Universitat Internacional de Catalunya
Facultat de Medicina i Ciències de la Salut -
Departament de Medicina

Tesi doctoral

VOICE ANALYSIS AS A METHOD FOR PREOPERATIVELY PREDICTING A DIFFICULT AIRWAY BASED ON MACHINE LEARNING ALGORITHMS.

Claudia Rodiera Clarens

DOCTORAL THESIS

Title:

VOICE ANALYSIS AS A METHOD FOR PREOPERATIVELY PREDICTING A DIFFICULT AIRWAY BASED ON MACHINE LEARNING ALGORITHMS

Author

Claudia Rodiera Clarens

Directors:

Dra. Adaia Valls

Dra. Rosa Borrás

PhD program: Healthcare

Universitat Internacional de Catalunya, 2023

To my dear family and parents,

This achievement would not have been possible without your unconditional love and support. Your dedication and encouragement have been my greatest inspiration on this academic journey. Through this achievement, I hope to honor and express my gratitude for everything you have done for me.

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to all those who have supported me throughout my PhD journey.

First and foremost, I am immensely grateful to my directors Dra. Borrás y Dra. Valls for their trust in me and for accepting to be my thesis advisors. Furthermore, I would also like to express my gratitude for their patience and understanding over these years, as it has sometimes been challenging for me to balance my residency in anesthesiology with the development of this research project. Moreover, I would like to add that their feedback throughout the project has been of immeasurable value and without it I would not be here today.

What is more, I would also like to extend my gratitude to my colleagues in the anesthesia department, including both anesthesia nurses and anesthesiologists at Teknon Medical center, my professional home. I would like to express my deep admiration for all of them, as they consistently demonstrate excellence in their daily practice and provide unparalleled care to our patients. Thank you so much for your involvement and collaboration. Without your immense support, despite the significant clinical workload and pressure, conducting this study would have been impossible.

Furthermore, I am deeply grateful to Nacho Rodriguez and Dr. Beatriz Giraldo for believing in the study from the very beginning, as well as for their selfless support and guidance, especially in the machine learning aspects of the project. Thank you sincerely for your time, patience and kindness and above all for providing me with a better understanding of machine learning. I would also like to thank you for your valuable input and constructive suggestions during the development of my research, which were a ray of light when I encountered technological difficulties that seemed insurmountable. Your assistance has been crucial.

Moreover, I would like to express my sincere gratitude to Helena, the biomedical engineer who collaborated on the project and with whom we worked hand in hand every day. Without her it would not have been possible. The success of this project is as much hers as it is mine. Thank you for your efforts and commitment, your willingness to help me and resolve my doubts at any time and to teach me everything I know about KNIME as well as for reassuring me that I could always count on you, you have been essential in the completion of the project. Thank you so much, sincerely.

On top of that I would like to express my utmost appreciation to my English teacher who has become a true friend over the years, Myles. I have been able to rely on his assistance in perfecting this document and he has been involved in helping me with English in other studies as well. Thank you for your dedication in teaching me critical thinking and the knowledge I have gained from it has significantly enhanced the quality of this study. My sincere thanks.

Finally, my heartfelt gratitude to my parents for always being there for me. Mami, thank you for your love and unwavering support, both in my professional endeavors and personal life. You are the pillar that keeps our family together. Thanks to my father who is also my boss. Papi, thank you for always being there for me. Thank you for being by my side in the operating room, ready to offer your help. Thank you for motivating me to pursue my doctoral thesis and to continue growing professionally as well as for providing me with excellent advice. You are the most brilliant person I know, and I feel incredibly fortunate to have had the opportunity to learn from you. You are a role model for me.

SUMMARY

Title: "Voice analysis as a method for preoperatively predicting a difficult airway based on machine learning algorithms"

Introduction: The occurrence of an unanticipated difficult airway is one of the greatest challenges faced by anesthesiologists as it can lead to severe complications. Its incidence ranges from 1.5-13.5%. An adequate preoperative assessment of the airway is a key factor in reducing the incidence of airway-related complications and should be routinely performed before any surgical procedure. Traditionally, airway evaluation has consisted of conducting screening tests based on anthropometric features of the patient as part of their physical examination. However, their clinical usefulness, accuracy and benefit remains unclear. For this reason, new methods of assessing the airway have been proposed with the aim of better assessing it such as the appearance of the voice as a new approach for the detection of a difficult airway and the emergence of artificial intelligence (AI) as a promising tool for clinical applications. Due to this, our study proposes the investigation of the voice for the prediction of a difficult airway using predictive algorithms based on machine learning.

Methods: The present study is observational, prospective, descriptive and multicentered. A total of N= 594 patients were enrolled in both Centro Medico Teknon and Institut Universitari Dexeus during the years 2019-2022. The study was approved by the ethics committee of Quironsalud group with the reference number: 62/2019. Two sub studies were performed including N=594 and N=313 patients for both the Mallampati and Cormack studies respectively. The inclusive criteria were adults over 18 years ASA physical status I-IV scheduled for a surgical procedure in need of orotracheal intubation by direct laryngoscopy. During the preanesthetic visit, variables regarding the clinical features of the patient and the traditional predictive tests were collected, as well as the voice recording. The headphones New Bee H360 were used for the recording.

Patients were asked to articulate the vocals "A, E, I, O, U" for 3 seconds in normal, flexion and extension positions. Then, the day of the surgery the Cormack grade was assessed by the anesthesiologist. For the data analysis the KNIME© analytical platform version 4.7 was used. Data was introduced into the classification algorithms and split into two groups: 70% training/validation and 30% testing for both the Cormack and the Mallampati study. To assess the model's performance, ROC curve values and other metrics were evaluated for all the different parameters and vocal combinations.

Results: For the Mallampati study, the combination harmonics + descriptive data with different combinations of the vocals E, I normal position and O flexion position vs extension positions only analyzing Mallampati I and IV cases, achieved a mean AUC of 0.97. In the Cormack study, two combinations generated the best results: first; harmonics + descriptive data all + vocal A in all positions only analyzing Cormack I and IV cases secured an AUC of 0.91; and second, voice parameters (including shimmer, jitter, HNR and others) + descriptive data + different combinations of vocals such as I flexion position, O normal position and O vocal all positions, only analyzing the Cormack I and IV cases, obtained a mean AUC of 0.90.

Conclusion: Acoustic parameters of the voice together with the clinical characteristics of the patients, when introduced into classification algorithms based on machine learning show promising signs of being able to predict a difficult intubation. Therefore, these algorithms have the potential to be used as a tool for the prediction of a difficult airway.

RESUMEN

Título: "Análisis de la voz como método de predicción preoperatorio de una vía aérea difícil mediante algoritmos de machine learning".

Introducción: la aparición de una vía aérea difícil no anticipada es uno de los mayores desafíos a los que se enfrenta un anestesiólogo, ya que puede conllevar complicaciones graves. Su incidencia varía entre 1,5% y 13.5%. Una evaluación preoperatoria adecuada de la vía aérea es un factor clave para reducir la incidencia de complicaciones relacionadas con la vía aérea y se debe realizar de manera rutinaria antes de cualquier procedimiento quirúrgico. Tradicionalmente, la evaluación de la vía aérea ha consistido en realizar test de screening basados en características antropométricas de los pacientes como parte de la exploración física de la evaluación anestésica preoperatoria. Sin embargo, su utilidad clínica, precisión y beneficio han sido cuestionados. Por este motivo, se han propuesto nuevos métodos de evaluación de la vía aérea con el objetivo de mejorar dicha evaluación. Concretamente, la voz ha surgido como un nuevo método para la detección de una vía aérea difícil. En los últimos años, la inteligencia artificial (IA) ha surgido como una herramienta prometedora para aplicaciones clínicas. En este estudio, proponemos investigar el uso de la voz para predecir una vía aérea difícil mediante algoritmos predictivos basados en machine learning.

Métodos: Se presenta un estudio observacional, prospectivo, descriptivo y multicéntrico en donde se incluyeron un total de N= 594 pacientes en el Centro Médico Teknon y el Institut Universitari Dexeus durante los años 2019-2022. El estudio fue aprobado por el comité de ética del grupo Quironsalud con el número de referencia: 62/2019. Se realizaron dos sub-estudios que incluyeron N= 594 y N=313 pacientes para los estudios de Mallampati y Cormack respectivamente. Los criterios de inclusión fueron: adultos mayores de 18 años, clasificados como ASA I-IV, programados para un procedimiento quirúrgico que requiriera de intubación orotraqueal

mediante laringoscopia directa. Durante la visita preanestésica, se recopilaron variables relacionadas con las características clínicas del paciente y los test predictivos tradicionales, así como la grabación de la voz. Para ello, se utilizaron los auriculares New Bee H360. Se pidió a los pacientes que pronunciaran las vocales "A, E, I, O, U" durante 3 segundos en posición normal, flexión y extensión. Posteriormente, en el día de la cirugía el grado de Cormack fue evaluado por el anestesiólogo. Para el análisis de datos se utilizó la plataforma analítica KNIME versión 4.7. Los datos se introdujeron en algoritmos de clasificación y se dividieron en dos grupos: 70% training/validación y 30% test para ambos estudios. Para la evaluación del rendimiento del modelo, se evaluaron los valores de la Curva ROC y otras métricas para los diferentes parámetros y combinaciones de vocales.

Resultados: Para el estudio de Mallampati, la combinación de armónicos + datos descriptivos de los pacientes con diferentes combinaciones de las vocales E, I en posición normal y O en posición flexión versus extensión, analizando únicamente los casos de Mallampati I y IV, se obtuvo una AUC de 0.97%. En el estudio de Cormack, dos combinaciones generaron los mejores resultados. Primero, armónicos+ datos descriptivos de los pacientes+ vocal A en todas las posiciones, analizando únicamente los casos Cormack I y IV, obtuvo una AUC de 0.91. Además, la combinación de parámetros de la voz (incluyendo Shimmer, Jitter, HNR y otros) + datos descriptivos de los pacientes + diferentes combinaciones de vocales como I en flexión, O posición normal y O todas las posiciones, analizando únicamente los casos Cormack I y IV, obtuvo una AUC de 0.90%.

Conclusión: Los parámetros acústicos de la voz junto con las características clínicas de los pacientes, cuando se introducen en algoritmos de clasificación basados en machine learning muestran signos prometedores de ser capaces de predecir una intubación difícil. Por lo tanto, estos algoritmos tienen el potencial de ser utilizados como herramienta para dicha predicción de manera preoperatoria.

PREFACE

Since beginning my journey in anesthesiology, having completed my residency just a few years ago, I have been confronted with increasingly complex and demanding scenarios. The dynamic nature of the field requires constant adaptation and learning as each patient presents a unique set of circumstances and considerations. As a young professional the magnitude of the challenges I face during my daily clinical practice can be both overwhelming and exhilarating. However, it is within these challenges that I discover the opportunity for growth and innovation.

In my time to date as a dedicated anesthesiologist, one particular area that has captured my interest and proved to be especially demanding is the management of difficult airways. Despite the significant advances in knowledge, development of new airway devices and predictive tools, combined with standardized guidelines from major global anesthesia societies, dealing with difficult airways remains a formidable challenge for an anesthesiologist.

The low prevalence of difficult airways adds to this complexity, making it a rare but critical situation that requires prompt recognition and effective management as otherwise the consequences can be severe or even life-threatening.

The main motivation for undertaking this doctoral thesis on the prediction of a difficult airway stems from a genuine concern for patient safety and a drive to improve clinical outcomes in the field of anesthesiology and airway management. As a doctor, I have always been fascinated by the potential of emerging technologies to improve healthcare outcomes. This interest led me to participate in the "Exponential Medicine Program" presented by Singularity University in San Diego, USA in 2017 during my first year of residency in anesthesiology. This university is a global community with a mission to educate, inspire and empower leaders to apply exponential technologies to help solve humanity's great challenges.

Annually, the world's experts in healthcare, technology, biopharma and innovative startup companies come together in this conference to explore and leverage the synergistic effects of rapidly advancing technologies in the transformative revolution of healthcare and medicine. It was during this conference where I found the inspiration to undertake this project as I had the privilege of attending a presentation done by a group of psychiatrists who had shared their groundbreaking work on voice analysis and its ability to predict episodes of depression and schizophrenia. They had developed an innovative phone application which showcased how the human voice could serve as a powerful diagnostic tool, capable of offering valuable insights into an individual's mental health status.

The profound impact of their research left a lasting impression on me and triggered a cascade of ideas regarding the potential application of voice in my field. Drawing from the strides made in voice-based predictive applications in psychiatry and upon realizing the evidence that the voice contains valuable diagnostic information, I contemplated the possibility that it could also provide insights into the anatomy of the upper airway. With this in mind, I became captivated by the possibility that voice analysis could potentially identify patients who are prone to difficult intubations. The idea of leveraging the unique characteristics of an individual's voice to detect potential difficult airways presented an exciting opportunity of research. This revelation inspired me to embark on this doctoral journey aiming to explore the predictive capabilities of voice analysis of a difficult airway.

The main objective of this thesis is to develop a predictive model based on the voice that could accurately identify patterns associated with a difficult intubation. By harnessing the advancements in data analysis such as those made with machine learning algorithms, this research seeks to pave the way for a non-invasive and straightforward approach for a preoperative detection of a difficult airway.

Throughout this thesis I will delve into the theoretical foundations of voice analysis and explore the relationship between voice characteristics and a difficult intubation. The methodology will encompass the collection and analysis of voice recordings from patients with varying degrees of airway

difficulty, with the aim of establishing meaningful correlations. Additionally, it will investigate the potential challenges and limitations of this approach while proposing strategies to overcome them.

It is with great enthusiasm and determination that I present this research and I would like to express my gratitude to my advisors, colleagues and mentors who have supported and guided me through this study. I am truly indebted to the psychiatrists whose work served as the catalyst for this thesis as well as all the patients who serve as the driving force that pushes us to continue improving.

INDEX

1. STATE OF THE ART 21

CHAPTER 1. ANESTHESIA 23

1.1. The anesthesia Management 23

1.1.1. Airway management 24

1.1.2. Difficult airway definition 27

1.1.3. Incidence and prevalence of a difficult airway 32

1.1.4. Difficult airway algorithms 32

1.1.5. Complications arising from a poor management of the airway 40

1.1.6. Factors involved in complications arising from airway management 43

1.2. Importance of the preoperatory airway evaluation 48

1.2.1. Preoperative airway evaluation predictive tests 48

1.2.2. Recent new airway evaluation predictive tests 58

CHAPTER 2. VOICE 61

2.1. Overview 61

2.2. Voice characteristics 62

2.2.1. Acoustic parameters of the voice 62

2.2.1.1. Frequency parameters 62

2.2.1.2. Intensity parameters 67

2.2.1.3. Noise parameters 69

2.3. Anatomical and physiological bases of voice formation 70

2.3.1. Main functions of the phonatory system 70

2.3.2. Musculoskeletal Framework 72

2.4. The vocal cords 78
2.5. The phonation cycle 79
2.6. Theories of voice formation 80
2.6.1. Myoleastic-aerodynamic theory 80

CHAPTER 3. ARTIFICIAL INTELLIGENCE 83
3.1. Machine learning (ML) and deep learning (DL) 83
3.1.1. Supervised learning 87
3.1.2. Reinforcement learning 87
3.1.3. Unsupervised learning 88
3.2. Machine learning methodology 88
3.3. Description of commonly used metrics in the evaluation of classification algorithms 92
3.4. Machine learning most commonly used algorithms 96
3.4.1. Naïve Bayes 96
3.4.2. Support vector machine 96
3.4.3. Decision tree 96
3.4.4. Random forest 97
3.4.5. K-nearest neighbors 97
3.4.6. Generalized linear 98
3.5. Machine learning healthcare applications 98

2. RESEARCH QUESTION 103

3. OBJECTIVES 107
3.1. Main objective 107
3.2. Secondary objectives 107

4. METHODS 111
4.1. Study design 111

4.2. Scope of the facilities 112
4.3. Sample size 114
4.4. Number of participants 115
4.5. Inclusion and exclusion criteria 115
4.5.1. Inclusion criteria 115
4.5.2. Exclusion criteria 116
4.6. Variables 116
4.6.1. Main variables 116
4.6.2. Secondary variables 118
4.7. Materials 122
4.8. Intervention 124
4.8.1. Preoperative phase of the study 124
4.8.2. Intraoperative phase of the study 129
4.8.3. Postoperative phase of the study 129
4.9. Ethics and legal aspects 136

5. RESULTS 141

5.1. Summary of the results 141
5.2. Mallampati study results 142
5.2.1. Descriptive data results of the participants included in both the Mallampati and the Cormack studies 142
5.2.2. Performance results of the classification algorithm for the Mallampati study 146
5.2.2.1. Results of accuracy metrics 146
5.2.2.2. Results of AUC and Pr AUC metrics 149
5.2.2.3. Results of precision and recall metrics 152
5.2.2.4. Results of F1 and Log loss values 155
5.2.2.5 Results of false positives and false negatives 158

5.3. Results of the Cormack Study 160

5.3.1. Descriptive data results of the participants included in the Cormack Study 160

5.3.2. Performance results of the classification algorithm for the Cormack study 164

5.3.2.1. Results of accuracy metrics 165

5.3.2.2. Results of Precision and recall metrics 167

5.3.2.3. Results of AUC and Pr AUC metrics 170

5.3.2.4. Results of F1 and log loss values 173

5.3.2.5. Results of False positives and false negatives 176

5.4. Selection of the best classification algorithms for both the Mallampati and the Cormack studies 179

5.5. Evaluation of the Mallampati performance as a predictor of the Cormack grade in our dataset 182

6. DISCUSSION OF THE RESULTS 187

7. CONCLUSIONS 201

8. BIBLIOGRAPHY 205

9. INDEX OF TABLES 173 227

10. INDEX OF FIGURES 233

11. ANNEXES 239

10.1. Annex 1. Roc Curve plot for traditional predictive tests for difficult laryngoscopy 239

10.2. Annex 2. Roc Curve plot for traditional predictive tests for difficult tracheal intubation 240

10.3. Annex 3. ASA difficult airway management algorithm 241

10.4. Annex 4. Research ethics committee report 242

10.5. Annex 5. Informed consent form 244
10.6. Annex 6. Data collection form 245
10.7. Annex 7. Full table 246
10.8. Annex 8. Full table 253
10.9. Annex 9. Full table 263
10.10. Annex 10. Full table 272
10.11. Annex 11. Full table 281
10.12. Annex 12. Full table 289
10.13. Annex 13. Full table 298
10.14. Annex 14. Full table 310
10.15. Annex 15. Full table 320
10.16. Annex 16. Full table 331
10.17. Annex 17. KNIME Workflow 340

1. STATE OF THE ART

CHAPTER 1. ANESTHESIA

1.1. ANESTHESIA MANAGEMENT

The perioperative process is defined as the time lapse surrounding the surgical act. It is subdivided into three stages: preoperative, intra-operatory and postoperative period, which are part of the surgical process. Furthermore, during this surgical process there is a comprehensive approach to the patient in which the relationship in each surgeon-anesthesiologist dyad is perhaps the most critical element of overall team performance (1). For this reason, the anesthesiologist is also a key component of the perioperative team which is responsible for patient safety and comfort throughout the surgical process.

During the preoperative phase, the anesthesiologist evaluates the patient's medical history and performs the physical examination as well as the necessary laboratory tests based on the preexisting medical conditions to determine the most appropriate anesthesia plan. Therefore, preanesthetic evaluation is considered a basic element of anesthesia care (2).

It is important to highlight the importance of the preoperative evaluation as not only does it allow the physician to assess the patient's overall health, but it also evaluates any potential risks for the anesthesia. For this reason, in certain patients it might be necessary to run evaluations and tests, which could be beneficial to the patient as its potential benefits may include a change in the content or timing of anesthetic management or a perioperative resource use that may improve the safety and effectiveness of the anesthetic processes (2).

Several studies have highlighted the importance of preoperative visits in reducing perioperative complications including anxiety, pain, hospital readmissions and even postponing the surgical procedure. Until the patient is optimized in order to face the surgery in their best conditions due to the fact that it provides the environment to detect certain conditions of

the patient which might have a negative effect in the postoperative period. For example, assessing the nutritional status of the patient at diagnosis as a nutritional supplement is known to have a positive effect on perioperative outcomes (3).

For this reason, the preoperative visit plays a vital role in supporting a successful surgical outcome and it should always be included in the preoperative surgical process. Moreover, another basic function of preanesthetic evaluation is to inform the patient of the procedure and to comply with the legal medical requirements regarding the signing of the informed consent. To elaborate further, during the intraoperative phase, the anesthesiologist administers and monitors the anesthesia to maintain the patient's physiological stability and emotional comfort. Furthermore, it is precisely in this phase when the anesthesiologist also fulfills a crucial function which is the compliance with the surgical security checklist before the beginning of the surgical incision as well as to ensure that all the information contained on it is correct. Finally, once the surgery is finished, the anesthesiologist continues to evaluate the patient for potential postoperative complications related to the anesthesia or the surgical procedure (4).

Overall, the perioperative surgical process is carefully tailored to each individual and the preanesthetic phase is crucial in order to ensure a safe and successful surgical outcome as it allows prehabilitation and preparation for the surgical procedure. In detail, it aims to augment physiological reserves and enhance functional capacity before surgery with nutritional, physical and psychosocial strategies to prepare patients to withstand the surgical stress response (5), as well as a faster return to the patient's daily life under the same conditions and maximised well-being.

1.1.1. Airway management

Airway management is a critical component of the surgical process, and it is essential for ensuring patient safety and optimal outcomes. In order to perform the surgical procedure without the patient experiencing discomfort during the procedure and in conditions of maximum safety for both the patient and the clinicians, the administration of the anesthesia and other medications

is often necessary. As a result, these drugs cause hypnosis, analgesia and neuromuscular relaxation and include agents such as opioids, inhaled anesthetic gases, propofol or neuromuscular blocking agents. However, the administration of the anesthesia and these other forms of medication may increase the risk of airway obstruction or respiratory depression. Therefore, it is crucial to carefully monitor the patient's airway and breathing throughout the procedure. For this reason, intubation to secure the airway might be needed in some cases. Specific indications for intubation include respiratory failure (Hypoxic or hypercapnic), apnea and a reduced level of consciousness among others (6). In particular the process of intubation consists of inserting a flexible plastic tube through a patient's nose or mouth and into their trachea. Its purpose, is to maintain an open airway and allow for the delivery of oxygen to the lungs either through mechanical ventilation or to prevent asphyxiation as well as airway obstruction. Precisely, the physician caring for the patient who decides to intubate is likely the person with the appropriate training to lead the team towards a successful intubation (7). Generally, in the operating room setting the intubation is carried out by the anesthesiologist.

Comprehensively, the process of intubation typically involves a number of steps including airway evaluation, position preparation, premedication, preoxygenation, face mask ventilation, induction of the anesthesia and the laryngoscopy to visualize the vocal cords as well as the confirmation of the tube placement using various methods such as auscultation or capnography. Once the tube is in place, the patient's airway is managed by adjusting the ventilation and monitoring for any signs of complications such as hypoxia, hypercapnia or airway obstruction (7).

Furthermore, the first step is to perform an airway evaluation, which should be typically done in the preanesthetic visit prior to the surgery. In detail, the pre-operative evaluation of adults undergoing elective noncardiac surgery updated guidelines recommends that screening for a difficult intubation should be carried out in all patients potentially requiring airway management for anesthesia or in the ICU (8). Furthermore, in the event that factors for a difficult airway management exist, there should be a primary plan and backup plans in the event of failure of securing the airway. Specifically, these include extra equipment preparation which consist of devices specially designed to

approach a difficult airway such as fiberscopes or videolaryngoscope due to the fact that the use of videolaryngoscopes reduce rates of failed intubation and result in higher rates of successful intubation on the first attempt with improved glottic views (9).

Moreover, there should be a designated team which is prepared and trained to manage a difficult airway with established roles and responsibilities during the procedure. On top of that, a stepwise algorithmic approach can be used to manage these situations, ensuring interventions are carried out in a systematic manner with appropriate backup plans in place if the initial attempt is unsuccessful. Moreover, in anesthetic practice, cognitive aids enhance performance and their use has been recommended in elective airway management (10). Once the airway evaluation has been performed the following step is the preparation of all the equipment that is required for intubation. This process involves gathering and setting up various tools and devices that are necessary for successfully performing an intubation procedure. Furthermore, this preparation is essential for ensuring that everything is in place and ready to be used when needed as otherwise, studies have shown that the lack of patient preparation, equipment checks or protocol deviations occur in up to half of critical incidents (11).

The next necessary step before administering sedative and paralytic medications is to ensure that the patient undergoes a pre-oxygenation phase with a face mask. The purpose of pre-oxygenation is to increase the amount of oxygen in the alveoli while reducing the concentration of nitrogen, which is achieved by using a high fraction of inspired oxygen (FiO2). Particularly, this is an essential step to reduce the rate of decline of the oxyhemoglobin during apnea, which is the moment in which the patient is not breathing on their own before the insertion of the endotracheal tube. It is important to highlight that performing an effective pre-oxygenation is crucial as it saves time for the anesthesiologist to secure the airway if the patient requires multiple attempts before successfully inserting the tube, minimizing the deleterious effects that oxygen deprivation causes in the patient. Moreover, preoxygenation broadens the safe apnea time period, described as the duration until a patient attains a saturation limit of 88% to 99% (12).

Following the steps outlined above, the anesthesiologist proceeds to intubation. For this purpose, a direct laryngoscope is used. The laryngoscope, should be held in the operator's left hand and inserted into the right side of the patient's mouth, pushing inward and upward at a 45-degree angle against the tongue. As the laryngoscope moves toward the back of the oropharynx, the operator may use the blade to shift the tongue towards the left side of the mouth to provide space for the endotracheal tube until all the structures of the oropharynx are visible and the vocal cords are exposed. In this situation, when the vocal cords are visible using the laryngoscope such as in Cormack grades I and IIa, successful intubation is usually anticipated (13).

Finally, the last step once the endotracheal tube has been placed is to verify that it is correctly positioned in the trachea, located proximal to the carina. This can be confirmed through two different methods: firstly, through an end-tidal carbon dioxide monitor which measures the amount of expired carbon dioxide with each breath allowing for the detection of esophageal intubation. In fact, Waveform capnography is the gold standard for confirming correct endotracheal tube placement (14). Secondly, to further ensure proper placement of the tube, the physician should auscultate for symmetrical bilateral breath sounds in the lungs (15). Following that, the anesthesiologist connects the patient to mechanical ventilation and the surgical procedure is carried out.

To conclude, potential complications also need to be considered. In detail, the most feared complication of intubation is hypoxemia and cardiac arrest that may be precipitated by multiple attempts with poor oxygenation and failed intubation (16,17). For this reason, it is extremely important to adequately follow all the steps of the process, including a correct preoperative evaluation of the airway, a correct preparation in the event of a difficult airway and a detailed plan in the event of a difficult intubation to avoid these serious complications.

1.1.2. Difficult airway definition

The occurrence of an unanticipated difficult airway is one of the greatest challenges faced by anesthesiologists. The American Society of

anesthesiologists defines the difficult airway as the clinical situation in which a conventionally trained anesthesiologist experiences difficulty with facemask ventilation of the upper airway, difficulty with tracheal intubation or both (18). However, a standard definition of the difficult airway cannot be found in the available literature.

The latest version of the American guidelines from the American Society of Anesthesiologists, published in 2022 presents fundamental recommendations based on a combination of current literature, expert opinions and the experience of the practitioner. These guidelines specify that a challenging airway pertains to a clinical scenario in which an anesthesiologist encounters anticipated or unanticipated difficulty or failure during airway management, which may involve any or all of the following: Facemask ventilation, laryngoscopy, supraglottic airway ventilation, tracheal intubation, extubation or invasive airway procedures (18).

With regard to difficult facemask ventilation, it is defined as the impossibility to maintain correct ventilation due to an inadequate mask seal, excessive leak, or elevated peak ventilation pressures to the entrance of air. Moreover, the guidelines differentiate between difficult laryngoscopy defined as the impossibility of visualizing any parts of the vocal cords after several attempts and a difficult tracheal intubation, defined as a tracheal intubation that either needs several attempts or it is unsuccessful despite multiple attempts. These guidelines also take into consideration the supraglottic devices as well as difficult supraglottic airway ventilation as the situation when it is not possible to adequately ventilate the patient due to a misplacement of the supraglottic device, when it requires multiple attempts or when there is an inadequate airway seal or excessive leakage of air and the pressure needed for the air entrance is high. Furthermore, it also mentions that a difficult airway can also appear during extubation. It refers to it as the inability to maintain an open airway and sufficient airflow after removing a tracheal tube or supraglottic device from a patient who is at risk of having a difficult airway. Moreover, another definition is a difficult or a failed invasive airway which refers to anatomical characteristics or abnormalities that decrease or prevent the likelihood of successfully inserting a surgical airway through the front of the neck into the trachea. Additionally, it also mentions that a difficult airway is

when there is an inadequate ventilation including absence of inadequate chest movement, the lack of exhaled carbon dioxide, clinical signs of severe obstruction, decreasing of oxygen saturation, hypoxemia and hypercapnia among many others (18).

Regarding the guidelines of the Catalan Society of Anesthesiology published in 2020 in the Spanish Journal of Anesthesiology and resuscitation, it can be seen that they follow the same framework as the American guidelines. Furthermore, the guidelines consider a difficult airway if an anesthesiologist who has already completed their training meet any of following criteria (19):

A) Suspicion or evidence of difficulty due to prior history or physical examination which induces to perform an airway management in spontaneous ventilation.

B) Difficulty in one of the planned sequences to secure the airway which requires two or more attempts or the change of device or operator.

C) Difficulty that causes a significant reduction in the oxygen saturation with or without complications which can cause injuries resulting from the instrumentalization of the airway.

In addition, difficult ventilation is defined as insufficient or impossible with a face mask that includes the following:

a) Excessive air leak or resistance to airflow.

b) Lack of chest movement and capnography waveform on the monitor or appearance of gastric distension.

c) Requirement of improvement maneuvers such as optimizing position, use of oro/ nasopharyngeal airway, 4- hand ventilation or deepening the neuromuscular blockade.

d) Compromises oxygenation and CO2 elimination.

With regard to a difficult laryngoscopy, it is defined as the following:

It refers to the inability to visualize the glottis fully or partially despite applying techniques such as external laryngeal pressure (BURP maneuver)

or changing patient's position, resulting in a Cormack Lehane grade 3 or 4. To elaborate further, Cormack and Lehane (1984) proposed a four-grade scoring system to describe the view at direct laryngoscopy (20) the assigned grades are the following: **Figure1**

- Cormack grade 1: Full view of the glottis.
- Cormack grade 2: Partial view of the glottis or arytenoids.
- Cormack grade 3: Only epiglottis visible.
- Cormack grade 4: neither glottis nor epiglottis visible.

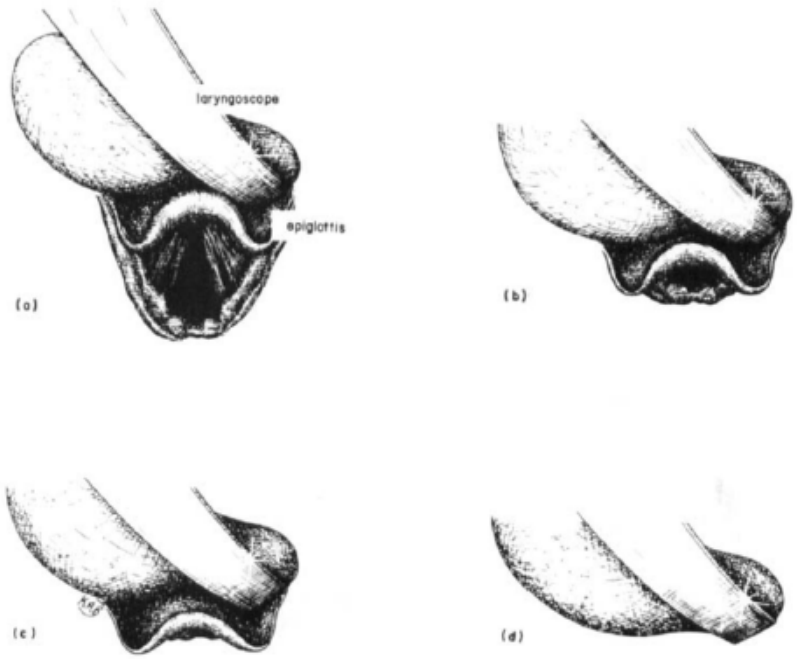


Figure 1. *Best views obtained at direct laryngoscopy, assuming the correct technique. Original from R.S Cormack and J Lehane, Difficult tracheal intubation in obstetrics. Anesthesia, 1984, Volume 39, pages 1105-1 I1* (20).

However, this scoring system has been modified since it was initially proposed by Cormack and Lehane. In 1998, Yentis and Lee subdivided grade 2 into two subcategories: 2a which corresponds with a partial view of the glottis and 2b which agrees with a view of the arytenoids or posterior section

of the vocal cords (21). In this study, Yentis et al, evaluated this modified version of the Cormack and Lehane scoring in a prospective study which consisted of 663 participants.

As a result, it concluded that the modified scoring system provided more information than the original as grade 2b laryngoscopy view is frequently encountered and is associated with difficulties advancing a tracheal tube. It is noteworthy that the importance of these studies lies in the fact that since their publication, the Cormack Lehane grade have been used almost universally to classify a difficult laryngoscopy.

On the other hand, difficult intubation can also be referred to as the difficulty in advancing the tracheal tube, requiring more than 2 attempts by one or more operators, the use of bougie or changing the initial plan. Moreover, the guidelines include two aspects to consider when it comes to difficult ventilation with supraglottic devices: 1. Difficulty in placing the device, requiring more than two maneuvers or more than 2 insertion attempts, 2. Insufficient ventilation due to the leakage or obstruction once the device is place, which does not allow efficient oxygenation or CO2 elimination requiring a change of supraglottic device. On top of that, it also considers difficult subglottic access and refers to it as difficulty in a) locating the cricothyroid membrane and or b) encountering difficulty in puncturing or inserting the cannula requiring more than one attempt or excessive time that does not avoid functional repercussions. Finally, it also considers difficult extubation in a patient with difficult airway referring to the suspicion of a) difficulty in maintaining sufficient ventilation/ oxygenation after extubation and or b) subsequent difficulty in managing the airway using any of the aforementioned modalities (19).

The above mentioned are the difficult airway definitions detailed in the most important guidelines. For this reason, although it is important to highlight that there is not a standard definition of a difficult airway it is still crucial for anesthesiologists to be familiar with the various definitions and criteria outlined in these guidelines and expert recommendations. This is due to the fact that despite slight variations in these definitions, there are many similarities in all of them.

Thus, by understanding and applying these definitions, providers can more effectively assess and manage patients with challenging airways and therefore, reduce the risk of adverse events, improve outcomes and enhancing patient safety.

1.1.3. Incidence and prevalence of a difficult airway

The incidence of a difficult airway varies widely depending on the patient population and the definition used. However, different studies have reported an incidence rate ranging from 1.5-13.5% (22). Nevertheless, it is commonly accepted that the incidence of a difficult airway in the surgical population is approximately around 5.8% (23). More specifically, some studies have focused on different subgroups of patients such as pediatrics where some studies show an incidence of a 3.6 % (24). With regard to the obstetrics population studies have shown an incidence of a 5% (25). Furthermore, another population which in particular may present a higher incidence of a difficult airway owing to the characteristics of the patients which may present with tumors or airway obstructions and the nature of the surgical procedure that in many cases involves the manipulation of the airway. In some studies, the prevalence in these kinds of patients of a difficult intubation is 8% (22)

1.1.4. Difficult airway management algorithms

Due to the relevance of a proper management of the difficult airway, the main existing societies of anesthesia have developed Clinical practice guidelines with the aim to provide direction to anesthesiologists with the management of patients with difficult airways, increase the likelihood of successful airway management on the first attempt, improve patient safety during airway management and reduce or prevent adverse events.

In the latest version of the American society of anesthesiologist guidelines, there is a section intended to provide evidence for and insight into airway management for adult and pediatric patients with either anticipated or unanticipated difficult airways, obstetric, intensive care and critically ill patients (18) **(Annex 3)**.

These guidelines provide strategies for two distinct types of situations. On the one hand, they provide direction for managing difficult airways in patients with an anticipated difficult airway. On the other hand, it details the strategy that should be followed with patients with unanticipated difficult airways and emergency airways.

In the first situation, it highlights the need of the development of a preformulated strategy considering factors such as the patient's conditions, the type of surgery and the anesthesiologist's skills and preferences. Furthermore, it notes that strategies for awake intubations while the patient is conscious, difficult ventilation or intubation, inability to ventilate or intubate and emergency invasive airway rescue should be identified. Moreover, awake intubation may be appropriate for patients who are at risk of aspiration or incapable of tolerating a brief apneic episode. As well as this, combination techniques may be used if individual techniques fail, highlighting the importance of limiting the number of attempts at tracheal intubation or supraglottic airway devices placement in order to avoid injuries and complications. Furthermore, alternative invasive interventions should be identified if the selected approach fails. However, these guidelines recommend that invasive airway procedures should be performed by individuals properly trained in invasive airway techniques, including ECMO which should be initiated when appropriate and available.

On the other hand, it also states the performance in the event of an unanticipated difficult airway. It clearly states that the airway management of an unanticipated airway consists of 7 interventions. The first step is calling for help. Secondly, it focuses on the need for optimizing oxygenation. Following that, it also notes the use of a cognitive aid or difficult airway algorithm. Moreover, it considers the use of noninvasive airway management devices and combination techniques if difficulty is encountered with individual techniques as well as to assess the benefit of waking and/or restoring spontaneous breathing. If a noninvasive device approach is selected there should be an established sequence of noninvasive devices for airway management. Finally, if none of the prior steps resolved the situation it evaluates the use of invasive airway management interventions and ECMO.

Additionally, the guidelines also provide examples of algorithms for airway management as seen in **Figure 2**.

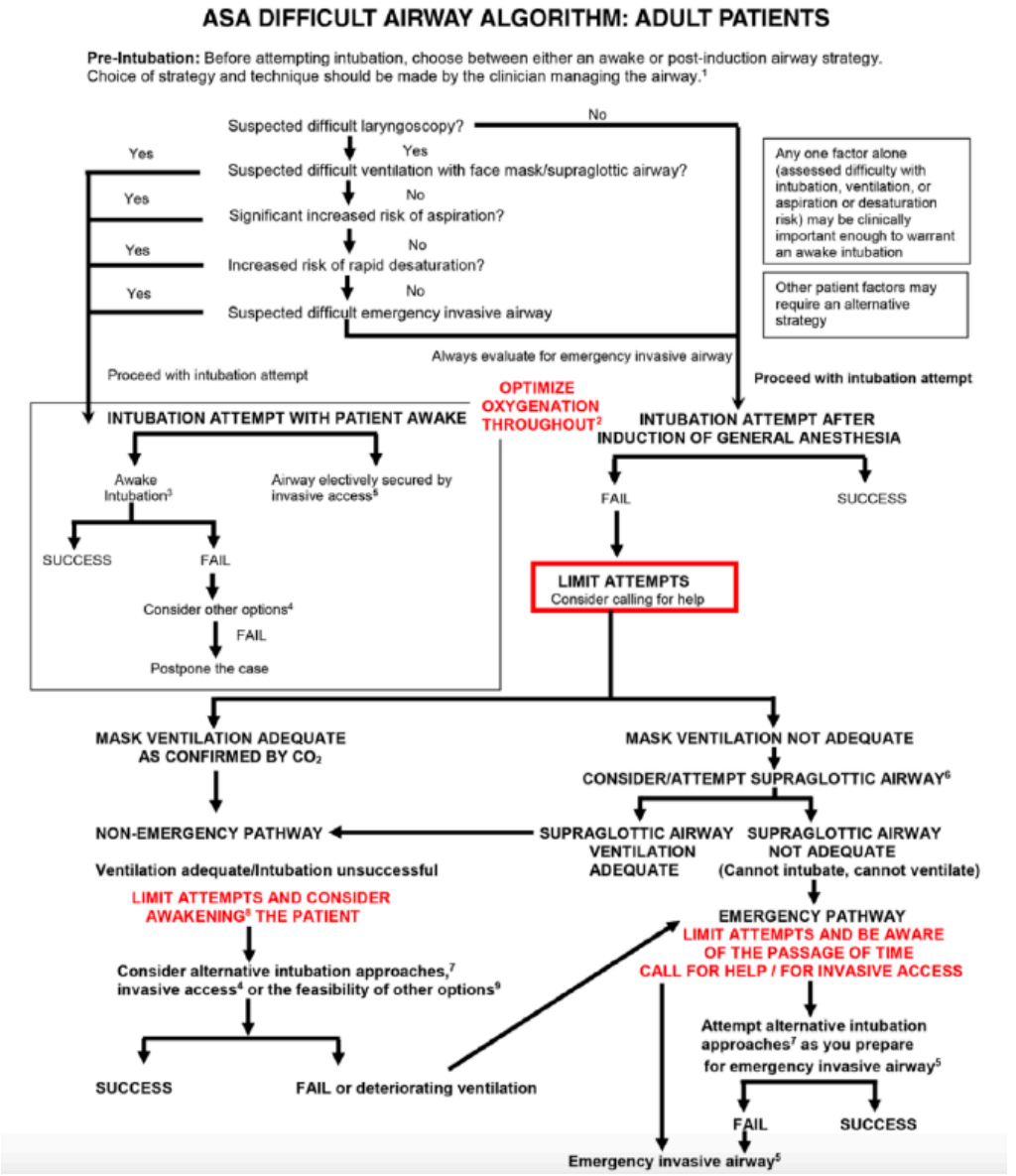


Figure 2. *Difficult airway algorithm: Adult patients* (18).

Overall, the ASA 2022 guidelines recommend the use of a systematic approach to airway management with a focus on identification of patients at risk for difficult airway management (18). In addition, the guidelines also stress the importance of appropriate preoxygenation and the use of supraglottic devices and videolaryngoscopy to improve first-attempt success rates during intubation.

What's more, in line with these results, other guidelines have also been developed with the aim to provide a comprehensive framework for the identification and management of patients at risk for a difficult airway. Another good example are the Canadian Guidelines (26-27). These guidelines are updated consensus-based recommendations. To elaborate further, the guidelines are divided into two parts. The first part focuses on strong and weak recommendations for and against based on three levels of evidence: A (High), B (Moderate) and C (Low) for difficult tracheal intubation encountered in an unconscious/ induced patient.

Specifically, it states that there should be a primary approach to tracheal intubation (Plan A), which is usually direct laryngoscopy. However, as a response to difficulty encountered with the first attempt, the following maneuvers can be made: external laryngeal pressure, followed by the use of a videolaryngoscope and the use of adjunct instruments such as a bougie. Furthermore, the importance of mask ventilation is also noted or optionally the placement of supraglottic devices. However, an alternative approach should be used after no more than two failed attempts at tracheal intubation using the primary approach and should employ a different device or operator. It is important to consider that if tracheal intubation is not successful after three attempts, it should be considered as a failed intubation situation and therefore, exit strategies should be considered as ineffective attempts at intubation could potentially harm the patient.

Having arrived at this moment, two different scenarios should be differentiated. Firstly, if the oxygenation is adequate the guidelines state the following exit strategies:

1) Awakening the patient.
2) Proceed with the surgery using a facemask or a supraglottic device.

3) Prepare equipment or expert help for a controlled attempt at tracheal intubation.
4) Proceeding with surgical access.

However, if the oxygenation fails it recommends promptly proceeding with a surgical transtracheal airway, most frequently, cricothyroidotomy.

The algorithm summarizing all these recommendations is found in **figure 3**.

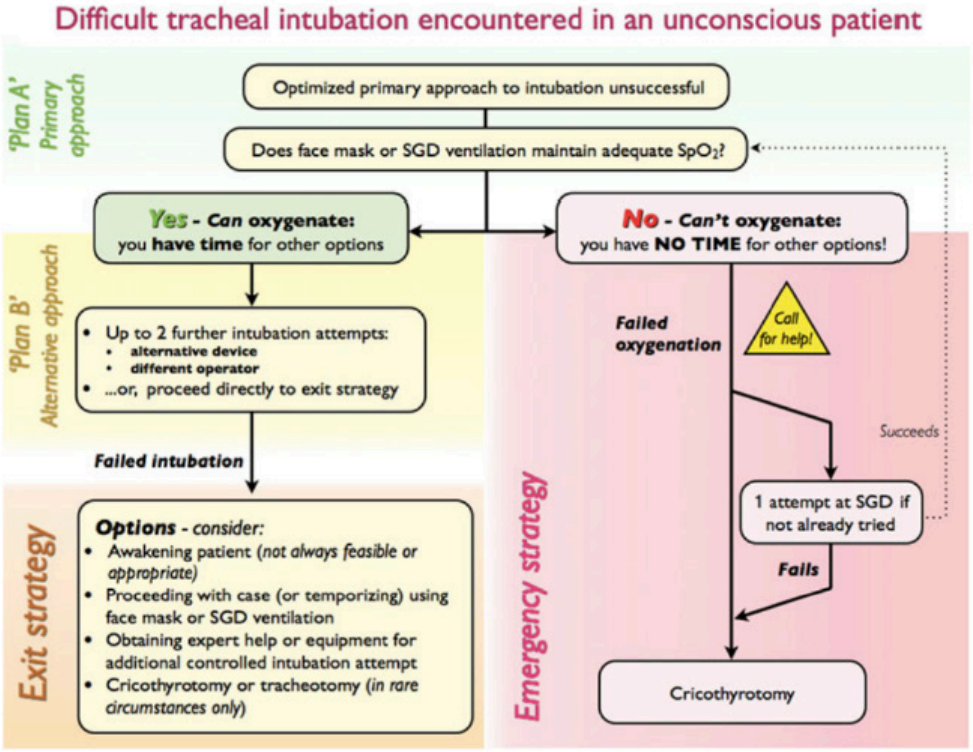


Figure 3. *Flow diagram: difficult tracheal intubation encountered in the unconscious patient. SGD = supraglottic device* (26).

On the other hand, the second part of these guidelines focus on the planning and implementation of safe management of the patient with an anticipated difficult airway (27). The article highlights the importance of early identification and assessment of these kinds of patients, the need for a multidisciplinary team approach to airway management and the use of appropriate equipment and techniques based on patient and clinical factors. Additionally, it also emphasizes the need for regular training and simulation-based education to improve patient safety and reduce the incidence of adverse events related to difficult airway management.

To conclude, it is important to note that in a more local environment the Catalan Society of Anesthesiology has also put effort into developing its own guidelines (19). These guidelines similarly to the American and Canadian ones also emphasize the difference between two different scenarios, the anticipated difficult airway and the non-anticipated difficult airway. However, the Catalan guidelines include safety criteria which a patient will either meet or not. If they do, they will be considered to have an airway with no difficulty. However, in the event that they fail these criteria they will be designated with a potentially difficult or anticipated difficult airway. These are not present in their American and Canadian counterparts. The results of this criteria will cause the anesthesiologist to vary their actions in the algorithm provided in **Figure 4**.

In detail the safety criteria are the following:

- Possible ventilation with facial mask or supraglottic devices.
- No bronchoaspiration risk.
- Low risk of rapid oxygen desaturation..
- High probability of intubation using a videolaryngoscope.
- Necessary help and instruments.

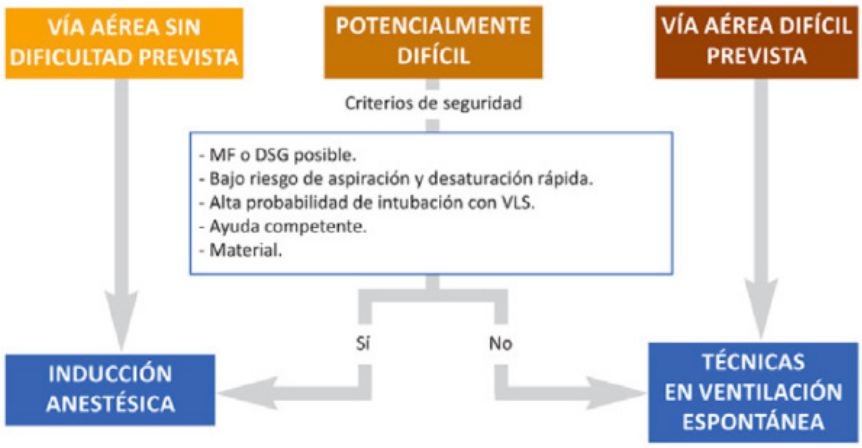


Figure 4. Different types of clinical situations based on the compliance with the safety criteria. *Recomendaciones para la evaluación y manejo de la vía aérea difícil prevista y no prevista basadas en la adaptación de guías de práctica clínica y consenso de expertos. Revista Española de Anestesiología y reanimación 2020*(19).

In addition, when it comes to the management of a difficult airway it also accentuates two different situations. Once again, these guidelines follow similar recommendations compared to ASA and Canadian ones.

In detail, it emphasizes the need of a plan A and B when approaching an anticipated difficult airway. With regard to plan A, depending on whether the patient meets the safety criteria. If it does an intubation is necessary and it considers the use of a videolaryngoscope whereas if intubation is not needed it recommends the use of a supraglottic device. However, If the patient does not meet the safety criteria, it recommends the use of intubation techniques with spontaneous ventilation. Additionally, if plan A doesn't work initially, it recommends you follow plan B which includes a surgical access to the airway through cricothyroidotomy or tracheotomy.

On top of that, it also states the algorithm to follow in case of an unanticipated difficult airway, similarly to other guidelines it also states a plan A, B, C and D.

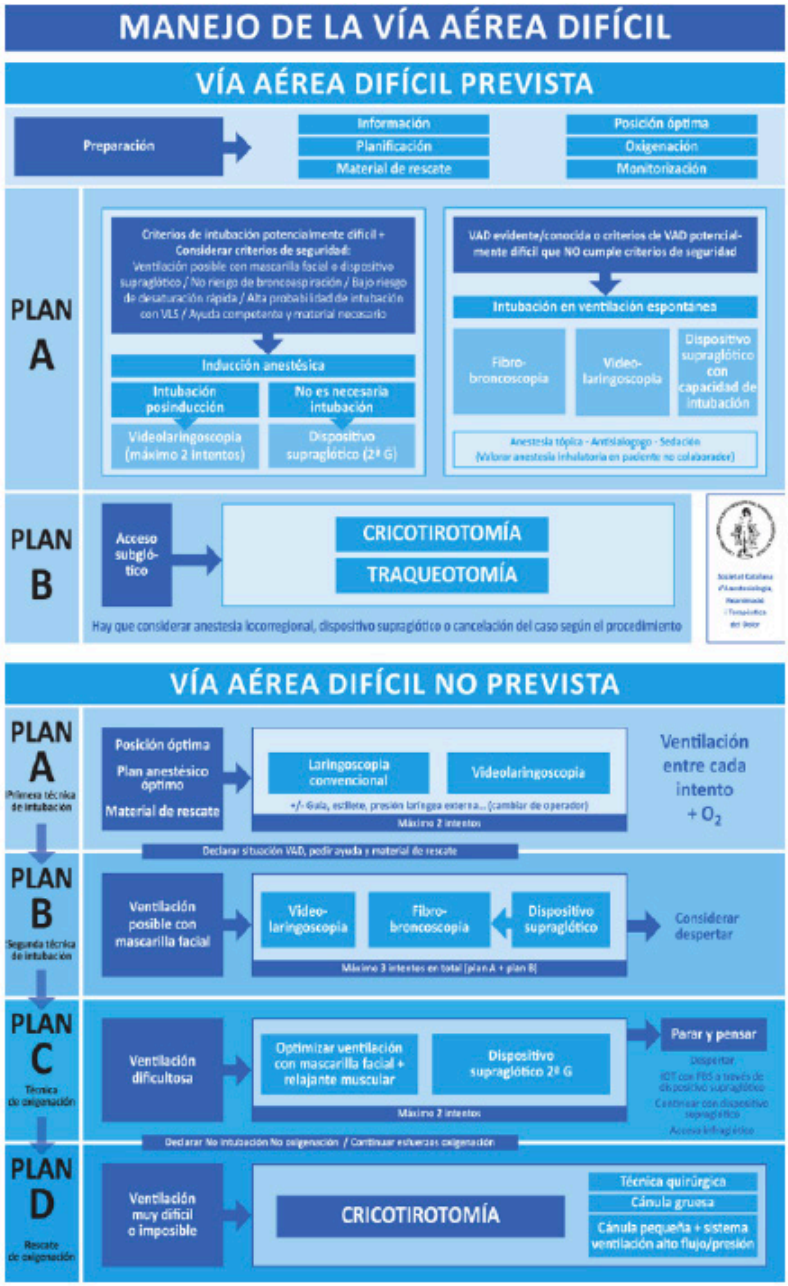


Figure 5. *Recommendations for the management of the airway. Revista Española de Anestesiología y reanimación SCARTD* (19).

Regarding plan A, it explains that the second intent should significantly improve the successful possibilities of an intubation with respect to the first one. This may include optimizing position, use of extra instruments such as a bougie or the use of a videolaryngoscope. If this plan fails it should be declared an unforeseen difficult airway situation and call for help as well as prepare extra equipment. Following that, plan B should be followed which includes the use of an alternative intubation technique such as the use of videolaryngoscope or assisted fiberoptic intubation through a supraglottic device. Furthermore, if this plan is not successful, the anesthesiologist should start plan C, which focuses on optimizing oxygenation through improving mask ventilation, deepening the neuromuscular blockade or try ventilation through a supraglottic device. Finally, if none of these plans work the physician should start plan D in which the situation can't oxygenate, can't ventilate should be declared and therefore a surgical infraglottic access should be rapidly initiated. **Figure 5**

In conclusion, the findings of the guidelines stated above demonstrate the complexity of managing difficult airways, especially the unpredicted difficult airway and highlights the importance of adhering to evidence-based guidelines to improve patient outcomes. Furthermore, the guidelines also emphasize the need for a standardized approach to airway management that can be adapted to the unique needs of the patient according to its risk factors and the surgical procedure. Ultimately, the use of appropriate strategies to prevent and manage complications associated with difficult airways is critical to ensure the safety and well-being of patients undergoing surgical procedures.

