Application of data mining to extract knowledge about the occurrence of fistulas after palatoplasty

Aplicação de mineração de dados para extração de conhecimento sobre ocorrência de fístulas após palatoplastia

PATRICK PEDREIRA SILVA[1*]
ELVIO GILBERTO DA SILVA[2]
VINICIUS SANTOS ANDRADE[1]
TELMA VIDOTTO DE SOUSA BROSCO[3]
GABRIELA APARECIDA PREARO[4]
MARIA INÊS PEGORARO-KROOK[1, 4]
JENIFFER DE CASSIA RILLO DUTKA[1, 4]

Institution: Hospital de Reabilitação de Anomalias Craniofaciais da Universidade de São Paulo, Bauru, SP, Brazil.

Article received: November 3, 2021.
Article accepted: September 13, 2022.

Conflicts of interest: none.

DOI: 10.5935/2177-1235.2023RBCP657-EN

■ **ABSTRACT**

Introduction: Data mining techniques expand access to important information for the decision-making process during health care. The objective the study proposes using data mining techniques to identify variables (surgical treatment protocols, patient characteristics, post-surgical complications) associated with fistulas after primary palatoplasty in patients with unilateral transforamen incisor cleft (UTIC). **Method:** A data set of 222 patients with UTIC without syndromes, operated by four surgeons with Furlow's or von Langenbeck's primary palatoplasty techniques, was analyzed for this study. Two models for detecting the outcome of surgery were induced using data mining techniques (Decision Tree and Apriori). **Results:** Five rules were selected from a decision tree pointing to some variables as predictors of fistulas associated with primary palatoplasty: infection, cough, hypernasality, and surgeon. Analysis of the model indicates that it correctly classifies 95.9% of occurrences between the absence and presence of fistulas. The second model indicates that the absence of post-surgical complications (infection and fever) and normal speech results (absent hypernasality, without suggestive of velopharyngeal dysfunction) are related to the absence of fistulas. Regarding surgical procedures, the Furlow technique and the Vomer flap were more frequent in patients with fistulas. **Conclusion:** Data mining techniques, as applied in the present study, pointed to infection and cough, hypernasality, and surgeon and surgical techniques as predictors of fistulas related to primary palatoplasty.

Keywords: Data mining; Health; Cleft palate; Oral fistula; Algorithms.

■ **RESUMO**

Introdução: As técnicas de mineração de dados ampliam o acesso a informações importantes para o processo de tomada de decisão durante os cuidados com a saúde. O objetivo do estudo propõe a utilização de técnicas de mineração de dados para identificar variáveis (protocolos de tratamento cirúrgico, características do paciente, intercorrências pós-cirúrgicas) associadas à ocorrência de fístulas após palatoplastia primária em pacientes com fissura transforame incisivo unilateral (FTIU). **Método:** Um conjunto de dados de 222 pacientes com FTIU sem síndromes, operados por quatro cirurgiões com as técnicas de palatoplastia primária de Furlow ou von Langenbeck, foi analisado para este estudo. Dois modelos para detecção do resultado da cirurgia foram induzidos usando técnicas de mineração de dados (Árvore de Decisão e Apriori). **Resultados:** Cinco regras

[1] Hospital de Reabilitação de Anomalias Craniofaciais, Universidade de São Paulo, Programa de Pós-Graduação em Ciências da Reabilitação, Bauru, SP, Brazil.
[2] Hospital de Reabilitação de Anomalias Craniofaciais, Universidade de São Paulo, Programa de Pós-Doutorado, Bauru, SP, Brazil.
[3] Hospital de Reabilitação de Anomalias Craniofaciais, Universidade de São Paulo, Departamento de Cirurgia Plástica, Bauru, SP, Brazil.
[4] Faculdade de Odontologia de Bauru, Universidade de São Paulo, Programa de Pós-Graduação em Fonoaudiologia, Bauru, SP, Brazil.

foram selecionadas de uma árvore de decisão apontando para algumas variáveis como preditivas de fístulas associadas à palatoplastia primária: infecção, tosse, hipernasalidade, cirurgião. A análise do modelo indica que ele classifica corretamente 95,9% das ocorrências entre ausência e presença de fístulas. O segundo modelo indica que a ausência de intercorrências pós-cirúrgicas (infecção e febre) e resultado de fala normal (hipernasalidade ausente, sem sugestivo de disfunção velofaríngea) estão relacionados à ausência de fístulas. Em relação aos procedimentos cirúrgicos, o uso da técnica de Furlow e retalho de Vomer foram mais frequentes nos pacientes com fístulas. **Conclusão:** Técnicas de mineração de dados, conforme aplicadas no presente estudo, apontaram para infecção e tosse, presença de hipernasalidade, cirurgião e técnica cirúrgica como preditores de fístulas relacionadas à palatoplastia primária.

