IJTMGH International Journal of Travel Medicine and Global Health

Int J Travel Med Glob Health. 2023 Mar;11(1):186-193
http://ijtmgh.com

doi 10.30491/IJTMGH.2022.364468.1316

Original Article Open Access

# Investigation of Factors Affecting Choice of Medical Travel Destination Using Data Mining Techniques

**Sevda Janalipour Jenizeh[1]*, Filiz Ersöz[1]**

[1] Department of Industrial Engineering, Karabuk University, Turkey.

**Corresponding Author:** Sevda Janalipour Jenizeha, Department of Industrial Engineering, Karabuk University, Turkey. Tel: +905468280875, Email: sevdajanalipoor@yahoo.com.

**Received** 2022-10-04**; Accepted** 2023-01-04**; Online Published** 2023-03-01

**Abstract**

**Introduction:** Medical tourism, one of the most profitable industries, has been growing rapidly in recent years. Especially Turkey, which has a high ranking among medical travel destinations, has some advantages that can become preferable for international patients. This study is among the first few studies which examine affecting factors in patients' medical travel destination choices with Data Mining techniques.

**Methods:** The data were obtained from patients who came to Ankara from abroad for treatment in May 2015 through a question-naire. Cross-industry Standard Process for data mining, known as the CRISP-DM method, is used in this study. After cleaning out the missing data, the models were created using classification algorithms.

**Results:** Models including Generalized Linear Model, Deep Learning, Decision Tree, Random Forest, Gradient Boosted Trees, and Support Vector Machine (SVM) were compared, and SVM reached the best performance with 0.2% Relative Error, 0.014 Root Mean Squared Error and 0.998 Correlation. As a result of the SVM model, effective attributes in patients' satisfaction level include low price advantage, advertisement, doctors with high-quality education, trained assistant staff, relatives living in Turkey, and high technology of medical equipment, respectively.

**Conclusion:** Special attention should be paid to these factors in developing plans and policies for the health tourism sector. However, the importance of related socio-demographic variables was indicated in detail. Eventually, some suggestions were presented to improve the weaknesses in the health tourism sector.

**Keywords:** Health Tourism, Medical Travel Destination, Data Mining, Support Vector Machine, Turkey

**Citation**: Janalipour Jenizeh S, Ersöz F. Investigation of Factors Affecting Choice of Medical Travel Destination Using Data Mining Techniques. Int J Travel Med Glob Health. 2023; 11(1):186-193. doi: 10.30491/IJTMGH.2022.364468.1316

## Background

Medical tourism has grown rapidly in recent years, especially due to high-cost treatment in advanced countries. Patients from developed countries are now traveling to developing countries that can access high-quality medical services at reasonable cost[1]. Since its inception, medical tourism has been one of the most profitable tourism sectors in developing countries due to its benefit both for medical tourists and the host country[2,3]. Health tourism is expected to keep attracting interest, not only because of its vast economic benefits but also because of potential implications for healthcare programs and public policy[4]. According to Patients beyond the Border, the worldwide medical tourism market is growing at a rate of 15-25%[5]. Turkey receives around 500,000 medical tourists annually for medical treatment[6]. Although India, Cuba, Costa Rica, Thailand, Singapore, Colombia, and Malaysia are the most popular developing countries in medical tourism[7]. Asia's medical tourism industry is growing, and Turkey is among the first. An assessment by the International Medical Travel Journal estimates the global market is worth about $10 billion, and Turkey is the third biggest beneficiary, with annual revenue of $600 million. In Table 1, the top 10 medical travel destinations have shown by value[8].

**Table 1.** Medical Travel destination rankings

| Rank | Country | USD (million) |
|---|---|---|
| 1 | United States | 3,500 |
| 2 | South Korea | 655 |
| 3 | Turkey | 600 |
| 4 | Thailand | 600 |
| 5 | Germany | 575 |
| 6 | India | 450 |
| 7 | United Kingdom | 350 |
| 8 | Malaysia | 350 |
| 9 | Mexico | 350 |
| 10 | Iran | 315 |
| | Others | 2,685 |
| | **Total** | **10,430** |

Turkey has the potential to become one of the world's leading medical tourism destinations[9]. Altin et al. mentioned the advantages of Turkey in health tourism, such as lower cost for medical treatment in comparison with the USA and European countries, well-trained doctors and staff ability to speak at least one foreign language, modern technical hospitals with high bed capacities, and proximity to Europe and the Middle East [10]. According to the TUIK (Turkish Statistical Institute), 33,036,774 foreign patients came to Turkey for treatment, and 1,020,134,000 $ expenditure was made in 2017. These figures were 38,710,099 patients and 1,110,843,000 $ for 2018. While in 2019 year, 44,208,886 patients spent 1,394,015,000 $ in Turkey accordingly it shows an upward trend[11]. Detailed data and tourists' purposes for coming to Turkey are presented in Table 2.

