Multichannel mixture models for time-series analysis and classification of engagement with multiple health services: an application to psychology and physiotherapy utilization patterns after traffic accidents

Nazanin Esmaili[a,b,*], Quinlan D. Buchlak[a], Massimo Piccardi[b], Bernie Kruger[d], Federico Girosi[c,e]

[a]*School of Medicine, University of Notre Dame Australia, Sydney, NSW, Australia*
[b]*Faculty of Engineering and IT, University of Technology Sydney, NSW, Australia*
[c]*Capital Markets Cooperative Research Centre (CMCRC), Sydney, NSW, Australia*
[d]*Transport Accident Commission, Geelong, VIC, Australia*
[e]*Translational Health Research Institute, Western Sydney University, Penrith, NSW, Australia*

Highlights

- Our study describes an effective machine learning method that combines supervised and unsupervised algorithms in an innovative way to yield useful clinical insights using complex time series data.
- De-identified compensation data was provided by the Australian Transport Accident Commission. Utilization of physiotherapy and psychology services was analysed. Datasets contained 788 psychology and 3,115 physiotherapy claimants and 22,522 and 118,453 episodes of service utilization, respectively.
- Data were processed to generate multidimensional sequences and time series clustering was applied using multichannel mixture of hidden Markov models. Combinations of hidden states and clusters were evaluated and optimised using the Bayesian information criterion.
- Cluster membership was investigated statistically using multinomial logistic regression and classified using gradient boosting machines (GBM), along with 5-fold cross-validation.
- Many administrative datasets contain similar timeseries data and, as such, the method presented has the potential to guide substantial further clinical research efforts. This work has the potential to be a well-cited article in the annals of clinical research as the field of applied machine learning continues to grow exponentially.

*Corresponding author
Email address: nazanin.esmaili@uts.edu.au (Nazanin Esmaili)

Multichannel mixture models for time-series analysis and classification of engagement with multiple health services: an application to psychology and physiotherapy utilization patterns after traffic accidents

Nazanin Esmaili[a,b,c,*], Quinlan D. Buchlak[a], Massimo Piccardi[b], Bernie Kruger[d], Federico Girosi[c,e]

[a]School of Medicine, University of Notre Dame Australia, Sydney, NSW, Australia
[b]Faculty of Engineering and IT, University of Technology Sydney, NSW, Australia
[c]Capital Markets Cooperative Research Centre (CMCRC), Sydney, NSW, Australia
[d]Transport Accident Commission, Geelong, VIC, Australia
[e]Translational Health Research Institute, Western Sydney University, Penrith, NSW, Australia

Abstract

Background. Motor vehicle accidents (MVA) represent a significant burden on health systems globally. Tens of thousands of people are injured in Australia every year and may experience significant disability. Associated economic costs are substantial. There is little literature on the health service utilization patterns of MVA patients. To fill this gap, this study has been designed to investigate temporal patterns of psychology and physiotherapy service utilization following transport-related injuries.

Method. De-identified compensation data was provided by the Australian Transport Accident Commission. Utilization of physiotherapy and psychology services was analysed. The datasets contained 788 psychology and 3,115 physiotherapy claimants and 22,522 and 118,453 episodes of service utilization, respectively. 582 claimants used both services, and their data were preprocessed to generate multidimensional time series. Time-series clustering was applied using a mixture of hidden Markov models to identify the main, different patterns of service utilization. Combinations of hidden states and clusters were evaluated and optimized using the Bayesian information criterion and interpretability. Cluster membership was further investigated using static covariates and multinomial logistic regression, and classified using high-performing classifiers (extreme gradient boosting machine, random forest and support vector machine) with 5-fold cross-validation.

Results. Four clusters of claimants were obtained from the clustering of the time series of service utilization. Service volumes and costs increased progressively from clusters 1 to 4. Membership of cluster 1 was positively associated with nerve damage and

*Corresponding author
Email address: nazanin.esmaili@uts.edu.au (Nazanin Esmaili)

negatively associated with severe ABI and spinal injuries. Cluster 3 was positively associated with severe ABI, brain/head injury and psychiatric injury. Cluster 4 was positively associated with internal injuries. The classifiers were capable of classifying cluster membership with moderate to strong performance (AUC: 0.62-0.96).

Conclusion. The available time series of post-accident psychology and physiotherapy service utilization were coalesced into four clusters that were clearly distinct in terms of patterns of utilization. In addition, pre-treatment covariates allowed prediction of a claimant's post-accident service utilization with reasonable accuracy. Such results can be useful for a range of decision-making processes, including the design of interventions aimed to improve the claimants' care and recovery.

Keywords: hidden Markov models, artificial intelligence, time-series analysis, traffic accidents, health service utilization

---

Introduction and background

Injuries from motor vehicle accidents represent a significant health burden worldwide [1, 2]. Tens of millions of people suffer non-fatal injuries yearly, with many experiencing a consequent disability [3]. In Australia, around forty-five thousand people are seriously injured every year and require hospitalization following transport accidents, with the age-standardized rate per person increasing by almost 0.9% annually [4, 5]. Additionally, while a large number of people endure minor injuries that do not require hospitalization, their ability to undertake regular daily activities, including work, is often reduced [4, 6]. The injury rate from road accidents is approximately 27 times the fatality rate in Australia, with injuries accounting for 40% of the total social cost of road accidents, spread over disability-related costs, medical expenses and out-of-work productivity losses [7, 8]. In 2016, the total cost came in at over $33 billion[1], a 22% increase from 2006 [5, 9]. According to a 2016 report from the Australian National University [9], the cost of road accidents amounts, on average, to $7.8 million per fatality, $310,094 per serious injury and $3,057 per minor injury.

Despite progress brought about by successful road safety programs, motor vehicle accidents (MVAs) still cause approximately 7, 800 serious hospitalizations per year in the Australian state of Victoria alone. MVAs are ranked as the second most important cause of hospitalization due to injury, after falls [10, 11]. The Transport Accident Commission (TAC) is a Victorian Government-owned organization that promotes road safety, funds treatment and rehabilitation services and provides financial

[1] In this paper, all costs are reported in Australian Dollars.

support for those who have been injured in transport-related accidents (including drivers, passengers, pedestrians, motorcyclists and cyclists) [12, 11, 4]. Unlike transport insurance schemes in many other jurisdictions that are tort-based, the scheme administered by the TAC is a no-fault injury compensation scheme, meaning that benefits may be paid to an injured person regardless of who caused the accident [12, 11, 4]. For the administration of this scheme, the TAC must collect data of the healthcare service payments made to its claimants. The data are organized as time series, where each time series is uniquely identified by a claim identification number and records all the service utilizations made by a claimant. In turn, each utilization episode includes the date of the service, its cost and its type, categorized according to three increasing levels of detail. In addition, the TAC possesses other data about claims commonly called "covariates" (or features) [13], such as the age and gender of the claimant, their mode of transportation at the time of accident, the type of injuries sustained, and others. These are grouped into the broader categories of demographic, injury and accident data.

Although accident and injury prevention should be the ultimate societal goal, much can be done to minimize the adverse impact of the accident injuries that do occur despite the best prevention efforts [14]. Appropriate post-accident care should aim to limit the suffering caused by the injury, and ensure the survivor's best possible recovery and quality of life [15, 14]. To this aim, the analysis of health service utilization following a transport-related injury may provide useful insight into a patient's health status months and years after the injury. Further, it may be possible to predict a patient's future health service utilization based on data collected at the time of the accident to facilitate care planning. This information may help healthcare providers to pre-empt patient care needs and tailor bespoke care plans for their patients to optimize clinical outcomes and improve the efficiency of service delivery. However, with the exception of severely injured patients [16, 17, 18, 19], we found little published information regarding patterns of healthcare utilization following transport injuries in the literature [12, 13]. In particular, to the best of our knowledge, no published work has addressed the utilization of *multiple* health services following transport injuries. In terms of methods of analysis, multivariate time-series clustering can glean structure and discover patterns in large time series datasets when little knowledge about potential underlying classes is available [20]. Such identification of structure may also prove useful as a preprocessing step for further analysis tasks [21]. Many algorithms and evaluation criteria have been developed for time-series clustering to date [22, 23, 24, 25], and they have found application in domains ranging from genomics to stock market analysis [21, 26]. In medicine, these methods have been applied, among other, to the analysis of magnetic resonance imaging data for brain mapping, cerebrovascular disease diagnosis and the identification of cancer [27, 28]. However, a review of the literature shows that the application of time-series clustering to health-care datasets has been more limited. Therefore, there is an opportunity to apply these methods to the healthcare domain and explore their potential for generating clinically-useful insights.

This research project has set out to investigate temporal patterns of post-accident utilization of multiple health services using de-identified compensation data provided by the TAC from the Australian state of Victoria. Claimants who lodged a claim in 2009 were included in the study, and utilization data were collected for the subsequent nine years. The measure of health service utilization was defined as the number of monthly visits to service providers such as psychologists, physiotherapists, chiropractors, physicians and other healthcare practitioners. Utilization was divided by type of service (details are provided in the following section), and the pool of claimants for a given type of service included all those who utilized it at least once. The main goals of this study were to (1) identify the main, distinct patterns of utilization for multiple health services and describe their main characteristics in terms of amount and speed of decay, and (2) describe the characteristics of, and classify, claimant groups in terms of covariates to understand client profiles associated with particular service utilization behavior patterns.

This work is primarily intended to present a general methodology that can be applied to healthcare problems beyond the specific research questions considered herein. Given the importance of mental health and physical recovery for the clients of the TAC, we have chosen psychology and physiotherapy as joint services to illustrate our approach. The healthcare literature suggests that these services exercise a synergistic influence on outcomes. For instance, studies [29, 30] have shown that the use of psychology services enhances physiotherapy outcomes. In its own right, physiotherapy is beneficial for a range of injuries and the physiotherapists' ability to encourage motivation in their patients can be key to their recovery [29]. In cases where patient motivation is low, psychological support services can prove beneficial in enhancing physiotherapy treatment [29, 30]. For these reasons, we believe that the joint analysis of these two services can shed some light on the nature of patient recovery, and assist with resource planning and intervention design.

Methodology

In this section, we describe the data and analytic methods used in this work. Subsection "Data" describes the raw data and the preprocessing. Subsection "Time-series clustering" describes the algorithm used for clustering the time series of service utilization, i.e. a mixture of hidden Markov models (MHMM). Subsection "Model selection" discusses the choice of hyperparameters for the model. Finally, after the clusters are identified, subsection "Characterizing the clusters with covariate analysis" analyzes the claimants' membership in the clusters using a set of available covariates.

Data

Two claims datasets for transport-related injuries were provided by the TAC regarding the claimants who sustained an accident and lodged their claim in 2009. The first dataset contained one record per claimant with all the information required for the management of their compensation claim, including demographic (gender, current age, age at accident), accident data (date, claimant security risk type, claim development month, number of claims, road user type), and injury classification (fatal, brain/head, severe acquired brain injury (ABI), concussion, degloving, burns, spinal injuries, amputation, quadriplegia, paraplegia, nerve damage, soft tissue injuries, dislocation, internal injuries, sprains/strains, limb fractures, general fractures, contusion/abrasion, sight/eyes, and general injuries, i.e. contusions, minor strains, minor sprains, minor lacerations and minor external bleeding). The second dataset included longitudinal data of service utilizations and payments for $9,328$ unique claimants (for a total of $1,048,576$ service utilizations), spanning January 2009-October 2017 (i.e., 106 consecutive months). Each service utilization was labeled with three service categories at an increasing level of detail, with 30, 75 and 226 unique values, respectively, across the dataset. The analysis deployed in this paper can be carried out using categories from any of the recorded levels and any combination thereof. However, since the third level was the most detailed, we chose to focus on it. The two datasets were linked and de-identified for the analysis, and the ethics committees of University of Technology Sydney (UTS) and the TAC approved its use (UTS Human Research Ethics Committee Application ETH182331).

For the two chosen services, "psychology" and "physiotherapy", the longitudinal data contained, respectively, 788 and $3,115$ claimants, for a respective total of $22,522$ and $118,453$ service utilizations. The average number of utilizations per claimant was 29 and 38, respectively. This showed that, in general, the number of claimants utilizing physiotherapy was higher than those using psychology. The number of claimants who used both services was 582, for a total of $18,472$ and $40,631$ psychology and physiotherapy service utilizations, respectively. The number of episodes of these two services was correlated. When the number of physiotherapy utilization episodes increased to 70 or above, the number of psychology episodes increased as well to approximately 32. This suggested that, on average, claimants who received more physiotherapy treatments also utilized more psychological services.