Descritores: Mineração de dados; Saúde; Fissura palatina; Fístula bucal; Algoritmos.

INTRODUCTION

One of the main objectives of primary surgery of the palate in cleft lip and palate (CLP) is the successful reconstruction of the levator muscle belt to provide a functional velopharyngeal mechanism for adequate speech production[1]. The presence of an oronasal fistula is one of the most significant complications after surgical repair of the palate since its implications can interfere with the individual's quality of life. The incidence of residual oronasal fistulas is one factor that indicates the success of primary surgical repair of the palate[2-4].

A fistula, as reported by Brosco[4,5], is a failure in healing or rupture of the primary surgical palate repair that can occur anywhere along the cleft closure line. The literature presents conflicting data regarding the occurrence of fistula[6-9]; for example, Salimi et al.[10] reported an incidence ranging between 0 and 78%. It is important to understand which variables (surgical treatment protocols, postoperative complications, speech results after surgery, and patient characteristics) are associated with fistulas to prevent and minimize these surgical complications.

The so-called "information age" is characterized by the growing expansion in the volume of data generated and stored, a phenomenon also reflected in the health area, which increases the possibility of obtaining important information supporting the decision-making process[11]. The patients' data and the surgery results are available in their medical records and can be used for clinical studies.

However, many times, the volume of data generated is so large that its use and manual analysis are difficult, demanding more sophisticated processes, such as, for example, automated processes, for the manipulation of such data. In this context of the overabundance of data, data mining emerged as a systematic, interactive, and iterative process of preparing and automatically extracting knowledge from databases[11,12].

In data mining, hypotheses are induced from a set of observed data, such as, for example, data on patients. Each patient is called an object, and different attributes are stored on each object (name, identification, gender, age, symptoms, etc.), which correspond to the different data of that patient. In one of the typical mining tasks, one seeks to learn ways to predict one of the attributes (this specific attribute of which one wants to make the prediction is called class or, simply, target attribute or output attribute). The other attributes that predict the target attribute are called predictors or input attributes.

From a set of data, we seek to create a model or hypothesis (represented by an algorithm or set of rules) capable of relating one or more attributes (predictors) to the target attribute (class). Through an inductive bias, each model identified from data mining uses a representation to describe the hypothesis induced from the data set.

OBJECTIVE

This work aims to use data mining techniques to automatically extract knowledge about variables (surgical treatment protocols, postoperative complications, speech results after surgery, and patient characteristics) associated with the occurrence of oronasal fistulas in patients with unilateral transforamen incisor cleft (UTIC).

METHOD

The investigation deals with descriptive, quantitative, experimental, and applied research, approved by the Research Ethics Committee of the Hospital for Rehabilitation of Craniofacial Anomalies of the University of São Paulo (opinion 1,753,467), carried out at that institution in September 2021. sample refers to a subset of medical records of patients with cleft lip and palate participating in a randomized clinical trial (RCT) with UTIC[13].

Data on the occurrence of fistulas were obtained for a total of 466 patients (infants). These patients were randomized (using a script–code written in a programming language developed at the University of Florida) to receive different surgical treatment protocols, including 1) primary cheiloplasty between 3 and 6 months of age with the Millard technique (M) or Spina (S); 2) early (9 to 12 months) or late (>12 months) palatoplasty; 3) primary palatoplasty with the von Langenbeck (VL) or Furlow (F) technique; and 4) to one of four possible surgeons (C1, C2, C3, C4).

Information about the occurrence of fistula after primary palatoplasty was of interest to the present study. To determine classes in data mining, Spina's[14] classification was used, grouping patients into two groups: SUCCESS (patients without fistula or with fistula in the pre-incisive foramen region); FAILURE (patients with fistulas in the post-incisive foramen region or transforamen fistulas). The incisive foramen marks the limits of the primary palate (central part of the upper lip and premaxilla) and secondary palate (hard and soft palate).