**Table 2.** Tourists' purposes for coming to Turkey

| | Patients number in 2017 | % | Patients number in 2018 | % | Patients number in 2019 | % |
|---|---|---|---|---|---|---|
| A | 19,389,968 | 50.6 | 25,355,412 | 65.5 | 29,956,670 | 67.7 |
| B | 8,436,850 | 25.5 | 8,050,784 | 20.8 | 8,712,806 | 19.7 |
| C | 1,780,820 | 5.4 | 1,902,089 | 4.9 | 1,850,208 | 4.2 |
| D | 1,505,756 | 4.6 | 1,433,776 | 3.7 | 1,632,818 | 3.7 |
| E | 433,29 | 1.3 | 551,75 | 1.4 | 662,087 | 1.5 |
| F | 104,90 | 0.3 | 114,04 | 0.3 | 135,930 | 0.3 |
| G | 27,01 | 0.1 | 29,07 | 0.1 | 80,643 | 0.2 |
| H | 20,59 | 0.1 | 55,15 | 0.1 | 94,272 | 0.2 |
| I | 1,337,588 | 4.0 | 1,218,028 | 3.2 | 1,074,452 | 3,00 |
| J | 33,036,774 | 100.0 | 38,710,099 | 100.0 | 44,208,886 | 100.0 |

**A**: Travel, entertainment, sports, and cultural activities; **B**: Visiting relatives and friends; **C**: Business (conference, meeting, task, etc); **D**: Shopping; **E**: Health and medical reasons; **F**: Education, internship; **G**: Religious; **H**: Transit; **I**: Other; **J**: Total

The use of data mining has been growing rapidly in recent years due to the production of vast amounts of data and the necessity for extracting useful information and knowledge from such data. The information obtained can be used for applications ranging from business management, production control, and market analysis to emerging design and science exploration and health data analysis[12]. Institutions should pay special attention to their internal structure, data set, target, and source. The most essential condition in the successful application of data mining is determining the purpose it is being used for in the institution [13]. Data mining techniques are efficient and accurate research tools for the medical tourism sector to identify and explore valuable patterns and relationships among a large number of variables.

Here, we mention some of the latest articles in which medical tourism has been studied. In the study of Cham et al., the factors affecting Chinese medical tourists' decision to choose Malaysia as a destination for medical tourism have been investigated. They used self-administered questionnaires for collecting data and analyzed them using a structural equation modeling approach in AMOS and SPSS[14]. Moghadam et al. examined the importance of medical and cultural sensitivities in the development of medical tourism marketing by collecting data from people working in the medical universities of Iran[15]. The study conducted by Malik aims to explore the perception of medical tourists in choosing a destination for medical treatment. Structured equation modeling and hierarchical regression analysis have been performed in this research[16]. A study conducted in Malaysia showed that experience quality is a crucial factor in increasing tourists' perceptions by reducing the negative effect of perceived sacrifice in medical tourism, such as perceived risks and perceived fees[17]. Esiyok et al. examined the relationship between the origin countries of foreign patients and cultural distance, along with other variables of interest, by regression analysis and found a quadratic relationship between cultural distance and the number of foreign patients[18]. In the model proposed by Aydın and Karamehmet, it is found that factors such as costs, cultural distance, political and/or economic stability, regulations and legal framework, overall quality of care, and trust significantly affect health tourism in Turkey[19]. The study by Ganguli and Ebrahim takes an in-depth qualitative method to determine the factors that make Singapore a competitive destination for medical tourism [20]. Luboweicki-Vikuk and Kurkowaik aims to fill the gap in the research on the process of shaping the image of a medical tourism destination, with a particular focus on the medical tourism potential in the central and Eastern parts of Europe (CEE). In this model, ward's algorithm was used to construct the data structure agglomerative clustering in Statistica software[21]. A qualitative research in Turkey aims to identify the factors that affect medical tourism development in a private hospital [22]. The study of Han and Hyun constructs a model that describes foreign medical travelers' intention formation by considering the influence of quality, satisfaction, trust, and reasonableness of prices[23]. In another study conducted in Malaysia, medical staff quality reflects the functional aspect of care, which has the highest mean score among medical tourism service quality dimensions and can predict patient satisfaction, perceived value, and future intention for treatment[24]. The study of Guiry and vequist aimed to determine South Korea's medical tourism destination personality by quantitative analyses based on 1588 US consumers via an online survey [25]. In a novel study, contributory factors to the competitiveness of Iran's medical tourism were investigated and identified using importance-performance analysis. They showed in the results that "medical services" were more important than the two components, "special factors of tourism" and "characteristics of medical tourism destinations"[26]. In another research conducted by Aminmansour et al. the challenges of medical tourism after the prevalence of Covid-19 in the field of neurosurgery were determined. For this purpose, 500 patients with neurosurgical diseases registered in medical tourism companies were identified and included in the study[27].