We built 582 time series for this population of claimants by aggregating their utilizations by month. We formed the time series of service utilization by using the date of the accident as the first month/slot. In recent work [13], Esmaili et al. used the month of first service utilization as the first month of the time series and only used the number of days between the accident date and the first service utilization as a covariate for cluster analysis. However, in our paper we chose to use the accident date as the first element of the time series to capture the effect of the time elapsed between the accident date and that of the first utilization in the clustering. In addition, our analysis covered multiple services and a common time reference was required. Each time series thus started from the month of the claimant's accident and ended in October 2017. Therefore, the time series had variable length, from a minimum of 97 months to a maximum of 106. The time series were not trimmed or normalized to a common length to avoid curtailing or distorting their information. The average total number of service utilizations per claimant was approximately 31 for psychology and 69 for physiotherapy. The number of service utilizations per month ranged between 0 and 24 for psychology and between 0 and 30 for physiotherapy. For ease of reference, utilization is displayed with unique colors throughout the paper (e.g., zero is white, one is yellow, and so on). Fig 2 shows a "stacked plot" visualization of the 582 time series for both services. The height of each colored bar is proportional to the number of time series with the corresponding number of utilizations. Table 1 shows the main characteristics of the claimants who used psychology services, physiotherapy services and both services in terms of demographic, accident and injury covariates. This table reports injuries for which the corresponding population has at least 1% coverage. As such, some injuries (e.g., quadriplegia, amputation, burns and fatal) do not appear. In addition, Table 1 shows that the descriptive statistics of the 9,328 claimants in the dataset of service utilization (the second dataset) are comparable to those in the complete dataset (the first dataset), so that the former can be regarded as a representative sample of the entire population.

Time-series clustering

The monthly utilizations of health services for a given claimant were treated as random variables and are noted in the following as x_t, $t = 1 \ldots T$. In turn, each observation, x_t, consisted of multiple *channels*, x_t^c, $c = 1 \ldots C$, one per type of service. For instance, in our case a notation such as $x_7 = \{x_7^1 = 2, x_7^2 = 4\}$ means that the claimant received two psychological consultations and four physiotherapy sessions in the seventh month since the date of their accident. Given the time series of service utilizations of a population of claimants, our goal was to group them into clusters that could be as homogeneous as possible and describe the main, distinct patterns of service utilization.

Table 1: **Descriptive statistics for demographic, accident and injury covariates.**

Variables	Psychology	Physiotherapy	Both	First dataset	Second dataset
Sample size	788	3,115	582	93,633	9,327
Accident injuries (percentage of cases)					
General injuries	95%	91%	95%	85%	85%
Contusion abrasion	67%	58%	66%	57%	61%
Soft tissue	53%	50%	58%	49%	49%
Limb fractures	37%	40%	40%	25%	28%
General fractures	38%	27%	38%	20%	21%
Internal injuries	29%	18%	28%	16%	15%
Sprains strains	24%	24%	28%	17%	18%
Brain head injury	24%	11%	23%	9%	10%
Dislocation	19%	19%	21%	8%	10%
Concussion	19%	11%	18%	9%	10%
Spinal	8%	5%	9%	4%	4%
Severe ABI	7%	2%	8%	1%	1%
Psych flag	3%	1%	3%	<1%	<1%
Degloving	2%	2%	2%	1%	1%
Nerve damage	2%	1%	2%	1%	1%
Paraplegia	1%	1%	1%	<1%	<1%
Role in transport accident					
Driver	47%	40%	49%	47%	44%
Passenger	22%	14%	21%	19%	21%
(Motor/) Cyclist	16%	34%	16%	24%	26%
Pedestrian	12%	10%	12%	8%	9%
Witness	3%	1%	2%	1%	1%
Gender					
Male	48%	56%	53%	54%	54%
Female	52%	44%	47%	46%	46%
Age group (at accident)					
< 30-year-old	35%	31%	17%	33%	34%
30-40-year-old	22%	21%	19%	18%	19%
40-50-year-old	21%	22%	24%	17%	17%
50-60-year-old	13%	14%	22%	14%	13%
> 60-year-old	9%	13%	18%	19%	16%

To tackle data clustering with a solid statistical foundation, we used a *mixture distribution* which is a probability distribution obtained from the combination of multiple component distributions. If each component is set into correspondence with a distinct pattern of utilization, the mixture describes all patterns of utilizations of an entire population. Formally, the probability distribution of a time series of T months, $x_{1:T}$, in a mixture with M components is defined as:

$$p_{mixture}(x_{1:T}) = \sum_{m=1}^{M} p(x_{1:T}|m)p(m) \tag{1}$$

where $p(x_{1:T}\,|m)$ is the m-th component distribution and $p(m)$ is its prior probability. Eq (1) can be easily proved using Bayes' rule and marginalization [31]. Given a generic time series, $x_{1:T}$, a simple inference step can be used to assign it to its most similar pattern in the mixture:

$$m^*(x_{1:T}) = \underset{m=1\ldots M}{\operatorname{argmax}}\, p(x_{1:T}\,|m)p(m) \tag{2}$$

In this way, the mixture distribution models the patterns of the entire population and, at the same time, it partitions its samples into the patterns. However, it remains to be chosen how to define the component distribution, $p(x_{1:T}\,|m)$, of each pattern, $m = 1 \ldots M$. A powerful model to describe time series is the hidden Markov model (HMM), a probabilistic graphical model that assumes that each observation, x_t, $t = 1 \ldots T$, in a time series is generated by a latent, discrete-valued "state", y_t. In turn, the sequence of the states, $y_{1:T}$, forms a Markov chain that embeds all the sequential properties of the model. An HMM is referred to as a "factorized model" that defines the probability of the observations as:

$$p_{HMM}(x_{1:T}) = \sum_{y_{1:T}} p(y_1) \prod_{t=2}^{T} p(y_t|y_{t-1}) \prod_{t=1}^{T} p(x_t|y_t) \tag{3}$$

The terms on the right hand side of Eq (3) fully define the HMM and include: 1) the probability of the initial state, $p(y_1)$; 2) the probability of transitioning from the state at time $t-1$ to the state at time t, $p(y_t|y_{t-1})$; and 3) the probability of observation x_t when in state y_t, $p(x_t|y_t)$. Such terms are commonly referred to as initial-state, transition and observation probabilities. Each state variable is a latent categorical variable with an arbitrary number of values; let us say, S: therefore both $p(y_1)$ and $p(y_t|y_{t-1})$ can be modeled by conventional categorical distributions. Conversely, each observation, x_t, can consist of any combination of numerical and categorical values and $p(x_t|y_t)$ is chosen accordingly. In our case, each observation is a multichannel vector with C values, which we assume independent conditional on the state. Please note that this does not equate to the values being assumed independent: the values are modeled jointly through the dependencies accounted for by the state variable. This is a rather general hypothesis that models well many real multichannel time series. To model $p(x_t|y_t)$ we have used a multinomial distribution simply to limit further assumptions.

$$p(x_t|y_t) = \prod_{c=1}^{C} p(x_t^c|y_t) \tag{4}$$

With their assumptions, HMMs often well reflect real sequential data, and for this reason they have found wide adoption in fields as diverse as computer vision [32], signal processing [33], natural language processing [34], financial prediction [35], gene finding [36] and inpatient journey analysis [37].

By using an HMM to model each component distribution, the mixture distribution specializes as a *mixture of hidden Markov models* (MHMM). Such a mixture simultaneously a) describes the multiple patterns of utilization and b) properly models each pattern by taking into account its temporal nature. The overall probability distribution is obtained by incorporating Eqs (3,4) in the mixture model (Eq 1):

$$\begin{aligned} p_{MHMM}(x_{1:T}) &= \sum_{m=1}^{M} p_{HMM}(x_{1:T}|m)p(m) \\ &= \sum_{m=1}^{M} \sum_{y_{1:T}} \left[p(y_1|m) \prod_{t=2}^{T} p(y_t|y_{t-1}, m) \prod_{t=1}^{T} \prod_{c=1}^{C} p(x_t^c|y_t) \right] p(m) \end{aligned} \tag{5}$$

Based on the model in Eq 5, it is also possible to infer the sequence of states with highest probability for a given time series and a given cluster by replacing the marginalization of the states with a maximization:

$$y^*_{1:T}(x_{1:T}, m) = \left[p(y_1|m) \prod_{t=2}^{T} p(y_t|y_{t-1}, m) \prod_{t=1}^{T} \prod_{c=1}^{C} p(x_t^c|y_t) \right] p(m) \tag{6}$$

The inferability of the states is a key property of this model since the inferred states can be used to extensively characterize and contrast the found clusters.

Operationally, the probability distribution in Eq (5) has a number of parameters that need to be set before it can be evaluated. Such parameters include (1) the prior probability of each component, $p(m)$ ($M - 1$ free parameters); (2) the initial state probabilities of each component ($M(S_m - 1)$ free parameters, where S_m notes the number of states of the m-th component distribution); (3) the transition probabilities of each component ($MS_m(S_m - 1)$ free parameters), and (4) the observation probabilities of each component ($MS_mC(V - 1)$ free parameters, where V is the number of distinct values for the observations of each channel). By noting them collectively as Θ, we can make them explicit in the probability as $p_{MHMM}(x_{1:T}|\Theta)$. Given a set of time series, $X = \{x^i_{1:T}\}, i = 1 \ldots N$, the parameters can then be estimated with a common maximum-likelihood criterion [38]:

$$\Theta^* = \underset{\Theta}{\operatorname{argmax}} \sum_{i=1}^{N} \ln p_{MHMM}(x^i_{1:T_i}|\Theta) \tag{7}$$

Efficient solvers exist for Eq (7) and we use the expectation-maximization algorithm for the experiments in this paper [38, 39]. To implement the model, we have used R and its seqHMM and RColorBrewer packages [40, 41, 42, 43, 39]. However, two further parameters need to be set outside of the maximum-likelihood framework: the number of HMMs in the mixture, M, and the number of states in each HMM, $S_m, m = 1 \ldots M$. These parameters cannot be set using the maximum-likelihood criterion because the estimation would lead to overly-complex models, and for this reason they are often referred to as "hyperparameters". The usual approach for setting a model's hyperparameters is to minimize a trade-off between its likelihood and its complexity such as the Bayesian information criterion (BIC) [44] or the Akaike information criterion (AIC) [45], or to utilize external knowledge such as the judgment of domain experts. This selection is discussed in the following subsection.

Model selection

To find optimal hyperparameters (i.e., the number of clusters and the number of states in each cluster), we first used the TAC's expertise to determine a plausible range for the clusters (i.e. $m \in \{2, \cdots, 6\}$) and the states (i.e. $S_m \in \{3, \cdots, 9\}$). Choosing the number of clusters and hidden states is a model selection problem, and although this problem has been vastly studied for both mixture models and HMM, the optimal selection of the number of hyperparameters for these models is still partially unresolved [46, 47, 48].

In general, model selection consists of choosing a suitable trade-off between fitting the available data and preventing over-fitting. In this context, over-fitting means fitting the available sample "too tightly" with a model with too many degrees of freedom, rather than modelling its true generating distribution. This trade-off can be achieved by using a criterion that balances the model's goodness of fit and complexity such as the BIC, the AIC and their variants. In order to find the best model, we have run our analysis using BIC, AIC, consistent AIC (CAIC), corrected AIC (AICc), AIC3, AICu and the Hannan-Quinn criterion (HQC) (for details of these criteria, the reader is referred to [49]). Based on the TAC's opinion on the interpretability of the resulting models, we have chosen the BIC as the primary selection criterion for the number of the states. The BIC consists of two terms: a term that measures the negative log likelihood of the data, and a term that grows logarithmically with model complexity; therefore, lower values of BIC are preferable. To avoid biasing the mixture model toward any components, we only considered models with the same number of hidden states in each cluster We have chosen to use the same number of hidden states for each cluster (e.g., six clusters with five states each, $(5, 5, 5, 5, 5, 5)$) to avoid favoring some clusters with more degrees of freedom than others, and therefore biasing the number of members toward them. For the number of the clusters, we have simply accepted the judgment of the TAC's domain experts on the results that offered the best interpretability.

Characterizing the clusters with covariate analysis

Interpretation of the clusters produced by an unsupervised clustering algorithm requires external information and validation criteria. To this aim, we were able to exploit the set of demographic, accident and injury covariates available for each individual claimant. The first analysis has been conducted with multinomial logistic regression, using the covariates as inputs and the claimants' memberships to the clusters inferred from the MHMM as outputs. Multinomial logistic regression is likely the most widely adopted method for the analysis of factors of influence of a categorical variable. The approach uses a simple multinomial logistic regression classifier to predict a categorical variable based on a set of inputs commonly referred to as the "covariates". It repeatedly fits the classifier onto a subset of the available training data and applies a statistical test of stability to the estimated coefficients, rating the coefficients as stable or unstable, and the corresponding inputs as significant or not significant. Cluster 1 was used as the multinomial reference. Covariate reference categories were driver, general injuries, female, and age <30. Covariates were removed from the multinomial logistic regression model if they demonstrated a z-score < 0.02 and the model was subsequently retrained [50, 51]. A standard significance threshold of $p < 0.05$ was used. This analysis was conducted using the statsmodels [52] package in Python 3.6.

In addition to the multinomial regression analysis, we carried out an exploration of clusters membership using supervised classifiers. Extreme gradient boosting machine (XGBM), random forest (RF) and support vector machine (SVM) classifiers were trained to predict membership of each of the derived clusters. As inputs to the classifiers, we used the features included in the final multinomial logistic regression model. As target outputs, we used the claimant memberships to the clusters inferred from the MHMM. The synthetic minority oversampling technique (SMOTE) was used to address class imbalance [53]. XGBM is a classifier based on the combination (ensemble learning) of multiple “weak” classifiers, trained in a sequential manner so that each classifier attempts to “boost” the performance of the previous. RF is another ensemble-learning classifier that combines the outputs of multiple decision trees, often displaying better generalization than the individual trees. SVM is a classifier that is able to empirically maximize the separation between the various classes in a kernel space. In turn, SMOTE is an oversampling method that creates new, synthetic samples for the minority classes by leveraging the k nearest neighbors of the actual samples. All of these approaches have reported a very strong empirical performance in a number of domains and data classification contexts [54, 55, 56, 50, 57, 58, 59, 60] and were adopted as a consequence. Optimal parameters for the supervised classifiers were found using grid search with 5-fold cross-validation. The performance of the final models was reported as the highest over the 5-fold cross-validation, averaged over three runs for each model. The performance was measured in terms of the area under the receiver operating characteristic (AUC) metric. Early stopping was used to mitigate over-fitting. In addition, Shapley additive explanations (SHAP) values were calculated and plotted to determine feature importance for the XGBM classification models (SHAP is an approach based on game theory that can be used to elucidate feature importance in tree-based machine learning models and can assist in explaining their output [61, 62, 63]). This analysis was conducted using the scikitlearn [64], xgboost [65] and shap [61] packages in Python 3.6. Figure 1 shows an overview of our overall methodology.

