injury Database (iBID) for England and Wales: descriptive analysis of burn injuries 2003–2011," *BMJ Open*, vol. 5, no. 2, Feb. 2015, [Online]. Available: http://bmjopen.bmj.com/content/5/2/e006184.abstract.

[26]   A. M. Epstein. R. S. Stern, J. Tognetti, C. B. Begg, R. M. Hartley, E. Cumella, and J. Z. Ayanian, "The Association of Patients' Socioeconomic Characteristics with the Length of Hospital Stay and Hospital Charges within Diagnosis-Related Groups," *N. Engl. J. Med.*, vol. 318, no. 24, pp. 1579–1585, Jun. 1988, doi: 10.1056/NEJM198806163182405.

[27]   L. Poole, E. Leigh, T. Kidd, A. Ronaldson, M. Jahangiri, and A. Steptoe, "The combined association of depression and socioeconomic status with length of post-operative hospital stay following coronary artery bypass graft surgery: Data from a prospective cohort study," *J. Psychosom. Res.*, vol. 76, no. 1, pp. 34–40, Jan. 2014, doi: 10.1016/j.jpsychores.2013.10.019.

[28]   L. Moore, B. Cisse, B. L. B. Kuimi, H. T. Stelfox, A. F. Turgeon, F. Lauzier, J. Clément, and G. Bourgeois, "Impact of socio-economic status on hospital length of stay following injury: A multicenter cohort study Quality, performance, safety and outcomes," *BMC Health Serv. Res.*, vol. 15, no. 1, Jul. 2015, doi: 10.1186/s12913-015-0949-2.

[29]   N. E. Alden, P. Q. Bessey, A. Rabbitts, P. J. Hyden, and R. W. Yurt, "Tap water scalds among seniors and the elderly: Socio-economics and implications for prevention," *Burns*, vol. 33, no. 5, pp. 666–669, Aug. 2007, doi: 10.1016/j.burns.2006.10.003.

[30]   J. Delgado, M. E. Ramírez-Cardich, R. H. Gilman, R. Lavarello, N. Dahodwala, A. Bazán, V. Rodríguez, R. I. Cama, M. Tovar, and A. Lescano, "Risk factors for burns in children: Crowding, poverty, and poor maternal education," *Inj. Prev.*, vol. 8, no. 1, pp. 38–41, Mar. 2002, doi: 10.1136/ip.8.1.38.

[31]   A. Hussain and K. Dunn, "Burn related mortality in Greater Manchester: 11-year review of Regional Coronial Department Data," *Burns*, vol. 41, no. 2, pp. 225–234, Mar. 2015, doi: 10.1016/j.burns.2014.10.008.

[32]   L. Warda, M. Tenenbein, and M. E. K. Moffatt, "House fire injury prevention update. Part I. A review of risk factors for fatal and non-fatal house fire injury," *Injury Prevention*, vol. 5, no. 2. pp. 145–150, 1999, doi: 10.1136/ip.5.2.145.

[33]   D. for C. and L. G. DCLG, "The English Index of Multiple Deprivation (IMD) 2015 – Guidance," 2015,   Accessed: Jan.   28,   2018.   [Online].   Available: https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015.

[34]   A. Katainen, "Social class differences in the accounts of smoking - striving for distinction?," *Social. Health Illn.*, vol. 32, no. 7, pp. 1087–1101, Nov. 2010, doi: 10.1111/j.1467-

9566.2010.01267.x.

[35]   J. Walburn, K. Vedhara, M. Hankins, L. Rixon, and J. Weinman, "Psychological stress and wound healing in humans: a systematic review and meta-analysis Psychological stress and wound healing in humans: A systematic review and meta-analysis ☆," *J. Psychosom. Res.*, vol. 67, no. 3, pp. 253–271, 2009, doi: 10.1016/j.jpsychores.2009.04.002.

[36]   K. H. O'Brien and V. Lushin, "Examining the impact of psychological factors on hospital length of stay for burn survivors: A systematic review," *J. Burn Care Res.*, vol. 40, no. 1, pp. 12–20, Jan. 2019, doi: 10.1093/jbcr/iry040.