Data mining and classification approaches are used in various areas. In one of these studies, the C5.0 Decision Tree algorithm was used, and the data were analyzed by Rapidminer software for finding the influential factors on the hospitalization of patients subject to chronic obstructive pulmonary disease [28]. In the study of Senturk, the Artificial Neural Network method has been applied to detect Diabetic Retinopathy [29]. Massaro et al. showed that the Artificial Neural Network-based Deep Learning algorithm is the best tool for forecasting the sales of a large retail chain. High correlation and low absolute and relative errors were found in this method. The Gradient Boosted Trees algorithms also offers good performance and could be an alternative approach [30]. The paper of Naik and Samant conducted a correlation review of classification algorithms. Indian liver patient dataset was used to classify the people with or without liver disorder. K-Nearest Neighbor's accuracy was considered to be higher as compared with Decision Tree and Naive Bayes[31].

There is little research focusing on medical tourism with the data mining approach; hence more studies are needed to investigate the medical tourism issues with one of the data mining approaches such as classification, clustering, and association. Thereore, this work aims to shed some light on different sensitivities affecting medical tourism using data mining techniques. For this purpose, the data was collected from hospitals in Turkey through a questionnaire that has been responded to by 105 patients from 11 different countries. This study attempts to determine various factors that influence the future behavior of medical tourists when traveling to Turkey for medical assistance. The Results of the present study can be a reference for medical and tourism industries from which future promotional strategies can be developed.

## 2. Material and Methods

The success of data mining techniques in developing a predictive model highly depends on the comprehensive of the selected methodology [28]. In this work, a cross-industry standard process for data mining known as the CRISP-DM method was applied. This method breaks the process of data mining into six major phases, Fig.1.

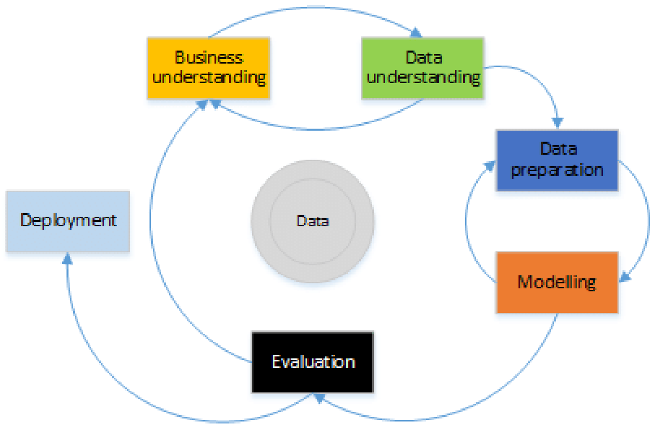


**Figure 1.** CRISP-DM methodology phases

### 2.1. Business understanding

This study, aimed to determine the factors that affect the patients' choice of medical travel destination, detect the current problems of the sector, and make suggestions for its development.

### 2.2. Data understanding

Data collection was conducted in two stages. In the first stage, a conceptual framework was created with the literature review. In the second stage, the hypotheses were determined based on this framework, and the application was put forward by collecting data. The research was carried out in May 2015 in the form of a questionnaire responded to by 105 patients from abroad for treatment to some education and research hospitals and private hospitals providing health-care services in Ankara. In the first part of the questionnaire, socio-demographic questions such as gender, age, marital status, education level, nationality, country/city of residence, disease to be treated, and monthly income of the international patients taking part in the survey were asked. In the second part, factors that enable patients to prefer Turkey for treatment and closed-ended questions consisting of 22 statements were prepared due to the information obtained from the literature review. In this section, the reasons for choosing Turkey to be treated, the sufficiency of the hospitals, and the problems and suggestions according to the patients were tried to be determined. Five-point Likert scale, one of the metric scale types, was used for the evaluation. On a scale, "1: Strongly Disagree, 2: Disagree, 3: Neutral, 4: Agree, 5: Strongly Agree" represents expressions.

In this section socio-demographic characteristics of the patients participating in the survey were examined in detail. Gender, marital status, nationality, insurance, and disease were specified as nominal variables in statistical analysis. While age, income, education level were specified as ordinal. Descriptive statistics regarding the socio-demographic characteristics of the patients are given in Table 3.