Experiments and results

As a preliminary analysis of the data, Table 1 shows that the three most common injuries in this sample were general injuries, contusion/abrasion and soft tissue injuries. Injuries related to the head such as brain/head injury, concussion and severe ABI were more common among the psychology-only sample and the sample that used both services, than the physiotherapy-only sample. The two most common roles in the accident for the psychology-only sample were driver and passenger, while they were driver and motorcyclist/cyclist for the physiotherapy-only sample. Women were slightly more likely to use psychology services; usage amongst females was 4% higher than males. Males were more likely to use physiotherapy services; usage amongst males was 12% higher than females. The age distributions for these two samples were similar, but the sample that had received both services was significantly older.

For further inspection, Fig 3 shows the weighted directed graph of injury co-occurrence for the 582 claimants. The most frequently co-occurring injuries are displayed as a set of vertices connected by edges. For each injury type, the graph has an edge pointing toward the injury that most frequently occurs with it, with the edge

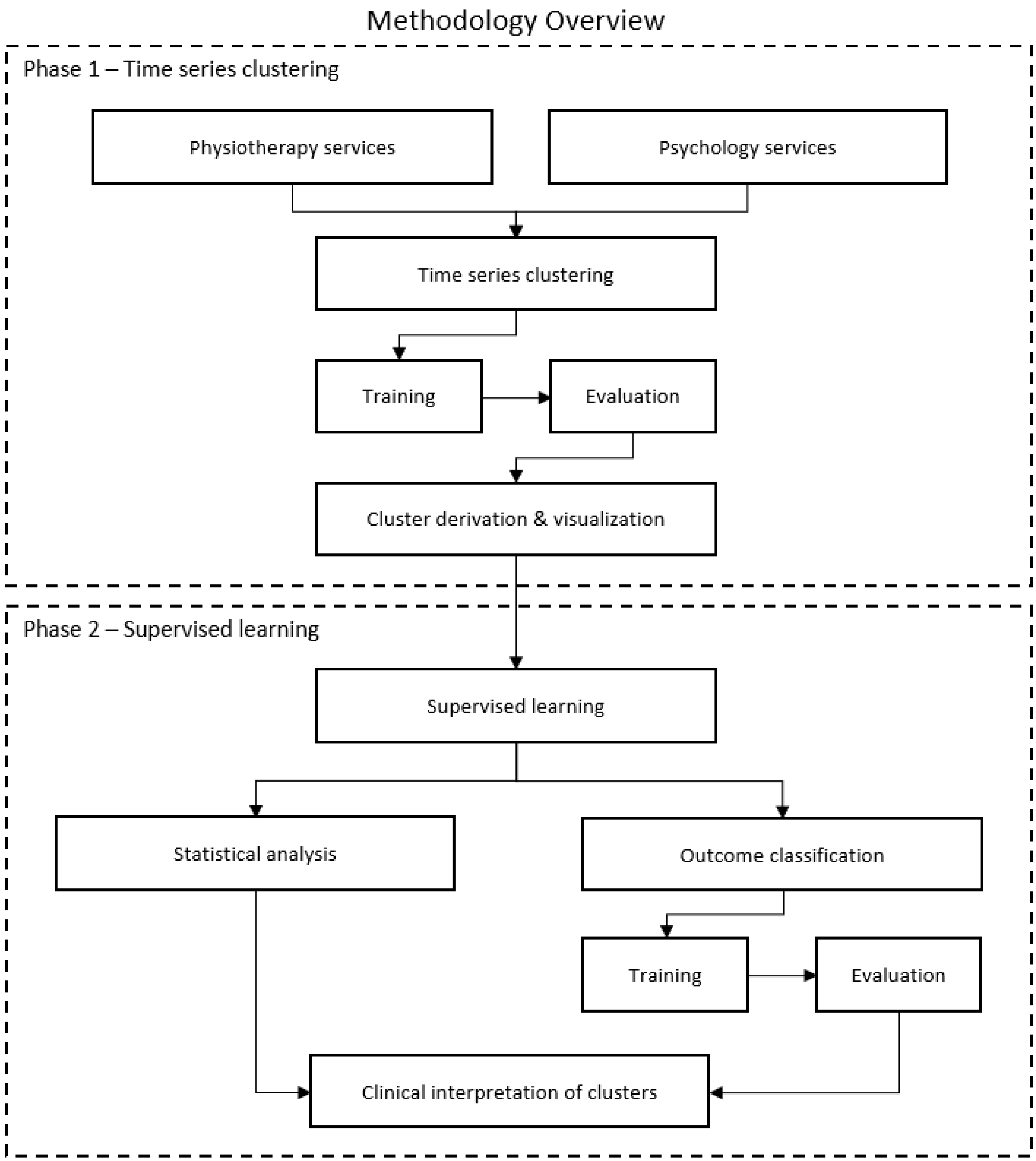

Figure 1: **Methodology overview.**

Table 2: **BIC for four clusters within a plausible range of hidden states (i.e. $S_m \in \{3, \cdots, 9\}$). Each row shows BIC of an MHMM model with four clusters and a different number of hidden states. The MHMM model with five hidden states has the minimum BIC, therefore, it is the optimal model.**

MHMM model	BIC
MHMM (2,2,2,2)	130999.9
MHMM (3,3,3,3)	122572.3
MHMM (4,4,4,4)	120437.2
MHMM (5,5,5,5)	120311.3
MHMM (6,6,6,6)	121420.6
MHMM (7,7,7,7)	123385.7
MHMM (8,8,8,8)	124945.8
MHMM (9,9,9,9)	127958.7

thickness proportional to the frequency of co-occurrence. For example, for quadriplegia and paraplegia, the most frequent co-occurring injury was general fractures. In turn, the most frequent co-occurring injury with general fractures was general injuries. General injuries were by far the most frequently co-occurring injuries overall, followed by general fractures and contusion/abrasion.

Results from the joint analysis of psychology and physiotherapy service utilization

For the model selection of the MHMM, the results with four clusters offered the best interpretability according to the TAC's domain experts. For the number of the states, we chose instead the value with the minimum BIC in the explored range. Table 2 reports the BIC values with four clusters and different numbers of hidden states, showing that the optimal value was attained with five hidden states. The resulting clusters have been named simply as clusters 1, 2, 3 and 4, and ended up containing 180, 116, 226 and 60 members, respectively (roughly 31%, 20%, 39% and 10% of the entire sample). For each time series, the sequence of hidden states is determined jointly from both services. To attach a meaningful interpretation to the inferred states, we first looked at the range of service utilizations associated with each of the 4 * 5 = 20 states in the MHMM. We then named the various levels of service utilizations for each individual service as "zero", "low", "medium", and "high". Such levels have been defined by sub-dividing the range of utilizations into approximately uniform intervals. When looking at the levels of service utilization for both psychology and physiotherapy for all 20 states and across the entire population, we observed the following 10 combinations of service level: (zero, zero), (zero, low), (zero, medium), (zero, high), (low, zero), (low, low), (medium, zero), (medium, high),(high, medium), and (high, high). In the rest of this section, we use these labels to refer to the 10 distinguishable states in our MHMM. The existence of some states of comparable service level across the various clusters is not in itself a sign of an overly-complex model, since these states have different transition probabilities and evolve from and toward different states. In addition, for ease of visualization, we use three different color shades to visualize the different states in all following plots and figures: orange for (*non-zero*, zero), blue for (zero, *non-zero*) and green for (*non-zero*, *non-zero*).

Figures 4 and 5 show stacked plots of the number of monthly utilizations of the psychology and physiotherapy services and stacked “state paths” for all four clusters. The right legend lists the numbers of monthly service utilizations while the bottom legend lists the state labels. The stacked “state paths” show the traversed sequences of states for the claimants in each cluster. For further insight, Figure 6 shows five random individual trajectories (utilizations for both services and inferred states) in each cluster. These figures show that the four clusters manifestly correspond to different typical trends of joint service utilization. In the following paragraphs, we study the various attributes (balance, amount, decay), pattern and behavior of monthly service utilizations for both services separately, and we leverage the hidden states to explore the latent relationship between the utilization of these two services.

Figures 4 and 6 show that cluster 1 contains psychology time series with mostly 0, 1 or 2 utilizations, and physiotherapy time series in a medium to high range of utilizations, with both decaying very fast to zero. The hidden states, too, rapidly decay toward state (zero, zero), reflecting the fast decay of the joint time series. The non-zero hidden states for this cluster are (low, zero), (zero, low) and (zero, high), showing that these claimants usually do not utilize both services. The trend in this cluster is for claimants to start with a high utilization of physiotherapy and to receive one or two psychology consultations after their first physiotherapy treatment. After that, they mainly receive only physiotherapy treatments. After 36 months from the accident, only 25% of these claimants use either service. The individual trajectories in Fig 6 confirm these observations. Overall, we believe that this is a group of claimants with rapid recovery and that it would be desirable to design interventions that could move more claimants to this cluster.

The trend in cluster 2 is more intense compared to cluster 1 and the duration of utilization is longer. At month 12, more than 60% of the claimants have used either of these services while the same figure is 30% for cluster 1. Based on the non-zero hidden states (i.e., (zero, low), (zero, medium) and (low, low)), we observe that most of the claimants either did not use the psychology service at all or they used low levels while instead using physiotherapy. Based on the amounts of non-zero hidden states, we can conclude that claimants in this cluster have, on average, used physiotherapy for much longer than those in cluster 1. Figure 7 illustrates the weighted directed graph of the most likely co-occurring injuries for each cluster. This graph shows two extra injuries in cluster 2 compared to cluster 1 – amputation and burns – which are both significant injuries. Also, the edges of some of the significant injuries such as spinal are thicker. This gives some evidence that the claimants in cluster 2 have likely suffered from more major injuries than those in cluster 1. The differences between clusters 1 and 2 are even more pronounced in the samples in Fig 6 (trajectories 1 to 5 vs trajectories 6 to 10). The imbalance between physiotherapy and psychology services may suggest that seeking more and early psychology services should be considered as a possible intervention for the members of this cluster.

In cluster 3, physiotherapy and psychology service utilizations seem intense in the first year after the accident. This is the only cluster that starts with 100% non-zero hidden states at the beginning of the time series. At month 6, more than 80% of the claimants have used at least either service. Unlike the previous two clusters, physiotherapy utilization decays faster than psychology. However, the physiotherapy time series are still longer than those in clusters 1 and 2. This may suggest that claimants in this cluster use psychology alongside physiotherapy to enhance their recovery. Fig 5 shows that the majority of the utilizations take place in the first three years after the accident and that they decay at a rate comparable with cluster 2 after the third year. The hidden states in this cluster are (low, zero), (zero, low), (zero, medium) and (medium, high), reflecting the higher average utilization.

Cluster 4 is the most distinctive cluster. As a general trend, its members used both psychology and physiotherapy more extensively than in any other cluster, and for a sustained period of time. There are claimants in this cluster that have used a high level of psychology or/and physiotherapy even eight years after the accident. In the sample in Fig 6, trajectories 16, 17 and 19 used long and relatively high levels of psychology services. Trajectory 18, in particular, intensely used psychology and physiotherapy years after the accident. In trajectory 20, the utilization of both services started long after the accident, possibly indicating a late onset of problems. Overall, the common pattern in this cluster is the use of both services long after the accident. The graph in Fig 7 shows the presence of most severe injuries such as spinal and brain/head injuries. This is the only cluster where brain/head injuries are the most frequently co-occurring injuries with general injuries (for other clusters, they are contusion/abrasion).

Table 3 shows the cost analysis per claimant for psychology and physiotherapy services. The overall number of service utilizations for this pool of 582 claimants was $397,229$. The total service cost increases progressively from cluster 1 through cluster 4. The average cost of psychology services for cluster 4 claimants is substantially higher than that seen in the other clusters, and it is also much higher than the average physiotherapy cost in that cluster. The average psychology cost is also higher than the physiotherapy cost in cluster 3. This is in line with the more frequent use of psychology services in this cluster. Conversely, in clusters 1 and 2, physiotherapy service costs outweigh the cost of psychology services.

Finally, we discuss the complexity of the model with respect to the available data. Like for any model, the number of parameters of our model is independent of the number of claimants; however, their ratio can be easily computed. Our model has $1,179$ total parameters, accounting for the parameters of the transition probabilities, initial state probabilities, prior probabilities of the clusters, and observation probabilities, which are, in order: $M \times S_m \times (S_m - 1)$, $M \times (S_m - 1)$, $M - 1$, and $M \times S_m \times C \times (V - 1)$, where $M (= 4)$, $S_m (= 5)$, $C (= 2)$ and $V (= 28)$ are the number of clusters, the number of states in each cluster, the number of channels, and the number of distinct values for the observations of each channel. Since the number of claimants is 582, the total number of parameters to be estimated per claimant is $1,179/582 = 2.03$. More so, the total number of observations is $59,103$, including $18,472$ and $40,631$ for psychology and physiotherapy, respectively, so we have $59,103/1,259 = 50.12$ observations per parameter. All in all, the number of parameters seems reasonable, and the model, effective and easily computable.

