

## PREDICTING 72-HOUR REATTENDANCE IN EMERGENCY DEPARTMENTS USING DISCRIMINANT ANALYSIS VIA MIXED INTEGER PROGRAMMING WITH ELECTRONIC MEDICAL RECORDS

FANWEN MENG\*, KIOK LIANG TEOW AND KELVIN WEE SHENG TEO

Department of Health Services and Outcomes Research  
National Healthcare Group, 138543, Singapore

CHEE KHEONG OOI AND SEOW YIAN TAY

Department of Emergency Medicine  
Tan Tock Seng Hospital, 308433, Singapore

(Communicated by Jie Sun)

**ABSTRACT.** The proportion of patients who reattended emergency department (ED) within 72 hours is an important indicator of quality of care. This study develops a practical framework to predict patients who will reattend ED in 72 hours from a clinical perspective. We analyze 328,733 ED patients from 1 January 2011 to 31 December 2013, with an average of 4.6% reattendances. We feature over 100 factors including demographics, diagnosis, patient acuity, chief complaints, selected laboratory tests, summarized vital signs. Using univariate analysis, a pool of risk variables is selected for subsequent factor selection. We then apply filter methods to derive a set of candidate factors. With these factors in combination with suggestions from ED clinicians, a mixed integer programming model based on discriminant analysis is proposed to determine a classification rule for 72-hour reattendance. In numerical experiments, various small subsets of risk factors are used for classification and prediction. The results show that favorable predicting performances can be achieved in both training and test sets.

**1. Introduction.** Emergency department (ED) attendances throughout the world have been steadily increasing [22, 23, 26, 35, 37]. This puts a strain on healthcare infrastructure and frontline staff. The proportion of patients who reattended ED within 72 hours is an important indicator of quality of care [13, 28, 36]. ED reattendance can be related to the nature of the disease, medical errors, and shortcomings in care during their initial treatment [13, 27, 36]. An in-depth review of the literature discussing ED discharge processes [5] suggested elements of high-quality discharges and risk factors for suboptimal discharges. Early reattendances to the ED may involve patients who are in a high-risk population [1, 12, 15]. However, other factors, such as high ED utilization and ED overcrowding resulting in decreasing efficiency,

---

2010 *Mathematics Subject Classification.* Primary: 90C11, 90C90; Secondary: 62P10.

*Key words and phrases.* Mixed integer programming, discriminant analysis, filter methods, emergency department reattendance, predict.

\* Corresponding author.

may contribute to this issue as well [24]. For hospital managers and clinicians, one of the main concerns is to reduce these early reattendances as much as possible.

In the past decades, a number of research work related to ED reattendance have been conducted from different ED settings with different perspectives. A reader is referred to [1, 3, 9, 10, 14, 16, 21, 25, 29, 33, 40] in this regard, just to name a few. Risk factors studied in the literature include diagnosis, patient chief complaints, patient demographic factors, and clinical measures. Amongst, some studies focused on specific areas, such as pediatric mental health care [25], acute pulmonary embolism [3], chronic obstructive pulmonary disease (COPD) exacerbations [9], and chest pain [29]. Most research work in this regard were retrospective observational studies using historical patient administration data while some was prospective observational study [29] or based on a population survey [16], for instance. In addition to analyze potential risk factors concerning reattendance, it is also important to predict patient early reattendance based on patient characters associated with identified factors of interest. This study seeks to investigate these two issues based on historical ED patient administration data of the local context.

Decision trees and regression analysis were most commonly used for predicting ED early reattendance in the past, see [11, 31, 34] just for example. In particular, studies using linear regression models indicated that errors in medical care or patient education might increase the risk of early reattendances. However, as pointed out in [19], the models may fail to accurately predict a reattendance visit since the association between the binary value of reattendance and risk factors is unlikely to a simple linear relationship. Recently, Lee et al [19] designed a clinical decision tool using discriminant analysis via mixed integer programming (DAMIP) to predict patient reattendance within 72 hours in two pediatric emergency departments (PEDs) of United States (U.S.). Using patient administration data of two PEDs, numerical results demonstrated that this approach remarkably outperformed these traditional methods in predicting patient early reattendance. As a machine learning model, DAMIP establishes a classification rule based on a training set which could achieve much higher blind predictive accuracy on an independent set of patients. This favorable feature was further elaborated in a recent multi-site evidence-based study [17] involved with more than 700 healthcare sites in U.S.

Classification under consideration is performed via the DAMIP approach first developed by Gallagher *et al* [8], which realizes the optimal parameters of Anderson's classification model [2]. It seeks to derive a rule for classifying entities into several groups, each being known a priori. A classification rule is basically formed with the data collected on a sample of entities with known group classifications, which can be generated by solving an established optimization problem such as mixed integer programming (MIP) and linear programming (LP) [8, 20]. According to the obtained rule, we can classify a new entity to a certain group. In the literature, such optimization-based prediction and classification methods were extensively applied to study various types of medicine or biology data [18]. The classification rule, which was generated based on only a small set of risk factors (around 10 factors), can perform very well in prediction as stated in [19]. For the problem under consideration, this feature is particularly enjoyable and important to clinicians and managers for better understanding several key factors of interest, facilitating managers to improve current treatment and care process. However, regression analysis approach usually relies on relatively larger number of risk factors in prediction while and decision tree is often criticized due to its instability and complexity and so on.

These limitations would undermine their transformed implementation to some extent from practical perspective. On the other hand, clinical decision tool proposed in [19] was with a specific focus on pediatric emergency care. In the literature, there has been little study to predict ED early reattendance for the adult population by using the DAMIP approach and to further improve general emergency care. Motivated by the above, the present study aims to fill this gap.

This study is a research collaboration with clinicians from an emergency department of a tertiary urban hospital where the data under investigation is very broad consisting of patient demographic information; patient medical information like medical history, diagnosis, vital signs, and lab tests; and patient social information, and so on. Slightly different from the framework presented in [19] or conventional classification method, in this paper we are interested to propose a practical framework to predict 72-hour reattendance in combination with clinical perspectives. Following advices from ED clinicians, one purpose of this study is to help clinicians analyze some clinical factors they might be unaware or neglect in routine emergency care. Clinicians like to understand the potential role of these factors in terms of early reattendance, rather than *commonly known* risk factors in ED like patient acuity category, gender, and age. Under current emergency care procedure, clinicians generally have put necessary attention and extra efforts when treat patients related to such *known* risk factors. As such, clinicians are more interested to explore any potential to improve emergency care by studying early reattendance including those *unfamiliar* risk factors of interest. Hence, this study seeks to introduce a practical framework to predict early reattendance with a set of patient features consisting of risk factors obtained through feature selection and clinical factors recommended by clinicians. The process of feature selection under consideration is a hybrid approach of heuristic and actual method to some extent.