**Table 3.** Socio-demographic characteristics of the patients

| Variables | Subgroup Variables | n | % |
|---|---|---|---|
| Gender | Female | 37 | 35.2 |
| | Male | 68 | 64.8 |
| Marital Status | Married | 88 | 83.8 |
| | Divorced | 9 | 8.6 |
| | Single | 8 | 7.6 |
| Age | 18-30 | 20 | 19.0 |
| | 31-40 | 52 | 49.5 |
| | 41-50 | 17 | 16.2 |
| | 51-60 | 1 | 1.0 |
| | ≥61 | 15 | 14.3 |
| Education Level | Elementary school | 28 | 26.7 |
| | Intermediate school | 20 | 19.0 |
| | High school | 8 | 7.6 |
| | Undergraduate | 23 | 21.9 |
| | Graduated | 26 | 24.8 |
| Nationality | Syria | 12 | 11.4 |
| | Sudan | 1 | 1.0 |
| | Indonesia | 2 | 1.9 |
| | Turkish republics | 1 | 1.0 |
| | Iraq | 45 | 42.9 |
| | Libya | 23 | 21.9 |
| | Kosovo | 7 | 6.7 |
| | Pakistan | 7 | 6.7 |
| | Afghanistan | 7 | 6.7 |
| Insurance | State | 31 | 29.5 |
| | Private | 40 | 38.1 |
| | No | 34 | 32.4 |
| Disease | Eye diseases | 13 | 12.4 |
| | Kidney diseases and dialysis | 1 | 1.0 |
| | Gynecological | 29 | 27.6 |
| | Cancer treatment | 1 | 1.0 |
| | Orthopedic | 17 | 16.2 |
| | Aesthetic and plastic operations | 1 | 1.0 |
| | Heart diseases | 11 | 10.5 |
| | Infertility (IVF) | 15 | 14.3 |
| | Respiratory diseases | 3 | 2.9 |
| | Diabetes | 7 | 6.7 |
| | Dental | 7 | 6.7 |
| Income | 1001-2000 $ | 40 | 38.1 |
| | 2001-5000 $ | 21 | 20.0 |
| | 5001-10000 $ | 37 | 35.2 |
| | 10001-50000 $ | 7 | 6.7 |

### 2.3. Data Preparation

To detect outliers in data, we used a distance-based outlier detection model in the Rapidminer program. Two variables, “other satisfied points” and “other dissatisfied points,” have missing values and should be removed before modeling. Also, 12 variables, such as age, gender, marital status, education level, nationality, hospital type, income, etc., were transformed. We changed the type of these variables from number to category. Since there is no continuous data, outlier and extreme values were not detected.

### 2.4. Modeling

The data obtained through the questionnaire within the scope of the research were analyzed by classification techniques in Rapidminer software.

#### 2.4.1. Classification Techniques

Data classification algorithm is a method in which a hypothesis is chosen from a set of alternatives that best matches a series of observations. The process of classifying data comprises two steps [31]:

i. Building the classification model: here, the classification is developed by learning the training set and the class labels assigned to it.
ii. Classification using classifier: the classifier is used for classification in this step. The test data were used here to estimate the accuracy of the rules for the classification.

After data preparation, mentioned data were loaded. Here, we want to predict the satisfaction level. Subsequently, it has 44 independent variables and one class variable. Six data mining techniques, including Generalized Linear Model, Deep Learning, Decision Tree, Random Forest, Gradient Boosted Trees, and Support Vector Machine, were run.

### 2.5. Evaluation

As a result of comparing classification models, Table 4 was obtained.

**Table 4.** Performance comparison of different classification models

| Model | Relative Error (%) | Root Mean Squared Error | Correlation |
|---|---|---|---|
| Generalized Linear Model | 1.0 | 0.047 | 0.989 |
| Deep Learning | 0.9 | 0.042 | 0.982 |
| Decision Tree | 0.6 | 0.031 | 0.996 |
| Random Forest | 1.8 | 0.091 | 0.967 |
| Gradient Boosted Trees | 1.0 | 0.06 | 0.968 |
| Support Vector Machine | 0.2 | 0.014 | 0.998 |

As demonstrated, the Support Vector Machine (SVM) model has low Relative and Root Mean Squared Errors and high Correlation.

### 2.5.1. Support Vector Machine model

The SVM was first proposed by Cortes and Vapnik in 1995. It is a new machine learning method based on statistical learning theory, which follows the principle of structural risk minimization [32]. It has a solid theoretical basis compared to traditional learning technologies. Its performance is higher than conventional methods, according to a many applications. So far, it has involved many fields, such as pattern recognition, regression estimation, data mining, information retrieval, intelligent signal processing, etc. [33]. SVMs are based on a hyperplane in the form of $w.X + b = 0$ that optimally separates a set of n-dimensional vectors ($X_i \in R_n$) into two categories. This optimal hyperplane has the farthest distance from support vectors and the nearest data points from each class. Finding w is equivalent to solving a quadratic programming problem. To solve this problem, a tradeoff parameter ($c > 0$) needs to be determined. To categorize vectors that are not linearly separable, a kernel function such as degree-d polynomial, radial basis, or hyperbolic tangent and Linear is used to map the observed multidimensional vectors to a space with higher dimensions [34]. Fig. 2 shows the schematic of SVM.