Results of the covariate analysis

After clustering the joint time series of psychology and physiotherapy services, we applied various classification algorithms to predict cluster membership based on demographic, injury and accident role covariates. Using external covariates to predict cluster membership allowed us to characterize and understand claimant profiles associated with differing service utilization patterns. This analysis also allowed us to determine whether the claimant covariate profile available at the time of the accident could be used as an appropriate and effective indicator of the extent of future service utilization over the lifetime of the patient's claim.

Multinomial logistic regression elucidated associations between covariates and cluster membership (Table 4). The multinomial logistic regression model was statistically significant ($p < 0.01$). Marginal effects derived from the multinomial logistic regression model (Table 5) suggested that severe ABI and other spinal injuries were significantly and negatively associated with membership of cluster 1, while nerve injuries were significantly and positively associated with membership of this cluster. No covariates were significantly associated with membership of cluster 2. Membership of cluster 3 was significantly and positively associated with severe ABI, brain/head injuries and psychiatric injury. Membership of cluster 4 was significantly and positively

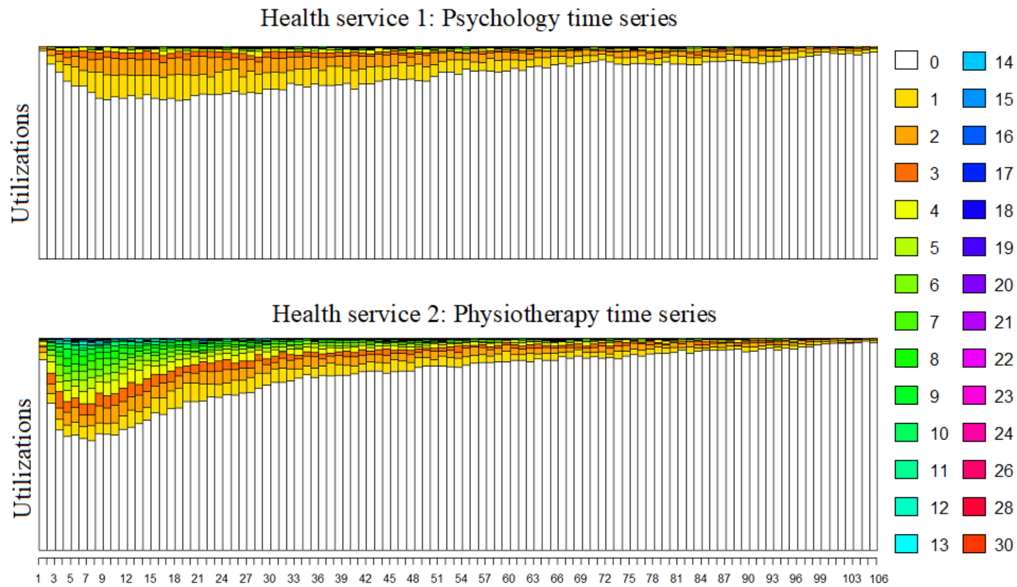

Figure 2: **"Stacked plot" of the number of monthly utilizations of psychology and physiotherapy services.** Each figure shows the 582 time series as a "stacked plot". The height of the colored bars is proportional to the number of time series with the corresponding number of utilizations.

Table 3: **Cost analysis per claimant.**

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Num. of claimants	180	116	226	60
Ave. psych. cost	1,874	1,755	4,042	18,449
Ave. physio. cost	2,710	3,684	2,818	5,256
Ave. total psych. and physio. cost	4,584	5,439	6,860	23,705
Percentage of psych. and physio. to total cost	4.80%	8.93%	13.58%	14.44%

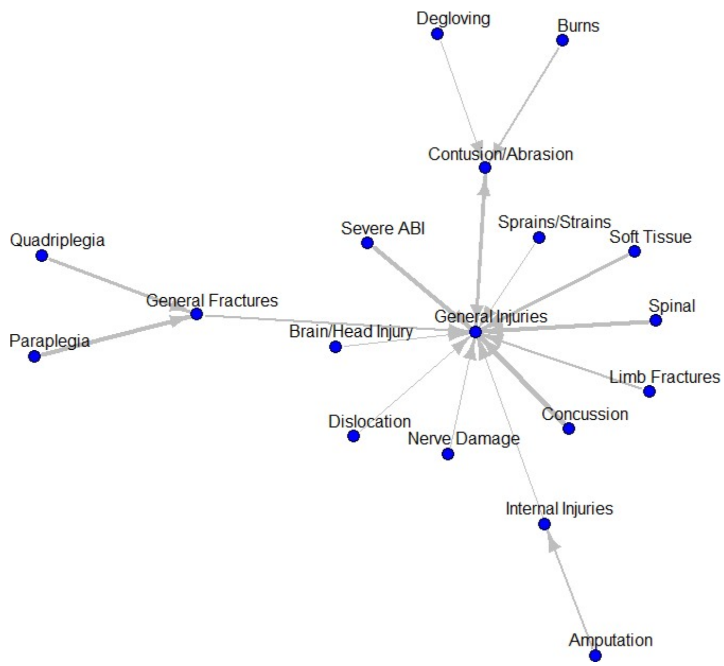

Figure 3: **"Weighted directed graph" of injury co-occurrence for the 582 claimants.** This graph shows the most frequent injury co-occurrences as a set of vertices connected by edges. For each injury type, the graph has an edge showing the injury that most frequently occurs with it. Example: for quadriplegia and paraplegia, the most frequent co-occurring injury is general fractures; the most frequent co-occurring injury with general fractures is general injuries. The thickness of each edge is proportional to the frequency of co-occurrence. For a balanced layout, the graph is centered on the most frequently co-occurring injury overall (general injuries).

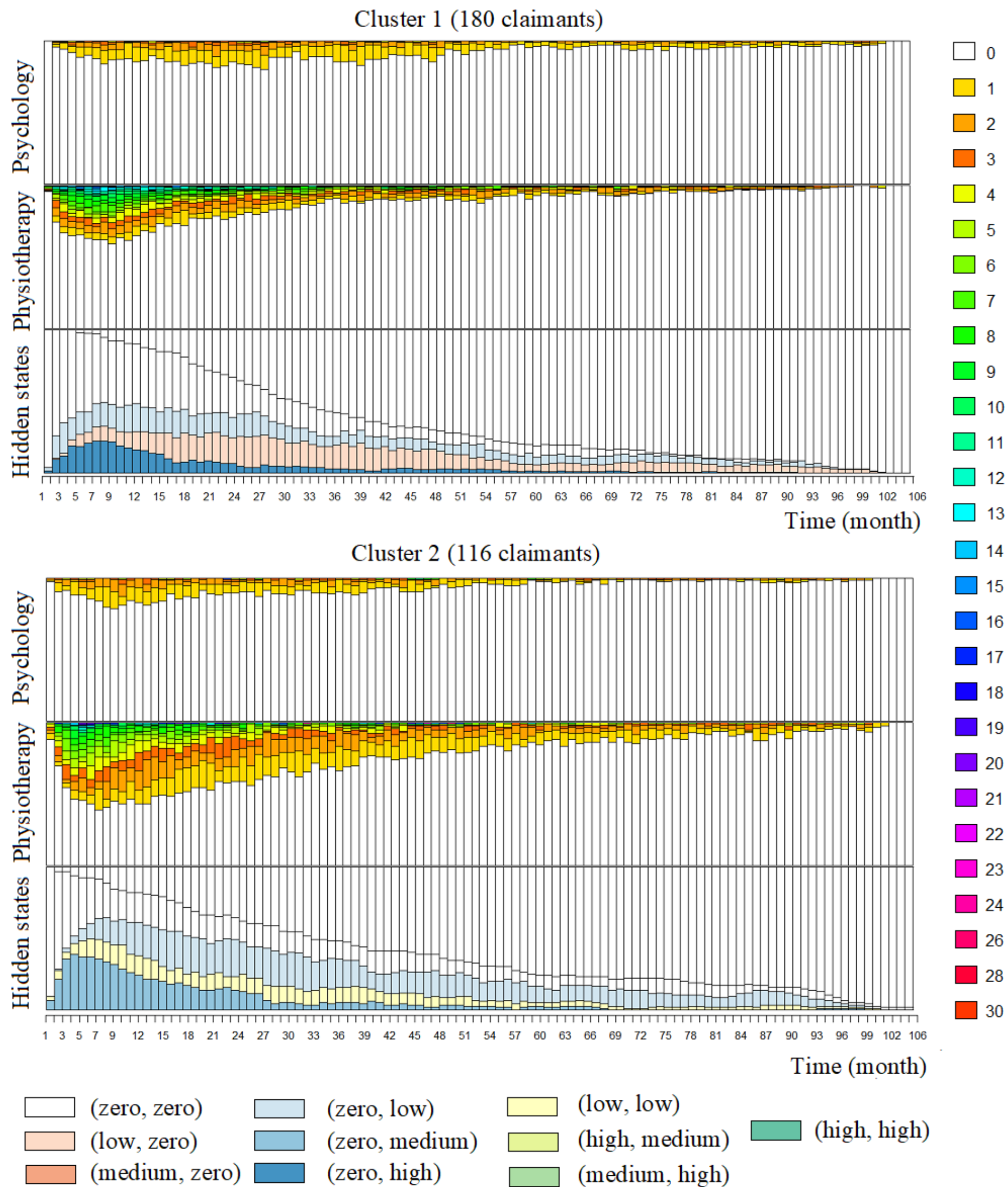

Figure 4: **Stacked plots of the number of monthly utilizations of psychology and physiotherapy services and stacked "state paths" for cluster 1 (180 claimants) and cluster 2 (116 claimants).** The stacked "state paths" chart shows the traversed sequences of states of the claimants in each cluster. The legend to the right shows monthly service utilization in the time series and the legend below shows the interpretation of the hidden states in the clusters.

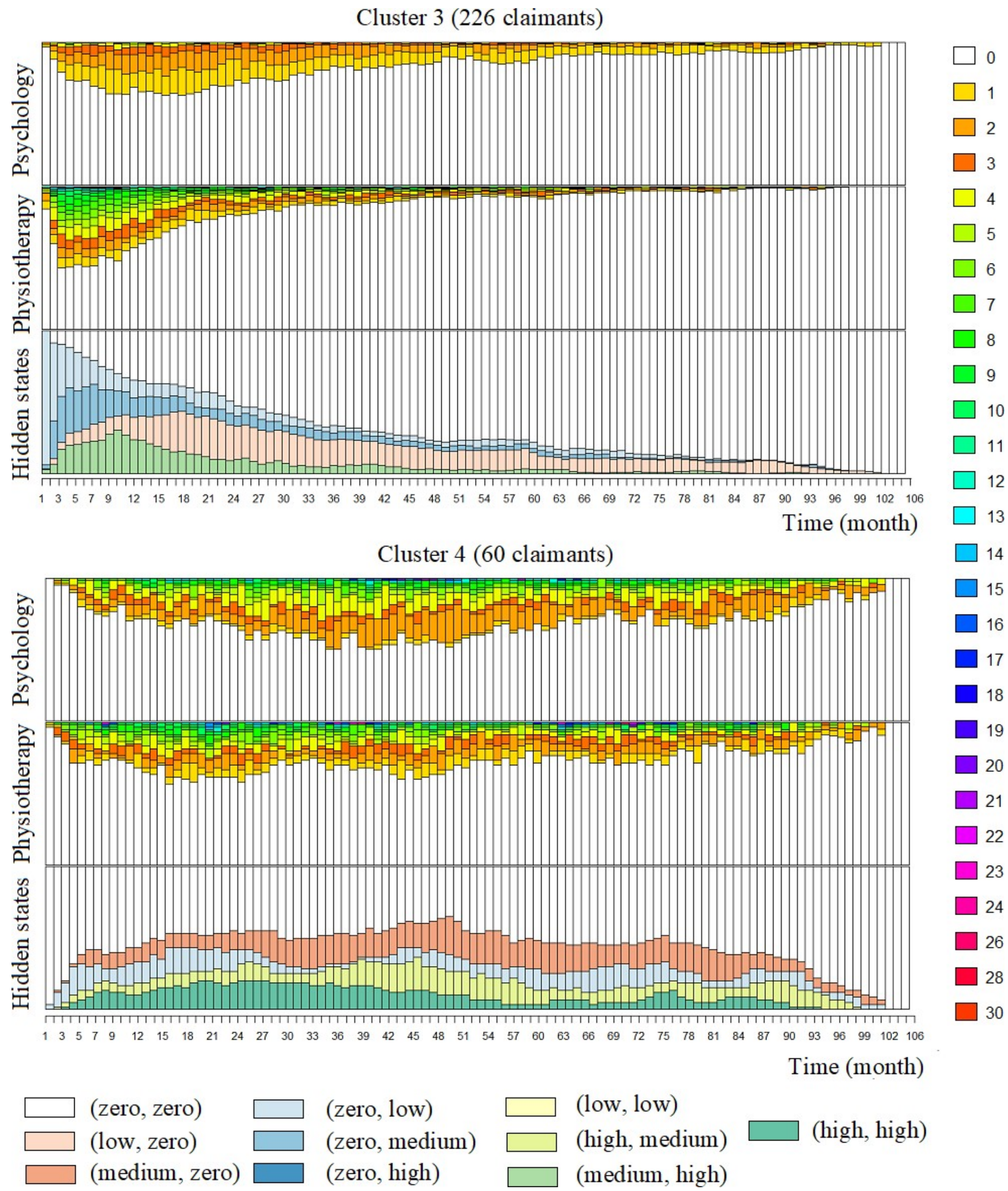

Figure 5: **Stacked plots of the number of monthly utilizations of psychology and physiotherapy services and stacked "state paths" for cluster 3 (226 claimants) and cluster 4 (60 claimants).**