Specifically, in this paper we first conducted univariate analysis on electronic medical records (EMR) of all ED attendances during 1 January 2011 to 31 December 2013 and chose a pool of potential risk factors for subsequent analysis. Second, in feature selection, we applied commonly used filter methods [6,32] to choose candidate risk factors among these potential risk factors. The factors were chosen according to their occurrence frequency and rankings associated with the underlying filter methods. Third, using a set of selected factors by incorporating advices from ED clinicians, we applied DAMIP model to determine a classification rule for predicting patient likely to reattend within 72 hours. Predictions with a training set and a test set were conducted to evaluate the performance of the proposed model. The prediction results showed that the DAMIP approach is favorable, compared with logistic regression. Of particular interest in this study, sensitivity was achieved at around 40% using a small number of risk factors (around 10 factors), which did not include some *known* key risk factors following clinicians' suggestion. This result is satisfying and remarkable because the overall 72-hour reattendance rate was very low (i.e., about 4.6%). The obtained results may inform ED clinicians and hospital managers for further improving the quality of emergency patient care. This study is an exercise of the use of operations research in emergency care management from a clinical perspective. We believe this work would enrich and complement their applications in health services research.

The rest of this paper is organized as follows. Section 2 presents model development of the study in detail, including the setting, filter methods, and mixed integer

programming for discriminant analysis. Numerical experiments are presented in Section 3, followed by discussions in Section 4. Section 5 concludes the paper.

**2. Model development.** In this section, we first detail data preparation process, followed by introducing a practical framework for feature selection, classification and reattendance prediction by employing filter methods and the DAMIP approach. At the end of this section, we present a tractable model reformulation for computational implementation.

**2.1. Setting.** The hospital in this study is a tertiary urban hospital in Singapore with about 1,500 beds. We collected 328,733 ED attendances that met the inclusion criteria of a repeat visit within 72 hours of the index visit during the period from 1 January 2011 to 31 December 2013. Patients were identified from the patient information system of ED, among which 15,171 patients (about 4.6%) reattended ED within 72 hours. The patients were classified into two groups, i.e., patients who reattended ED within 72 hours; and patients who were discharged but did not revisit ED within 72 hours.

The electronic medical records contained about 140 factors for each patient, including patient demographic information; chief complaints; ICD 9-CM (International Classification of Diseases, 9th Revision, Clinical Modification) descriptions [7]; medical history; lab tests; handover; vital signs; patient information at triage such as language, mobility, time, date, patient acuity category (PAC), and social issues.

Patient demographic information included age, sex, race, nationality, residential status. Patient chief complaints were classified to 18 categories according to advices from clinicians. In Singapore, there are four levels of patient acuity category, with PAC1 being the most serious and PAC4 the least. In this study, we considered the number of handovers between doctors (e.g., when a doctor leaves a shift) a patient had as a potential risk factor. Medical history was defined by the total number of known medical conditions a patient had. In this study, social issues and national service were included in the analysis. Social issues basically follow ICD-9-CM descriptions such as patients with generalised weakness, severe social and financial issues, no home, social care reasons due to family violence, and so on. National service, a compulsory military draft, might be a unique feature of patients in Singapore, compared with other countries or regions. All male citizens and second-generation permanent residents of Singapore are required to serve 2 years of full-time national service. In this paper, we were interested to explore whether this factor would play a potential role in early reattendance and thereby included this factor in the study data.

**2.2. A practical framework to predict ED reattendance.** We present a practical framework to investigate ED early reattendance, consisting of four steps, i.e., data preprocessing, factor/feature selection, classification and prediction as follows.

**Step 1. Pre-selection of risk factors.** Use univariate analysis to derive statistically significant factors associated with ED 72-hour reattendance.

**Step 2. Factor selection.** For the group of factors derived in Step 1, use filter methods, which will be elaborated in Section 2.2.1, to generate the corresponding sets of risk factors.

Choose a set of candidate factors consisting of most frequently occurred factor among top rankings in the sets of selected factors and some clinical factors recommended by ED clinicians.

**Step 3. Classification.** Derive a classification rule by solving DAMIP model with the set of discriminatory factors obtained in Step 2.

**Step 4. Prediction.** Using the classification rule obtained in Step 3, predict the possibility of a new patient whether he/she would reattend ED in 72 hours.

In what follows, we discuss factor selection using filter methods and the development of DAMIP model.

2.2.1. *Filter methods.* We analyzed the set of risk factors derived from univariate analysis by applying feature selection technique. Note that the set of selected factors used for classification and prediction is high dimensional data set, which is a significant challenge in data mining. As pointed out by Brown *et al* [6], in this circumstance, some of factor information might be redundant in the context of others, resulting in important issues like over-fitting or computational burden of processing many similar features. Hence, it is important and necessary to identify meaningful and smaller subset of these factors using the process of *feature selection* [6, 30, 32, 38, 39].

In this paper, we used different filter methods for feature selection discussed in Brown *et al* [6] and the underlying filter criteria are *JMI*(joint mutual information), *MIFS*(mutual information feature selection), *CMIM*(conditional mutual information maximisation), *MRMR* (max-relevance min-redundancy), *ICAP*(iteration capping), *CIFE*(conditional infomax feature extraction), *DISR*(double input symmetrical relevance), *CMI*(conditional mutual information), and *ConDred*(conditional redundancy). For each filter method, with patient EMR data consisting of potential risk factors derived in Step 1, the top 15 factors were chosen by implementing the MATLAB toolbox *FEAST* developed by Brown *et al*, and we then obtained 9 corresponding sets of selected factors. We selected a subset of factors, which are of high frequency of occurrences in these 9 sets. This set of selected factors was used as candidate discriminatory factors for classification and prediction. In this paper, we chose 10 candidate factors. We also included some factors recommended by clinicians. Their advices were made usually from the perspective of clinical treatment together with the management of care process. The underlying process of factor selection is different from traditional feature selection in machine learning. Instead, it is flexible by including factors according to the inputs from experts' point of view. Although this process is nonstandard, it appears to be useful and important from the practical perspective. In particular, in the situation where care process is effectively performed at areas related to important but known risk factors like patient acuity category and age, clinicians should draw much attention to risk factors of interest, which were not selected as top risk factors by usual filtering methods. We think the underlying factor selection process incorporating clinical perspectives is very important and with potential high impact in practice, which however is usually neglected in academia.