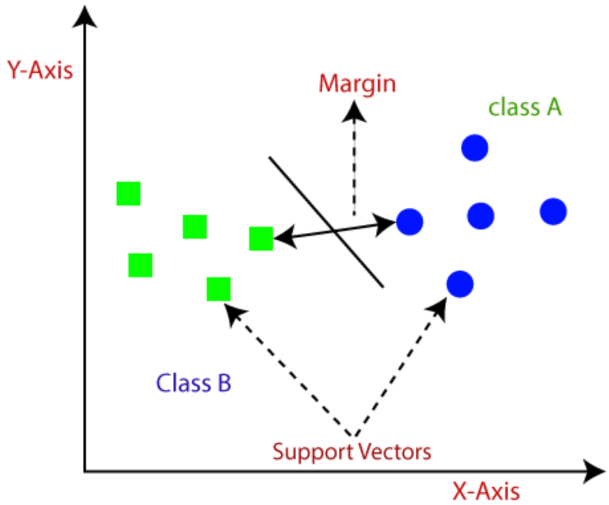


**Figure 2.** Diagram of Support Vector Machine

In the modeling, we used radial kernel type. The Radial Kernel is determined by

$$\exp(-g \, \| x - y \| \wedge 2)$$

Where g is the gamma, it is stated by the kernel gamma parameter. The adjustable parameter gamma plays a significant role in the kernel's performance and should be carefully tuned to the problem. This kernel type has another parameter indicated by C. It is the SVM complexity constant that sets the misclassification tolerance, where higher C values allow for 'softer' boundaries and lower values create 'harder' boundaries. A large complexity constant may lead to over-fitting, while over-generalization results from small values [35]. The process of the SVM model in this study for medical tourism data using Rapidminer is shown in Fig. 3.

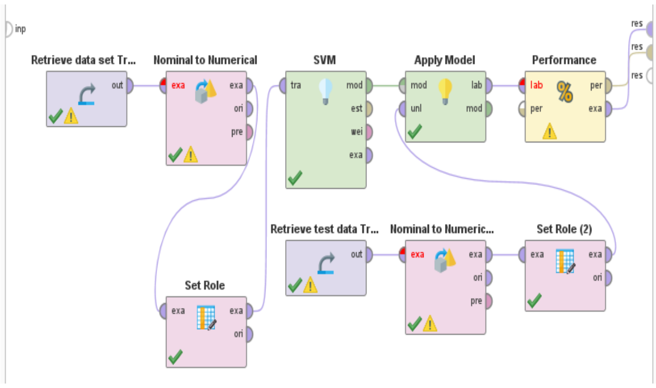


**Figure 3.** Process of the model

The data set is loaded first. The nominal to a numerical operator is then applied to convert its nominal attributes to numerical form. This step is necessary because the SVM operator cannot take nominal attributes; only numerical attributes can be used for classifying. In the next step, the set role operator is used to determine the label attribute. The model generated from the SVM operator is then applied to the 'Test Data'. Nominal to the numerical operator was applied to this data set as well. This is necessary because the testing and training data set should be in the same format, and label attribute is determined for the test data set. The statistical performance of this model is measured using the performance operator [35]. It is also possible to select some features to be ignored so that it can be observed how much the remaining features affect the performance [36].

## 3. Result and Discussion

In the result of model analysis, the kernel model with a total number of support vectors= 63, Bias= 3.580, Kernel Gamma= 0.00, and C= 100 has optimal performance. In the following chart, the true values and predictions for satisfaction in the SVM model can be seen.

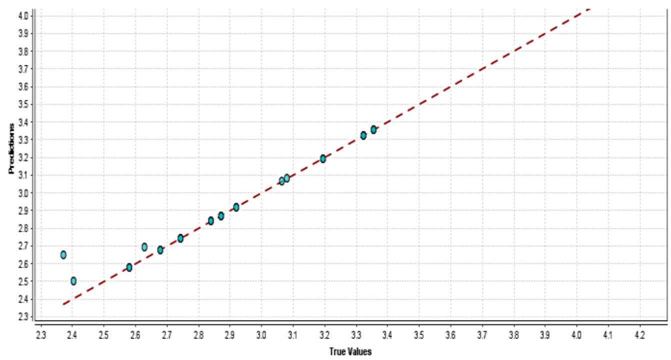


**Figure 4.** Predictions chart

Attributes with high weights in the SVM model, listed in Table 5.

**Table 5.** Important attributes in SVM model

| Attribute | Weight |
|---|---|
| Low price advantage | 0.328 |
| Advertisement | 0.316 |
| Doctors with high quality of education | 0.309 |
| Trained assistant staff | 0.289 |
| Relatives living in Turkey | 0.243 |
| High technology of medical equipment | 0.206 |

According to Table 4, factors that have an essential role in customers' decisions to choose Turkey as a treatment destination have been determined. However, the importance of socio-demographic variables is demonstrated in Table 6.