Table 1 presents the information identified in the patient's records for this study. The definition of the variables of interest is directly related to some factors, including 1) Surgical treatment protocols (a surgical technique in cheiloplasty and palatoplasty, surgeon, use of surgery modifications such as relaxing incision and vomer flap, duration of palatoplasty in minutes); 2) Patient characteristics (age at palatoplasty, duration of palatoplasty); 3) Post-surgical complications (whether there was an infection in the palatoplasty - at the site - or elsewhere after the primary palatoplasty; whether there was vomiting or coughing in the postoperative period of the palatoplasty); 4) Speech results after surgery (whether there was a symptomatic diagnosis of velopharyngeal dysfunction, presence of hypernasality - recorded in spontaneous or directed conversation); results of nasal air emission, hypernasality, and hyponasality tests (observed during the repetition of 10 words); 5) Result of the surgery regarding the occurrence of fistula (SUCCESS or FAILURE). The variables of interest are listed in the "attribute name" column.

Variables associated with surgical treatment protocols and patient characteristics can help indicate whether there is a greater propensity to develop fistulas (even before the surgical procedure). In contrast, variables related to post-surgical complications and speech results can be indicative of clinically relevant fistulas after surgery.

In the management of CLP, the SUCCESS of the treatment occurs in the absence of a fistula and absence of speech disorders. For the present study, a fistula in the region posterior to the incisive foramen and the presence of velopharyngeal dysfunction were interpreted as indicative of treatment FAILURE. The guiding question for data mining involved checking which factors (surgical treatment protocols, patient characteristics, post-surgical complications, and speech results after surgery) would be associated with the occurrence or not of fistulas. Therefore, this study aimed to identify whether some of the analyzed variables could be used as predictors of the occurrence of fistulas on the palate or as indicators of clinically relevant fistulas after palatoplasty.

To compute the results of the experiment, the algorithm C4.5 (J48) was used, which generates decision trees to find the relationship between the characteristics considered and the results of the surgeries, as well as the "Apriori algorithm" (association) for the rule generation. Decision trees allow variables or attributes to be categorical (qualitative) or numerical (quantitative). It can be used simultaneously by the model (which proved adequate considering the different types of

Table 1. Definition of variables (attributes) of interest for this study.

Variables (Attribute Name)	Categories (Values)
Surgical technique in cheiloplasty	Millard, Spina
Palatoplasty time	Early (9-12m), Late (>12m)
Age at palatoplasty	months (m)
Surgical technique in palatoplasty	Furlow, von Langenbeck
Surgeon	C1, C2, C3, C4
Relaxing incision	No incision, unilateral, bilateral
Vomer flap	yes, no
Duration of palatoplasty	Minutes
Infection in palatoplasty surgery, at another location	There was not; at the site of
Postoperative vomiting Palatoplasty	yes, no
Cough after surgery of palatoplasty	yes, no
Fever	yes, no
Suggestive of velopharyngeal dysfunction	yes, no
Hypernasality	yes, no
Air emission test	[1-10]
Hypernasality test	[1-10]
Hyponasality test	[1-10]
*Occurrence of fistula	SUCCESS, FAILURE

m=months; C=surgeon; *target attribute (class) variables in the database used in the investigation). The Apriori algorithm deals only with qualitative variables. Both models induce a hypothesis through a model represented by rules ("if...then").

In this analysis, the variables of interest were treated as attributes in the WEKA software. Considering a typical mining task, the experiment was divided into four stages: data pre-processing, feature extraction, classification, and description of results. The procedure was performed considering the occurrence of a fistula after palatoplasty as the primary result. Pre-processing was carried out semi-automatically. Data from the medical records made available in the ".XLS" file format (Excel® spreadsheet) were converted to the ".ARFF" format (used by WEKA) using the Excel2ArffConverter software. Before conversion, the attributes were identified as described in Table 1.

RESULTS

Only patients with complete data were selected for analysis, considering the parameters described in Table 1. After discarding patients with incomplete data for any variables, 222 patients were selected for analysis. Due to the possibility of bias in the base, it was decided not to estimate the missing values[7]. Information on the occurrence of some type of fistula was identified in the medical records of 222 (47.6%) of the 466 patients studied, and data from these patients were mined for this article.