We adopted 9 filter methods to determine a set of candidate risk factors. In fact, people may use any filter method of interest in feature selection. Following similar arguments in Lee *et al* [19], we chose 10 factors generated by filter methods in this paper. Actually, the number factors basically depends on the choice of modelers or decision-makers from practical perspectives. One may choose 15 or

20 top factors in analysis. However, it might not be useful or even impractical to improve hospital care operations if choose too many factors like over 30 risk factors in prediction. Evidently, it will greatly increases the difficulty and infeasibility to improve emergence care process concerning too many risk factors, constrained by the limited capacity and resource in emergency department.

**2.2.2. DAMIP model.** In this section, we present a general model for discriminant analysis and its approximation using mixed integer programming. A detailed discussion in this regard can be found in [18–20]. Given a training sample of entities each having different features/factors and with known group classification, the classification model seeks to find a function to classify entities into groups based on the selected set of features. Some notations such as parameters and variables used in the model are listed below. For ease of reading and the completeness, we include necessary mathematical expositions concerning mathematical models in the Appendix.

Let  $R_g$  denote the region containing group  $g$  entities,  $g = 1, \dots, G$ , and  $R_0$  denote the reserved-judgment region. The model under consideration is based on Anderson's method [2] for finding a partition  $\{R_0, R_1, \dots, R_G\}$  that maximizes the probability of correct classification subject to constraints on the misclassification probability limits. Since it is difficult or even impossible to solve this primary problem in practice, we resort to its approximation counterpart for numerical computation under some standard assumptions in discriminant analysis. That is, with the assumptions of multivariate normal group-conditional distributions with a common covariance matrix and equal prior probabilities for all groups, we estimate conditional density functions  $f_g$  and prior probabilities  $\pi_g$ ,  $g = 1, \dots, G$ , for a given training data. As we shall see in Appendix, the classifier is defined using the discriminant function  $L_g(x)$  where the coefficients were derived by solving the underlying mixed integer programming. Further, these coefficients are actually the unique Lagrangian multipliers of the corresponding optimization problem. Then, we can derive a classifier for prediction using the DAMIP approach. In literature, there has been a number of research work concerning classification and pattern recognition, for example the excellent work concerning linear support vector machine by Bennet and Mangasarian [4], where a linear classifier was studied by solving a linear programming problem.

For the ED reattendance problem under consideration, we choose a training data of ED patients with selected discriminatory factors and two known group classifications, i.e., 72-hour return and non-return. We derive the corresponding MIP model (3) with underlying  $k$  factors for classification. It is known that the model can be solved efficiently using the well developed commercial software like CPLEX, which is popular and widely used in the community of numerical optimization and operations research. In addition, the input parameters involved in (3), such as  $\hat{p}_1$  and  $\hat{p}_2$ , the sample covariance matrix and its inverse, can be derived readily using some software such as MATLAB.

### 3. Results.

**3.1. Data preparation and model implementation.** The patient data of this study were extracted from the ED Web system of a public hospital in Singapore from 1 January 2011 to 31 December 2013. There were more than 100 factors

involved in this study. The factors included diagnosis, PAC, chief complaints, demographic, vital signs (such as average pulse, average temperature, diastolic blood pressure range), laboratory parameters (for example albumin, basophils, creatinine, red blood cell count, white blood cell count, sodium, potassium, urea within normal, borderline, or abnormal range), medical history (hypertension, hyperlipidemia, diabetes, asthma, heart problem, stroke, gastric problem, etc), handover, together with derived factors like national service and social issue. We excluded the factor of radiology since there was no available results in the data. Respiratory rate was excluded in the study as the result was based on estimations by the staff.

During data preparation, we merged all information from various sources such as tables of vital signs, laboratory results, handover forms, and diagnosis, into a single table using Excel spreadsheet according to a deidentified patient identifier in the ED Web system.

Numerical experiment was performed by implementing codes in SPSS 17.0, MATLAB 7.8.0 and CPLEX 12.4. Following the proposed integrated framework, we first conducted univariate analysis using SPSS and derived a set of potential risk factors. After that, we applied filter methods by running the *FEAST* toolbox in MATLAB to select a group of risk factors for classification. Finally, we implemented the CPLEX build-in solver for solving the DAMIP model (3) to determine the classifier as desired.

**3.2. Univariate analysis.** Using univariate analysis, we derived a total of 47 potential risk factors ( $p < 0.001$ ) as shown in Table 1, each being assigned an index for ease of reference.

Index	Factor Name	Index	Factor Name
1	Age	24	Diagnosis: COPD <sup>3</sup>
2	Gender	25	Diagnosis: fracture upper limb (FUL)
3	Race	26	Diagnosis: fracture lower limb (FLL)
4	Registration day of week (DOW)	27	Diagnosis: open wound upper limb (OWUL)
5	Registration public holiday and Sunday (HAS)	28	Diagnosis: superficial injuries (SIN)
6	CC <sup>1</sup> orthopaedic (ORTHO)	29	Diagnosis: intestinal diseases (ID)
7	CC <sup>1</sup> gastrointestinal (GI)	30	Diagnosis: rheumatism (RHE)
8	CC <sup>1</sup> respiratory (RES)	31	Diagnosis: diseases of ear and mastoid (DEM)
9	CC <sup>1</sup> general & minor	32	Diagnosis: open wound of head, neck, and trunk (OW of HNT)
10	CC <sup>1</sup> external causes of morbidity	33	Diagnosis: neurotic disorders (ND)
11	CC <sup>1</sup> neurologic	34	Diagnosis: viral diseases (VD)
12	CC <sup>1</sup> cardiovascular (CV)	35	Diagnosis: complications of care (COMC)
13	CC <sup>1</sup> skin	36	Diagnosis: open wound lower limb (OWLL)
14	CC <sup>1</sup> eye complaints (EC)	37	Diagnosis: others in ICD-9-CM <sup>4</sup> groups
15	CC <sup>1</sup> ear, neck, throat	38	PAC <sup>5</sup> status
16	CC <sup>1</sup> substance misuse (SM)	39	Social issues (SI)
17	CC <sup>1</sup> others	40	National service full-time (NSTF)
18	EDTC <sup>2</sup> category	41	Handover
19	Diagnosis: general symptoms (GM)	42	Medical history diabetes
20	Diagnosis: acute respiratory infection (ARI)	43	Medical history asthma
21	Diagnosis: sprains and strains (SS)	44	Medical history cardiovascular disease
22	Diagnosis: minor head injury (MHI)	45	Average temperature (ave temp.)
23	Diagnosis: diseases of oesophagus, stomach, duodenum	46	Average pulse
		47	Diastolic blood pressure (DBP)

<sup>1</sup> CC stands for chief complaint  
<sup>2</sup> EDTC stands for emergency diagnostic and therapeutic centre. Certain emergency patients are observed at EDTC for up to 24 hours before being admitted or discharged  
<sup>3</sup> COPD stands for chronic obstructive pulmonary disease  
<sup>4</sup> ICD-9-CM stands for international classification of diseases, 9th revision, clinical modification  
<sup>5</sup> PAC stands for patient acuity category