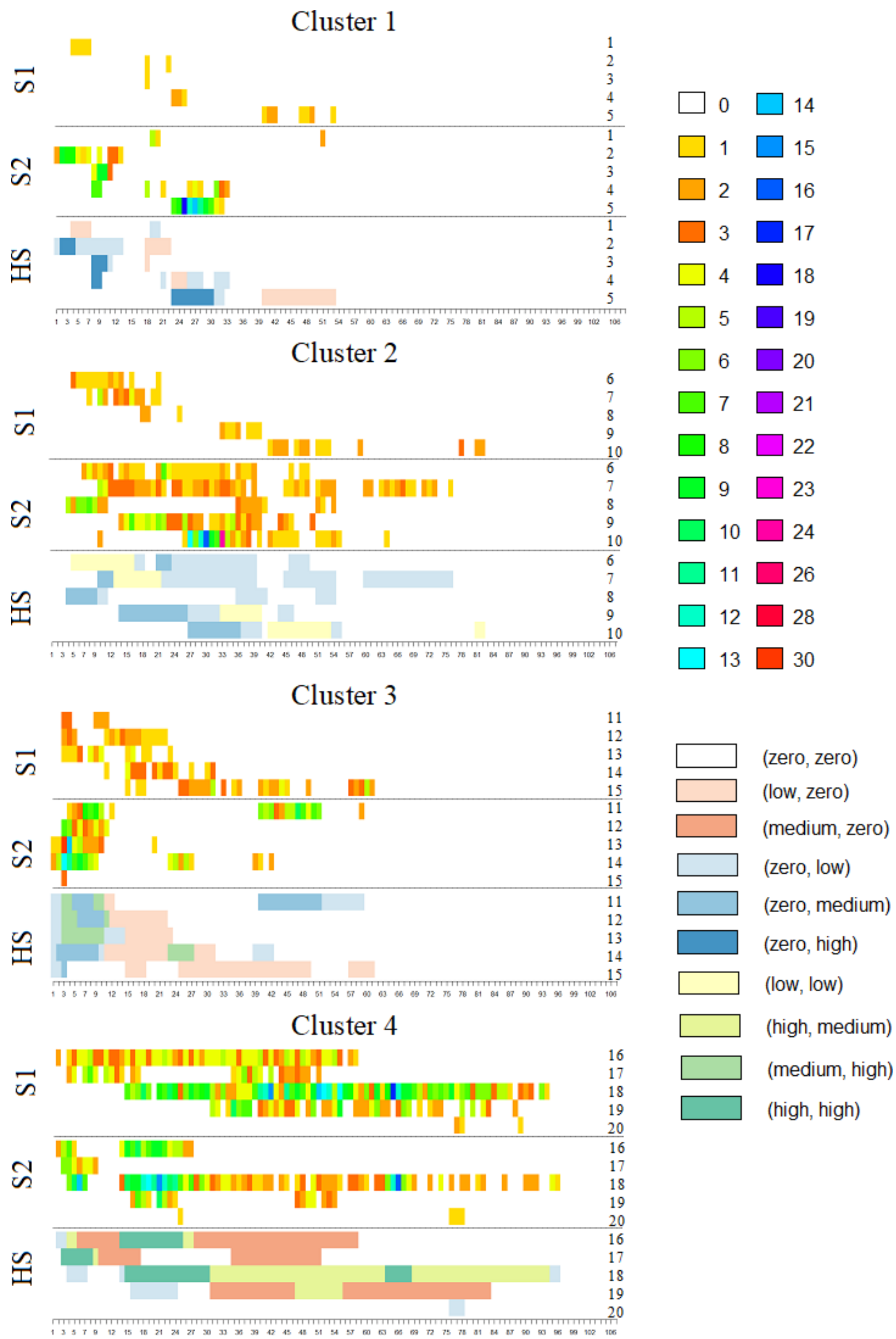

Figure 6: **Individual trajectories of the number of monthly and "state paths" for five random claimants per cluster (20 claimants in total)**. These plots confirm the different utilization patterns in the four clusters. The plots show the good correspondence between the transitions in amounts of service utilizations and the transitions in inferred states, giving further evidence to the effective empirical fit of the model. HS, Hidden States. S1, Service 1 (psychology). S2, Service 2 (physiotherapy).

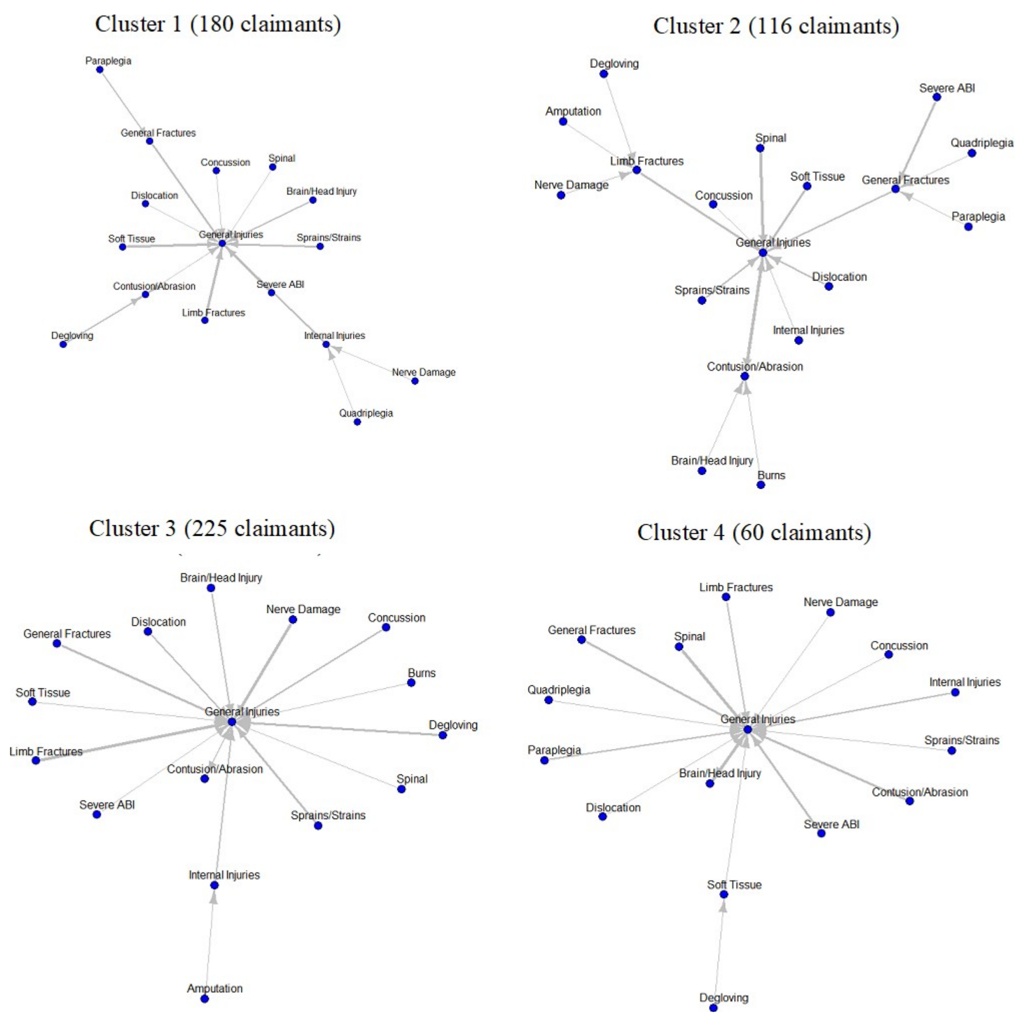

Figure 7: **"Weighted directed graph" of the most likely injuries for all clusters**

associated with internal injuries and negatively associated with dislocation injuries.

Odds ratios derived using multivariate logistic regression demonstrated that patients with nerve damage were almost 7 times more likely to be members of cluster 1 than patients with general injuries. Patients with severe ABI were 9.5 times more likely to be members of cluster 3 than patients with general injuries. Patients with severe brain/head injury were 9.9 times more likely to be members of cluster 3 than patients with general injuries. Patients with a psychiatric injury were almost 5 times more likely to be members of cluster 3 than patients with general injuries. Patients with internal injuries were almost twice as likely to be a member of cluster 4 than patients with other general injuries.

XGBM, RF and SVM classifiers, coupled with SMOTE, were capable of predicting cluster membership, based on demographic, injury and accident covariates, with moderate to strong performance (Table 6). Mean AUC statistics ranged from 0.62 to 0.96. The 470 clusters were formed based on a particular sequential model and time series, and it was not unexpected that the predictive accuracy based on the given (static) covariates may vary significantly across clusters. Results demonstrated that demographic, injury and accident covariates could be used to effectively predict future psychology and physiotherapy service utilization patterns after a traffic accident, with the potential to facilitate service planning. The SHAP plots of Figs 8 and 9 visualize feature importance for each of the four XGBM classifiers. The SHAP results were consistent with the statistical results gleaned from the multinomial logistic regression analysis. For example, the SHAP plot of cluster 4 shows that internal injuries exerted a heavy positive influence on the classification of cluster 4 membership. Correspondingly, the internal injuries covariate was significantly and positively associated with membership of cluster 4 in the multinomial logistic regression analysis.

Discussion, future work and conclusions

The purpose of this research was to identify the main, distinct patterns of joint service utilization behavior of multiple health services from a sample of MVA sufferers. The resulting analysis can inform policy makers and facilitate resource planning, intervention design and be used to improve MVA claimant recovery trajectories. To analyze claimant behaviors in terms of multiple service utilizations over time, we have used an approach known as the multichannel mixture of hidden Markov models (MHMM). To the best of our knowledge, this is the first application of this model to the joint analysis of the utilization of multiple healthcare services.

Given the synergistic relationship between physiotherapy and psychology services [29, 30], we have chosen these two health services to illustrate the capabilities of our joint analysis. However, the same approach can be applied to any other combination

Table 4: **Multinomial logistic regression results demonstrating statistical associations.**

Note: Cluster 1 used as reference. ABI, acquired brain injury.