1.1.5. Complications arising from a poor management of the airway

The occurrence of an unanticipated difficult airway is one of the greatest challenges faced by anesthesiologists as a situation with both difficult mask ventilation and difficult intubation is potentially life-threatening (28). An unanticipated difficult airway is an extremely serious situation that can lead to severe complications and have fatal consequences thus, being a major contributor to morbidity and mortality of patients from anesthesia causes. In

fact, not achieving adequate airway management is one of the main causes related to the development of anesthetics complications (29).

The closed claim analysis carried out by M. Joffe et al (29) analyzed claims with difficult tracheal intubation as the primary damaging event. Their findings showed that despite updated practice guidelines and improved airway techniques potentially preventable complications still occur with difficult or failed tracheal intubation. Insufficient or inadequate airway assessment, failure to anticipate and situation awareness were identified as one of the contributors to a failed or a difficult intubation in this and other studies (29).

Furthermore, the latest version of the fourth National Audit project (NAP4) was a comprehensive study carried out by Cook et al (30) in 2011. This study was conducted in the United Kingdom to evaluate major complications related to airway management during anesthesia. Specifically, it was led by an expert group of anesthesiologists and was one of the largest studies of its kind with the participation of 309 NHS hospitals. The primary aim of the study was to identify the incidence and nature of major complications associated with airway management during anesthesia, and to investigate the factors that contributed to these complications. Moreover, it also identified specific outcomes associated with airway related complications during anesthesia. These include brain damage, death, the need for emergency surgical intervention and admission to ICU. In particular, the study identified 133 reports related to anesthesia related complications. From this, 16 of the patients the final outcome was death, is an overall mortality rate of 12%. Additionally, 3 patients presented with brain damage. An emergency surgical approach of the airway such as tracheostomy or cricothyroidotomy was needed in 58 of the 133 cases. With regard to ICU admission, it was necessary in 100/133 patients. The main causes of anesthesia-related causes of admission in the ICU were the following: manage airway swelling or trauma and aspiration of gastric content. **Figure 6.**

Table 5 Final outcome: narrative outcome and NPSA classification (Table 2)

	All cases (*n*=184)	Anaesthesia (*n*=133)
Final outcome (narrative)		
Death	38	16
Brain damage	8	3
Other partial recovery	10	6
Full recovery	124	106
Unrelated death	4	2
Final outcome (NPSA definitions)		
Death	38	16
Severe	10	5
Moderate	126	103
Low	7	6
None	3	3

Figure 6. *Final outcome: narrative outcome and NPSA classification. Fourth National Audit project of the Royal College of Anesthetist and the difficult airway society* (31).

Additionally, this study also specifies that problems with tracheal intubation were the most frequently recorded primary event accounting for 39% of all events during anesthesia such as: difficult or delayed intubation, failed intubation and the situation cannot ventilate, cannot intubate and aspiration of gastric contents. These outcomes are detailed in **figure 7.**

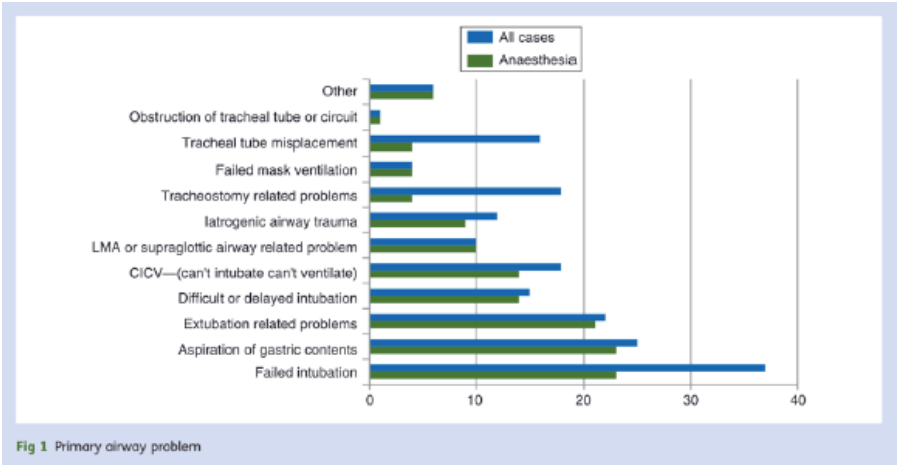


Figure 7. *List of recorded primary events problems. Fourth National Audit project of the Royal College of Anesthetist and the difficult Airway Society.*

Ultimately, The Fourth National Audit project of the Royal College of Anesthetists and the difficult Airway Society, highlights the importance of proper airway management providing important insights into the incidence, nature and risk factors for airway- related complications during anesthesia.

Additionally, the importance of the complications derived from a poor management of a difficult airway are such that the studies explained above are not the only ones that have dedicated themselves to studying the outcomes that may derive from an inadequate difficult airway management. For example, more recently the Canadian Medical protection association also carried out a study led by Crosby et al in 2021 in which anesthesiology airway-related medicolegal cases were investigated (32). In this study, 406 legal cases records involving anesthesiologists were investigated. It is important to note that of these, 46 (11%) of the cases involving 47 patients identified complications related to airway management. The result derived from poor airway management emphasizes that up to 30% of patients had moderate to severe harm of which 11% had anoxic drain damage and 52% of the patients died.

Furthermore, in the same line as these larger studies there are many smaller studies that have been dedicated to studying the implications derived from a difficult airway. It is important to note that most of them include minor complications, but sometimes, as seen in the aforementioned studies it can lead to extremely severe complications counting dysrhythmias, laryngospasms, surgical access to the airway, hypotension, cerebral ischemia and cardiac arrest and death (16,33-35). For this reason, not ensuring the airway is considered a crisis situation.

1.1.6. Factors involved in complications arising from airway management

Despite advances in airway management techniques and equipment as well as the development of Clinical practice guidelines, complications do still occur. For this reason, and due to the fact that these complications can have significant implications for patient safety, clinical outcomes and healthcare costs, some studies have put effort into better understanding

the factors involved in the development of complications including patient-related factors, Clinician-related factors, and system-related factors as well as to identify strategies to reduce its incidence and therefore improve patient safety and quality of care. It is important to highlight that despite the fact that in all the available clinical guidelines for a difficult airway the preoperative assessment of the airway and the preparation of all the necessary equipment is considered a crucial step prior to any surgical procedure and management of the airway, studies show that along with other factors, a lack of preoperative assessment of the airway is significantly involved in major airway related complications.

To expand on this point, in the closed claims analysis carried out by Joffre et al 2019(36) it has been concluded that inadequate airway planning and judgment errors were contributors to patient harm. To be more precise a total N=195 from the period 1993 until 2012 claims related to difficult intubation were investigated. According to this study, it clearly states that an inadequate preoperative or airway evaluation was related to 17% of the claims related to difficult intubation. **Figure 8.**

Moreover, it is striking that in 21% of the claims inadequate support and equipment as well as a lack of structured communication were found. These findings emphasize the importance of a proper airway evaluation as an adequate airway assessment would allow for better equipment preparation and support. Therefore, it may contribute to avoiding or reducing those complications.

Table 3. Predictors of Difficult Tracheal Intubation and Judgment Failures in Airway Management

Question	Claims, No.	%
Indicate any predictors of difficult tracheal intubation (whether known/recognized at the time or not) or factors that contributed to difficult airway management.		
Airway obstruction from any cause*	31	30%
Past history of difficult intubation	21	21%
Mallampati grade 3 to 4	19	19%
Limited cervical spine extension	16	16%
Limited mouth opening	13	13%
Secretions/blood in airway	12	12%
Short neck	10	10%
Swollen tongue	6	6%
Short thyromental distance	6	6%
Thick or bull neck	6	6%
History of neck irradiation	5	5%
Preeclampsia	2	2%
Prominent teeth	1	1%
Number of predictors		
0	24	24%
1	36	35%
2 to 6	42	41%
Inappropriate difficult airway management		
Failure to use supraglottic airway as a bridge ($\kappa = 0.552$)	27	26%
Perseveration ($\kappa = 0.489$)	25	25%
Failure to plan for difficult tracheal intubation (induction; $\kappa = 0.627$)	23	23%
Delayed calling for, or did not call for, a surgical airway ($\kappa = 0.436$)	20	20%
Inadequate preoperative or airway evaluation ($\kappa = 0.664$)	17	17%
No backup plan for difficult reintubation (extubation; $\kappa = 0.664$)	14	14%
Number of judgment failures (n = 97)†		
None of the above (appropriate management)	26	27%
1	34	35%
2 to 5	37	38%

Figure 8. *Predictors of Difficult tracheal Intubation and Judgment Failures in Airway Management* (36).

In accordance with these results more evidence that a major contributor of morbidity and mortality related to airway management is a lack of proper preoperative assessment can be found in the analysis of deaths related to anesthesia in the period 1996-2004 from closed claims registered by the Danish Patient Insurance Association (37). This study examined a total of 1.256 claims related to an injury from the field of anesthesia. To be more

specific, there were 24 claims which included death out of the total number of claims. To clarify further, 16 % of these claim deaths (4/24) were due to airway handling, meaning a significant amount of claims and therefore, highlighting the importance of proper airway management. Furthermore, that importance is also emphasized in the study's conclusions where it states that these fatalities may potentially have been preventable if an adequate airway algorithm including thorough preoperative evaluation, training and the use of standardized protocols for treatment had been used.

Furthermore, another recent example which is available in the literature is the study carried out by the Canadian Medical Protection Association where it analyzes anesthesiology-related medicolegal cases lead by Crosby et al in 2021(32) This study collected 406 claims from which 46/406 (11%) were airway related. To provide more details, from these 46 airway-related claims in 27/46 (59%) of them there was an inadequate preoperative airway evaluation of risk factors such as difficult airway history, comorbidities or inadequate or incomplete airway assessment. **Figure 9.**

Moreover, surprisingly a complete airway examination was lacking in 10/27 (37%) of the surgical cases.

Table 4 Peer expert and analyst-identified judgement failures for CMPA closed cases, 2007–2016 ($n = 46$ cases)

	N (%)
Inappropriate difficult airway management	
Failure to use supraglottic airway as a bridge	1 (2)
Perseveration	5 (11)
Failure to plan for difficult airway management (induction)	7 (15)
Delay in calling for, or failure to call for, a surgical airway	2 (4)
Inadequate preoperative or airway evaluation	27 (59)
Number of judgement failures	
None of the above (appropriate management)	16 (35)
1	20 (43)
2–5	10 (22)

CMPA = Canadian Medical Protective Association.

Figure 9. *Peer expert and analyst-identified judgment failures for CMPA closed cases, 2007–2016 (n = 46 cases)* (32).

Additionally, further evidence in line with these results can be found in the literature. For example, in the analysis of claims for compensation after injuries related to airway management carried out by Fornebo et al in 2017,

which stated that more than half of the severe cases were caused by failed intubation or misplaced endotracheal tube (38).

As well as the lessons learned from the courts of South Korea carried out by Jae Hoon Lee et al 2021, studies which showed medical malpractices and severe complications related to endotracheal intubation in south Korea in which the most common problem was failed or delayed intubation (39).

However, it is important to note that human factors are also significantly involved in the development of adverse outcomes related to anesthesia. For example, Flin et al 2013 carried out a study called: Human factors in the development of complications of airway management. Preliminary evaluation of an interview tool. In this study, Flin used the human factors investigation tool to collect information about 28 types of human and organizational factors, related to four components of an accident trajectory figure 10. The study concludes that the most frequently mentioned factors were aspects of situation awareness such as failure to anticipate or make the wrong decision, followed by job factors like time pressure and finally factors related to the person, that is, hunger or tiredness (40).

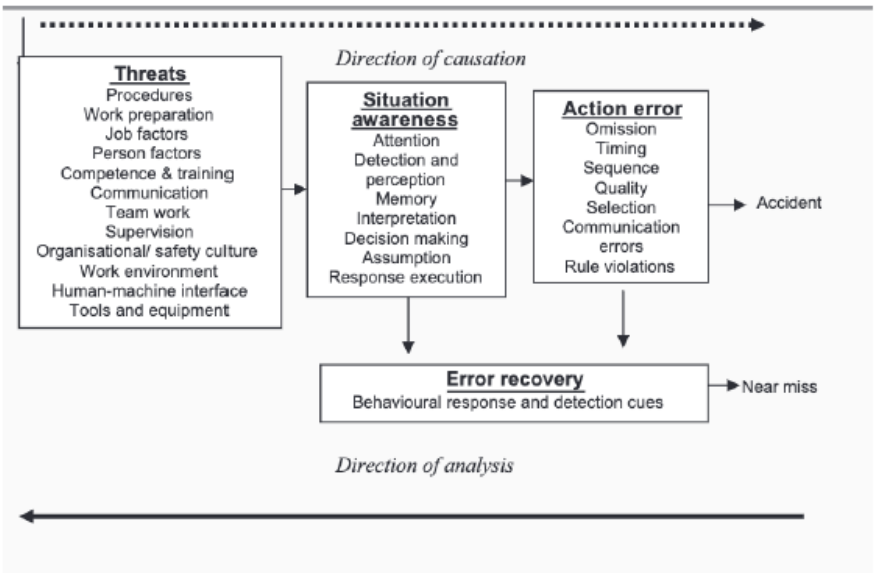


Figure 10. *Human Factors Investigation Tool model of incident causation and direction of analysis* (40).

To conclude, in light of these results, it is essential to emphasize the importance of an adequate preoperative assessment of the airway in reducing the incidence of airway-related complications. The above studies analyzing the claims related to airway management revealed that a significant number of cases could have been avoided by a thorough evaluation of the patient's airway prior to the surgical procedure.

For this reason, the evaluation of the airway in a preoperative phase of the perioperative process is a priority and should be routinely performed before any surgical procedure as such intervention could notably reduce the number of claims related to airway complications and therefore, improve patient outcomes.

1.2. IMPORTANCE OF THE PREOPERATIVE EVALUATION

1.2.1. Preoperative airway evaluation predictive tests

The difficult airway is a potentially life-threatening situation. For this reason, traditionally airway assessment tests have been described for screening a difficult airway preoperatively as part of the physical examination. However, it is important to note that the difficult airway is focused on the upper respiratory tract. Moreover, this difficulty is not solely related to a single anatomical feature of the patient's appearance but a multifaceted combination of patient anatomy, clinical conditions and healthcare provider expertise. For this reason, despite many predictive traditional tests, an adequate assessment of a difficult airway still remains a challenge as their clinical utility, accuracy and benefit remains unclear (41,42). Nevertheless, due to the fact that in some situations an unanticipated difficult airway can appear in a patient without obvious anatomical peculiarity and appear to have normal airway there is still a risk of encountering an unexpected difficult airway. For this reason, the search for adequate diagnostic screening tests has continued and bedside tests are routinely performed to identify potential difficulties and therefore, help plan alternative airway management.

These tests can be completed in a few seconds and are usually easy to perform and not-time consuming. Furthermore, these tests are based on the preoperative evaluation of particular physical features of patients which have linked to a difficult intubation as part of the physical examination such as the following: (36)

- Limited mouth opening.
- Neck diameter.
- Limited spine extension.
- Thyromental distance.
- Jaw protrusion.
- Visibility of the uvula.

Firstly, one of the most commonly used tests was the modified Mallampati test which was developed by S.R Mallampati in 1983. He suggested that the difficulty of the intubation could be predicted by the size of the tongue base in relation to the oropharyngeal cavity. Additionally, he stated that the proper assessment of this clinically should be done with the patient in the upright position meanwhile opening their mouth and maximally protruding their tongue. Furthermore, he carried out a study involving 210 patients in which studied the correlation between this clinical sign and the intubation difficulty (43). As a result, he described a simple grading system which involves assessing the preoperative ability to visualize the faucial pillars, soft palate and base of uvula as a way to predict the intubation difficulty. The three categories were the following:

- Class I: Fully visualize the faucial pillars, soft palate and uvula.
- Class II: Full visualization of the faucial pillars and soft palate. Unable to fully visualize the uvula.
- Class III: Only soft palate can be visualized.

However, later on, Samson and Young added a new category which referred to the inability to visualize the faucial pillars, uvula or soft palate known as class IV (44). As a result, nowadays the modified Mallampati score including four categories is commonly used and includes as following: **Figure 11**

Class I: Soft palate, uvula and pillars are visible.

Class II: Soft palate and uvula are visible.

Class III: Soft palate and uvula are visible.

Class IV: Only the hard palate is visible.

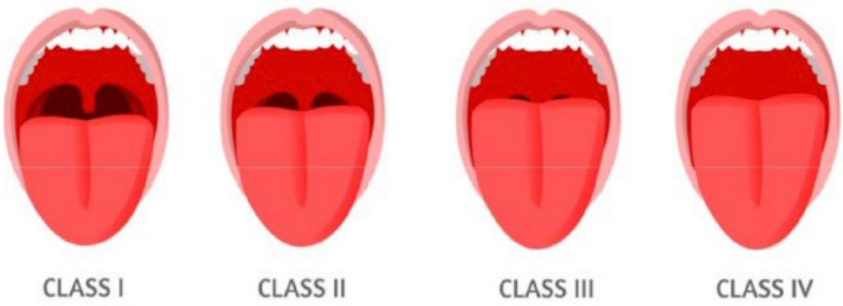


Figure 11. *The modified Mallampati classification* (45).

Following its introduction in the 1980s the Mallampati classification rapidly became one of the most commonly used bedside tests as part of the standard preoperative physical examination for the prediction of a difficult airway. However, since then many studies have analyzed and examined this predictor (46). For example, Roth et al carried out a large systematic review which included a total of 133 studies in 2018.(47) When it comes to predicting a difficult laryngoscopy, in the Mallampati test the obtained mean sensitivity was 0.4 [0.16-0.71] with 95% Confidence interval (CI) whereas the Modified Mallampati test slightly improved these results, obtaining a mean sensitivity of 0.53% [0.47-0.59] (CI 95%). **(Annex 1)** Similarly, this score shows comparable findings when predicting a difficult intubation obtaining a sensitivity for the modified Mallampati test of 0.51 [0.4-0.61] (CI 95%). **(Annex 2).** Furthermore, other meta-analyses such as that carried out by Lee et al (48) obtained a sensitivity for the original Mallampati score for a difficult laryngoscopy and the Modified Mallampati score ranging 0.05 to 1 and 0.12-1 respectively. In this same study, parallel findings were found for a difficult intubation where the sensitivity for both the original Mallampati score and the Modified Mallampati score ranged from 0.34 to 0.66 and

0-0.88. Additionally, another large meta-analysis carried out by Shiga et al (49) found comparable results with regard to Mallampati score sensitivity ranging from 49% to 86%, although no distinction was made between the original Mallampati score and the modified Mallampati score. Overall, the performance in terms of sensitivity for the prediction of either a difficult laryngoscopy or a difficult intubation for both the Mallampati and the modified Mallampati score ranges from poor to good and are characterized by significant variability and high heterogeneity among studies (50). Moreover, it has been shown to undergo notable variation in certain ethnic groups as there are differences in anatomical measurements commonly used to predict a difficult airway between the Indian and non-Indian population (51) Furthermore, in particular situations such as pregnancy, the physiological and anatomical changes in pregnancy contribute to difficult laryngoscopy and visualization of the larynx (25).

On the other hand, with regard to specificity for the Mallampati score, similar data can be found in the literature as a high variability of outcomes exists. Nevertheless, the values for specificity of this score are relatively higher compared to sensitivity values. For example, Roth et al obtained for the original Mallampati score a mean specificity for a difficult laryngoscopy of 0.89 (CI 95%) ranging from 0.75 to 0.96 and a mean specificity of 0.80 (CI 95%) ranging from 0.74 to 0.85. Convergent results were found for both the original Mallampati score and the Modified Mallampati score for the prediction of a difficult intubation.

Furthermore, another commonly used anthropometric measure which has been previously related to an increased intubation difficulty is a short thyromental distance. This is measured from the anterior larynx to the mandible and is viewed as an indicator of the mandibular space and can also be used to determine the ease or difficulty of displacing the tongue with a laryngoscope blade. In general, a thyromental distance of greater than or equal to 6 centimeters or the width of 3 fingerbreadths is acceptable. However, a thyromental distance of less than 6 centimeters has been linked to a difficult intubation. **Figure 12**.

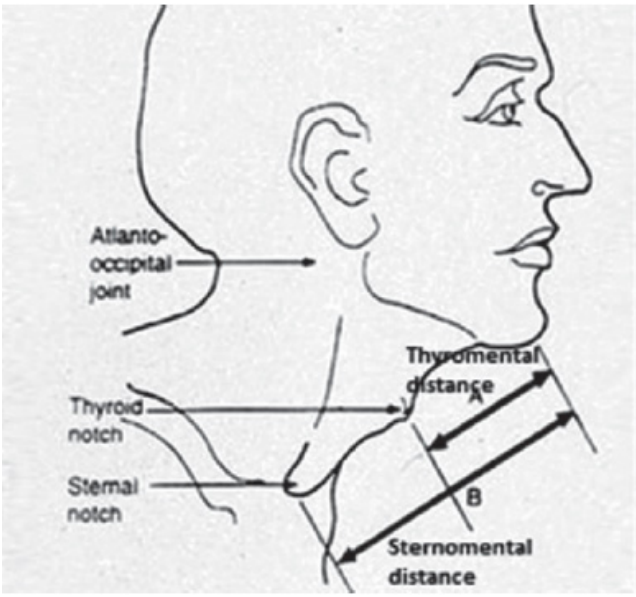


Figure 12. *A. thyromental distance, B. Sternomental distance* (52).

With regard to this clinical sign, Roth et al (47) showed to some extent low sensitivity values. Concretely, it obtained a mean sensitivity value of 0.37 and 0.24 for both the prediction of a difficult laryngoscopy and a difficult intubation respectively. Matching conclusions were obtained for Shiga et al (49) as in their findings the sensitivity for the thyromental distance was approximately 20%. On the other hand, with regard to the specificity the thyromental distance yielded better results. For example, in the Roth et al meta-analysis, the specificity procured for a thyromental distance for the prediction of a difficult laryngoscopy was of 0.89 [0.84-0.93] (CI 95%) and of 0.90 [0.80-0.96] (CI 95%) for a difficult intubation. Moreover, corresponding outcomes according to this bedside test were also found in the Shiga study where it achieved a specificity for a difficult intubation of 94%.

On top of that, another anatomical measurement which has previously been related to difficulties with intubation and airway management is the Sternomental distance. It refers to the distance between the bony prominence of the sternum and the mandible with the head fully extended. For this reason, a shorter Sternomental distance is associated with a higher risk of difficult intubation as it is an indicator of a smaller oropharynx which may make it

more challenging to maneuver the laryngoscope and therefore, to achieve a successful intubation. Roth et al, encountered that this measurement yielded a reasonably low sensitivity of 0.33 [0.16-0-56] (CI 95%) whereas other studies such as the one carried out by Shiga et al found disparate outcomes as it reached a sensitivity of 62% which could be considered moderate. On the contrary, in terms of specificity, the literature shows relatively high values such as 0.92 (47) and 0.82 (49). However, in most of the meta-analysis few studies including this measurement were comprised due to the fact that only a limited number of studies analyze Sternomental distance. For this reason, it is not possible to conclude an overall diagnostic performance of this test.

What's more, the mouth opening test has also been used and studied as a bedside difficult intubation prediction test. Furthermore, this prediction is based upon the fact that mouth opening indicates movement of the temporomandibular joint and therefore, it makes it more difficult to visualize the larynx in patients with a limited mouth opening. Thus, restricted mouth opening may make access to and control of the difficult airway (53). Besides, the interincisal distance or maximum mouth opening is classified into three grades: **Figure 13.**

- ≥ 5cm
- ≤ 3,5 cm. and < 5 cm
- < 3,5 cm

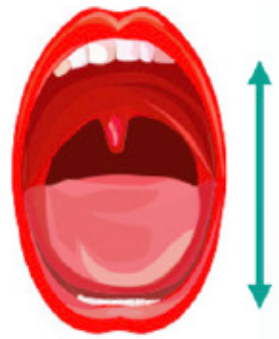


Figure 13. *Mouth opening or interincisal gap (Source: Preanestesia2.anestalia.com).*

Additionally, for this particular measurement, Roth et al, and Shiga et al, encountered similar findings in terms of sensitivity. To explore more in detail, for a difficult laryngoscopy and a difficult intubation Roth et al found a sensitivity of 0.22 [0.13-0.33] and 0.27 [0.16-0.41] (CI 95%) respectively.

Parallelly, Shiga et al also found a mean sensitivity of 22%, meaning that the sensitivity procured in both analyses is low. However, resembling the outcomes according to specificity found in the literature of other bedside tests, the analysis revealed high values of specificity of 0.94 [0.9-0.97] (CI 95%) and 97% for both studies, respectively.

Moreover, the upper lip bite test is one of the various bedside tests used for prediction of difficult laryngoscopy intubation. However, its usefulness is still not very clear, and there is controversy regarding its accuracy (54). This test is based on the patient's ability to bite the upper lip with the lower incisors. If a patient experiences difficulties in doing so may be due to a limited mouth opening, protruding incisors or a short thyromental distance, all of which are related to difficult intubation. **Figure 14.**

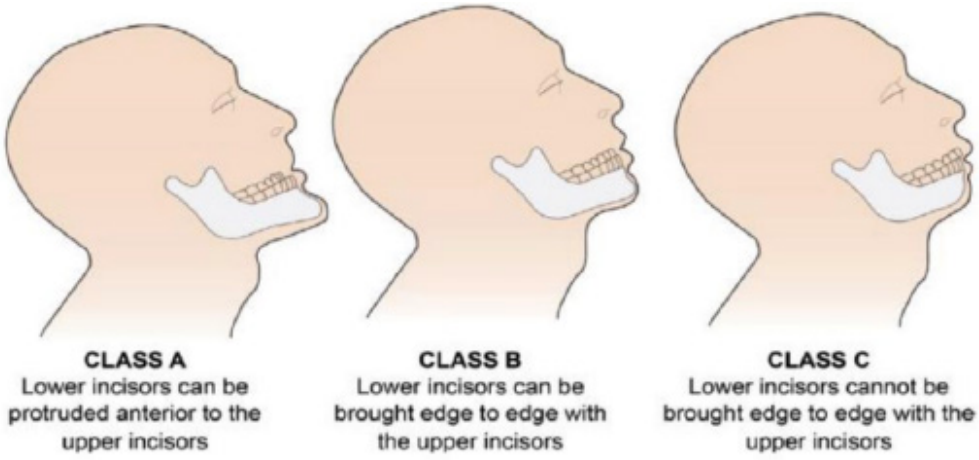


Figure 14. *Jaw protrusion grade A, B, C respectively* (55).

Several studies have reported on the sensitivity of the upper lip test bite for difficult intubation. Overall, although the sensitivity of this test appears to vary across studies, contrarily to other bedside tests, this one yielded relatively moderate sensitivity up to 0.67 [0.45-0.83] (CI 95%) according to Roth et al (47). In addition, when it comes to specificity, it achieves high scores of 0.92 [0.86-0.95] (CI 95%)

In light of these results, it becomes evident that the currently available screening bedside test for a difficult intubation has only poor to moderate

discriminative power when used alone as compared to a combination of tests (56). Therefore, they might not be sufficiently accurate and may not be well suited for detecting an unanticipated difficult airway.

For this reason, the best way to predict a difficult airway is by using an index or a multivariate analysis that combine several predictors as it has been shown to improve the sensitivity and specificity when compared to the value of each test alone (57-60).

What's more, the most used scores systems are the Wilson risk score and the Arne test (61-62)

Firstly, the Arne test is a predictive tool for difficult intubation based on a multivariable index. This test was validated in a population of N= 1090 patients and achieved a sensitivity and specificity of 94% and 96% in general surgery respectively. Moreover, it was also evaluated in other kinds of surgeries such as non- cancer ENT in which it achieved a sensitivity of 90% and a specificity of 93%. It was also assessed in ENT cancer surgery in which it achieved a 92% and 66% sensitivity and specificity respectively (62). **Figure 15.**

Table 1: *Number of patients undergoing different surgical procedures in two studies, initial and validation studies. J. Arné et al 1998. British Journal of Anaesthesia* (62).

Table 1 Number of patients undergoing different surgical procedures in the two studies (initial and validation studies)

Surgical population	Initial study (*n*)	Validation study (*n*)
ENT surgery		
Cancer	187	83
Non-cancer	425	290
General surgery		
Oral	312	421
Neurosurgery	102	
Abdominal	92	53
Gynaecology	29	53
Orthopaedic	28	32
Urology	25	88
Cardiac		70
Total	1200	1090

Furthermore, the ROC analysis provided a high and comparable discriminant power for predicting difficult tracheal intubation with the area under the curve (AUC) 0.95. The cut-off value of 11 was found to provide the best balance between sensitivity and specificity.

Risk factors	"Points" of the exact score	"Points" of the simplified score
Previous knowledge of difficult intubation		
No	0	0
Yes	3.28	10
Pathologies associated with difficult intubation		
No	0	0
Yes	1.63	5
Clinical symptoms of airway pathology		
No	0	0
Yes	0.98	3
Inter-incisor gap (IG) and mandible luxation (ML)		
IG ⩾ 5 cm or ML > 0	0	0
3.5 < IG < 5 and ML = 0	1.09	3
IG < 3.5 cm and ML < 0	4.12	13
Thyromental distance		
⩾ 6.5 cm	0	0
< 6.5 cm	1.36	4
Maximum range of head and neck movement		
Above 100°	0	0
About 90° (90° ± 10°)	0.65	2
Below 80°	1.46	5
Mallampati's modified test		
Class 1	0	0
Class 2	0.66	2
Class 3	1.93	6
Class 4	2.52	8
Total possible	15.35	48

Figure 15. *Arné test Index. J. Arné et al 1998. British Journal of Anaesthesia* (62).

With regard to the Wilson score, M.E Wilson et al published it in the B.J Anesthesia in 1988. **Figure 16**

Initially, he selected a group of patients including those with a known difficult intubation and measurements were taken to establish descriptive statistics for the development of a simple rule. What's more, the tool was tested and confirmed using the same data. Finally, the rule was validated in another set of patients(N=633) (61).

TABLE IV. *The three levels of the final five risk factors. (* 5 cm is approximately three fingers' breadth)*

Risk factor	Level	
Weight	0	< 90 kg
	1	90–110 kg
	2	> 110 kg
Head and neck movement	0	Above 90°
	1	About 90° (ie. ± 10°)
	2	Below 90°
Jaw movement	0	IG ⩾ 5 cm* or SLux > 0
	1	IG < 5 cm and SLux = 0
	2	IG < 5 cm and SLux < 0
Receding mandible	0	Normal
	1	Moderate
	2	Severe
Buck teeth	0	Normal
	1	Moderate
	2	Severe

Figure 16. *Wilson risk score. M.E Wilson et al, B.J. Anesthesia, 1988.*

Furthermore, further studies have shown that the sensitivity and specificity for a difficult intubation for this risk score is approximately 0.51 (0.40-0.61) and 0.95 (0.88-0.98) respectively (47).

Even so, the diagnostic accuracy and feasibility of these tests or index may be questioned because despite being performed routinely preoperatively in most patients the number of unpredicted difficult intubations remain considerably high. However, systematic evidence-based and consistent airway assessment may reduce the incidence of unanticipated difficult airway management (63,64). In a study published by Norkov et al in 2015 on the diagnostic accuracy of anesthesiologists' prediction of a difficult airway, revealed that only half of difficult intubations were predicted (64). In light of these results, it seems evident that there is a need to find better tests or methods to improve our capacity as anesthesiologists to predict a difficult intubation preoperatively.

1.2.2. Recent methods of airway assessment

New methods of assessing the airway have been proposed with the aim of better assessing the airway. Ultrasound is one of the most explored new methods of airway assessment due to its usefulness when assessing subhyoid parameters which are otherwise difficult to evaluate with the other traditional bedside tests as they are limited to the external suprahyoid evaluation such as the mouth opening. Concretely, accessing this anterior soft thickness tissue via ultrasound may help us identify patients with potentially difficult laryngoscopy due to consistency of soft tissues without any external abnormality (65). Due to the fact that this airway evaluation tool is gaining importance in recent years, studies investigating the accuracy of the ultrasound have been carried out (66-68). For example, a large systematic review and meta-analysis carried out by Carsetti et al in 2022(69) in order to assess whether preoperative upper airway ultrasound can predict a difficult airway in adult patients without clear anatomical evidence of difficult airway on the traditional preoperative physical examination. As a result, it was shown that ultrasound assessment and more specifically the measurement of the distance -skin-epiglottis (DSE) with a cut-off value of 2.36 centimeters procured a AUC-ROC of 0.87 (95% CI,0.84-0.90). Moreover, it also showed high specificity and sensitivity for the prediction of a difficult laryngoscopy. Besides, other studies report anatomical findings obtained through ultrasound comparable to Cormack grade obtained during direct laryngoscopy (70). These outcomes are similar to the ones obtained in other studies, which is why it appears that the use of ultrasound for airway assessment may represent a potential tool to improve the performance of a difficult airway bedside prediction test and at the same time reducing interobserver variability. However, as a disadvantage, the use of ultrasound may be probe-dependent and it's a methodology which might require a learning curve.

Additionally, other authors propose other imaging tests which involve the measuring of the diameters of the upper respiratory tract and then correlate them with the degree of difficulty in intubation such as functional magnetic resonance imaging or computerized tomography. These techniques, enable an exhaustive evaluation of airway pathology such as airway trauma, fractures, tumors and help to figure out the dimensionality of the

airway especially, CT with 3- dimensional reconstruction, permits excellent visualization of the pharyngeal airway (71). Finally, acoustic reflectometry is another described method with great potential for expansion but currently underutilized which is based on the measuring of the volumes of the upper airway (72).

Furthermore, newly developed techniques such as nasal endoscopy, virtual endoscopy and virtual laryngoscopy are becoming commonly used.

These techniques create virtual images and 3D reconstructions through software from CT scans. It achieves the acquisition of accurate images, airway mapping and even dynamic videos. Concretely, virtual endoscopy is a tool that can detail intraluminal anatomical "fly-through" information in a format visually similar to the flexible endoscopic views. (73) Moreover, this technology helps to assess the anatomy of difficult airways and to assist in the formulation of the most optimal airway management strategy in such patients (74).

What's more, due to the fact of being a non-invasive diagnostic tool and their high accuracy and safety, these techniques have emerged as a promising tool for assessing challenging airway conditions such as patients with deformities, facial asymmetries or airway anomalies where conventional methods might not be sufficient for a proper airway evaluation (75-77).

What's more, it is noteworthy that in recent years artificial intelligence (AI) has increasingly gained importance in the medical field. More concretely, advances have been made for applying this technology, specifically facial image recognition with the aim of improving the diagnosis of a difficult airway. With this purpose, many studies have been carried out willing to create AI models which use the patient's facial images to predict the difficulty of intubation (78,79). One study worth mentioning is the study led by Tatsuya et al 2021 (80) which achieved to develop an AI model for classifying intubation difficulties from 16 different images of each patient from a total of 205 participants. The results showed the following: accuracy of 80.55%; sensitivity, of 81.8%; specificity,83.3%; AUC 0.864; (CI 95% 0.73-0.96). In light of these results, it seems evident that AI has emerged as a promising method and will only continue expanding its potential.

In recent years, the voice has emerged as a new potential method for the detection of a difficult airway preoperatively. (81-83). These studies are built on the fact that voice parameters can reflect the anatomical characteristics of the upper airway. Our study shows that acoustic parameters of the voice together with the clinical characteristics of the patients when introduced into classification algorithms based on machine learning have promising signs of being able to predict a difficult intubation. More evidence can be found in the literature, for example, a study carried out by Valls et al, aimed to evaluate the impact of orthognathic surgery on speech, addressing in particular the effects of skeletal and airway changes on the resonance characteristics of the voice and on articulatory function (84). Furthermore, in previous studies, Carvalho et al 2021 already showed a significant association between three formants and Cormack-Lehane scale classification (82).

To further elaborate, this study included 453 participants. Each of them were recorded and were asked to pronounce the vowels a,e,i,o,u preoperatively. After initial analysis five models' performance was evaluated showing a significant association between three formants and Cormack-Lehane classification. The model containing Mallampati, and formants procured the following results: The area under the curve was 91.8%; sensitivity 87.5%; specificity, 82.7%. To conclude, Carvalho et al, suggested that voice can have a role in difficult airway prediction. Other authors such as Shuang Cao et al, have revealed that voice parameters differed significantly between patients with difficult mask ventilation compared to patients with easy mask ventilation (83) In these previous studies formants were analyzed. However, other voice parameters such as jitter or shimmer might be also potentially relevant biomarkers for predicting a difficult airway. Thus, the authors conclude that voice parameters may be considered as potential predictors of a difficult airway. However, prospective validation and additional studies are needed to confirm these findings and therefore, to assess the applicability of this novel approach.

CHAPTER 2. VOICE

2.1. OVERVIEW:

The voice is created through the collaboration of many parts of the body working together in a coordinated manner. These include the respiratory system, digestive system, vocal tract, and various muscles throughout the body. Although none of these components are dedicated solely to voice production, each plays an essential role in the process.

Furthermore, the thoracic cavity serves as the primary source of air for energy generation, while the larynx is responsible for regulating the frequency and producing vibrations. Concretely, through a nonlinear interaction between the airstream and the collapsible segment of the airway wall also known as soft tissue, it becomes possible to sustain vibrations through which as a result, the fundamental frequency (F0) is produced, as well as a spectrum of higher frequencies (85).

Moreover, the vibrating tissue allows the propagation of the acoustic waves which travel from the larynx and are retained in the airway in the form of multiple reflections from irregular boundaries. Besides, a small portion of the sound is modified by the resonance cavities including the hypopharynx, oropharynx, oral cavity and nasal passages which give the unique harmonics of human speech.

Furthermore, amplitude and frequency modulation of the fundamental frequency and higher partials are added to allow rhythmics and melodic patterns (85). It is important to highlight that voice output is a psychophysiological response that is part of the human integrative psychophysiological stress system, which is a complex integration of sympathetic and parasympathetic control (86). All these processes are regulated by both the central and peripheral nervous systems, as well as the auditory system. Moreover, as a curiosity, literature provides compelling evidence that the human vocal apparatus has been, at least partly, shaped by sexual selection (87).

2.2. VOICE CHARACTERISTICS

2.2.1. Acoustic parameters of the voice

Voice parameters are quantitative measures used to assess the acoustic characteristics of the voice and can provide valuable information about the voice production. Thus, being helpful in the diagnosis, treatment of voice disorders and different pathologies of the airway as well as contribution to the understanding of mechanisms underlying normal and disordered voice production. Concretely, acoustic analysis of voice has become widely used for correct diagnosis of dysphonia (88). Moreover, the most commonly used voice parameters used for such assessment can be classified into three groups and include the following: Frequency parameters, intensity parameters and noise parameters.

2.2.1.1. *Frequency parameters*

Fundamental frequency (F0):

The fundamental frequency corresponds to the lowest frequency component of the signal. It represents the number of times that vocal cords open and close per second and is expressed in cycles per second or Hertz (Hz). The fundamental frequency at which vocal cords vibrate can be influenced by several factors, such as changes in the magnitude of the vocal cord due to edema or masses which can cause a reduction of the F0; the cordal viscoelasticity and the suglottic pressure (89). Due to this fact, situations in which the thickness of the vocal cord is increased such as edema, tumors or masses can result in a reduction in the F0 causing a deeper voice.

Moreover, a pathology which causes an increase in length or tension of the vocal cords results in increases in the F0 and therefore leading to a higher-pitched voice (90). It is noteworthy that F0 values vary depending on the individual's gender, the normal values being approximately two times higher in women than in men. Concretely, the normal values are 125Hz for men and 250Hz for women (91). Moreover, it has been suggested that changes in F0 coincide with variations in the vertical thickness of the vocal fold medial surface. Therefore, causing alterations in the spectral properties

of the produced voice (89). Moreover, the thickness of the medial surface is conducted by a coordinated activity between the cricothyroid and thyroarytenoid muscles as well as lung pressures as it has been shown that a change in lung pressures can modify the F0 frequency by 100-200 Hz due to the fact that dynamic stretch increases the thyroarytenoid fibers (92). Due to the fact that these muscles share the innervation, its common activation is frequent, thus frequently leading to an increase of vocal fold approximation.
Figure 17

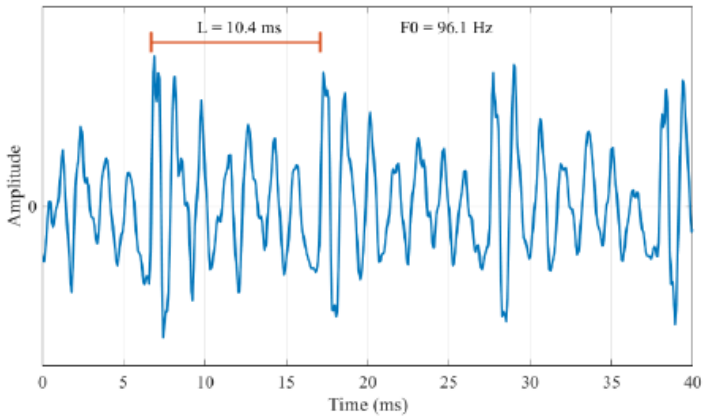


Figure 17. *Segment of a speech signal with the period length L, and fundamental frequency F0= 1/L* (93).

There are several methods that allow the analysis of the fundamental frequency such as subjective methods used by speech therapists; electromyography used to measure muscle activity during voice production through the placement of electrodes on the laryngeal muscles which record the electrical activity during phonation (94). With regard to acoustic analysis, two software programs are commonly used in clinical and research settings. Firstly, the MDVP (Multidimensional voice program) is the most used and cited acoustic analysis software, which was developed by Kay Elemetrics Corporation in the United States (95).

Furthermore, the second most used program is called Praat, which translates into "speak" in Dutch. Praat is likely the most extensive set of tools for phonetic investigation accessible globally and it was designed by Paul

Boersma and David Weenink from the Phonetic sciences Department of the University of Amsterdam and first released in 1995(96). Praat software is distributed for free and supported by many clinicians and scientists (97).

Besides, both software extracts a set of acoustic parameters many of which are defined similarly (98). This acoustic software is able to do the quantitative analysis of F0, Jitter, Shimmer and NHR (Noise-to-harmonics ratio). However, the discriminatory power of these different voice analysis programs when differentiating voices has been questioned, as voice outcomes reported by one analysis acoustic software program might not be comparable to those obtained by another software (99).Moreover, another study supported this suggestion as its results indicate statistically significant differences between the two systems and programs, with the multi-dimensional voice program yielding consistently higher measures than Praat (100). For this reason, the validity and reliability of acoustic analysis performed using different tools have already been shown to be affected by several factors, including the type of microphone, ambient noise levels, data acquisition system, and the software used for the analysis (98, 101).

Finally, another method to assess the F0 as well as other larynx pathologies is laryngostroboscopy. This methodology allows the visualization of the larynx with high resolution and its mucosal wave patterns with greater precision aids in the better understanding of its anatomy and function(102). Moreover, laryngostroboscopy is the state of the art diagnostic tool and provides valuable information about the nature of the vibration and a visual image that can be used for both immediate analysis and as a permanent record for comparison of repeated examination at a later date of vocal cord lesions(103).

Additionally, it is important to consider other parameters in analyzing the human voice such as the formants. To note, a formant is defined by the range of frequencies in which there is absolute or relative maximum in the sound spectrum. That is to say that the frequency at the maximum is the formant frequency(104). Furthermore, the fundamental frequency is the basic but human voice also consists of a resonance system such as the jaws, teeth, lips, tongue, etc. **Figure 18**

As a result, the pharynx shapes the voice signal, and the length and shape of the larynx is ultimately the aspect responsible for amplifying certain harmonics and attenuating the rest. Hence, formants are essentially the peaks in the frequency spectrum that have a high degree of energy. However, only the first four formants are relevant to human hearing (105).

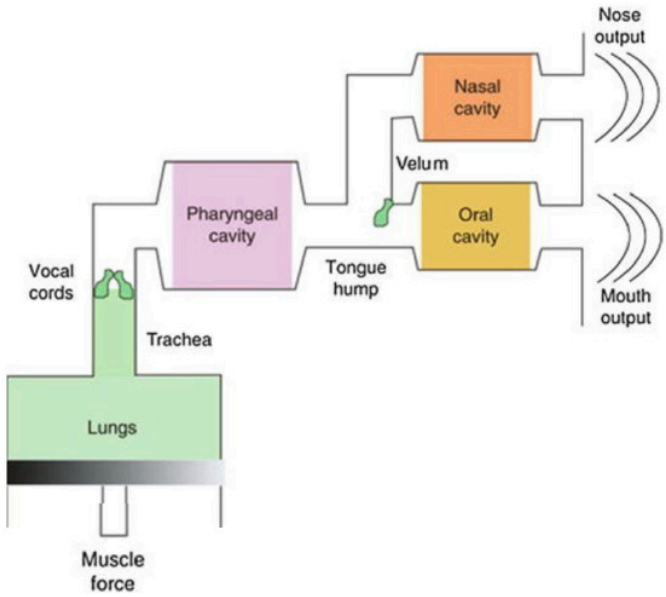


Figure 18. *Voice formation diagram. Pathway from its origin to its exit through the resonating elements* (106).

As previously stated, the shape of the vocal tract can affect the frequencies of the formants, in addition to its length, for this reason, constrictions or dilations in the vocal tract which are mainly influenced by the position of the jaw and tongue, can impact the frequency of all formants in different ways. Moreover, the mandibular opening primarily affects the frequency of the first formant, which increases with opening. The shape of the tongue's body plays a significant role in the frequency of the second formant, while the position of the lingual apex affects the frequency of the third formant (107).

Additionally, another important characteristic of the voice is the harmonics. To further elaborate, the pitch of a sound is determined by its fundamental frequency which is the lowest frequency component of a complex sound

wave. The secondary frequency components that are present in the sound are called harmonics. The first harmonic corresponds to the fundamental frequency and contains the most energy or sound power that we can hear. Then, the frequency of each harmonic is a multiple of the fundamental frequency and are higher pitched sounds than the fundamental frequency (108).

Jitter

Jitter is defined as the parameter of frequency variation from cycle to cycle and has proven to be useful in describing the vocal characteristics. Jitter represents a measure of the stability of phonation and is mainly affected by the lack of control of vibration of the vocal cords; the voices of patients with pathologies often have a higher percentage of jitter (89). For example, a perturbation of F0 higher than usual in a patient can be due to neurological, insufficient glottal closure, or mechanical such as tumors or edema which cause a significant increase in jitter. Moreover, jitter varies greatly between genders as female voices normally display less shimmer but more jitter than male voices (109).

The values for jitter can be measured in different parameters: Absolute, relative, relative average perturbation and the period perturbation quotient.

Jitter absolute:

It is defined as the cycle-to-cycle variation of the fundamental frequency. It is influenced by the F0 and for this reason, it varies depending on the gender.

The formula is the following (89): **Figure 19**

$$jitta = \frac{1}{N-1}\sum_{i=1}^{N-1}\left|T_i - T_{i-1}\right|$$

Figure 19. *Jitter absolute formula* (89).

Relative Jitter:

Is the average absolute difference between consecutive periods, divided by the average period. It is expressed as a percentage (89): **Figure 20**

$$jitter(relative) = \frac{\frac{1}{N-1}\sum_{i=1}^{N-1}|T_i - T_{i-1}|}{\frac{1}{N}\sum_{i=1}^{N}T_i} \times 100$$

Figure 20. *Jitter relative formula* (89).

Jitter (ppq5):

Also known as five-point period perturbation quotient, computed as the average of it and its four closest neighbors divided by the average period. It is also expressed as a percentage (89). **Figure 21**

$$ppq5 = \frac{\frac{1}{N-1}\sum_{i=2}^{N-2}\left|T_i - \left(\frac{1}{5}\sum_{n=i-2}^{i+2}T_n\right)\right|}{\frac{1}{N}\sum_{i=1}^{N}T_i} \times 100$$

Figure 21. *Jitter (ppq5) Formula* **(89).**

2.2.1.2. *Intensity parameters*

The intensity of a sound is the power of the sound in watts divided by the area the sound covers in square meters. That is, the loudness of a sound relates the intensity of any given sound to the intensity at the threshold of hearing and it is measured in (dB)(110). For a normal adult, the intensity of speech during conversation is 75-70 decibels and its values depend mainly on the amplitude of vocal fold vibration and subglottic pressure. Several parameters indicate the level of perturbation of the intensity such as the following:

Shimmer: (dB) is expressed as the variability of the peak-to -peak amplitude in decibels; that is the average absolute base-10 logarithm of

the difference between the amplitude of consecutive periods, multiplied by 20 (89). Furthermore, it is also an acoustic characteristic of a voice signal caused by irregular vocal fold vibration. On top of that, shimmer changes with the reduction of glottal resistance and mass lesions on the vocal cords and is correlated with the presence of noise emission and breathiness (89). Additionally, it can also be influenced by a variety of factors such as loudness, language or gender and it can be altered in pathologies with higher values in voices with severe deviations (111). For example, shimmer may be used to diagnose presence of hoarseness (111). This characteristic is most clearly detected from long sustained vowels (93). **Figure 22**

$$ShdB = \frac{1}{N-1}\sum_{i=1}^{N-1}\left|20 * \log\left(\frac{A_{i+1}}{A_i}\right)\right|$$

Figure 22. *Shimmer (dB) formula. Ai is the extracted peak-to-peak amplitude data and N is the number of extracted fundamental frequency periods* (89).

Shimmer relative is defined as the average absolute difference between the amplitudes of consecutive periods, divided by the average amplitude, expressed as a percentage (89): **Figure 23, figure 24.**

$$Shim = \frac{\frac{1}{N-1}\sum_{i=1}^{N-1}\left|A_i - A_{i+1}\right|}{\frac{1}{N}\sum_{i=1}^{N}A_i} \times 100$$

Figure 23. *Shimmer relative formula. Aiare the extracted peak-to-peak amplitude data andNis the number of extracted fundamental frequency periods* (89).

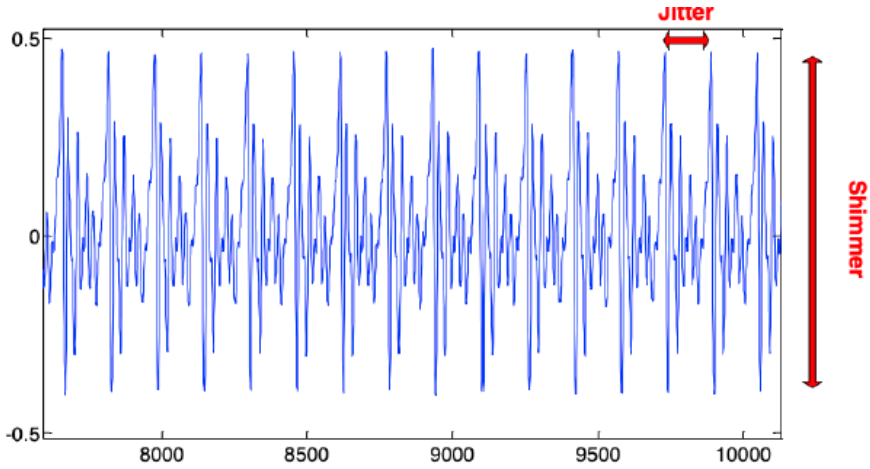


Figure 24. *Jitter and Shimmer perturbation measures in speech signal* (112).

2.2.1.3. *Noise parameters*

Noise measurements are used as indicators of air escape and inadequate glottal closure (111). The most used noise parameters are the following:

HNR (mean Harmonic to Noise Ratio): The Harmonic-to-noise ratio is the measure that quantifies the amount of additional noise in the voice signal. It is quantified in decibels (88) and is the logarithm in base 10 of the ratio between the periodic energy and the energy corresponding to the noise, multiplied by 10 (113). The comparison of the two components provides a measure of speech efficiency, indicating how much of the expelled air from the lungs is converted into energy for vocal cord vibration. Furthermore, a higher HNR indicates a more efficient voice, typically associated with sonorous and harmonic sounds. Conversely, a low HNR denotes an asthenic voice and dysphonia (112)

NNE (Normalized noise energy): This parameter transforms the noise intensity values into a normal distribution. This transformation results in negative values with those closer to zero being more pathological. Moreover, since NNE primarily measures the turbulent noise caused by the closing insufficiency of the glottis during the phonation, it is very useful for the detection of laryngeal disease (114)

2.3. ANATOMICAL AND PHYSIOLOGICAL BASES OF VOICE FORMATION

2.3.1. Main functions of the phonatory system:

The larynx is situated in the upper part of the airway and is an essential part of the respiratory system. Furthermore, it is also referred to as the organ of phonation because of its unique characteristics that enables it to produce voice. Moreover, the larynx essentially functions as a valve with three roles which are the biologically primarily ones: the respiratory and the sphincteric function. Firstly, it allows air to pass through during respiration. Secondly, that of a partially closed valve whose orifice can be modulated in phonation and finally that of a closed valve protecting the trachea and bronchial tree during deglutition (115).

From the embryological point of view, the larynx develops from a double outline: the supraglottic region originated from a buccopharyngeal protrusion, while the glottic and subglottic regions develop from a tracheobronchial protrusion (116).

What's more, the laryngeal cavity is characterized by the presence of two pairs of folds. The upper folds are called vestibular folds whereas the lower folds are called the vocal folds; both folds are separated by the laryngeal ventricles which entail the space between both of them.