[37]   J. A. Wisely, E. Wilson, R. T. Duncan, and N. Tarrier, "Pre-existing psychiatric disorders, psychological reactions to stress and the recovery of burn survivors," *Burns*, vol. 36, no. 2, pp. 183–191, Mar. 2010, doi: 10.1016/j.burns.2009.08.008.

[38]   N. Tarrier, L. Gregg, J. Edwards, and K. Dunn, "The influence of pre-existing psychiatric illness on recovery in burn injury patients: the impact of psychosis and depression," *Burns*, vol. 31, no. 1, pp. 45–49, Feb. 2005, doi: 10.1016/j.burns.2004.06.010.

[39]   C. C. Berry, T. L. Patterson, T. L. Wachtel, and H. A. Frank, "Behavioural factors in burn mortality and length of stay in hospital," *Burns*, vol. 10, no. 6, pp. 409–414, Aug. 1984, doi: 10.1016/0305-4179(84)90080-9.

[40]   E. R. H. Wilson, J. A. Wisely, A. J. Wearden, K. W. Dunn, J. Edwards, and N. Tarrier, "Do illness perceptions and mood predict healing time for burn wounds? A prospective, preliminary study," *J. Psychosom. Res.*, vol. 71, no. 5, pp. 364–366, Nov. 2011, doi: 10.1016/j.jpsychores.2011.05.009.

[41]   L. P. Kamolz, H. Andel, A. Schmidtke, D. Valentini, G. Meissl, and M. Frey, "Treatment of patients with severe burn injuries: The impact of schizophrenia," *Burns*, vol. 29, no. 1, pp. 49–53, Feb. 2003, doi: 10.1016/S0305-4179(02)00207-3.

[42]   V. Lorant, D. Deliège, W. Eaton, A. Robert, P. Philippot, and M. Ansseau, "Socioeconomic inequalities in depression: A meta-analysis," *Am. J. Epidemiol.*, vol. 157, no. 2, pp. 98–112, Jan. 2003, doi: 10.1093/aje/kwf182.

[43]   T. Takizawa, T. Kondo, S. Sakihara, M. Ariizumi, N. Watanabe, and H. Oyama, "Stress buffering effects of social support on depressive symptoms in middle age: Reciprocity and community mental health," *Psychiatry Clin. Neurosci.*, vol. 60, no. 6, pp. 652–661, Dec. 2006, doi: 10.1111/j.1440-1819.2006.01579.x.

[44]   iBID, "International Burn Injury Database Annual Report (2014 / 15)," 2014.

[45]   V. L. Shavers, "Measurement of Socioeconomic Status in Health Disparities Research," *J. Natl. Med. Assoc.*, vol. 99, no. 9, pp. 1013–1023, 2007, Accessed: Jan. 25, 2018. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2575866/pdf/jnma00208-0045.pdf.

[46]   G. Kaplan, "Part III summary: What is the role of the social environment in understanding inequalities in health?," *Ann.NY Acad.Sci.*, vol. 896, pp. 116–119, 1999.

[47]   World Health Organization, "Closing the gap in a generation Health equity through action on the social determinants of health," 2008, Accessed: Feb. 03, 2018. [Online]. Available: http://apps.who.int/iris/bitstream/10665/43943/1/9789241563703_eng.pdf.

[48]   Welsh Government, "Welsh Index of Multiple Deprivation (WIMD) 2014," 2014, Accessed: Jan. 28, 2018. [Online]. Available: http://gov.wales/docs/statistics/2015/150812-wimd-2014-revised-en.pdf.

[49]   B. W. J H Penninx, A. T. F. Beekman, J. Ormel, D. M. W. Kriegsman, A. J. P. Boeke, J. TH. M. Van Eijk, and D. J. H. Deeg, "Psychological Status Among Elderly People With Chronic Diseases: Does Type Of Disease Play A Part?," 1996.

[50]   World Health Organization, *The ICD-10 classification of mental and behavioural disorders : clinical descriptions and diagnostic guidelines*. Geneva: World Health Organization, 1992.

[51]   M. A. Morgan *et al.*, "Influence of socio-economic deprivation on outcomes for patients diagnosed with oesophageal cancer," *Scand. J. Gastroenterol.*, 2007, doi: 10.1080/00365520701320471.