**Table 6.** Importance of socio-demographic variables

| Residence | Weight |
|---|---|
| Pakistan | 0.235 |
| Afghanistan | 0.164 |
| Libya | 0.147 |
| Kosovo | 0.106 |
| Iraq | 0.087 |
| **Nationality** | |
| Afghanistan | 0.182 |
| Libya | 0.133 |
| Pakistan | 0.114 |
| Iraq | 0.010 |
| **Age** | |
| 41-50 | 0.162 |
| ≥61 | 0.144 |
| **Disease** | |
| Dental | 0.134 |
| Diabetes | 0.122 |
| Gynecological | 0.029 |
| Orthopedic | 0.022 |
| **Gender** | |
| Male | 0.073 |
| **Marital status** | |
| Married | 0.064 |
| **Insurance** | |
| No | 0.039 |
| **Education level** | |
| Graduated | 0.032 |
| Undergraduated | 0.021 |
| **Income** | |
| 5000-10000$ | 0.0172 |

Patients that are residents of Pakistan, Afghanistan, Libya, Kosovo, and Iraq are among the most influential customers in satisfaction level. Also, patients with Afghanistan, Libya, Pakistan, and Iraq nationalities are more effective than others. Married men between 41-50 and above 61 age range with 5000-10000$ income have high weights in the model compared to others. Patients with graduate and undergraduate levels that don't have any insurance prefer a hospital in Ankara. Treating dental, diabetes, gynecological, and orthopedic diseases can influence the satisfaction level of international patients more.

Weaknesses in Turkey's health tourism and suggestions for solutions to these weaknesses can be summarized as follows:

- ✓ First of all, to eliminate the shortage of qualified personnel in the health sector in Turkey, it is necessary to increase the number of all health personnel, especially doctors and nurses, and to train sufficient personnel who know the required foreign languages and different cultures to work in health tourism.
- ✓ Considering the importance of advertisement in patient satisfaction, it is necessary to present practical introduction to Turkey's health sector.
- ✓ To ensure that more health tourism patients come from abroad, public and private hospitals should have up-to-date functional medical devices, tools, and equipment.
- ✓ Reliable statistical data is unavailable due to the lack of a health database in medical tourism. A "Health Tourism Database" should be established by the Turkish Statistical Institute, the Ministry of Health, and the Ministry of Culture and Tourism to access reliable data.
- ✓ In terms of developing the sector, cooperation between the Ministry of Health and the Ministry of Culture and Tourism could not be achieved enough, and both ministries carry out independent studies within the scope of their duties and responsibilities. Regular meetings should be provided to consider the health tourism platform, which was established between the Ministry of Health and the Ministry of Culture and Tourism with a participatory approach to encourage the participation of other relevant institutions and organizations.

## 4. Conclusion

As health tourism is a rapidly developing sector in the world, international standards, quality, and information sharing should be considered extraordinary. Also, customer-oriented services and branding should be attached to health tourism. The benefits of health tourism include the contribution of income from foreign tourists to economic prosperity, the increase of information sharing between countries and the development of strategic partnerships, the contribution of technology and knowledge transfer between countries, sharing of social and cultural experiences, developing global marketing in medical trade, providing competitive advantages due to international competition and following that, increasing customer satisfaction. Besides the benefits of health tourism, there are also negative aspects. Since some governments and health insurances do not pay for the health services received from abroad, patients pay for it themselves; thus partnerships should be established between insurance companies in countries of origin and destination. The patient usually returns to his country a few days after the operation and the needed postoperative care is met in the patient's own country; consequently coordinated and mutual information system should be established between hospitals.

Turkey has sufficient infrastructure and superstructure activities and social and cultural opportunities to increase medical tourism revenues. In addition, Turkey's climate, historical and cultural riches and natural beauties, modern health facilities, and professional staff improve the quality of health tourism. The number of accredited hospitals serving medical tourism is high, and the majority of health services are provided by private sector organizations. However, it does not get a sufficient share of health tourism income on the world scale. As a result of the research, this situation is evaluated that it is caused by incomplete, unplanned advertising and promotion activities and the low number of health personnel who speak foreign languages. In addition, it is crucial to ensure the coordination between the Ministry of Health, the Ministry of Culture and Tourism, the private health sector, and all stakeholders related to the private tourism sector, to strengthen the intermediary institutions and to see medical tourism as a country policy, in terms of making the contribution of the sector to the country's economy more effective. Although the Covid-19 pandemic had a negative impact on the industry in the world and Turkey, this process should be seen as an opportunity for countries to introduce themselves and take measures for the future. It is foreseen that this process may continue with different viruses in the future. Therefore, emergency action plans should be prepared for these private and state institutions and ready for implementation. Strategies that will provide international competition in health tourism should be determined. It is expected that tourists who realize health tourism mobility will prefer to receive services in accredited health institutions and health tourism institutions with a high awareness of brands are experts in their fields, and minimize risk factors.

**Research Highlights**

**What Is Already Known?**

Turkey, which has a high ranking among medical travel destinations, has some advantages that can become preferable for international patients.

