Received: Nov 17, 2023 Revised: Jan 23, 2024 Accepted: Feb 28, 2024

Comparative Study of Factors Influencing the Adoption or Non-Adoption of Healthcare Services at Newly Opening Private Hospitals

การศึกษาเปรียบเทียบปัจจัยจำแนกการเลือกใช้หรือไม่ต่อบริการด้านสุขภาพของโรงพยาบาลเอกชนที่จะเปิดใหม่

Attakrai Punpukdee[1*], Chitinout Wattana[2], Wachira Punpairoj[3],
Udomlak Srichuachom[4], Pimsara Yaklai[5], Suparawadee Trongtortam[6]
อรรถไกร พันธุ์ภักดี[1*], จิตตินุช วัฒนะ[2], วชิรา พันธุ์ไพโรจน์[3],
อุดมลักษณ์ ศรีชัวชม[4], พิมสหรา ยาคล้าย[5], สุภรวดี ตรงต่อธรรม[6]
Faculty of Business, Economics, and Communication Naresuan University[1-6]
คณะบริหารธุรกิจเศรษฐศาสตร์ และการสื่อสารมหาวิทยาลัยนเรศวร[1-6]
usathailand@gmail.com[1*]
ผู้ประพันธ์บรรณกิจ (Corresponding Author) [1*]

Abstract

A comparative study using logistic regression and discriminant analysis examines the importance of both methods in health services. The study compares both methods' data categorization using 2 favorite Statistics to determine their relevance in healthcare. The study focuses on Chiang Rai province in northern Thailand and the factors that influence the selection of newly established private hospitals. The study examines service quality, social insurance entitlements, and patient behavior during illness. The study helps emerging private healthcare facilities understand the complex relationship between service quality and patient response, use statistically sound methods for accurate predictions, and improve health services.

Keywords: Discriminant analysis, Logistic regression, Service quality, Data Mining, Classification

บทคัดย่อ

การศึกษาเปรียบเทียบโดยใช้การถดถอยโลจิสติกและการวิเคราะห์จำแนกจะตรวจสอบความสำคัญของทั้งสองวิธีในการให้บริการด้านสุขภาพ การศึกษาจะเปรียบเทียบการจัดหมวดหมู่ข้อมูลของทั้งสองวิธีโดยใช้สถิติยอดนิยม 2 โปรแกรม เพื่อระบุความเกี่ยวข้องในการดูแลสุขภาพ การศึกษามุ่งเน้นไปที่จังหวัดเชียงรายทางภาคเหนือของประเทศไทยและปัจจัยที่มีอิทธิพลต่อการเลือกโรงพยาบาลเอกชนที่จัดตั้งขึ้นใหม่ การศึกษานี้ตรวจสอบคุณภาพการบริการ สิทธิประกันสังคม และพฤติกรรมของผู้ป่วยระหว่างเจ็บป่วย การศึกษานี้ช่วยให้สถานพยาบาลเอกชนที่เกิดขึ้นใหม่เข้าใจความสัมพันธ์ที่ซับซ้อนระหว่างคุณภาพการบริการและการตอบสนองของผู้ป่วย ใช้วิธีการที่ถูกต้องทางสถิติเพื่อการคาดการณ์ที่แม่นยำ และปรับปรุงบริการด้านสุขภาพ

คำสำคัญ: การวิเคราะห์จำแนกดิสคริมินันท์, การถดถอยโลจิสติก, คุณภาพการบริการ, การทำเหมืองข้อมูล, การจำแนกประเภท

Introduction

Customers choose new private hospitals based on healthcare provider quality and other factors. Trust, satisfaction, perceived competence, word of mouth, patient-centered care, facilities, technology, and healthcare experience affect patient preferences. Treatment benefits—rights, payment options, coverage schemes, and insurance—also affect healthcare adoption. Prior minor, severe, and emergency hospitalizations help new private hospitals succeed. To tailor services and grow sustainably, these hospitals must understand patient preferences and experiences. Logistic regression, discriminant analysis, and segmented data aid medical decision-making. Healthcare data mining and Chiang Rai private hospital growth are covered. Logistic/Discriminant Regression Logistic regression and discriminant analysis are popular for categorical predictions with many explanatory variables. Many medical and sociological studies use these methods, according to Horváthová and Mokriová (2020) and Tillmanns and Krafft (2021).

Data utilization requires analysis, interpretation, and reporting, which this study simplifies. Logistic regression's dichotomous, discrete, and categorical dependent variables aid in medical diagnosis. Discriminant analysis is a method for classifying natural group observations (Tavassoli and Saen, 2019; Mehrolia et al., 2021). Medical Data Mining and Management By 2025, healthcare will generate 36% of global data (Dash et al., 2019; Dhungana, 2021). Medical organizations benefit from segmented data for clinical decision-making, accurate diagnosis, effective therapy, and drug interaction mitigation (Sohrabi et al., 2019). Medical Scene Details Meechaiyo and Guo (2019) say Chiang Rai private hospitals are better connected to Myanmar, Laos, and southern China.

This study examines northern Thai patients' private hospital choices. Industrial attractiveness, competition, and service quality affect hospital preferences (Owusu-Frimpong et al., 2010). ASEAN sociological research illuminates private hospital growth, consumer income, medical tourism, and local medical supply competition. Healthcare data mining, new hospital Chiang Kasemrad has 13 public hospitals, including Chiangrai Prachanukroh, Mae Fah Luang University, Mae Sai, Wiang Pa Pao, Doi Tung, Phaya Mengrai, Mae Suai, Mae Chan, and Mae Sriphat Medical Centre, and four private hospitals: Bangkok Hospital Chiang Rai, McCormick Hospital, and Sribur, the largest hospital with 773 beds and many services. Other public hospitals are well-equipped and provide quality care. Private hospitals charge more but provide more. Thai hospital openings are examined. Success factors, healthcare market research, and data mining. Mousavi et al. (2014), Jothi and Husain (2015), Ushakov (2019), Punpukdee (2021), and (2023) illuminate this field. Discuss patient behavior, healthcare provider quality, treatment rights, and vital data mining. The article analyzes Chiang Rai's healthcare using statistics, regional economic factors, innovative approaches, and technology. This research examines healthcare theory, practice, and regional differences.