Additionally, the vocal cords serve as a reference to separate the laryngeal cavity into three well-defined regions. **Figure 25.**

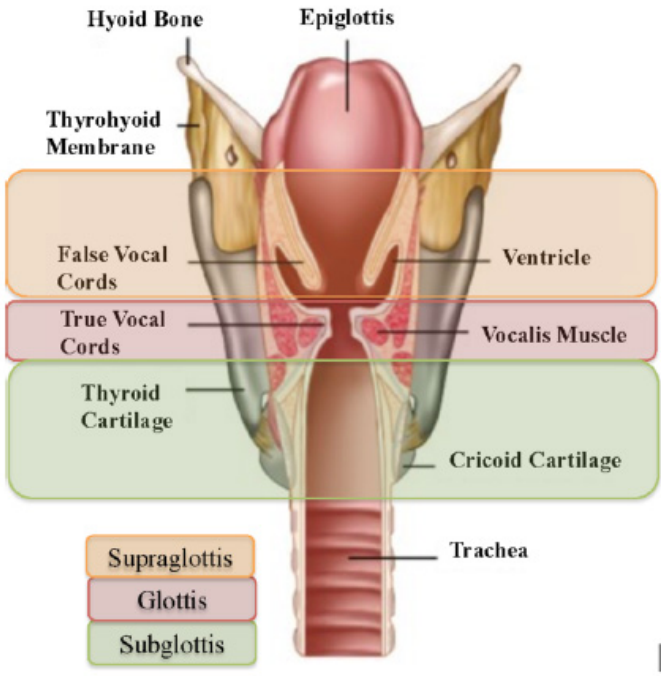


Figure 25. *Regions of the laryngeal cavity* (117).

1) The supraglottis: which communicates with the pharynx. Its upper limit is given by the upper border of the epiglottis, the aryepiglottic folds, the arytenoid cartilages and the interarytenoid fold. The lower border is delimited by the lateral border of the laryngeal ventricle.
2) The glottis: which contains the vocal cords. Their free edges are subdivided into a membranous part which corresponds to the vocal ligament and a cartilaginous part which corresponds to the arytenoids (116).
3) The subglottis: The subglottis extends from 1cm below the free edge of the vocal cords to the lower border of the cricoid cartilage. Its walls are lined by respiratory mucosa and are supported by the cricothyroid ligament above and the cricoid cartilage below (115).

2.3.2. Musculoskeletal framework

The skeletal framework of the larynx is formed by a series of cartilages connected among themselves and with the underlying structures by ligaments and fibrous membranes. **Figure 26.**

Cartilages:

The epiglottis: unlike the rest of the cartilage, it is a fibrocartilaginous structure (118). At its lower end, it is attached to the inner surface of the thyroid cartilage by the thyroepiglottic ligament. The major function of the epiglottis is to assist in preventing aspiration during swallowing.

Thyroid cartilage: It is the largest of the laryngeal cartilages. It has a shield-like or partially open book shape with its concavity facing backward, to protect the laryngeal cavity. The angle formed by the two thyroid cartilage plates show sexual dimorphism. It formats a more acute angle in males than in females which explains the reason why the thyroid cartilage is more anteriorly projected in the neck of males, as well as the greater anteroposterior diameter of the glottis compared to females, which as a result tend to have higher-pitched voices than in men. It is important to mention the presence of the thyroid foramen through which the superior laryngeal vessels together with the internal and external laryngeal nerves in the majority of subjects. (118).

Cricoid cartilage: It is located in the lower part of the larynx, in continuity with the trachea. It is the only complete cartilage of the airway, and for this reason. It acts as a support for the rest of the cartilaginous elements of the larynx. It also corresponds to the narrowest section of the airway. (118)

Arytenoid cartilages: they have a triangular shape and connect with the cricoid cartilage through their articular facet on the lower surface. They have two distinct processes. On the one hand, laterally there is the muscular process and on the other medial and anteriorly there are the vocal ligaments and thyroarytenoid muscles which cause vocal cord movements such as adduction or abduction as well as opening and closure of the vocal cords.

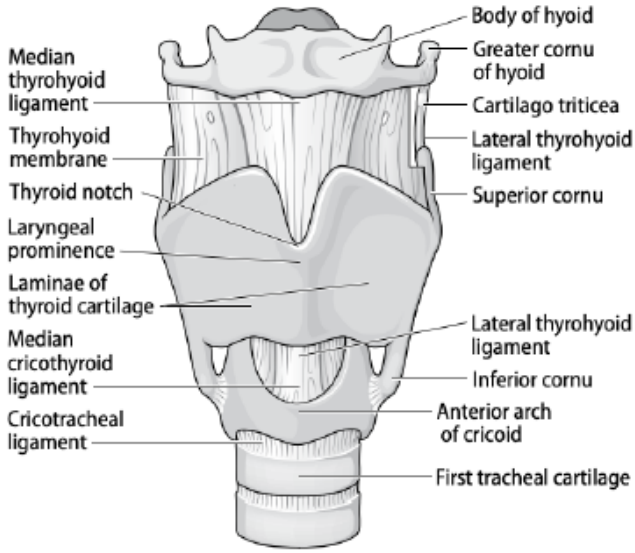


Figure 26. *Anterolateral view of the laryngeal cartilages and ligaments* (118).

Membranes and ligaments:

The larynx also includes membranous and ligamentous structures that participate in the connection of the cartilages to each other and to adjacent structures. These structures can be classified into two categories: the extrinsic ligaments and the intrinsic ligaments.

Firstly, the extrinsic ligaments functionality is to connect the cartilages in order to form the laryngeal structure. These are the following (119):

- Thyrohyoid membrane: which extends from the hyoid bone to the thyroid notch and is part of the anterior boundary of the preepiglottic space.
- Hypoepiglottic membrane: delimits the pre-epiglottic space with the thyroid cartilage and the vallecula.
- Aryepiglottic ligament: contributes to the structure of the vestibule.
- Cricothyroid membrane: which extends from the superior border of the cricoid cartilage to the inferior border of the thyroid cartilage.
- Cricotracheal ligament: which extends from the inferior border of the cricoid cartilage to the first tracheal ring.

Secondly, anterior ligaments which play a key role in the laryngeal function are the intrinsic ligaments which connect the cartilages of the larynx to each other. These ligaments are the following (119).

Quadrangular membrane: it is located between the lateral surfaces of the arytenoid cartilages and the epiglottis. The upper edge of the membrane forms the aryepiglottic ligament which is part of the aryepiglottic fold and contains the corniculate cartilages also known as accessory cartilages. Its lower edge is formed by the vestibular ligament which is part of the vestibular fold.

Elastic cone: which ascends from the cricoid cartilage to the vocal cords.

Vocal ligament: located in the border of the vocal cord, residing between the mucosa and the underlying muscle.

Larynx muscles:

The muscles of the larynx can be divided into intrinsic and extrinsic muscles with two well-differentiated functions.

Firstly, the intrinsic muscles are responsible for altering the length, tension, shape and spatial position of the vocal folds by changing the orientation of the muscular and vocal processes of the arytenoid cartilages with the fixed anterior commissure. As a result of these movements, the glottis is opened during inspiration, closed during phonation and closed with supraglottic reinforcement during deglutition (115). These are the following: Cricothyroid, cricoarytenoids, intrarytenoid, aryepiglottic, thyroarytenoid, thyroepiglottic muscles. Additionally, depending on their functions the muscles be classified into: **Figure 27**

- Adductors muscles: lateral cricoarytenoid, interarytenoid, lateral throarytenoid and are the muscles responsible for the approximation of the vocal cords.

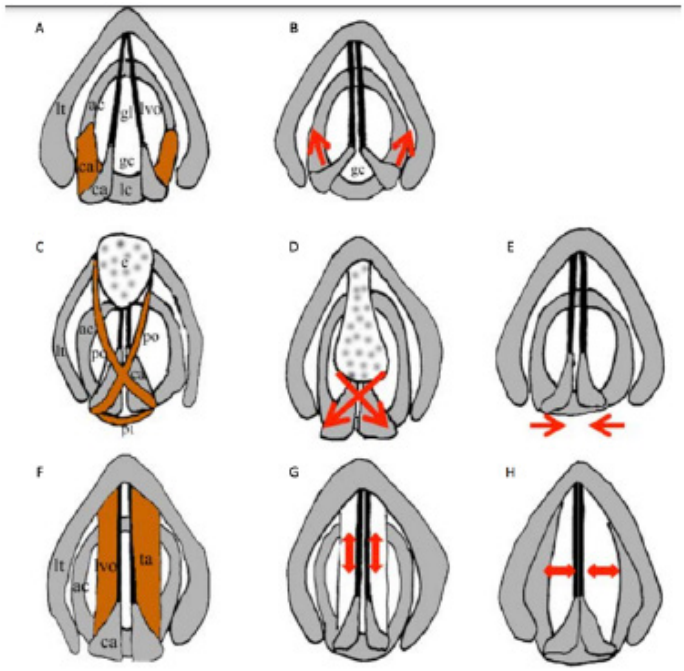


Figure 27. *Top view of the muscles on the left, location and on the right action. Lateral cricoarytenoid muscles (A and B), arytenoid muscles (C and D) and thyroarytenoid muscles (F and H). ac, cricoid arch; ca, arytenoid cartilage; cal, lateral cricoarytenoid muscle; e, epiglottis; gc, cartilaginous glottis; gl, ligamentous glottis; lc, cricoid lamina; lt, thyroid lamina; lvo, vocal ligament; po, oblique portion; pt, transverse portion; ta, thyroarytenoid muscle* (118).

- Abductor muscles: Posterior cricoarytenoid which is responsible for the separation of the vocal cords.
- Approximation of thyroid and cricoid cartilages: Cricothyroid muscle.

With regard to the extrinsic muscles of the larynx it is important to note that its main function is to connect the larynx to the adjacent structures and are responsible for moving it vertically during phonation and swallowing. This extrinsic muscular apparatus is composed of the suprahyoid muscles (digastric, stylohyoid, mylohyoid, geniohyoid and hypoglossus) which are involved in the elevation of the larynx during swallowing and the infrahyoid muscles such as: thyrohyoid, sternothyroid, omohyoid, sternohyoid that depresses the larynx during phonation. **Figure 28**

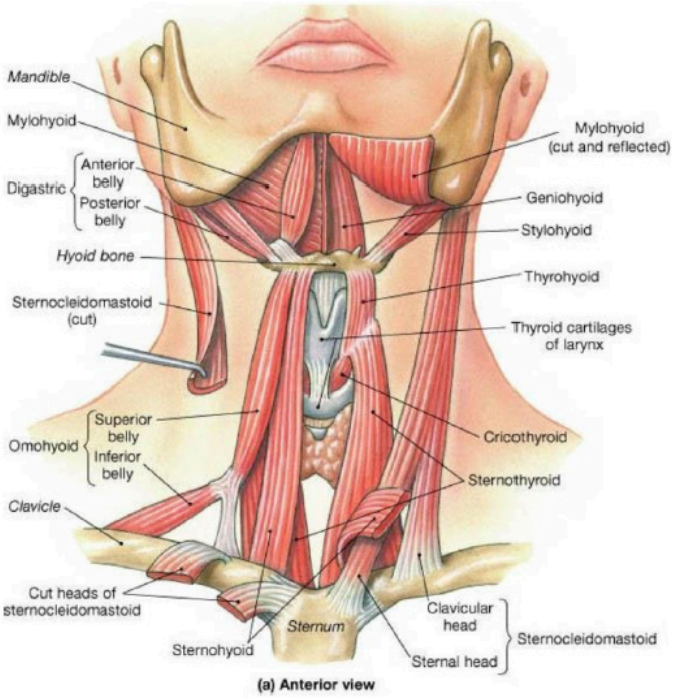


Figure 28. *Extrinsic muscles of the larynx* (120).

Innervation of the larynx (116):

The larynx is innervated through the external branch of the superior laryngeal nerve, as well as the recurrent nerves which are branches of the vagus nerve. Additionally, the superior laryngeal nerve divides into two branches. An internal branch responsible for sensory innervation above the glottis and an external branch which supplies motor fibers to the cricothyroid muscle. Moreover, the recurrent nerve innervates the ipsilateral intrinsic muscles of the larynx and provides contralateral fibers to the interarytenoid muscle. In addition, it also carries sensory fibers for the laryngeal mucosa below the glottis. What's more both recurrent nerves enter the larynx at the level of the inferior horn of the thyroid cartilage. **Figure 29.**

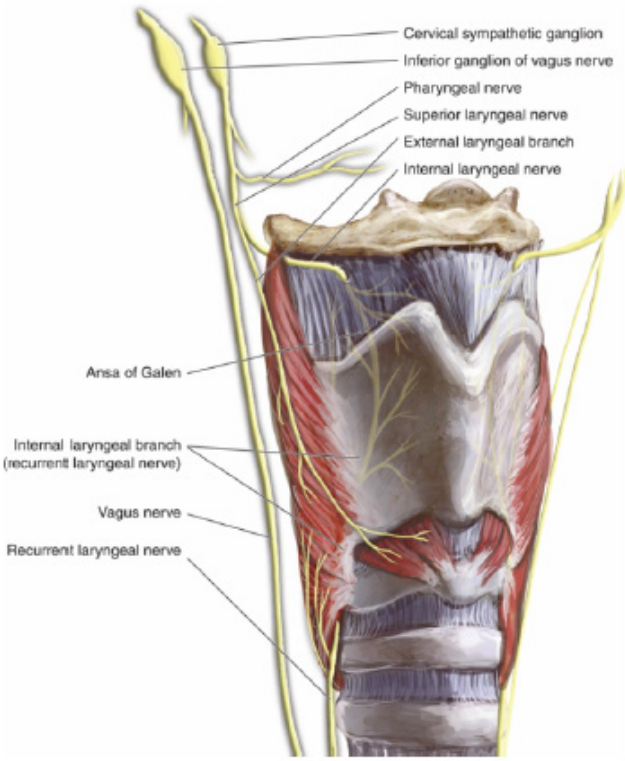


Figure 29. *Innervation of the larynx* (121).

Larynx irrigation

The larynx receives its blood supply primarily from two arteries: the superior laryngeal artery, originating from the superior thyroid artery, branch of the external carotid artery, and the inferior laryngeal artery, which arises from the inferior thyroid artery, branch of the thyrocervical trunk from the subclavian artery. The venous drainage from the larynx is carried out through the superior thyroid vein and drains to the internal jugular vein (115). **Figure 30.**

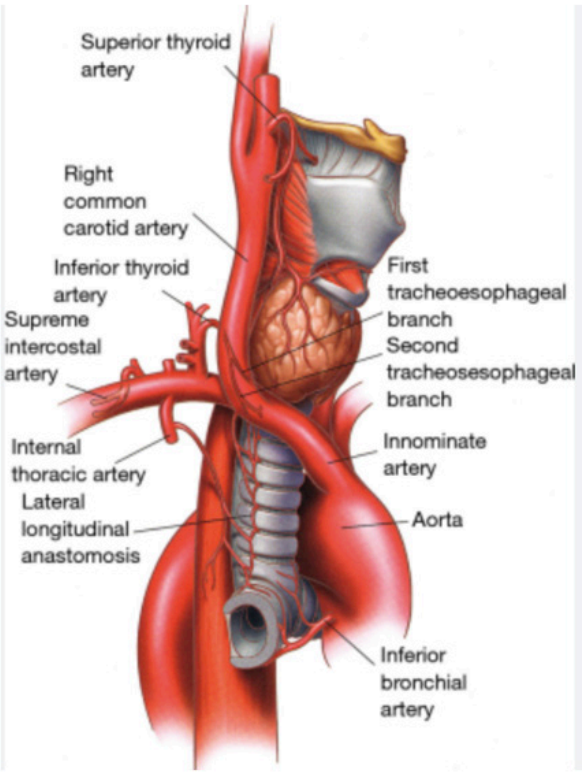


Figure 30. *Irrigation of the larynx* (122).

2.4. THE VOCAL CORDS

Mammalian vocal cords are generally multi-layered in their tissue construction. The epithelial structure and organization of the vocal cords are highly suitable for facilitating vibration, providing protection and preserving the shape influenced by the intrinsic muscles of the larynx (118). The epithelium recovers a structure known as the lamina propia which contains three layers. The superficial layer, the intermediate layer and a deep layer (85). Furthermore, the lamina propia is built of collagen fibers with interstitial fluids such as proteoglycans and glycoproteins which have important biological and biomechanical effects (123). From these proteoglycans mostly hyaluronic acid and, to a lesser extent, fibronectin (124) are the main contributors to viscosity and therefore, play a key role in oscillation tissue properties. **Figure 31**

Furthermore, the presence of a mucociliary layer on the vocal cord serves the dual purpose of protecting it from the strain of vibration and facilitating it through lubrication. Moreover, two mechanisms for self-sustained vocal fold vibration have been proposed. Firstly, a mucosal wave propagates along the medial surface of the vocal cords and secondly a vocal tract that offers inertive reactance (125). As the person ages, the extracellular matrix experiences a decline in viscosity and elasticity due to a decrease in protein turnover. Besides, the surface of the epithelium becomes thinner and altered in surface architecture with aging. These surfaces are less robust and more vulnerable to the effects of atrophy of the mucus glands (126).

The vocal cords are situated within the larynx and each vocal fold is about 11-15mm long in adult women and 17-21mm in men (127).

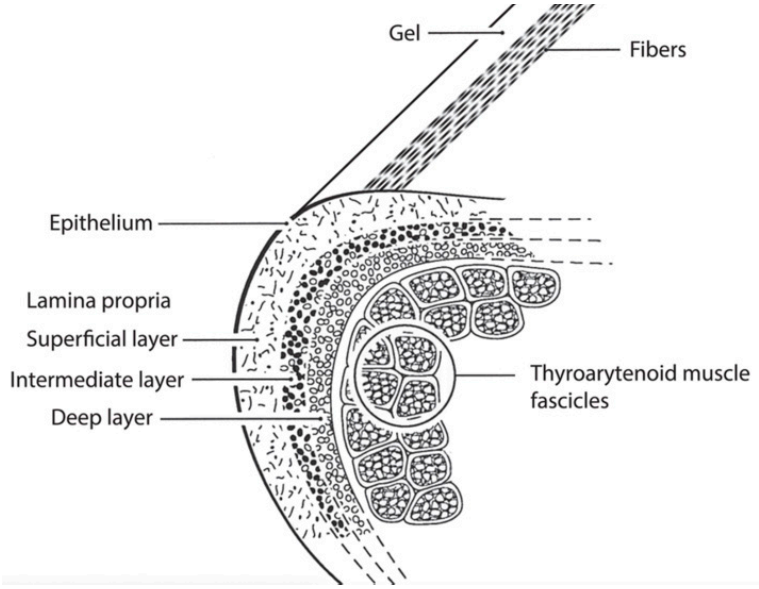


Figure 31. *Vocal folds structure* (128).

2.5. THE PHONATION CYCLE

The phonation cycle can be divided into two different phases: firstly, the open phase and then the closed phase. With regard to the open phase, the glottis continuously opens and closes whereas in the closed phase the glottis is either completely closed or remains partially open when the closure is

incomplete. Furthermore, in the open phase, the glottal flow increases and decreases, while in the closed one, the glottal flow remains at zero or at a minimum for incomplete closure (127).

2.6. THEORIES OF VOICE FORMATION

Some theories exist which explain the different aspects of the sound formation during phonation such as the myoelastic theory, the aerodynamic theory and the Bernoulli one. These theories complement each other and are combined in order to facilitate a better understanding of the sound production of the human voice.

2.6.1. Myoelastic- aerodynamic theory

This theory was first described by Van den Berg in 1958 (129). It suggests that the fundamental frequency of vocal fold vibration is determined by three factors: firstly, the mass of the vocal cords; secondly, the viscoelasticity of the vocal folds and finally, subglottic pressure. Moreover, they propose that the vocal folds vibrate due to a series of forces explained by the Bernoulli principle which evaluates the amount of mechanical energy stored at a particular point among the streamline (130).

The myoelastic aspect of phonation control refers to the neuromuscular control of vocal fold tension and elasticity. According to this theory, the vocal folds approximate, contract and tense during phonation to regulate their elasticity. Besides, the coordination of subglottic pressure and vocal fold elasticity is believed to be key in regulating voice production as in addition to regulating vocal fold tension and elasticity, neuromuscular control also adjusts the configuration of the glottal opening. Additionally, the tridimensional structure of the glottis allows variations in subglottic and supraglottic areas which is an essential component of the motor aerodynamic force (118).

The aerodynamic aspect helps explain the role of fluid dynamics in the initiation of vocal emission and is based on three principles:

In the first place, the air flows from areas of high pressure to areas of low pressure; Secondly, according to Bernoulli's principle, the pressure of an

incompressible fluid decreases as the velocity of its molecules increases; and finally, the velocity of molecules in an incompressible fluid confined within a duct increases as the cross-sectional area narrows, following the equation of continuity. Additionally, in order for phonation to be initiated the vocal cords must come together to form a narrow channel and the air needs to be expelled from the lungs., therefore, increasing the pressure between the vocal folds sufficiently so that the air pressure separates both cords allowing the trachea airflow through the opening and phonation is initiated. As the airflow passes through the channel, the velocity of its molecules increases which causes at the same time a reduction of pressure ultimately closing the vocal folds again. It is important to note that during all this process of voice formation the vocal folds' vibration and deformability are an important piece of input that contributes to aerodynamics (131) of voice formation. Concretely, the fluid-structure interaction between the laryngeal airflow and the vocal folds is accountable (131) and seems to be involved in the characterization of laryngeal airflow and the vocal fold vibration mechanism.

CHAPTER 3. ARTIFICIAL INTELLIGENCE (AI)

3.1. MACHINE LEARNING (ML) AND DEEP LEARNING (DL)

The term artificial intelligence (AI) was first coined by McCarthy, a twenty-eight-year-old undergraduate at the Faculty of Mathematics in Dartmouth in 1955(132). The definition of artificial intelligence is still controversial mainly due to the fact that its central nucleus, that is; intelligence, is still difficult to define. Firstly, the difficulty lies in the fact that intelligence is complex and has a multidimensional nature, involving the capacity to comprehend, reason, learn, adapt and ability to solve problems. However, if we consider its multidimensional nature, the definition of intelligence might encompass a border sense depending on each context or approach. For example, human intelligence is built on a combination of the different dimensions of intelligence such as cognitive, emotional and social skills which contribute to increasing the difficulty in finding an accurate definition. Nevertheless, artificial intelligence has taken another nuance and focused on two efforts: one scientific and another practical. On the scientific side, AI researchers are investigating the mechanisms of intelligence by trying to embed it in computers. On the practical side, they simply want to create computer programs that perform tasks as well as or better than humans without worrying about whether these programs are actually thinking in the way humans think (132).

The ongoing process in technology and advancements in artificial intelligence and as machines and algorithms increasingly achieve sophistication in performance of cognitive tasks and problem-solving abilities, there is a growing debate surrounding this technology and its capability classification as genuine intelligence or simply imitations of it. As an example of this debate and the potential scope of this technology is captured in the following book: " The singularity is nearer" published in 2024 and written by Ray Kurzweil, a noted inventor and futurist in which states his

vision of singularity "When AI empowered by its ability to improve itself and learn on his own will quickly reach and then exceed human-level intelligence" (133).For this reason, it is important to clarify the types of intelligences that exist within this technology, the scope and numerous applications such of each of them. Moreover, it is important to note that in recent years, AI has gained importance in the healthcare field, specifically in diagnosis and drug development. For example, a computation field known as "virtual screening" has emerged in the past decades to aid experimental drug discovery studies by statistically estimating unknown bio-interactions between compounds and biological targets (134). **Figure 32**

Furthermore, in order to fully understand the magnitude of this technology and its applications it is important to focus on and explain in detail and in-depth what these types of technologies consist of. Firstly, machine and deep learning are revolutionary fields in the computer science area and are widely used in business applications. What's more, machine learning is an approach to train computers and machines to learn from past data so it can determine future data or behavior. (135)

Generally speaking, AI comprises any technique that enables computers to mimic human behavior and reproduce or excel over human decision-making to solve complex tasks independently or with minimal human intervention (136). In short, AI's aim is to create intelligent machines that can simulate human thinking capability behavior and ultimately, human intelligence. The artificial intelligence systems do not require to be pre-programmed, alternatively they use machine learning algorithms capable of using their own intelligence. Moreover, this type of intelligence is nourished from machine learning algorithms and its applicability can be found in Siri, AlphaGo and Google among others.

Furthermore, AI can be classified into three categories:

1) Weak AI.
2) General AI.
3) Strong AI.

However, currently only the use of weak and general AI is available. Strong AI is thought to be developed in the near future and it may very well overcome human intelligence.

Additionally, AI foundations are based on data analytics such as big data which is generally considered linkable information that has large data volumes and complex data structures such as social media, mobile phone call records and commercial website data (137). More concretely, data mining plays a crucial role in the field of AI as it incorporates techniques for the analysis of these large volumes of data which is the infrastructure that allows AI functions such as pattern recognition or training of AI models. For this reason, data mining is considered the core stage of the knowledge discovery process that aims to extract interesting and potentially useful information from data (138).

Besides, in early stages of AI the main emphasis was on utilizing hard-coded statements written in formal languages which could be processed by applying logical inference rules to reason and draw conclusions. However, due to the fact that humans have intuitive or implicit knowledge and skills that may be difficult to communicate explicitly, this might be limiting the ability of systems to comprehend certain tasks that rely on such tacit knowledge (139). Nevertheless, machine learning overcomes this limitation as it enables machines to learn from past data or experiences and improves its performance with experience. This is achieved by applying algorithms that iteratively learn from problem-specific training data, which allows computers to find hidden insights and patterns (139), therefore, enabling a computer system to make predictions.

Furthermore, machine learning is limited to specific domains. For instance, if we train a machine learning model to recognize dog pictures it will only provide accurate results for dog images. However, if we introduce new data, such as a cat image it won't be able to recognize it, due to the fact that it only works with data that has been shown before. Furthermore, its main applications can be found in Google search algorithms, email spam filters and Facebook's automatic friend tagging's suggestions.

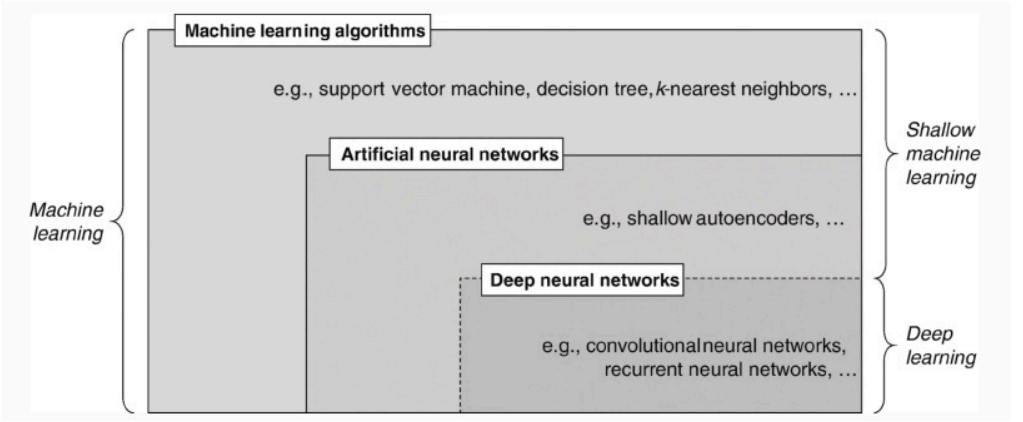


Figure 32. *machine learning concepts and classes* (139).

On top of that, the main difference between AI and machine learning is that machine learning is a subset of AI whose goal is to train a machine to perform a task and allow them to give an accurate output focusing on pattern recognition, whereas AI's aim is to make smart computer systems which solve complex problems and enables them to perform any task like a human, entailing a wide range of scope.

Another distinguishing factor between both of them is that AI includes reasoning, learning and self-correction conversely, machine learning doesn't include reasoning, only learning when introduced with new data. Finally, the two technologies can also be distinguished based on the kind of data that they can process; while AI is able to deal with unstructured data, machine learning can only analyze structured or semi-structured data.

With regard to machine learning, it is important to highlight that considering the given problem that needs to be solved and the available data, this technology can be classified into three distinct types.

- Supervised learning.
- Reinforcement learning.
- Unsupervised learning.

3.1.1. Supervised learning:

Supervised learning is a machine learning paradigm for acquiring input-output relationship information of a system based on a given set of paired input-output training samples (140) The objective of supervised learning is to construct an artificial system capable of learning the relationship between input and output and making predictions on new inputs. Additionally, the main difference between supervised and unsupervised learning is that the first one needs supervised or labeled information whereas in the second one labeling is not necessary. The foremost advantage of supervised learning is that it can be easily used for discriminative pattern classification and data regression. On the other hand, as a drawback it may experience difficulties with managing a great amount of data as it makes labeling difficult (140).

3.1.2. Reinforcement learning:

Deep reinforcement learning is a branch of AI which combines reinforcement learning with neural networks. For example, it is used when it is preferable to learn directly from measurements of raw video or image data without any hand-engineered features or domain rules (141). Instead of using the traditional learning algorithms, it uses multiple layers of neuronal networks that replicate the structure of a human brain. Besides, this learning process involves the utilization of trial-and-error techniques in which the ability to make optimal decisions in interactive environments is acquired.

Artificial neural networks (ANN) are the roots of deep learning, and it encompasses units that act in serial or in parallel receiving inputs from other units. Overall, ANN forms a hierarchical structure, from neurons, to layers and ultimately a network formation. Additionally, its depth depends on how many layers the network is composed of. Traditionally, if the network is formed by more than five layers it is possible to coin the term “deep”. Moreover, ANNs are able to reproduce the dynamic interaction of multiple factors simultaneously, allowing the study of complexity; they can also draw conclusions on an individual basis and not as average trends (142). **Figure 33**

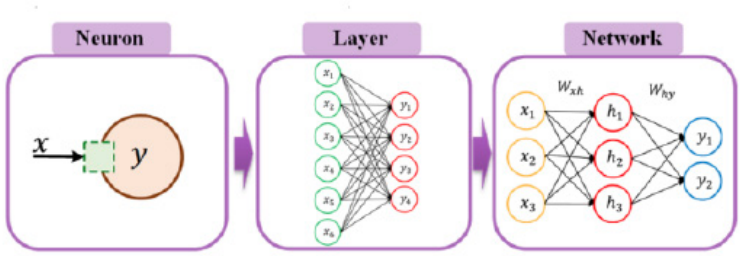


Figure 33. *Hierarchical structure of an ANN* (141).

3.1.3. Unsupervised learning

Unsupervised learning involves the training of a model in an unlabeled dataset. The model learns on its own by learning the features of the training dataset and based on those makes predictions on test data (143). In detail, it means that in these kinds of techniques there is a lack of information of the variable to be predicted. Furthermore, in order to achieve this, these procedures explore structures, patterns or features within the source data that can be replicated in new datasets. Several unsupervised learning approaches and algorithms range from clustering, K-means and principal component analysis (PCA).

3.2. MACHINE LEARNING METHODOLOGY

In the methodology of creating predictive models, it is necessary to have a specific framework on the process of analytical model building programming in order to achieve a model capable of predicting with high accuracy and reliability.

In the first place, the process starts defining the requirements for the model, determining the datasets to be utilized for training and validation purposes and selecting the performance metrics that will be employed to evaluate the performance of the final model. Thus, the key component of the model building is to break down the whole process into six phases: business understanding, data understanding, data preparation, modeling, evaluation and deployment (144). Besides, it is important to note that the process involves iteration, where the knowledge gained from previous phases informs and improves subsequent iterations. **Figure 34.**

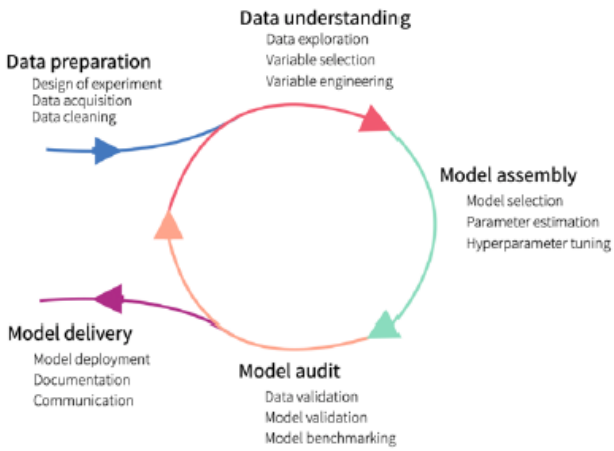


Figure 34. *The life-cycle of a predictive model* (145).

The first step is data preparation which involves the identification of various data sources, their collection, acquisition and integration. However, this might not be a straightforward process, as extracting patterns and relationships by hand would exceed the cognitive capacity of human operators which is why algorithmic support for the processing of data of different types, such as cross-modal learning is crucial when dealing with large databases (139). These algorithms are useful for data cleaning and sample selection which may include tasks like dealing with outliers or unbalanced or heterogeneous data through oversampling the minority class, undersampling or segmentation (144). As a result of this preparation of a consistent set of data, also known as a dataset, is obtained.

Secondly, the next step is data understanding. During this stage, some tasks are needed to understand the relationship between dependent and explanatory variables, deal with the missing values or duplicate data as well as feature selections.

Data understanding is necessary in order to decide which variables shall be included in the model (144). This is important as the goal of feature extraction is to retain valuable discriminatory information which is significant to the overall learning objective and therefore, being able to eliminate noise which can sometimes confuse the model. Moreover, it may require the

encoding or the normalization of some variables for a better interpretation and comparison of the importance of different features. **Figure 35**. In detail, hot encoding is a data preparation practice that makes certain kinds of data easier to work with or actually readable by an algorithm and refers to the process of expanding and splitting categorical data (146).

"Nationality" Feature		
Categorical	Integer Encoding	One-Hot Encoding
UK	0	[1, 0, 0]
French	1	[0, 1, 0]
USA	2	[0, 0, 1]

Figure 35. *Example of a different encoding of a categorical feature* (146).

Furthermore, there is the model assembly phase. In this phase some parameters through specific strategies need to be set in order to construct the model.

Finally, the last step is the model audit in which the data and the model are validated. In order to do that, the dataset needs to be divided into three different subsets: training set 60%, validation set 20% and test set 20%. On the one hand the training set is used to adjust the different machine learning classification algorithms and to directly improve the model's parameters. Once this has been achieved the validation subset is used to evaluate a model's performance while optimizing the model's hyperparameters. Basically, the aim of this step is to assess the quality of the model by maximizing the specific metric that is most significant in our particular scenario. For example, the ROC curve. Typically, this training-validation process is iterated multiple times using randomized subsets, known as k-folds. **Figure 36**

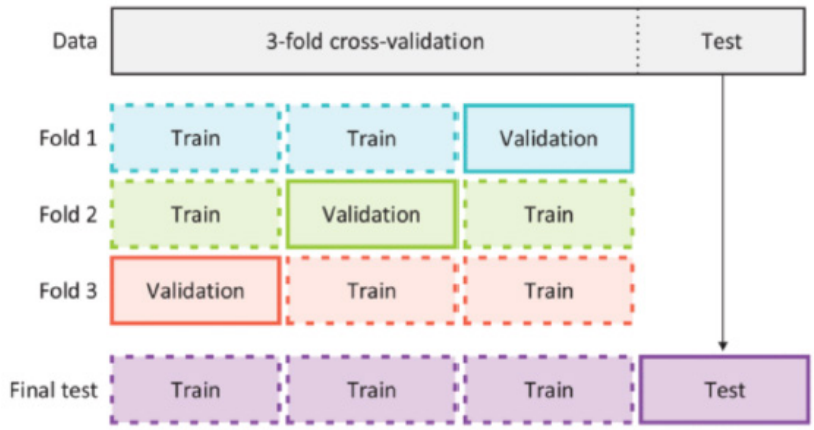


Figure 36. *Example of 3-folds cross- validation* (147).

The cross validation consists of dividing the data into different K parts (referred to as k-folds). One part (fold) is held out as the validation set. The model is trained on the remaining K-1 parts and then applied to the validation set and records its predictive performance (148). This process is iterated K times ensuring that each portion serves as a validation set once. The purpose is to optimize the internal parameters of the algorithm, evaluate the model's robustness and assess whether our model is experiencing overfitting, which is to see how well our model might perform on data it has not seen before. Concretely, overfitting means that your algorithm works well with the training dataset but not well with the test dataset.

What's more, the cross-validation approach ensures that the results are reliable and not influenced by the specific division between the training and testing data.

Finally, once the final model is completed the model needs to be tested to make sure that it performs as expected with data that has not been used for its construction and validation.

3.3. DESCRIPTION OF COMMONLY USED METRICS IN THE EVALUATION OF CLASSIFICATION ALGORITHMS:

In classification studies, many researchers employ performance metrics to demonstrate their success. However, there is a prevalent confusion in the literature regarding the terminology and a lack of understanding of the fundamental aspects underlying these metrics (149). The main reason for this confusion is that comparing the different classifiers using the traditional metrics such as sensitivity, specificity may lead to some difficulties as these metrics are not sensitive to imbalance data and might ignore the performance of the minority class.

Therefore, other metrics such as the ROC curve are important for an appropriate assessment of the classification model.

Firstly, when assessing the quality of a binary classification model there are several metrics, we can use such as the following:

Accuracy is defined as the division between the number of correct answers and the total of the answers. Basically, it provides us with the number of times our model accurately classifies an item in our dataset in relation to the overall total. However, it is well known that accuracy is an inappropriate metric for rare event classification problems such as medical diagnosis and fraud detection (150). This is mainly due to the fact that when classes are unevenly distributed as accuracy is the ratio of correct. This is mainly due to the fact that accuracy is the ratio of correct answers to the total and in the case that the classes are unevenly distributed our model might wrongly classify each sample in the dataset, thus failing to predict the minority class. For this reason, alternative metrics better tailored to imbalanced classification such as the F1 are employed. Nevertheless, when dealing with balanced datasets accuracy is a sensible metric to use. **Figure 37.**

$$Acc = \frac{TP + TN}{TP + TN + FP + FN}$$

Figure 37. *Accuracy formula. In where: TP True positive), TN (true negative), FP (False positive), FN (False negative)* (151).

Sensitivity is also called as true positive rate or recall and represents the positive correctly classified samples to the total number of positive samples (151). **Figure 38**

$$TPR = \frac{TP}{TP + FN} = \frac{TP}{P}$$

Figure 38. *Sensitivity or True positive rate formula* (151).

Specificity, which is also called true negative rate, precision or inverse recall and is expressed as the ratio of the correctly classified negative samples to the total number of negative samples(151). In short, this metric tells us how often we are correct when we classify a class as a positive. **Figure 39**

$$precision = \frac{(TP)}{(TP + FP)}$$

Figure 39. *Specificity or True negative rate formula* (151).

As it is difficult to have a model with both high precision and recall as the two metrics are complementary, another metric is used called F1 score.

F1 score combines precision and recall into one metric and is the most used metric for evaluating binary classification models. Furthermore, it represents the harmonic mean of precision and recall. The value of F-measure is ranged from zero to one, and high values of F-measure indicate high classification performance (151). **Figure 40**

$$F - measure = \frac{(\beta^2 + 1) * precision * recall}{\beta^2 * precision + recall}$$

Figure 40. *F1 score formula* (152).

Log loss score: is used to measure the performance of a model by using the probability of the expected outcome. The higher the probability of the actual class is, the higher the log loss will be (153). That is to say, this metric measures the differences between the probabilities of the model's predictions and the probabilities of observed reality. Besides, the obtention of lower scores indicate better performance of the model.

Receiver operating characteristics (ROC) curve metric, is a two-dimensional graph in which the true positive rate (TPR) represents the y-axis, and the false positive rate (FPR) is the x-axis. Moreover, it has been used to evaluate many systems such as diagnostic systems, medical decision-making systems, and machine learning systems (151). On top of that, the assessment of a classifier depends on the position it occupies in the ROC curve. Therefore, any classifier which appears in the lower right triangle performs worse than the classifier that appears in the upper left triangle (151). **Figure 41.**

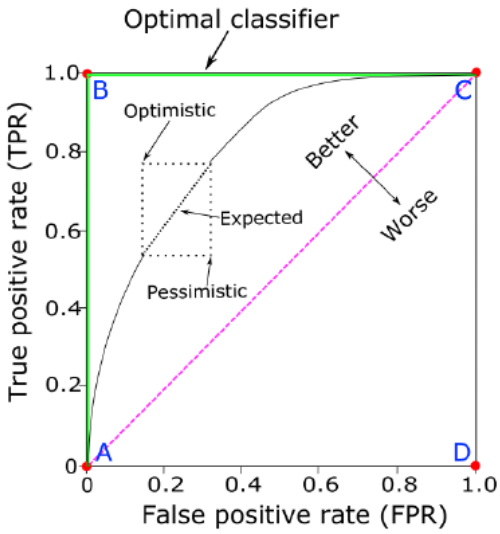


Figure 41. *ROC curve* (151).

In summary, the AUC provides a single measure of a classifier's performance for evaluating which model is better on average (154). Moreover, some authors state that the ROC curve appears to be one of the best ways to evaluate a classifier's performance on a dataset as it exhibits a number of desirable properties when compared to overall accuracy (155).

Finally, the confusion matrix is a tool for predictive analysis in machine learning which is used to assess the performance of a classification model. It summarizes the classification performance of a classifier with respect to some test data in a two-dimensional matrix. It is able to show how many predictions are correct and incorrect per class and it helps in understanding the classes that are being confused by the model as another class (156). **Figure 42.**

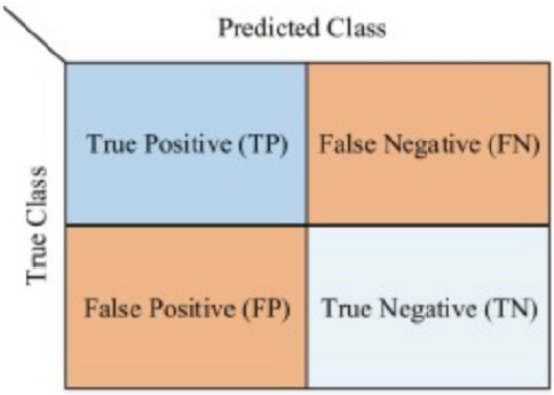


Figure 42. *Confusion matrix* (157).

In the confusion matrix we can also see the other classification metrics such as the following:

- True Negatives (TN): consists of items which have been correctly classified as negative.
- True Positives (TP): consists of items which have been correctly classified as positives.
- False Negatives (FN): consists of items which have been incorrectly classified as negative and are therefore actually positive.
- False Positives (FP): consists of items which have been incorrectly classified as positive and are therefore actually negative.

3.4. MOST COMMONLY USED ALGORITHMS IN MACHINE LEARNING

Figure 43

3.4.1. Naïve Bayes:

This is a popular algorithm for classification and regression in predictive modeling and is based on the Bayes theorem, therefore, it involves a statistical-based approach to learning (143). In detail, the algorithm is used to determine the probability of an event occurring based on previous events occurring which is called posterior probability (153).

3.4.2. Support vector machines (SVM)

Support Vector Machines is a supervised learning discriminative algorithm that uses a hyperplane to separate the training data to classify future predictions. The hyperplanes divide a dataset into two classes and they are decision boundaries that help classify the data points (153). Furthermore, it has become very popular since it is both used for regression and classification and its strengths include its ability to provide a good out-of-sample generalization (158). On top of that, it is an excellent tool when working with high dimensional data (143).

3.4.3. Decision Tree (DT)

Decision tree (DT) is a supervised learning classification technique that creates a model which anticipates the value of a target variable depending on input values (158). Moreover, it can be used for both classification and regression problems. Precisely, DT is a graphical representation of the outcome in which the roots (nodes) represent the tests and attributes, the branches show the results of the tests, and the leaves represent the class distributions (158).

Figure 43. *Widely used supervised classification algorithms* (143).

3.4.4. Random Forest

Random forest is a supervised learning algorithm that is seen to be an improvement on the decision tree model (153). The reason for this is that contrary to a single decision tree which may have the tendency to overfit data, it uses multiple decision trees independently trained on a randomly selected subset of the data set and then their results are aggregated for a more accurate prediction. Particularly, each individual tree in the random forest gives a class prediction and the class with the most votes among the trees, becomes the model's predicted class (158).

3.4.5. K-Nearest neighbors (KNN)

K-Nearest Neighbors (KNN) algorithm is another supervised learning method. It makes an assumption that similar things exist in close proximity which implies that similar things are close to each other. For this reason, the most frequent label for data classification is then considered the winning class for the data sample selected (158). Furthermore, the K-Nearest neighbor is

a very simple, highly efficient and effective technique, which is used in many applications like data mining, text categorization or object recognition (159).

3.4.6 Generalized linear models (GLM)

Generalized linear models are used for classification tasks, linear regression and logistic regression. Additionally, the linear model assumes that the conditional expectation of the dependent variable Y is equal to a linear combination of the explanatory variables X. However, this assumption may not hold in many practical scenarios. For this reason, generalized linear models overcome this limitation by introducing a link function which enables the modeling of nonlinear relationships or different variable distributions (160).

3.5. MACHINE LEARNING AND HEALTHCARE APPLICATIONS

In recent years, there has been a boom and transformation in the field of healthcare due to the incorporation of technologies such as machine learning with predictive algorithms. These algorithms are recently being implemented and gradually becoming consolidated as useful tools that assist us and have numerous applications in our daily practice, particularly as a diagnostic tool.

Concretely, in the field of anesthesia machine learning has emerged as a solid resource to use. Connor et al, carried out an integrative review in which it explains that the introduction of artificial intelligence and machine learning into anesthetic practice will be the routine intraoperative management of patients which will be handed off to closed-loop control algorithms (161). In detail, Connor et al states that utilizing a stable anesthetic as an initial implementation is advantageous since the algorithms are not required to make diagnosis but rather to identify deviations from the predetermined control parameters established by the anesthesiologist. Additionally, other researchers have also started to use this technology to improve their clinical practice or to solve a particular problem. For instance, Lee et al (162) developed a deep learning model which predicted bispectral index during target-controlled infusion of propofol and remifentanil more accurately compared to the traditional model and stated that the deep

learning approach in anesthetic pharmacology seemed promising due to its excellence performance and extensibility.

Furthermore, it is important to emphasize the landmark of this statement as it prompted its mention in another article carried out by Gambus et al,(163) which questioned whether the application of machine learning would mean the end of clinical pharmacology. This question is justified by the evidence that machine learning offers a stronger and more comprehensive approach in contrast to the complex modeling methods that clinical pharmacologists have traditionally depended on for many years.

What's more, focusing more specifically on the field of airway management more scenarios of the applicability of this technology can be found. For example, Kim et al (164) aimed to create and validate a machine learning model that utilizes neck circumference and thyromental height as predictors of a difficult laryngoscopy. Moreover, Yamanaka et at (165) also applied machine learning to predict difficult airways and first-pass success in a multicenter prospective study in emergency departments. Besides, Connor et al (166) derived a computer model to predict difficult intubation based on the analysis of photographs of patient's faces and concluded that this kind of analysis was able to classify easy versus difficult intubation and outperformed popular clinical predictive tests.

On top of that, the use of facial analysis to predict the airway has also been the objective of other researchers such as Tavolara et al (167) who proposed an ensemble of convolutional neural networks which learned robust features of multiple facial regions with the aim to develop a model for the identification of difficult to intubate patients. As a result, the proposed model outperformed conventional bedside tests.

Furthermore, Hatib et al (168) explored another application of this technology through the development of a machine-learning algorithm to predict hypotension based on high-fidelity arterial pressure waveform analysis, which was able to detect the incipient onset of hypotension up to 15 minutes before its onset.

In conclusion, these are just some examples that illustrate the potential of artificial intelligence in healthcare. As the technology continues to develop its

ability to enhance medical diagnosis, solve medical challenges and improve treatment is likely to become increasingly evident. For this reason, artificial intelligence has the potential to positively transform healthcare.

2. RESEARCH QUESTION

2. RESEARCH QUESTION

Is it possible to develop a voice-based tool using machine learning algorithms that can classify patients based on their Cormack and Mallampati grades and thereby predict a difficult intubation?

3. OBJECTIVES

3. OBJECTIVES

3.1. MAIN OBJECTIVE:

Objective 1: Develop a tool using machine learning algorithms to classify patients according to their Cormack and Mallampati grade and therefore predict a difficult intubation.

3.2. SECONDARY OBJECTIVES

Objective 2: Demonstrate the relationship between the anatomy of the upper respiratory tract and the following voice parameters: harmonics, jitter, shimmer, HNR and spectral.

Objective 3: Demonstrate the correspondence between the analysis of acoustic voice parameters and the grade of the Cormack Classification scale.

Objective 4: Demonstrate the correlation of the voice parameters with traditional predictive tests such as the Mallampati.

Objective 5: Obtain a combination based on voice parameters which is able to achieve high AUC values for Cormack prediction together with both the clinical and anatomical characteristics of the patient such as age, sex, weight, comorbidities and other preexisting patient pathologies.

Objective 6: Obtain a combination based on voice parameters which is able to achieve high AUC values for Mallampati prediction together with both the clinical and anatomical characteristics of the patient such as age, sex, weight, comorbidities and other preexisting patient pathologies.

Objective 7: Define to what extent the voice parameters contribute to the prediction of a difficult intubation in terms of AUC values, when introduced into the different algorithm combinations in comparison to the descriptive data.

Objective 8: Demonstrate if it is possible to find a tool based on voice and descriptive data of the patients with comparable predictive power to the Mallampati test which may enable its replacement during the airway assessment in the preanesthetic visit.

Objective 9: Use the knowledge obtained from this study to develop an app that allows for preoperative online or telemedicine assessment of the airway.

Objective 10: To explore the potential of future applications such as telematic preanesthetic consultations.

Objective 11: Determine the usefulness and effectiveness of this newly developed tool and its need for further validation.

4. METHODS

4. METHODS

4.1. STUDY DESIGN

The present study is observational, prospective, descriptive and multicentered. Furthermore, two different groups of patients were enrolled in two different sub-studies:

- The Mallampati study: N=594 patients were enrolled in the Mallampati sub-study. 27 of them were included in Hospital Universitari Dexeus, and the rest 567 were recruited in Centro Medico Teknon)
- The Cormack study: N=313 patients were enrolled in the Cormack sub-study. 25 of them were included in Hospital Universitari Dexeus and the rest 288 patients were recruited in Centro Medico Teknon)

This is an open study design as the type of intervention being evaluated did not need the use of blinding techniques. The sequence of assignments to the study groups was performed sequentially and randomly as patients were recruited from the pre-anesthetic consultation previously referred by the surgeon. Subsequently, in the pre-anesthetic consultation, patients were assigned an anonymous alphanumeric code for their follow-up and inclusion in the study. **Figure 44.**

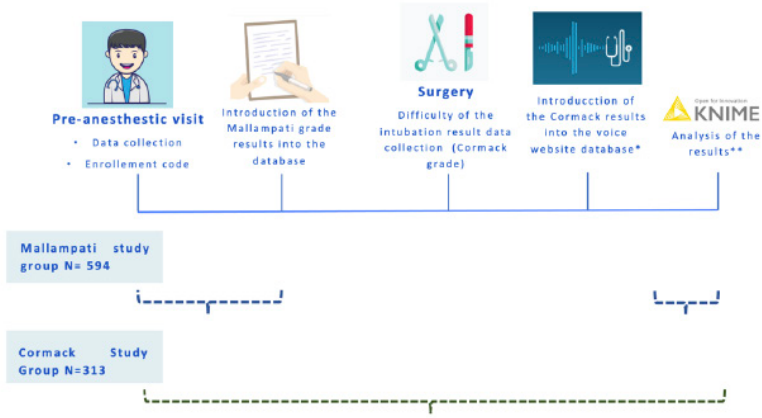


Figure 44. *Study design and follow-up of participants. * voice.anestestalia.com database, ** KNIME®: Open-source platform built to productionize data science.*

4.2. SCOPE OF THE FACILITIES

The present study was ~~a~~ multicentered conducted in two different hospitals. The first one, is the Centro Medico Teknon, Quironsalud group in Barcelona, Spain. The second hospital is Hospital Universitari Dexeus, Quironsalud group also in Barcelona. Both hospitals are private and belong to Quironsalud Group, the leading company in Spanish private healthcare which represents one of the most important groups of private hospitals nationally. Furthermore, both hospitals are highly regarded for their quality management and commitment to patient safety, recently reaffirmed by their reaccreditation from the Joint Commission International. Precisely, the Joint commission international, is an independent, not-for-profit organization which defines, measures and shares best practices in quality and patient safety around the world(169) .

Moreover, Centro Médico Teknon includes 225 beds, 31-day hospital boxes and 20 operating rooms. It is well-equipped to handle a wide range of medical specialties and serve patients of all ages. This center holds approximately 36.000 anesthesia procedures each year and during the year 2022, approximately 38.300 anesthetic procedures were performed, out of which 18.300 were surgical interventions. In detail, these interventions were divided among specialties in the following manner: firstly, the medical

specialization with the highest number of surgeries was traumatology and orthopedics accounting with 5.180 surgeries, followed by general surgery with 2.943 procedures, urology with 1.759 surgeries, gynecology with 1.474 interventions and cardiac catheterization laboratory with 1.100 procedures among the notable specialties. It is also important to note, the significant number of very high, high and moderate complexity surgeries being 1.32%, 19.47% and 69.50% respectively. Finally, only 9.5% of the procedures were of low complexity. **Figure 45.**

Figure 45. *Centro Médico Teknon, Quirón Salud Group, Barcelona, Spain.*

Moving onto Hospital Universitari Dexeus, it also stands out for being a benchmark in many specialties and one of the first private university hospitals in Spain. Particularly, it is important to highlight its significant dedication to the training of postgraduates and aspiring medical residents (MIR), as it offers a residency training program in anesthesiology, obstetrics and orthopedics. Moreover, in this hospital more than 10.000 surgical procedures and 3.000 deliveries are performed each year. On top of that, it offers a range of facilities such as the following: 126 individual rooms and 4 royal suites, 13 operating rooms equipped with the latest technology, an obstetric area with 7 delivery rooms, 2 operating rooms and a separate emergency area, 10 individual boxes of day hospital for outpatient treatments and an intensive care unit of 10 beds and neonatal intensive care unit including 21 beds, among others. These facilities ensure that the hospital is well-equipped to provide comprehensive medical services. **Figure 46.**

Figure 46. *Hospital Universitari Dexeus; Quironsalud Group Barcelona, Spain.*

4.3. SAMPLE SIZE

In the field of artificial intelligence research sample size calculations are not commonly carried out and are not as standardized as in traditional research studies based on statistical parameters. For this reason, many studies involving machine learning do not directly predict a sample size. This is mainly due to the fact that learning algorithms deal with large datasets rather than collecting a specific sample as such algorithms learn patterns and make predictions based on the available data to generalize patterns and relationships. Moreover, high quality and representativeness of the inputs that form up the dataset rather than the specific sample size are prioritized. It is worth mentioning here that due to the nature of the functioning of these algorithms and technologies, larger databases are not always necessarily better as sometimes an excess of information might lead to confusion of the model and therefore end up being detrimental.