**What Does This Study Add?**

This study is among the first few studies which examine affecting factors in patients' medical travel destination choices with Data Mining techniques. Effective attributes in patients' satisfaction level determined. Eventually, some suggestions were presented to improve the weaknesses in the health tourism sector.


## Authors' Contributions

SJJ: study design, search strategy, data analysis, preparing the draft

FE: study design, data collection, project supervision, revising and editing the manuscript.

All the authors read and approved the final manuscript.

## Conflict of Interest Disclosures

The authors have no conflicts of interest.

## Ethical Approval

Patient data are permitted to be used in a scientific study.

## Funding/ Support

None.


## References


1. Lee, M., Han, H., & Lockyer, T. (2012). Medical tourism—attracting Japanese tourists for medical tourism experience. Journal of Travel & Tourism Marketing, 29(1), 69-86. Doi: 10.1080/10548408.2012.638564.
2. Cham, T. H., Lim, Y. M., Aik, N. C., & Tay, A. G. M. (2016). Antecedents of hospital brand image and the relationships with medical tourists' behavioral intention. International Journal of Pharmaceutical and Healthcare Marketing. Vol. 10 No. 4, pp. 412-431. Doi: 10.1108/IJPHM-02-2016-0012.
3. Pan, T. J., & Chen, W. C. (2014). Chinese medical tourists–Their perceptions of Taiwan. Tourism Management, 44, 108-112. Doi: 10.1016/j.tourman.2014.02.008.
4. Loh, C. P. A. (2015). Trends and structural shifts in health tourism: Evidence from seasonal time-series data on health-related travel spending by Canada during 1970–2010. Social science & medicine, 132, 173-180. doi: 10.1016/j.socscimed.2015.03.036.
5. PBB (2018). Medical tourism statistics & facts, available from: https://www.patientsbeyondborders.com/ (accessed March 1, 2018).
6. TURSAB (2017). Association of Turkish Travel Agencies, available from: www.tursab.org.tr/dosya/12186/saglikturizmiraporu_12186_5485299.pdf (accessed December 25, 2017).
7. Aksu, A.A., Gürsoy, D. and Aksu, G.M. (2016), "Sağlık turizminde keşfedildikce büyüyen ülke Türkiye: Pazar analizi ve izlenmesi gereken stratejiler Türkiye turizm yatırımcıları derneği/The Turkish Tourism Association, İstanbul.
8. IMTJ (2018). International Medical Travel Journal via the Thaiger, available from: https://thethaiger.com/thai-life/thailand-third-biggest-medical-tourism-destination-in-the-world (accessed July 17, 2020).

9. Sag, I., & Zengul, F. D. (2019). Why medical tourists choose Turkey as a medical tourism destination. Journal of Hospitality and Tourism Insights, 2(3). Doi: 10.1108/JHTI-05-2018-0031.

10. Altın, U., Bektaş, G., Antep, Z., & İrban, A. (2012). The international patient's portfolio and marketing of Turkish health tourism. Procedia-Social and Behavioral Sciences, 58, 1004-1007. Doi: 10.1016/j.sbspro.2012.09.1080.

11. TUIK (2020), Turkish Statistical Institute, available from: https://biruni.tuik.gov.tr/medas/?kn=74&locale=tr (accessed September 5, 2020).

12. Aggarwal, C. C., & Zhai, C. (2012). A survey of text clustering algorithms mining text data (pp. 77-128): Springer. Doi: 10.1007/978-1-4614-3223-4_4.

13. Ersoz, F. (2019). Data mining and text mining with big data: review of differences, International Journal of Recent Advances in Multidisciplinary Research, Vol. 06, Issue 01, pp.4391-4396. Available from: https://www.researchgate.net/publication/354960220_Data_Mining_and_Text_Mining_with_Big_Data_Review_of_Differences.

14. Cham, T. H., Lim, Y. M., Sia, B. C., Cheah, J. H., & Ting, H. (2020). Medical tourism destination image and its relationship with the intention to revisit: A study of Chinese medical tourists in Malaysia. Journal of China Tourism Research, 1-29. Doi: 10.1080/19388160.2020.1734514.

15. Moghadam, F. N., Masoudi ASL, I., Hessam, S., & Farahani, M. M. (2020). In search a medical tourism marketing pattern in Iran: The case of cultural sensitivities. International Journal of Healthcare Management, 1-6. Doi: 10.1080/20479700.2020.1732647.

16. Malik, G. (2019). The creation of a medical tourist destination: Exploring determinants of perception. International Journal of Healthcare Management, 1-11. Doi: 10.1080/20479700.2019.1677023.

17. Habibi, A., & Ariffin, A. A. M. (2019). Value as a medical tourism driver interacted by experience quality. Anatolia, 30(1), 35-46. Doi: 10.1080/13032917.2018.1496122.