	Covariate	coef	std err	z	P>[z]	[0.025	0.975]
Cluster 2	Male	0.11	0.24	0.45	0.66	-0.37	0.59
	Age 30-40 years	0.16	0.34	0.47	0.64	-0.51	0.83
	Age 41-50 years	-0.22	0.33	-0.67	0.50	-0.87	0.43
	Age 51-60 years	-0.05	0.35	-0.14	0.89	-0.74	0.64
	Age >60 years	-0.08	0.44	-0.18	0.86	-0.95	0.79
	Injury: Severe ABI	1.57	0.99	1.58	0.11	-0.37	3.51
	Injury: Brain Head	0.28	0.35	0.82	0.42	-0.40	0.96
	Injury: Other Spinal	0.49	0.46	1.08	0.28	-0.41	1.40
	Injury: Internal Injuries	-0.04	0.30	-0.12	0.90	-0.63	0.55
	Injury: Degloving	-0.50	0.87	-0.57	0.57	-2.21	1.21
	Injury: Limb Fractures	-0.03	0.29	-0.12	0.90	-0.60	0.53
	Injury: Other Fractures	-0.19	0.28	-0.68	0.50	-0.74	0.36
	Injury: Dislocation	0.30	0.29	1.01	0.31	-0.28	0.87
	Injury: Soft Tissue	-0.23	0.27	-0.84	0.40	-0.76	0.31
	Injury: Sprains Strains	0.16	0.26	0.62	0.53	-0.35	0.67
	Injury: Concussion	-0.10	0.35	-0.29	0.77	-0.78	0.58
	Injury: Nerve Damage	-1.87	1.13	-1.67	0.10	-4.08	0.33
	Injury: Contusion Abrasion	0.26	0.27	0.95	0.34	-0.27	0.79
	Injury: Psych Flag	-0.83	1.64	-0.51	0.61	-4.05	2.39
	Pedestrian	0.52	0.36	1.46	0.15	-0.18	1.22
	Motorcyclist	0.26	0.40	0.63	0.53	-0.53	1.04
	Car passenger	0.17	0.33	0.50	0.61	-0.48	0.82
	Cyclist	-0.36	0.65	-0.56	0.58	-1.62	0.91
	Constant	-0.86	0.40	-2.12	0.03*	-1.65	-0.07
Cluster 3	Male	-0.41	0.43	-0.95	0.34	-1.25	0.43
	Age 30-40 years	-0.70	0.56	-1.24	0.21	-1.80	0.40
	Age 41-50 years	-0.35	0.52	-0.67	0.50	-1.37	0.67
	Age 51-60 years	-1.07	0.69	-1.55	0.12	-2.42	0.29
	Age >60 years	-1.35	0.92	-1.48	0.14	-3.14	0.44
	Injury: Severe ABI	3.38	0.93	3.64	0.00**	1.56	5.20
	Injury: Brain Head	2.42	0.50	4.83	0.00**	1.43	3.40
	Injury: Other Spinal	1.04	0.59	1.75	0.08	-0.13	2.20
	Injury: Internal Injuries	-0.11	0.48	-0.22	0.83	-1.05	0.84
	Injury: Degloving	-1.36	1.71	-0.79	0.43	-4.72	2.00
	Injury: Limb Fractures	0.20	0.45	0.43	0.67	-0.69	1.08
	Injury: Other Fractures	0.14	0.49	0.28	0.78	-0.82	1.09
	Injury: Dislocation	0.20	0.48	0.42	0.68	-0.74	1.13
	Injury: Soft Tissue	-0.05	0.46	-0.11	0.91	-0.96	0.86
	Injury: Sprains Strains	0.00	0.46	0.01	1.00	-0.90	0.91
	Injury: Concussion	-0.73	0.53	-1.36	0.17	-1.77	0.32
	Injury: Nerve Damage	-0.75	1.10	-0.68	0.49	-2.90	1.40
	Injury: Contusion Abrasion	-0.63	0.44	-1.44	0.15	-1.48	0.23
	Injury: Psych Flag	1.24	1.32	0.94	0.35	-1.34	3.82
	Pedestrian	-0.40	0.67	-0.60	0.55	-1.71	0.91
	Motorcyclist	-0.39	0.71	-0.55	0.59	-1.79	1.01
	Car passenger	-0.26	0.53	-0.49	0.62	-1.31	0.78
	Cyclist	-0.98	0.95	-1.03	0.30	-2.85	0.88
	Constant	-1.79	0.64	-2.78	0.01*	-3.05	-0.53
Cluster 4	Male	0.26	0.22	1.18	0.24	-0.18	0.70
	Age 30-40 years	-0.02	0.31	-0.05	0.96	-0.62	0.58
	Age 41-50 years	-0.37	0.29	-1.25	0.21	-0.94	0.21
	Age 51-60 years	-0.57	0.34	-1.69	0.09	-1.23	0.09
	Age >60 years	-0.24	0.39	-0.62	0.53	-1.00	0.52
	Injury: Severe ABI	1.63	0.93	1.75	0.08	-0.19	3.46
	Injury: Brain Head	0.14	0.33	0.44	0.66	-0.50	0.79
	Injury: Other Spinal	0.86	0.40	2.15	0.03*	0.08	1.64
	Injury: Internal Injuries	0.60	0.26	2.29	0.02*	0.09	1.11
	Injury: Degloving	0.03	0.70	0.04	0.97	-1.34	1.40
	Injury: Limb Fractures	0.01	0.27	0.03	0.98	-0.51	0.53
	Injury: Other Fractures	-0.38	0.26	-1.44	0.15	-0.89	0.14
	Injury: Dislocation	-0.47	0.30	-1.58	0.11	-1.06	0.11
	Injury: Soft Tissue	-0.11	0.25	-0.44	0.66	-0.61	0.39
	Injury: Sprains Strains	-0.08	0.24	-0.33	0.75	-0.56	0.40
	Injury: Concussion	-0.02	0.32	-0.07	0.94	-0.65	0.60
	Injury: Nerve Damage	-2.51	1.11	-2.26	0.02*	-4.69	-0.33
	Injury: Contusion Abrasion	-0.11	0.24	-0.45	0.65	-0.57	0.36
	Injury: Psych Flag	-0.59	1.38	-0.43	0.67	-3.29	2.11
	Pedestrian	-0.02	0.36	-0.06	0.95	-0.73	0.69
	Motorcyclist	0.11	0.38	0.30	0.77	-0.63	0.86
	Car passenger	0.21	0.29	0.71	0.48	-0.36	0.78
	Cyclist	-0.17	0.57	-0.31	0.76	-1.29	0.94
	Constant	-0.09	0.36	-0.25	0.81	-0.80	0.62

Table 5: **Marginal effects associated with membership of each of the four clusters.** No marginal effects were significant for cluster 2. ABI, acquired brain injury.

	Covariate	dy/dx	std err	z	P>—z—	[0.025	0.975]
Cluster 1	Male	-0.03	0.04	-0.76	0.45	-0.11	0.05
	Age 30-40 years	0.00	0.06	0.02	0.98	-0.11	0.12
	Age 41-50 years	0.07	0.06	1.22	0.22	-0.04	0.18
	Age 51-60 years	0.09	0.06	1.48	0.14	-0.03	0.21
	Age >60 years	0.06	0.07	0.81	0.42	-0.08	0.20
	Injury: Severe ABI	-0.38	0.19	-2.04	0.04*	-0.75	-0.01
	Injury: Brain Head	-0.09	0.06	-1.46	0.15	-0.20	0.03
	Injury: Other Spinal	-0.16	0.08	-2.06	0.04*	-0.31	-0.01
	Injury: Internal Injuries	-0.06	0.05	-1.28	0.20	-0.16	0.03
	Injury: Degloving	0.06	0.14	0.46	0.64	-0.21	0.33
	Injury: Limb Fractures	0.00	0.05	-0.03	0.97	-0.10	0.10
	Injury: Other Fractures	0.06	0.05	1.19	0.23	-0.04	0.15
	Injury: Dislocation	0.03	0.05	0.49	0.63	-0.08	0.13
	Injury: Soft Tissue	0.03	0.05	0.70	0.48	-0.06	0.13
	Injury: Sprains Strains	0.00	0.05	-0.10	0.92	-0.09	0.08
	Injury: Concussion	0.02	0.06	0.41	0.68	-0.09	0.14
	Injury: Nerve Damage	0.46	0.17	2.66	0.01*	0.12	0.80
	Injury: Contusion Abrasion	0.00	0.05	0.05	0.96	-0.09	0.09
	Injury: Psych Flag	0.11	0.28	0.40	0.69	-0.44	0.67
	Pedestrian	-0.03	0.07	-0.51	0.61	-0.16	0.10
	Motorcyclist	-0.03	0.07	-0.38	0.70	-0.17	0.11
	Car passenger	-0.03	0.06	-0.60	0.55	-0.14	0.08
	Cyclist	0.07	0.11	0.65	0.52	-0.14	0.27
Cluster 3	Male	-0.03	0.02	-1.32	0.19	-0.07	0.01
	Age 30-40 years	-0.04	0.03	-1.38	0.17	-0.10	0.02
	Age 41-50 years	-0.01	0.03	-0.30	0.77	-0.06	0.04
	Age 51-60 years	-0.04	0.04	-1.24	0.21	-0.12	0.03
	Age >60 years	-0.07	0.05	-1.38	0.17	-0.16	0.03
	Injury: Severe ABI	0.13	0.03	4.35	0.00**	0.07	0.18
	Injury: Brain Head	0.12	0.03	4.68	0.00**	0.07	0.18
	Injury: Other Spinal	0.03	0.03	1.08	0.28	-0.03	0.09
	Injury: Internal Injuries	-0.02	0.02	-0.73	0.46	-0.07	0.03
	Injury: Degloving	-0.07	0.09	-0.74	0.46	-0.24	0.11
	Injury: Limb Fractures	0.01	0.02	0.47	0.64	-0.03	0.06
	Injury: Other Fractures	0.02	0.03	0.72	0.47	-0.03	0.07
	Injury: Dislocation	0.02	0.02	0.69	0.49	-0.03	0.06
	Injury: Soft Tissue	0.00	0.02	0.12	0.91	-0.04	0.05
	Injury: Sprains Strains	0.00	0.02	-0.02	0.98	-0.05	0.05
	Injury: Concussion	-0.04	0.03	-1.39	0.16	-0.09	0.02
	Injury: Nerve Damage	0.04	0.06	0.68	0.50	-0.07	0.15
	Injury: Contusion Abrasion	-0.04	0.02	-1.58	0.11	-0.08	0.01
	Injury: Psych Flag	0.09	0.04	2.13	0.03*	0.01	0.18
	Pedestrian	-0.03	0.03	-0.84	0.40	-0.10	0.04
	Motorcyclist	-0.03	0.04	-0.74	0.46	-0.10	0.05
	Car passenger	-0.02	0.03	-0.77	0.44	-0.07	0.03
	Cyclist	-0.04	0.05	-0.91	0.36	-0.14	0.05
Cluster 4	Male	0.05	0.04	1.36	0.18	-0.02	0.13
	Age 30-40 years	0.00	0.05	0.03	0.98	-0.10	0.11
	Age 41-50 years	-0.05	0.05	-0.99	0.32	-0.15	0.05
	Age 51-60 years	-0.09	0.06	-1.43	0.15	-0.21	0.03
	Age >60 years	-0.01	0.07	-0.20	0.84	-0.15	0.12
	Injury: Severe ABI	0.15	0.13	1.17	0.24	-0.11	0.41
	Injury: Brain Head	-0.04	0.06	-0.72	0.47	-0.15	0.07
	Injury: Other Spinal	0.12	0.07	1.79	0.07	-0.01	0.25
	Injury: Internal Injuries	0.12	0.05	2.74	0.01*	0.04	0.21
	Injury: Degloving	0.07	0.13	0.51	0.61	-0.19	0.32
	Injury: Limb Fractures	0.00	0.05	-0.01	0.99	-0.09	0.09
	Injury: Other Fractures	-0.07	0.05	-1.40	0.16	-0.16	0.03
	Injury: Dislocation	-0.12	0.05	-2.23	0.03*	-0.22	-0.01
	Injury: Soft Tissue	-0.01	0.05	-0.15	0.88	-0.10	0.08
	Injury: Sprains Strains	-0.03	0.04	-0.59	0.55	-0.11	0.06
	Injury: Concussion	0.02	0.06	0.31	0.76	-0.09	0.13
	Injury: Nerve Damage	-0.36	0.22	-1.68	0.09	-0.79	0.06
	Injury: Contusion Abrasion	-0.02	0.04	-0.56	0.57	-0.11	0.06
	Injury: Psych Flag	-0.09	0.21	-0.44	0.66	-0.51	0.32
	Pedestrian	-0.03	0.06	-0.44	0.66	-0.15	0.10
	Motorcyclist	0.01	0.07	0.22	0.83	-0.12	0.15
	Car passenger	0.04	0.05	0.70	0.48	-0.07	0.14
	Cyclist	0.01	0.10	0.09	0.93	-0.19	0.21

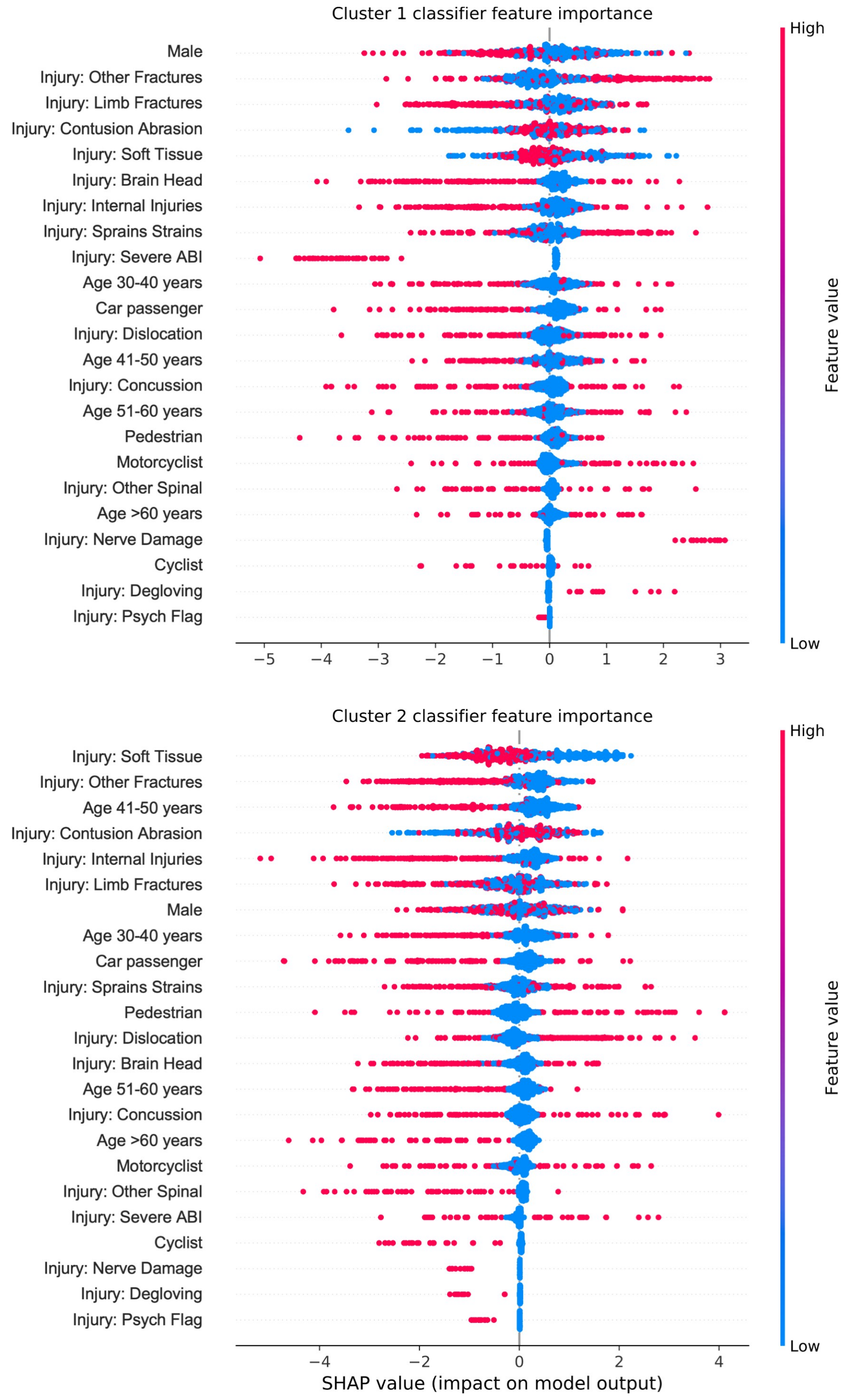

Figure 8: **SHAP feature importance plots for the Cluster 1 and 2 XGBM classifiers.** The feature values are binary, so the values of each appear as either blue (zero, absent) or red (one, present) points, with the vertical spread on each line indicating relative case frequency. Features most strongly related to classification of the cluster in question appear at the top of each plot. The plots show that the presence (red) or absence (blue) of a covariate increases or decreases (or has little or no effect on) the likelihood of classifying the cluster membership, with positive or negative SHAP values, respectively, displayed on the X-axis. ABI: acquired brain injury.