Objectives of the study

The research aims to determine the key factors that influence the selection of health services in a newly established private hospital in Thailand's Chiang Rai province. This will be achieved through the application of logistic regression and discriminant analysis. Additionally, it assesses the predictive accuracy of logistic regression and discriminant analysis algorithms, evaluating their performance in making predictions.

Literature review

Understanding patient preferences and prior experience with minor, severe, and emergency illnesses will help new private hospitals succeed. Sustainable growth requires prioritizing specialized care and building a reputation for excellent care, emergency response, and accessibility.

1) The quality of healthcare providers affected new private hospital adoption: healthcare quality, whether in public or private hospitals, strongly influences service adoption. Trust, satisfaction, perceived competence, word of mouth, patient-centered care, facilities and technology, hospital choice, and healthcare experience all influence choices. Patients trust healthcare providers who provide high-quality services, which influences their decision to use them. Staff skills, cleanliness, and communication affect patient satisfaction. In addition to technology and facilities, service competence and quality also have an impact on patients' perceptions. Providers also benefit from word-of-mouth and reputation. Thus, patient perceptions and choices depend on healthcare quality. P. Li et al. (2022)

2) Treatment benefit types and their perceived pros and cons can affect whether new private hospitals offer healthcare. Individuals' healthcare service adoption decisions, especially at newly opened private hospitals, depend on treatment benefit types. These include health insurance, treatment rights, self-payment options, universal coverage, social security, government officer rights, and state enterprise officer rights. Respecting patients' rights to choose medical treatments, informed consent, and refusal boosts autonomy and satisfaction. Transparent and affordable private hospitals can attract more people. Harmonizing or supplementing universal coverage schemes can attract more patients seeking affordable, comprehensive healthcare. Individual health insurance coverage must be examined to appeal to different insurance preferences. K. Nasri et al. (2022).

3) Minor illness admissions can affect whether new private hospitals offer healthcare. JZ Lock et al. (2023) said new private hospitals can be affected by previous minor illness treatment admissions due to factors like buying pharmacy drugs, visiting clinics, public hospitals, or private hospitals. Those who purchase over-the-counter medications for minor illnesses may prefer self-medication. For complex medical conditions, private hospitals should prioritize specialized care. Clinic visits may require specialized services and shorter wait times for fast and convenient care. While cost, accessibility, and perceived care quality may influence public hospitals, shorter wait times, personalized care, and superior service quality may attract patients seeking alternatives. Finally, new private hospitals must understand these preferences and behaviors to tailor their services and grow.

4) Previous serious illness can affect whether new private hospitals offer care: Aiello, T. F., et al. (2022) found that previous severe illness admission behavior, such as visits to drug stores, clinics, public hospitals, and private hospitals, affects new private hospital success and viability. Financial constraints, accessibility issues, and illness underestimation can affect patients' decisions. New private hospitals should emphasize immediate medical care awareness, affordable services, convenient locations, and educational campaigns. Clinic visits may indicate a preference for convenience, quality, and speed. Public hospitals must provide fast, high-quality care despite longer wait times. Private hospitals should prioritize quality, efficiency, and personalization to attract patients seeking a better healthcare experience. New private hospitals must understand these preferences to tailor their services, succeed, and grow.

5) Previous emergency hospitalization may affect new private hospital service: Essex, R., et al. (2023) found that emergency illness admission preferences and behaviors affect new private hospitals' success and viability. Patients prefer proximity, accessibility, and perceived preparedness during clinic

visits. New private hospitals should prioritize efficient and well-prepared emergency services to attract urgent patients. Reputation, accessibility, and emergency response quality may lead to public hospital selection. New private hospitals should compete with public hospitals by establishing a strong reputation for emergency response, accessibility, and quality care. Advanced technology, expert staff, and personalized emergency care may attract patients seeking high-quality care.

6) The literature review emphasizes logistic regression and discriminant analysis. IBM SPSS Statistics vs. RapidMiner explains modern data analytics. This analysis emphasizes these tools' importance in private hospital health service provider decision-making. Linear variables differentiate DA object or event classes (Punpukdee et al., 2021). This statistical method analyzes interval predictors and categorical dependent variables. Category-dependent variables have multiple categories, unlike interval predictors (Tabachnick et al., 2013). Statistical discriminant function analysis classifies. It aids independent-dependent relationship research (Stella, 2019). We build models from a representative subset of groups using a discriminant function. This function classifies new instances without group membership using measured predictor variables (Mazanec and Bartosova, 2021). Parametrically, Cacoullos (2014) found quantitative variable weightings that outperform random chance in differentiating multiple groups of examples. Logistic regression and discriminant analysis are popular categorical variable classification methods. Logistic regression explains categorical variables with continuous independent variables (Tabachnick et al., 2013). Discriminant analysis is popular for handling different assumptions and data types (Nikita & Nikitas, 2020). Both methods estimate category and continuous variable coefficients differently (Tabachnick et al., 2013). Discriminant analysis improves observation classification, say Tillmanns and Krafft (2021). This applies to discrete categorical variables across five categories. Logistic regression assumes no explanatory factor distribution, linear dependent variables, or group variance. Pohar et al. (2004) and Bewick (2005) found strong results without assumptions.

Classification Accuracy of Prediction Models and Data Mining Algorithms

A classification rule's hypothetical unit classification accuracy can be evaluated using various criteria (Hofmann and Klinkenberg, 2016). The error or misclassification rate is used in most approach comparisons (Worth and Cronin, 2003). The algorithm's hit rate indicates classification. According to O'Connell (2006), probabilities include ideal hits, conditional hits, and predicted error rates (Rossi, 1990). Hastie et al. (2009) call cross-validation a simple prediction error estimator. The accuracy and precision of algorithms affect private hospital health service selection. Emanet et al. (2014) define accuracy as 0–100. Accuracy Measurement: Emanet et al. described that the accuracy of an algorithm can be expressed as Accuracy = (TP + TN) / (TP + FN + FP + TN) *100. (TP = number of positive labeled data that are correctly classified; FP = number of negative labeled data that are incorrectly classified positive; FN = number of positive labeled data that are incorrectly classified negative; and TN = number of negative labeled data that are correctly classified.) Naser (2023).