TABLE 1. List of Potential Risk Factors

**3.3. Factor selection.** We used filter methods to select a subset of risk factors listed in Table 1. The filter methods were *JMI*, *MIFS*, *CMIM*, *MRMR*, *ICAP*, *CIFE*, *DISR*, *CMI*, *ConDred*, which are commonly used in feature selection. We implemented each filter method using the *FEAST* toolbox in MATLAB on the data consisting of 47 factors for each patient as shown in Table 1. Then, factors with 15 top rankings generated with each filter method were selected for further analysis. These factors were summarized in Table 2. We calculated their frequency of occurrences associated with the underlying 9 filter methods. Namely, the frequency of each factor is a percentage of its total number of occurrences divided by 9. For example, PAC status was one of top 15 factors generated using all filter methods. Then, its occurrence frequency would be equal to 100%.

In this study, we chose the top 10 factors by considering their high occurrence frequencies and top rankings. Further, we also included clinical measures like average temperature, average pulse and diastolic BP range for predicting patient early reattendance, as recommended by ED clinicians. Then, we derived the total of 13 factors denoted by  $S_0$  as follows, where the first ten factors were arranged according to their rankings from the highest to the lowest.

$$S_0 = \{ \text{Social issues; Diagnosis: COPD; CC substance misuse; Diagnosis: neurotic disorders; Diagnosis: open wound lower limb; Handover; CC respiratory; Diagnosis: open wound upper limb; Diagnosis: fracture upper limb; Diagnosis: open wound of HNT; Average temperature; Average pulse; Diastolic BP} \}.$$

As discussed previously and following the suggestions from ED, some familiar risk factors to clinicians like PAC status, age and gender were not included although they were of high occurrence frequency and top ranking. Note that these factors were not directly associated with treatment and care process, unlike clinical factors shown in  $S_0$ .

**3.4. Classification and prediction.** Discriminant analysis was then focused on patient data with risk factors in  $S_0$  together with known patient classification groups (i.e., return and non-return). As we shall see later, we also used the following subsets of  $S_0$  to derive different classification rules for prediction.

$$\begin{aligned} S_1 &= S_0 \setminus \{ \text{Social issues} \}; & S_2 &= S_1 \setminus \{ \text{Diagnosis: open wound of HNT} \}; \\ S_3 &= S_2 \setminus \{ \text{Diagnosis: fracture upper limb} \}; \\ S_4 &= S_3 \setminus \{ \text{Diagnosis: open wound upper limb} \}. \end{aligned}$$

Clearly,  $S_4 \subset S_3 \subset S_2 \subset S_1 \subset S_0$ . Also,  $|S_0| = 13$ ,  $|S_1| = 12$ ,  $|S_2| = 11$ ,  $|S_3| = 10$ ,  $|S_4| = 9$ , where  $|S|$  denotes the Cardinality of a set  $S$ , i.e., number of elements of the set  $S$ . In our analysis, the training set consists of 109,674 patients and the remaining patient data were used for blind prediction. We included both 10-fold cross-validation and the blind prediction results to reflect the consistency of predictive power of the underlying classification rules. The values of input parameters in the DAMIP model (3) were chosen as  $\alpha = \alpha_{12} = \alpha_{21} = 0.15$ ,  $M = 100$ ,  $\varepsilon = 0.01$ ,  $\beta = 2$ , and  $\gamma = 3$ . Numerical study was conducted by implementing codes in CPLEX 12.4 to solve the MIP problem (3).

Further data analysis indicated that the 72-hour reattendance rate for patients with social issues was as high as 65.6%, compared with the low reattendance rate of 4.6% on average. However, the total number of patients in this group was very small, accounting for only 0.33% of total patients of the study. This implied that

JMI	MIFS	CMIM	MRMR	ICAP	CIFE	DISR	CMI	ConDred
SI	SI	SI	SI	SI	SI	SI	SI	SI
PAC	PAC	PAC	PAC	PAC	PAC	diag: COPD	PAC	PAC
DBP	diag: ID	DBP	diag: ID	DBP	ave pulse	CC stance	ave SM	DBP
ave pulse	reg HAS	ave pulse	reg HAS	ave pulse	DBP	diag: ND	DBP	EDTC
diag: COPD	diag: COMC	diag: COPD	diag: OWLL	diag: COPD	age	diag: OWLL	reg DOW	CC others
age	diag: OWLL	age	diag: OW of HNT	age	reg DOW	handover	age	diag: MHI
CC RES	diag: OW of HNT	handover	diag: COMC	gender	race	CC RES	race	age
gender	diag: VD	gender	diag: ND	handover	diag: COPD	diag: OWUL	ave temp.	NSFT
ave temp.	diag: DEM	ave temp.	diag: DEM	ave temp.	ave temp.	diag: FLL	gender	ave pulse
handover	CC SM	CC RES	ave temp.	CC GI	gender	diag: OW of HNT	CC ORTHO	CC ORTHO
race	diag: RHE	CC EC	diag: FLL	race	CC RES	PAC	diag: GM	diag: GM
EDTC	diag: SIN	CC ORTHO	diag: RHE	NSFT	handover	EDTC	diag: others in ICD9	CC CV
NSFT	diag: FLL	race	diag: OWUL	reg DOW	NSFT	diag: DEM	CC RES	diag: ARI
CC ORTHO	diag: FUL	NSFT	diag: SIN	diag: ND	diag: GM	diag: ID	CC GI	CC RES
reg DOW	diag: OWUL	diag: ND	gender	diag: GM	medical history asthma	CC EC	NSFT	diag: SS

TABLE 2. Top 15 Risk Factors Selected Using Filter Methods

Set of risk factors	Training set			Test set		
	sensitivity	specificity	overall	sensitivity	specificity	overall
$S_0$	29.4%	86.2%	83.5%	29.3%	85.8%	83.3%

TABLE 3. Prediction Results Including Factor of Social Issues

the factor of social issues might not play a key role in ED 72-hour reattendance overall. We then conducted further analysis without the factor of social issues in prediction using four sets of risk factors  $S_1$ ,  $S_2$ ,  $S_3$ , and  $S_4$  as defined above. The purpose was to achieve a satisfying prediction result using risk factors as few as possible. In other words, we would like to find the possible smallest set of risk factors, with which one could attain certain good prediction results. This is different from classical prediction methods like decision tree and regression. Prediction results were reported in Tables 3 and 4. We also compared with prediction results conducted using logistics regression as shown in Table 5.