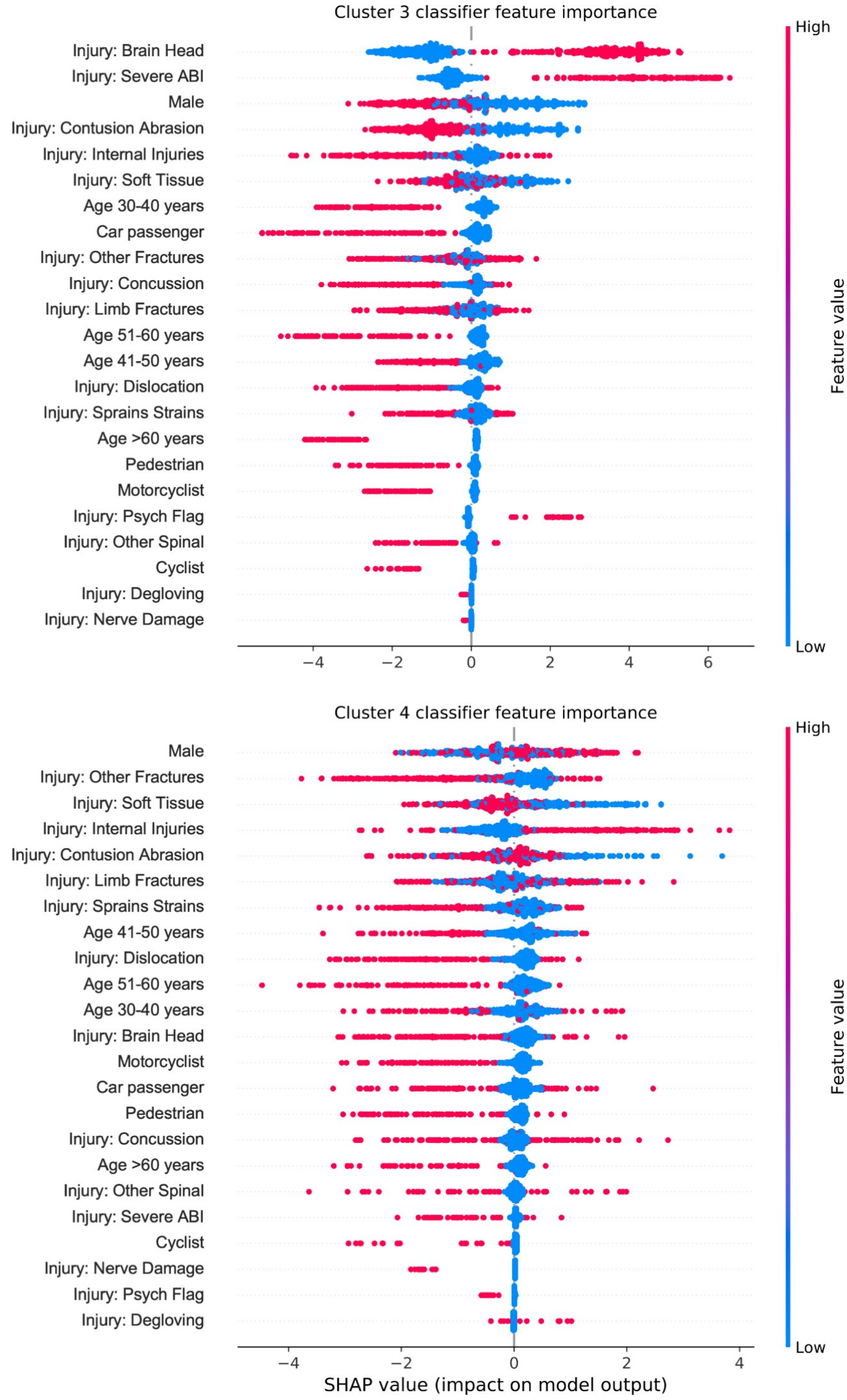

Figure 9: **SHAP feature importance plots for the Cluster 3 and 4 XGBM classifiers.** The feature values are binary, so the values of each appear as either blue (zero, absent) or red (one, present) points, with the vertical spread on each line indicating relative case frequency. Features most strongly related to classification of the cluster in question appear at the top of each plot. The plots show that the presence (red) or absence (blue) of a covariate increases or decreases (or has little or no effect on) the likelihood of classifying the cluster membership, with positive or negative SHAP values, respectively, displayed on the X-axis. ABI, acquired brain injury.

of services. After an extensive analysis with nearly nine years of compensation data for these two services, we determined that four clusters of utilization patterns, each with five hidden states, comprised the most suitable model. We have interpreted the hidden state of the MHMM as the latent “level of service” of the joint physiotherapy and psychology services. By examining the relationship between hidden states and observations across the entire population, we have identified four levels for the individual services (“zero”, “low”, “medium” and “high”) and ten distinct values for the joint services ((zero, zero), (zero, low), (zero, medium), (zero, high), (low, zero), (low, low), (medium, zero), (medium, high), (high, medium), and (high, high)). We have used these ten values as the distinct values of the hidden state, and employed them extensively to characterize and contrast the found clusters. While alternative models and factorizations could be used for this task [66, 67, 68, 69], the adopted model has displayed good fitting and interpretability.

It was evident that a small number of injury covariates were significantly related to membership of three of the four derived clusters. These significant relationships offer insight into the clinical theme of these clusters and, from a clinical perspective, are logically supported by the services and levels of service accessed by patients. The first found cluster contained claimants who used psychology services occasionally soon after their accident, and physiotherapy services for a relatively short period of time. Based on the analysis of co-occurring injuries, the members of this cluster seemed to be those who suffered less serious injuries than in other clusters. Cluster 1 was defined by nerve damage but not ABI or spinal injuries, suggesting a group of lower acuity patients with relatively minor injuries. Patients in this cluster required the lowest level of psychology and physiotherapy support.

The second found cluster contained claimants who used services more intensely and for a longer duration than those in cluster 1. Its graph of co-occurring injuries showed two injuries – amputation and burns – that did not appear in cluster 1, and a higher co-frequency of spinal injuries. However, statistical relationships with covariates were not forthcoming and did not suggest whether these injuries were major or minor. Assuming that injuries were generally more serious than those in cluster 1, the imbalance between physiotherapy and psychology services may suggest that seeking more psychology services should be explored as a possible early intervention for the members of this cluster.

The third cluster contained claimants who sought substantial psychological help alongside their physiotherapy treatment. These claimants used psychology services moderately but persistently. Unlike in the two previous clusters, physiotherapy service utilization decayed more rapidly than psychology service utilization. Cluster 3 was defined statistically by severe ABI, brain/head injuries and psychiatric injury.

These higher acuity patients required a higher level of care to treat their accident-related injuries than those in clusters 1 and 2. A number of explanations may exist for this pattern of behavior: one is that seeking more substantial psychological treatment may lead to a quicker physical recovery. On the other hand, this population is likely to have suffered from more psychological harmful events than members of clusters 1 and 2.

The fourth and final cluster contained the claimants who used both services intensely even years after the accident. This cluster was defined primarily by internal injuries, which may have included severe events like splenic, vascular or bowel rupture, renal or liver laceration or cardiothoracic injuries [70]. Cluster 4 appeared to describe patients with high-severity injuries and these people required a much higher level of care over a longer period of time than those in any of the other clusters. The injury co-occurrence graph showed that the most frequently co-occurring injury with general injuries (which, by themselves, occured in 95% of cases) are brain injuries. From a cost perspective, the average cost per claimant for psychology and physiotherapy services in this cluster is at least four times that of any of the other clusters. Since the number of claimants in this cluster was small, it may be possible to contact the members individually to acquire additional information about their case and design tailored interventions to facilitate more rapid recovery.

Next, this paper deployed supervised machine learning methods (XGBM, RF and SVM) to investigate claimant profiles associated with each cluster. The supervised classifiers, coupled with SMOTE, were able to effectively predict cluster membership using data available at the time of the accident. Statistical relationships were supported by the calculated SHAP values elucidating feature importance within the XGBM models. Results demonstrated that demographic, injury and accident covariates could be used to effectively predict future psychology and physiotherapy service utilization patterns after a traffic accident. These methods can be used to predict the long term care needs of patients and facilitate care planning [71, 72].

With regard to limitations, the scope of the service analysis was undeniably narrow. The focus on physiotherapy and psychology services provides some indication of injury severity and patient trajectory. However, the inclusion of medical services in the analysis would be useful to facilitate the interpretation of the identified clusters and guide clinical practice and policy improvements with greater confidence and precision. The analysis of other combinations of clinical services is an opportunity for future research projects. Another limitation of this project was that the supervised learning analysis yielded mixed results, and these models are therefore of mixed practical utility.

With regard to future work, we plan to use these methods to investigate additional clinical research questions. Time-series clustering, coupled with supervised learning approaches, appear to be useful in generating insights that may beneficially guide policy development and clinical care. It may be clinically useful to investigate the prevalence of post-traumatic stress disorder (PTSD) among claimants who have suffered a whiplash injury. It is reported that whiplash injury and PTSD following motor vehicle accidents are comorbid, and there is evidence that symptoms of PTSD are prevalent in whiplash-injured individuals following an accident, with approximately 25% meeting the diagnostic criteria for PTSD [73, 74]. There is an ongoing debate between researchers regarding prognostic factors; however, psychological factors including PTSD have been confirmed as indicators of a poor prognosis[73]. A multichannel MHMM analysis may shed light on the service utilization of this group of claimants and may improve guidance of their recovery. Furthermore, we plan to apply the clustering approach to all health services that a claimant might use. In this case, the number of time series for each claimant will be different and the number of channels will therefore be different for each claimant, requiring slight adaptations to our model. All in all, our primary goal is to better describe the behaviors of claimants to help them to "get back on track".

Conflict of interest statement

The authors declare that this article was composed in the absence of commercial or financial relationships that could be construed as a conflict of interest.

References

[1] C. Staton, J. Vissoci, E. Gong, N. Toomey, R. Wafula, J. Abdelgadir, Y. Zhou, C. Liu, F. Pei, B. Zick, et al., Road traffic injury prevention initiatives: a systematic review and metasummary of effectiveness in low and middle income countries, PLoS One 11 (1) (2016).

[2] M. Brussoni, M. A. George, A. Jin, O. Amram, R. McCormick, C. E. Lalonde, Hospitalizations due to unintentional transport injuries among aboriginal population of British Columbia, Canada: incidence, changes over time and ecological analysis of risk markers, PloS one 13 (1) (2018).

[3] W. H. O. (WHO), Road traffic injuries. March, 2018 (2018).

[4] S. E. Gray, B. Hassani-Mahmooei, I. D. Cameron, E. Kendall, J. Kenardy, A. Collie, Patterns and predictors of failed and sustained return-to-work in transport injury insurance claimants, Journal of occupational rehabilitation 28 (4) (2018) 740–748.

[5] G. Henley, J. Harrison, et al., Trends in serious injury due to land transport accidents, Australia 2000-01 to 2007-08., Injury Research and Statistics (56) (2010).

[6] R. Lilley, G. Davie, S. Ameratunga, S. Derrett, Factors predicting work status 3 months after injury: results from the prospective outcomes of injury study, BMJ open 2 (2) (2012) e000400.

[7] G. Henley, J. E. Harrison, Trends in injury deaths, Australia: 1999-00 to 2009-10 (2015).

[8] T. Bureau of Infrastructure, R. E. (BITRE)., Road trauma Australia 2018 statistical summary. september, 2018.

[9] F. Litchfield, The cost of road crashes in Australia 2016: an overview of safety strategies, The Australian National University (2016).

[10] M. State Government of Victoria, 2015 Victorian road trauma: analysis of fatalities and serious injuries; 2016 (2016).

[11] M. Transport Accident Commission, TAC annual report; 2017 (2017).

[12] R. Ruseckaite, B. Gabbe, A. P. Vogel, A. Collie, Health care utilisation following hospitalisation for transport-related injury, Injury 43 (9) (2012) 1600–1605.