Since this is an exploratory study and a pilot test in order to calculate the sample size, the sample size used in similar studies available in the literature that investigated the difficulty of intubation using machine learning technology was taken as a reference.

4.4. NUMBER OF PARTICIPANTS

Initially, a total of N=719 patients were included in the study. However, N=125 patients were excluded. The main reason for this exclusion was firstly due to an incorrect voice recording and secondly, due to the fact that in some cases preoperative data was missing. As a result, a total of N=594 patients were selected for the Mallampati study and N=313 patients were recruited for the Cormack study. **Figure 47.**

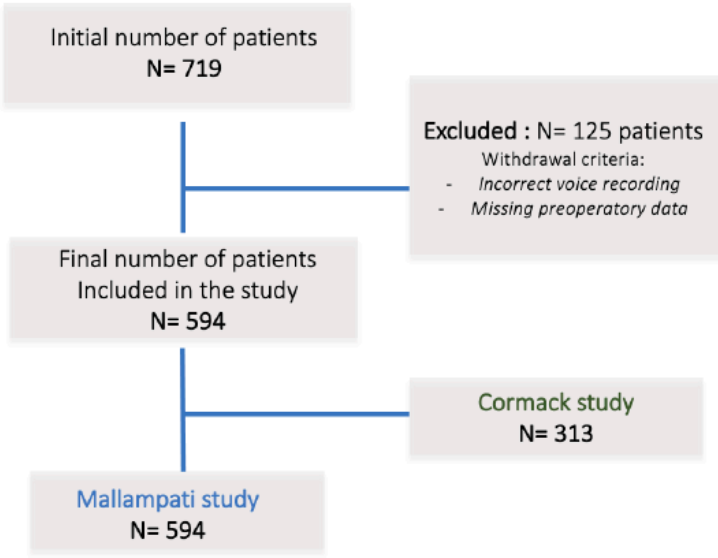


Figure 47. *Flowchart of patient selection.*

4.5. INCLUSION AND EXCLUSION CRITERIA OF PARTICIPANTS

4.5.1. Inclusion criteria:

- Adults over 18 years.
- ASA physical status I-VI.
- Scheduled for an intervention or surgical procedure in need of orotracheal intubation by direct laryngoscopy.

4.5.2. Exclusion criteria:

- ASA physical status >IV.
- Patients aged under 18 years old.
- Emergency procedures.
- Procedures without need of orotracheal intubation.
- Patients who refused to participate in the study.
- Patients who refused to sign the informed consent.

4.6. VARIABLES

4.6.1. Main variables

The main variables are shown in **table 2.**

Table 2. *Main variables*

Variables	Characteristics / Units
Voice recording: Harmonics (20 harmonics / vocal/ position) (Hz). Quantitative continuous	• A Normal position: H1, H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, H16, H17, H18, H19, H20 • A Extension position: H1, H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, H16, H17, H18, H19, H20 • A flexion position: H1, H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, H16, H17, H18, H19, H20 • E Normal position: H1, H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, H16, H17, H18, H19, H20 • E Extension position: H1, H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, H16, H17, H18, H19, H20 • E Flexion position: H1, H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, H16, H17, H18, H19, H20 • I Normal position: H1, H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, H16, H17, H18, H19, H20 • I Extension position: H1, H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, H16, H17, H18, H19, H20 • I Flexion position: H1, H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, H16, H17, H18, H19, H20

Table 2. *Main variables*. Continuation

Variables	Characteristics / Units
Voice recording: Harmonics (20 harmonics / vocal/ position) (Hz). Quantitative continuous	• O Normal position: H1, H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, H16, H17, H18, H19, H20 • O Extension position: H1, H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, H16, H17, H18, H19, H20 • O Flexion position: H1, H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, H16, H17, H18, H19, H20 • U Normal position: H1, H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, H16, H17, H18, H19, H20 • U Extension position: H1, H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, H16, H17, H18, H19, H20 • U flexion position: H1, H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, H16, H17, H18, H19, H20
Mallampati scale grade Qualitative, ordinal	Grade I
	Grade II
	Grade III
	Grade IV
Cormack scale grade Qualitative, ordinal.	Grade I
	Grade IIa
	Grade IIb
	Grade III
	Grade IV

4.6.2 Secondary variables:

The secondary variables were collected during the preanesthetic visit and are shown in the following **table 3:**

Table 3

Variable	Type of variable	Classification / units
Age	Quantitative discrete	Years
Weight	Quantitative continuous	Kilograms (kg)
Gender	Qualitative nominal dichotomous	Male/ Female
Height	Quantitative continuous	Centimeters (cm)
Body Mass Index (BMI)	Quantitative continuous	Kg/m^2
American Society of Anesthesiologists status (ASA)	Qualitative ordinal	I
		II
		III
		IV
Smoker	Qualitative nominal Dichotomous	Yes / No
Sleep Apnea Syndrome	Qualitative nominal Dichotomous	Yes / No
COPD	Qualitative nominal Dichotomous	Yes / No
Diabetes	Qualitative nominal Dichotomous	Yes / No
Thyroid disorders	Qualitative nominal Dichotomous	Yes / No
History of previous difficult intubation	Qualitative nominal Dichotomous	Yes / No

Airway pathology	Qualitative nominal Dichotomous	Yes / No
Inter-incisor gap/ mouth opening	Qualitative ordinal	>=5 cm
		3.5-5 cm
		< 3.5 cm
Jaw protrusion grade	Qualitative ordinal	A
		B
		C
Thyromental distance	Qualitative ordinal	>=6.5 cm
		<6.5 cm
Neck extension range	Qualitative ordinal	>100°
		90+/- 10°
		<80°
Arne Index	Qualitative ordinal	>= 11
		<11

Voice recordings		
Fundamental frequency	Quantitative continuous	Fundamental frequency (F0) (Hz) • f0A_ median_1: Fundamental frequency vocal A in normal position • f0A_median_2: Fundamental frequency vocal A in extension position • f0A_median_3: Fundamental frequency vocal A in flexion position • f0E_ median_1: Fundamental frequency vocal E in normal position • f0E_median_2: Fundamental frequency vocal E in extension position • f0E_median_3: Fundamental frequency vocal E in flexion position

Table 3. Continuation

Voice recordings		
Fundamental frequency	Quantitative continuous	Fundamental frequency (F0) (Hz) • f0I_ median_1: Fundamental frequency vocal I in normal position • f0I_median_2: Fundamental frequency vocal I in extension position • f0I_median_3: Fundamental frequency vocal I in flexion position • f0O_ median_1: Fundamental frequency vocal O in normal position • f0O_median_2: Fundamental frequency vocal O in extension position • f0O_median_3: Fundamental frequency vocal O in flexion position • f0U_ median_1: Fundamental frequency vocal U in normal position • f0U_median_2: Fundamental frequency vocal U in extension position • f0U_median_3: Fundamental frequency vocal U in flexion position • f0A_iqr_1: range interquartile A normal position • f0A_iqr_2: range interquartile A extension position • f0A_iqr_3: range interquartile A flexion position • f0E_iqr_1: range interquartile E normal position

Table 3. Continuation

Voice recordings		
Jitter	Quantitative continuous	• jitterabsA_1: A normal position • jitterabsA_2: A extension position • jitterabsA_3: A flexion position • jitterabsE_1: E normal position • jitterabsE_2: E extension position • jitterabsE_3: E flexion position • jitterabsI_1: I normal position • jitterabsI_I: I extension position • jitterabsI_I: I flexion position • jitterabsO_1: O normal position • jitterabsO_2: O extension position • jitterabsO_3: O flexion position • jitterabsU_1: U normal position • jitterabsU_2: U extension position • jitterabsU_3: flexion position
Shimmer	Quantitative continuous	• shimmerA_1: A normal position (dB) • shimmerA_2: A extension position (dB) • shimmerA_3: A flexion position (dB) • shimmerE_1: E normal position (dB) • shimmerE_2: E extension position (dB) • shimmerE_3: E flexion position (dB) • shimmerI_1: I normal position (dB) • shimmerI_2: I extension position (dB) • shimmerI_3: I flexion position (dB) • shimmerO_1: O normal position (dB) • shimmerO_2: O extension position (dB) • shimmerO_3: O flexion position (dB) • shimmerU_1: U normal position (dB) • shimmerU_2: U extension position (dB) • shimmerU_3: U flexion position (dB)

Table 3. Continuation

Voice recordings		
HNR (Harmonic to noise ratio)	Quantitative continuous	• HNRA_1: A normal position (dB) • HNRA_2: A extension position (dB) • HNRA_3: A flexion position (dB) • HNRE_1: E normal position (dB) • HNRE_2: E extension position (dB) • HNRE_3: E flexion position (dB) • HNRI_1: I normal position (dB) • HNRI_2: I extension position (dB) • HNRI_3: I flexion position (dB) • HNRO_1·: O normal position (dB) • HNRO_2: O extension position (dB) • HNRO_3: O flexion position (dB) • HNRU_1: U normal position (dB) • HNRU_2: U extension position (dB) • HNRU_3: U flexion position (dB)

4.7. MATERIALS

The following is a detailed description of the materials used during the different phases of the study. Multiple instruments were used for data collection during the preoperative visit, as well as the intraoperative period and the subsequent data analysis.

Firstly, the online preoperative visit platform developed by the anesthesia department was used to create the patient's own medical record. Accordingly, in this platform all the information and the demographic data of the patients was collected. **Figure 48**

Figure 48. *Preanesthetic visit platform. Source Preanestesia2.anestalia.com*

Secondly, for the voice recording and the storage of the results after intubation the voice.anestalia.com website was used. This website was developed by the biomedical engineering team of the department specifically for voice study. **Figure 49**

Figure 49. *Home menu website: voice.anestalia.com*

Moreover, for the voice recording during the preanesthetic visit the headphones New Bee H360 model equipped with a 3.5mm connector and detachable USB plug, compatible with PC, laptop, Android phones, supporting Windows 2000/7/8/10/XP/ vista, Mac OSX, iOS, Android were used, providing a sample frequency from 44100 to 48000Hz. **Figure 50**

Figure 50. *New Bee H360 headphones model.*

In order to carry out the data analysis, the KNIME® analytics platform version 4 was used. KNIME® analytics platform is an open-source software with an intuitive visual interface that allows users to build analysis of any level of complexity focused on predictive modeling and machine learning without any coding needed (170). **Figure 51**

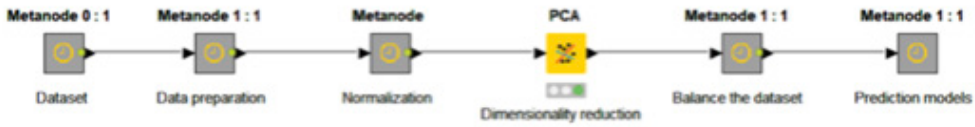


Figure 51. *The developed KNIME® workflow.*

4.8. INTERVENTION

4.8.1. Preoperative phase

During the preoperative phase, patients arrived at the pre-anesthetic consultation referred by the surgeon. During the preoperative visit, the standard clinical interview for the preoperative assessment was conducted using the department's preanesthetic platform for medical records. Patients were asked about their age, weight, height and demographic characteristics

as well as relevant medical conditions and previous surgical interventions. Additionally, a physical examination and evaluation of the airway were routinely performed using traditional bedside tests prior to the surgical procedure. **Figure 52, 53**

Figure 52. *Preanesthetic evaluation form (Example). Source: Preanestesia2.anestalia.com.*

Figure 53. *Airway physical assessment performed during the preanesthetic visit.*

On top of that, all valuable data and variables needed for the clinical study were automatically extracted and stored in a database called MySQL. This process of storing data containing any patient information was done in a completely anonymous manner and did not allow directly or indirectly the identification or association of the data with the patient's identity.

At the end of the clinical interview if they met the inclusion criteria, patients were asked to participate in the study. The participants were provided with all

the information related to the study both orally and written and any doubts or questions that arise in relation to the study were resolved. What's more, if they agreed to participate, they were asked to sign the informed consent and were included in the study.

Next the selected participants of the study were marked in the pre-anesthesia platform indicating their participation in the "Voice" study. Afterwards, in order to ensure traceability while preserving patients' privacy at the same time, an alphanumeric code was automatically assigned to every patient which served to track them throughout the rest of the phases of the study. **Figure 54. (Annex 4)**


UNITAT DE PREANESTÈSIA
Vilana 12, Edifici Marquesa
Despatx 61-62
08022 Barcelona
Ext. 4611 / Telf. 93 290 64 61
preoperatori@anestalia.com

Servei Central d'Anestesiologia
CENTRO MÉDICO TEKNON

CÓDIGO DEL ESTUDIO:
TK0179

Cirujano: Dr./Dra. No Especificado
Fecha de la intervención: 04/02/2020

CONSENTIMIENTO INFORMADO DOCUMENTO DE INFORMACIÓN AL SUJETO

PARTICIPANTE DEL ESTUDIO DE INVESTIGACIÓN

CONSENTIMIENTO INFORMADO

Código del estudio: VOICE2019
Versión del protocolo: V02
Fecha de la versión: 14 septiembre 2019
Fecha de la presentación: 19 septiembre 2019
Título del Proyecto: "Análisis de la voz como método de predicción de una vía aérea difícil de manera preoperatoria"

Directoría del Proyecto: Dra. Borrás, Dra. Valls
Investigador/a: Dra. Claudia Rodiera
Departamento: Anestesiología y reanimación

Yo, el Sr./la Sra: .....[redacted].....

Figure 54. *Informed consent of the study. It contains an alphanumeric code which was assigned to each patient included in the study.*

At this point, the anesthesiologist or the anesthesia nurse conducting the pre-anesthetic visit was automatically redirected to the website platform: *voice.anestalia.com* in order to proceed with the voice recording.

To continue with the voice recording, the patient was asked to sit down with a straight back position, facing the computer screen with any mal position of the body. They were asked to put on the headphones and locate the speaker at a 5cm distance from the mouth. **Figure 55.**

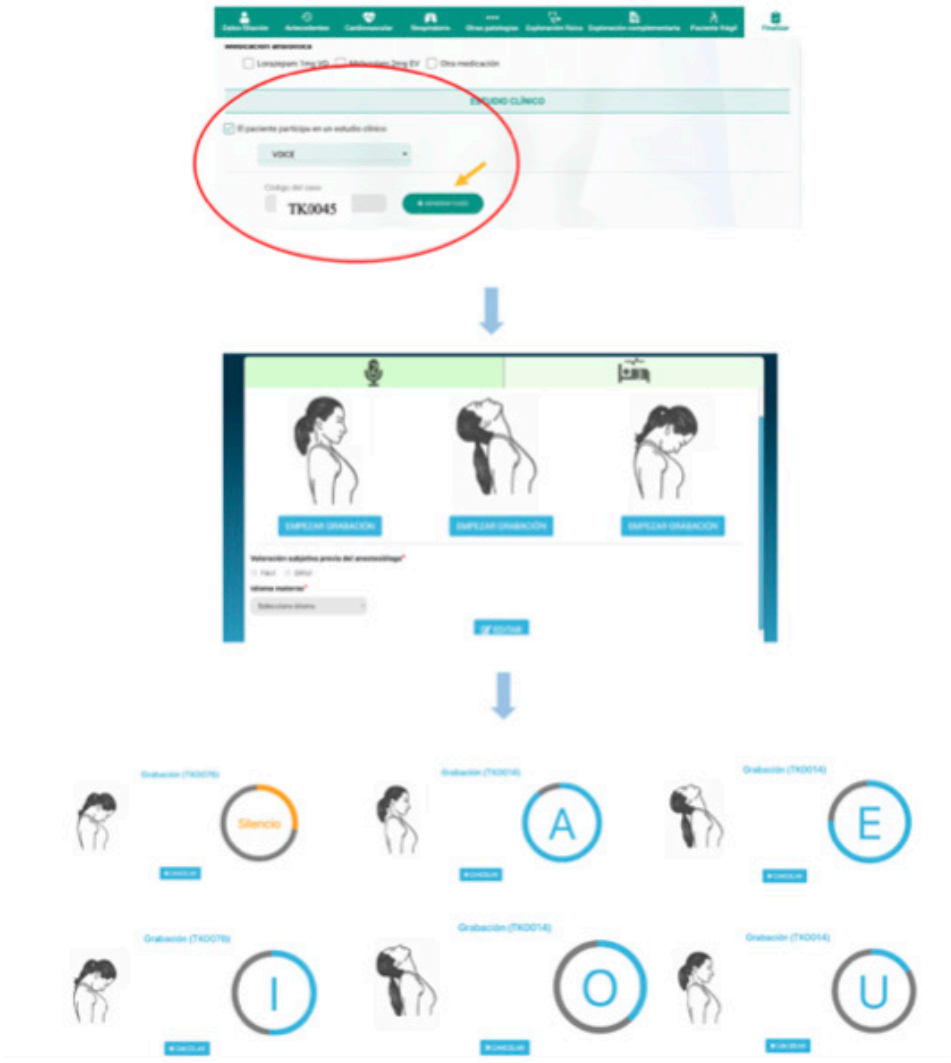


Figure 55. *Voice recording flowchart. First picture: case code generation for the inclusion of the patient in the study. Second picture: website voice.anestalia.com (Own production) designed for the voice recording. Third picture: different voice recording positions and vocals.*

At this moment, prior to the initiation of the voice recording a calibration of the system was automatically carried out and the patients were instructed to remain silent for 5 seconds during the adjustment of the system. Shortly afterwards, the participants were encouraged to articulate the vocals A, E, I, O, U for 3 seconds at three different positions of the head: normal, extension and flexion. Moreover, an automatic authentication system was configured to detect those audios poorly recorded, thus ensuring that the stored audios met the necessary requirements in order to be adequately analyzed later. If an error in the recording was detected, the system automatically indicated the need for a re-recording of the audio until it was properly recorded and stored. This recording was done without any interruption and in a closed room free from noise. If any interruption or noise occurred (for example, patient coughing or a knock on the door), the recording was repeated until ensuring that it was done properly. **Figure 56** Then, the voice recordings were encoded and stored in a secured and anonymous database in which only the principal investigator and the biomedical engineer had access.

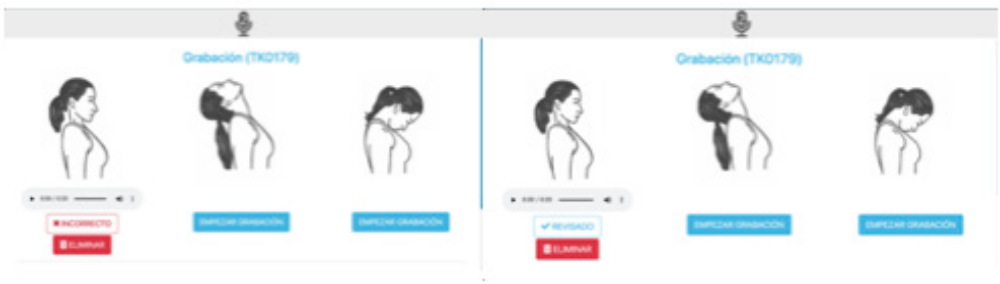


Figure 56. *Automatic filter which detects the poorly recorded audios and identifies them as 'Incorrect" indicates the need to re-record them until they meet the necessary requirements. Identified as 'reviewed'.*

Following that, the voice signal was processed, and parameters were extracted using R2022b Matlab® and Simulink® by an expert in signal processing biomedical engineering.

4.8.2. Intraoperative phase of the study

During the intraoperative phase of the study, the Cormack-Lehane grade was recorded during the intubation phase of the anesthesia.

Firstly, all the safety checklists were completed, and all the equipment was verified and the anesthesia plan was organized. Afterwards, premedication with a sedative was administered and a proper preoxygenation of the patient was performed. The anesthesia-inducing drugs were then added, and adequate neuromuscular relaxation was ensured. Prior to ventilation with a facemask, intubation was performed using direct laryngoscopy with a good alignment of the axes and a correct head position. The Cormack grade was evaluated for the anesthesiologist and recorded in the data collection sheet. **Figure 57** Afterwards, the data collection sheet was deposited in the operating room mailbox. **(Annex 5).**


ESTUDIO VOICE

CÓDIGO DEL ESTUDIO: TK001 CirujanoDr/Dra:
Fecha intervención:

1. Marcar con un círculo el valor de Cormack obtenido mediante
LARINGOSCOPIA DIRECTA CON **PALA METÁLICA**

I IIA IIB III IV

Figure 57. *Data collection form.*

4.8.3. Post operatory phase:

In this phase, the main investigator collected all the data sheets from the mailbox and introduced the data in the *voice.anestalia.com* platform in order to store the information electronically. **Figure 58** This data was stored according to the code in the data sheet and was linked to the voice recording of the patient with the same code.

Figure 58. *Data collection platform. Voice.anestalia.com*

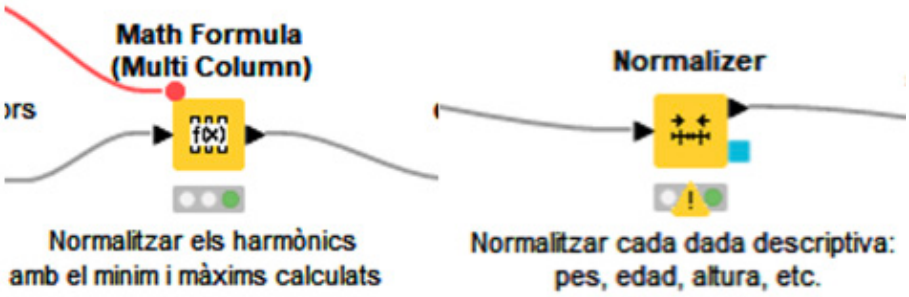
Once all the data was collected, the results were introduced and analyzed. Firstly, the voice signal was processed using MATLAB® during which the frequency amplitude of each vowel was extracted with a sampling frequency of 48.000Hz. Since the minimum number of samples calculated for all recorded vowels was 24481 samples, a duration of 0.5s was analyzed for each vowel. Subsequently, when the signal was transformed from the time domain to the frequency domain, the analyzed frequencies were half the sampling frequency (24000 Hz). Therefore, 12243 samples of the power spectrum of the signal in the frequency domain were obtained. This information was saved in a folder with 15 different Excel files (one for each vowel). An attempt was made to analyze this data with KNIME® platform. However, this platform does not have enough capacity to process such a large volume of data. For this reason, another attempt was made to store the information in another database, but it was also unable to process such a high volume of data. For this reason, it was decided to return to MATLAB® and extract the information of each vowel from the frequency spectrum, but with a reduced sampling frequency of 9600 Hz. The minimum number of samples for each vowel is 4897 samples and 500ms of each vowel were analyzed.

However, when the signal was transformed from the time domain to the frequency domain, the analyzed frequencies were up to half the sampling frequency (4800 Hz). A total of 2451 samples of the frequency spectrum amplitude of each vowel were obtained from 0 to 4800 Hz. This reduction in the sampling frequency allowed for much fewer data points for each vowel. This information was stored in 15 different Excel files.

Nevertheless, when using the KNIME® program it was found that even with the reduced number of data points the reading of 15 Excel files hindered smooth manipulation of the program. All the data was imported to MySQL in 15 tables through the KNIME® program. Once the data was stored in the MySQL database, a new KNIME® file was created, and the frequency spectrum data was read through MySQL. As a result, the program was able to operate faster without the presence of the Excel files.

A workflow was created using the KNIME® program with the aim to build a predictive classification model for the assessment of the intubation difficulty of the patient.

In the first place, when preparing any machine learning model, it is important to preprocess the raw data and transform it into a set of variables that machine learning algorithms can use so that the data can be correctly interpreted for the model. Specifically, we performed the normalization of the data as in our database each variable had different ranges and magnitudes, therefore, making it possible to bring all the features to the same scale and allowing the machine learning algorithm to treat them equally. With respect to the frequency data amplitude the minimum value was set to 0 and the maximum value was set to 1, the remaining middle values were distributed between 0 and 1. Furthermore, the rest of the variables such as age, weight and Mallampati grade were also normalized to ensure that all variables had a similar range and could be compared. **Figure 59** and **60.**



Figures 59 and 60. *KNIME® node that allows normalizing each column from 0 to 1 for both the Harmonics and the descriptive data. (Extracted from the final workflow, own production).*

Once all the data was normalized, different classification models were used to predict intubation difficulty such as the followings: K-nearest neighbors, Naïve Bayes, Neural networks, random forest, support machine vector, Adaboost.

The datasets were divided into two groups for both studies for both the Cormack and the Mallampati study.

The training/validation dataset, which is commonly used to train the model and serves as the foundation of their learning, was composed of 70% of the data (N=416 / 594 patients) and (N=219/ 313 patients) for both the Mallampati and the Cormack studies respectively.

The test dataset which is commonly used to evaluate the performance of the final model and to determine which model performs better, consisted of the 30% of the data (N=178/594 patients) and (N=94/313 patients) for both the Mallampati and the Cormack studies. **Figure 61**

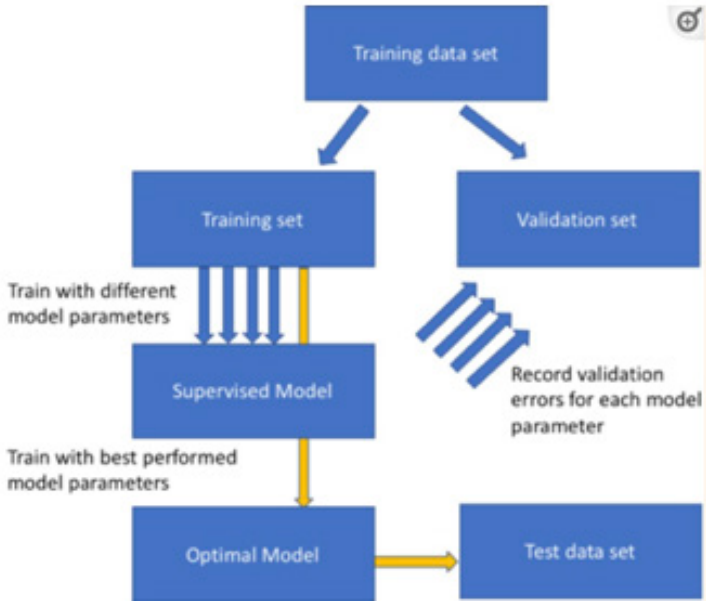


Figure 61. *Partitioning of the dataset. 70% of the participants formed the training and validation dataset and 30% of them formed the test dataset for both studies* (171).

It is worth mentioning at this stage that the database contained more straightforward cases than difficult ones and therefore our database was imbalanced. Specifically, the number of Mallampati 3 and 4 patients was lower compared to the number of patients with Mallampati 1 or 2, meaning easy cases in the Mallampati study. Similarly, for the Cormack study we also had fewer patients classified as difficult (Cormack 3 and 4) and more patients classified as easy (Cormack 1 and 2)

In order to address this problem different approaches were required such as upweight minority class and oversample minority class. These techniques were applied depending on the type of model, in order to balance the data and to ensure that each category in the dataset was properly represented.

Moreover, the cross-validation method was also used with the objective to provide realizable results for model prediction. This method involved evaluating and testing the performance multiple times using different training and validation sets to ensure that the results are reliable and not

dependent on the specific partition of the data into training and testing sets. For this reason, the number of folds in the mode configuration was set at 10. This means that the dataset was divided into ten equal subsets, then the model was trained in nine subsets and then validated on the remaining subset. This process was repeated ten times with a different subset used for validation each time. With regard to the Mallampati study, since we only had a few patients with Mallampati grade 4, the number of validations was reduced to 2.

What's more, in our dataset we had more columns than rows, thus creating an excessively high volume of data that far from contributing to the improvement of the model's prediction it ended up having a detrimental effect on the overall performance as it confused the model. In order to avoid that, it is common to use techniques of dimensionality reduction such as Principal Component Analysis technique (PCA). For this reason, we decided to apply the PCA function in the workflow in order to mitigate the negative impact of dealing with an excessive volume of data, as it assists in reducing the number of variables while retaining the most valuable information of the dataset. **Figure 62**

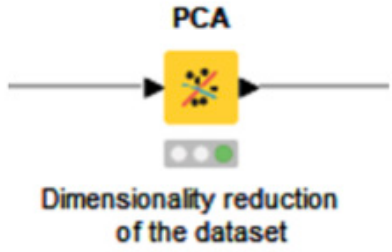


Figure 62. *PCA node configuration in KNIME workflow.*

Furthermore, a hyperparameter tuning of the models was performed to enhance the performance of machine learning classifiers. The goal of adjusting these parameters was to optimize models and improve their classification performance. As an example, an adjustment of hyperparameters can be done by determining the depth of the decision trees. **Figure 63**

Row ID	I depthT...	I numTree	D Mean(Log Loss)	D Standard deviation(Log Loss)	D Mean(Accuracy)	D Standard deviation(Accuracy)	D
Row0#0	10	50	0.496	0.131	0.825	0.059	0.
Row0#1	10	100	0.499	0.125	0.821	0.067	0.
Row0#2	10	150	0.499	0.126	0.823	0.075	0.
Row0#3	12	50	0.495	0.14	0.823	0.067	0.
Row0#4	12	100	0.5	0.131	0.821	0.067	0.
Row0#5	12	150	0.502	0.133	0.814	0.069	0.
Row0#6	14	50	0.496	0.142	0.823	0.067	0.
Row0#7	14	100	0.5	0.132	0.821	0.067	0.
Row0#8	14	150	0.503	0.134	0.814	0.069	0.
Row0#9	16	50	0.495	0.142	0.823	0.067	0.
Row0#10	16	100	0.499	0.132	0.824	0.062	0.
Row0#11	16	150	0.504	0.135	0.814	0.069	0.
Row0#12	18	50	0.495	0.141	0.823	0.067	0.
Row0#13	18	100	0.499	0.132	0.824	0.062	0.
Row0#14	18	150	0.504	0.135	0.814	0.069	0.
Row0#15	20	50	0.495	0.141	0.823	0.067	0.
Row0#16	20	100	0.499	0.132	0.824	0.062	0.
Row0#17	20	150	0.504	0.135	0.814	0.069	0.

Figure 63. *Decision three hyperparameters adjustment: set of the depth and the number of threes.*

At this point we inserted different variable combinations to incorporate into the models. For example, selecting the type of vowel and its position, including the data from all the patients or only selecting those with Cormack grades 1 and 4 etc. Additionally, the target variable to be predicted such as Cormack or Mallampati grades was also modified in the workflow configuration depending on the type of study to which the patients belonged. **Figure 64** and **65**.

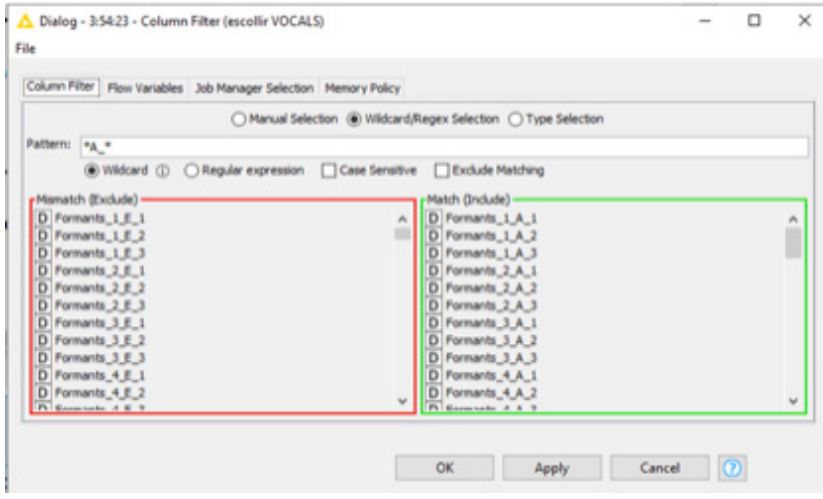


Figure 64. *Vowel selection in KNIME workflow.*

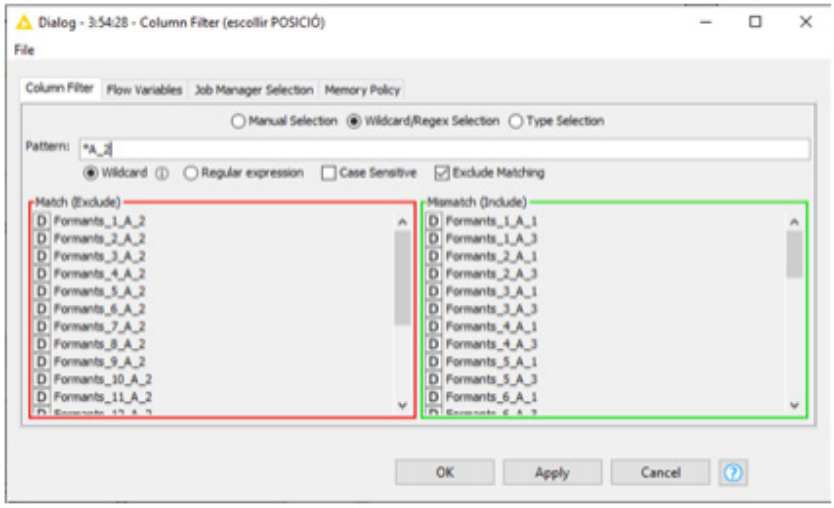


Figure 65. *Position for each vowel selection in KNIME workflow.*

Following this, all the combinations were done and all the models were trained and validated, we proceeded to test the model's performance. In order to do that, the following metrics were extracted for each model: F1 score, ROC AUC curve, Log-loss value, recall, precision, rate of false positive, rate of false negatives, rates of true positives and rates of true negatives as well as the overall accuracy obtained by the different models for both the Cormack and the Mallampati studies.

Finally, the best model, combination of variables and parameters yielding the best results for both the Mallampati and Cormack study was selected, respectively.

4.9. ETHICS AND LEGAL ASPECTS OF THE STUDY

The study was conducted in accordance with the Declaration of Helsinki (current version: Fortaleza, Brazil, October 2013) and strictly adhered to international ethical recommendations for medical research. The approval from the ethics committee of Quironsalud group came on the 26th September 2019 with the following reference number: 62/2019 (Protocol code VOICE2019).

During the intervention of the study, there were no additional risks for the participants as all patients received the same standard care and during the intraoperative process the usual clinical practice without any deviations was followed by the medical team. On top of that, the anesthetic plan and the devices used during the anesthesia were identical as if they had not been included in the study. The obtained data was stored in a coded format to ensure the anonymity of the participating patients in a dedicated anonymous database, accessible only to the principal investigators guaranteeing confidentiality. Accordingly, the identifiable patient data was kept separate from their clinical data. Furthermore, this procedure was done in compliance with the regulations of the Organic Law 3/2018 of December 4," Protection of Personal Data and Guarantee of Digital Rights" and the General Data Protection Regulation (EU)2016/679 of April 27, 2016, which came into effect on May 24, 2018.

All the participants were informed about the purpose and procedures of the study and any potential risks or benefits were communicated to them. Moreover, their enrollment was completely voluntary, and they had the freedom to withdraw from the study at any time without facing any repercussions. In addition, the informed consent of all the participants of the study was obtained before their inclusion in the study. **(Annex 6)** Finally, researchers committed not to use the data for future related studies without prior re-evaluation of the project by the Ethics Committee.

5. RESULTS

5. RESULTS

5.1. SUMMARY OF THE RESULTS

In this chapter the results of the thesis will be presented. Firstly, the outcomes of both the Mallampati and the Cormack study will be accurately displayed according to this order:

Firstly, the descriptive statistics results of all the participants' dataset included in both studies will be exposed in the form of tables. These will include the main clinical features, the variables collected of the patients as well as the number and percentage of patients which correspond to the assigned feature or class.

Secondly, the performance results of the classification algorithms for both the Mallampati and the Cormack studies will be presented according to specific statistical parameters such as area under the curve value (AUC), accuracy, precision, recall and number of true positives, true negatives, false positives and false negatives.

Finally, 3 and 4 combinations of parameters which obtained the best results of the Mallampati, and the Cormack study were selected and will be carefully explained.

The results will be structured as the following:

- Results of the Mallampati study:
  - Descriptive data of the total number of participants included in both the Mallampati and the Cormack studies.
  - Performance results of the classification algorithms of the Mallampati study.
- Results of the Cormack study:
  - Descriptive statistical results of all the patient's dataset included in the Cormack study.

- Performance results of the classification algorithms of the Cormack study
- Best selected parameters combinations for the Mallampati and the Cormack study.
- Evaluation of the Mallampati performance as a predictor of a difficult intubation in our dataset. ROC Curve

5.2. THE MALLAMPATI STUDY RESULTS

5.2.1. Descriptive data results of the total participants included in both the Mallampati and Cormack studies

A total of N= 719 patients were collected for the study. Of these, N= 125 were excluded due to poorly recorded audio signals or missing clinical data of the patients and were not included in the database. The sample for this study consisted of 594 participants who were included in both studies of which, N= 284 (47.81%) were male and N=310 (52.18%) were female. It is important to note that gender was self-reported by the participants and the study did not collect any additional information regarding gender identity. The mean age of the participants was 48.8 years with a standard deviation of 14.8. The distribution of weight and body mass index (BMI) was also analyzed for the sample population. The average weight of the participants was 73.9 kilograms with a standard deviation of 16.6. The average height was 169.1 centimeters with a standard deviation of 9.3. The average BMI was 25.7 with a standard deviation of 4.8. Using the World Health Organization's classification N= 304 (51.1%) of the participants were classified as overweight (BMI ≥ 25). The remaining N= 290 (48.8%) of the participants had a BMI within the normal range (BMI ≤ 25).

The ASA Physical status classification was recorded for each participant in the study to evaluate the overall health status of the sample. Most of the participants N=392 (65.9%) were classified as ASA II indicating mild systemic disease. N=180 (30.3%) of the participants were classified as ASA I, indicating normal healthy patients while the remaining N=22 (3.7) was classified as ASA III, indicating severe systemic disease that limits activity but

is not incapacitating. It is important to highlight that none of the participants were classified as ASA IV, which indicates severe systemic disease that is a constant threat to life. These results suggest that the sample population was generally healthy and suitable for the surgical procedures that were performed.

When it comes to airway evaluation the majority of the patients N=475 (80%) did not test positive for the bedside screening tests. However, N=119 (20%) of the patients were tested positive for the physical features collected in the traditional difficult airway evaluation which predicts increased airway difficulty. The most frequent clinical feature which predicts increased airway difficulty among the patients included in the study was a short thyromental distance with a total of N=115 (19.3%) of patients with a thyromental distance (Patil's test) of < 6.5 centimeters.

The distribution of the rest of the clinical features was as follows: N=2 (0.3%) of patients had an inter-incision gap or mouth opening < 3.5 centimeters. Likewise, the same number of patients N=2 (0.3%) had a jaw protrusion grade C. None of the patients reported having a reduced range of neck extension movement of < 80°. Moreover, N=37 (6.2%) obtained a total of ≥ 11 points in the Arne test indicating a difficult airway. Finally, only N=4 (0.67%) patients reported having the antecedent of a difficult intubation. In addition, only N=4 (0.67%) of patients reported having a pathology associated with a difficult intubation. From these results, it can be inferred that the percentage of patients with physical characteristics which suggest a difficult airway in the entire sample is low and will be considered when analyzing the results.

In this study we also examined the prevalence of comorbidities which have been previously associated with a difficult intubation such as diabetes, sleep apnea syndrome, thyroid disorders and COPD. Among all the participants a total of N= 139 (23%) participants presented one of those comorbidities as being the most common thyroid disorders affecting N=53 (8.9%) of the patients. The second more frequent comorbidity observed in our sample was sleep apnea syndrome present in N= 47 (7.9%) of patients. Another significantly high comorbidity was diabetes affecting N=31 (5.2%)

of patients. Finally, a minority of the patients N=8 (1.3%) presented COPD. It is important to note that N=191 (32.1%) of patients defined themselves as smokers. **Table 4** and **5.**

Table 4. *Descriptive results of all the participants included in the Mallampati study.*

CHARACTERISTICS		N (patients)	% (percentage)
Total number of patients		594	100,00
Gender	Male	284	47,81
	Female	310	52,19
BMI	>=25	304	51,18
	<25	290	48,82
Smoker		191	32,15
Sleep apnea syndrome		47	7,91
COPD		8	1,35
Diabetes		31	5,22
Thyroid disorders		53	8,92
History of difficult intubation		4	0,67
Pathologies associated with difficult intubation		4	0,67
Airway pathology symptoms		2	0,34
Inter-incisors gap/ Mouth opening	>=5 cm	514	86,53
	3,5 cm - 5 cm	77	12,96
	<3,5 cm	2	0,34
Jaw protrusion grade	>0	533	89,73
	0	58	9,76
	<0	2	0,34

Table 4. *Descriptive results of all the participants included in the Mallampati study.* Continuation

CHARACTERISTICS		N (patients)	% (percentage)
Thyromental distance	>=6,5 cm	479	80,64
	<6,5 cm	115	19,36
Neck extension	>100°	548	92,26
	90° ± 10°	46	7,74
	<80°	0	0,00
ASA	ASA I	180	30,30
	ASA II	392	65,99
	ASA III	22	3,70
	ASA IV	0	0,00
Arne Index	>=11	37	6,23
	<11	557	93,77
Mallampati grade	I	258	43,43
	II	211	35,52
	III	113	19,02
	IV	12	2,02

Table 5. *Main features of the patients included in the Mallampati study.*

	Mean	Standard deviation
Weight (kg)	74,00	16,63
Height (cm)	169,12	9,34
Age (Years)	48,86	14,90
BMI	25,76	4,89

5.2.2. Performance results of classification algorithms of the Mallampati study

In this study, we evaluated the performance of several classification models in predicting the presence of Mallampati grades III or IV (Difficult intubation prediction group) on a set of different clinical, demographic and voice parameter variables combinations. Our sample consisted of N= 594 patients. In which, the Mallampati grades were distributed as follows:

- Mallampati grade I: N= 258 patients
- Mallampati grade II: N= 211 patients
- Mallampati grade III: N=113 patients
- Mallampati grade IV: N= 12 patients

Subsequently, different combinations of clinical and demographic dataset as well as voice parameters were analyzed. Moreover, in order to proceed with the analysis of the data through different classification algorithms, patients in the dataset were divided into two groups:

- Training group (70% of N= 594 patients): 415.8 patients
- Testing group (30% of N= 594 patients): 178.2 patients

The performance of the different classification algorithms is displayed below:

5.2.2.1. *Results of accuracy metrics*

The mean accuracy of the classification algorithms for the prediction of the Mallampati of the different combinations is summarized as the following and shown in detail in **table 6**. (Simplified). The full table is shown in **Annex 7.**

- The combination harmonics + descriptive data + different vocals in different positions + all cases obtained a mean accuracy of 0.72%.
- The combination of harmonics + descriptive data + different vocals in different positions + only analyzing Mallampati I and IV cases obtained a mean accuracy of 0.98%.
- Only descriptive data (without any voice parameters) + all cases obtained an accuracy of 0.74%

- Only harmonics of different vocals in different position (without any descriptive data) + all cases obtained a mean accuracy of 0.27%
- Spectral of different vocals in different positions + descriptive data + all cases obtained a mean accuracy of 0.70%
- The combination of voice parameters (such as shimmer, jitter, HNR) + descriptive data + all cases obtained a mean accuracy of 0.69%
- The combination of voice parameters (Such as shimmer, jitter, HNR) + descriptive data + only analyzing Mallampati I and IV cases obtained a mean accuracy of 0.97%

Table 6. *Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in both training and testing groups, Accuracy values for the training group, Accuracy values for the testing group (Simplified).*

ACCURACY					
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data All vocals		415/179	H2O GRADIENT BOOSTING	0,79	0,69
HARMONICS + Descriptive data I flexion (_3) vs I extension (_2)		415/179	H2O GENERALISED LINEAR MODEL	0,81	0,75
HARMONICS + Descriptive data I all positions		415/179	H2O GENERALISED LINEAR MODEL	0,82	0,75
HARMONICS + Descriptive data U normal (_1) vs U extension (_2)		415/179	H2O GENERALISED LINEAR MODEL	0,82	0,74
HARMONICS + Descriptive data U flexion (_3) vs U extension (_2)		415/179	H2O GENERALISED LINEAR MODEL	0,77	0,76
HARMONICS + Descriptive data U all positions		415/179	H2O GENERALISED LINEAR MODEL	0,81	0,75
HARMONICS + Descriptive data All vocals	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,95	0,95

Table 6. *Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in both training and testing groups, Accuracy values for the training group, Accuracy values for the testing group (Simplified).* Continuation

ACCURACY					
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data A normal (_1)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,97	0,98
HARMONICS + Descriptive data A all positions	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,95	0,95
Only Descriptive data		415/179	H2O GENERALISED LINEAR MODEL	0,75	0,74
Descriptive data + HARMONICS A normal		415/179	H2O GENERALISED LINEAR MODEL	0,35	0,34
HARMONICS A extension (_2)		415/179	H2O RANDOM FOREST	0,46	0,21
Spectral + descriptive data All vocals		415/179	H2O GENERALISED LINEAR MODEL	0,77	0,72
Spectral+ descriptive data A normal		415/179	H2O GRADIENT BOOSTING	0,78	0,79
Voice parameters+ descriptive data E flexion vs E extension		415/179	H2O NAIVE BAYES	0,76	0,59
Voice parameters+ descriptive data O flexion vs O extension		415/179	H2O NAIVE BAYES	0,78	0,68
Voice parameters+ descriptive data All vocals	Only Mallampati 1 and 4 cases	189/81	H2O RANDOM FOREST	0,39	0,95
Voice parameters+ descriptive data O flexion vs O extension	Only Mallampati 1 and 4 cases	189/81	H2O GRADIENT BOOSTING	0,97	0,98

5.2.2.2. *Results of AUC and Pr AUC metrics*

AUC values:

The mean AUC values of the classification algorithms for the prediction of the Mallampati of the different combinations is summarized as the following and shown in detail in **table 7** (Simplified). The full table is shown in **Annex 8**

- The combination harmonics + descriptive data + different vocals in different positions + all cases obtained a mean AUC of 0.74%.
- The combination of harmonics + descriptive data + different vocals in different positions + only analyzing Mallampati I and IV cases obtained a mean AUC of 0.98%.
- Only descriptive data (without any voice parameters) + all cases obtained an AUC value of
- 0.73%
- Only harmonics of different vocals in different position (without any descriptive data) + all cases obtained a mean AUC of 0.48%
- Spectral of different vocals in different positions +descriptive data + all cases obtained a mean AUC of 0.70%
- The combination of voice parameters (such as shimmer, jitter, HNR) + descriptive data + all cases obtained a mean AUC of 0.71%
- The combination of voice parameters (such as shimmer, jitter, HNR) + descriptive data + only analyzing Mallampati I and IV cases obtained a mean AUC value of 0.95%

Pr AUC values:

- The combination harmonics + descriptive data + different vocals in different positions + all cases obtained a mean Pr AUC of 0.40%.
- The combination of harmonics + descriptive data + different vocals in different positions + only analyzing Mallampati I and IV cases obtained a mean Pr AUC of 0.43%.
- Only descriptive data (without any voice parameters) + all cases obtained an Pr AUC of 0.41%

- Only harmonics of different vocals in different position (without any descriptive data) + all cases obtained a mean Pr AUC of 0.20 %
- Spectral of different vocals in different positions + descriptive data + all cases obtained a mean Pr AUC of 0.37%
- The combination of voice parameters (such as shimmer, jitter, HNR) + descriptive data + all cases obtained a mean Pr AUC of 0.36%
- The combination of voice parameters (Such as shimmer, jitter, HNR) + descriptive data + only analyzing Mallampati I and IV cases obtained a mean Pr AUC 0.42%

Table 7. *Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in both training and testing groups, AUC values for the training group, AUC values for the testing group, Pr AUC values for the testing group.*

				AUC		Pr AUC	
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data U normal (_1) vs U extension (_2)		415/179	H2O GENERALISED LINEAR MODEL	0,76	0,74	0,53	0,41
HARMONICS + Descriptive data U flexion (_3) vs U extension (_2)		415/179	H2O GENERALISED LINEAR MODEL	0,76	0,74	0,52	0,40
HARMONICS + Descriptive data A normal (_1) vs A extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,87	0,97	0,26	0,56
HARMONICS + Descriptive data E flexion (_2) vs E extension (_3)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,85	0,98	0,20	0,58
HARMONICS + Descriptive data U flexion (_3) vs U extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,85	0,97	0,19	0,50

Table 7. *Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in both training and testing groups, AUC values for the training group, AUC values for the testing group, Pr AUC values for the testing group.* Continuation

				AUC		Pr AUC	
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data U all positions	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,91	0,98	0,23	0,29
Spectral+ descriptive data A flexion vs A extension		415/179	H2O GENERALISED LINEAR MODEL	0,72	0,71	0,45	0,39
Spectral + descriptive data E flexion vs E extension		415/179	H2O GRADIENT BOOSTING	0,72	0,70	0,42	0,34
Voice parameters+ descriptive data All vocals		415/179	H2O GRADIENT BOOSTING	0,60	0,66	0,30	0,36
Voice parameters+ descriptive data A flexion vs A extension	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,53	0,98	0,15	0,46
Voice parameters+ descriptive data E flexion vs E extension	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,89	0,97	0,30	0,50
Voice parameters+ descriptive data O flexion vs O extension	Only Mallampati 1 and 4 cases	189/81	H2O GRADIENT BOOSTING	0,77	0,93	0,19	0,36

5.2.2.3. *Results of Precision and Recall metrics*

The results for all parameters combinations for both the testing and training groups according to precision and recall of the Mallampati are shown in detail in **Table 8** (*simplified*), the full table can be found in **Annex 9**.