Research conceptual framework

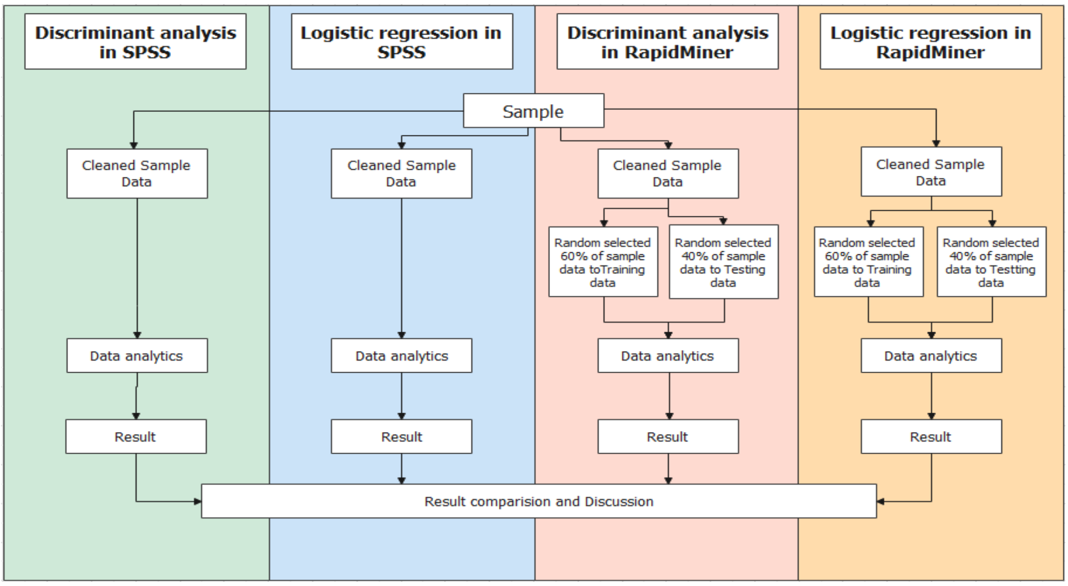

Figure 1 The infographic shows conceptual research framework in a comparative study of logistic regression and discriminant analysis

Source: Wondershare. (2023). EdrawMax (Version 13.0.0) [Software]. Wondershare. http://www.edrawsoft.com/

Research methodology

By 2025, healthcare will account for 36% of global data, up from 30%, according to Dash et al. (2019) and Dhungana (2021). This data segmentation can improve clinical decision-making, diagnosis, and drug-food interactions in medical organizations (Sohrabi et al., 2019). Case study of Chiang Rai, northern Thailand, patients' private hospital selection. Kiani et al. (2019) predicted outcomes using discriminant analysis and logistic regression. Researchers evaluated these methods' empirical and theoretical evidence to choose the best.

Hospitalized patients participated in the quantitative study. Chiang Rai, Thailand samples had a minimum count of 400 in Cochran's sample (Kaur & Kaur, 2023). Bujang et al. (2018) recommend a 500-sample minimum for parameter statistics in large population logistic regression observational studies. Sample size and predictor variable ratio affect discriminant analysis. Numerous studies show 21 predictor variables. Each independent variable should have five observations, per Liu (2015). RapidMiner randomly splits the sample into two groups, per László and Ghous (2020). Training datasets teach machine learning predictive model creation. The testing dataset assesses model performance. The characteristics suggest RapidMiner can perform logistic regression and discriminant analysis with the sample size. Researchers advised a 1,000-sample incidental sample study.

Five main factors influence people's private hospital selection decisions. Variables used for classification which shows 21 variables (5 Main variables) and 2 grouping variables. Data collection provided this study's analysis data. The questionnaire used in this study was based on previous research. In addition, the researcher added a new questionnaire question. Thus, the questionnaire had five parts:

This study uses SPSS and RM to classify data and predict outcomes using logistic regression and linear discriminant analysis (Tillmanns and Krafft, 2021). The paper discusses medical and sociological research using these methods. IBM SPSS Statistics analyzes statistical models like logistic regression and discriminant analysis, while RapidMiner (RM) analyzes conceptual research models without statistical assumptions. Logistic regression uses independent variables to predict outcome probabilities based on the dependent or criteria variable (Boateng and Abaye, 2019; Nikita and Nikitas, 2020). It is best for categorical predictions and variable distribution independence (Mehrolia et al., 2021). By identifying the criteria that distinguish two or more natural groups, discriminant analysis classifies observations. This method estimates orthogonal discriminant functions using linear combinations of standardized independent predictor variables (Tavassoli and Saen, 2019).

Data collecting

The decision to choose health services: variables for the predictive modeling

Table 1 Variables for the predictive modeling of decision to choose health services In upcoming private hospital

The grouping variable (2 groups)

1) **Choose** to use future private hospital health services.
2) **Not choose** to use future private hospital health services.

Classification variables (21 variables)

1) **Provider health service quality**:

1.1) Service quality of medical clinic: evaluates the perceived quality of medical clinic services. It includes staff skills, waiting times, cleanliness, communication, and clinic patient satisfaction.

1.2) Services quality of operating public hospitals: this measures public hospital service quality. It includes staff competence, waiting times, cleanliness, modern equipment, and patient experience in public hospitals.

1.3) Services quality of operating private hospitals: private hospital service quality is assessed by this variable. It includes medical professionals' expertise, technology, facilities, personalized care, and the patient's experience in private hospitals.

2) **Treatment benefit types**:

2.1) Various treatment rights: this variable focuses on individuals' rights to choose various medical treatments. It may include informed consent, treatment choice, and treatment refusal.

2.2) self-payment: individuals pay for healthcare services without insurance or third-party payers. Financial capacity, affordability, and service value may affect it.

2.3) Universal coverage scheme (gold/30baht): a government-sponsored healthcare coverage scheme offering comprehensive (gold) or basic plans for a fixed payment of 30 baht. Understanding attitudes and experiences with these schemes is crucial.

2.4) Social security scheme: this variable pertains to the system's healthcare offering. It may include insurance coverage, medical facility access, and social security system healthcare support satisfaction.

2.5) Government officer rights: examining healthcare benefits for government officers. Insurance, hospital access, and government-provided healthcare privileges may be included.

2.6) State enterprise officer rights: this variable addresses state enterprise employees' healthcare benefits, like government officers. It includes insurance, medical facility access, and state enterprise employee healthcare privileges.

2.7) Health insurance: examines individual health insurance coverage. Questions may include insurance type, coverage limits, and plan satisfaction.

3) **Treatment admission for slightly ill**:

3.1) Buy pharmacy drugs for slightly ill: assesses the behavior of individuals who choose to purchase over-the-counter drugs from pharmacies for minor illnesses.