Set of risk factors	Training set			Test set		
	sensitivity	specificity	overall	sensitivity	specificity	overall
$S_1$	41.1%	74.2%	72.5%	40.3%	74.1%	72.4%
$S_2$	40.6%	74.5%	72.8%	40.1%	74.1%	72.6%
$S_3$	40.6%	74.4%	72.7%	40.1%	74.2%	72.7%
$S_4$	40.3%	74.4%	72.8%	39.9%	74.4%	72.8%

TABLE 4. Prediction Results Excluding Factor of Social Issues

	Training set: 211,160			Test set: 117,573		
Hosmer-Lemeshow test	$p < 0.001$			$p < 0.001$		
C-statistic of ROC curve	0.67			0.66		
Predictive accuracy	sensitivity	specificity	overall	sensitivity	specificity	overall
70% cutoff	1.7%	99.9%	95.4%	1.0%	99.9%	95.5%
50% cutoff	4.6%	99.9%	95.5%	2.4%	99.9%	95.5%
20% cutoff	11.9%	99.3%	95.2%	7.2%	99.4%	95.2%
10% cutoff	18.5%	96.3%	92.7%	14.1%	96.1%	92.4%
5% cutoff	51.6%	72.2%	71.2%	51.0%	70.8%	69.9%

TABLE 5. Prediction Results Using Logistic Regression

**4. Discussion.** This paper introduces a practical framework which consists of univariate analysis, filter methods and discriminant analysis in combination with clinicians' advices to predict ED patient 72-hour reattendance. Briefly, filter methods were applied to select risk factors of interest and discriminant analysis was then employed to determine the classification rule. Different from traditional methods such as decision tree and regression, discriminant analysis under consideration was based on Bayesian discriminant rule with some prior distributional assumptions, where clinical perspectives were incorporated. We used a mixed integer programming model to approximate the associated discriminant analysis problem, for a given sample of ED patients.

In this study, a relatively small number ranging from 9 to 13 of discriminatory factors were used to predict ED early reattendance. As discussed previously, these factors were mostly related to patient diagnosis and treatment except the factor of social issues. Clinicians would draw particular attention on these factors once needed in patient treatment and care. Table 4 indicated that the proposed model predicted 72-hour reattendance with a satisfying accuracy of more than 72.6% overall. Of particular interest in this study, sensitivity, which describes the prediction accuracy for patients who would reattend ED within 72 hours, was about 40% in the four scenarios under consideration. In contrast to the overall reattendance rate as low as 4.6% and considering risk factors used for prediction without including some known key risk factors, the obtained results were favorable. Also, the respective sensitivity and specificity results between training and test sets were very close in all scenarios under consideration. This indicated that the proposed model appears to be stable and consistent. We also performed prediction analysis using logistic regression as demonstrated in Table 5. Logistic regression can achieve very high overall prediction accuracy of more than 90%. In particular, the predictive accuracy of specificity was as high as over 96% in most cases under consideration. However, this approach performed poor with the predictive accuracy of sensitivity as low as 1.7% in training set and 1.0% in test set. And, only 4.6% predictive accuracy was achieved in training set when choosing 50% cutoff probability. In addition, over 50% sensitivity prediction accuracy could be attained at the cost of choosing 5% cutoff.

On the other hand, sensitivity and overall prediction results in Lee *et al* [19] were higher than those derived in this study, where prediction results were 70.5% and 82.2% for two PEDs under consideration. There might be a couple of reasons leading to these different prediction performances. One main reason could be that our analysis included clinical factors clinicians suggested, but also excluded those well known factors like PAC status, age and gender from practical perspectives, which inevitably affected prediction performance of the DAMIP approach. Another reason might be the nature of different ED settings. In this paper, we studied an adult emergency department with approximately 500 patient visits per day while Lee's

study focused on pediatric emergency department with a relatively low daily visits (about 280 attendances per day). Patient conditions and care in adult emergency departments appear to be much more complicated than in pediatric emergency departments, which could make more challenging to predict early reattendance and thereby affect prediction accuracies as well.

**4.1. Limitations.** This work was a retrospective study. The data collection may be incomplete. However, as the data was collected from multiple sources, this issue might be minimized. We anticipate that we may not have any variables that would completely fulfil the criteria to be a potential predictor or to derive a clinical decision rule. In this situation, we will highlight the variables as being associated with the outcome and seek to validate them in future studies.

**5. Conclusion.** In this paper, we proposed an integrated framework to study 72-hour reattendance in emergency department from clinical perspectives, including pre-selection of factors using univariate analysis, factor selection using filter methods, and classification and prediction using DAMIP model. A relatively small number of discriminatory factors were used in prediction. The obtained prediction results were satisfying in both training and test sets. In the future, we are interested to explore the implementation of the model in practice. Furthermore, we would like to apply DAMIP approach to investigate early reattendance focusing on specific group of ED patients of interest, for example, patients with the diagnosis of chronic obstructive pulmonary disease since in general it was a large contribution to ED reattendances [9].

**Acknowledgments.** The authors would like to thank the two anonymous reviewers for their constructive suggestions and comments, which help to improve the presentation of the paper.

## Appendix.

**Discriminant analysis.** First, we list some notations such as parameters and variables used in the model development.

- $N$ : number of entities of the training sample
- $G$ : number of groups
- $k$ : number of feature attributes of each group
- $n_g$ : total number of entities in group  $g$  of the training sample,  $g = 1, \dots, G$
- $f_g$ : group conditional density function,  $g = 1, \dots, G$
- $\pi_g$ : prior probability,  $g = 1, \dots, G$
- $x^{gj} \in \mathfrak{R}^k$ :  $k$  attributes of entities in group  $g$ ,  $g = 1, \dots, G$ ,  $j = 1, \dots, n_g$
- $\alpha_{hg} \in (0, 1)$ : misclassification limit concerning  $n_g$  entities of group  $g$ , that is, proportion of group  $g$  training entities classified to region  $R_h$  should be no more than  $\alpha_{hg}$ ,  $h, g = 1, \dots, G$ ,  $h \neq g$
- $R_g$ : region containing group  $g$  entities,  $g = 1, \dots, G$
- $R_0$ : reserved-judgment region

The underlying model is based on Anderson's method [2] for finding a partition  $\{R_0, R_1, \dots, R_G\}$  that maximizes the probability of correct classification subject to constraints on the misclassification probability limits as follows.

$$\max \sum_{g=1}^G \pi_g \int_{R_g} f_g(x) dx$$

$$\text{s. t.} \quad \int_{R_g} f_h(x)dx \leq \alpha_{hg}, \quad h, g = 1, \dots, G; \quad h \neq g. \tag{1}$$