Of the 222 patients considered for this study, 98 (44.1%) were female, and 124 (55.9%) were male. The mean age at primary palatoplasty was 12.8 months (σ=3.2). In this group, 114 (51.3%) received the Millard procedure in primary cheiloplasty, while 108 (48.7%) received the Spina procedure. One hundred twelve patients (50.4%) received the Furlow technique in primary palatoplasty, while 110 (49.6%) received von Langenbeck. Of the patient sample, 182 (81.9%) belonged to the SUCCESS group and 40 (18.1%) to the FAILURE group.

Through constructing a decision tree, 37 rules were generated from the complete patient data set. However, in this article, we chose to display only the 5 rules with the highest value for the coverage metric of each final result of the surgery (SUCCESS or FAILURE). The coverage metric is the ratio of correctly classified data to the total sample data for that class. Information about the rule's accuracy metric (probability of the final result conditional on the attributes, i.e., the model's ability to avoid false positives) was also considered.

The mean accuracy of the rules associated with surgical SUCCESS is 97.26% (σ=4.59). The five rules together present coverage of about 77.5%; that is, if applied to the data, they manage to detect 77.5% of the cases of SUCCESS. As for the FAULT class, the average accuracy of the associated rules is 84.32% (σ=9.40). The coverage of the five rules is 62.5%, that is, the number of FAILURE cases that the rules can detect if applied to the database.

The rule with greater coverage and accuracy for predicting a good result indicates that the main factors involved are: infection ("absence"), hypernasality tests ("≤6") and hyponasality (">9"), and the surgical technique ("von Langenbeck"). The interpretation of this rule indicates, therefore, that patients submitted to the "von Langenbeck" procedure, without infection and with hyponasality test results with values greater than or equal to 9 and hypernasality test with values less than or equal to 6 are more likely to have obtained SUCCESS as the final result of the surgery. As for FAILURE, according to the two rules with greater precision and coverage, the factors involved are related to post-surgical complications and speech results and include infection ("absence or elsewhere"), hypernasality tests ("greater than 6"), air emission ("greater than 9") and fever ("yes"). The rules are shown in Table 2.

When analyzing the global performance of the model (generated decision tree) concerning its predictive capacity, it is observed that it correctly classifies 95.9% of the patients and incorrectly only 4.1%. Considering each category individually, the model manages to hit 90.0% of the cases in which a FAILURE result occurs. As for the other class, the model manages to hit 97.3% of the cases in which a SUCCESS result occurs.

The correlations found using the Apriori algorithm were obtained using the support (minimum of 60%) and confidence (minimum of 90%) metrics. The objective was to find frequent (high support value) rules in the database with a high degree of confidence (directly related to rule validity). Four rules were found with an average confidence of 90.75% (σ=0.5) and an average support of 69.45% (σ=0.49), which meet the requirements above, as shown in Table 3.

Considering only the 40 patients in the FAILURE group, the results show the six rules found with a minimum support of 67.5% and a minimum confidence of 100% (Table 3). The rules have average support of 72.08% for this group. In the SUCESSO group, the model indicates the absence of post-surgical intercurrences (cough and infection) and speech results with absent hypernasality. Patients in the FAILURE group also had no cough and no infection.

Table 4 summarizes the relationship between the duration of the palatoplasty and the result regarding the occurrence of fistulas. It is observed that surgery times vary from 25 to 140 minutes.

The algorithms allow data-based exploration of non-linear relationships and interactions between

Table 2. Surgery result.

Number	Rule	Result (class)	Coverage	Precision
1	If "infection=none" and "hypernasality test≤6" and "cough=no" and "surgical technique=von Langenbeck" and "hyponasality test>9"	SUCCESS	77	100%
2	If "infection=none" and "test hypernasality≤6" and "cough=no" and "surgical technique=Furlow" and "fissure width=regular"	SUCCESS	33	96.9%
3	If "infection=none" and "test in hypernasality≤6" and "cough=no" and "surgical technique=Furlow" and "cleft width=wide" and "surgeon=C3"	SUCCESS	19	89.4%
4	If "infection=none" and "test in hypernasality≤6" and "cough=no" and "surgical technique=Furlow" and "fissure width=wide" and "surgeon=C2"	SUCCESS	7	100%
5	If "infection=none" and "test in hypernasality≤6" and "cough=no" and "surgical technique=Furlow" and "fissure width=wide" and "surgeon=C1" and "relaxing incision=no"	SUCCESS	5	100%
6	If "infection=none" and "hypernasality test>6" and "air emission test>9" and "fever=yes"	FAILURE	6	83.3%
7	If "infection=occurred elsewhere"	FAILURE	6	83.3%
8	If "infection=none" and "hypernasality test>6" and "air emission test>9" and "fever=no" and "relaxing incision=bilateral" and "vomit=no" and "surgeon=C3"	FAILURE	5	80.0%
9	If "infection=none" and "test hypernasality≤6" and "cough=no" and "surgical technique=Furlow" and "fissure width=wide" and "surgeon=C4" and "air emission test>2"	FAILURE	4	100%
10	If "infection=none" and "test in hypernasality>6" and "air emission test>9" and "fever=no" and "relaxing incision=no"	FAILURE	4	75.0%