3.2) Slightly ill will go to a clinic: this variable examines clinic visits by minor illness patients. It considers accessibility, perceived care quality, and convenience.

3.3) Slightly ill will go to a public hospital: this measures the likelihood of minor illness patients choosing public hospitals for medical care, like the clinic variable. Cost, accessibility, and perceived care quality are considered.

3.4) Slightly ill will go to a private hospital: This characteristic predicts mild illness patients choosing private hospitals. Excellent service, easy access, and tailored care are factors.

4) **Severe illness admission**:

4.1) severely ill will go to drug store: This variable examines severely ill patients who visit drug stores instead of doctors. Financial constraints, accessibility, and illness severity may affect it.

4.2) severely ill will go to a clinic: Evaluates the likelihood of individuals with severe illnesses seeking medical care at clinics. Accessibility, perceived care quality, and convenience are factors.

4.3) severely ill will go to a public hospital: Like the clinic variable, this measures the likelihood of severe illness patients choosing public hospitals. Cost, accessibility, and perceived care quality are considered.

4.4) severely ill will go to a private hospital: This variable measures the likelihood of severe illness patients choosing private hospitals. Higher service quality, faster access, and personalized care are factors.

5) **Emergency illness admission**:

5.1) Emergency Ill Will Go to a Clinic: This variable examines emergency clinic visits. It considers proximity, accessibility, and clinics' emergency preparedness.

5.2) Emergency Ill Will Go to Public Hospitals: This variable measures the likelihood of emergency patients choosing public hospitals, like the clinic variable. It evaluates public hospitals' reputation, accessibility, and emergency response.

5.3) Emergency Ill Will Go to Private Hospitals: This variable measures the likelihood of emergency patients choosing private hospitals. Private hospitals are perceived to have better services, faster access, and better emergency response.

Results

1. The elements that influence people's decisions for choosing health services from an upcoming private hospital in Thailand's Chiang Rai province.

1.1 Logistic regression result from using IBM SPSS Statistics (SPSS)

Out of the 1252 examples in the data set, 638 plan to use the services of a new private hospital opening in Chiang Rai, Thailand, while the other 614 do not.

Table 2 Logistic regression model summary

Model summary			
Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	1224.145[a]	.335	.447

Initial -2 Log Likelihood: 1735.180, in the iteration history table, compare the -2 Log Likelihood at step 1 with the -2 Log Likelihood at step 0. The decreasing -2 Log Likelihood indicates that the logistic regression equation is valid.

Table 3 Predicted result of logistic regression in SPSS

Predicted group membership: Health service selection from an upcoming private hospital				
Step 1		Do (638)	Do not (614)	Percentage Correct
	DO	486	206	76.2
	Do not	152	408	66.4
	Overall Percentage			71.4

Table 4 Variables in the logistic regression equation

Variables in the equation					
Classification variables	B	S.E.	Wald	df	Sig.
Service quality of medical clinic	-.597	.206	8.366	1	.004
Services quality of operating public hospitals	-.563	.188	8.926	1	.003
Services quality of operating private hospitals	1.029	.164	39.603	1	.000
Various treatment rights	-1.077	.353	9.331	1	.002
Self-payment	-.958	.319	9.026	1	.003
Gold patent	-2.430	.369	43.286	1	.000
Social security rights	-2.114	.519	16.582	1	.000
Employees' rights in state enterprises	1.502	.441	11.587	1	.001

From using IBM SPSS Statistics (SPSS), the logistic regression results reveal that factors such as service quality in medical clinics, public hospitals, private hospitals, various treatment rights, self-payment, having a gold patent, social security rights, and employees' rights in state enterprises can influence the likelihood of a specific outcome. Lower service quality in medical clinics increases the likelihood of experiencing the outcome. Similarly, lower service quality in public hospitals and higher service quality in private hospitals also increase the likelihood of experiencing the outcome. Having various treatment rights, self-payment, having a gold patent, having social security rights, and having employees' rights in state enterprises also increase the likelihood of experiencing the outcome.

1.2 Discriminant analysis result from using IBM SPSS Statistics (SPSS)

638 of the 1252 instances in the data set are interested in using the services of the new private hospital opening in Chiang Rai, Thailand, while the remaining 614 are not. This is the same data set utilized for logistic regression analysis.

Table 5 Predicted result of discriminant analysis in SPSS

Predicted group membership: Health service selection from an upcoming private hospital				
Step 1		Do (638)	Do not (614)	Percentage Correct
	DO	452	206	70.85
	Do not	186	408	66.45
	Overall Percentage			71.4

Table 6 Variables in the discriminant analysis equation

Variables in the equation						
Eigenvalue	% of Variance	Cumulative %				Canonical Correlation
.254	100.0	100.0				.450
Wilks' Lamba	Chi-square	df				Sig.
.789	281.525	10				<.001
Classification variables	Wilks' Lamba	F	df	df2	Sig.	Canonical discriminant function coefficient
Services Quality of operating public hospitals	.958	54.275	1	1250	.000	-.414
Severely ill will go to a private hospital	.996	5.302	1	1250	.021	.066
Severely ill will go to a public hospital	.977	29.640	1	1250	.000	-.306
Emergency ill will go to a public hospitals	.995	5.914	1	1250	.015	-.137
Various treatment rights	.996	4.441	1	1250	.035	-.118
Gold patent	.995	6.085	1	1250	.014	.139
Social security rights	.949	67.197	1	1250	.000	.460
Civil servant rights	.993	8.554	1	1250	.004	.164

From using IBM SPSS Statistics (SPSS), The study examines the impact of various variables on the discriminant function, focusing on the services quality of operating public hospitals, the preference for private hospitals for severely ill individuals, the preference for emergency cases in public hospitals, the influence of various treatment rights, the role of social security rights, and the role of civil servant rights. The results show that lower service quality in public hospitals negatively contributes to the discriminant function, as individuals with lower service quality are more likely to belong to a specific group. Severely ill individuals may prefer private hospitals due to their perceived higher quality of care, specialized services, or faster access to medical attention. Emergency cases in public hospitals have a negative impact, possibly due to the perception that public hospitals are more equipped to handle emergency cases. Possessing a gold patent also has a positive impact on the discriminant function, suggesting that individuals with such rights exhibit distinct healthcare-seeking behavior.