[52]   G. A. Abel, M. E. Barclay, and R. A. Payne, "Adjusted indices of multiple deprivation to enable comparisons within and between constituent countries of the UK including an illustration using mortality rates," *BMJ Open*, vol. 6, no. 11, Nov. 2016, doi: 10.1136/bmjopen-2016-012750.

[53]   J. Handl and J. Knowles, "An Investigation of Representations and Operators for Evolutionary Data Clustering with a Variable Number of Clusters," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 4193 LNCS, Springer Verlag, 2006, pp. 839–849.

[54]   A. K. Jain and R. C. Dubes, *Algorithms for clustering data.* Englewood Cliffs, NJ: Prentice Hall, 1988.

[55]   D. Müllner, "Modern hierarchical, agglomerative clustering algorithms," Sep. 2011, Accessed: Apr. 05, 2020. [Online]. Available: http://arxiv.org/abs/1109.2378.

[56]   J. C. Gower, "A General Coefficient of Similarity and Some of Its Properties," *Biometrics*, vol. 27, no. 4, p. 857, Dec. 1971, doi: 10.2307/2528823.

[57]   L. Kaufman and P. J. Rousseeuw, *Finding groups in data : an introduction to cluster analysis*. Wiley, 2005.

[58]   T. Caliñski and J. Harabasz, "A Dendrite Method For Cluster Analysis," *Commun. Stat.*, vol. 3, no. 1, pp. 1–27, 1974, doi: 10.1080/03610927408827101.

[59]   R. Tibshirani, "Regression Shrinkage and Selection via the Lasso," *Journal of the Royal Statistical*

Society. Series B (Methodological), vol. 58. WileyRoyal Statistical Society, pp. 267–288, 1996, doi: 10.2307/2346178.

[60]   B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, "Least angle regression," *Ann. Stat.*, vol. 32, no. 2, pp. 407–499, Apr. 2004, doi: 10.1214/009053604000000067.

[61]   J. Friedman, T. Hastie, and R. Tibshirani, "Regularization paths for generalized linear models via coordinate descent," *J. Stat. Softw.*, vol. 33, no. 1, pp. 1–22, Feb. 2010, doi: 10.18637/jss.v033.i01.

[62]   M. J. Krob, F. J. Dʾamico, and D. L. Ross, "Do Trauma Scores Accurately Predict Outcomes for Patients with Burns?," *J. Burn Care Rehabil.*, vol. 12, no. 6, pp. 560–563, Nov. 1991, doi: 10.1097/00004630-199111000-00011.

[63]   J. R. Saffle, B. Davis, and P. Williams, "Recent Outcomes in the Treatment of Burn Injury in the United States: A Report From the American Burn Association Patient Registry," *J. Burn Care Rehabil.*, vol. 16, no. 3, pp. 219–232, May 1995, doi: 10.1097/00004630-199505000-00002.

[64]   M. K. Wong and R. C. K. Ngim, "Burns mortality and hospitalization time — a prospective statistical study of 352 patients in an Asian National Burn Centre," *Burns*, vol. 21, no. 1, pp. 39–46, Feb. 1995, doi: 10.1016/0305-4179(95)90780-4.

[65]   C.-S. Yang, C.-P. Wei, C.-C. Yuan, and J.-Y. Schoung, "Predicting the length of hospital stay of burn patients: Comparisons of prediction accuracy among different clinical stages," *Decis. Support Syst.*, vol. 50, no. 1, pp. 325–335, Dec. 2010, doi: 10.1016/J.DSS.2010.09.001.

[66]   S. Meshulam-Derazon, S. Nachumovsky, D. Ad-El, J. Sulkes, and D. J. Hauben, "Prediction of morbidity and mortality on admission to a burn unit," *Plast. Reconstr. Surg.*, vol. 118, no. 1, pp. 116–120, 2006, doi: 10.1097/01.prs.0000221111.89812.ad.

[67]   J. A. Fauerbach, J. Lawrence, J. Haythornthwaite, M. McGuire, and A. Munster, "Preinjury psychiatric illness and postinjury adjustment in adult burn survivors," *Psychosomatics*, vol. 37, no. 6, pp. 547–555, 1996, doi: 10.1016/S0033-3182(96)71518-9.