It is known [2] that under some conditions there exists  $\lambda_{ig} \geq 0, i, g \in \{1, \dots, G\}, i \neq g$ , such that

$$\begin{aligned} R_0 &= \{x \in \mathfrak{R}^k : L_g(x) \leq 0, \forall g = 1, \dots, G\}, \\ R_g &= \{x \in \mathfrak{R}^k : L_g(x) \geq 0, \quad L_g(x) \geq L_h(x), \forall h = 1, \dots, G\}, \quad g = 1, \dots, G, \end{aligned}$$

where  $L_g(x) = \pi_g f_g(x) - \sum_{i=1, i \neq g}^G \lambda_{ig} f_i(x), g = 1, \dots, G$ . Let  $L_0(x) \equiv 0$ . As we shall see later, the nonnegative coefficients  $\lambda_{ig}$  involved can be derived by solving the subsequent DAMIP model. For simplicity, let

$$\begin{aligned} p_g(x) &= f_g(x) / \sum_{i=1}^G f_i(x), \quad g = 1, \dots, G, \\ L_g(x) &= \pi_g p_g(x) - \sum_{i=1, i \neq g}^G \lambda_{ig} p_i(x), \quad g = 1, \dots, G. \end{aligned}$$

For a given set of solutions  $\lambda_{ig}$  and a new entity with feature attribute vector  $x \in \mathfrak{R}^k$ , let  $h := \arg \max_g \{L_g(x) : g = 0, 1, \dots, G\}$ . Then, we call this entity is classified to group  $h$ . In practice, it is difficult or even impossible to solve problem (1) directly. In what follows, we discuss its approximation counterpart for numerical implementation.

**The DAMIP approximation.** Under standard assumptions in discriminant analysis, namely, multivariate normal group-conditional distributions with a common covariance matrix and equal prior probabilities for all groups, we estimated conditional density functions  $f_g$  and prior probabilities  $\pi_g, g = 1, \dots, G$ , for a given training data. Specifically, we used the normal group conditional density  $\hat{f}_g$  as an estimate,

$$\hat{f}_g(x) = \frac{1}{(2\pi)^{k/2} |S|^{1/2}} \exp \left\{ -\frac{1}{2} (x - \bar{x}_g)^T S^{-1} (x - \bar{x}_g) \right\},$$

where  $\bar{x}_g \in \mathfrak{R}^k$  denotes the sample mean vector for group  $g, S \in \mathfrak{R}^{k \times k}$  is the pooled sample covariance matrix as follows

$$S = \frac{\sum_{g=1}^G (n_g - 1) S_g}{\sum_{g=1}^G (n_g - 1)} = \frac{\sum_{g=1}^G (n_g - 1) S_g}{N - G},$$

where  $S_g \in \mathfrak{R}^{k \times k}$  is the sample covariance matrix for group  $g$  and

$$S_g = \frac{1}{n_g - 1} \sum_{i=1}^{n_g} (x^{g,i} - \bar{x}^g)(x^{g,i} - \bar{x}^g)^T, \quad \bar{x}^g = \frac{1}{n_g} \sum_{i=1}^{n_g} x^{g,i},$$

where  $\{x^{g,1}, \dots, x^{g,n_g}\}$  is the  $k$ -attribute sample of entities in group  $g$ . Accordingly, we derived the following approximation,  $\hat{p}_g(x)$ , of  $p_g(x)$

$$\hat{p}_g(x) = \frac{\hat{f}_g(x)}{\sum_{i=1}^G \hat{f}_i(x)}, \quad g = 1, \dots, G.$$

In addition, we used  $\hat{\pi}_g = 1/G$  to estimate  $\pi_g$  for  $g = 1, \dots, G$ . We also introduced a 0-1 binary decision variable,  $u_{hgj}$ , to indicate whether or not the  $j$ th entity from

group  $g$  is allocated to group  $h$ . Then, model (1) can be solved by the following approximation (Lee, 2007).

$$\begin{aligned}
 \max \quad & \sum_{g=1}^G \sum_{j=1}^{n_g} \beta u_{ggj} - \sum_{g=1}^G \sum_{j=1}^{n_g} \gamma y_{gj} \\
 \text{s. t.} \quad & L_{hgj} = \hat{\pi}_h \hat{p}_h(x^{gj}) - \sum_{i=1, \dots, G, i \neq h} \lambda_{ih} \hat{p}_i(x^{gj}), \quad h, g = 1, \dots, G, \quad j = 1, \dots, n_g, \\
 & y_{gj} \geq L_{hgj}, \quad h, g = 1, \dots, G, \quad j = 1, \dots, n_g, \\
 & y_{gj} - L_{ggj} \leq M(1 - u_{ggj}), \quad g = 1, \dots, G, \quad j = 1, \dots, n_g, \\
 & y_{gj} - L_{hgj} \geq \varepsilon(1 - u_{hgj}), \quad h, g = 1, \dots, G, \quad h \neq g; \quad j = 1, \dots, n_g, \\
 & \sum_{j=1}^{n_g} u_{hgj} \leq \lfloor \alpha_{hg} n_g \rfloor, \quad h, g = 1, \dots, G, \quad h \neq g, \\
 & -\infty < L_{hgj} < \infty, \quad y_{gj} \geq 0, \quad h, g = 1, \dots, G, \quad j = 1, \dots, n_g, \\
 & \lambda_{hg} \geq 0, \quad h, g = 1, \dots, G, \quad h \neq g, \\
 & u_{hgj} \in \{0, 1\}, \quad h, g = 1, \dots, G, \quad j = 1, \dots, n_g,
 \end{aligned} \tag{2}$$

where  $\beta$  and  $\gamma$  are positive constants,  $M$  is a large positive number, and  $\varepsilon$  is a small positive number. It is evident that (2) is a mixed linear programming, consisting of the total of  $3GN + G(G - 1)$  constraints,  $2GN + N + G(G - 1)$  variables, and  $GN$  0-1 binary variables. In next section, we reformulate model (2) with specific to the underlying ED reattendance prediction.