1.3 Logistic regression result from using RapidMiner (RM)

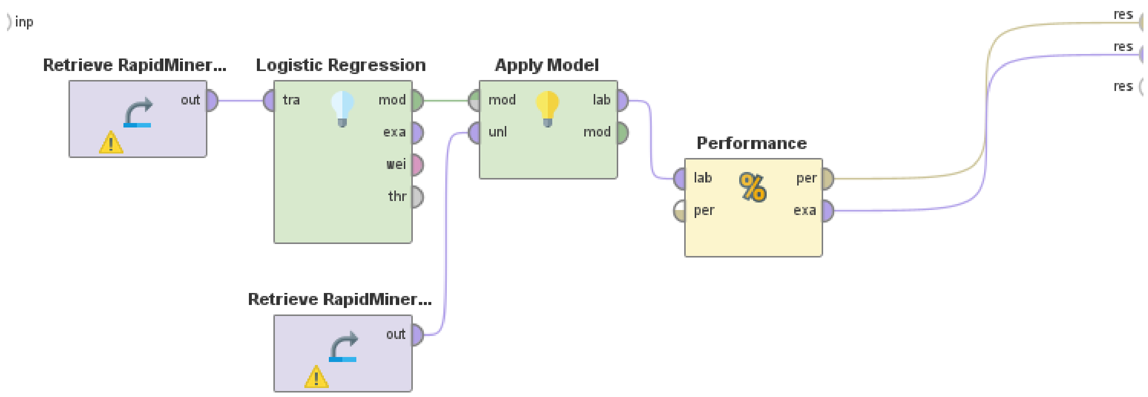

Figure 2 RapidMiner's logistic regression design in a comparative study
Source: RapidMiner. (2023). RapidMiner Studio (Version 10.1) [Software]. RapidMiner. http://www.rapidminer.com/

Table 7 Predicted result of discriminant analysis in RM

Predicted group membership: Health service selection from an upcoming private hospital				
Step 1		Do (638)	Do not (614)	Percentage Correct
	DO	448	154	70.22
	Do not	190	460	74.92
	Overall Percentage			71.4

Table 8 Variables in logistic regression equation

Attribute	Coefficient	Std. Coefficient	Std. Error	z-Value	p-Value
Services Quality of operating public hospitals	.868	.663	.120	4.134	.000
Social security rights	-1.618	-.541	.476	-3.397	.001
Health insurance	1.860	.583	.514	3.621	.0000
Buy pharmacy drugs for slightly ill	2.072	.988	.339	6.111	.000
Slightly ill will go to a clinic	2.420	.864	.438	5.526	.000
Severely ill will go to a clinic	-1.410	-.588	.414	-3.406	.001
Severely ill will go to a private hospital	-1.337	-.561	.429	-3.114	.002

From using RapidMiner, the logistic regression study reveals that the likelihood of a specific outcome is influenced by several factors. These include the quality of services provided by public hospitals, the availability of social security rights for certain demographics, the availability of health insurance, the preference for purchasing pharmacy drugs for minor illnesses, the preference for seeking medical attention in clinics, the preference for immediate medical attention in severe cases, and the preference for private hospitals for severe illnesses. These factors contribute to a comprehensive understanding of the factors influencing the likelihood of a specific outcome, thereby enhancing the overall health outcomes of individuals. The study provides a comprehensive understanding of these factors.

1.4 Discriminant analysis result from using RapidMiner (RM)

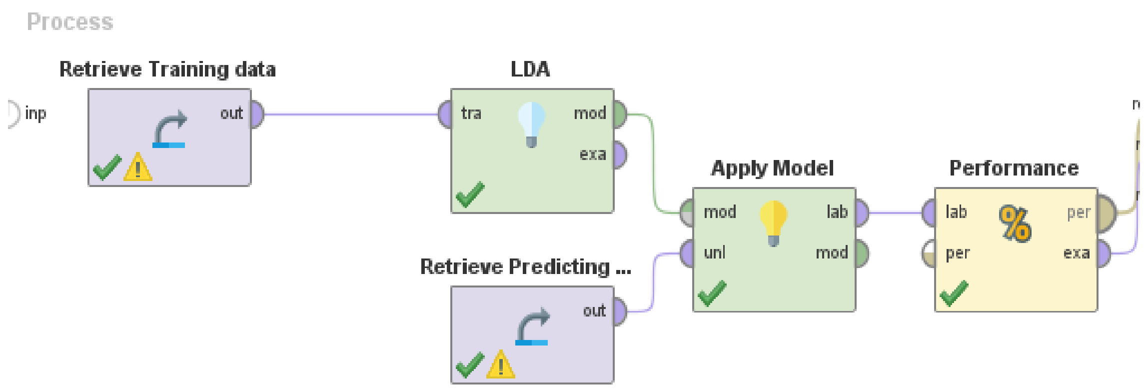

Figure 3 RapidMiner designed discriminant analysis (LDA) in a comparative study
Source: RapidMiner. (2023). RapidMiner Studio (Version 10.1) [Software]. RapidMiner. http://www.rapidminer.com/

Table 9 Predicted result of logistic regression in RM

Predicted group membership: Health service selection from an upcoming private hospital				
Step 1		Do (638)	Do not (614)	Percentage Correct
	DO	458	180	71.79
	Do not	120	494	80.46
Overall Percentage				71.4

Table 10 Variables in discriminant analysis equation

Attribute	Coeff.	Std. Coefficient	Std. Error	z-Value	p-Value
Service quality of medical clinic	-.413	-.223	.202	-2.049	.040
Services Quality of operating private hospitals	-.478	-.294	.189	-2.527	.011
Services Quality of operating public hospitals	.799	.609	.147	5.431	.000
Gold patent	-.717	-.355	.217	-3.303	.001
Social security rights	-1.994	-.672	.297	-6.713	.000
Health insurance	1.337	.421	.341	3.920	.000
Buy pharmacy drugs for slightly ill	2.094	.998	.218	9.618	.000
Slightly ill will go to a clinic	2.338	.853	.294	7.945	.000
Severely ill will go to a clinic	-1.494	-.627	.292	-5.111	.000
Severely ill will go to a private hospital	-1.343	-.564	.300	-4.473	.000
Emergency ill will go to a public hospitals	.585	.282	.286	2.050	.040

From using RapidMiner, this study explores the impact of service quality, socio-economic factors, healthcare access, and individual behavior on healthcare outcomes. It suggests that higher service quality is associated with better medical care, more comfortable environments, and improved patient experience. Public hospitals may have higher service quality due to government funding, better infrastructure, and a focus on serving a broader population. Possession of a gold patent and social security rights may indicate a specific socio-economic status, potentially causing inequalities in healthcare access. Health insurance can provide better healthcare access and timely interventions. Proactive healthcare-seeking behavior, such as purchasing pharmacy drugs and visiting clinics, may lead to better health outcomes. Hospital choices for severely ill individuals may reflect variations in healthcare quality and accessibility.