Precision results:

- The combination harmonics + descriptive data + different vocals in different positions + all cases obtained a mean precision of 0.41%.
- The combination of harmonics + descriptive data + different vocals in different positions + only analyzing Mallampati I and IV cases obtained a mean precision of 0.82%.
- Only descriptive data (without any voice parameters) + all cases obtained a precision of 0.43%
- Only harmonics of different vocals in different position (without any descriptive data) + all cases obtained a mean precision of 0.23 %
- Spectral of different vocals in different positions + descriptive data + all cases obtained a mean precision 0.40%
- The combination of voice parameters (such as shimmer, jitter, HNR) + descriptive data + all cases obtained a mean precision of 0.38%
- The combination of voice parameters (such as shimmer, jitter, HNR) + descriptive data + only analyzing Mallampati I and IV cases obtained a mean precision of 0.71%

Recall:

- The combination harmonics + descriptive data + different vocals in different positions + all cases obtained a mean recall of 0.66%.
- The combination of harmonics + descriptive data + different vocals in different positions + only analyzing Mallampati I and IV cases obtained a mean recall of 0.81%.
- Only descriptive data (without any voice parameters) + all cases obtained a recall of 0.61%
- Only harmonics of different vocals in different position (without any descriptive data) + all cases obtained a mean recall of 0.99 %

- Spectral of different vocals in different positions + descriptive data + all cases obtained a mean recall 0.67%
- The combination of voice parameters (such as shimmer, jitter, HNR) + descriptive data + all cases obtained a mean recall of 0.70%
- The combination of voice parameters (Such as shimmer, jitter, HNR) + descriptive data + only analyzing Mallampati I and IV cases obtained a mean recall of 0.71%

Table 8. *Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in both training and testing groups, Precision values for the training group, Recall values for the testing group (Simplified)*

				Precision		Recall	
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data U normal (_1) vs U extension (_2)		415/179	H2O GENERALISED LINEAR MODEL	0,65	0,42	0,55	0,66
HARMONICS + Descriptive data U flexion (_3) vs U extension (_2)		415/179	H2O GENERALISED LINEAR MODEL	0,59	0,45	0,65	0,61
HARMONICS + Descriptive data U all positions		415/179	H2O GENERALISED LINEAR MODEL	0,60	0,43	0,60	0,61
HARMONICS + Descriptive data All vocals	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,65	0,50	0,63	1,00
HARMONICS + Descriptive data A normal (_1)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,88	0,75	0,47	0,75
HARMONICS + Descriptive data E extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,67	1,00	0,73	0,75

Table 8. *Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in both training and testing groups, Precision values for the training group, Recall values for the testing group (Simplified).* Continuation

				Precision		Recall	
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data E flexion(_3)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,88	0,75	0,47	0,75
HARMONICS + Descriptive data I normal (_1)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,88	1,00	0,47	0,75
Només DADES DESCRIPTIVES		415/179	H2O GENERALISED LINEAR MODEL	0,56	0,43	0,73	0,61
Only HARMONICS A normal (_1)		415/179	H2O GENERALISED LINEAR MODEL	0,24	0,24	0,91	0,97
Only HARMÒNICS A extension (_2)		415/179	H2O RANDOM FOREST	0,27	0,21	0,78	1,00
Spectral + descriptive data All vocals		415/179	H2O GENERALISED LINEAR MODEL	0,52	0,40	0,61	0,66
Spectral+ descriptive data A normal		415/179	H2O GRADIENT BOOSTING	0,54	0,51	0,43	0,55
Spectral+ descriptive data O extension		415/179	H2O GRADIENT BOOSTING	0,61	0,31	0,44	0,76
Voice parameters+ descriptive data All vocals	Only Mallampati 1 and 4 cases	189/81	H2O RANDOM FOREST	0,08	0,50	0,80	0,50
Voice parameters+ descriptive data A normal	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,75	0,75	0,37	0,75

5.2.2.4. *Results of F1 and Log Loss metrics*

The results for all parameter combinations for both the testing and training groups according to precision and recall of the Mallampati are shown in detail in **Table 9.** (*Simplified*) The full table can be found in **Annex 10.**

Log loss:

- The combination harmonics + descriptive data + different vocals in different positions + all cases obtained a mean Log loss of 0.53%.
- The combination of harmonics+ descriptive data + different vocals in different positions + only analyzing Mallampati I and IV cases obtained a mean Log loss of 0.43%.
- Only descriptive data (without any voice parameters) + all cases obtained a F1 of 0.50%
- Only harmonics of different vocals in different position (without any descriptive data) + all cases obtained a mean Log loss of 0.54 %
- Spectral of different vocals in different positions + descriptive data + all cases obtained a mean Log loss of 0.51%
- The combination of voice parameters (such as shimmer, jitter, HNR) + descriptive data + all cases obtained a mean Log loss of 0.50%
- The combination of voice parameters (Such as shimmer, jitter, HNR) + descriptive data + only analyzing Mallampati I and IV cases obtained a mean Log loss of 0.19%

F1:

- The combination harmonics + descriptive data + different vocals in different positions + all cases obtained a mean F1 of 0.50%.
- The combination of harmonics + descriptive data + different vocals in different positions + only analyzing Mallampati I and IV cases obtained a mean F1 of 0.79%.
- Only descriptive data (without any voice parameters) + all cases obtained a F1 of 0.50%
- Only harmonics of different vocals in different position (Without any descriptive data) + all cases obtained a mean F1 of 0.37 %

- Spectral of different vocals in different positions + descriptive data + all cases obtained a mean F1 of 0.49%
- The combination of voice parameters (such as shimmer, jitter, HNR) + descriptive data + all cases obtained a mean F1 of 0.49%
- The combination of voice parameters (Such as shimmer, jitter, HNR) + descriptive data + only analyzing Mallampati I and IV cases obtained a mean F1 of 0.75%

Table 9. *Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in both training and testing groups, F1 and Log loss values for the training group and the testing group respectively.*

				Log loss		F1	
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data All vocals		415/179	H2O GRADIENT BOOSTING	0,58	0,50	0,47	0,50
HARMONICS + Descriptive data A normal (_1)		415/179	H2O NAIVE BAYES	0,80	1,00	0,57	0,48
HARMONICS + Descriptive data All vocals	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,33	0,70	0,60	0,67
HARMONICS + Descriptive data A normal (_1)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,37	0,09	0,58	0,75
HARMONICS + Descriptive data E extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,42	0,81	0,70	0,86
HARMONICS + Descriptive data O flexion(_3)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,40	0,72	0,64	0,75

Table 9. *Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in both training and testing groups, F1 and Log loss values for the training group and the testing group respectively.* Continuation

				Log loss		F1	
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data U normal (_1)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,42	0,79	0,70	0,75
Only Descriptive data		415/179	H2O GENERALISED LINEAR MODEL	0,58	0,50	0,58	0,50
Only HARMÒNICS A normal (_1)		415/179	H2O GENERALISED LINEAR MODEL	0,73	0,53	0,37	0,38
Spectral + descriptive data E flexion vs E e xtension		415/179	H2O GRADIENT BOOSTING	0,73	0,49	0,52	0,45
Voice parameters+ descriptive data All vocals		415/179	H2O GRADIENT BOOSTING	0,60	0,52	0,40	0,47
Voice parameters+ descriptive data A normal		415/179	H2O GENERALISED LINEAR MODEL	0,58	0,50	0,57	0,52
Voice parameters+ descriptive data O extension		415/179	H2O GRADIENT BOOSTING	0,60	0,48	0,46	0,50
Voice parameters+ descriptive data O flexion		415/179	H2O NAIVE BAYES	0,90		0,56	0,53
Voice parameters+ descriptive data E flexion vs E extension	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,26	0,11	0,56	
Voice parameters+ descriptive data O flexion vs O extension	Only Mallampati 1 and 4 cases	189/81	H2O GRADIENT BOOSTING	0,15	0,10	0,54	

5.2.2.5. *Results of false positives and false negatives*

Results of TN, FN, FP, TP metrics.

The results for the true negatives, false negatives, false positives and true positive values of the classification algorithms for the different combinations in the testing group are summarized in the following text.

Furthermore, the results of both the training and testing groups and shown in detail in **table 10** (Simplified) the full Table is found in **annex 11**.

- The combination harmonics + descriptive data + different vocals in different positions + all cases obtained a mean TN of 104; FN: 13; FP:36; TP:24.9.
- The combination of harmonics + descriptive data + different vocals in different positions + only analyzing Mallampati I and IV cases obtained a mean TN of 76; FN: 0.77; FP:1; TP:3.2.
- Only descriptive data (without any voice parameters) + all cases obtained a TN of 110; FN: 15; FP:31; TP:23.
- Only harmonics of different vocals in different positions (without any descriptive data) + all cases obtained a mean TN of 11.5; FN: 0.5; FP: 129.5; TP 37.5.
- Spectral of different vocals in different positions + descriptive data + all cases obtained a mean TN of 100; FN: 12.5; FP: 40; TP 25.5.
- The combination of voice parameters (such as shimmer, jitter, HNR) + descriptive data + all cases obtained a mean TN of 96; FN: 11.2; FP: 45; TP 26.7.
- The combination of voice parameters (such as shimmer, jitter, HNR) + descriptive data + only analyzing Mallampati I and IV cases obtained a mean TN of 75.8; FN: 1.1; FP: 1.1; TP 2.8.

Table 10. *Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in testing group. TN: True negatives, FN: False negatives, FP: False positives, TP: True positives. (Simplified)*

TESTING							
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TN	FN	FP	TP
HARMONICS + Descriptive data All vocals		179	H2O GRADIENT BOOSTING	96	10	45	28
HARMONICS + Descriptive data U normal (_1) vs U extension (_2)		179	H2O GENERALISED LINEAR MODEL	107	13	34	25
HARMONICS + Descriptive data U flexion (_3) vs U extension (_2)		179	H2O GENERALISED LINEAR MODEL	113	15	28	23
HARMONICS + Descriptive data U all positions		179	H2O GENERALISED LINEAR MODEL	111	15	30	23
HARMONICS + Descriptive data All vocals	Only Mallampati 1 and 4 cases	81	H2O NAIVE BAYES	73	0	4	4
HARMONICS + Descriptive data E flexion (_2) vs E extension (_3)	Only Mallampati 1 and 4 cases	81	H2O GENERALISED LINEAR MODEL	77	1	0	3
HARMONICS + Descriptive data E all positions	Only Mallampati 1 and 4 cases	81	H2O GENERALISED LINEAR MODEL	76	1	1	3
HARMONICS + Descriptive data I normal (_1) vs I flexion (_3)	Only Mallampati 1 and 4 cases	81	H2O GENERALISED LINEAR MODEL	77	1	0	3
Only HARMONICS A normal (_1)		179	H2O GENERALISED LINEAR MODEL	23	1	118	37
Only HARMONICS A extension (_2)		179	H2O RANDOM FOREST	0	0	141	38
Spectral+ descriptive data A flexion vs A extension		179	H2O GENERALISED LINEAR MODEL	107	13	34	25
Spectral + descriptive data E flexion vs E e xtension		179	H2O GRADIENT BOOSTING	87	11	54	27
Voice parameters+ descriptive data All vocals		179	H2O GRADIENT BOOSTING	100	14	41	24
Voice parameters+ descriptive data O flexion vs O extension		179	H2O NAIVE BAYES	92	9	49	29
Voice parameters+ descriptive data All vocals	Only Mallampati 1 and 4 cases	81	H2O RANDOM FOREST	75	2	2	2

Table 10. *Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in testing group. TN: True negatives, FN: False negatives, FP: False positives, TP: True positives. (Simplified).* Continuation

TESTING							
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TN	FN	FP	TP
Voice parameters+ descriptive data E flexion vs E extension	Only Mallampati 1 and 4 cases	81	H2O GENERALISED LINEAR MODEL	76	1	1	3
Voice parameters+ descriptive data O flexion vs O extension	Only Mallampati 1 and 4 cases	81	H2O GRADIENT BOOSTING	76	1	1	3

5.3. RESULTS OF THE CORMACK STUDY

5.3.1. Descriptive statistics of the participants included in the Cormack study

A total of N= 313 (52%) of all participants recruited for the studies were selected and included in the Cormack study. The distribution of the patients according to gender was as follows:

N= 146 (46.64%) were male and N= 167 (53.35%) were female. The mean age of the participants was 49.16 years with a standard deviation of 14.6 The distribution of weight and body mass index (BMI) was also analyzed for the sample population. The average weight of the participants was 74.53 kilograms with a standard deviation of 17.0 The average height was 168.74 centimeters with a standard deviation of 9.7 the average BMI was 26.09 with a standard deviation of 5.2 using the World Health Organization's classification N= 151 (48.24%) of the participants were classified as overweight (BMI ≥ 25). The remaining N= 162 (51.75%) of the participants had a BMI within the normal range (BMI ≤ 25)

The ASA Physical status classification was also recorded for each participant in the study to evaluate the overall health status of the sample. Most of the participants N= 210 (67.09%) were classified as ASA II indicating mild systemic disease. N= 86 (27.47%) of the participants were classified as ASA

I, indicating normal healthy patients while the remaining N=17 (5.43) were classified as ASA III, indicating severe systemic disease that limits activity but is not incapacitating. It is important to highlight that none of the participants were classified as ASA IV which indicates severe systemic disease that is a constant threat to life. These results suggest that the sample population was generally healthy and suitable for the surgical procedures that were performed.

These results are similar to those obtained in the Mallampati study which means that both sample groups have much the same characteristics.

When it comes to airway evaluation a considerable percentage of patients N =133 (43%) did test positive for at least one of the bedside screening tests for a difficult airway. Moreover N= 22 (7%) of patients accumulated a score of ≥ 11 points of the Arne Test Score. The remaining N=158 (50%) of patients did not have any predictive anthropometric features for a difficult airway in the bedside screening tests. It is noteworthy that in this Cormack study, the percentage of patients which were positive for at least one parameter for a prediction of a difficult airway is almost double the other patients included in the Mallampati study 43% vs. 20% of patients in the Cormack study.

The most frequent clinical feature which predicts increased airway difficulty among the patients included in the study was a short thyromental distance with a total of N=68 (21.7%) of patients with a thyromental distance (Patil's test) of ≤ 6.5 centimeters. When it comes to the most clinical feature for a prediction of increased airway difficulty the results are very similar to the ones found in the Mallampati study 20% vs 19.3%. The remaining clinical features were distributed in the following manner: N=1 (0.3%) of patients had an inter-incision gap or mouth opening < 3.5 centimeters.

Equally, the same number of patients N=1 (0.3%) had a jaw protrusion grade C. None of the patients reported having a reduced range of neck extension movement of < 80°. It should be emphasized that there is a great degree of similarity between these results and the ones found for the Mallampati study.

Moreover, N=22 (7%) obtained a total of ≥ 11 points in the Arne test indicating a difficult airway. Finally, only N=4 (1.27%) of patients reported

having the antecedent of a difficult intubation. In addition, only N=2 (0.63%) of patients reported having a pathology associated with a difficult intubation.

Furthermore, we also examined the occurrence of comorbidities that have been previously associated with a difficult intubation as part of our research in this Cormack study such as diabetes, sleep apnea syndrome, thyroid disorders and COPD. Among all the participants a total of N= 84 (27%) participants presented one of those comorbidities as being the most common thyroid disorders affecting N=34 (10.86%) of the patients. The second more frequent comorbidity observed in our sample was sleep apnea syndrome present in N= 28 (8.94%) of patients.

In addition to this, we identified that diabetes was also significantly elevated affecting N=19 (6.07%) of patients. Finally, it should be noted that only a minority of the patients N=3 (0.95%) presented with COPD. Of note, a substantial number of patients N=105 (33.54%) of patients reported themselves as smokers. Of significance, both the distribution of comorbidities among this Cormack study and the prevalence of said comorbidities in the Mallampati study are strikingly similar. All the findings are summarized in **Table 11** and **Table 12**.

Table 11. *Descriptive statistics of all the participants included in the Cormack study.*

CHARACTERISTICS		N (patients)	% (percentage)
Total number of patients		313	100,00
Gender	Male	146	46,65
	Female	167	53,35
BMI	>=25	151	48,24
	<25	162	51,76
Smoker		105	33,55
Sleep apnea syndrome		28	8,95
COPD		3	0,96
Diabetes		19	6,07
Thyroid disorders		34	10,86
History of difficult intubation		4	1,28

Table 11. *Descriptive statistics of all the participants included in the Cormack study.* Continuation

CHARACTERISTICS		N (patients)	% (percentage)
Pathologies associated with difficult intubation		2	0,64
Airway pathology symptoms		2	0,64
Inter-incisors gap/ Mouth opening	>=5 cm	271	86,58
	3,5 cm - 5 cm	41	13,10
	<3,5 cm	1	0,32
Jaw protrusion grade	A	289	92,33
	B	23	7,35
	C	1	0,32
Thyromental distance	>=6,5 cm	245	78,27
	<6,5 cm	68	21,73
Neck extension	>100°	288	92,01
	90° ± 10°	25	7,99
	<80°	0	0,00
ASA	ASA I	86	27,48
	ASA II	210	67,09
	ASA III	17	5,43
	ASA IV	0	0,00
Mallampati grade	1	139	44,41
	2	111	35,46
	3	59	18,85
	4	4	1,28
Arne Index	>=11	22	7,03
	<11	291	92,97
Cormack grade	I	124	39,62
	IIa	98	31,31
	IIb	28	8,95
	III	45	14,38
	IV	18	5,75

Table 12. *Main features of the participants included in the Cormack study.*

	Mean	Standard deviation
Weight (kg)	74,54	17,01
Height (cm)	168,75	9,72
Age (years)	49,17	14,69
BMI	26,09	5,27

5.3.2. Performance results of the classification algorithms of the Cormack study

In this study, we evaluated the performance of several classification models in predicting the presence of a Cormack grades III or IV (Difficult intubation group) on a set of different clinical, demographic and voice parameter variable combinations. Our sample consisted of N= 313 patients. In which, the Cormack grades were distributed as following:

1) Cormack grade I: N= 124 patients.
2) Cormack grade IIa: N= 98 patients.
3) Cormack grade IIb: N= 28 patients.
4) Cormack grade III: N= 45 patients.
5) Cormack grade IV: N= 18 patients.

Subsequently, different combinations of clinical and demographic datasets as well as voice parameters were analyzed. Moreover, in order to proceed with the analysis of the data through different classification algorithms, the patients in the dataset were divided into two groups:

- Training group (70% of N= 313 patients): 219 patients.
- Testing group (30% of N= 313 patients): 94 patients.

The performance of the different classification algorithms is displayed as the following:

5.3.2.1. *Results of accuracy metrics*

The mean accuracy of the classification algorithms for the different combinations is summarized as the following and shown in detail in **table 13** (*simplified*). The full table is shown in **Annex 12**

- The combination harmonics + descriptive data + different vocals in different positions + all cases obtained a mean accuracy of 0.59%.
- The combination of harmonics + descriptive data + different vocals in different positions + overweight Cormack grade 4 cases obtained a mean accuracy of 0.48%.
- The combination of harmonics + descriptive data + different vocals in different positions + only analyzing Cormack I and IV cases obtained a mean accuracy of 0.82%.
- Only descriptive data (without any voice parameters) + all cases obtained an accuracy of 0.61%.
- Only harmonics of different vocals in different positions (without any descriptive data) + all cases obtained a mean accuracy of 0.54%.
- Spectral of different vocals in different positions + descriptive data + all cases obtained a mean accuracy of 0.64%.
- The combination of voice parameters (such as shimmer, jitter, HNR) + Descriptive data + all cases obtained a mean accuracy of 0.57%.
- The combination of voice parameters (such as shimmer, jitter, HNR) + Descriptive data + only analyzing Cormack I and IV cases obtained a mean accuracy of 0.84%.

Table 13. *Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in both training and testing groups, for accuracy values for the training group and the testing group respectively.*

ACCURACY					
DATASET (PARAMETERS COMBINATION)	TYPE OF CASES	BEST MODEL	N PATIENTS	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data All vocals		H2O GENERALISED LINEAR MODEL	219/94	0,70	0,49
HARMONICS + Descriptive data U all positions		H2O NAIVE BAYES	219/94	0,75	0,60
HARMONICS + Descriptive data E all positions	Overweight Cormack 4	H2O NAIVE BAYES	219/94	0,75	0,61
HARMONICS + Descriptive data I normal (_1) vs I flexion (_3)	Overweight Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,67	0,49
HARMONICS + Descriptive data U normal (_1)	Cormack 1 and 4 cases	H2O GRADIENT BOOSTING	100/43	0,71	0,88
HARMONICS + Descriptive data U extension 2)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,87	0,84
ONLY descriptive data		H2O RANDOM FOREST	219/94	0,69	0,61
HARMONICS I flexion (_3)		H2O NAIVE BAYES	219/94	0,47	0,55
HARMONICS O normal (_1)		H2O GRADIENT BOOSTING	219/94	0,56	0,56
HARMONICS I normal (_1) vs I extension (_2)		H2O NAIVE BAYES	219/94	0,49	0,65
Spectral+ Descriptive data all vocals		H2O GRADIENT BOOSTING	219/94	0,66	0,64
Spectral+ Descriptive data A flexion vs A extension		H2O RANDOM FOREST	219/94	0,66	0,61
Spectral+ Descriptive data E flexion vs E extension		H2O NAIVE BAYES	219/94	0,72	0,65
Voice parameters + Descriptive data all vocals		H2O GENERALISED LINEAR MODEL	219/94	0,65	0,45
Voice parameters + Descriptive data A normal		H2O GENERALISED LINEAR MODEL	219/94	0,75	0,65
Voice parameters + Descriptive data E flexion vs E extension	Only Cormack 1 and 4	H2O GRADIENT BOOSTING	100/43	0,71	0,91
Voice parameters + Descriptive data O flexion vs O extension	Only Cormack 1 and 4	H2O NAIVE BAYES	100/43	0,93	0,81

5.3.2.2. *Results of precision and recall metrics*

The mean precision and recall of the classification algorithms for the different combinations in the testing group is summarized in the following text.

Furthermore, the results of both the training and testing groups are shown in detail in **table 14** (*simplified*) the full table is found in **annex 13.**

Precision results:

- The combination harmonics + descriptive data + different vocals in different positions + all cases obtained a mean precision of 0.32%.
- The combination of harmonics + descriptive data+ different vocals in different positions + overweight Cormack grade 4 cases obtained a mean precision of 0.27%.
- The combination of harmonics + descriptive data + different vocals in different positions + only analyzing Cormack I and IV cases obtained a mean precision of 0.41%.
- Only descriptive data (without any voice parameters) + all cases obtained a precision of 0.30%.
- Only Harmonics of different vocals in different position (without any descriptive data) + all cases obtained a mean precision of 0.29%.
- Spectral of different vocals in different positions + descriptive data + all cases obtained a mean precision of 0.33%.
- The combination of voice parameters (such as shimmer, jitter, HNR) + descriptive data + all cases obtained a mean precision of 0.30%
- The combination of voice parameters (such as shimmer, jitter, HNR) + descriptive data + only analyzing Cormack I and IV cases obtained a mean precision of 0.42%.

Recall results

- The combination harmonics + descriptive data + different vocals in different positions + all cases obtained a mean recall of 0.86%.
- The combination of harmonics + descriptive data+ different vocals in different positions + overweight Cormack grade 4 cases obtained a mean recall of 0.91%.

- The combination of harmonics + descriptive data + different vocals in different positions + only analyzing Cormack I and IV cases obtained a mean recall of 0.90%.
- Only descriptive data (without any voice parameters) + all cases obtained a recall of 0.68 %.
- Only Harmonics of different vocals in different position (without any descriptive data) + all cases obtained a mean recall of 0.78%.
- Spectral of different vocals in different positions + descriptive data + all cases obtained a mean recall of 0.75%.
- The combination of voice parameters (such as shimmer, jitter, HNR) + Descriptive data + all cases obtained a mean recall of 0.80%.
- The combination of voice parameters (such as shimmer, jitter, HNR) + Descriptive data + only analyzing Cormack I and IV cases obtained a mean recall of 0.80%.

Table 14. *Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in both training and testing groups, for precision and recall values for the training group and the testing group respectively. (Simplified).*

DATASET (PARAMETERS COMBINATION)	TYPE OF CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				Precision	Recall	Precision	Recall
HARMONICS + Descriptive data All vocals		H2O GENERALISED LINEAR MODEL	219/94	0,40	0,68	0,27	0,89
HARMONICS + Descriptive data O all positions		H2O GRADIENT BOOSTING	219/94	0,60	0,49	0,31	0,89
HARMONICS + Descriptive data U normal (_1) vs U flexion (_3)		H2O RANDOM FOREST	219/94	0,38	0,78	0,29	0,95
HARMONICS + Descriptive data A extension (_2)	Overweight Cormack 4	H2O GRADIENT BOOSTING	219/94	0,50	0,52	0,30	0,84

Table 14. *Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in both training and testing groups, for precision and recall values for the training group and the testing group respectively. (Simplified).*

Continuation

DATASET (PARAMETERS COMBINATION)	TYPE OF CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				Precision	Recall	Precision	Recall
HARMONICS + Descriptive data I normal (_1) vs I extension (_2)	Overweight Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,36	0,89	0,28	1,00
HARMONICS + Descriptive data U normal (_1) vs U extension (_2)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,59	0,78	0,40	0,80
HARMONICS + Descriptive data U flexion (_3) vs U extension (_2)	Cormack 1 and 4 cases	H2O RANDOM FOREST	100/43	0,40	0,89	0,33	0,80
HARMONICS + Descriptive data U all positions	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,81	0,83	0,36	0,80
ONLY descriptive data		H2O RANDOM FOREST	219/94	0,46	0,67	0,30	0,68
HARMONICS All vocals		H2O NAIVE BAYES	219/94	0,26	0,79	0,27	0,84
HARMONICS U all positions		H2O RANDOM FOREST	219/94	0,21	0,21	0,21	0,21
Spectral+ Descriptive data all vocals		H2O GRADIENT BOOSTING	219/94	0,34	0,64	0,32	0,68
spectral+ Descriptive data A normal		H2O RANDOM FOREST	219/94	0,41	0,77	0,35	0,79
Voice parameters + Descriptive data all vocals		H2O GENERALISED LINEAR MODEL	219/94	0,33	0,70	0,27	1,00
Voice parameters + Descriptive data O flexion vs O extension		H2O NAIVE BAYES	219/94	0,62	0,47	0,28	0,89
Voice parameters + Descriptive data all vocals	Only Cormack 1 and 4	H2O RANDOM FOREST	100/43	0,68	0,73	0,44	0,80
Voice parameters + Descriptive data A normal	Only Cormack 1 and 4	H2O GENERALISED LINEAR MODEL	100/43	0,35	0,69	0,36	0,80

5.3.2.3. *Results of AUC and Pr AUC metrics*

The AUC values and Pr AUC values of the classification algorithms for the different combinations in the testing group are summarized in the following text.

Furthermore, the results of both the training and testing groups and shown in detail in **table 15** (*simplified*) the full table is found in **annex 14**.

AUC values results:

- The combination harmonics + descriptive data + different vocals in different positions + all cases obtained a mean AUC value of 0.68%.
- The combination of harmonics + descriptive data+ different vocals in different positions + overweight Cormack grade 4 cases obtained a mean AUC value of 0.61%.
- The combination of harmonics + descriptive data + different vocals in different positions + only analyzing Cormack I and IV cases obtained a mean AUC value of 0.83%.
- Only descriptive data (without any voice parameters) + all cases obtained an AUC value of 0.56%.
- Only Harmonics of different vocals in different position (without any descriptive data) + all cases obtained a mean AUC value of 0.61%
- Spectral of different vocals in different positions + descriptive data + all cases obtained a mean AUC value of 0.67%.
- The combination of voice parameters (such as shimmer, jitter, HNR) + descriptive data + all cases obtained a mean AUC value of 0.63%.
- The combination of voice parameters (such as shimmer, jitter, HNR) + descriptive data + only analyzing Cormack I and IV cases obtained a mean AUC value of 83%.

Pr AUC values results:

- The combination harmonics + descriptive data + different vocals in different positions + all cases obtained a mean Pr AUC value of 0.28%.

- The combination of harmonics + descriptive data+ different vocals in different positions + overweighting Cormack grade 4 cases obtained a mean Pr AUC value of 0.25%.
- The combination of harmonics + descriptive data + different vocals in different positions + only analyzing Cormack I and IV cases obtained a mean Pr AUC value of 0.27%.
- Only descriptive data (without any voice parameters) + all cases obtained a Pr AUC value of 0.22%.
- Only harmonics of different vocals in different position (without any descriptive data) + all cases obtained a mean Pr AUC value of 0.26%.
- Spectral of different vocals in different positions + descriptive data + all cases obtained a mean Pr AUC value of 0.26%.
- The combination of voice parameters (such as shimmer, jitter, HNR) + descriptive data + all cases obtained a mean Pr AUC value of 0.24%
- The combination of voice parameters (such as shimmer, jitter, HNR) + descriptive data + only analyzing Cormack I and IV cases obtained a mean Pr AUC value of 0.28%.

Table 15. *Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in both training and testing groups, for AUC and Pr AUC values for the training group and the testing group respectively (Simplified).*

DATASET (PARAMETERS COMBINATION)	CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				AUC	Pr AUC	AUC	Pr AUC
HARMONICS + Descriptive data O all positions		H2O GRADIENT BOOSTING	219/94	0,58	0,30	0,70	0,30
HARMONICS + Descriptive data U normal (_1) vs U flexion (_3)		H2O RANDOM FOREST	219/94	0,67	0,31	0,59	0,22
HARMONICS + Descriptive data A normal (_1) vs A flexion (_3)	Overweight Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,69	0,32	0,63	0,27
HARMONICS + Descriptive data E flexion (_2) vs E extension (_3)	Overweight Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,70	0,34	0,67	0,29

Table 15. *Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in both training and testing groups, for AUC and Pr AUC values for the training group and the testing group respectively (Simplified).* Continuation

DATASET (PARAMETERS COMBINATION)	CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				AUC	Pr AUC	AUC	Pr AUC
HARMONICS + Descriptive data E normal (_1)	Cormack 1 and 4 cases	H2O RANDOM FOREST	100/43	0,77	0,16	0,83	0,32
HARMONICS + Descriptive data U all positions	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,90	0,31	0,82	0,22
ONLY descriptive data		H2O RANDOM FOREST	219/94	0,69	0,34	0,56	0,22
HARMONICS All vocals		H2O NAIVE BAYES	219/94	0,50	0,20	0,63	0,29
HARMONICS U all positions		H2O RANDOM FOREST	219/94	0,42	0,17	0,59	0,23
Spectral+ Descriptive data all vocals		H2O GRADIENT BOOSTING	219/94	0,56	0,22	0,60	0,27
Spectral+ Descriptive data O flexion		H2O NAIVE BAYES	219/94	0,70	0,31	0,67	0,23
Spectral+ Descriptive data A flexion vs A extension		H2O RANDOM FOREST	219/94	0,68	0,30	0,69	0,29
Voice parameters + Descriptive data A normal		H2O GENERALISED LINEAR MODEL	219/94	0,71	0,31	0,62	0,24
Voice parameters + Descriptive data O extension		H2O GRADIENT BOOSTING	219/94	0,65	0,35	0,64	0,25
Voice parameters + Descriptive data all vocals	Only Cormack 1 and 4	H2O RANDOM FOREST	100/43	0,76	0,14	0,75	0,25
Voice parameters + Descriptive data A normal	Only Cormack 1 and 4	H2O GENERALISED LINEAR MODEL	100/43	0,69	0,18	0,82	0,26

5.3.2.4. *Results of F1 and log loss metrics*

The F1 values and Log loss values of the classification algorithms for the different combinations in the testing group are summarized in the following text.

Furthermore, the results of both the training and testing groups and shown in detail in **table 16** (*simplified*) the full table is found in **annex 15**.

F1 values results:

- The combination harmonics + descriptive data + different vocals in different positions + all cases obtained a mean F1 value of 0.46.
- The combination of harmonics + descriptive data + different vocals in different positions + overweight Cormack grade 4 cases obtained a mean F1 value of 0.41.
- The combination of harmonics + descriptive data + different vocals in different positions + only analyzing Cormack I and IV cases obtained a mean F1 value of 0.52.
- Only descriptive data (without any voice parameters) + all cases obtained a F1 value of 0.41.
- Only harmonics of different vocals in different positions (without any descriptive data) + all cases obtained a mean F1 value of 0.42.
- Spectral of different vocals in different positions + descriptive data + all cases obtained a mean F1 value of 0.46.
- The combination of voice parameters (such as shimmer, jitter, HNR) + descriptive data + all cases obtained a mean F1 value 0.43.
- The combination of voice parameters (such as shimmer, jitter, HNR) + descriptive data + only analyzing Cormack I and IV cases obtained a mean F1 value of 0.54.

Log loss values results:

- The combination harmonics + descriptive data + different vocals in different positions + all cases obtained a mean Log loss value of 0.60.

- The combination of harmonics + descriptive data+ different vocals in different positions + overweighting Cormack grade 4 cases obtained a mean Log loss value of 0.60.
- The combination of harmonics + descriptive data + different vocals in different positions + only analyzing Cormack I and IV cases obtained a mean log loss value of 0.38.
- Only descriptive data (without any voice parameters) + all cases obtained a Log loss value of 0.52.
- Only harmonics of different vocals in different positions (without any descriptive data) + all cases obtained a Log loss value of 0.56.
- Spectral of different vocals in different positions + descriptive data + all cases obtained a Log loss value of 0.50.
- The combination of voice parameters (such as shimmer, jitter, HNR) + Descriptive data + all cases obtained a mean log loss value of 0.61.
- The combination of voice parameters (such as shimmer, jitter, HNR) + descriptive data + only analyzing Cormack I and IV cases obtained a mean Log loss value of 0.42.

Table 16. *Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in both training and testing groups, for F1 and Log loss values for the training group and the testing group respectively (Simplified).*

DATASET (PARAMETERS COMBINATION)	CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				F1	Log Loss	F1	Log Loss
HARMONICS + Descriptive data U normal (_1) vs U extension (_2)		H2O RANDOM FOREST	219/94	0,50	0,49	0,43	0,50
HARMONICS + Descriptive data U all positions		H2O NAIVE BAYES	219/94	0,52	0,88	0,42	
HARMONICS + Descriptive data O all positions	Overweight Cormack 4		219/94				
HARMONICS + Descriptive data U normal (_1) vs U flexion (_3)	Overweight Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,50	0,58	0,44	0,50

Table 16. *Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in both training and testing groups, for F1 and Log loss values for the training group and the testing group respectively (Simplified).*
Continuation

DATASET (PARAMETERS COMBINATION)	CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				F1	Log Loss	F1	Log Loss
HARMONICS + Descriptive data U all positions	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,78	0,47	0,50	
ONLY descriptive data		H2O RANDOM FOREST	219/94	0,50	0,60	0,41	0,52
HARMONICS All vocals		H2O NAIVE BAYES	219/94	0,36	0,56	0,41	
HARMONICS A normal (_1)		H2O NAIVE BAYES	219/94	0,35	0,73	0,47	0,48
HARMONICS A extension (_2)		H2O GENERALISED LINEAR MODEL	219/94	0,46	0,68	0,37	0,55
HARMONICS E all positions		H2O NAIVE BAYES	219/94	0,39		0,44	0,63
HARMONICS I normal (_1) vs I flexion (_3)		H2O NAIVE BAYES	219/94	0,40	0,86	0,41	0,52
Spectral+ Descriptive data O extension		H2O RANDOM FOREST	219/94	0,53	0,63	0,43	0,49
Voice parameters + Descriptive data all vocals		H2O GENERALISED LINEAR MODEL	219/94	0,44	0,65	0,42	0,51
Voice parameters + Descriptive data O flexion		H2O RANDOM FOREST	219/94	0,58	0,53	0,47	0,48
Voice parameters + Descriptive data all vocals	Only Cormack 1 and 4	H2O RANDOM FOREST	100/43	0,58	0,35	0,57	0,33
Voice parameters + Descriptive data O extension	Only Cormack 1 and 4	H2O GENERALISED LINEAR MODEL	100/43	0,47	0,50	0,53	0,43

5.3.2.5. *Results of TN, FN, FP, TP metrics*

The results for the true negatives, false negatives, false positives and true positive values of the classification algorithms for the different combinations in the testing group are summarized in the following text.

Furthermore, the results of both the training and testing groups and shown in detail in **table 17** (*simplified*) the full Table is found in **annex 16**.

TN, FN, FP, TP metrics:

- The combination harmonics + descriptive data + different vocals in different positions + all cases obtained a mean TN of 39.5; FN: 2.6; FP:35.4; TP:16.3.
- The combination of harmonics + descriptive data + different vocals in different positions + overweighting Cormack grade 4 cases obtained a mean TN of 27.4; FN: 1.6; FP:47.5; TP:17.3.
- The combination of harmonics + descriptive data + different vocals in different positions + only analyzing Cormack I and IV cases obtained a mean TN of 31.3; FN: 1.05; FP:6.6; TP:3.9.
- Only descriptive data (without any voice parameters) + all cases obtained a TN of 44; FN: 6; FP:31; TP:13.
- Only Harmonics of different vocals in different positions (without any descriptive data) + all cases obtained a mean TN of 35.6; FN: 3.8; FP: 39.3; TP 15.1.
- Spectral of different vocals in different positions + descriptive data + all cases obtained a mean TN of 45.5; FN: 4.6; FP: 29.5; TP 14.3.
- The combination of voice parameters (such as shimmer, jitter, HNR) + Descriptive data + all cases obtained a mean TN of 38.5; FN: 3.7; FP: 36.4; TP 15.2.
- The combination of voice parameters (such as shimmer, jitter, HNR) + Descriptive data + only analyzing Cormack I and IV cases obtained a mean TN of 32.1; FN: 1; FP: 5.85; TP 4.

Table 17. *Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in both training and testing groups, for true negatives, false negatives, false positives, true positive values for the training group and the testing group respectively (Simplified).*

DATASET (PARAMETERS COMBINATION)	CASES	BEST MODEL	N PATIENTS	TESTING 30%			
				TN	FN	FP	TP
HARMONICS + Descriptive data U normal (_1) vs U flexion (_3)		H2O RANDOM FOREST	219/94	30	1	45	18
HARMONICS + Descriptive data U normal (_1) vs U extension (_2)		H2O RANDOM FOREST	219/94	39	4	36	15
HARMONICS + Descriptive data U flexion (_3) vs U extension (_2)		H2O NAIVE BAYES	219/94	50	8	25	11
HARMONICS + Descriptive data U all positions		H2O NAIVE BAYES	219/94	42	5	33	14
HARMONICS + Descriptive data U Flexion (_3)	Overweight Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	26	0	49	19
HARMONICS + Descriptive data A normal (_1) vs A flexion (_3)	Overweight Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	26	0	49	19
HARMONICS + Descriptive data A normal (_1) vs A extension (_2)	Overweight Cormack 4	H2O NAIVE BAYES	219/94	28	1	47	18
HARMONICS + Descriptive data U normal (_1) vs U extension (_2)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	32	1	6	4
HARMONICS + Descriptive data U flexion (_3) vs U extension (_2)	Cormack 1 and 4 cases	H2O RANDOM FOREST	100/43	30	1	8	4
HARMONICS + Descriptive data U all positions	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	31	1	7	4
ONLY descriptive data		H2O RANDOM FOREST	219/94	44	6	31	13
HARMONICS All vocals		H2O NAIVE BAYES	219/94	31	3	44	16
HARMONICS A normal (_1)		H2O NAIVE BAYES	219/94	42	3	33	16
HARMONICS A extension (_2)		H2O GENERALISED LINEAR MODEL	219/94	14	1	61	18
HARMONICS A flexion (_3)		H2O NAIVE BAYES	219/94	53	6	22	13
HARMONICS E normal (_1)		H2O GENERALISED LINEAR MODEL	219/94	48	5	27	14
HARMONICS E extension (_2)		H2O RANDOM FOREST	219/94	38	5	37	14
HARMONICS E flexion (_3)		H2O GRADIENT BOOSTING	219/94	43	6	32	13

Table 17. *Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in both training and testing groups, for true negatives, false negatives, false positives, true positive values for the training group and the testing group respectively (Simplified).* Continuation

DATASET (PARAMETERS COMBINATION)	CASES	BEST MODEL	N PATIENTS	TESTING 30%			
				TN	FN	FP	TP
HARMONICS I normal (_1)		H2O NAIVE BAYES	219/94	38	5	37	14
HARMONICS I extension (_2)		H2O GRADIENT BOOSTING	219/94	35	5	40	14
HARMONICS I flexion (_3)		H2O NAIVE BAYES	219/94	36	3	39	16
Spectral+ Descriptive data A normal		H2O RANDOM FOREST	219/94	47	4	28	15
Voice parameters + Descriptive data E flexion vs E extension		H2O NAIVE BAYES	219/94	31	2	44	17
Voice parameters + Descriptive data O flexion vs O extension		H2O NAIVE BAYES	219/94	31	2	44	17
Voice parameters + Descriptive data all vocals	Only Cormack 1 and 4	H2O RANDOM FOREST	100/43	33	1	5	4
Voice parameters + Descriptive data A normal	Only Cormack 1 and 4	H2O GENERALISED LINEAR MODEL	100/43	31	1	7	4
Voice parameters + Descriptive data O extension	Only Cormack 1 and 4	H2O GENERALISED LINEAR MODEL	100/43	32	1	6	4

5.4. SELECTION OF THE BEST CLASSIFICATION ALGORITHMS FOR BOTH THE CORMACK AND THE MALLAMPATI STUDY

Finally, the best combinations of parameters and binary classification models that obtained the best results were selected. **Table 18. Annex 17**

In total, 3 combinations from the Mallampati and 4 from the Cormack study yielded the best results when predicting the Mallampati and the Cormack respectively.

In detail, according to the Mallampati study, the combination harmonics + descriptive data with different combinations of the vocals E, I normal position and O flexion position vs extension positions only analyzing Mallampati I and IV, achieved a mean AUC of 0.97, a mean F1 of 0.86, mean recall of 0.75, log loss values of 0.09 as well as Pr AUC of 0.58 and a false negative of 1.

Furthermore, in relation to the Cormack study, two combinations generated best results. On the one hand, harmonics + descriptive data all + vocal A in all positions only analyzing Cormack I and IV secured an AUC of 0.91, F1 of 0.67, recall of 0.60, precision of 0.75, log loss values of 0.24, 2 false negatives and an overall accuracy of 0.93%. On the other hand, voice parameters (including shimmer, jitter, HNR and others) + descriptive data + different combinations of vocals such as I flexion position, O normal position and O vocal all positions, only analyzing the Cormack I and IV cases, obtained a mean AUC of 0.90, mean F1 value of 0.67, mean recall of 0.70, a mean precision of 0.67, mean log loss values of 0.27, mean false negative values of 1.5 and an overall accuracy of 0.92%.

Table 18. *Best combinations of parameters and binary classifications models which obtained the best results for the Mallampati and the Cormack study for the training and testing datasets.*

MALLAMPATI STUDY		TRAINING							TESTING										
Combination	Model	Accuracy	Precision	Recall	F1	Log Loss	AUC	Pr AUC	Log Loss	Accuracy	Precision	Recall	F1	AUC	Pr AUC	TN	FN	FP	TP
HARMONICS+ DESCRIPTIVE DATA E normal position only Mallampati I and IV cases	H20 generalized linear	0,97	0,88	0,47	0,58	0,38	0,86	0,23	0,09	0,99	1,00	0,75	0,86	0,97	0,58	77	1	0	3
HARMONICS+ DESCRIPTIVE DATA I normal position only Mallampati I and IV cases	H20 generalized linear	0,97	0,88	0,47	0,58	0,38	0,86	0,23	0,09	0,99	1,00	0,75	0,86	0,97	0,58	77	1	0	3
HARMONICS+ DESCRIPTIVE DATA O flexion vs O extension position only Mallampati I and IV cases	H20 generalized linear	0,98	1,00	0,47	0,63	0,38	0,86	0,29	0,09	0,99	1,00	0,75	0,86	0,97	0,58	77	1	0	3

Table 18. *Best combinations of parameters and binary classifications models which obtained the best results for the Mallampati and the Cormack study for the training and testing datasets.* Continuation

CORMACK STUDY		TRAINING							TESTING										
Combination	Model	Accuracy	Precision	Recall	F1	Log Loss	AUC	Pr AUC	Log Loss	Accuracy	Precision	Recall	F1	AUC	Pr (AUC)	TN	FN	FP	TP
HARMONICS+ DESCRIPTIVE DATA A All positions only Cormack I and IV cases	H20 generalized linear	0,83	0,39	0,62	0,44	0,50	0,64	0,18	0,24	0,93	0,75	0,60	0,67	0,91	0,40	37	2	1	3
VOICE PARAMETERS+ DESCRIPTIVE DATA I flexion position only Cormack I and IV cases	H20 Random Forest	0,80	0,40	0,83	0,51	0,39	0,74	0,22	0,26	0,88	0,50	0,80	0,62	0,88	0,34	34	1	4	4
VOICE PARAMETERS+ DESCRIPTIVE DATA O normal position only Cormack I and IV cases	H20 Random Forest	0,55	0,29	1,00	0,43	0,44	0,66	0,18	0,24	0,93	0,67	0,80	0,73	0,90	0,45	36	1	2	4
VOICE PARAMETERS+ DESCRIPTIVE DATA O All positions only Cormack I and IV cases	H20 generalized linear	0,75	0,67	0,73	0,54	0,49	0,68	0,23	0,34	0,93	0,75	0,60	0,67	0,89	0,44	37	2	1	3

5.5. EVALUATION OF THE MALLAMPATI PERFORMANCE AS A PREDICTOR OF THE CORMACK GRADE IN OUR DATASET

What is more, we also analyzed the predictive power of Mallampati to predict Cormack in our population. In detail, the Mallampati obtained an AUC of 0.70, a F1 of 0.41, log loss value of 0.45, a precision and recall of both 0.42%, a Pr AUC of 0.44, 11 false negatives and an overall accuracy of 0.76%. **Table 19. Figure 66, 67.**

Table 19. *Metrics results of Mallampati vs Cormack.*

Mallampati vs Cormack H20 Generalised linear																
TRAINING 70% (N=219)							TESTING (30%) (N=94)									
Accuracy	Precision	Recall	F1	Log loss	AUC	Pr AUC	Accuracy	Precision	Recall	F1	Log loss	AUC	Pr AUC	TN	FN	FP	TP
0.69	0.39	0.69	0.48	0.62	0.68	0.30	0.76	0.42	0.42	0.41	0.45	0.70	0.44	64	11	11	8

Row ID	L Fàcil	L Difícil	D Error	S Rate
Fàcil	64	11	0.147	11 / 75
Difícil	11	8	0.579	11 / 19
Totals	75	19	0.234	22 / 94

Figure 66. *Confusion matrix (Testing sample 30% N= 94) Mallampati vs Cormack. True negatives = 64, False negatives = 11, True positives =8, False positives=11.*

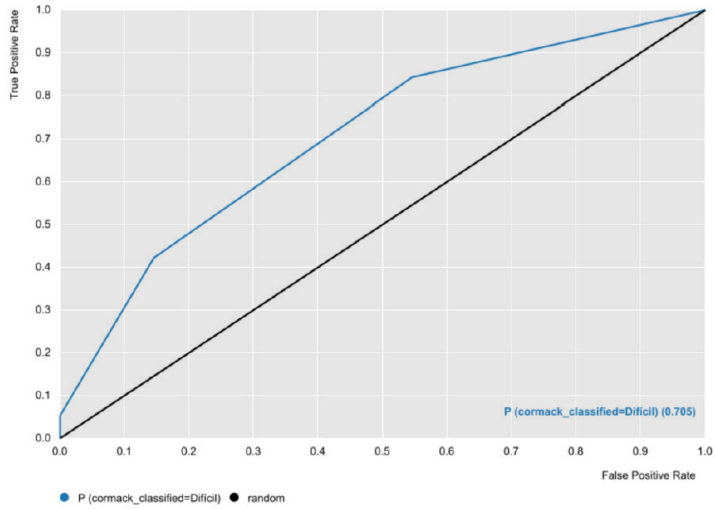


Figure 67. *ROC curve Mallampati as a Cormack predictor obtains a value of 0.70%.*

6. DISCUSSION

6. DISCUSSION

The preoperative detection of a difficult airway is crucial to ensure a safe airway management during surgical procedures as the identification of a difficult airway aids in the prevention of complications arising from a difficult airway (28,29). An airway assessment plays a paramount role in achieving optimal conditions of patient safety ensuring the highest standards in patient care. In the present study, an innovative approach to predicting a difficult airway has been explored, focusing on voice analysis as a potential predictor. This is not the first study to propose investigating this predictor as previous studies carried out by Carvalho et al in 2019 and Xia et al in 2021 also analyzed the relationship between voice parameters and a difficult airway management (81-83). However, these authors focused mainly on the voice formants as a potential predictor of a difficult mask ventilation and only one of these studies proposed the analysis of the voice as a predictor of a difficult laryngoscopy (82). Nevertheless, in these studies a significant relationship between the voice and a difficult airway was already demonstrated. In detail, Xia et al concluded that the first four formants presented a correlation with difficult mask ventilation (83). Carvalho et al also found significant association between three formants and Cormack-Lehane classification.

To elaborate further, Carvalho found a model with Mallampati and voice formants which achieved an AUC of 0.91 % for the prediction of a difficult laryngoscopy. On top of that, Xia et al, constructed a predictive model for a difficult mask ventilation, which included 20 voice parameters such as formants, pitch, bandwidth and obtained an AUC of 0.77%. Both authors conclude that voice parameters provide valuable information in airway evaluation and are a potential tool for the diagnosis of a difficult airway.

Additionally, their findings are consistent with the results found in this study as our outcomes support the initially hypothesized relationship between voice parameters and the presence of a difficult airway. However,

our findings differ in several aspects, as we have decided to take a step further and explore different approaches in terms of the analysis method used such as KNIME, the voice parameters utilized such as the harmonic instead of the formants and the introduction of additional clinical characteristics into the predictive algorithm such as smoking status or thyroid disorders. Building on this, we have aimed to improve the accuracy and efficiency of the voice as a potential predictor of a difficult airway.

Firstly, given the lack of a standard definition of a difficult airway in the literature (18), we have chosen to define a difficult intubation as obtaining a Cormack-Lehane grade III or IV. Yet, we are aware that the difficulty of the intubation is not solely determined by the degree of glottis structure visualizations as defined by the Cormack-Lehane grading scale and that other factors also come into play, such as the use of additional devices or the need for extra maneuvers to facilitate intubation (18). Nonetheless, we have decided to limit the definition to a Cormack-Lehane grade III or IV in order to standardize the data collection and enhance the reproducibility of the study.

Furthermore, unlike the previous studies that focused on formants and typical voice characterization parameters such as jitter and shimmer, we aimed to explore whether the analysis of harmonics might play a relevant role in improving the prediction of a difficult airway compared to formants. Moreover, formants contain important information about an individual's voice and female and male typical vocal folds vary in size, thus resulting in female voices with higher pitch and formants than male voices (105). On top of that, vowel production has often been characterized by measurements of the formants, especially the formant frequencies (172). However, it has been demonstrated that the measurement of formants is not straightforward due to the fact that formant frequencies are biased by the fundamental frequency particularly when fundamental frequency is high and/or the first formant is low (172). Furthermore, studies have reported formant measurement errors and inflation of formant variability in a wide range of fundamental frequency and formant combinations mostly due to F0 bias (173). Therefore, errors in formant identification are frequent and its measurement has high variability and may show inaccuracies in some cases. Furthermore, the pitch

of each subject varies and may even differ from sample to sample for the same person. Hence, absolute frequency varies across the samples and makes it difficult to extract any meaningful information. For this reason, we decided to focus on the harmonic amplitude data, defined as a multiple of the fundamental frequency, which might come up with a solution to identify patterns between patients. By normalizing the amplitude of each harmonic, it is possible to align them and compare samples. It should be mentioned at this stage that we have focused our efforts on analyzing the final output of the voice rather than the specific location where it has been generated. Our aim was to find a parameter with minimal variability which allows us to compare patients in order to identify a pattern and classify them as easy or difficult. On top of that, in contrast to the other authors we decided to explore the potential of analyzing the entire raw spectrogram without extracting any parameters, thus avoiding the biases derived from this extraction.

Furthermore, KNIME, the data analysis tool used, is able to deal with a great amount of data and has given us the opportunity to do so. Additionally, the Matlab tool was chosen over Praat, the most commonly used software, to manually extract the different voice parameters of interest owing to the fact that Praat does not allow personalized extraction whereas Matlab does (99). Furthermore, diverging from previous research we decided to explore the potential of incorporating additional clinical characteristics of the patients as well as voice pathologies associated with a difficult intubation such as diabetes, or thyroid disorders in a single prediction algorithm.

These conditions, specifically thyroid disorders have been associated with an incidence of a difficult tracheal intubation around 10% (174). Moreover, it is estimated that difficult tracheal intubation is ten times higher in patients suffering from long-term diabetes as compared to those not suffering from diabetes (175). For this reason, we thought that this approach not only may have the potential to improve our algorithms predictions, but also the importance of this particular analysis lies in the fact that currently there is no validated index that includes as many variables as the ones that we extract from the perioperative data. Particularly, there is no index that simultaneously analyzes and includes patient anthropometric variables such as age or weight along with concomitant pathologies related to difficult intubation such as

Thyroid disorders, COPD, sleep apnea or smoking habit, together with the traditional physical parameters of intubation difficulty such as the Mallampati score, degree of mouth opening and thyromental distance among others.

There were two main reasons for introducing combinations into the predictive algorithms in our study without including the voice. Firstly, we aimed to investigate the potential of a predictive algorithm that included multiple variables which had not yet been validated in the literature. Secondly, we intended to analyze the weight or the relevance of the voice within this algorithm.

This approach of considering these clinical characteristics in addition to voice parameters aligns with the findings of other studies which suggest that the predictive values for the model containing Mallampati and formants were higher than the values for single- variable models (82). This corresponds with the consensus among studies that airway evaluation should not rely solely on a single characteristic as currently available screening bedside tests have only poor to moderate discriminative power when used alone as compared to a combination of tests (56).

Furthermore, regarding our population sample in the Mallampati and the Cormack studies, the high patient attrition rate is striking. We attribute this high rate primarily to the fact that, despite developing a voice recording system that automatically detected any issues with the audio recording, it did not completely prevent a significant portion of the samples from being discarded due to incorrect recording because of external noise interferences. Another factor which contributed to data loss was missing clinical information that was not collected during the clinical interview during the preanesthetic evaluation. We associate this fact to the high workload in our preanesthetic evaluation service which often led to questions being overlooked. Additionally, if we focus on the analysis of the characteristics of the patients included in the study, our results show that the wide majority of the patients were relatively healthy as 96.2% and 94.5% of the patients in both the Mallampati and the Cormack study respectively, were considered ASA I or II. However, the mean obtained BMI index was 25.7 which indicated a tendency of the population in the sample to be overweight.

On top of that, there was a slightly higher percentage of women compared to men in both studies, 47.81% of males and 52.18% of women in the Mallampati study; and 46.65% of men and 53.35% of women in the Cormack study. Furthermore, when it comes to airway evaluation a considerable percentage of patients were tested positive for at least one parameter collected in traditional airway evaluation which predicts increased airway difficulty. However, this percentage was almost double in the Mallampati study than in the Cormack study, 43% of patients compared to 20% respectively. It is important to note at this juncture that the number of patients which meet the target class for the difficult airway prediction group classified as Mallampati III or IV was only 21.04%.

Besides, similar results were found in the Cormack study where only 20,13% of the patients were classified as Cormack III or IV. Therefore, in our sample we obtained fewer patients from the class we aimed to predict, in this case, the difficult ones and conversely, a higher number of easy patients, resulting in an imbalanced dataset. This fact may have contributed to an increased difficulty for the predictive algorithms to identify patterns as there are fewer patients in the target group.

Additionally, it has compelled us to employ analysis techniques such as using only extreme cases, specifically groups I and IV in both studies to find greater differences between the groups. Furthermore, we have also implemented specific machine learning techniques to address imbalanced datasets such as overweight groups Mallampati and Cormack IV. For this reason, we believe that despite having an imbalanced dataset any additional bias may not have been introduced in our study since we have employed all the techniques to effectively minimize it. However, we cannot affirm that these factors have not influenced the performance of the models to some extent. Furthermore, due to the fact that difficult airway cases have a low incidence in the general population of around 5.8% (23) it is very difficult to find an equal number of easy and difficult patients in the study dataset and it is common to find these inequalities in all the studies related to the prediction of a difficult airway.

Additionally, in relation to the Mallampati study, our findings revealed noteworthy observations. Firstly, the outcomes obtained from the machine

learning classification models in this study offer valuable insights into the capabilities of these models for predicting a Mallampati class III or IV. The results indicate that the developed models achieved robust performance across different evaluation metrics, including AUC curve values, F-1 score, accuracy, precision and recall in most of the combinations. This shows the potential of machine learning algorithms to effectively classify and predict the target class based on the selected set of features. However, when analyzing the combinations in which only the harmonics without any descriptive data are introduced into the model, the performance of evaluation metrics decreases notably. Furthermore, these results are consistent across all metrics and combinations, indicating that voice analysis alone might not be a reliable method for classifying patients and additional descriptive data is required to improve the predictions. For this reason, it is important to highlight that our results show that adding voice parameters to the descriptive data enhances predictions in the majority of combinations. Therefore, we can assert that the voice does contribute positively to the classification of patients as easy or difficult in both studies.