Table 3. Rules with high support and confidence values.

Characteristics	Result	Support	Confidence
Absence of cough and infection without suggestive of dysfunction velopharyngeal	SUCCESS	69.8%	91.0%
Absence of cough and infection with absent hypernasality	SUCCESS	69.8%	91.0%
Absence of cough and infection without suggestive of dysfunction velopharyngeal with absent hypernasality	SUCCESS	69.8%	91.0%
Absence of cough and infection and no fever	SUCCESS	68.4%	90.0%
absence of cough	FAILURE	77.5%	100.0%
absence of infection	FAILURE	77.5%	100.0%
Furlow's surgical technique	FAILURE	72.5%	100.0%
Use of vomer flap	FAILURE	70.0%	100.0%
absence of vomiting	FAILURE	67.5%	100.0%
Absence of cough and infection	FAILURE	67.5%	100.0%

Table 4. Relationship between duration of palatoplasty and classes (SUCCESS and FAILURE).

Duration: Minutes	No	Average	Standard deviation	Minimum	Maximum
Duration of palatoplasty – All groups	222	65.62	24.43	25	140
Duration of palatoplasty (group SUCCESS)	182	62.57	22.89	25	125
Duration of palatoplasty (group FAILURE)	40	79.5	26.62	25	140

many variables, generating easy interpretation models. However, as a weakness of the method, the unbalance between the two groups (SUCCESS and FAILURE) and the full use of the sample for the induction of the models can be pointed out, which can cause overfitting of the data, impairing the extrapolation of the findings (rules) to other databases.

DISCUSSION

Specifically, concerning fistulas, the rules found with a high degree of precision and coverage can show useful standards on which variables, among surgical treatment protocols, patient characteristics, speech results after surgery, and post-surgical intercurrences, are determinant for the success or failure of the palatoplasty. The opportunity to adopt data mining on patients undergoing palatoplasty can provide a better understanding of the specificities that may occur with the group of patients, thus expanding the professional's knowledge in identifying the conduct to be adopted.

In this specific study, the visibility given to some factors (Table 1) allows health professionals to identify patterns of association of variables, with the proper analysis of this set of discoveries, which can give meaning to diagnostic and therapeutic actions. In the same way, as in other previous studies, this investigation opted for combining different types of data mining tasks to carry out the experiment or identify patterns[15-19].

Despite the initial availability of data referring to 466 patients, we chose to use 222 (considering only the complete ones). This may have limited the rules obtained and not have evidenced other associations of the factors related to the final results of the palatoplasty. This decision follows the guidelines of other works[20]. The entire database can be used for future studies, as some algorithms can deal with missing data[11].

Another limitation related to the base is the fact that the two classes considered are unbalanced; however, as they reflect the real situation in which SUCCESS results are more common than FAILURES, it was decided to maintain the natural proportion of the data. This presence of majority classes much more frequently than other minority classes makes algorithms respond well to majority classes to the detriment of minority ones. In future works, the experiment can be repeated using random resampling techniques of the data in order to generate balanced sets[21].

The fact that the entire database was used for induction and testing of the model may generate a bias to fit the data. Any mining method is subject to generating a model that overfits itself to the data on which it was induced (overfitting) but cannot generalize the learned knowledge, not obtaining a good performance when confronted with data from another base. However, this approach was chosen as the purpose of this experiment is not to induce a model to automate the classification process of surgeries but rather to extract rules that can be evaluated by humans, evidencing useful patterns.