Based on what was discovered, classifying factors that influence to choose or not choose the service in an upcoming private hospital with conditions analyzed with Logistic regression and discriminant analysis with SPSS and RapidMiner programs. It is reasonable to assume that various analysis techniques and software packages were used to create various classification variables. “Service quality of operating public hospitals” and “social security rights” are both classification variables. While classification variables that do not influence the decision to choose or not choose the service include: “slightly ill will go to a public hospital”, “slightly ill will go to a private hospital”, “severely ill will go to a drug store”, “emergency ill will go to a clinic” and “emergency ill will go to a clinic”.

2. To compare the accuracy of a prediction algorithm that predicts decision to choose health services in an upcoming private hospital.

While it is feasible to compare the outcomes of logistic regression and discriminant analysis between SPSS and RapidMiner, it is important to note that there are several constraints and obligatory requirements. To conduct a comprehensive comparison of the outcomes obtained from logistic regression and discriminant analysis in both SPSS and RapidMiner, it is advisable to adopt a methodical and structured approach. This encompasses various stages of the machine learning process, such as data preprocessing, model setup, model training, evaluation of performance, interpretability analysis, randomization techniques, review of documentation, ensuring version consistency, validation and cross-validation procedures, consulting support, participation in communities, and iterative comparison of results. Ensuring a fair and reliable assessment can be achieved through consistent data preparation, model configuration, model training, evaluation metrics, interpretability, randomization, documentation review, version consistency, cross-validation procedures, and consulting support and communities. Iterative comparison can be performed by making modifications based on observed disparities, refining model setups, data preprocessing procedures, or addressing platform-specific intricacies. By employing this methodical approach, the ability to compare results is improved and the evaluation of both platforms becomes more dependable. Based on the reasons, the author aims to assess the predictive accuracy of logistic regression and discriminant analysis models individually for each software.

2.1 To compare the accuracy of prediction model between logistic regression and discriminant analysis result from using IBM SPSS Statistics (SPSS) software.

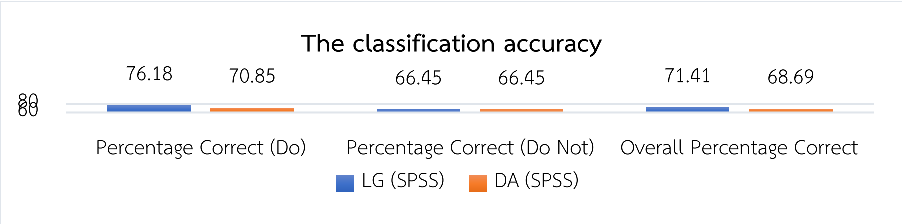

Figure 4 A comparative study of logistic regression and discriminant analysis classification accuracy using IBM SPSS

Source: RapidMiner. (2023). RapidMiner Studio (Version 10.1) [Software]. RapidMiner. http://www.rapidminer.com/

From IBM SPSS Statistics, the logistic regression had a higher percentage of overall correct decisions to choose health services in an upcoming private hospital than the discriminant analysis.

2.2 To compare the accuracy of prediction model between logistic regression and discriminant analysis result from using RapidMiner (RM) software.

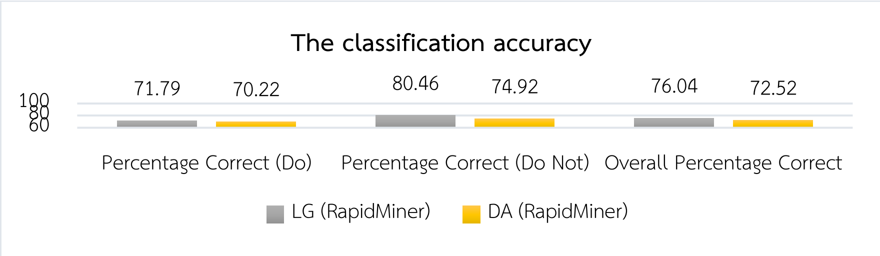

Figure 5 Comparison of logistic regression and discriminant analysis classification accuracy using RM
Source: RapidMiner. (2023). RapidMiner Studio (Version 10.1) [Software]. RapidMiner. http://www.rapidminer.com/

From RapidMiner software, the logistic regression had a higher percentage of overall correct decisions to choose health services in an upcoming private hospital than the discriminant analysis.

2.3 Summaries of comparison of data mining classification algorithms: the classification accuracy of decision to choose health services in upcoming private hospital.

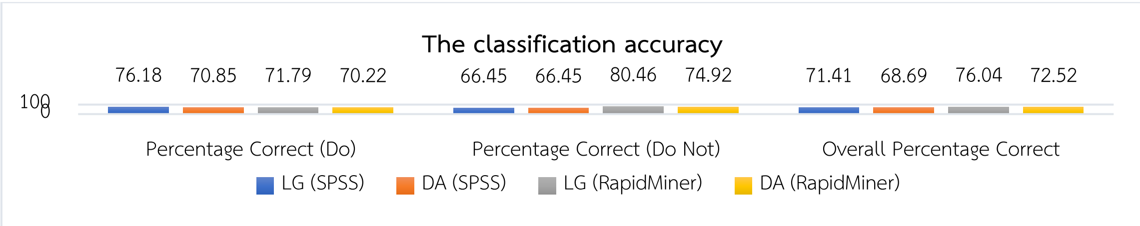

Figure 6 The study compared the classification accuracy of various data mining algorithms.
Source: RapidMiner. (2023). RapidMiner Studio (Version 10.1) [Software]. RapidMiner. http://www.rapidminer.com/

To summarize, the results indicate that the accuracy of selecting health services in an upcoming private hospital can be ranked from highest to lowest percentage overall correct: The order is as follows, from highest to lowest: RapidMiner offers two types of analyses: 1) logistic regression, and 2) discriminant analysis. When using SPSS, you can perform two types of analyses: 1) logistic regression, and 2) discriminant analysis.