According to the Mallampati study, the best combinations found were the ones which combined both Harmonics and descriptive data in E normal position, I normal position and O flexion vs O extension only in patients with Mallampati I and IV.

Furthermore, it is not surprising that in the vast majority of combinations of metrics, the results obtained by only analyzing extreme cases yielded better results. This might be due to the fact that intermediate values such as Mallampati II or III could have been subjected to bias from the anesthesiologists who assessed them as in some patients it might be difficult to differentiate between one category or the other, therefore leading to a potential mislabeling of data. For this reason, patients belonging to extreme categories are easier to classify for anesthesiologists.

Therefore, it is less likely that extreme values such as Mallampati I and IV are prone to this type of error, allowing high quality data to be incorporated into the algorithm. This explains the attainment of better results in cases where only patients from Mallampati I and IV were included as the analyzed

data contained more relevant information and it was of a better quality. This is important as if poor data is passed into an algorithm, the output data can be erroneous (176).

Additionally, our results show an excellent discriminative power and generalization capabilities, which is desirable in classification problems in all the three selected combinations which yielded best results. Moreover, the obtained results are nearly identical among the three combinations showing values of an AUC of 0.97 in the testing group indicating excellent performance. Besides, AUC is a metric which assesses the model's ability to differentiate between positive and negative classes. Consequently, obtaining an AUC of 0.97 suggests that the model has a high probability of ranking a randomly chosen positive case higher than a negative one, therefore making highly reliable predictions. Furthermore, the results obtained in the training group were also high, achieving an AUC of 0.86. Accordingly, it is important to highlight that the results obtained regarding the AUC in both the training and the testing group are high.

To elaborate further, we believe that the possibility of overfitting, which is when the model does not generalize well from observed data to unseen data (177) is unlikely due to the small difference between the two results and those results being higher in the testing group (unseen data). Additionally, measures to avoid overfitting such as cross validation in the training group have been applied.

We have also considered the evaluation of other metrics which have provided us with a more comprehensive view of the model. All the combinations yielded similar results, achieving great accuracy values from 0.97-0.98%, which means that the models correctly predicted the outcome for the majority of the cases in the dataset.

This accuracy represents the proportion of correct predictions out of the total number of predictions made by the model. Nevertheless, accuracy might not be the main metric to focus our attention on owing to the fact that it alone does not distinguish between correctly predicting positives or negatives as it analyzes the proportion of correct predictions on the whole. To elaborate further, in our daily clinical practice it is more important to detect

positive cases due to the serious consequences of not identifying a patient with a difficult airway which cannot be intubated, rather than the negatives, where intubation most likely will be achieved on the first attempt.

Therefore, in our case it is crucial that the model detects the positive cases accurately. For this reason, additional metrics are needed to evaluate the models, such as the number of false negatives. In light of this, having a very low number of false negatives is essential in our clinical context due to the severe clinical consequences that could be derived from a positive case to go unnoticed. Besides, in our models only one positive case was incorrectly identified as negative, which indicates that the model has high sensitivity in detecting positive cases and a relatively low error rate in identifying these cases. Moreover, the recall values obtained were also moderately high of 0.75 suggesting that the model successfully captured a significant portion of the positive instances. However, in the training group the recall result values were only 0.47%. This suggests that the model performed better identifying positive cases in the unseen testing data compared to the training data and these discrepancies might mainly be due to the fact that the sample size is small. In addition, in all the three combinations the same results for precision of 1 in the testing group were obtained, meaning that all the instances predicted as positive by the model were indeed true positive cases, suggesting that there were no false positive predictions.

Moreover, having no false positive predictions is not crucial but highly desirable in the clinical context as it would allow us to be highly confident in the correctness of that prediction, allowing for a better optimization of resources by not requiring extra material or personnel resources for a patient who can be easily intubated using the standard clinical protocol. What is more, the log loss results obtained of 0.09 in the testing group and 0.38 in the training group suggest that the model has a low level of uncertainty in its predictions, indicating that the predicted probabilities align closely with the true labels in the testing data. In addition, the high F1 values of 0.86 in the testing group results indicate a balanced performance between precision and recall. However, these values are significantly higher than the ones obtained in the training group 0.57-0.63, which suggests the model is able to generalize well to unseen data or indicate different data distribution between

the training and the testing groups which may cause the model's performance to vary. Furthermore, when analyzing the results of Pr AUC values in the three combinations, we also obtained better results in the testing group than in the training one indicating relatively low model performance on the training data. However, it improves notably in the testing group, which might be due to the generalization capabilities of the model.

Overall, the obtained results show that the models are capable of generalizing to new data and that the achieved results are reliable. Moreover, although it is known that variable outcomes for sensitivity and specificity have been reported in the literature for Mallampati (41), it is still widely used by anesthesiologists in their clinical practice. The relevance of our findings lies in the fact that predicting Mallampati with a tool based on the voice may replace the physical examination needed when traditionally assessing the Mallampati metric, thus offering an alternative approach to Mallampati evaluation. Finding a predictor substitute for Mallampati could provide a more accessible and convenient method for evaluating the airway. Ultimately this could reduce the need for a physical examination in certain cases and, contribute to improving efficiency and patient outcomes.

On the other hand, the results obtained from the predictive model algorithm in the Cormack study have also provided valuable insights. Firstly, the findings indicate that the developed algorithm has the potential to be a reliable tool for the prediction of a difficult intubation. In detail, four combinations which yielded the best results were selected. Similarly, to the Mallampati study, the ones that performed best were those which only analyzed extreme cases such as Cormack I and IV.

However, contrary to the Mallampati study three out of the four combinations included voice parameters such as jitter and shimmer along with descriptive data and only in one of the selected combinations, harmonics were included together with descriptive voice parameters. On the other hand, in the Mallampati study, all the best combinations included the harmonics and not the voice parameters. Moreover, all the models yielded moderate to excellent accuracy (0.88-0.91%) obtaining the best results in the harmonics combination and vocal A in all positions. It is important to note that comparably to the

Mallampati study differences of the achieved results in the metrics evaluations differ notably from the training to the testing group being significantly higher in the second one.

As mentioned before, the small sample size could explain these results. Moreover, the results indicate that the developed models achieved robust performance across different evaluation metrics such as AUC values. Regarding this metric, in all the combinations an AUC value ranging from 0.61-0.83 was obtained, which indicates a reasonably strong to high level of discrimination showing consistent and reliable performance. However, in the combination of only descriptive data the model obtained an AUC of 0.56 suggesting that its predictive performance is slightly better than random chance. In light of these results and similarly to the Mallampati study, we can affirm that the voice might play an important role in predicting a difficult airway as it significantly increases the AUC when introduced into the different algorithm combinations compared to only analyzing the descriptive data without the voice.

Furthermore, the models also obtained moderate to good performance in terms of identifying positive instances correctly and capturing a reasonable proportion of actual positive instances as the results obtained for precision and recall in the four selected combinations ranged from 0.50-0.75% and 0.60-0.80% respectively.

On top of that, these results are aligned with obtaining an average F1-scores results in the testing group ranging from 0.62 to 0.73, which also indicate a moderate to good performance of the model. F1 score is a metric that combines precision and recall into a single value, pointing to the fact that the model achieves a reasonable balance between precision and recall successfully identifying a significant proportion of positive cases while minimizing both false positives and false negatives.

Finally, the outcomes achieved indicate that the model's predicted probabilities align with the actual outcomes as log loss values for the selected combinations ranged from 0.24-0.34 suggesting that the probabilistic predictions are relatively accurate, as the closer the log loss values are to 0, the better the match between predicted probabilities and true labels.

Additionally, other results worth mentioning are the analysis of the predictive power of the Mallampati test assessment in our population. When comparing the Mallampati test evaluated by the anesthesiologists in our department and then the correlation of these predictions to the Cormack grades obtained in our sample, it is observable that the AUC result values obtained are 0.73%. Therefore, this indicates a moderate discriminatory power to correctly identify the Cormack of the patients.

Nevertheless, the ability of our developed models which introduce the harmonics + the descriptive data succeed in significantly improving this prediction reaching an AUC of 0.91%. This indicates excellent performance in terms of its ability to discriminate between positive and negative instances. In the view of these results, it can be suggested that the predictive capabilities of the algorithm are confident, making it a valuable asset in the practical scenario of predicting a difficult airway. It is important to highlight that in this particular context the choice of the best models for both studies were based on a joint evaluation of all metrics which were considered relevant to the clinical problem rather than on a single metric. However, priority has been given to those models performing with high AUC as the ROC curve is used to assess the overall diagnostic performance of a test and to compare the performance of two or more diagnostic tests (178). However, when choosing the best models, a balance of all the metrics used to evaluate a binary classification model have been considered.

Furthermore, it is paramount to acknowledge the limitations and potential sources of bias in the study. One of the main limitations observed is that the sample size is small and it would be advisable for the study to undergo external validation to ensure the reproducibility of the results. Consequently, further research and validation on a larger dataset is necessary to confirm the generalizability and robustness of the developed algorithms. It is possible that if the sample size is increased and new data is introduced into the algorithm the obtained results will be different from the ones found in this study. Hence, new combinations that yield better results compared to the current ones can be found.

Another limitation is that voice parameters exhibit significant individuals' variability. Voice parameters have been shown to be influenced by a variety

of factors such as gender or altered in pathologies with higher values in voices with severe deviations (111). This variability may have affected our data analysis. In addition to this fluctuation in the voice parameters, another limitation to consider is the fact that the range of extension or flexion of the neck was not measured. Thus, patients could have been subjected to different degrees of neck movement meaning there might have been a variation in terms of the level of stress imposed on their voices, which may well have altered the results.

Furthermore, for future studies, it would be interesting to monitor and observe how this phenomenon affects the results. Besides, voice recordings were performed in a sitting position, whereas intubation occurs in supine position. For this reason, it should be noted that when the patients are intubated, they are relaxed under the effects of neuromuscular relaxant and other medications, which contribute to the relaxation of all pharyngeal structures. Thus, this causes changes in the patient's baseline conditions, which might feasibly affect the study outcomes.

Moreover, it should be also mentioned that improving our data collection and voice recordings method would have allowed us to have fewer patient losses and increase the total sample size, enabling us to test our hypothesis in a larger proportion of patients.

To conclude, our findings contribute to the existing body of knowledge by offering a novel approach and advancing in the field preoperatively predicting a difficult airway. The results obtained so far present promising opportunities for future applications such as a telematic preanesthetic consultation. However, further exploration and refinement of the algorithm to enhance its performance and utility should be performed.

7. CONCLUSIONS

7. CONCLUSIONS

1. The predictive algorithms based on machine learning are useful as a tool for predicting a difficult intubation, due to the fact that when the acoustic parameters of the voice along with the descriptive data such as the clinical characteristics of the patients when introduced into classification algorithms are able to predict a difficult intubation.
2. There is a relationship between the anatomy of the upper respiratory tract and the following voice parameters: Harmonics, jitter, shimmer, HNR and spectral.
3. There is a correspondence between the analysis of acoustic voice parameters and the grade of the Cormack Classification scale as it enables the classification of patients into different groups based on the degree of observed laryngeal structures according to the Cormack classification.
4. There is a correlation between the voice parameters with the Mallampati test, as it enables the classification of patients into different groups according to the Mallampati Class.
5. The model's combinations which yielded the best results for the prediction of Cormack III or IV were the ones including Harmonic + descriptive data of patients in vocal A all positions and Voice parameters + descriptive data vocals I flexion position, O normal position and O all positions respectively when only including patients with Cormack III and IV achieving an AUC value of 0.91%
6. The model's combinations which yielded the best results for the prediction of Mallampati were the ones including Harmonics + descriptive data of the patients in vocals E and I in normal position and O flexion vs O extension when only including patients with Mallampati III and IV, reaching an AUC value of 0.97%

7. The voice parameters such as jitter, shimmer and harmonics contribute to the prediction of a difficult intubation as it significantly increases the AUC when introduced into the different algorithm combinations in comparison to the ones in which only the descriptive data without the voice is analyzed.
8. Our findings show that it is possible to find a tool based on the voice and descriptive data of the patient with comparable predictive power to Mallampati test which may potentially allow the replacement of the physical presence of the patient when assessing the Mallampati grades during the preanesthetic visit.
9. The outcomes derived from this study demonstrate the potential utility of predictive algorithms as a non-invasive, voice-based tool for predicting a difficult intubation preoperatively.
10. The results of this study present promising opportunities for future applications such as telematic pre-anesthetic consultation. What could mean a significant change in our daily clinical practice as it would allow for the online visits low-risk patients, as the only limitation to this type of visits currently is the physical examination and concretely the airway evaluation. Therefore, this would allow for better resource management, thus improving the efficiency of healthcare processes and a change in the current paradigm of the anesthesiologist's daily practice.
11. Further validation of this developed tool is needed including studies with a larger sample size of patients to validate it and obtain more reliable results, as well as to analyze its usefulness and effectiveness.

8. BIBLIOGRAPHY

8. BIBLIOGRAPHY

1. Cooper JB. Critical Role of the Surgeon-Anesthesiologist Relationship for Patient Safety. Anesthesiology [Internet]. 2018 [cited 2023 May 22];129(3):402–5. Available from: https://pubmed.ncbi.nlm.nih.gov/30045093/
2. Practice Advisory for Preanesthesia EvaluationAn Updated Report by the American Society of Anesthesiologists Task Force on Preanesthesia Evaluation. Anesthesiology [Internet]. 2012 Mar 1 [cited 2023 May 4];116(3):522–38. Available from: https://pubs.asahq.org/anesthesiology/article/116/3/522/13033/Practice-Advisory-for-Preanesthesia-EvaluationAn
3. Ashok A, Niyogi D, Ranganathan P, Tandon S, Bhaskar M, Karimundackal G, et al. The enhanced recovery after surgery (ERAS) protocol to promote recovery following esophageal cancer resection. Surg Today. 2020 Apr 1;50(4):323–34.
4. Harris K, Søfteland E, Moi AL, Harthug S, Storesund A, Jesuthasan S, et al. Patients' and healthcare workers' recommendations for a surgical patient safety checklist - a qualitative study. BMC Health Serv Res [Internet]. 2020 Jan 16 [cited 2023 May 4];20(1). Available from: https://pubmed.ncbi.nlm.nih.gov/31948462/
5. Gillis C, Ljungqvist O, Carli F. Prehabilitation, enhanced recovery after surgery, or both? A narrative review. Br J Anaesth [Internet]. 2022 Mar 1 [cited 2023 May 4];128(3):434–48. Available from: https://pubmed.ncbi.nlm.nih.gov/35012741/
6. Avva U, Lata JM, Kiel J. Airway Management. Surg Perioper Manag Patients with Anat Anomalies [Internet]. 2023 Jan 2 [cited 2023 May 4];57–78. Available from: https://www.ncbi.nlm.nih.gov/books/NBK470403/
7. Alvarado AC, Panakos P. Endotracheal Tube Intubation Techniques. StatPearls [Internet]. 2022 Jul 13 [cited 2023 May 4]; Available from: https://www.ncbi.nlm.nih.gov/books/NBK560730/
8. De Hert S, Staender S, Fritsch G, Hinkelbein J, Afshari A, Bettelli G, et al. Pre-operative evaluation of adults undergoing elective noncardiac surgery: Updated guideline from the European Society of Anaesthesiology. Eur J Anaesthesiol [Internet]. 2018 Jun 1 [cited 2023 May 11];35(6):407–65. Available from: https://pubmed.ncbi.nlm.nih.gov/29708905/

9. Hansel J, Rogers AM, Lewis SR, Cook TM, Smith AF. Videolaryngoscopy versus direct laryngoscopy for adults undergoing tracheal intubation. Cochrane database Syst Rev [Internet]. 2022 Apr 4 [cited 2023 May 5];4(4). Available from: https://pubmed.ncbi.nlm.nih.gov/35373840/

10. Ahmad I, El-Boghdadly K, Bhagrath R, Hodzovic I, McNarry AF, Mir F, et al. Difficult Airway Society guidelines for awake tracheal intubation (ATI) in adults. Anaesthesia [Internet]. 2020 Apr [cited 2023 May 5];75(4):509–28. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/anae.14904

11. Higgs A, Mcgrath BA, Goddard C, Rangasami J. UK 10 , 6 Anaesthesia & Intensive Care Medicine, Countess of Chester Hospital NHS Foundation Trust, Chester, UK 11 and 7 Anaesthesia and Intensive Care Medicine, Royal United Hospitals Bath NHS Foundation Trust, Bath, UK 12. London North West Healthc NHS Trust. 2018;120(2):323–52.

12. Azam Danish M. Preoxygenation and Anesthesia: A Detailed Review. Cureus [Internet]. 2021 Feb 9 [cited 2023 May 5];13(2). Available from: https://pubmed.ncbi.nlm.nih.gov/33728189/

13. Hung TY, Lin LW, Yeh YH, Su YC, Lin CH, Yang TF. The evaluation of a better intubation strategy when only the epiglottis is visible: a randomized, crossover mannequin study. BMC Anesthesiol [Internet]. 2019 Oct 1 [cited 2023 May 24];19(1). Available from: /pmc/articles/PMC6329169/

14. Russotto V, Rahmani LS, Parotto M, Bellani G, Laffey JG. Tracheal intubation in the critically ill patient. Eur J Anaesthesiol. 2022 May 1;39(5):463–72.

15. Sakles JC, Ross C, Kovacs G. Preventing unrecognized esophageal intubation in the emergency department. J Am Coll Emerg Physicians open [Internet]. 2023 Jun 29 [cited 2023 May 5];4(3):e12951. Available from: http://www.ncbi.nlm.nih.gov/pubmed/37128296

16. Russotto V, Myatra SN, Laffey JG, Tassistro E, Antolini L, Bauer P, et al. Intubation Practices and Adverse Peri-intubation Events in Critically Ill Patients From 29 Countries. JAMA [Internet]. 2021 Mar 23 [cited 2023 May 5];325(12):1164–72. Available from: https://pubmed.ncbi.nlm.nih.gov/33755076/

17. Driver BE, Semler MW, Self WH, Ginde AA, Trent SA, Gandotra S, et al. Effect of Use of a Bougie vs Endotracheal Tube With Stylet on Successful Intubation on the First Attempt Among Critically Ill Patients Undergoing Tracheal Intubation: A Randomized Clinical Trial. JAMA [Internet]. 2021 Dec 28 [cited 2023 May 5];326(24):2488–97. Available from: https://pubmed.ncbi.nlm.nih.gov/34879143/

18. Apfelbaum JL, Hagberg CA, Connis RT, Abdelmalak BB, Agarkar M, Dutton RP, et al. 2022 American Society of Anesthesiologists Practice Guidelines for Management of the Difficult Airway. Anesthesiology [Internet]. 2022 Jan 1 [cited 2022 Oct 4];136(1):31–81. Available from: https://pubmed.ncbi.nlm.nih.gov/34762729/

19. López AM, Belda I, Bermejo S, Parra L, Áñez C, Borràs R, et al. Recomendaciones para la evaluación y manejo de la vía aérea difícil prevista y no prevista de la Societat Catalana d'Anestesiologia, Reanimació i Terapèutica del Dolor, basadas en la adaptación de guías de práctica clínica y consenso de expertos. Rev Esp Anestesiol Reanim [Internet]. 2020 Jun 1 [cited 2022 Oct 4];67(6):325–42. Available from: https://www.elsevier.es/es-revista-revista-espanola-anestesiologia-reanimacion-344-articulo-recomendaciones-evaluacion-manejo-via-aerea-S003493561930249X

20. CORMACK RS, LEHANE J. Difficult tracheal intubation in obstetrics. Anaesthesia [Internet]. 1984 Nov [cited 2020 Apr 17];39(11):1105–11. Available from: http://doi.wiley.com/10.1111/j.1365-2044.1984.tb08932.x

21. Yentis SM, Lee DJH. Evaluation of an improved scoring system for the grading of direct laryngoscopy. Anaesthesia [Internet]. 1998 Nov [cited 2023 May 9];53(11):1041–4. Available from: https://pubmed.ncbi.nlm.nih.gov/10023271/

22. Başpinar ŞM, Günüşen İ, Sergin D, Sargin A, Balcioğlu ST. Evaluation of anthropometric measurements and clinical tests in the diagnosis of difficult airway in patients undergoing head and neck surgery. Turkish J Med Sci [Internet]. 2022 [cited 2023 May 7];52(3):730–40. Available from: https://pubmed.ncbi.nlm.nih.gov/36326344/

23. Muzien SJ, Gelgelo KG, Woldeyohans M. Incidence of difficult airway among adult patients of undiagnosed obstructive sleep apnea who are undergoing elective surgery: a prospective cohort study. Int J Surg Glob Heal [Internet]. 2022 Jul [cited 2023 May 24];5(4):e76–e76. Available from: https://journals.lww.com/ijsgh/Fulltext/2022/07010/Incidence_of_difficult_airway_among_adult_patients.2.aspx

24. Klucka J, Kosinova M, Kratochvil M, Marecek L, Kovalcikova P, Urik M, et al. Difficult airway prediction in paediatric anaesthesia (Diffair): Prospective observational study. Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub [Internet]. 2021 Sep 1 [cited 2023 May 7];165(3):298–304. Available from: https://pubmed.ncbi.nlm.nih.gov/32424374/

25. Lee YL, Leanne Lim M, Ling Leong W, Lew E. Difficult and failed intubation in Caesarean general anaesthesia: a four-year retrospective review. Singa-

pore Med J [Internet]. 2022 Mar 1 [cited 2023 May 7];63(3):152–6. Available from: https://pubmed.ncbi.nlm.nih.gov/32798358/

26. Law JA, Duggan L V., Asselin M, Baker P, Crosby E, Downey A, et al. Canadian Airway Focus Group updated consensus-based recommendations for management of the difficult airway: part 2. Planning and implementing safe management of the patient with an anticipated difficult airway. Can J Anaesth [Internet]. 2021 Sep 1 [cited 2022 Oct 4];68(9):1405. Available from: /pmc/articles/PMC8186352/
27. Law JA, Duggan L V., Asselin M, Baker P, Crosby E, Downey A, et al. Canadian Airway Focus Group updated consensus-based recommendations for management of the difficult airway: part 1. Difficult airway management encountered in an unconscious patient. Can J Anaesth [Internet]. 2021 Sep 1 [cited 2023 May 7];68(9):1373. Available from: /pmc/articles/PMC8212585/
28. Nørskov AK, Rosenstock C V., Wetterslev J, Astrup G, Afshari A, Lundstrøm LH. Diagnostic accuracy of anaesthesiologists' prediction of difficult airway management in daily clinical practice: a cohort study of 188064 patients registered in the Danish Anaesthesia Database. Anaesthesia [Internet]. 2015 Mar [cited 2019 Mar 10];70(3):272–81. Available from: http://doi.wiley.com/10.1111/anae.12955
29. Joffe AM, Aziz MF, Posner KL, Duggan L V., Mincer SL, Domino KB. Management of difficult tracheal intubation: A closed claims analysis. Anesthesiology. 2020 Oct 1;818–29.
30. Cook TM, Woodall N, Frerk C. Major complications of airway management in the UK: results of the Fourth National Audit Project of the Royal College of Anaesthetists and the Difficult Airway Society. Part 1: Anaesthesia †. Br J Anaesth [Internet]. 2011 May 1 [cited 2018 Oct 21];106(5):617–31. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0007091217332099
31. Cook TM, Woodall N, Frerk C. Major complications of airway management in the UK: results of the Fourth National Audit Project of the Royal College of Anaesthetists and the Difficult Airway Society. Part 1: Anaesthesia †. Br J Anaesth [Internet]. 2011 May [cited 2019 Apr 27];106(5):617–31. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0007091217332099
32. Crosby ET, Duggan L V., Finestone PJ, Liu R, De Gorter R, Calder LA. Anesthesiology airway-related medicolegal cases from the Canadian Medical Protection Association. Can J Anaesth [Internet]. 2021 Feb 1 [cited 2023 May 6];68(2):183–95. Available from: https://pubmed.ncbi.nlm.nih.gov/33200320/

33. Kheterpal S, Healy D, Aziz MF, Shanks AM, Freundlich RE, Linton F, et al. Incidence, Predictors, and Outcome of Difficult Mask Ventilation Combined with Difficult Laryngoscopy. Anesthesiology [Internet]. 2013 Dec 1 [cited 2019 Mar 6];119(6):1360–9. Available from: http://anesthesiology.pubs.asahq.org/Article.aspx?doi=10.1097/ALN.0000435832.39353.20

34. Caplan RA, Posner KL, Ward RJ, Cheney FW. Adverse respiratory events in anesthesia: A Closed Claims Analysis. Anesthesiology. 1990;72(5):828–33.

35. Cook TM, Macdougall-Davis SR. RESPIRATION AND THE AIRWAY Complications and failure of airway management. Br J Anaesth. 2012;109:i68–85.

36. Joffe AM, Aziz MF, Posner KL, Duggan L V., Mincer SL, Domino KB. Management of difficult tracheal intubation: a closed claims analysis. Anesthesiology [Internet]. 2019 [cited 2022 Oct 4];131(4):818. Available from: /pmc/articles/PMC6779339/

37. Hove LD, Steinmetz J, Christoffersen JK, Møller A, Nielsen J, Schmidt H. Analysis of deaths related to anesthesia in the period 1996-2004 from closed claims registered by the Danish Patient Insurance Association. Anesthesiology [Internet]. 2007 Apr [cited 2023 May 8];106(4):675–80. Available from: https://pubmed.ncbi.nlm.nih.gov/17413904/

38. Fornebo I, Simonsen KA, Bukholm IRK, Kongsgaard UE. Claims for compensation after injuries related to airway management: a nationwide study covering 15 years. Acta Anaesthesiol Scand [Internet]. 2017 Aug 1 [cited 2023 May 8];61(7):781–9. Available from: https://www.researchgate.net/publication/317275934_Claims_for_compensation_after_injuries_related_to_airway_management_A_nationwide_study_covering_15_years

39. Lee JH. Difficult intubation: lessons learned from the courts of South Korea. Korean J Anesthesiol [Internet]. 2021 Dec 1 [cited 2023 May 8];74(6):463. Available from: /pmc/articles/PMC8648503/

40. Flin R, Fioratou E, Frerk C, Trotter C, Cook TM. Human factors in the development of complications of airway management: preliminary evaluation of an interview tool. Anaesthesia [Internet]. 2013 Aug [cited 2023 May 8];68(8):817–25. Available from: https://pubmed.ncbi.nlm.nih.gov/23682749/

41. Roth D, Pace NL, Lee A, Hovhannisyan K, Warenits AM, Arrich J, et al. Airway physical examination tests for detection of difficult airway management in apparently normal adult patients. Cochrane Database Syst Rev [Internet]. 2018 May 15 [cited 2022 Nov 6];2018(5). Available from: /pmc/articles/PMC6404686/

42. Roth D, Pace NL, Lee A, Hovhannisyan K, Warenits AM, Arrich J, et al. Bedside tests for predicting difficult airways: an abridged Cochrane diagnostic test accuracy systematic review. Anaesthesia [Internet]. 2019 Jul 1 [cited 2022 Nov 6];74(7):915–28. Available from: https://pubmed.ncbi.nlm.nih.gov/30843190/

43. Mallampati SR, Gatt SP, Gugino LD, Desai SP, Waraksa B, Freiberger D, et al. A clinical sign to predict difficult tracheal intubation: a prospective study. Can Anaesth Soc J [Internet]. 1985 Jul [cited 2022 Nov 6];32(4):429–34. Available from: https://pubmed.ncbi.nlm.nih.gov/4027773/

44. SAMSOON GLT, YOUNG JRB. Difficult tracheal intubation: a retrospective study. Anaesthesia. 1987;42(5):487–90.

45. SAMSOON GLT, YOUNG JRB. Difficult tracheal intubation: a retrospective study. Anaesthesia [Internet]. 1987 [cited 2023 May 10];42(5):487–90. Available from: https://pubmed.ncbi.nlm.nih.gov/3592174/

46. Sepúlveda Haro E, Raigón Ponferrada A, Ramírez Aliaga M, Galache Laza M, Guerrero Orriach JL, Mañas JC. Mallampati class Zero airway: a narrative review. Minerva Anestesiol. 2022 May 1;88(5):390–5.

47. Roth D, Pace NL, Lee A, Hovhannisyan K, Warenits AM, Arrich J, et al. Airway physical examination tests for detection of difficult airway management in apparently normal adult patients. Vol. 2018, Cochrane Database of Systematic Reviews. John Wiley and Sons Ltd; 2018.

48. Lee A, Fan LTY, Gin T, Karmakar MK, Kee WDN. A systematic review (meta-analysis) of the accuracy of the mallampati tests to predict the difficult airway. Anesth Analg [Internet]. 2006 [cited 2023 May 10];102(6):1867–78. Available from: https://journals.lww.com/anesthesia-analgesia/Fulltext/2006/06000/A_Systematic_Review__Meta_Analysis__of_the.50.aspx

49. Shiga T, Wajima Z, Inoue T, Sakamoto A. Predicting Difficult Intubation in Apparently Normal PatientsA Meta-analysis of Bedside Screening Test Performance. Anesthesiology [Internet]. 2005 Aug 1 [cited 2022 Nov 23];103(2):429–37. Available from: https://pubs.asahq.org/anesthesiology/article/103/2/429/9268/Predicting-Difficult-Intubation-in-Apparently

50. Lundstrøm LH, Vester-Andersen M, Møller AM, Charuluxananan S, L'hermite J, Wetterslev J. Poor prognostic value of the modified Mallampati score: a meta-analysis involving 177 088 patients. 2011;

51. Prakash S, Kumar A, Bhandari S, Mullick P, Singh R, Gogia AR. Difficult laryngoscopy and intubation in the Indian population: An assessment of anatomical and clinical risk factors. Indian J Anaesth [Internet]. 2013 Nov [cited 2023 May 10];57(6):569. Available from: /pmc/articles/PMC3883391/

52. Badheka J, Doshi P, Vyas A, Kacha N, Parmar V. Comparison of upper lip bite test and ratio of height to thyromental distance with other airway assessment tests for predicting difficult endotracheal intubation. Indian J Crit Care Med. 2016 Jan 1;20(1):3–8.

53. Calder I, Picard J, Chapman M, O'Sullivan C, Crockard HA. Mouth opening: a new angle. Anesthesiology [Internet]. 2003 Oct 1 [cited 2022 Nov 6];99(4):799–801. Available from: https://pubmed.ncbi.nlm.nih.gov/14508309/

54. Faramarzi E, Soleimanpour H, Khan ZH, Mahmoodpoor A, Sanaie S. Upper lip bite test for prediction of difficult airway: A systematic review. Pakistan J Med Sci [Internet]. 2018 Jul 1 [cited 2022 Nov 6];34(4):1019–23. Available from: https://pubmed.ncbi.nlm.nih.gov/30190773/

55. Verghese S, Madhu D. Comparison of Validity of Mallampati Test and Lower Jaw Protrusion Maneuver in Predicting Difficult Laryngoscopy and Intubation. 10(03):2023.

56. Dhanger S, Gupta SL, Vinayagam S, Bidkar PU, Elakkumanan LB, Badhe AS. Diagnostic accuracy of bedside tests for predicting difficult intubation in Indian population: An observational study. Anesth Essays Res [Internet]. 2016 [cited 2022 Nov 7];10(1):54. Available from: /pmc/articles/PMC4767095/

57. Yadav U, Singh R, Chaudhari S, Srivastava S. Comparative Study of Preoperative Airway Assessment by Conventional Clinical Predictors and Ultrasound-Assisted Predictors. Anesth essays Res [Internet]. 2020 [cited 2022 Dec 6];14(2):213. Available from: https://pubmed.ncbi.nlm.nih.gov/33487818/

58. Liaskou C, Vouzounerakis E, Trikoupi A, Staikou C. Evaluation of bedside tests and proposal of a model for predicting difficult laryngoscopy: an observational prospective study. Brazilian J Anesthesiol [Internet]. 2020 Mar [cited 2022 Dec 6];70(2):125. Available from: /pmc/articles/PMC9678250/

59. Belda I, Ayuso MA, Sala-Blanch X, Luis M, Bergé R. Test predictivo de intubación difícil en microcirugía de laringe. Estudio de validación. Rev Esp Anestesiol Reanim [Internet]. 2017 Feb [cited 2019 Apr 27];64(2):71–8. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0034935616300767

60. Moustafa MA, El-Metainy S, Mahar K, Mahmoud Abdel-magied E. Defining difficult laryngoscopy findings by using multiple parameters: A machine learning approach. Egypt J Anaesth. 2017 Apr 1;33(2):153–8.

61. Wilson ME, Spiegelhalter D, Robertson JA, Lesser P. PREDICTING DIFFICULT INTUBATION. Br J Anaesth. 1988;61:211–6.

62. Arné J, Descoins P, Fusciardi J, Ferrier B, Boudigues D, Ariès J. Preoperative assessment for difficult intubation in general and ENT surgery: predictive value of a clinical multivariate risk index † [Internet]. Vol. 80, British Journal of Anaesthesia. 1998 [cited 2019 Apr 27]. Available from: https://bjanaesthesia.org/article/S0007-0912(17)40530-7/pdf

63. Nørskov AK, Rosenstock C V., Wetterslev J, Lundstrøm LH. Incidence of unanticipated difficult airway using an objective airway score versus a standard clinical airway assessment: The DIFFICAIR trial - trial protocol for a cluster randomized clinical trial. Trials. 2013 Oct 23;14(1):347.

64. Nørskov AK, Rosenstock C V., Wetterslev J, Astrup G, Afshari A, Lundstrøm LH. Diagnostic accuracy of anaesthesiologists' prediction of difficult airway management in daily clinical practice: a cohort study of 188064 patients registered in the Danish Anaesthesia Database. Anaesthesia [Internet]. 2015 Mar 1 [cited 2020 Apr 10];70(3):272–81. Available from: http://doi.wiley.com/10.1111/anae.12955

65. Ni H, Guan C, He G, Bao Y, Shi D, Zhu Y. Ultrasound measurement of laryngeal structures in the parasagittal plane for the prediction of difficult laryngoscopies in Chinese adults. BMC Anesthesiol [Internet]. 2020 Jun 2 [cited 2023 May 11];20(1). Available from: https://pubmed.ncbi.nlm.nih.gov/32487070/

66. Alessandri F, Antenucci G, Piervincenzi E, Buonopane C, Bellucci R, Andreoli C, et al. Ultrasound as a new tool in the assessment of airway difficulties: An observational study. Eur J Anaesthesiol. 2019 Jul 1;36(7):509–15.

67. Bhagavan S, Nelamangala K. Accuracy of Preoperative Ultrasonographic Airway Assessment in Predicting Difficult Laryngoscopies in Adult Patients. Cureus [Internet]. 2023 Mar 1 [cited 2023 May 11];15(3):e35652. Available from: http://www.ncbi.nlm.nih.gov/pubmed/37009359

68. Wu H, Wang H. Diagnostic Efficacy and Clinical Value of Ultrasonography in Difficult Airway Assessment: Based on a Prospective Cohort Study. Contrast Media Mol Imaging. 2022;2022.

69. Carsetti A, Sorbello M, Adrario E, Donati A, Falcetta S. Airway Ultrasound as Predictor of Difficult Direct Laryngoscopy: A Systematic Review and Meta-analysis. Anesth Analg [Internet]. 2022 Apr 1 [cited 2023 May 11];134(4):740–50. Available from: https://pubmed.ncbi.nlm.nih.gov/34914641/

70. Soltani Mohammadi S, Saliminia A, Nejatifard N, Azma R. Usefulness of Ultrasound View of Larynx in Pre-Anesthetic Airway Assessment: A Comparison With Cormack-Lehane Classification During Direct Laryngoscopy. An-

esthesiol Pain Med [Internet]. 2016 Aug 15 [cited 2018 Oct 3];6(6):e39566. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28975073

71. Hernández-Alfaro F, Guijarro-Martínez R, Mareque-Bueno J. Effect of Mono- and Bimaxillary Advancement on Pharyngeal Airway Volume: Cone-Beam Computed Tomography Evaluation. J Oral Maxillofac Surg [Internet]. 2011 Nov [cited 2018 Oct 3];69(11):e395–400. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21798648
72. Eckmann DM, Glassenberg R, Gavriely N. Acoustic reflectometry and endotracheal intubation. Anesth Analg [Internet]. 1996 Nov [cited 2018 Oct 3];83(5):1084–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8895291
73. El-Boghdadly K, Onwochei DN, Millhoff B, Ahmad I. The effect of virtual endoscopy on diagnostic accuracy and airway management strategies in patients with head and neck pathology: a prospective cohort study. Can J Anesth. 2017 Nov 1;64(11):1101–10.
74. Ormandy D, Kolb B, Jayaram S, Burley O, Kyzas P, Vallance H, et al. Difficult airways: a 3D printing study with virtual fibreoptic endoscopy. Br J Oral Maxillofac Surg. 2021 Feb 1;59(2):e65–71.
75. Ravindran B. Innovations in the Management of the Difficult Airway: A Narrative Review. Cureus [Internet]. 2023 Feb 17 [cited 2023 May 11];15(2). Available from: https://pubmed.ncbi.nlm.nih.gov/36945260/
76. Myatra SN, Dhawan I, D'Souza SA, Elakkumanan LB, Jain D, Natarajan P. Recent advances in airway management. Indian J Anaesth [Internet]. 2023 Jan 1 [cited 2023 May 11];67(1):48–55. Available from: http://www.ncbi.nlm.nih.gov/pubmed/36970487
77. Molins G, Valls A, Guijarro R, Hernández-Alfaro F. Mandibular hypoplasia and narrow airway in goldenhar syndrome: Anticipation of difficult intubation with cone-beam computed tomography. J Clin Anesth [Internet]. 2016 Nov [cited 2018 Oct 22];34:1–2. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27687336
78. Tavolara TE, Gurcan MN, Segal S, Niazi MKK. Identification of difficult to intubate patients from frontal face images using an ensemble of deep learning models. Comput Biol Med. 2021 Sep 1;136.
79. Connor CW, Segal S. Accurate Classification of Difficult Intubation by Computerized Facial Analysis. Anesth Analg [Internet]. 2011 Jan [cited 2019 Jun 25];112(1):84–93. Available from: https://insights.ovid.com/crossref?an=00000539-201101000-00015

80. Hayasaka T, Kawano K, Kurihara K, Suzuki H, Nakane M, Kawamae K. Creation of an artificial intelligence model for intubation difficulty classification by deep learning (convolutional neural network) using face images: an observational study. J intensive care [Internet]. 2021 May 6 [cited 2023 May 11];9(1):38. Available from: http://www.ncbi.nlm.nih.gov/pubmed/33952341

81. De Carvalho CC, Da Silva DM, De Carvalho AD, Nóbrega FJF, De Orange FA. Evaluation of the association between voice formants and difficult facemask ventilation. Eur J Anaesthesiol [Internet]. 2019 Dec 1 [cited 2022 Oct 11];36(12):972–3. Available from: https://pubmed.ncbi.nlm.nih.gov/31688299/

82. de Carvalho CC, da Silva DM, de Carvalho Junior AD, Santos Neto JM, Rio BR, Neto CN, et al. Pre-operative voice evaluation as a hypothetical predictor of difficult laryngoscopy. Anaesthesia. 2019;74(9):1147–52.

83. Cao S, Xia M, Zhou R, Wang J, Jin C-Y, Pei B, et al. Voice parameters for difficult mask ventilation evaluation: an observational study. Ann Transl Med [Internet]. 2021 Dec [cited 2022 Oct 11];9(23):1740–1740. Available from: /pmc/articles/PMC8743704/

84. Valls-Ontañón A, Ferreiro M, Moragues-Aguiló B, Molins-Ballabriga G, Julián-González S, Sauca-Balart A, et al. Impact of 3-dimensional anatomical changes secondary to orthognathic surgery on voice resonance and articulatory function: a prospective study. Br J Oral Maxillofac Surg [Internet]. 2023 Apr [cited 2023 May 26]; Available from: https://pubmed.ncbi.nlm.nih.gov/37210244/

85. Titze IR. Human Speech: A Restricted Use of the Mammalian Larynx. J Voice [Internet]. 2017 Mar 1 [cited 2023 May 12];31(2):135–41. Available from: https://pubmed.ncbi.nlm.nih.gov/27397113/

86. Van Puyvelde M, Neyt X, McGlone F, Pattyn N. Voice Stress Analysis: A New Framework for Voice and Effort in Human Performance. Front Psychol [Internet]. 2018 Nov 20 [cited 2023 May 13];9(NOV). Available from: https://pubmed.ncbi.nlm.nih.gov/30515113/

87. Leongómez JD, Pisanski K, Reby D, Sauter D, Lavan N, Perlman M, et al. Voice modulation: from origin and mechanism to social impact. Philos Trans R Soc B Biol Sci [Internet]. 2021 [cited 2023 May 12];376(1840). Available from: /pmc/articles/PMC8558768/

88. Cavallaro G, Di Nicola V, Quaranta N, Fiorella ML. Acoustic voice analysis in the covid-19 era. Acta Otorhinolaryngol Ital. 2021;41(1):1–5.

89. Paulo Teixeira J, Odete Fernandes P. Jitter, Shimmer and HNR Classification within Gender, Tones and Vowels in Healthy Voices. 2014 [cited 2019 Jan 19]; Available from: www.sciencedirect.com

90. Zhang Z. Effect of vocal fold stiffness on voice production in a three-dimensional body-cover phonation model. J Acoust Soc Am [Internet]. 2017 Oct [cited 2018 Jul 23];142(4):2311–21. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29092586

91. Casper J, Leonard R. Understanding voice problems: A physiological perspective for diagnosis and treatment. 2006 [cited 2023 May 12]; Available from: https://books.google.com/books?hl=es&lr=&id=1RwVcsw0KLAC&oi=fnd&pg=PR1&ots=eW6z-GH5KJ&sig=J1iQ1W2lgtcDvlojPAs_2vGs6wE

92. Yin J, Zhang Z. The influence of thyroarytenoid and cricothyroid muscle activation on vocal fold stiffness and eigenfrequencies. J Acoust Soc Am [Internet]. 2013 May [cited 2023 May 12];133(5):2972–83. Available from: https://pubmed.ncbi.nlm.nih.gov/23654401/

93. Fundamental frequency (F0) - Introduction to Speech Processing - Aalto University Wiki [Internet]. [cited 2023 May 13]. Available from: https://wiki.aalto.fi/pages/viewpage.action?pageId=149890776

94. Krasnodębska P, Szkiełkowska A, Czarkwiani-Woźniakowska L, Miaśkiewicz B, Sinkiewicz A, Skarżyński H. Laryngeal Electromyography in the Therapeutic Process of Patients with Vocal Fold Immobility or Dysmobility. Life (Basel, Switzerland) [Internet]. 2022 Mar 1 [cited 2023 May 13];12(3). Available from: https://pubmed.ncbi.nlm.nih.gov/35330142/

95. Nicastri M, Chiarella G, Gallo L V, Catalano M, Cassandro E. Multidimensional Voice Program (MDVP) and amplitude variation parameters in euphonic adult subjects. Normative study. Acta Otorhinolaryngol Ital [Internet]. 2004 Dec [cited 2018 Apr 8];24(6):337–41. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15952683

96. Boersma P, Van Heuven V. Speak and unSpeak with PRAAT RAAT.

97. Wei M, Du J, Wang X, Lu H, Wang W, Lin P. Voice disorders in severe obstructive sleep apnea patients and comparison of two acoustic analysis software programs: MDVP and Praat. Sleep Breath [Internet]. 2021 Mar 1 [cited 2022 Sep 18];25(1):433–9. Available from: https://pubmed.ncbi.nlm.nih.gov/32583274/

98. Amir O, Wolf M, Amir N. A clinical comparison between two acoustic analysis softwares: MDVP and Praat. Biomed Signal Process Control. 2009 Jul 1;4(3):202–5.

99. Keung LC, Richardson K, Sharp Matheron D, Martel-Sauvageau V. A Comparison of Healthy and Disordered Voices Using Multi-Dimensional Voice Program, Praat, and TF32. J Voice. 2022;

100. Maryn Y, Corthals P, De Bodt M, Van Cauwenberge P, Deliyski D. Perturbation measures of voice: a comparative study between Multi-Dimensional Voice Program and Praat. Folia Phoniatr Logop [Internet]. 2009 Sep [cited 2022 Sep 18];61(4):217–26. Available from: https://pubmed.ncbi.nlm.nih.gov/19590221/

101. Lovato A, De Colle W, Giacomelli L, Piacente A, Righetto L, Marioni G, et al. Multi-Dimensional Voice Program (MDVP) vs Praat for Assessing Euphonic Subjects: A Preliminary Study on the Gender-discriminating Power of Acoustic Analysis Software. J Voice. 2016 Nov 1;30(6):765.e1-765.e5.

102. Sachdeva K, Mittal N, Sachdeva N. Role of Video Laryngostroboscopy in Benign Disease of Larynx. Indian J Otolaryngol Head Neck Surg [Internet]. 2020 Jun 1 [cited 2023 May 13];72(2):267–73. Available from: https://pubmed.ncbi.nlm.nih.gov/32551287/

103. Banjara H, Mungutwar V, Singh D, Gupta A, Singh S. Demographic and videostroboscopic assessment of vocal pathologies. Indian J Otolaryngol Head Neck Surg [Internet]. 2012 Jun [cited 2023 May 13];64(2):150–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23730576

104. Titze IR, Baken RJ, Bozeman KW, Granqvist S, Henrich N, Herbst CT, et al. Toward a consensus on symbolic notation of harmonics, resonances, and formants in vocalization. J Acoust Soc Am [Internet]. 2015 May [cited 2023 May 14];137(5):3005. Available from: /pmc/articles/PMC5392060/

105. López S, Riera P, Assaneo MF, Eguía M, Sigman M, Trevisan MA. Vocal caricatures reveal signatures of speaker identity. Sci Reports 2013 31 [Internet]. 2013 Dec 3 [cited 2023 May 14];3(1):1–7. Available from: https://www.nature.com/articles/srep03407

106. Docio-Fernandez L, García Mateo C. Speech Production. Encycl Biometrics [Internet]. 2015 [cited 2023 May 26];1493–8. Available from: https://link.springer.com/referenceworkentry/10.1007/978-1-4899-7488-4_199

107. Propietario. Acústica del tracto vocal.

108. Blomgren M, Chen Y, Ng ML, Gilbert HR. Acoustic, aerodynamic, physiologic, and perceptual properties of modal and vocal fry registers. J Acoust Soc Am [Internet]. 1998 May [cited 2023 May 14];103(5 Pt 1):2649–58. Available from: https://pubmed.ncbi.nlm.nih.gov/9604359/

109. Brockmann M, Drinnan MJ, Storck C, Carding PN. Reliable Jitter and Shimmer Measurements in Voice Clinics: The Relevance of Vowel, Gender, Vocal Intensity, and Fundamental Frequency Effects in a Typical Clinical Task. J Voice [Internet]. 2011 Jan [cited 2018 Apr 9];25(1):44–53. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20381308

110. Sound Intensity & Loudness - Teachers (U.S. National Park Service) [Internet]. [cited 2023 May 14]. Available from: https://www.nps.gov/teachers/classrooms/sound-intensity-and-loudness.htm

111. Lopes LW, Cavalcante DP, da Costa PO. Severity of voice disorders: integration of perceptual and acoustic data in dysphonic patients. CoDAS [Internet]. 2014 [cited 2023 May 14];26(5):382–8. Available from: http://www.scielo.br/j/codas/a/kGTm3ryX49stcPVt9YvC5vS/?lang=en

112. Teixeira JP, Oliveira C, Lopes C. Vocal Acoustic Analysis – Jitter, Shimmer and HNR Parameters. Procedia Technol [Internet]. 2013 [cited 2023 May 13];9:1112–22. Available from: https://www.researchgate.net/publication/273822486_Vocal_Acoustic_Analysis_-_Jitter_Shimmer_and_HNR_Parameters

113. Vista de Análisis acústico de la voz: medidas temporales, espectrales y cepstrales en la voz normal con el Praat en una muestra de hablantes de español [Internet]. [cited 2023 May 14]. Available from: https://revistas.ucm.es/index.php/RLOG/article/view/58191/52369

114. Kasuya H, Ogawa S. Normalized noise energy as an acoustic measure to evaluate pathologic voice Related Content. J Acoust Soc Am [Internet]. 1986 [cited 2023 May 14];80:1329–34. Available from: https://doi.org/10.1121/1.394384

115. Anniko M, Bernal-Sprekelsen M, Bonkowsky V, Bradley P, Iurato S. Otorhinolaryngology, Head and Neck Surgery European Manual of Medicine.

116. Becker W, Naumann HH (Hans H, Pfaltz CR (Carl R. Otorrinolaringología : manual ilustrado. 1986;

117. Karlsson T. Radiotherapy and Voice Rehabilitation in Laryngeal Cancer Effects on Health-Related Quality of Life and Voice Function. 2015;

118. Campos G, Farías P, Friedrich G, Hess M, na Jackson-Menaldi Marc Remacle C, Cobeta Núñez S Fernández Jarmed IF. Ignacio Cobeta Faustino Núñez Secundino Fernández Ponencia ofi cial Sociedad Española de Otorrinolaringología y Patología Cérvico-Facial 2013 Patología de la voz Patología de la voz Con la colaboración de: Patología de la voz.

119. Lina D, López R, Mena N, Jesús V, Turpín I. CAPÍTULO 93. ANATOMÍA Y EMBRIOLOGÍA DE LA LARINGE EMBRIOLOGÍA DE LA LARINGE PRINCIPIOS GENERALES DEL DESARROLLO.

120. 8.1 Extrinsic muscles of larynx (muscles of face) | Download Scientific Diagram [Internet]. [cited 2023 May 15]. Available from: https://www.researchgate.net/figure/1-Extrinsic-muscles-of-larynx-muscles-of-face_fig6_301549536

121. regional-anesthesia-innervation-of-the-larynx - NYSORA | NYSORA [Internet]. [cited 2023 May 15]. Available from: https://www.nysora.com/techniques/head-and-neck-blocks/airway/regional-topical-anesthesia-awake-endotracheal-intubation/attachment/regional-anesthesia-innervation-of-the-larynx/

122. Drevet G, Conti M, Deslauriers J. Surgical anatomy of the tracheobronchial tree. J Thorac Dis [Internet]. 2016 Mar 1 [cited 2023 May 15];8(Suppl 2):S121–9. Available from: https://jtd.amegroups.com/article/view/6635/html

123. Gray SD, Titze IR, Chan R, Hammond TH. Vocal fold proteoglycans and their influence on biomechanics. Laryngoscope [Internet]. 1999 Jun [cited 2023 May 15];109(6):845–54. Available from: https://pubmed.ncbi.nlm.nih.gov/10369269/

124. Hamilton NJI. The life‐cycle and restoration of the human vocal fold. Laryngoscope Investig Otolaryngol [Internet]. 2023 Feb 1 [cited 2023 May 15];8(1):168. Available from: /pmc/articles/PMC9948577/

125. Titze IR. How can vocal folds oscillate with a limited mucosal wave? Jasa Express Lett [Internet]. 2022 Oct [cited 2023 May 15];2(10):105201. Available from: /pmc/articles/PMC9533394/

126. Cannito MP, Kahane JC, Chorna L. Vocal aging and adductor spasmodic dysphonia: Response to botulinum toxin injection. Clin Interv Aging [Internet]. 2008 [cited 2023 May 15];3(1):131. Available from: /pmc/articles/PMC2544359/

127. Zhang Z. Mechanics of human voice production and control. J Acoust Soc Am [Internet]. 2016 Oct [cited 2023 May 15];140(4):2614. Available from: /pmc/articles/PMC5412481/

128. Titze I, Riede T, Mau T. Predicting Achievable Fundamental Frequency Ranges in Vocalization Across Species. PLoS Comput Biol. 2016 Jun 1;12(6).

129. VAN DEN BERG J. Myoelastic-Aerodynamic Theory of Voice Production. J Speech Hear Res [Internet]. 1958 [cited 2023 May 15];1(3):227–44.

Available from: https://pubs.asha.org/doi/10.1044/jshr.0103.227?url_ver=Z39.88-2003&rfr_id=ori%3Arid%3Acrossref.org&rfr_dat=cr_pub++-0pubmed

130. Fluid Mechanics: Landau and Lifshitz: Course of Theoretical Physics, Volume 6 - L D Landau, E. M. Lifshitz - Google Libros [Internet]. [cited 2023 May 15]. Available from: https://books.google.es/books?hl=es&lr=&id=eOBbAwAAQBAJ&oi=fnd&pg=PP1&ots=2TE48fCMtR&sig=y2buQ5jT993cX_agG6oyQHA7PF0&redir_esc=y#v=onepage&q&f=false
131. Zhang LT, Yang J. Evaluation of aerodynamic characteristics of a coupled fluid-structure system using generalized Bernoulli's principle: An application to vocal folds vibration. J coupled Syst multiscale Dyn [Internet]. 2016 Mar 14 [cited 2023 May 15];4(4):241. Available from: /pmc/articles/PMC5841236/
132. Artificial Intelligence: A Guide for Thinking Humans - Melanie Mitchell - Google Libros [Internet]. [cited 2023 May 16]. Available from: https://books.google.es/books?hl=es&lr=&id=VfSIDwAAQBAJ&oi=fnd&pg=PT6&ots=-siqquJXz5r&sig=vYhSwyThyHkcXE44dwMnMwCsBmw&redir_esc=y#v=onepage&q&f=false
133. Kurzweil R, hernández CG. La Singularidad está cerca : Cuando los humanos transcendamos la biología.
134. Rifaioglu AS, Atas H, Martin MJ, Cetin-Atalay R, Atalay V, Doğan T. Recent applications of deep learning and machine intelligence on in silico drug discovery: methods, tools and databases. Brief Bioinform [Internet]. 2019 Sep 27 [cited 2023 May 16];20(5):1878–912. Available from: https://pubmed.ncbi.nlm.nih.gov/30084866/
135. Alaskar H, Saba T. Machine Learning and Deep Learning: A Comparative Review. 2021;143–50.
136. Artificial Intelligence A Modern Approach Third Edition.
137. Liu J, Li J, Li W, Wu J. Rethinking big data: A review on the data quality and usage issues. ISPRS J Photogramm Remote Sens. 2016 May 1;115:134–42.
138. Nguyen G, Dlugolinsky S, Bobák M, Tran V, López García Á, Heredia I, et al. Machine Learning and Deep Learning frameworks and libraries for large-scale data mining: a survey. Artif Intell Rev [Internet]. 2019 Jun 1 [cited 2023 May 16];52(1):77–124. Available from: https://link.springer.com/article/10.1007/s10462-018-09679-z
139. Janiesch C, Zschech P, Heinrich K. Machine learning and deep learning. Electron Mark [Internet]. 2021 Sep 1 [cited 2023 May 16];31(3):685–95.