18. Esiyok, B., Çakar, M., & Kurtulmuşoğlu, F. B. (2017). The effect of cultural distance on medical tourism. Journal of Destination Marketing & Management, 6(1), 66-75. Doi: 10.1016/j.jdmm.2016.03.001.

19. Aydın, G., & Karamehmet, B. (2017). Factors affecting health tourism and international health-care facility choice. International Journal of Pharmaceutical and Healthcare Marketing. Doi: 10.1016/j.jdmm.2016.03.001.

20. Ganguli, S., & Ebrahim, A. H. (2017). A qualitative analysis of Singapore's medical tourism competitiveness. Tourism Management Perspectives, 21, 74-84. Doi: 10.1016/j.tmp.2016.12.002.

21. Lubowiecki-Vikuk, A., & Kurkowiak, J. (2017). Medical tourism potential of Central and Eastern Europe: Attempt at classification. Paper presented at the CBU International Conference Proceedings. Doi: 10.12955/cbup.v5.940.

22. Ulaş, D., & Anadol, Y. (2016). A case study for medical tourism: Investigating a private hospital venture in Turkey. Anatolia, 27(3), 327-338. Doi: 10.1080/13032917.2016.1191763.

23. Han, H., & Hyun, S. S. (2015). Customer retention in the medical tourism industry: Impact of quality, satisfaction, trust, and price reasonableness. Tourism Management, 46, 20-29. Doi: 10.1016/j.tourman.2014.06.003.

24. Manaf, N. H. A., Hussin, H., Kassim, P. N. J., Alavi, R., & Dahari, Z. (2015). Medical tourism service quality: finally, some empirical findings. Total Quality Management & Business Excellence, 26(9-10), 1017-1028. Doi: 10.1080/14783363.2015.1068597.

25. Guiry, M., & Vequist IV, D. G. (2015). South Korea's medical tourism destination brand personality and the influence of personal values. Asia Pacific Journal of Tourism Research, 20(5), 563-584. Doi: 10.1080/10941665.2014.904804.

26. Ahangaran J, Bhartiya SP, Nadalipour Z, Salehi M. Determining contributory factors to the competitiveness of Iran’s medical tourism: an importance-performance analysis. Int J Travel Med Glob Health. 2022; 10(1):32-39. Doi: 10.34172/IJTMGH.2022.06.

27. Aminmansour B, Mahmoodkhani M, Shafiei M, Mokhtari A, Hematzadeh M, Sheibani-Tehrani D. (2022). Medical tourism challenges after the prevalence of Covid-19: the neurosurgery field. Int J Travel Med Glob Health. 10(2):83-87. Doi: 10.34172/ijtmgh.2022.15.

28. Kondori, M. (2018). A data mining algorithm for determination of influential factors on the hospitalization of patients subject to chronic obstructive pulmonary disease. Archives of Medical Laboratory Sciences, 4(2). Doi: 10.22037/amls.v4i2.25568.

29. Senturk, Z. (2020). Artificial neural networks-based decision support system for the detection of diabetic retinopathy. Sakarya Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 24(2), 424-431. Doi: 10.16984/saufenbilder.630482.

30. Massaro, A., Maritati, V., & Galiano, A. (2018). Data mining model performance of sales predictive algorithms based on RapidMiner workflows. International Journal of Computer Science & Information Technology (IJCSIT), 10(3), 39-56.doi: 10.5121/ijcsit.2018.10303.

31. Naik, A., & Samant, L. (2016). Correlation review of classification algorithm using data mining tool: WEKA, Rapidminer, Tanagra, Orange and Knime. Procedia Computer Science, 85, 662-668. Doi: 10.1016/j.procs.2016.05.251.

32. Du, Z., Yang, Q., He, H., Qiu, M., Chen, Z., Hu, Q. & Huang, Y. (2019). A comprehensive health classification model based on support vector machine for proseal laryngeal mask and tracheal catheter assessment in herniorrhaphy. Mathematical Biosciences and Engineering: MBE, 17(2), 1838-1854. doi: 10.3934/mbe.2020097.

33. Chen, Y. (2020). Mining of instant messaging data in the Internet of Things based on support vector machine. Computer Communications. Volume 154, pages 278-287. Doi: 10.1016/j.comcom.2020.02.080.

34. Khaledian, M. R., Isazadeh, M., Biazar, S. M., & Pham, Q. B. (2020). Simulating Caspian Sea surface water level by artificial neural network and support vector machine models. Acta Geophysica, 1-11. doi: 10.1007/s11600-020-00419-y.

35. Rapidminer (2020), Rapidminer documentation, available from: https://docs.rapidminer.com/latest/studio/operators/modeling/predictive/support_vector_machines/support_vector_machine.html (accessed July 16, 2020).

36. Ersöz, Filiz (2019). Veri madenciliği teknikleri ve uygulamaları, Seçkin yayınevi, Ankara. Available from: https://www.seckin.com.tr/kitap/574338851.