Logistic regression and discriminant analysis play crucial roles in decision-making processes within the fields of healthcare and sociology. An in-depth comparative analysis of these methodologies and their practical applications elucidates the decision-making process in Thai healthcare. The researcher's methodology will be contingent upon the evidence and its interpretation, thereby underscoring the significance of statistical methods in contemporary scholarship.

Discussion and Suggestions

Based on the given findings, we can identify two primary contributions to knowledge:

1. Determinants of Private Hospital Choice: The findings provide valuable insights into the determinants that significantly impact individuals' decision-making processes when choosing services from a recently established private hospital. The logistic regression and discriminant analysis models identify key variables that are pivotal in this decision-making process. Favorable coefficients for factors such as the service quality of private hospitals, enhanced employee rights in state companies, and

specific preferences (such as a preference for mildly ill individuals to visit clinics) suggest that these aspects have a positive influence on the probability of selecting services in a private hospital. The user did not provide any text. In contrast, when factors such as the service quality of medical clinics, the service quality of operating public hospitals, and specific patient preferences (such as a preference for severely ill individuals to seek care at clinics or public hospitals) have negative coefficients, it indicates that a decrease in these parameters reduces the probability of choosing private hospital services. The user did not provide any text. This comprehension enhances the overall understanding of consumer behavior in the healthcare industry, elucidating the particular factors that individuals prioritize when selecting their healthcare service provider.

2. Methodological Insights from Logistic Regression and Discriminant Analysis: The results demonstrate the successful application of statistical techniques, particularly logistic regression, and discriminant analysis, in revealing patterns and connections within the dataset. The positive and negative coefficients obtained from these models act as quantitative measures of the intensity and orientation of these relationships. The user did not provide any text. Logistic regression quantifies the influence of each predictor variable on the probability of selecting private hospital services, whereas discriminant analysis identifies variables that substantially contribute to the distinction between groups. The user did not provide any text. These statistical techniques provide methodological insights for researchers to analyze and interpret complex healthcare decision-making processes. This emphasizes the significance of simultaneously considering multiple variables and comprehending their combined impact on the desired outcome. Healthcare management and policy planning professionals can utilize these methodological insights to develop more focused and efficient strategies for enhancing private hospital services while considering the identified influential factors. Overall, the primary contributions of this study are the improved comprehension of the factors that impact the choice of private hospitals and the methodological insights gained from utilizing logistic regression and discriminant analysis. These findings make a valuable contribution to the wider domains of healthcare management, consumer behavior, and statistical methodology.

Suggestions for applying the research results.

Private hospitals can bolster their competitive edge by improving service quality, prioritizing employee rights and satisfaction, emphasizing premium patents and social security entitlements, and implementing targeted marketing and outreach strategies. Hospitals can increase patient numbers by fostering positive associations, such as improved public hospital service quality, health insurance availability, and the ability to choose clinics. To counteract negative perceptions, it is important to address negative correlations, such as the quality of medical clinic services, the operation of private hospitals, the absence of gold patents, and the lack of social security rights.

Customizing services to suit individual patient preferences can also be beneficial. Hospitals ought to periodically evaluate and modify their strategies, employing feedback mechanisms such as patient surveys and focus groups to obtain valuable insights. Hospitals can effectively adjust their strategies by staying updated on regulatory changes.

Executing these recommendations necessitates cooperation among hospital administration, healthcare practitioners, and marketing departments. To improve their competitiveness and meet the expectations of their target audience, private hospitals can strategically address the determinants identified in the study and utilize methodological insights from logistic regression and discriminant analysis.

Suggestions for further research

The dependent variable, research goals, and assumptions determine whether to use logistic regression or discriminant analysis. Logistic regression works for binary outcome variables, while discriminant analysis works for multi-level categorical dependent variables. Logistic regression requires a larger sample size than discriminant analysis. It assumes fewer independent variable distribution assumptions and no normal predictor distribution. Logistic regression produces odds ratios, which help explain how predictors affect event likelihood. However, it does not assume group variance-covariance matrices are equal. This method focuses on probabilities, while discriminant analysis distinguishes groups. Logistics regression is often better for studying the factors that influence the acceptance or rejection of healthcare services at newly established private hospitals.

Study limitations

The study utilized RapidMiner's free version, which does not include Turbo Prepare, a feature that enhances proficiency, tidiness, and effectiveness in data preparation. The "Split Data" operator in RapidMiner partitions the data into separate sets for training and testing purposes. This division can be tailored according to specific requirements using SPSS's random select data feature. Manual data splitting provides the advantages of control, simplicity, and rapid correction. However, when dealing with large datasets, it can introduce bias and require a significant amount of time. It is unjust to compare SPSS and RapidMiner because SPSS utilizes all available data to build and evaluate models, whereas RapidMiner divides the data into separate training and testing sets. The researcher conducted a comparison between different models instead of relying on the findings from analytic software.

Reference

Aiello, T. F., Puerta-Alcalde, P., Chumbita, M., Monzó, P., Lopera, C., Hurtado, J. C., ... & Garcia-Vidal, C. (2022). Infection with the Omicron variant of SARS-CoV-2 is associated with less severe disease in hospitalized patients with COVID-19. *Journal of Infection*, *85*(5), e152-e154.

Bewick, V., Cheek, L., & Ball, J. (2005). Statistics review 14: Logistic regression. *Critical care*, *9*(1), 1-7.

Boateng, E. Y., & Abaye, D. A. (2019). A review of the logistic regression model with emphasis on medical research. *Journal of data analysis and information processing*, *7*(4), 190-207.

Bujang, M. A., Sa'at, N., Bakar, T. M. I. T. A., & Joo, L. C. (2018). Sample size guidelines for logistic regression from observational studies with large population: emphasis on the accuracy between statistics and parameters based on real life clinical data. *The Malaysian journal of medical sciences: MJMS*, *25*(4), 122.

Cacoullos, T. (Ed.). (2014). *Discriminant analysis and applications*. Academic Press.

Dash, S., Shakyawar, S. K., Sharma, M., & Kaushik, S. (2019). Big data in healthcare: management, analysis and future prospects. *Journal of Big Data*, *6*(1), 1-25.

Dhungana, R. R., Pandey, A. R., & Shrestha, N. (2021). Trends in the prevalence, awareness, treatment, and control of hypertension in Nepal between 2000 and 2025: a systematic review and meta-analysis. *International journal of hypertension*, *2021*, 1-11.