Available from: https://link.springer.com/article/10.1007/s12525-021-00475-2

140. Liu Q, Wu Y. Supervised Learning. Encycl Sci Learn. 2012;3243–5.
141. Li SE. Deep Reinforcement Learning. Reinf Learn Seq Decis Optim Control [Internet]. 2023 [cited 2023 May 16];365–402. Available from: https://link.springer.com/chapter/10.1007/978-981-19-7784-8_10
142. Grossi E, Buscema M. Introduction to artificial neural networks. Eur J Gastroenterol Hepatol [Internet]. 2007 Dec [cited 2023 May 16];19(12):1046–54. Available from: https://www.researchgate.net/publication/5847739_Introduction_to_artificial_neural_networks
143. Dridi S. Supervised Learning - A Systematic Literature Review. [cited 2023 May 16]; Available from: https://osf.io/tysr4/
144. Biecek P. MODEL DEVELOPMENT PROCESS A PREPRINT. 2019 [cited 2023 May 17]; Available from: https://rviews.rstudio.com/2019/06/19/
145. Breiman L. Statistical modeling: The two cultures. Stat Sci. 2001;16(3):199–215.
146. How One-Hot Encoding Improves Machine Learning Performance | dl-question-bank – Weights & Biases [Internet]. [cited 2023 May 17]. Available from: https://wandb.ai/ayush-thakur/dl-question-bank/reports/How-One-Hot-Encoding-Improves-Machine-Learning-Performance--VmlldzoxOTkzMDk
147. Duran-Lopez L, Dominguez-Morales JP, Conde-Martin AF, Vicente-Diaz S, Linares-Barranco A. PROMETEO: A CNN-Based Computer-Aided Diagnosis System for WSI Prostate Cancer Detection. IEEE Access. 2020;8:128613–28.
148. Xu Y, Goodacre R. On Splitting Training and Validation Set: A Comparative Study of Cross-Validation, Bootstrap and Systematic Sampling for Estimating the Generalization Performance of Supervised Learning. J Anal Test [Internet]. 2018 Jul 1 [cited 2023 May 17];2(3):249–62. Available from: https://link.springer.com/article/10.1007/s41664-018-0068-2
149. Canbek G, Temizel TT, Sagiroglu S, Baykal N. Binary classification performance measures/metrics: A comprehensive visualized roadmap to gain new insights. 2nd Int Conf Comput Sci Eng UBMK 2017 [Internet]. 2017 Oct 31 [cited 2023 May 17];821–6. Available from: https://www.researchgate.net/publication/320829355_Binary_Classification_Performance_MeasuresMetrics_A_comprehensive_visualized_roadmap_to_gain_new_insights

150. Koyejo O, Natarajan N, Ravikumar P, Dhillon IS. Consistent Binary Classification with Generalized Performance Metrics.

151. Tharwat A. Classification assessment methods. [cited 2023 May 17]; Available from: https://www.emerald.com/insight/2210-8327.htm

152. Sokolova M, Japkowicz N, Szpakowicz S. Beyond accuracy, F-score and ROC: A family of discriminant measures for performance evaluation. AAAI Work - Tech Rep [Internet]. 2006 [cited 2023 May 17];WS-06-06:24–9. Available from: https://link.springer.com/chapter/10.1007/11941439_114

153. Churcher A, Ullah R, Ahmad J, Ur Rehman S, Masood F, Gogate M, et al. An Experimental Analysis of Attack Classification Using Machine Learning in IoT Networks. Sensors (Basel) [Internet]. 2021 Jan 2 [cited 2023 May 17];21(2):1–32. Available from: /pmc/articles/PMC7827441/

154. López V, Fernández A, García S, Palade V, Herrera F. An insight into classification with imbalanced data: Empirical results and current trends on using data intrinsic characteristics. Inf Sci (Ny). 2013 Nov 20;250:113–41.

155. Bradley AP. The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern Recognit. 1997;30(7):1145–59.

156. Tiwari A. Supervised learning: From theory to applications. Artif Intell Mach Learn EDGE Comput. 2022 Jan 1;23–32.

157. Demir F. Deep autoencoder-based automated brain tumor detection from MRI data. Artif Intell Brain-Computer Interface. 2022 Jan 1;317–51.

158. Akpan UI, Starkey A. Review of classification algorithms with changing inter-class distances. Mach Learn with Appl. 2021 Jun 15;4:100031.

159. Mulak P, Talhar N. Analysis of Distance Measures Using K-Nearest Neighbor Algorithm on KDD Dataset. Int J Sci Res [Internet]. 2013 [cited 2023 May 18];4:2319–7064. Available from: www.ijsr.net

160. Müller M. Generalized Linear Models. Handb Comput Stat [Internet]. 2012 [cited 2023 May 18];681–709. Available from: https://www.researchgate.net/publication/228396785_Generalized_Linear_Models

161. Connor CW. Artificial Intelligence and Machine Learning in Anesthesiology. Anesthesiology [Internet]. 2019 Dec 1 [cited 2022 Dec 6];131(6):1346–59. Available from: https://pubmed.ncbi.nlm.nih.gov/30973516/

162. Lee HC, Ryu HG, Chung EJ, Jung CW. Prediction of Bispectral Index during Target-controlled Infusion of Propofol and Remifentanil: A Deep Learning Approach. Anesthesiology [Internet]. 2018 Mar 1 [cited 2023 May 18];128(3):492–501. Available from: https://pubmed.ncbi.nlm.nih.gov/28953500/

163. Gambus P, Shafer SL. Artificial Intelligence for Everyone. Anesthesiology [Internet]. 2018 Mar 1 [cited 2023 May 18];128(3):431–3. Available from: https://pubmed.ncbi.nlm.nih.gov/29166324/

164. Kim JH, Kim H, Jang JS, Hwang SM, Lim SY, Lee JJ, et al. Development and validation of a difficult laryngoscopy prediction model using machine learning of neck circumference and thyromental height. BMC Anesthesiol. 2021 Dec 1;21(1).

165. Yamanaka S, Goto T, Morikawa K, Watase H, Okamoto H, Hagiwara Y, et al. Machine Learning Approaches for Predicting Difficult Airway and First-Pass Success in the Emergency Department: Multicenter Prospective Observational Study. Interact J Med Res [Internet]. 2022 Jan 25 [cited 2022 Dec 6];11(1):e28366. Available from: https://pubmed.ncbi.nlm.nih.gov/35076398/

166. Connor CW, Segal S. Accurate classification of difficult intubation by computerized facial analysis. Anesth Analg. 2011;112(1):84–93.

167. Tavolara TE, Gurcan MN, Segal S, Niazi MKK. Identification of difficult to intubate patients from frontal face images using an ensemble of deep learning models. Comput Biol Med [Internet]. 2021 Sep 1 [cited 2022 Dec 6];136:104737. Available from: /pmc/articles/PMC8440461/

168. Hatib F, Jian Z, Buddi S, Lee C, Settels J, Sibert K, et al. Machine-learning Algorithm to Predict Hypotension Based on High-fidelity Arterial Pressure Waveform Analysis. Anesthesiology. 2018;129(4):663–74.

169. JCI-Accredited Organizations | Joint Commission International [Internet]. [cited 2023 May 22]. Available from: https://www.jointcommissioninternational.org/who-we-are/accredited-organizations/#sort=%40aoname ascending

170. KNIME [Internet]. [cited 2023 May 22]. Available from: https://www.knime.com/

171. Xu Y, Goodacre R. On Splitting Training and Validation Set: A Comparative Study of Cross-Validation, Bootstrap and Systematic Sampling for Estimating the Generalization Performance of Supervised Learning. J Anal Test [Internet]. 2018 Jul 1 [cited 2023 May 17];2(3):249–62. Available from: https://link.springer.com/article/10.1007/s41664-018-0068-2

172. Shadle CH, Nam H, Whalen DH. Comparing measurement errors for formants in synthetic and natural vowels. J Acoust Soc Am [Internet]. 2016 Feb [cited 2023 May 14];139(2):713. Available from: /pmc/articles/PMC4752539/

173. Chen W-R, Whalen ; D H, Shadle CH, Whalen) D H. F0-induced formant measurement errors result in biased variabilities. J Acoust Soc Am [Internet]. 2019 May 1 [cited 2023 Jun 11];145(5):EL360–6. Available from: /asa/jasa/article/145/5/EL360/671477/F0-induced-formant-measurement-errors-result-in

174. Zhou CM, Wang Y, Xue Q, Yang JJ, Zhu Y. Predicting difficult airway intubation in thyroid surgery using multiple machine learning and deep learning algorithms. Front public Heal [Internet]. 2022 Aug 10 [cited 2022 Dec 6];10. Available from: https://pubmed.ncbi.nlm.nih.gov/36033770/

175. Sagün A, Özdemir L, Melikoğulları SB. The assessment of risk factors associated with difficult intubation as endocrine, musculoskeletal diseases and intraoral cavity mass: A nested case control study. Turkish J Trauma Emerg Surg. 2022 Sep 1;28(9):1270–6.

176. McDonald A. Data Quality Considerations for Petrophysical Machine-Learning Models. Petrophysics. 2021 Dec 1;62(6):585–613.

177. Decuyper M, Stockhoff M, Vandenberghe S, - al, Ying X. An Overview of Overfitting and its Solutions. J Phys Conf Ser [Internet]. 2019 Feb 1 [cited 2023 Jun 15];1168(2):022022. Available from: https://iopscience.iop.org/article/10.1088/1742-6596/1168/2/022022

178. Nahm FS. Receiver operating characteristic curve: overview and practical use for clinicians. Korean J Anesthesiol [Internet]. 2022 Feb 1 [cited 2023 Jun 16];75(1):25–36. Available from: https://pubmed.ncbi.nlm.nih.gov/35124947/

9. INDEX OF TABLES

9. INDEX OF TABLES

Table 1. Number of patients undergoing different surgical procedures in two studies, initial and validation studies. J. Arné et al 1998. British Journal of Anaesthesia (62)..... 55

Table 2. Main variables..... 116

Table 3. Secondary variables..... 118

Table 4. Descriptive results of all the participants included in the Mallampati study. 144

Table 5. Main features of the patients included in the Mallampati study 145

Table 6. Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in both training and testing groups, Accuracy values for the training group, Accuracy values for the testing group..... 147

Table 7. Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in both training and testing groups, AUC and Pr AUC values for the training group, AUC and Pr AUC values for the testing group 150

Table 8. Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in both training and testing groups, Precision values for the training group, Recall values for the testing group..... 153

Table 9. Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in both training and testing groups, F1 and Log loss values for the training group and the testing group respectively. 156

Table 10. Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in testing group. TN: True negatives, FN: False negatives, FP: False positives, TP: True positives. 159

Table 11. Descriptive statistics of all the participants included in the Cormack study. 162

Table 12. Main features of the participants included in the Cormack study. 164

Table 13. Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in both training and testing groups, for accuracy values for the training group and the testing group respectively. 166

Table 14. Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in both training and testing groups, for precision and recall values for the training group and the testing group respectively. (Simplified)..... 168

Table 15. Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in both training and testing groups, for AUC and Pr AUC values for the training group and the testing group respectively (Simplified)..... 171

Table 16. Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in both training and testing groups, for F1 and Log loss values for the training group and the testing group respectively (Simplified)..... 174

Table 17. Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in both training and testing groups, for true negatives, false negatives, false positives, true positive values for the training group and the testing group respectively (Simplified)..... 177

Table 18. Metrics results of Mallampati vs Cormack..... 180

Table 19. Best combinations of parameters and binary classifications models which obtained the best results for the Mallampati and the Cormack study for the training and testing datasets. ... 182

10. INDEX OF FIGURES

10. INDEX OF FIGURES

Figure 1. Best views obtained at direct laryngoscopy, assuming the correct technique. Original from R.S Cormack and J Lehane, Difficult tracheal intubation in obstetrics. Anaesthesia, 1984, Volume 39, pages 1105-1 I1 (1)..... 30

Figure 2. Difficult airway algorithm: Adult patients (2)..... 34

Figure 3. Flow diagram: difficult tracheal intubation encountered in the unconscious patient. SGD = supraglottic devices..... 36

Figure 4. Different types of clinical situations based on the compliance with the safety criteria. Recomendaciones para la evaluación y manejo de la vía aérea difícil prevista y no prevista basadas en la adaptación de guías de práctica clínica y consenso de expertos. Revista Española de Anestesiología y reanimación 2020 38

Figure 5. Recommendations for the management of the airway. Revista Española de Anestesiología y reanimación SCARTD..... 39

Figure 6. Final outcome: narrative outcome and NPSA classification. Fourth National Audit project of the Royal College of Anesthetist and the difficult airway society (3). 42

Figure 7. List of recorded primary events problems. Fourth National Audit project of the Royal College of Anesthetist and the difficult Airway Society. 42

Figure 8. Predictors of Difficult tracheal Intubation and Judgment Failures in Airway Management(4)with a higher proportion of death in 2000 to 2012 claims (73%; n = 74 of 102 vs. 42%; n = 39 of 93 in 1993 to 1999 claims; P < 0.001 adjusted for multiple testing..... 45

Figure 9. Peer expert and analyst-identified judgement failures for CMPA closed cases, 2007–2016 (n = 46 cases) (32) 46

Figure 10. Human Factors Investigation Tool model of incident causation and direction of analysis (40) 47

Figure 11. The modified Mallampati classification (45)..... 50

Figure 12. A. thyromental distance, B. Sternomental distance (52)..... 52

Figure 13. Mouth opening or interincisal gap (Source: Preanestesia2.anestalia.com)..... 53

Figure 14. Jaw protrusion grade A, B, C respectively (55)..... 54

Figure 15. Arné test Index. J. Arné et al 1998. British Journal of Anaesthesia (62)..... 56

Figure 16. Wilson risk score. M.E Wilson et al, B.J. Anesthesia, 1988. 57

Figure 17. Segment of a speech signal with the period length L, and fundamental frequency F0= 1/L (93)..... 63

Figure 18. Voice formation diagram. Pathway from its origin to its exit through the resonating elements (106)..... 65

Figure 19. Jitter absolute formula (89)..... 66

Figure 20. Jitter relative formula (89) 67

Figure 21. Jitter (ppq5) Formula (89) 67

Figure 22. Shimmer (dB) formula. Ai is the extracted peak-to-peak amplitude data and N is the number of extracted fundamental frequency periods (89) 68

Figure 23. Shimmer relative formula. Aiare the extracted peak-to-peak amplitude data andNis the number of extracted fundamental frequency periods (89) 68

Figure 24. Jitter and Shimmer perturbation measures in speech signal (112) 69

Figure 25. Regions of the laryngeal cavity (117)..... 71

Figure 26. Anterolateral view of the laryngeal cartilages and ligaments(118)..... 73

Figure 27. Top view of the muscles in the left, location and in the right action. Lateral cricoarytenoid muscles (A and B), arytenoid muscles (C and D) and thyroarytenoid muscles (F and H). ac, cricoid arch; ca, arytenoid cartilage; cal, lateral cricoarytenoid muscle; e, epiglottis; gc, cartilaginous glottis; gl, ligamentous glottis; lc, cricoid lamina; lt, thyroid lamina; lvo, vocal ligament; po, oblique portion; pt, transverse portion; ta, thyroarytenoid muscle (118)..... 75

Figure 28. Extrinsic muscles of the larynx (120) 76

Figure 29. Innervation of the larynx (121) 77

Figure 30. Irrigation of the larynx (122) 78

Figure 31. Vocal folds structure (128) 79

Figure 32. machine learning concepts and classes (139) 86

Figure 33. Hierarchical structure of an ANN (141) 88

Figure 34. The life-cycle of a predictive model (145) 89

Figure 35. Example of a different encoding of a categorical feature (146) 90

Figure 36. Example of 3-folds cross- validation(147) 91

Figure 37. Accuracy formula. In where: TP True positive), TN (true negative), FP (False positive), FN (False negative) (151) 93

Figure 38. Sensitivity or True positive rate formula (151) 93

Figure 39. Specificity or True negative rate formula (151) 93

Figure 40. F1 score formula (152) 93

Figure 41. ROC curve (151) 94

Figure 42. Confusion matrix (157) 95

Figure 43. Widely used supervised classification algorithms (143) 97

Figure 44. Study design and follow-up of participants. * voice. anestestalia.com database, ** KNIMEÒ: Open-source platform built to productionize data science 112

Figure 45. Centro Médico Teknon, Quironsalud group, Barcelona, Spain. 113

Figure 46. Hospital Universitari Dexeus; Quironsalud Group Barcelona, Spain. 114

Figure 47. Flowchart of patient selection 115

Figure 48. Preanesthetic visit platform. Source: Preanestesia2. anestestalia.com 123

Figure 49. Home menu. Source: website: voice.anestalia.com 123

Figure 50. New Bee H360 headphones model. Source: anewbee.com 124

Figure 51. KNIMEÒ general workflow example Source: KNIMEÒ platform, own production 124

Figure 52. Preanesthetic evaluation form (Example). Source: Preanestesia2.anestalia.com 125

Figure 53. Airway physical assessment performed during the preanesthetic visit. Source: Preanestesia2.anestalia.com 125

Figure 54. Informed consent of the study. It contains an alphanumeric code which was assigned to each patient included in the study. Source: own production 126

Figure 55. Voice recording flowchart. First picture: case code generation for the inclusion of the patient in the study. Second picture: website voice.anestalia.com (Own production) designed for the voice recording. Third picture: different voice recording positions and vocals..... 127

Figure 56. Automatic filter which detects the poorly recorded audios and identifies them as 'Incorrect' and indicates the need to re-record them until they meet the necessary requirements. Identified as 'reviewed 'Source: Preanestesia2.anestalia.com 128

Figure 57. Data collection form. Own production 129

Figure 58. Data collection platform. Voice.anestalia.com..... 130

Figures 59 and 60. KNIME® node that allows normalizing each column from 0 to 1 for both the Harmonics and the descriptive data. (Extracted from the final workflow, own production) 132

Figure 61. Partitioning of the dataset. 70% of the participants formed the training and validation dataset and 30% of them were formed the test dataset for both studies(3). 133

Figure 62. PCA node configuration in KNIME workflow. 134

Figure 63. Decision three hyperparameters adjustment: set of the depth and the number of threes. Source: KNIMEÒ workflow own production 135

Figure 64. Vowel selection in KNIME workflow. 135

Figure 65. Position for each vowel selection in KNIME workflow..... 136

Figure 66. Confusion matrix Mallampati vs Cormack..... 182

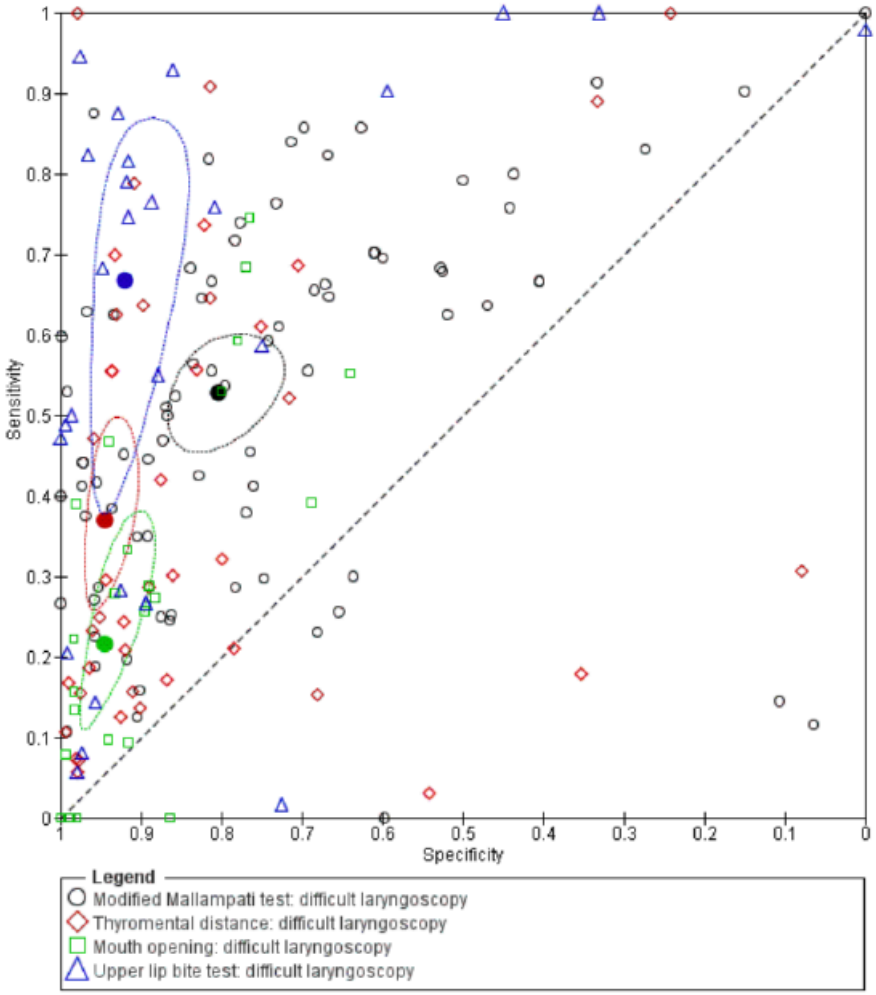
Figure 67. ROC AUC curve Mallampati vs Cormack..... 183

11. ANNEXES

11. ANNEXES

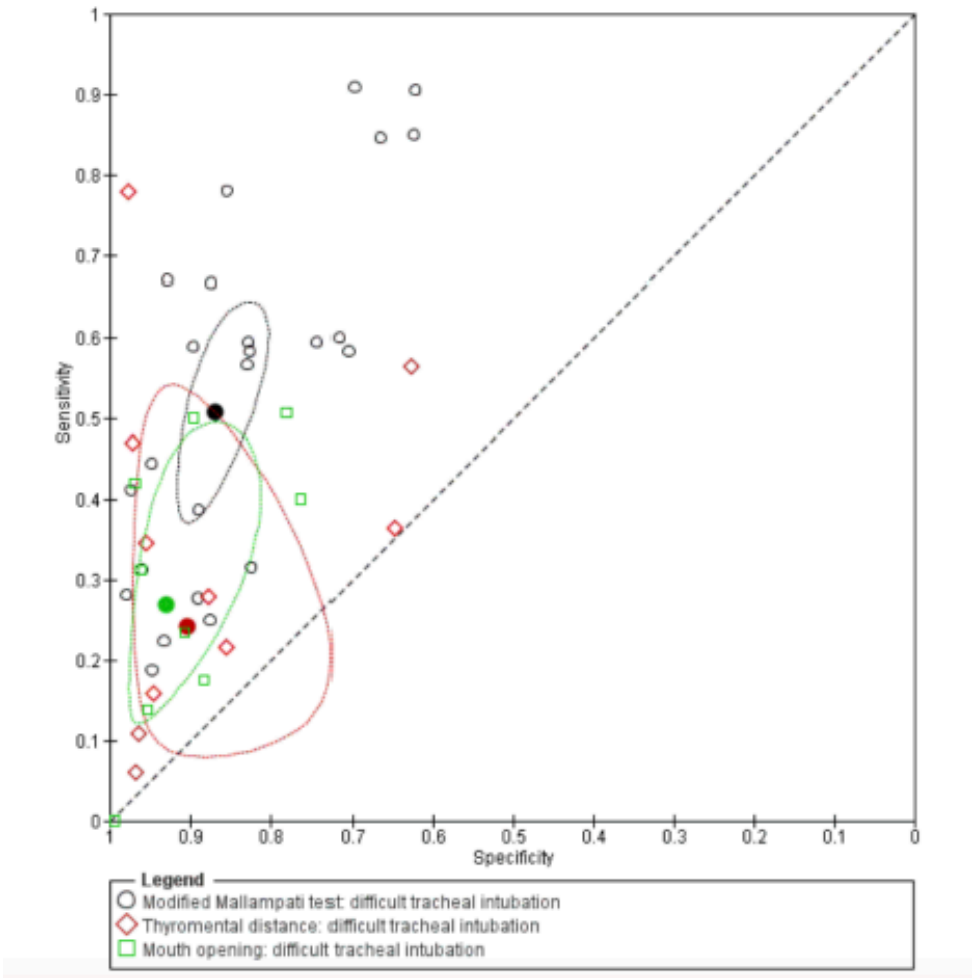
11.1. ANNEX 1: Summary receiver operating characteristic ROC plot for different traditional predictive test for a difficult laryngoscopy (47)

Figure 10. Summary receiver operating characteristic (ROC) plot of modified Mallampati test, thyromental distance, mouth opening, and upper lip bite test for difficult laryngoscopy. For each index test, the summary point with the 95% confidence region is displayed.

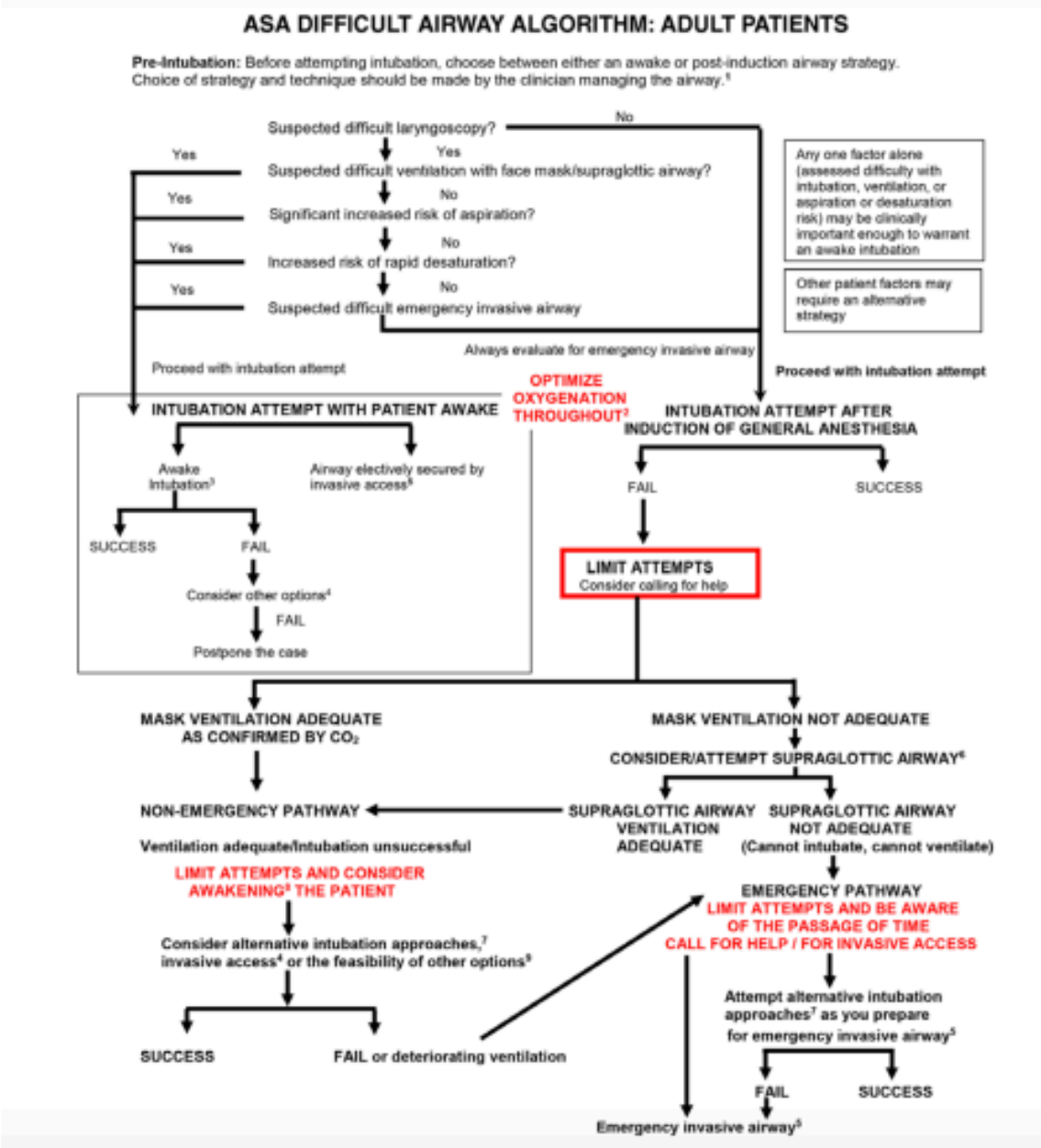


11.2. ANNEX 2: Summary receiver operating characteristic (ROC) plot for the traditional predictive test for a difficult tracheal intubation. (47)

Figure 14. Summary receiver operating characteristic (ROC) plot of modified Mallampati test, thyromental distance, and mouth opening test for difficult tracheal intubation. For each index test the summary point with the 95% confidence region is displayed.



11.3 ANNEX 3: American Society of Anesthesiologists difficult airway management algorithm in adults (18)



11.4. ANNEX 4: Research Ethics committee report

quirónsalud
La salud persona a persona

Título: **Análisis de la voz como método para predecir una vía aérea difícil de manera preoperatoria.**
Código Protocolo: **VOICE2019**
Código Interno: **2019/78-ANE-DEX**
Investigador Principal: **Claudia Rodiera**
Servicio: **Anestesia y Reanimación**
Centro de Trabajo: **Hospital Universitari Dexeus**

Se hace constar que:

En la reunión celebrada el día 26 de septiembre de 2019 (ACTA 62), se decidió emitir el informe al protocolo de referencia.

En dicha reunión se cumplieron los requisitos establecidos en la legislación vigente para que la decisión del citado CEIm sea válida.

El CEIm del Grupo Hospitalario Quirónsalud en Barcelona, tanto en su composición como en los PNT cumple con las normas BPC (CPMP/ICH/135/95). Los miembros participantes en la evaluación del estudio son:

Vicepresidente	Dr. Joan Costa Pagés. Farmacología Clínica
Secretaria	Sra. Montse Granados Plaza. Enfermera
Vocales	Dr. Fernando Cereto Castro. Medicna Interna
	Dr. Joan Albert Arnaiz Gargallo. Farmacología Clínica
	Sra. Mercedes Gozalbo Mestres. Unidad Investigación y Docencia
	Dr. Jordi Peláez de Loño. Farmacéutico Atención Primaria
	Sr. Ignacio Rodríguez García. Epidemiología y Estadística
	Dr. Josep Rodiera Olivé. Anestesiología y Reanimación

Para que conste donde proceda,

En Barcelona, a 30 de septiembre de 2019.

GRANADOS PLAZA MONTSERRAT AMELIA - 45474810S

Montserrat Granados Plaza
Secretaria CEIm Grupo Hospitalario Quirónsalud en Barcelona

quirónsalud
La salud persona a persona

INFORME DEL COMITÉ ÉTICO DE LA INVESTIGACIÓN

MONTSERRAT GRANADOS PLAZA, Secretaria Técnica del COMITÉ ÉTICO DE INVESTIGACIÓN con MEDICAMENTOS (CEIm) Grupo Hospitalario Quirónsalud en Barcelona, sito en C/ Vilana 12, 08022 Barcelona,

CERTIFICA

Que este Comité ha evaluado en fecha **26/09/2019 (acta nº62/2019)** la propuesta correspondiente al proyecto:
Título: **Análisis de la voz como método para predecir una vía aérea difícil de manera preoperatoria.**
Código Protocolo: **VOICE2019**
Código Interno: **2019/78-ANE-DEX**

Se cumplen los requisitos necesarios de idoneidad del protocolo en relación con los objetivos del estudio y están justificados los riesgos y molestias previsibles para el sujeto.

La capacidad del investigador y los medios disponibles son apropiados para llevar a cabo el estudio.

Son adecuados los procedimientos previstos para obtener el Consentimiento Informado.

El alcance de las compensaciones económicas previstas no interfiere con el respeto a los postulados éticos.

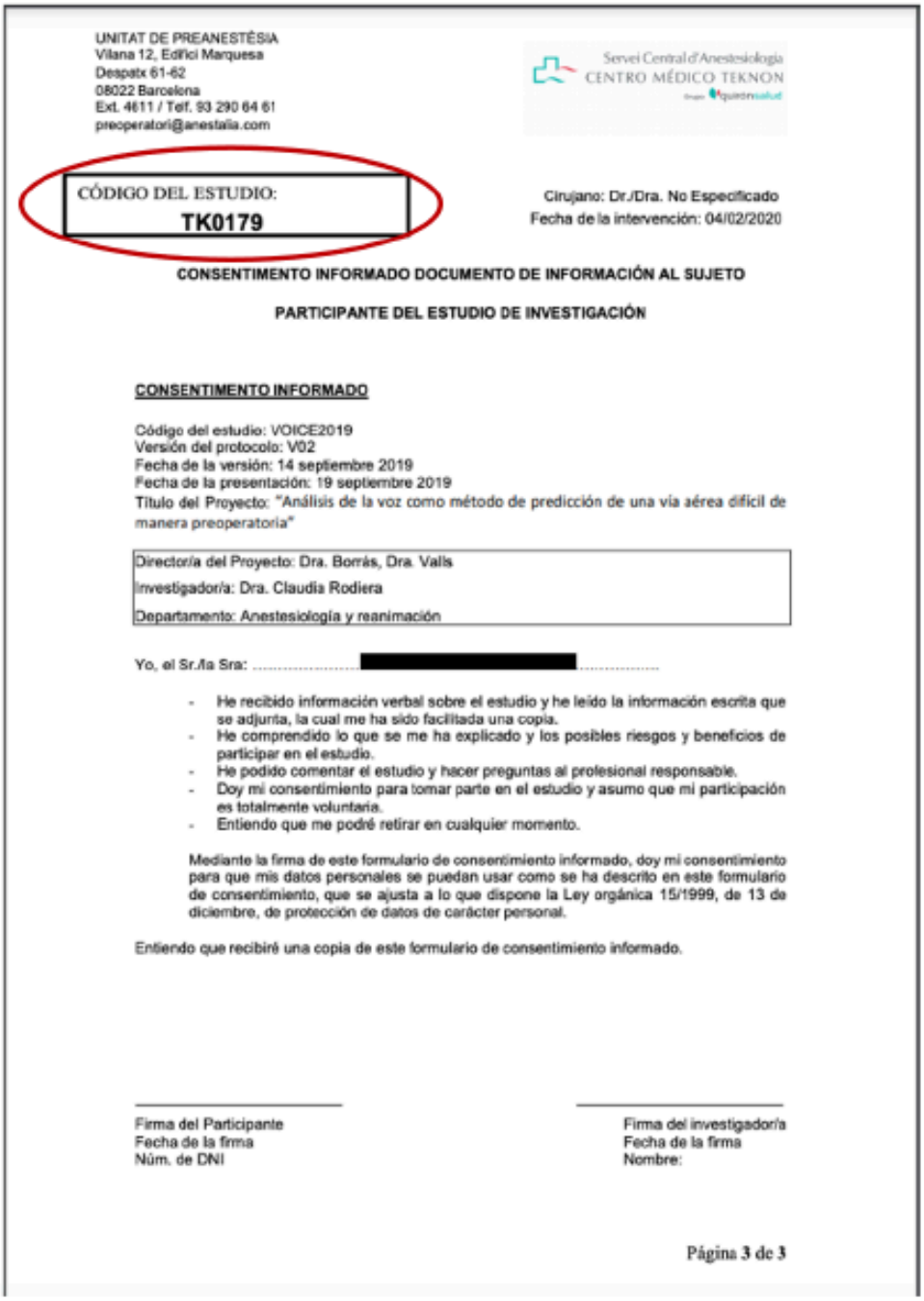
Y que este Comité emite **DICTAMEN FAVORABLE** para que dicho estudio sea realizado por **Claudia Rodiera** del **servicio de Anestesia y Reanimación** como Investigadora Principal en **Hospital Universitari Dexeus.**

En Barcelona, a 30 de septiembre de 2019.

GRANADOS PLAZA MONTSERRAT AMELIA - 45474810S

Montserrat Granados Plaza
Secretaria CEIm Grupo Hospitalario Quirónsalud en Barcelona

11.5. ANNEX 5: Informed consent of the participants



UNITAT DE PREANESTÈSIA
Vilana 12, Edifici Marquesa
Despatx 61-62
08022 Barcelona
Ext. 4611 / Telf. 93 290 64 61
preoperatori@anestalia.com

Servei Central d'Anestesiologia
CENTRO MÉDICO TEKNON

CÓDIGO DEL ESTUDIO:
TK0179

Cirujano: Dr./Dra. No Especificado
Fecha de la intervención: 04/02/2020

CONSENTIMIENTO INFORMADO DOCUMENTO DE INFORMACIÓN AL SUJETO

PARTICIPANTE DEL ESTUDIO DE INVESTIGACIÓN

CONSENTIMIENTO INFORMADO

Código del estudio: VOICE2019
Versión del protocolo: V02
Fecha de la versión: 14 septiembre 2019
Fecha de la presentación: 19 septiembre 2019
Título del Proyecto: "Análisis de la voz como método de predicción de una vía aérea difícil de manera preoperatoria"

Director/a del Proyecto: Dra. Borrás, Dra. Valls
Investigador/a: Dra. Claudia Rodiera
Departamento: Anestesiología y reanimación

Yo, el Sr./la Sra: .....[redacted].....

- He recibido información verbal sobre el estudio y he leído la información escrita que se adjunta, la cual me ha sido facilitada una copia.
- He comprendido lo que se me ha explicado y los posibles riesgos y beneficios de participar en el estudio.
- He podido comentar el estudio y hacer preguntas al profesional responsable.
- Doy mi consentimiento para tomar parte en el estudio y asumo que mi participación es totalmente voluntaria.
- Entiendo que me podré retirar en cualquier momento.

Mediante la firma de este formulario de consentimiento informado, doy mi consentimiento para que mis datos personales se puedan usar como se ha descrito en este formulario de consentimiento, que se ajusta a lo que dispone la Ley orgánica 15/1999, de 13 de diciembre, de protección de datos de carácter personal.

Entiendo que recibiré una copia de este formulario de consentimiento informado.

Firma del Participante
Fecha de la firma
Núm. de DNI

Firma del investigador/a
Fecha de la firma
Nombre:

Página 3 de 3

11.6. ANNEX 6: Data collection form

ESTUDIO VOICE

CÓDIGO DEL ESTUDIO: TK001	CirujanoDr/Dra: Fecha intervención:

1. Marcar con un círculo el valor de Cormack obtenido mediante LARINGOSCOPIA DIRECTA CON **PALA METÁLICA**

I IIA IIB III IV

2. Marque si ha necesitado hacer alguna de las siguientes maniobras:

- ☐ Uso videolaringoscopio (Especificar cuál: Glidescope/ McGraph
- ☐ Uso fibroscopio
- ☐ Guía
- ☐ BURP

3. Nombre anestesiólogo o enfermera que ha intubado : .....

RECUERDA DEPOSITAR HOJA EN BUZÓN DE FACTURACIÓN

11.7. ANNEX 7: Full table 6. The mean accuracy of the classification algorithms for the prediction of the Mallampati of the different combinations

ACCURACY					
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data All vocals		415/179	H2O GRADIENT BOOSTING	0,79	0,69
HARMONICS + Descriptive data A normal (_1)		415/179	H2O NAIVE BAYES	0,79	0,69
HARMONICS + Descriptive data A extension (_2)		415/179	H2O GENERALISED LINEAR MODEL	0,81	0,66
HARMONICS + Descriptive data A flexion (_3)		415/179	H2O GENERALISED LINEAR MODEL	0,81	0,74
HARMONICS + Descriptive data E normal (_1)		415/179	H2O GENERALISED LINEAR MODEL	0,82	0,73
HARMONICS + Descriptive data E extension (_2)		415/179	H2O GRADIENT BOOSTING	0,79	0,73
HARMONICS + Descriptive data E flexion(_3)		415/179	H2O GENERALISED LINEAR MODEL	0,82	0,75
HARMONICS + Descriptive data I normal (_1)		415/179	H2O GENERALISED LINEAR MODEL	0,76	0,74
HARMONICS + Descriptive data I extension (_2)		415/179	H2O GENERALISED LINEAR MODEL	0,82	0,75
HARMONICS + Descriptive data I flexion(_3)		415/179	H2O GENERALISED LINEAR MODEL	0,75	0,74
HARMONICS + Descriptive data O normal (_1)		415/179	H2O GENERALISED LINEAR MODEL	0,82	0,72
HARMONICS + Descriptive data O extension (_2)		415/179	H2O GENERALISED LINEAR MODEL	0,82	0,76
HARMONICS + Descriptive data O flexion(_3)		415/179	H2O GENERALISED LINEAR MODEL	0,82	0,72
HARMONICS + Descriptive data U normal (_1)		415/179	H2O GENERALISED LINEAR MODEL	0,81	0,74
HARMONICS + Descriptive data U extension 2)		415/179	H2O GENERALISED LINEAR MODEL	0,83	0,73
HARMONICS + Descriptive data U Flexion (_3)		415/179	H2O GENERALISED LINEAR MODEL	0,80	0,74
HARMONICS + Descriptive data A normal (_1) vs A flexion (_3)		415/179	H2O NAIVE BAYES	0,79	0,71
HARMONICS + Descriptive data A normal (_1) vs A extension (_2)		415/179	H2O NAIVE BAYES	0,79	0,68

ACCURACY					
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data A flexion (_3) vs A extension (_2)		415/179	H2O GRADIENT BOOSTING	0,80	0,70
HARMONICS + Descriptive data A all positions		415/179	H2O GRADIENT BOOSTING	0,82	0,56
HARMONICS + Descriptive data E normal (_1) vs E flexion (_3)		415/179	H2O GRADIENT BOOSTING	0,82	0,75
HARMONICS + Descriptive data E normal (_1) vs E extension(_2)		415/179	H2O GENERALISED LINEAR MODEL	0,82	0,73
HARMONICS + Descriptive data E flexion (_2) vs E extension (_3)		415/179	H2O GENERALISED LINEAR MODEL	0,82	0,72
HARMONICS + Descriptive data E all positions		415/179	H2O GENERALISED LINEAR MODEL	0,81	0,72
HARMONICS + Descriptive data I normal (_1) vs I flexion (_3)		415/179	H2O GENERALISED LINEAR MODEL	0,81	0,74
HARMONICS + Descriptive data I normal (_1) vs I extension(_2)		415/179	H2O GENERALISED LINEAR MODEL	0,82	0,74
HARMONICS + Descriptive data I flexion (_3) vs I extension (_2)		415/179	H2O GENERALISED LINEAR MODEL	0,81	0,75
HARMONICS + Descriptive data I all positions		415/179	H2O GENERALISED LINEAR MODEL	0,82	0,75
HARMONICS + Descriptive data O normal (_1) vs O flexion (_3)		415/179	H2O GENERALISED LINEAR MODEL	0,76	0,73
HARMONICS + Descriptive data O normal (_1) vs O extension (_2)		415/179	H2O NAIVE BAYES	0,79	0,70
HARMONICS + Descriptive data O flexion(_3) vs O extension (_2)		415/179	H2O GENERALISED LINEAR MODEL	0,77	0,72
HARMONICS + Descriptive data O all positions		415/179	H2O GENERALISED LINEAR MODEL	0,81	0,74
HARMONICS + Descriptive data U normal (_1) vs U flexion (_3)		415/179	H2O GENERALISED LINEAR MODEL	0,81	0,74
HARMONICS + Descriptive data U normal (_1) vs U extension (_2)		415/179	H2O GENERALISED LINEAR MODEL	0,82	0,74
HARMONICS + Descriptive data U flexion (_3) vs U extension (_2)		415/179	H2O GENERALISED LINEAR MODEL	0,77	0,76
HARMONICS + Descriptive data U all positions		415/179	H2O GENERALISED LINEAR MODEL	0,81	0,75
HARMONICS + Descriptive data All vocals	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,95	0,95

ACCURACY					
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data A normal (_1)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,97	0,98
HARMONICS + Descriptive data A extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,97	0,99
HARMONICS + Descriptive data A flexion (_3)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,96	0,98
HARMONICS + Descriptive data E normal (_1)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,97	0,99
HARMONICS + Descriptive data E extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,97	0,99
HARMONICS + Descriptive data E flexion(_3)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,97	0,98
HARMONICS + Descriptive data I normal (_1)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,97	0,99
HARMONICS + Descriptive data I extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,98	0,99
HARMONICS + Descriptive data I flexion(_3)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,97	0,99
HARMONICS + Descriptive data O normal (_1)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,97	0,99

ACCURACY					
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data O extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,97	0,99
HARMONICS + Descriptive data O flexion(_3)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,96	0,98
HARMONICS + Descriptive data U normal (_1)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,97	0,98
HARMONICS + Descriptive data U extension 2)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,97	0,98
HARMONICS + Descriptive data U Flexion (_3)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,96	0,99
HARMONICS + Descriptive data A normal (_1) vs A flexion (_3)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,97	0,98
HARMONICS + Descriptive data A normal (_1) vs A extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,98	0,99
HARMONICS + Descriptive data A flexion (_3) vs A extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,97	0,99
HARMONICS + Descriptive data A all positions	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,95	0,95
HARMONICS + Descriptive data E normal (_1) vs E flexion (_3)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,97	0,98

ACCURACY					
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data E normal (_1) vs E extension(_2)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,97	0,99
HARMONICS + Descriptive data E flexion (_2) vs E extension (_3)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,97	0,99
HARMONICS + Descriptive data E all positions	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,97	0,98
HARMONICS + Descriptive data I normal (_1) vs I flexion (_3)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,98	0,99
HARMONICS + Descriptive data I normal (_1) vs I extension(_2)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,97	0,96
HARMONICS + Descriptive data I flexion (_3) vs I extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,97	0,98
HARMONICS + Descriptive data I all positions	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,97	0,98
HARMONICS + Descriptive data O normal (_1) vs O flexion (_3)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,96	0,94
HARMONICS + Descriptive data O normal (_1) vs O extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,98	0,99
HARMONICS + Descriptive data O flexion(_3) vs O extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,98	0,99

ACCURACY					
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data O all positions	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,98	0,98
HARMONICS + Descriptive data U normal (_1) vs U flexion (_3)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,96	0,98
HARMONICS + Descriptive data U normal (_1) vs U extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,97	0,98
HARMONICS + Descriptive data U flexion (_3) vs U extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,97	0,98
HARMONICS + Descriptive data U all positions	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,96	0,98
Només DADES DESCRIPTIVES		415/179	H2O GENERALISED LINEAR MODEL	0,75	0,74
Només HARMÒNICS A normal (_1)		415/179	H2O GENERALISED LINEAR MODEL	0,35	0,34
Només HARMÒNICS A extensió (_2)		415/179	H2O RANDOM FOREST	0,46	0,21
Spetral + descriptive data All vocals		415/179	H2O GENERALISED LINEAR MODEL	0,77	0,72
Spectral+ descriptive data A normal		415/179	H2O GRADIENT BOOSTING	0,78	0,79
Spectral+ descriptive data O extension		415/179	H2O GRADIENT BOOSTING	0,80	0,58
Spectral+ descriptive data O flexion		415/179	H2O GENERALISED LINEAR MODEL	0,81	0,75
Spectral+ descriptive data A flexion vs A extension		415/179	H2O GENERALISED LINEAR MODEL	0,79	0,74
Spectral + descriptive data E flexion vs E e xtension		415/179	H2O GRADIENT BOOSTING	0,80	0,64
Voice parameters+ descriptive data All vocals		415/179	H2O GRADIENT BOOSTING	0,69	0,69

ACCURACY					
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%
Voice parameters+ descriptive data A normal		415/179	H2O GENERALISED LINEAR MODEL	0,80	0,74
Voice parameters+ descriptive data O extension		415/179	H2O GRADIENT BOOSTING	0,79	0,73
Voice parameters+ descriptive data O flexion		415/179	H2O NAIVE BAYES	0,78	0,70
Voice parameters+ descriptive data A flexion vs A extension		415/179	H2O NAIVE BAYES	0,79	0,66
Voice parameters+ descriptive data E flexion vs E extension		415/179	H2O NAIVE BAYES	0,76	0,59
Voice parameters+ descriptive data O flexion vs O extension		415/179	H2O NAIVE BAYES	0,78	0,68
Voice parameters+ descriptive data All vocals	Only Mallampati 1 and 4 cases	189/81	H2O RANDOM FOREST	0,39	0,95
Voice parameters+ descriptive data A normal	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,97	0,98
Voice parameters+ descriptive data O extension	Only Mallampati 1 and 4 cases	189/81	H2O GRADIENT BOOSTING	0,96	0,98
Voice parameters+ descriptive data O flexion	Only Mallampati 1 and 4 cases	189/81	H2O GRADIENT BOOSTING	0,96	0,98
Voice parameters+ descriptive data A flexion vs A extension	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,97	0,98
Voice parameters+ descriptive data E flexion vs E extension	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,96	0,98
Voice parameters+ descriptive data O flexion vs O extension	Only Mallampati 1 and 4 cases	189/81	H2O GRADIENT BOOSTING	0,97	0,98

11.8. ANNEX 8: Full table 7. The mean AUC values of the classification algorithms for the prediction of the Mallampati of the different combinations

				AUC		Pr AUC	
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data All vocals		415/179	H2O GRADIENT BOOSTING	0,68	0,74	0,39	0,36
HARMONICS + Descriptive data A normal (_1)		415/179	H2O NAIVE BAYES	0,76	0,72	0,46	0,39
HARMONICS + Descriptive data A extension (_2)		415/179	H2O GENERALISED LINEAR MODEL	0,75	0,73	0,52	0,40
HARMONICS + Descriptive data A flexion (_3)		415/179	H2O GENERALISED LINEAR MODEL	0,75	0,73	0,52	0,41
HARMONICS + Descriptive data E normal (_1)		415/179	H2O GENERALISED LINEAR MODEL	0,76	0,74	0,52	0,42
HARMONICS + Descriptive data E extension (_2)		415/179	H2O GRADIENT BOOSTING	0,70	0,75	0,37	0,43
HARMONICS + Descriptive data E flexion(_3)		415/179	H2O GENERALISED LINEAR MODEL	0,76	0,74	0,52	0,41
HARMONICS + Descriptive data I normal (_1)		415/179	H2O GENERALISED LINEAR MODEL	0,76	0,74	0,52	0,41
HARMONICS + Descriptive data I extension (_2)		415/179	H2O GENERALISED LINEAR MODEL	0,76	0,74	0,52	0,41
HARMONICS + Descriptive data I flexion(_3)		415/179	H2O GENERALISED LINEAR MODEL	0,75	0,74	0,52	0,41
HARMONICS + Descriptive data O normal (_1)		415/179	H2O GENERALISED LINEAR MODEL	0,76	0,74	0,52	0,41
HARMONICS + Descriptive data O extension (_2)		415/179	H2O GENERALISED LINEAR MODEL	0,76	0,74	0,52	0,41

				AUC		Pr AUC	
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data O flexion(_3)		415/179	H2O GENERALISED LINEAR MODEL	0,76	0,74	0,52	0,40
HARMONICS + Descriptive data U normal (_1)		415/179	H2O GENERALISED LINEAR MODEL	0,76	0,74	0,53	0,41
HARMONICS + Descriptive data U extension 2)		415/179	H2O GENERALISED LINEAR MODEL	0,76	0,74	0,52	0,41
HARMONICS + Descriptive data U Flexion (_3)		415/179	H2O GENERALISED LINEAR MODEL	0,76	0,73	0,52	0,40
HARMONICS + Descriptive data A normal (_1) vs A flexion (_3)		415/179	H2O NAIVE BAYES	0,75	0,72	0,44	0,39
HARMONICS + Descriptive data A normal (_1) vs A extension (_2)		415/179	H2O NAIVE BAYES	0,76	0,71	0,45	0,37
HARMONICS + Descriptive data A flexion (_3) vs A extension (_2)		415/179	H2O GRADIENT BOOSTING	0,67	0,74	0,38	0,41
HARMONICS + Descriptive data A all positions		415/179	H2O GRADIENT BOOSTING	0,69	0,71	0,32	0,35
HARMONICS + Descriptive data E normal (_1) vs E flexion (_3)		415/179	H2O GRADIENT BOOSTING	0,71	0,74	0,43	0,41
HARMONICS + Descriptive data E normal (_1) vs E extension(_2)		415/179	H2O GENERALISED LINEAR MODEL	0,76	0,74	0,52	0,41
HARMONICS + Descriptive data E flexion (_2) vs E extension (_3)		415/179	H2O GENERALISED LINEAR MODEL	0,75	0,74	0,52	0,41