**Model reformulation.** We chose a training data of ED patients with selected discriminatory factors and two known group classifications, i.e., 72-hour return and non-return. Then,  $G = 2$ . Let  $g = 1$  represent the group of ED 72-hour reattendance,  $g = 2$  otherwise. Let  $n_1 = m$ ,  $n_2 = N - m$ . So, among the total of  $N$  patients,  $m$  patients were in group 1 while  $N - m$  patients were in group 2. We assume there are  $k$  factors selected for classification. Without loss of generality, we arranged the training data matrix  $X$  with two classified groups in the following way.

$$X = \begin{pmatrix} X^1 \\ X^2 \end{pmatrix} \in \mathbb{R}^{N \times k}, \quad X^1 = (x^1, \dots, x^m)', \quad X^2 = (x^{m+1}, \dots, x^N)',$$

where  $x^i \in \mathbb{R}^k$  is a column vector representing the values of  $k$  factors of the  $i$ th patient in the training data,  $i = 1, \dots, N$ . For brevity in description, we introduce the following notations.

$$\begin{aligned}
 u &= (u_{111}, \dots, u_{11m}, u_{121}, \dots, u_{12(N-m)}), \\
 v &= (v_{211}, \dots, v_{21m}, v_{221}, \dots, v_{22(N-m)}), \\
 y &= (y_{11}, \dots, y_{1m}, y_{21}, \dots, y_{2(N-m)}), \\
 \mu &= (L_{111}, \dots, L_{11m}, L_{121}, \dots, L_{12(N-m)}), \\
 \eta &= (L_{211}, \dots, L_{21m}, L_{221}, \dots, L_{22(N-m)}), \quad \lambda = (\lambda_{12}, \lambda_{21}).
 \end{aligned}$$

It is evident that  $u, v, y, \mu, \eta \in \mathbb{R}^N$ . The model (2) can be rewritten as

$$\begin{aligned}
 \max \quad & \beta(u_1 + \dots + u_m) + \beta(v_{m+1} + \dots + v_N) - \gamma(y_1 + \dots + y_N) \\
 \text{s. t.} \quad & \mu_j = \hat{\pi}_1 \hat{p}_1(x^j) - \lambda_2 \hat{p}_2(x^j), \quad j = 1, \dots, N,
 \end{aligned}$$

$$\begin{aligned}
\eta_j &= \hat{\pi}_2 \hat{p}_2(x^j) - \lambda_1 \hat{p}_1(x^j), \quad j = 1, \dots, N, \\
y_j &\geq \mu_j, \quad j = 1, \dots, N, \\
y_j &\geq \eta_j, \quad j = 1, \dots, N, \\
y_j - \mu_j &\leq M(1 - u_j), \quad j = 1, \dots, m, \\
y_j - \eta_j &\leq M(1 - v_j), \quad j = m + 1, \dots, N, \\
y_j - \mu_j &\geq \varepsilon(1 - u_j), \quad j = m + 1, \dots, N, \\
y_j - \eta_j &\geq \varepsilon(1 - v_j), \quad j = 1, \dots, m, \\
\sum_{j=m+1}^N u_j &\leq \lfloor \alpha_{12}(N - m) \rfloor, \\
\sum_{j=1}^m v_j &\leq \lfloor \alpha_{21}m \rfloor, \\
y_j &\geq 0, \quad j = 1, \dots, N, \\
\mu_j, \eta_j &\in \mathfrak{R}, \quad u_j, v_j \in \{0, 1\}, \quad j = 1, \dots, N, \\
\lambda_1 &\geq 0, \quad \lambda_2 \geq 0.
\end{aligned} \tag{3}$$

Problem (3) is a mixed integer programming which can be solved efficiently by the well developed software like CPLEX. The inputs needed in model (3) like  $\hat{p}_1$ ,  $\hat{p}_2$ , the sample covariance matrix and its inverse, can be derived readily using some software such as MATLAB.

#### REFERENCES

- [1] E. A. Alessandrini, J. M. Lavelle, S. M. Grenfell, C. R. Jacobstein and K. N. Shaw, Return visits to a pediatric emergency department, *Pediatr. Emerg. Care*, **20** (2004), 166–171.
- [2] J. A. Anderson, Constrained discrimination between  $k$  populations, *J. R. Stat. Soc. Series B*, **31** (1969), 123–139.
- [3] D. Aujesky, M. K. Mor, M. Geng, R. A. Stone, M. J. Fine and S. A. Ibrahim, Predictors of early hospital readmission after acute pulmonary embolism, *Arch. Intern. Med.*, **169** (2009), 287–293.
- [4] K. P. Bennett and O. L. Mangasarian, Robust linear programming discrimination of two linearly inseparable sets, *Optim. Method Softw.*, **1** (1992), 23–34.
- [5] R. T. Boonyasai, O. M. Ijagbemi and J. C. Pham et al, *Improving the Emergency Department Discharge Process*, Agency for Healthcare Research and Quality, Rockville, Maryland, United States. 2014.
- [6] G. Brown, A. Pocock, M.-J. Zhao and M. Luján, Conditional likelihood maximisation: A unifying framework for information theoretic feature selection *J. Mach. Learn. Res.*, **13** (2012), 27–66.
- [7] CDC, ICD-9-CM official guidelines for coding and reporting, [http://www.cdc.gov/nchs/data/icd/icd9cm\\_{-}guidelines\\_{-}2011.pdf](http://www.cdc.gov/nchs/data/icd/icd9cm_{-}guidelines_{-}2011.pdf) (accessed on 20 March 2016).
- [8] R. J. Gallagher, E. K. Lee and D. A. Patterson, Constrained discriminant analysis via 0/1 mixed integer programming, *Ann. Oper. Res.*, **74** (1997), 65–88.
- [9] E. Goksu, C. Oktay, M. Kartal, A. Oskay and A. V. Sayrac, Factors affecting revisit of COPD exacerbated patients presenting to emergency department, *Eur. J. Emerg. Med.*, **17** (2010), 283–285.
- [10] C. Y. Han, L. C. Chen, A. Barnard, C. C. Lin, Y. C. Hsiao and H. E. Liu et al, Early revisit to the emergency department: An integrative review, *J. Emerg. Nurs.*, **41** (2015), 285–295.
- [11] S. Hao, B. Jin and A. Y. Shin et al, Risk prediction of emergency department revisit 30 days post discharge: A prospective study, *PLoS One*, **10** (2015), e0117633.
- [12] S. C. Hu, Analysis of patient revisits to the emergency department, *Am. J. Emerg. Med.*, **10** (1992), 366–370.
- [13] I. Insuwan, Characteristics of unscheduled emergency department return visit patients within 48 hours in Thammasat university hospital, *J. Med. Assoc. Thai.*, **94** (2011), S73–S80.