Emanet, N., Öz, H. R., Bayram, N., & Delen, D. (2014). A comparative analysis of machine learning methods for classification type decision problems in healthcare. *Decision Analytics*, *1*, 1-20.

Essex, R., Ahmed, S., Elliott, H., Lakika, D., Mackenzie, L., & Weldon, S. M. (2023). The impact of strike action on healthcare delivery: a scoping review. *The International Journal of Health Planning and Management*, *38*(3), 599-627.

Google. (n.d.). [*The hospitals in Chaing Rai - Google Maps*]. Retrieved October 1, 2023, from https://www.google.com/maps/@19.8852803,99.7534968,12z?entry=ttu

Gupta, M. K., & Chandra, P. (2020). A comprehensive survey of data mining. *International Journal of Information Technology*, *12*(4), 1243-1257.

Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction* (Vol. 2, pp. 1-758). New York: springer.

Hofmann, M., & Klinkenberg, R. (Eds.). (2016). *RapidMiner: Data mining use cases and business analytics applications*. CRC Press.

Horváthová, J., & Mokrišová, M. (2020). Comparison of the results of a data envelopment analysis model and logit model in assessing business financial health. *Information*, *11*(3), 160.

Jothi, N., & Husain, W. (2015). Data mining in healthcare–a review. *Procedia computer science*, *72*, 306-313.

Kaur, R., & Kaur, G. (2023). Managerial leadership and turnover intentions with the mediating role of employee happiness: an empirical study based on private universities. *The TQM Journal*, ahead-of-print.

Kiani, J., Camp, C., & Pezeshk, S. (2019). On the application of machine learning techniques to derive seismic fragility curves. *Computers & Structures*, *218*, 108-122.

László, K., & Ghous, H. (2020). Efficiency comparison of Python and RapidMiner. *Multidiszciplináris Tudományok*, *10*(3), 212-220.

Li, P., Luo, Y., Yu, X., Mason, E., Zeng, Z., Wen, J., ... & Jalali, M. S. (2022). Readiness of healthcare providers for e-hospitals: a cross-sectional analysis in China before the COVID-19 period. *BMJ open*, *12*(2), e054169.

Liu, X. (2015). *Methods and applications of longitudinal data analysis*. Elsevier.

Lock, J. Z., Khoo, Z. X., & Pek, J. H. (2023). Paediatric one-day admission: why and is it necessary?. *Singapore Medical Journal*.

Mazanec, J., & Bartosova, V. (2021). Prediction model as sustainability tool for assessing financial status of non-profit organizations in the Slovak Republic. *Sustainability*, *13*(17), 9721.

Meechaiyo, N., & Guo, J. (2019). Strategic Development and Comparative Advantages of Mengla Border Special Economic Zone, China and Chiang Rai Special Economic Zone, Thailand. *Journal of Mekong Societies*, *15*(2), 55-70.

Mehrolia, S., Alagarsamy, S., & Solaikutty, V. M. (2021). Customers response to online food delivery services during COVID‐19 outbreak using binary logistic regression. *International journal of consumer studies*, *45*(3), 396-408.

Mousavi, S. M. H., Agharahimi, Z., Daryabeigi, M., & Rezaei, N. (2014). Implementation of clinical governance in hospitals: challenges and the keys for success. *Acta Medica Iranica*, 493-495.

Naser, M. Z. (2023). Machine learning for all! Benchmarking automated, explainable, and coding-free platforms on civil and environmental engineering problems. *Journal of Infrastructure Intelligence and Resilience*, *2*(1), 100028.

Nasri, K., Amara, M., & Helmi, I. (2022, September). The Landscape of Social Protection in Tunisia. Economic Research Forum (ERF).

Nikita, E., & Nikitas, P. (2020). Sex estimation: a comparison of techniques based on binary logistic, probit and cumulative probit regression, linear and quadratic discriminant analysis, neural networks, and naïve Bayes classification using ordinal variables. *International journal of legal medicine*, *134*(3), 1213-1225.

O'Connell, A. A. (2006). *Logistic regression models for ordinal response variables* (Vol. 146). sage.

Pohar, M., Blas, M., & Turk, S. (2004). Comparison of logistic regression and linear discriminant analysis: a simulation study. *Metodoloski zvezki*, *1*(1), 143.

Punpukdee, A. (2023). Considering Consumer Behavior in Thailand's Traditional and Digital Markets. *Journal of Contemporary Issues and Thought*, *13*(1), 25-34.

Punpukdee, A., Wattana, C., Punpairoj, W., Srichuachom, U., Yaklai, P., & Trongtortam, S. (2021). Key Factors Affecting Consumer Behavior in the Thailand Offline Marketing: Research Synthesis by Systematic Literature Review and Data Mining Techniques. *Journal of Humanities and Social Sciences Thonburi University*, *15*(2), 21-32.

Rossi, J. S. (1990). Statistical power of psychological research: What have we gained in 20 years?. *Journal of consulting and clinical psychology*, *58*(5), 646.

Sohrabi, B., Raeesi Vanani, I., Nikaein, N., & Kakavand, S. (2019). A predictive analytics of physicians prescription and pharmacies sales correlation using data mining. *International Journal of Pharmaceutical and Healthcare Marketing*, *13*(3), 346-363.

Stella, O. (2019). Discriminant analysis: An analysis of its predictship function. *Journal of Education and Practice*, *10*(5), 50-57.

Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2013). *Using multivariate statistics* (Vol. 6, pp. 497-516). Boston, MA: pearson.

Tavassoli, M., & Saen, R. F. (2019). Predicting group membership of sustainable suppliers via data envelopment analysis and discriminant analysis. *Sustainable Production and Consumption*, *18*, 41-52.

Tillmanns, S., & Krafft, M. (2021). Logistic regression and discriminant analysis. In *Handbook of market research* (pp. 329-367). Cham: Springer International Publishing.

Ushakov, D. S., Yushkevych, O. O., Ovander, N. L., Tkachuk, H. Y., & Vyhovskyi, V. H. (2019). The strategy of Thai medical services promotion at foreign markets and development of medical tourism. *Geo Journal of Tourism and Geosites*, *27*(4), 1429-1438.

Worth, A. P., & Cronin, M. T. (2003). The use of discriminant analysis, logistic regression and classification tree analysis in the development of classification models for human health effects. *Journal of Molecular Structure: THEOCHEM*, *622*(1-2), 97-111.