[13] N. Esmaili, M. Piccardi, B. Kruger, F. Girosi, Analysis of healthcare service utilization after transport-related injuries by a mixture of hidden Markov models, PLoS one 13 (11) (2018).

[14] W. H. O. (WHO)., Care of the injured. January, 2019.

[15] W. H. O. (WHO)., Delivering post-crash care. March, 2018.

[16] K.-H. Prang, R. Ruseckaite, A. Collie, Healthcare and disability service utilization in the 5-year period following transport-related traumatic brain injury, Brain injury 26 (13-14) (2012) 1611–1620.

[17] S. Guilcher, S. E. Munce, C. Couris, K. Fung, B. C. Craven, M. Verrier, S. Jaglal, Health care utilization in non-traumatic and traumatic spinal cord injury: a population-based study, Spinal Cord 48 (1) (2010) 45–50.

[18] P. M. Ullrich, R. K. Lincoln, M. J. Tackett, S. Miskevics, B. M. Smith, F. M. Weaver, Pain, depression, and health care utilization over time after spinal cord injury., Rehabilitation psychology 58 (2) (2013) 158.

[19] I. A. Harris, D. F. Murgatroyd, I. D. Cameron, J. M. Young, M. J. Solomon, The effect of compensation on health care utilisation in a trauma cohort, Medical Journal of Australia 190 (11) (2009) 619–622.

[20] G. Das, K.-I. Lin, H. Mannila, G. Renganathan, P. Smyth, Rule discovery from time series., in: KDD, Vol. 98, 1998, pp. 16–22.

[21] S. Aghabozorgi, A. S. Shirkhorshidi, T. Y. Wah, Time-series clustering–a decade review, Information Systems 53 (2015) 16–38.

[22] T. W. Liao, Clustering of time series dataa survey, Pattern recognition 38 (11) (2005) 1857–1874.

[23] T. Zhongda, L. Shujiang, W. Yanhong, S. Yi, A prediction method based on wavelet transform and multiple models fusion for chaotic time series, Chaos, Solitons & Fractals 98 (2017) 158–172.

[24] Z. Tian, Chaotic characteristic analysis of network traffic time series at different time scales, Chaos, Solitons & Fractals 130 (2020) 109412.

[25] Z. Tian, Echo state network based on improved fruit fly optimization algorithm for chaotic time series prediction, Journal of Ambient Intelligence and Humanized Computing (2020) 1–20.

[26] E. Keogh, S. Chu, D. Hart, M. Pazzani, Segmenting time series: A survey and novel approach, in: Data mining in time series databases, World Scientific, 2004, pp. 1–21.

[27] A. Wismüller, O. Lange, D. R. Dersch, G. L. Leinsinger, K. Hahn, B. Pütz, D. Auer, Cluster analysis of biomedical image time-series, International Journal of Computer Vision 46 (2) (2002) 103–128.

[28] M. Van Den Heuvel, R. Mandl, H. H. Pol, Normalized cut group clustering of resting-state fmri data, PloS one 3 (4) (2008) e2001.

[29] V. Harding, A. C. Williams, Applying psychology to enhance physiotherapy outcome (1995).

[30] C. J. Barron, J. A. Klaber Moffett, M. Potter, Patient expectations of physiotherapy: definitions, concepts, and theories, Physiotherapy Theory and Practice 23 (1) (2007) 37–46.

[31] G. J. McLachlan, D. Peel, Finite mixture models, John Wiley & Sons, 2004.

[32] A. Bargi, R. Y. Da Xu, M. Piccardi, AdOn HDP-HMM: an adaptive online model for segmentation and classification of sequential data, IEEE Transactions on Neural Networks and Learning Systems 29 (9) (2017) 3953–3968.

[33] M. Gales, S. Young, et al., The application of hidden Markov models in speech recognition, Foundations and Trends® in Signal Processing 1 (3) (2008) 195–304.

[34] G. Zhou, J. Su, Named entity recognition using an HMM-based chunk tagger, in: proceedings of the 40th Annual Meeting on Association for Computational Linguistics, Association for Computational Linguistics, 2002, pp. 473–480.

[35] M. R. Hassan, B. Nath, Stock market forecasting using hidden Markov model: a new approach, in: 5th International Conference on Intelligent Systems Design and Applications (ISDA'05), IEEE, 2005, pp. 192–196.

[36] K. Munch, A. Krogh, Automatic generation of gene finders for eukaryotic species, BMC bioinformatics 7 (1) (2006) 263.

[37] Z. Huang, W. Dong, F. Wang, H. Duan, Medical inpatient journey modeling and clustering: a Bayesian hidden Markov model based approach, in: AMIA Annual Symposium Proceedings, Vol. 2015, American Medical Informatics Association, 2015, p. 649.

[38] L. R. Rabiner, A tutorial on hidden Markov models and selected applications in speech recognition, Proceedings of the IEEE 77 (2) (1989) 257–286.

[39] S. Helske, J. Helske, Mixture hidden Markov models for sequence data: The seqHMM package in R, arXiv preprint arXiv:1704.00543 (2017).

[40] R. C. Team, et al., R: A language and environment for statistical computing (2013).

[41] G. Csardi, T. Nepusz, et al., The igraph software package for complex network research, InterJournal, complex systems 1695 (5) (2006) 1–9.

[42] H. Wickham, ggplot2: elegant graphics for data analysis, Springer, 2016.

[43] E. Neuwirth, R. C. Brewer, Colorbrewer palettes, R package version (2014) 1–1.

[44] G. Schwarz, et al., Estimating the dimension of a model, The annals of statistics 6 (2) (1978) 461–464.

[45] H. Akaike, A new look at the statistical model identification, IEEE transactions on automatic control 19 (6) (1974) 716–723.

[46] M. Costa, L. De Angelis, Model selection in hidden Markov models: a simulation study (2010).

[47] N. Dridi, M. Hadzagic, Akaike and Bayesian information criteria for hidden markov models, IEEE Signal Processing Letters 26 (2) (2018) 302–306.

[48] S. Akogul, M. Erisoglu, A comparison of information criteria in clustering based on mixture of multivariate normal distributions, Mathematical and Computational Applications 21 (3) (2016) 34.

[49] S. Bacci, S. Pandolfi, F. Pennoni, A comparison of some criteria for states selection in the latent Markov model for longitudinal data, Advances in Data Analysis and Classification 8 (2) (2014) 125–145.

[50] F. Farrokhi, Q. D. Buchlak, M. Sikora, N. Esmaili, M. Marsans, P. McLeod, J. Mark, E. Cox, C. Bennett, J. Carlson, Investigating risk factors and predicting complications in deep brain stimulation surgery with machine learning algorithms, World Neurosurgery 134 (2020) e325–e338.

[51] Q. D. Buchlak, V. Yanamadala, J.-C. Leveque, A. Edwards, K. Nold, R. Sethi, The seattle spine score: predicting 30-day complication risk in adult spinal deformity surgery, Journal of Clinical Neuroscience 43 (2017) 247–255.

[52] S. Seabold, J. Perktold, Statsmodels: Econometric and statistical modeling with Python, in: Proceedings of the 9th Python in Science Conference, Vol. 57, Scipy, 2010, p. 61.

[53] N. V. Chawla, K. W. Bowyer, L. O. Hall, W. P. Kegelmeyer, Smote: synthetic minority over-sampling technique, Journal of artificial intelligence research 16 (2002) 321–357.

[54] Y. Zhang, A. Haghani, A gradient boosting method to improve travel time prediction, Transportation Research Part C: Emerging Technologies 58 (2015) 308–324.

[55] J. Lu, D. Lu, X. Zhang, Y. Bi, K. Cheng, M. Zheng, X. Luo, Estimation of elimination half-lives of organic chemicals in humans using gradient boosting machine, Biochimica et Biophysica Acta (BBA)-General Subjects 1860 (11) (2016) 2664–2671.

[56] S. Touzani, J. Granderson, S. Fernandes, Gradient boosting machine for modeling the energy consumption of commercial buildings, Energy and Buildings 158 (2018) 1533–1543.

[57] R. Rawi, R. Mall, K. Kunji, C.-H. Shen, P. D. Kwong, G.-Y. Chuang, Parsnip: sequence-based protein solubility prediction using gradient boosting machine, Bioinformatics 34 (7) (2018) 1092–1098.

[58] J. Fan, X. Wang, L. Wu, H. Zhou, F. Zhang, X. Yu, X. Lu, Y. Xiang, Comparison of support vector machine and extreme gradient boosting for predicting daily global solar radiation using temperature and precipitation in humid subtropical climates: A case study in china, Energy Conversion and Management 164 (2018) 102–111.

[59] C. Bentéjac, A. Csörgő, G. Martínez-Muñoz, A comparative analysis of xgboost, arXiv preprint arXiv:1911.01914 (2019).

[60] A. Fernández, S. Garcia, F. Herrera, N. V. Chawla, Smote for learning from imbalanced data: progress and challenges, marking the 15-year anniversary, Journal of artificial intelligence research 61 (2018) 863–905.

[61] S. M. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions, in: Advances in neural information processing systems, 2017, pp. 4765–4774.

[62] S. M. Lundberg, G. Erion, H. Chen, A. DeGrave, J. M. Prutkin, B. Nair, R. Katz, J. Himmelfarb, N. Bansal, S.-I. Lee, From local explanations to global under-

standing with explainable ai for trees, Nature machine intelligence 2 (1) (2020) 2522–5839.

[63] S. M. Lundberg, B. Nair, M. S. Vavilala, M. Horibe, M. J. Eisses, T. Adams, D. E. Liston, D. K.-W. Low, S.-F. Newman, J. Kim, et al., Explainable machine-learning predictions for the prevention of hypoxaemia during surgery, Nature biomedical engineering 2 (10) (2018) 749–760.

[64] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al., Scikit-learn: Machine learning in Python, Journal of Machine Learning Research 12 (Oct) (2011) 2825–2830.

[65] T. Chen, C. Guestrin, Xgboost: A scalable tree boosting system, in: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, 2016, pp. 785–794.

[66] F. Bartolucci, Likelihood inference for a class of latent markov models under linear hypotheses on the transition probabilities, Journal of the Royal Statistical Society: Series B (Statistical Methodology) 68 (2) (2006) 155–178.

[67] F. Bartolucci, A. Farcomeni, F. Pennoni, Latent Markov models for longitudinal data, CRC Press, 2012.

[68] L. Catania, R. Di Mari, Hierarchical Markov-switching models for multivariate integer-valued time-series, Journal of Econometrics (2020).

[69] C. Keribin, Consistent estimation of the order of mixture models, Sankhyā: The Indian Journal of Statistics, Series A (2000) 49–66.

[70] J. M. Glover, M. F. Waychoff, M. Casmaer, M. D. April, C. J. Hunter, S. T. Trexler, L. H. Blackbourne, Association between seatbelt sign and internal injuries in the contemporary airbag era: A retrospective cohort study, The American journal of emergency medicine 36 (4) (2018) 545–550.

[71] Q. D. Buchlak, N. Esmaili, J.-C. Leveque, F. Farrokhi, C. Bennett, M. Piccardi, R. K. Sethi, Machine learning applications to clinical decision support in neurosurgery: an artificial intelligence augmented systematic review, Neurosurgical review (2019) 1–19.

[72] Q. D. Buchlak, N. Esmaili, J.-C. Leveque, C. Bennett, M. Piccardi, F. Farrokhi, Ethical thinking machines in surgery and the requirement for clinical leadership, The American Journal of Surgery (2020).

[73] T. E. Andersen, A. Elklit, L. Vase, The relationship between chronic whiplash-associated disorder and post-traumatic stress: attachment-anxiety may be a vulnerability factor, European Journal of Psychotraumatology 2 (1) (2011) 5633.

[74] M. Sterling, J. Hendrikz, J. Kenardy, Compensation claim lodgement and health outcome developmental trajectories following whiplash injury: a prospective study, Pain 150 (1) (2010) 22–28.

Table 6: **Mean 5-fold cross-validation performance of the supervised classifiers trained to classify membership of each of the four clusters.** Mean AUC is based on three sets of 5-fold cross-validation AUC statistics. AUC, area under the receiver operating characteristic curve.

	Mean 5-fold cross validation AUC		
Cluster index/classifier	XGBM	RF	SVM
Cluster 1	0.65	0.62	0.69
Cluster 2	0.76	0.68	0.79
Cluster 3	0.95	0.90	0.96
Cluster 4	0.71	0.65	0.74

Multichannel mixture models for time-series analysis and classification of engagement with multiple health services: an application to psychology and physiotherapy utilization patterns after traffic accidents

Nazanin Esmaili[a,b,*], Quinlan D. Buchlak[a], Massimo Piccardi[b], Bernie Kruger[d], Federico Girosi[c,e]

[a]School of Medicine, University of Notre Dame Australia, Sydney, NSW, Australia
[b]Faculty of Engineering and IT, University of Technology Sydney, NSW, Australia
[c]Capital Markets Cooperative Research Centre (CMCRC), Sydney, NSW, Australia
[d]Transport Accident Commission, Geelong, VIC, Australia
[e]Translational Health Research Institute, Western Sydney University, Penrith, NSW, Australia

Conflict of interest statement

The authors declare that this article was composed in the absence of commercial or financial relationships that could be construed as a conflict of interest.

*Corresponding author
Email address: nazanin.esmaili@uts.edu.au (Nazanin Esmaili)