The analysis of Table 2 indicates that the SUCCESS results are associated with post-surgical complications such as the absence of infection and cough; in addition, the patients presented a hypernasality test below or equal to 6 (on a scale that goes up to 10). In the case of large fissures associated with the Furlow surgical technique, in addition to the complications highlighted, the surgeon's factor influences the final result.

In the case of the FAILURE results, the presence of infection seems to be an important factor; however, it is not decisive. Due to the similarity between rules 9, 3, and 4 (Table 2), the decisive factor for obtaining a FAILURE result is linked to the surgeon. Under the same conditions, surgeons C2 and C3 obtained SUCCESSFUL results; however, surgeon C4 obtained SUCCESS in only 50% of the surgeries, which may indicate the influence of the surgeon factor. In the case of speech results, values of hypernasality tests greater than 6 are indicative of a possible FAILURE.

In the same way as the rules of the decision tree, the rules presented by the model induced by the Apriori algorithm must be evaluated by a professional to validate them against reality. The Apriori algorithm does not deal with quantitative attributes, only with categorical ones, which requires excluding some attributes or even their transformation to non-numerical data (discretization process); this strategy was used in some processing carried out in this work. Thus, to avoid this limitation in future work, other algorithms may be experimented with, such as AprioriTid, SETM, and AprioriHybrid[22].

The analysis of Table 3 indicates that, in general, the absence of post-surgical complications (infection and fever) and speech results with absent hypernasality, as well as patients without suggestive of velopharyngeal dysfunction, present SUCCESS after primary palatoplasty. Concerning surgical procedures, there are indications that the Furlow technique and the Vomer flap are frequent in the FAILURE group. Observations such as the absence of cough, vomiting, or infection alone cannot be used as parameters to rule out a possible FAILURE.

The analysis of Table 4 shows that a palatoplasty in the group of patients who had a result of FAILURE lasts an average of 79.5 minutes; for the group of patients with SUCCESS results, the average drops to 62.57 minutes. There are indications, therefore, that longer surgeries tend to cause worse results.

Finally, it is recognized that this study offers only a punctual perspective of reality through the analysis of models induced by data mining techniques in the considered database since it reveals only a few factors associated with the results of palatoplasty from the point of view of the mining algorithms, with the need for validation by health professionals.

CONCLUSION

Data analysis revealed that the absence of some post-surgical complications (fever, cough, infection) together with speech results after surgery (hypernasality, suggestive of velopharyngeal dysfunction) and with characteristics associated with surgical treatment protocols (technique, the flap of the vomer, surgeon) could help to predict the success or failure of the palatoplasty.

COLLABORATIONS

PPS Analysis and/or data interpretation, Data Curation, Formal Analysis, Investigation, Writing - Original Draft Preparation.

EGS Analysis and/or data interpretation, Data Curation, Methodology, Writing - Original Draft Preparation.

VSA Data Curation, Final manuscript approval, Writing - Original Draft Preparation, Writing - Review & Editing.

TVSB Final manuscript approval, Visualization, Writing - Original Draft Preparation, Writing - Review & Editing.

GAP Resources, Writing - Original Draft Preparation, Writing - Review & Editing.

MIPK Visualization, Writing - Original Draft Preparation, Writing - Review & Editing.

JCRD Final manuscript approval, Project Administration, Supervision, Writing - Review & Editing.