- [14] Y. Jenab, S. Haghani, A. Jalali and F. Darabi, [Unscheduled return visits and leaving the chest pain unit against medical advice](#), *Iran Red Crescent Med. J.*, **17** (2015), e18320.
- [15] K. D. Keith, J. J. Bocka, M. S. Kobernick, R. L. Krome and M. A. Ross, [Emergency department revisits](#), *Ann. Emerg. Med.*, **18** (1989), 964–968.
- [16] M. Ko, Y. Lee, C. Chen, P. Chou and D. Chu, [Incidence of and predictors for early return visits to the emergency department: a population-based survey](#), *Medicine*, **94** (2015), e1770.
- [17] E. K. Lee, Y. Wang, M. S. Hagen, X. Wei, R. A. Davis and B. M. Egan, [Machine learning: Multi-site evidence-based best practice discovery](#), in *Machine Learning, Optimization, and Big Data. MOD 2016. Lecture Notes in Computer Science* (eds. P. Pardalos, P. Conca, G. Giuffrida and G. Nicosia), Springer, Cham, (2016), 1–15.
- [18] E. K. Lee, [Large-scale optimization-based classification models in medicine and biology](#), *Ann. Biomed. Eng.*, **35** (2007), 1095–1109.
- [19] E. K. Lee, F. Yuan, D. Hirsh, M. Mallory and H. K. Simon, [A clinical decision tool for predicting patient care characteristics: patients returning within 72 hours in the emergency department](#), *AMIA Annu. Symp. Proc.*, (2012), 495–504.
- [20] E. K. Lee, R. J. Gallagher and D. A. Patterson, [A linear programming approach to discriminant analysis with a reserved-judgment region](#), *INFORMS J. Comput.*, **15** (2003), 23–41.
- [21] S. J. Liaw, M. J. Bullard, P. M. Hu, J. C. Chen and H. C. Liao, [Rates and causes of emergency department revisits within 72 hours](#), *J. Formos Med. Assoc.*, **98** (1999), 422–425.
- [22] J. A. Lowthian, A. J. Curtis and P. A. Cameron et al, [Systematic review of trends in emergency department attendances: an Australian perspective](#), *Emerg. Med. J.*, **28** (2011), 373–377.
- [23] F. Meng, C. K. Ooi, C. K. K. Soh, K. L. Teow and P. Kannapiran, [Quantifying patient flow and utilization with patient flow pathway and diagnosis of an emergency in Singapore](#), *Health Syst.*, **5** (2016), 140–148.
- [24] F. Meng, J. Qi, M. Zhang, J. Ang, S. Chu and M. Siim, [A robust optimization model for managing elective admission in a public hospital](#), *Oper. Res.*, **63** (2015), 1452–1467.
- [25] A. S. Newton, S. Ali, D. W. Johnson, C. Haines, R. J. Rosychuk, R. A. Keaschuk, P. Jacobs, M. Cappelli and T. P. Klassen, [Who comes back? Characteristics and predictors of return to emergency department services for pediatric mental health care](#), *Acad. Emerg. Med.*, **17** (2010), 177–186.
- [26] Y. Y. Ng, [Optimal use of emergency services](#), *SFP*, **40** (2014), supplement 8–13.
- [27] S. Nunéz, A. Hexdall and A. Aguirre-Jaime, [Unscheduled returns to the emergency department: An outcome of medical errors?](#) *Qual. Saf. Health Care.*, **15** (2006), 102–108.
- [28] K. O’Loughlin, K. A. Hacking, N. Simmons, W. Christian, R. Syahanee, A. Shamekh and N. J. Prince, [Paediatric unplanned reattendance rate: A&E clinical quality indicators](#), *Arch. Dis. Child.*, **98** (2013), 211–213.
- [29] L. Pereira, C. Choquet and A. Perozziello et al, [Unscheduled return visits after an emergency department \(ED\) attendance and clinical link between both visits in patients aged 75 years and over: a prospective observational study](#), *PLoS One*, **10** (2015), e0123803.
- [30] H. Qiu, X. Chen, W. Liu, G. Zhou, Y. Wang and J. Lai, [A fast11-solver and its applications to robust face recognition](#), *J. Ind. Manag. Optim.*, **8** (2012), 163–178.
- [31] J. Ryan, J. Hendler and K. P. Bennett, [Understanding emergency department 72-hour revisits among medicaid patients using electronic healthcare records](#), *Big Data*, **3** (2015), 238–248.
- [32] Y. Saeys, I. Inza and P. Larrañaga, [A review of feature selection techniques in bioinformatics](#), *Bioinformatics*, **23** (2007), 2507–2517.
- [33] G. Sauvin, Y. Freund, K. Saidi, B. Riou and P. Hausfater, [Unscheduled return visits to the emergency department: Consequences for triage](#), *Acad. Emerg. Med.*, **20** (2013), 33–39.
- [34] Y. Sun, B. H. Heng, S. Y. Tay and K. B. Tan, [Unplanned 3-day re-attendance rate at Emergency Department and hospital’s bed occupancy rate](#), *Int. J. Emerg. Med.*, **8** (2015).
- [35] C. R. Trivedy and M. W. Cooke, [Unscheduled-return-visits \(URV\) in adults to the emergency department \(ED\): A rapid evidence assessment policy review](#), *Emerg. Med. J.*, **32** (2015), 324–329.
- [36] S. J. J. Verelst, S. Pierloot, D. Desruelles, J.-B. Gillet and J. Bergs, [Short-term unscheduled return visits of adult patients to the emergency department](#), *J. Emerg. Med.*, **47** (2014), 131–199.
- [37] A. K. C. Wai, C. M. Chor, A. T. C. Lee, Y. Sittambunka, C. A. Graham and T. H. Rainer, [Analysis of trends in emergency department attendances, hospital admissions and medical staffing in a Hong Kong university hospital: 5-year study](#), *Int. J. Emerg. Med.*, **2** (2009), 141–148.

- [38] Y. Wang, G. Zhou, L. Caccetta and W. Liu, [An alternative Lagrange-dual based algorithm for sparse signal reconstruction](#), *IEEE Trans. Sig. Process.*, **59** (2011), 1895–1901.
- [39] C. Wei and X. Tang, [A fuzzy group decision making approach based on entropy, similarity measure](#), *Int. J. Inf. Technol. Decis. Mak.*, **10** (2011), 1111–1130.
- [40] C. L. Wu, F. T. Wang, Y. C. Chiang, Y. F. Chiu, T. G. Lin and L. F. Fu et al, [Unplanned emergency department revisits within 72 hours to a secondary teaching referral hospital in Taiwan](#), *J. Emerg. Med.*, **38** (2010), 512–517.

Received October 2017; revised January 2018.

*E-mail address:* [fanwen-meng@nhg.com.sg](mailto:fanwen-meng@nhg.com.sg)

*E-mail address:* [kiok-liang-teow@nhg.com.sg](mailto:kiok-liang-teow@nhg.com.sg)

*E-mail address:* [kelvinwsteo@gmail.com](mailto:kelvinwsteo@gmail.com)

*E-mail address:* [chee-kheong-ooi@ttsh.com.sg](mailto:chee-kheong-ooi@ttsh.com.sg)

*E-mail address:* [seow-yian-tay@ttsh.com.sg](mailto:seow-yian-tay@ttsh.com.sg)