REFERENCES

1. Dutka JCR, Pegoraro-Krook MI. Avaliação e tratamento das disfunções velofaríngeas. In: Marchesan IQ, Silva HJ, Tomé MC, eds. Tratado das Especialidades em Fonoaudiologia. São Paulo: Grupo Gen-Editora Roca; 2014. p. 363-8.
2. de Agostino Biella Passos V, de Carvalho Carrara CF, da Silva Dalben G, Costa B, Gomide MR. Prevalence, cause, and location of palatal fistula in operated complete unilateral cleft lip and palate: retrospective study. Cleft Palate Craniofac J. 2014;51(2):158-64.
3. Deshpande GS, Campbell A, Jagtap R, Restrepo C, Dobie H, Keenan HT, et al. Early Complications After Cleft Palate Repair. J Craniofac Surg. 2014;25(5):1614-8.
4. Brosco TVS. Fístula de palato após reparo da fissura labiopalatina em um estudo clínico randomizado [Tese]. Bauru: Hospital de Reabilitação de Anomalias Craniofaciais, Universidade de São Paulo; 2017. 168 p.
5. Brosco TVS, Prearo GA, Silva HLA, Dutka JCR. Brosco-Dutka classification system for palate fistulas. Rev Bras Cir Plást. 2021;36(2):164-72.
6. Bykowski MR, Naran S, Winger DG, Losee JE. The Rate of Oronasal Fistula Following Primary Cleft Palate Surgery: A Meta-Analysis. Cleft Palate Craniofac J. 2015;52(4):e81-7.
7. Hardwicke JT, Landini G, Richard BM. Fistula incidence after primary cleft palate repair: a systematic review of the literature. Plast Reconstr Surg. 2014;134(4):618e-27e.
8. Timbang MR, Gharb BB, Rampazzo A, Papay F, Zins J, Doumit G. A systematic review comparing Furlow double-opposing Z-plasty and straight-line intravelar veloplasty methods of cleft palate repair. Plast Reconstr Surg. 2014;134(5):1014-22.
9. Park MS, Seo HJ, Bae YC. Incidence of fistula after primary cleft palate repair: a 25-year assessment of one surgeon's experience. Arch Plast Surg. 2022;49(1):43-9.
10. Salimi N, Aleksejūnienė J, Yen EH, Loo AY. Fistula in Cleft Lip and Palate Patients-A Systematic Scoping Review. Ann Plast Surg. 2017;78(1):91-102.
11. Castro LN, Ferrari DG. Introdução à mineração de dados: Conceitos básicos, algoritmos e aplicações. São Paulo: Saraiva; 2017.
12. Netto AV, Berton L, Takahata AK. Ciência de Dados e a Inteligência Artificial na Área da Saúde. São Paulo: Editora dos Editores; 2021.
13. Dutka JCR. Estudo clínico randomizado - Projeto Florida (ECR-PF- ementa fase 2): função velofaríngea para a fala e estudo do crescimento da face e dos arcos dentários após a palatoplastia primária. Projeto em andamento e com aprovação Ética do CEP/CONEP desde 02/09/2016. Pesquisador Responsável: Jeniffer de Cássia Rillo Dutka. CAAE: 57727416.9.0000.5441. Instituição Proponente: Hospital de Reabilitação de Anomalias Craniofaciais da USP.
14. Spina V, Psillakis JM, Lapa FS, Ferreira MC. Classificação das fissuras lábio-palatais: sugestão de modificação. Rev Hosp Clín Fac Med (São Paulo). 1972;27:5-6.
15. West D, Mangiameli P, Rampal R, West V. Ensemble strategies for a medical diagnostic decision support system: A breast cancer diagnosis application. Eur J Oper Res. 2005;162(2):532-51.
16. Steiner MTA, Soma NY, Shimizu T, Nievola JC, Steiner Neto PJ. Abordagem de um problema médico por meio do processo de KDD com ênfase à análise exploratória dos dados. Gest Prod. 2006;13(2):325-37.
17. Bodini Junior AC. Utilização de técnicas de data mining na detecção de outliers em auxílio à auditoria operacional com um estudo de caso com dados do sistema de informações hospitalares [Tese]. Rio de Janeiro: Universidade Federal do Rio de Janeiro; 2009.
18. Kuretzki CH. Técnicas de mineração de dados aplicadas em bases de dados da saúde a partir de protocolos eletrônicos [Dissertação]. Curitiba: Universidade Federal do Paraná; 2009. 98 p.

19. Salarte Martínez GR, Castro Bermúdez YV. Modelo híbrido para el diagnóstico de enfermedades cardiovasculares basado en inteligencia artificial. Tecnura. 2012;16(33):35-52.
20. Carvalho DR, Moser AD, Silva VA, Dallagassa MR. Mineração de dados aplicada à fisioterapia. Fisioter Mov. 2012;(3):595-605.
21. Goldschmidt R, Passos E, Bezerra E. Data Mining: Conceitos, técnicas, algoritmos, orientações e aplicações. 2ª ed. Rio de Janeiro: Elsevier; 2015.
22. Agrawal R, Imieliński T, Swami A. Mining association rules between sets of items in large databases. ACM SIGMOD Rec. 1993;22(2):207-16.

***Corresponding author: Patrick Pedreira Silva**
R. Silvio Marchione, 3-20, Vila Nova, Cidade Universitaria, Bauru, SP, Brazil.
Zip code: 17012-900
E-mail: patrickpsilva@alumni.usp.br