				AUC		Pr AUC	
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data E all positions		415/179	H2O GENERALISED LINEAR MODEL	0,75	0,74	0,52	0,41
HARMONICS + Descriptive data I normal (_1) vs I flexion (_3)		415/179	H2O GENERALISED LINEAR MODEL	0,76	0,74	0,53	0,41
HARMONICS + Descriptive data I normal (_1) vs I extension(_2)		415/179	H2O GENERALISED LINEAR MODEL	0,76	0,74	0,53	0,41
HARMONICS + Descriptive data I flexion (_3) vs I extension (_2)		415/179	H2O GENERALISED LINEAR MODEL	0,76	0,74	0,53	0,41
HARMONICS + Descriptive data I all positions		415/179	H2O GENERALISED LINEAR MODEL	0,76	0,74	0,53	0,41
HARMONICS + Descriptive data O normal (_1) vs O flexion (_3)		415/179	H2O GENERALISED LINEAR MODEL	0,75	0,74	0,50	0,41
HARMONICS + Descriptive data O normal (_1) vs O extension (_2)		415/179	H2O NAIVE BAYES	0,75	0,73	0,45	0,39
HARMONICS + Descriptive data O flexion(_3) vs O extension (_2)		415/179	H2O GENERALISED LINEAR MODEL	0,75	0,74	0,51	0,41
HARMONICS + Descriptive data O all positions		415/179	H2O GENERALISED LINEAR MODEL	0,75	0,74	0,50	0,42
HARMONICS + Descriptive data U normal (_1) vs U flexion (_3)		415/179	H2O GENERALISED LINEAR MODEL	0,76	0,74	0,52	0,40
HARMONICS + Descriptive data U normal (_1) vs U extension (_2)		415/179	H2O GENERALISED LINEAR MODEL	0,76	0,74	0,53	0,41

				AUC		Pr AUC	
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data U flexion (_3) vs U extension (_2)		415/179	H2O GENERALISED LINEAR MODEL	0,76	0,74	0,52	0,40
HARMONICS + Descriptive data U all positions		415/179	H2O GENERALISED LINEAR MODEL	0,76	0,74	0,53	0,41
HARMONICS + Descriptive data All vocals	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,89	0,97	0,29	0,23
HARMONICS + Descriptive data A normal (_1)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,86	0,97	0,23	0,50
HARMONICS + Descriptive data A extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,85	0,99	0,25	0,38
HARMONICS + Descriptive data A flexion (_3)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,87	0,96	0,27	0,49
HARMONICS + Descriptive data E normal (_1)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,86	0,97	0,23	0,58
HARMONICS + Descriptive data E extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,88	0,99	0,24	0,38
HARMONICS + Descriptive data E flexion(_3)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,87	0,96	0,23	0,49
HARMONICS + Descriptive data I normal (_1)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,86	0,97	0,23	0,58

				AUC		Pr AUC	
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data I extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,86	0,97	0,26	0,58
HARMONICS + Descriptive data I flexion(_3)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,90	0,99	0,24	0,38
HARMONICS + Descriptive data O normal (_1)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,88	0,99	0,24	0,37
HARMONICS + Descriptive data O extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,90	0,99	0,24	0,38
HARMONICS + Descriptive data O flexion(_3)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,86	0,99	0,22	0,31
HARMONICS + Descriptive data U normal (_1)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,88	0,99	0,24	0,31
HARMONICS + Descriptive data U extension 2)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,90	0,98	0,24	0,29
HARMONICS + Descriptive data U Flexion (_3)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,88	0,99	0,22	0,37
HARMONICS + Descriptive data A normal (_1) vs A flexion (_3)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,83	0,98	0,32	0,52
HARMONICS + Descriptive data A normal (_1) vs A extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,87	0,97	0,26	0,56

				AUC		Pr AUC	
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data A flexion (_3) vs A extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,86	0,97	0,20	0,56
HARMONICS + Descriptive data A all positions	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,85	0,98	0,21	0,23
HARMONICS + Descriptive data E normal (_1) vs E flexion (_3)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,86	0,97	0,20	0,50
HARMONICS + Descriptive data E normal (_1) vs E extension(_2)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,86	0,98	0,21	0,58
HARMONICS + Descriptive data E flexion (_2) vs E extension (_3)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,85	0,98	0,20	0,58
HARMONICS + Descriptive data E all positions	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,86	0,97	0,22	0,50
HARMONICS + Descriptive data I normal (_1) vs I flexion (_3)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,86	0,97	0,26	0,58
HARMONICS + Descriptive data I normal (_1) vs I extension(_2)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,89	0,98	0,24	0,25
HARMONICS + Descriptive data I flexion (_3) vs I extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,89	0,99	0,24	0,31
HARMONICS + Descriptive data I all positions	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,89	0,99	0,24	0,34

				AUC		Pr AUC	
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data O normal (_1) vs O flexion (_3)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,85	0,97	0,18	0,21
HARMONICS + Descriptive data O normal (_1) vs O extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,87	0,97	0,27	0,58
HARMONICS + Descriptive data O flexion(_3) vs O extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,86	0,97	0,29	0,58
HARMONICS + Descriptive data O all positions	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,87	0,97	0,28	0,40
HARMONICS + Descriptive data U normal (_1) vs U flexion (_3)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,88	0,99	0,22	0,30
HARMONICS + Descriptive data U normal (_1) vs U extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,87	0,99	0,22	0,34
HARMONICS + Descriptive data U flexion (_3) vs U extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,85	0,97	0,19	0,50
HARMONICS + Descriptive data U all positions	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,91	0,98	0,23	0,29
Només DADES DESCRIPTIVES		415/179	H2O GENERALISED LINEAR MODEL	0,75	0,73	0,52	0,41
Només HARMÒNICS A normal (_1)		415/179	H2O GENERALISED LINEAR MODEL	0,50	0,51	0,21	0,21
Només HARMÒNICS A extensió (_2)		415/179	H2O RANDOM FOREST	0,47	0,44	0,23	0,18

				AUC		Pr AUC	
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%	TRAINING 70%	TESTING 30%
Spetral + descriptive data All vocals		415/179	H2O GENERALISED LINEAR MODEL	0,73	0,72	0,42	0,37
Spectral+ descriptive data A normal		415/179	H2O GRADIENT BOOSTING	0,64	0,69	0,35	0,39
Spectral+ descriptive data O extension		415/179	H2O GRADIENT BOOSTING	0,69	0,65	0,41	0,32
Spectral+ descriptive data O flexion		415/179	H2O GENERALISED LINEAR MODEL	0,76	0,74	0,50	0,41
Spectral+ descriptive data A flexion vs A extension		415/179	H2O GENERALISED LINEAR MODEL	0,72	0,71	0,45	0,39
Spectral + descriptive data E flexion vs E e xtension		415/179	H2O GRADIENT BOOSTING	0,72	0,70	0,42	0,34
Voice parameters+ descriptive data All vocals		415/179	H2O GRADIENT BOOSTING	0,60	0,66	0,30	0,36
Voice parameters+ descriptive data A normal		415/179	H2O GENERALISED LINEAR MODEL	0,75	0,74	0,49	0,40
Voice parameters+ descriptive data O extension		415/179	H2O GRADIENT BOOSTING	0,66	0,72	0,35	0,37
Voice parameters+ descriptive data O flexion		415/179	H2O NAIVE BAYES	0,75	0,75	0,41	0,36
Voice parameters+ descriptive data A flexion vs A extension		415/179	H2O NAIVE BAYES	0,71	0,68	0,42	0,35
Voice parameters+ descriptive data E flexion vs E extension		415/179	H2O NAIVE BAYES	0,76	0,69	0,44	0,31

				AUC		Pr AUC	
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%	TRAINING 70%	TESTING 30%
Voice parameters+ descriptive data O flexion vs O extension		415/179	H2O NAIVE BAYES	0,75	0,72	0,45	0,34
Voice parameters+ descriptive data All vocals	Only Mallampati 1 and 4 cases	189/81	H2O RANDOM FOREST	0,66	0,81	0,04	0,16
Voice parameters+ descriptive data A normal	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,80	0,97	0,23	0,41
Voice parameters+ descriptive data O extension	Only Mallampati 1 and 4 cases	189/81	H2O GRADIENT BOOSTING	0,70	0,97	0,12	0,48
Voice parameters+ descriptive data O flexion	Only Mallampati 1 and 4 cases	189/81	H2O GRADIENT BOOSTING	0,75	0,99	0,20	0,56
Voice parameters+ descriptive data A flexion vs A extension	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,53	0,98	0,15	0,46

				AUC		Pr AUC	
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%	TRAINING 70%	TESTING 30%
Voice parameters+ descriptive data E flexion vs E extension	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,89	0,97	0,30	0,50
Voice parameters+ descriptive data O flexion vs O extension	Only Mallampati 1 and 4 cases	189/81	H2O GRADIENT BOOSTING	0,77	0,93	0,19	0,36

11.9. ANNEX 9. Full Table 8: Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in both training and testing groups, Precision values for the training group, Recall values for the testing group

				Precision		Recall	
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data All vocals		415/179	H2O GRADIENT BOOSTING	0,52	0,38	0,45	0,74
HARMONICS + Descriptive data A normal (_1)		415/179	H2O NAIVE BAYES	0,51	0,37	0,64	0,68
HARMONICS + Descriptive data A extension (_2)		415/179	H2O GENERALISED LINEAR MODEL	0,59	0,35	0,63	0,74
HARMONICS + Descriptive data A flexion (_3)		415/179	H2O GENERALISED LINEAR MODEL	0,62	0,42	0,60	0,63
HARMONICS + Descriptive data E normal (_1)		415/179	H2O GENERALISED LINEAR MODEL	0,62	0,41	0,56	0,66
HARMONICS + Descriptive data E extension (_2)		415/179	H2O GRADIENT BOOSTING	0,51	0,41	0,50	0,63
HARMONICS + Descriptive data E flexion(_3)		415/179	H2O GENERALISED LINEAR MODEL	0,63	0,44	0,57	0,63
HARMONICS + Descriptive data I normal (_1)		415/179	H2O GENERALISED LINEAR MODEL	0,57	0,42	0,72	0,63
HARMONICS + Descriptive data I extension (_2)		415/179	H2O GENERALISED LINEAR MODEL	0,61	0,44	0,58	0,63
HARMONICS + Descriptive data I flexion(_3)		415/179	H2O GENERALISED LINEAR MODEL	0,56	0,43	0,72	0,61
HARMONICS + Descriptive data O normal (_1)		415/179	H2O GENERALISED LINEAR MODEL	0,63	0,40	0,56	0,66

				Precision		Recall	
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data O extension (_2)		415/179	H2O GENERALISED LINEAR MODEL	0,60	0,45	0,59	0,61
HARMONICS + Descriptive data O flexion(_3)		415/179	H2O GENERALISED LINEAR MODEL	0,63	0,40	0,55	0,66
HARMONICS + Descriptive data U normal (_1)		415/179	H2O GENERALISED LINEAR MODEL	0,60	0,43	0,59	0,63
HARMONICS + Descriptive data U extension 2)		415/179	H2O GENERALISED LINEAR MODEL	0,64	0,41	0,56	0,66
HARMONICS + Descriptive data U Flexion (_3)		415/179	H2O GENERALISED LINEAR MODEL	0,56	0,42	0,64	0,63
HARMONICS + Descriptive data A normal (_1) vs A flexion (_3)		415/179	H2O NAIVE BAYES	0,50	0,39	0,66	0,66
HARMONICS + Descriptive data A normal (_1) vs A extension (_2)		415/179	H2O NAIVE BAYES	0,51	0,36	0,65	0,71
HARMONICS + Descriptive data A flexion (_3) vs A extension (_2)		415/179	H2O GRADIENT BOOSTING	0,61	0,38	0,46	0,68
HARMONICS + Descriptive data A all positions		415/179	H2O GRADIENT BOOSTING	0,62	0,31	0,46	0,84
HARMONICS + Descriptive data E normal (_1) vs E flexion (_3)		415/179	H2O GRADIENT BOOSTING	0,61	0,44	0,45	0,63
HARMONICS + Descriptive data E normal (_1) vs E extension(_2)		415/179	H2O GENERALISED LINEAR MODEL	0,60	0,41	0,59	0,63

				Precision		Recall	
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data E flexion (_2) vs E extension (_3)		415/179	H2O GENERALISED LINEAR MODEL	0,62	0,40	0,59	0,66
HARMONICS + Descriptive data E all positions		415/179	H2O GENERALISED LINEAR MODEL	0,59	0,40	0,58	0,66
HARMONICS + Descriptive data I normal (_1) vs I flexion (_3)		415/179	H2O GENERALISED LINEAR MODEL	0,59	0,43	0,59	0,63
HARMONICS + Descriptive data I normal (_1) vs I extension(_2)		415/179	H2O GENERALISED LINEAR MODEL	0,61	0,43	0,59	0,63
HARMONICS + Descriptive data I flexion (_3) vs I extension (_2)		415/179	H2O GENERALISED LINEAR MODEL	0,59	0,44	0,63	0,63
HARMONICS + Descriptive data I all positions		415/179	H2O GENERALISED LINEAR MODEL	0,60	0,44	0,59	0,63
HARMONICS + Descriptive data O normal (_1) vs O flexion (_3)		415/179	H2O GENERALISED LINEAR MODEL	0,54	0,41	0,69	0,68
HARMONICS + Descriptive data O normal (_1) vs O extension (_2)		415/179	H2O NAIVE BAYES	0,51	0,39	0,65	0,68
HARMONICS + Descriptive data O flexion(_3) vs O extension (_2)		415/179	H2O GENERALISED LINEAR MODEL	0,56	0,40	0,65	0,68
HARMONICS + Descriptive data O all positions		415/179	H2O GENERALISED LINEAR MODEL	0,58	0,42	0,60	0,63
HARMONICS + Descriptive data U normal (_1) vs U flexion (_3)		415/179	H2O GENERALISED LINEAR MODEL	0,60	0,42	0,63	0,63

				Precision		Recall	
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data U normal (_1) vs U extension (_2)		415/179	H2O GENERALISED LINEAR MODEL	0,65	0,42	0,55	0,66
HARMONICS + Descriptive data U flexion (_3) vs U extension (_2)		415/179	H2O GENERALISED LINEAR MODEL	0,59	0,45	0,65	0,61
HARMONICS + Descriptive data U all positions		415/179	H2O GENERALISED LINEAR MODEL	0,60	0,43	0,60	0,61
HARMONICS + Descriptive data All vocals	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,65	0,50	0,63	1,00
HARMONICS + Descriptive data A normal (_1)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,88	0,75	0,47	0,75
HARMONICS + Descriptive data A extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,63	1,00	0,63	0,75
HARMONICS + Descriptive data A flexion (_3)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,54	0,75	0,47	0,75
HARMONICS + Descriptive data E normal (_1)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,88	1,00	0,47	0,75
HARMONICS + Descriptive data E extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,67	1,00	0,73	0,75
HARMONICS + Descriptive data E flexion(_3)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,88	0,75	0,47	0,75
HARMONICS + Descriptive data I normal (_1)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,88	1,00	0,47	0,75
HARMONICS + Descriptive data I extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	1,00	1,00	0,47	0,75

				Precision		Recall	
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data I flexion(_3)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,67	1,00	0,73	0,75
HARMONICS + Descriptive data O normal (_1)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,67	0,80	0,73	1,00
HARMONICS + Descriptive data O extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,67	1,00	0,73	0,75
HARMONICS + Descriptive data O flexion(_3)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,58	0,75	0,73	0,75
HARMONICS + Descriptive data U normal (_1)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,67	0,75	0,73	0,75
HARMONICS + Descriptive data U extension 2)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,67	0,75	0,73	0,75
HARMONICS + Descriptive data U Flexion (_3)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,58	1,00	0,73	0,75
HARMONICS + Descriptive data A normal (_1) vs A flexion (_3)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,71	0,75	0,63	0,75
HARMONICS + Descriptive data A normal (_1) vs A extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	1,00	1,00	0,47	0,75
HARMONICS + Descriptive data A flexion (_3) vs A extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,88	1,00	0,47	0,75
HARMONICS + Descriptive data A all positions	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,47	0,50	0,73	1,00
HARMONICS + Descriptive data E normal (_1) vs E flexion (_3)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,88	0,75	0,47	0,75

				Precision		Recall	
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data E normal (_1) vs E extension(_2)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,80	1,00	0,47	0,75
HARMONICS + Descriptive data E flexion (_2) vs E extension (_3)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,88	1,00	0,47	0,75
HARMONICS + Descriptive data E all positions	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	1,00	0,75	0,37	0,75
HARMONICS + Descriptive data I normal (_1) vs I flexion (_3)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	1,00	1,00	0,47	0,75
HARMONICS + Descriptive data I normal (_1) vs I extension(_2)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,67	0,57	0,73	1,00
HARMONICS + Descriptive data I flexion (_3) vs I extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,67	0,75	0,73	0,75
HARMONICS + Descriptive data I all positions	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,67	0,67	0,73	1,00
HARMONICS + Descriptive data O normal (_1) vs O flexion (_3)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,55	0,44	0,63	1,00
HARMONICS + Descriptive data O normal (_1) vs O extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	1,00	1,00	0,47	0,75
HARMONICS + Descriptive data O flexion(_3) vs O extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	1,00	1,00	0,47	0,75
HARMONICS + Descriptive data O all positions	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	1,00	0,75	0,47	0,75

				Precision		Recall	
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data U normal (_1) vs U flexion (_3)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,58	0,67	0,73	1,00
HARMONICS + Descriptive data U normal (_1) vs U extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,62	0,67	0,73	1,00
HARMONICS + Descriptive data U flexion (_3) vs U extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,88	0,75	0,47	0,75
HARMONICS + Descriptive data U all positions	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,58	0,75	0,73	0,75
Només DADES DESCRIPTIVES		415/179	H2O GENERALISED LINEAR MODEL	0,56	0,43	0,73	0,61
Només HARMÒNICS A normal (_1)		415/179	H2O GENERALISED LINEAR MODEL	0,24	0,24	0,91	0,97
Només HARMÒNICS A extensió (_2)		415/179	H2O RANDOM FOREST	0,27	0,21	0,78	1,00
Spetral + descriptive data All vocals		415/179	H2O GENERALISED LINEAR MODEL	0,52	0,40	0,61	0,66
Spectral+ descriptive data A normal		415/179	H2O GRADIENT BOOSTING	0,54	0,51	0,43	0,55
Spectral+ descriptive data O extension		415/179	H2O GRADIENT BOOSTING	0,61	0,31	0,44	0,76
Spectral+ descriptive data O flexion		415/179	H2O GENERALISED LINEAR MODEL	0,58	0,44	0,56	0,68
Spectral+ descriptive data A flexion vs A extension		415/179	H2O GENERALISED LINEAR MODEL	0,52	0,42	0,55	0,66
Spectral + descriptive data E flexion vs E e xtension		415/179	H2O GRADIENT BOOSTING	0,61	0,33	0,47	0,71

				Precision		Recall	
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%	TRAINING 70%	TESTING 30%
Voice parameters+ descriptive data All vocals		415/179	H2O GRADIENT BOOSTING	0,34	0,37	0,50	0,63
Voice parameters+ descriptive data A normal		415/179	H2O GENERALISED LINEAR MODEL	0,54	0,42	0,61	0,66
Voice parameters+ descriptive data O extension		415/179	H2O GRADIENT BOOSTING	0,52	0,41	0,44	0,63
Voice parameters+ descriptive data O flexion		415/179	H2O NAIVE BAYES	0,49	0,40	0,66	0,79
Voice parameters+ descriptive data A flexion vs A extension		415/179	H2O NAIVE BAYES	0,52	0,34	0,64	0,63
Voice parameters+ descriptive data E flexion vs E extension		415/179	H2O NAIVE BAYES	0,47	0,32	0,69	0,82
Voice parameters+ descriptive data O flexion vs O extension		415/179	H2O NAIVE BAYES	0,50	0,37	0,66	0,76
Voice parameters+ descriptive data All vocals	Only Mallampati 1 and 4 cases	189/81	H2O RANDOM FOREST	0,08	0,50	0,80	0,50
Voice parameters+ descriptive data A normal	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,75	0,75	0,37	0,75
Voice parameters+ descriptive data O extension	Only Mallampati 1 and 4 cases	189/81	H2O GRADIENT BOOSTING	0,75	0,75	0,37	0,75
Voice parameters+ descriptive data O flexion	Only Mallampati 1 and 4 cases	189/81	H2O GRADIENT BOOSTING	0,50	0,75	0,27	0,75

				Precision		Recall	
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%	TRAINING 70%	TESTING 30%
Voice parameters+ descriptive data A flexion vs A extension	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	1,00	0,75	0,37	0,75
Voice parameters+ descriptive data E flexion vs E extension	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,69	0,75	0,70	0,75
Voice parameters+ descriptive data O flexion vs O extension	Only Mallampati 1 and 4 cases	189/81	H2O GRADIENT BOOSTING	1,00	0,75	0,37	0,75

11.10. ANNEX 10. Full Table 9: Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in both training and testing groups, F1 and Log loss values for the training group and the testing group respectively

				Log loss		F1	
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data All vocals		415/179	H2O GRADIENT BOOSTING	0,58	0,50	0,47	0,50
HARMONICS + Descriptive data A normal (_1)		415/179	H2O NAIVE BAYES	0,80	1,00	0,57	0,48
HARMONICS + Descriptive data A extension (_2)		415/179	H2O GENERALISED LINEAR MODEL	0,59	0,50	0,59	0,48
HARMONICS + Descriptive data A flexion (_3)		415/179	H2O GENERALISED LINEAR MODEL	0,59	0,50	0,58	0,51
HARMONICS + Descriptive data E normal (_1)		415/179	H2O GENERALISED LINEAR MODEL	0,58	0,50	0,58	0,51
HARMONICS + Descriptive data E extension (_2)		415/179	H2O GRADIENT BOOSTING	0,69	0,48	0,50	0,49
HARMONICS + Descriptive data E flexion(_3)		415/179	H2O GENERALISED LINEAR MODEL	0,59	0,50	0,58	0,52
HARMONICS + Descriptive data I normal (_1)		415/179	H2O GENERALISED LINEAR MODEL	0,58	0,50	0,59	0,51
HARMONICS + Descriptive data I extension (_2)		415/179	H2O GENERALISED LINEAR MODEL	0,58	0,50	0,58	0,52

				Log loss		F1	
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data I flexion(_3)		415/179	H2O GENERALISED LINEAR MODEL	0,59	0,50	0,58	0,50
HARMONICS + Descriptive data O normal (_1)		415/179	H2O GENERALISED LINEAR MODEL	0,58	0,50	0,59	0,50
HARMONICS + Descriptive data O extension (_2)		415/179	H2O GENERALISED LINEAR MODEL	0,58	0,50	0,59	0,52
HARMONICS + Descriptive data O flexion(_3)		415/179	H2O GENERALISED LINEAR MODEL	0,59	0,50	0,58	0,50
HARMONICS + Descriptive data U normal (_1)		415/179	H2O GENERALISED LINEAR MODEL	0,58	0,50	0,59	0,51
HARMONICS + Descriptive data U extension 2)		415/179	H2O GENERALISED LINEAR MODEL	0,58	0,50	0,59	0,51
HARMONICS + Descriptive data U Flexion (_3)		415/179	H2O GENERALISED LINEAR MODEL	0,59	0,50	0,58	0,51
HARMONICS + Descriptive data A normal (_1) vs A flexion (_3)		415/179	H2O NAIVE BAYES	0,80	0,97	0,57	0,49
HARMONICS + Descriptive data A normal (_1) vs A extension (_2)		415/179	H2O NAIVE BAYES	0,83		0,57	0,48
HARMONICS + Descriptive data A flexion (_3) vs A extension (_2)		415/179	H2O GRADIENT BOOSTING	0,72	0,63	0,50	0,49
HARMONICS + Descriptive data A all positions		415/179	H2O GRADIENT BOOSTING	0,67	0,49	0,52	0,45

				Log loss		F1	
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data E normal (_1) vs E flexion (_3)		415/179	H2O GRADIENT BOOSTING	0,57	0,46	0,51	0,52
HARMONICS + Descriptive data E normal (_1) vs E extension(_2)		415/179	H2O GENERALISED LINEAR MODEL	0,59	0,50	0,59	0,50
HARMONICS + Descriptive data E flexion (_2) vs E extension (_3)		415/179	H2O GENERALISED LINEAR MODEL	0,59	0,50	0,59	0,50
HARMONICS + Descriptive data E all positions		415/179	H2O GENERALISED LINEAR MODEL	0,59	0,50	0,58	0,50
HARMONICS + Descriptive data I normal (_1) vs I flexion (_3)		415/179	H2O GENERALISED LINEAR MODEL	0,58	0,50	0,58	0,51
HARMONICS + Descriptive data I normal (_1) vs I extension(_2)		415/179	H2O GENERALISED LINEAR MODEL	0,58	0,50	0,58	0,51
HARMONICS + Descriptive data I flexion (_3) vs I extension (_2)		415/179	H2O GENERALISED LINEAR MODEL	0,58	0,50	0,59	0,52
HARMONICS + Descriptive data I all positions		415/179	H2O GENERALISED LINEAR MODEL	0,58	0,50	0,59	0,52
HARMONICS + Descriptive data O normal (_1) vs O flexion (_3)		415/179	H2O GENERALISED LINEAR MODEL	0,59	0,50	0,57	0,51
HARMONICS + Descriptive data O normal (_1) vs O extension (_2)		415/179	H2O NAIVE BAYES	0,84		0,57	0,50
HARMONICS + Descriptive data O flexion(_3) vs O extension (_2)		415/179	H2O GENERALISED LINEAR MODEL	0,59	0,50	0,58	0,50

				Log loss		F1	
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data O all positions		415/179	H2O GENERALISED LINEAR MODEL	0,59	0,49	0,58	0,51
HARMONICS + Descriptive data U normal (_1) vs U flexion (_3)		415/179	H2O GENERALISED LINEAR MODEL	0,58	0,50	0,59	0,51
HARMONICS + Descriptive data U normal (_1) vs U extension (_2)		415/179	H2O GENERALISED LINEAR MODEL	0,58	0,50	0,59	0,52
HARMONICS + Descriptive data U flexion (_3) vs U extension (_2)		415/179	H2O GENERALISED LINEAR MODEL	0,59	0,50	0,58	0,52
HARMONICS + Descriptive data U all positions		415/179	H2O GENERALISED LINEAR MODEL	0,59	0,50	0,59	0,51
HARMONICS + Descriptive data All vocals	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,33	0,70	0,60	0,67
HARMONICS + Descriptive data A normal (_1)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,37	0,09	0,58	0,75
HARMONICS + Descriptive data A extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,34	0,64	0,63	0,86
HARMONICS + Descriptive data A flexion (_3)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,38	0,10	0,50	0,75
HARMONICS + Descriptive data E normal (_1)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,38	0,09	0,58	0,86
HARMONICS + Descriptive data E extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,42	0,81	0,70	0,86

				Log loss		F1	
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data E flexion(_3)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,39	0,09	0,58	0,75
HARMONICS + Descriptive data I normal (_1)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,38	0,09	0,58	0,86
HARMONICS + Descriptive data I extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,39	0,09	0,63	0,86
HARMONICS + Descriptive data I flexion(_3)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,43	0,80	0,70	0,86
HARMONICS + Descriptive data O normal (_1)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,46	0,74	0,70	0,89
HARMONICS + Descriptive data O extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,42	0,77	0,70	0,86
HARMONICS + Descriptive data O flexion(_3)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,40	0,72	0,64	0,75
HARMONICS + Descriptive data U normal (_1)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,42	0,79	0,70	0,75
HARMONICS + Descriptive data U extension 2)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,44	0,82	0,70	0,75
HARMONICS + Descriptive data U Flexion (_3)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,42	0,56	0,64	0,86
HARMONICS + Descriptive data A normal (_1) vs A flexion (_3)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,28	0,47	0,67	0,75
HARMONICS + Descriptive data A normal (_1) vs A extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,37	0,09	0,63	0,86

				Log loss		F1	
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data A flexion (_3) vs A extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,37	0,09	0,58	0,86
HARMONICS + Descriptive data A all positions	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,59	0,74	0,57	0,67
HARMONICS + Descriptive data E normal (_1) vs E flexion (_3)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,38	0,09	0,58	0,75
HARMONICS + Descriptive data E normal (_1) vs E extension(_2)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,38	0,09	0,55	0,86
HARMONICS + Descriptive data E flexion (_2) vs E extension (_3)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,36	0,09	0,58	0,86
HARMONICS + Descriptive data E all positions	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,35	0,09	0,54	0,75
HARMONICS + Descriptive data I normal (_1) vs I flexion (_3)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,39	0,09	0,63	0,86
HARMONICS + Descriptive data I normal (_1) vs I extension(_2)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,36	0,76	0,70	0,73
HARMONICS + Descriptive data I flexion (_3) vs I extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,42	0,80	0,70	0,75
HARMONICS + Descriptive data I all positions	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,43	0,76	0,70	0,80
HARMONICS + Descriptive data O normal (_1) vs O flexion (_3)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,70	0,89	0,58	0,62

				Log loss		F1	
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data O normal (_1) vs O extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,37	0,09	0,63	0,86
HARMONICS + Descriptive data O flexion(_3) vs O extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,38	0,09	0,63	0,86
HARMONICS + Descriptive data O all positions	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,36	0,09	0,63	0,75
HARMONICS + Descriptive data U normal (_1) vs U flexion (_3)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,44	0,65	0,64	0,80
HARMONICS + Descriptive data U normal (_1) vs U extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,50	0,73	0,67	0,80
HARMONICS + Descriptive data U flexion (_3) vs U extension (_2)	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,37	0,09	0,58	0,75
HARMONICS + Descriptive data U all positions	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,35	0,86	0,64	0,75
Només DADES DESCRIPTIVES		415/179	H2O GENERALISED LINEAR MODEL	0,58	0,50	0,58	0,50
Només HARMÒNICS A normal (_1)		415/179	H2O GENERALISED LINEAR MODEL	0,73	0,53	0,37	0,38
Només HARMÒNICS A extensió (_2)		415/179	H2O RANDOM FOREST	0,73	0,55	0,39	0,35
Spetral + descriptive data All vocals		415/179	H2O GENERALISED LINEAR MODEL	0,56	0,49	0,53	0,50

				Log loss		F1	
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%	TRAINING 70%	TESTING 30%
Spectral+ descriptive data A normal		415/179	H2O GRADIENT BOOSTING	0,62	0,54	0,46	0,53
Spectral+ descriptive data O extension		415/179	H2O GRADIENT BOOSTING	0,57	0,59	0,49	0,44
Spectral+ descriptive data O flexion		415/179	H2O GENERALISED LINEAR MODEL	0,58	0,47	0,56	0,54
Spectral+ descriptive data A flexion vs A extension		415/179	H2O GENERALISED LINEAR MODEL	0,60	0,48	0,52	0,52
Spectral + descriptive data E flexion vs E e xtension		415/179	H2O GRADIENT BOOSTING	0,73	0,49	0,52	0,45
Voice parameters+ descriptive data All vocals		415/179	H2O GRADIENT BOOSTING	0,60	0,52	0,40	0,47
Voice parameters+ descriptive data A normal		415/179	H2O GENERALISED LINEAR MODEL	0,58	0,50	0,57	0,52
Voice parameters+ descriptive data O extension		415/179	H2O GRADIENT BOOSTING	0,60	0,48	0,46	0,50
Voice parameters+ descriptive data O flexion		415/179	H2O NAIVE BAYES	0,90		0,56	0,53
Voice parameters+ descriptive data A flexion vs A extension		415/179	H2O NAIVE BAYES	0,80		0,57	0,44
Voice parameters+ descriptive data E flexion vs E extension		415/179	H2O NAIVE BAYES	0,86		0,56	0,46
Voice parameters+ descriptive data O flexion vs O extension		415/179	H2O NAIVE BAYES	0,92		0,56	0,50

				Log loss		F1	
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TRAINING 70%	TESTING 30%	TRAINING 70%	TESTING 30%
Voice parameters+ descriptive data All vocals	Only Mallampati 1 and 4 cases	189/81	H2O RANDOM FOREST		0,17	0,14	
Voice parameters+ descriptive data A normal	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,36	0,10	0,49	
Voice parameters+ descriptive data O extension	Only Mallampati 1 and 4 cases	189/81	H2O GRADIENT BOOSTING		0,48	0,47	
Voice parameters+ descriptive data O flexion	Only Mallampati 1 and 4 cases	189/81	H2O GRADIENT BOOSTING	0,39	0,10	0,34	
Voice parameters+ descriptive data A flexion vs A extension	Only Mallampati 1 and 4 cases	189/81	H2O NAIVE BAYES	0,48	0,26	0,54	
Voice parameters+ descriptive data E flexion vs E extension	Only Mallampati 1 and 4 cases	189/81	H2O GENERALISED LINEAR MODEL	0,26	0,11	0,56	
Voice parameters+ descriptive data O flexion vs O extension	Only Mallampati 1 and 4 cases	189/81	H2O GRADIENT BOOSTING	0,15	0,10	0,54	

11.11. ANNEX 11. Full Table 10. Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in testing group. TN: True negatives, FN: False negatives, FP: False positives, TP: True positives.

TESTING							
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TN	FN	FP	TP
HARMONICS + Descriptive data All vocals		179	H2O GRADIENT BOOSTING	96	10	45	28
HARMONICS + Descriptive data A normal (_1)		179	H2O NAIVE BAYES	97	12	44	26
HARMONICS + Descriptive data A extension (_2)		179	H2O GENERALISED LINEAR MODEL	90	10	51	28
HARMONICS + Descriptive data A flexion (_3)		179	H2O GENERALISED LINEAR MODEL	108	14	33	24
HARMONICS + Descriptive data E normal (_1)		179	H2O GENERALISED LINEAR MODEL	105	13	36	25
HARMONICS + Descriptive data E extension (_2)		179	H2O GRADIENT BOOSTING	106	14	35	24
HARMONICS + Descriptive data E flexion(_3)		179	H2O GENERALISED LINEAR MODEL	111	14	30	24
HARMONICS + Descriptive data I normal (_1)		179	H2O GENERALISED LINEAR MODEL	108	14	33	24
HARMONICS + Descriptive data I extension (_2)		179	H2O GENERALISED LINEAR MODEL	111	14	30	24
HARMONICS + Descriptive data I flexion(_3)		179	H2O GENERALISED LINEAR MODEL	110	15	31	23
HARMONICS + Descriptive data O normal (_1)		179	H2O GENERALISED LINEAR MODEL	104	13	37	25

TESTING							
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TN	FN	FP	TP
HARMONICS + Descriptive data O extension (_2)		179	H2O GENERALISED LINEAR MODEL	113	15	28	23
HARMONICS + Descriptive data O flexion(_3)		179	H2O GENERALISED LINEAR MODEL	103	13	38	25
HARMONICS + Descriptive data U normal (_1)		179	H2O GENERALISED LINEAR MODEL	109	14	32	24
HARMONICS + Descriptive data U extension 2)		179	H2O GENERALISED LINEAR MODEL	105	13	36	25
HARMONICS + Descriptive data U Flexion (_3)		179	H2O GENERALISED LINEAR MODEL	108	14	33	24
HARMONICS + Descriptive data A normal (_1) vs A flexion (_3)		179	H2O NAIVE BAYES	102	13	39	25
HARMONICS + Descriptive data A normal (_1) vs A extension (_2)		179	H2O NAIVE BAYES	94	11	47	27
HARMONICS + Descriptive data A flexion (_3) vs A extension (_2)		179	H2O GRADIENT BOOSTING	99	12	42	26
HARMONICS + Descriptive data A all positions		179	H2O GRADIENT BOOSTING	69	6	72	32
HARMONICS + Descriptive data E normal (_1) vs E flexion (_3)		179	H2O GRADIENT BOOSTING	110	14	31	24
HARMONICS + Descriptive data E normal (_1) vs E extension(_2)		179	H2O GENERALISED LINEAR MODEL	107	14	34	24
HARMONICS + Descriptive data E flexion (_2) vs E extension (_3)		179	H2O GENERALISED LINEAR MODEL	104	13	37	25
HARMONICS + Descriptive data E all positions		179	H2O GENERALISED LINEAR MODEL	104	13	37	25
HARMONICS + Descriptive data I normal (_1) vs I flexion (_3)		179	H2O GENERALISED LINEAR MODEL	109	14	32	24

TESTING							
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TN	FN	FP	TP
HARMONICS + Descriptive data I normal (_1) vs I extension(_2)		179	H2O GENERALISED LINEAR MODEL	109	14	32	24
HARMONICS + Descriptive data I flexion (_3) vs I extension (_2)		179	H2O GENERALISED LINEAR MODEL	111	14	30	24
HARMONICS + Descriptive data I all positions		179	H2O GENERALISED LINEAR MODEL	111	14	30	24
HARMONICS + Descriptive data O normal (_1) vs O flexion (_3)		179	H2O GENERALISED LINEAR MODEL	104	12	37	26
HARMONICS + Descriptive data O normal (_1) vs O extension (_2)		179	H2O NAIVE BAYES	100	12	41	26
HARMONICS + Descriptive data O flexion(_3) vs O extension (_2)		179	H2O GENERALISED LINEAR MODEL	102	12	39	26
HARMONICS + Descriptive data O all positions		179	H2O GENERALISED LINEAR MODEL	108	14	33	24
HARMONICS + Descriptive data U normal (_1) vs U flexion (_3)		179	H2O GENERALISED LINEAR MODEL	108	14	33	24
HARMONICS + Descriptive data U normal (_1) vs U extension (_2)		179	H2O GENERALISED LINEAR MODEL	107	13	34	25
HARMONICS + Descriptive data U flexion (_3) vs U extension (_2)		179	H2O GENERALISED LINEAR MODEL	113	15	28	23
HARMONICS + Descriptive data U all positions		179	H2O GENERALISED LINEAR MODEL	111	15	30	23
HARMONICS + Descriptive data All vocals	Only Mallampati 1 and 4 cases	81	H2O NAIVE BAYES	73	0	4	4
HARMONICS + Descriptive data A normal (_1)	Only Mallampati 1 and 4 cases	81	H2O GENERALISED LINEAR MODEL	76	1	1	3
HARMONICS + Descriptive data A extension (_2)	Only Mallampati 1 and 4 cases	81	H2O NAIVE BAYES	77	1	0	3

TESTING							
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TN	FN	FP	TP
HARMONICS + Descriptive data A flexion (_3)	Only Mallampati 1 and 4 cases	81	H2O GENERALISED LINEAR MODEL	76	1	1	3
HARMONICS + Descriptive data E normal (_1)	Only Mallampati 1 and 4 cases	81	H2O GENERALISED LINEAR MODEL	77	1	0	3
HARMONICS + Descriptive data E extension (_2)	Only Mallampati 1 and 4 cases	81	H2O NAIVE BAYES	77	1	0	3
HARMONICS + Descriptive data E flexion(_3)	Only Mallampati 1 and 4 cases	81	H2O GENERALISED LINEAR MODEL	76	1	1	3
HARMONICS + Descriptive data I normal (_1)	Only Mallampati 1 and 4 cases	81	H2O GENERALISED LINEAR MODEL	77	1	0	3
HARMONICS + Descriptive data I extension (_2)	Only Mallampati 1 and 4 cases	81	H2O GENERALISED LINEAR MODEL	77	1	0	3
HARMONICS + Descriptive data I flexion(_3)	Only Mallampati 1 and 4 cases	81	H2O NAIVE BAYES	77	1	0	3
HARMONICS + Descriptive data O normal (_1)	Only Mallampati 1 and 4 cases	81	H2O NAIVE BAYES	76	0	1	4
HARMONICS + Descriptive data O extension (_2)	Only Mallampati 1 and 4 cases	81	H2O NAIVE BAYES	77	1	0	3
HARMONICS + Descriptive data O flexion(_3)	Only Mallampati 1 and 4 cases	81	H2O NAIVE BAYES	76	1	1	3
HARMONICS + Descriptive data U normal (_1)	Only Mallampati 1 and 4 cases	81	H2O NAIVE BAYES	76	1	1	3
HARMONICS + Descriptive data U extension 2)	Only Mallampati 1 and 4 cases	81	H2O NAIVE BAYES	76	1	1	3
HARMONICS + Descriptive data U Flexion (_3)	Only Mallampati 1 and 4 cases	81	H2O NAIVE BAYES	77	1	0	3

TESTING							
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TN	FN	FP	TP
HARMONICS + Descriptive data A normal (_1) vs A flexion (_3)	Only Mallampati 1 and 4 cases	81	H2O NAIVE BAYES	76	1	1	3
HARMONICS + Descriptive data A normal (_1) vs A extension (_2)	Only Mallampati 1 and 4 cases	81	H2O GENERALISED LINEAR MODEL	77	1	0	3
HARMONICS + Descriptive data A flexion (_3) vs A extension (_2)	Only Mallampati 1 and 4 cases	81	H2O GENERALISED LINEAR MODEL	77	1	0	3
HARMONICS + Descriptive data A all positions	Only Mallampati 1 and 4 cases	81	H2O NAIVE BAYES	73	0	4	4
HARMONICS + Descriptive data E normal (_1) vs E flexion (_3)	Only Mallampati 1 and 4 cases	81	H2O GENERALISED LINEAR MODEL	76	1	1	3
HARMONICS + Descriptive data E normal (_1) vs E extension(_2)	Only Mallampati 1 and 4 cases	81	H2O GENERALISED LINEAR MODEL	77	1	0	3
HARMONICS + Descriptive data E flexion (_2) vs E extension (_3)	Only Mallampati 1 and 4 cases	81	H2O GENERALISED LINEAR MODEL	77	1	0	3
HARMONICS + Descriptive data E all positions	Only Mallampati 1 and 4 cases	81	H2O GENERALISED LINEAR MODEL	76	1	1	3
HARMONICS + Descriptive data I normal (_1) vs I flexion (_3)	Only Mallampati 1 and 4 cases	81	H2O GENERALISED LINEAR MODEL	77	1	0	3
HARMONICS + Descriptive data I normal (_1) vs I extension(_2)	Only Mallampati 1 and 4 cases	81	H2O NAIVE BAYES	74	0	3	4
HARMONICS + Descriptive data I flexion (_3) vs I extension (_2)	Only Mallampati 1 and 4 cases	81	H2O NAIVE BAYES	76	1	1	3
HARMONICS + Descriptive data I all positions	Only Mallampati 1 and 4 cases	81	H2O NAIVE BAYES	75	0	2	4
HARMONICS + Descriptive data O normal (_1) vs O flexion (_3)	Only Mallampati 1 and 4 cases	81	H2O NAIVE BAYES	72	0	5	4

TESTING							
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TN	FN	FP	TP
HARMONICS + Descriptive data O normal (_1) vs O extension (_2)	Only Mallampati 1 and 4 cases	81	H2O GENERALISED LINEAR MODEL	77	1	0	3
HARMONICS + Descriptive data O flexion(_3) vs O extension (_2)	Only Mallampati 1 and 4 cases	81	H2O GENERALISED LINEAR MODEL	77	1	0	3
HARMONICS + Descriptive data O all positions	Only Mallampati 1 and 4 cases	81	H2O GENERALISED LINEAR MODEL	76	1	1	3
HARMONICS + Descriptive data U normal (_1) vs U flexion (_3)	Only Mallampati 1 and 4 cases	81	H2O NAIVE BAYES	75	0	2	4
HARMONICS + Descriptive data U normal (_1) vs U extension (_2)	Only Mallampati 1 and 4 cases	81	H2O NAIVE BAYES	75	0	2	4
HARMONICS + Descriptive data U flexion (_3) vs U extension (_2)	Only Mallampati 1 and 4 cases	81	H2O GENERALISED LINEAR MODEL	76	1	1	3
HARMONICS + Descriptive data U all positions	Only Mallampati 1 and 4 cases	81	H2O NAIVE BAYES	76	1	1	3
Només DADES DESCRIPTIVES		179	H2O GENERALISED LINEAR MODEL	110	15	31	23
Només HARMÒNICS A normal (_1)		179	H2O GENERALISED LINEAR MODEL	23	1	118	37
Només HARMÒNICS A extensió (_2)		179	H2O RANDOM FOREST	0	0	141	38
Spetral + descriptive data All vocals		179	H2O GENERALISED LINEAR MODEL	103	13	38	25
Spectral+ descriptive data A normal		179	H2O GRADIENT BOOSTING	121	17	20	21
Spectral+ descriptive data O extension		179	H2O GRADIENT BOOSTING	75	9	66	29
Spectral+ descriptive data O flexion		179	H2O GENERALISED LINEAR MODEL	108	12	33	26

TESTING							
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TN	FN	FP	TP
Spectral+ descriptive data A flexion vs A extension		179	H2O GENERALISED LINEAR MODEL	107	13	34	25
Spectral + descriptive data E flexion vs E e xtension		179	H2O GRADIENT BOOSTING	87	11	54	27
Voice parameters+ descriptive data All vocals		179	H2O GRADIENT BOOSTING	100	14	41	24
Voice parameters+ descriptive data A normal		179	H2O GENERALISED LINEAR MODEL	107	13	34	25
Voice parameters+ descriptive data O extension		179	H2O GRADIENT BOOSTING	107	14	34	24
Voice parameters+ descriptive data O flexion		179	H2O NAIVE BAYES	96	8	45	30
Voice parameters+ descriptive data A flexion vs A extension		179	H2O NAIVE BAYES	95	14	46	24
Voice parameters+ descriptive data E flexion vs E extension		179	H2O NAIVE BAYES	75	7	66	31
Voice parameters+ descriptive data O flexion vs O extension		179	H2O NAIVE BAYES	92	9	49	29
Voice parameters+ descriptive data All vocals	Only Mallampati 1 and 4 cases	81	H2O RANDOM FOREST	75	2	2	2
Voice parameters+ descriptive data A normal	Only Mallampati 1 and 4 cases	81	H2O GENERALISED LINEAR MODEL	76	1	1	3
Voice parameters+ descriptive data O extension	Only Mallampati 1 and 4 cases	81	H2O GRADIENT BOOSTING	76	1	1	3
Voice parameters+ descriptive data O flexion	Only Mallampati 1 and 4 cases	81	H2O GRADIENT BOOSTING	76	1	1	3
Voice parameters+ descriptive data A flexion vs A extension	Only Mallampati 1 and 4 cases	81	H2O NAIVE BAYES	76	1	1	3
Voice parameters+ descriptive data E flexion vs E extension	Only Mallampati 1 and 4 cases	81	H2O GENERALISED LINEAR MODEL	76	1	1	3

TESTING							
DATASET (PARAMETERS COMBINATION)	CASES	N PATIENTS	BEST MODEL	TN	FN	FP	TP
Voice parameters+ descriptive data O flexion vs O extension	Only Mallampati 1 and 4 cases	81	H2O GRADIENT BOOSTING	76	1	1	3

11.12. ANNEX 12. Full table 13: Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in both training and testing groups, for accuracy values for the training group and the testing group respectively.

ACCURACY					
DATASET (PARAMETERS COMBINATION)	TYPE OF CASES	BEST MODEL	N PATIENTS	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data All vocals		H2O GENERALISED LINEAR MODEL	219/94	0,70	0,49
HARMONICS + Descriptive data A normal (_1)		H2O GENERALISED LINEAR MODEL	219/94	0,71	0,62
HARMONICS + Descriptive data A extension (_2)		H2O GRADIENT BOOSTING	219/94	0,81	0,69
HARMONICS + Descriptive data A flexion (_3)		H2O GRADIENT BOOSTING	219/94	0,80	0,62
HARMONICS + Descriptive data E normal (_1)		H2O GRADIENT BOOSTING	219/94	0,82	0,60
HARMONICS + Descriptive data E extension (_2)		H2O NAIVE BAYES	219/94	0,79	0,57
HARMONICS + Descriptive data E flexion (_3)		H2O NAIVE BAYES	219/94	0,77	0,60
HARMONICS + Descriptive data I normal (_1)		H2O RANDOM FOREST	219/94	0,76	0,61
HARMONICS + Descriptive data I extension (_2)		H2O RANDOM FOREST	219/94	0,78	0,60
HARMONICS + Descriptive data I flexion (_3)		H2O NAIVE BAYES	219/94	0,78	0,64
HARMONICS + Descriptive data O normal (_1)		H2O RANDOM FOREST	219/94	0,80	0,53
HARMONICS + Descriptive data O extension (_2)		H2O RANDOM FOREST	219/94	0,78	0,61
HARMONICS + Descriptive data O flexion (_3)		H2O NAIVE BAYES	219/94	0,76	0,52
HARMONICS + Descriptive data U normal (_1)		H2O RANDOM FOREST	219/94	0,82	0,64
HARMONICS + Descriptive data U extension 2)		H2O RANDOM FOREST	219/94	0,80	0,59

ACCURACY					
DATASET (PARAMETERS COMBINATION)	TYPE OF CASES	BEST MODEL	N PATIENTS	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data U Flexion (_3)		H2O RANDOM FOREST	219/94	0,74	0,60
HARMONICS + Descriptive data A normal (_1) vs A flexion (_3)		H2O RANDOM FOREST	219/94	0,75	0,69
HARMONICS + Descriptive data A normal (_1) vs A extension (_2)		H2O RANDOM FOREST	219/94	0,78	0,65
HARMONICS + Descriptive data A flexion (_3) vs A extension (_2)		H2O RANDOM FOREST	219/94	0,78	0,62
HARMONICS + Descriptive data A all positions		H2O RANDOM FOREST	219/94	0,79	0,55
HARMONICS + Descriptive data E normal (_1) vs E flexion (_3)		H2O NAIVE BAYES	219/94	0,78	0,59
HARMONICS + Descriptive data E normal (_1) vs E extension(_2)		H2O RANDOM FOREST	219/94	0,77	0,51
HARMONICS + Descriptive data E flexion (_2) vs E extension (_3)		H2O NAIVE BAYES	219/94	0,84	0,57
HARMONICS + Descriptive data E all positions		H2O NAIVE BAYES	219/94	0,78	0,59
HARMONICS + Descriptive data I normal (_1) vs I flexion (_3)		H2O GRADIENT BOOSTING	219/94	0,83	0,65
HARMONICS + Descriptive data I normal (_1) vs I extension(_2)		H2O GRADIENT BOOSTING	219/94	0,83	0,67
HARMONICS + Descriptive data I flexion (_3) vs I extension (_2)		H2O RANDOM FOREST	219/94	0,79	0,62
HARMONICS + Descriptive data I all positions		H2O NAIVE BAYES	219/94	0,75	0,57
HARMONICS + Descriptive data O normal (_1) vs O flexion (_3)		H2O GRADIENT BOOSTING	219/94	0,77	0,53
HARMONICS + Descriptive data O normal (_1) vs O extension (_2)		H2O GENERALISED LINEAR MODEL	219/94	0,67	0,68
HARMONICS + Descriptive data O flexion(_3) vs O extension (_2)		H2O GRADIENT BOOSTING	219/94	0,71	0,52
HARMONICS + Descriptive data O all positions		H2O GRADIENT BOOSTING	219/94	0,72	0,57
HARMONICS + Descriptive data U normal (_1) vs U flexion (_3)		H2O RANDOM FOREST	219/94	0,63	0,51

ACCURACY					
DATASET (PARAMETERS COMBINATION)	TYPE OF CASES	BEST MODEL	N PATIENTS	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data U normal (_1) vs U extension (_2)		H2O RANDOM FOREST	219/94	0,61	0,57
HARMONICS + Descriptive data U flexion (_3) vs U extension (_2)		H2O NAIVE BAYES	219/94	0,73	0,65
HARMONICS + Descriptive data U all positions		H2O NAIVE BAYES	219/94	0,75	0,60
HARMONICS+ descriptive data all vocals	Overweight Cormack 4	H2O RANDOM FOREST	219/94	0,74	0,47
HARMONICS + Descriptive data A normal (_1)	Overweight Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,65	0,50
HARMONICS + Descriptive data A extension (_2)	Overweight Cormack 4	H2O GRADIENT BOOSTING	219/94	0,76	0,56
HARMONICS + Descriptive data A flexion (_3)	Overweight Cormack 4	H2O RANDOM FOREST	219/94	0,66	0,46
HARMONICS + Descriptive data E normal (_1)	Overweight Cormack 4	H2O RANDOM FOREST	219/94	0,70	0,49
HARMONICS + Descriptive data E extension (_2)	Overweight Cormack 4		219/94		
HARMONICS + Descriptive data E flexion(_3)	Overweight Cormack 4		219/94		
HARMONICS + Descriptive data I normal (_1)	Overweight Cormack 4	H2O GRADIENT BOOSTING	219/94	0,75	0,46
HARMONICS + Descriptive data I extension (_2)	Overweight Cormack 4	H2O GRADIENT BOOSTING	219/94	0,79	0,36
HARMONICS + Descriptive data I flexion(_3)	Overweight Cormack 4		219/94		
HARMONICS + Descriptive data O normal (_1)	Overweight Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,68	0,52
HARMONICS + Descriptive data O extension (_2)	Overweight Cormack 4	H2O GRADIENT BOOSTING	219/94	0,76	0,38
HARMONICS + Descriptive data O flexion(_3)	Overweight Cormack 4	H2O GRADIENT BOOSTING	219/94	0,85	0,44
HARMONICS + Descriptive data U normal (_1)	Overweight Cormack 4	H2O NAIVE BAYES	219/94	0,73	0,46
HARMONICS + Descriptive data U extension 2)	Overweight Cormack 4	H2O RANDOM FOREST	219/94	0,78	0,33

ACCURACY					
DATASET (PARAMETERS COMBINATION)	TYPE OF CASES	BEST MODEL	N PATIENTS	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data U Flexion (_3)	Overweight Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,64	0,48
HARMONICS + Descriptive data A normal (_1) vs A flexion (_3)	Overweight Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,69	0,48
HARMONICS + Descriptive data A normal (_1) vs A extension (_2)	Overweight Cormack 4	H2O NAIVE BAYES	219/94	0,69	0,49
HARMONICS + Descriptive data A flexion (_3) vs A extension (_2)	Overweight Cormack 4	H2O NAIVE BAYES	219/94	0,75	0,45
HARMONICS + Descriptive data A all positions	Overweight Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,63	0,48
HARMONICS + Descriptive data E normal (_1) vs E flexion (_3)	Overweight Cormack 4	H2O NAIVE BAYES	219/94	0,80	0,51
HARMONICS + Descriptive data E normal (_1) vs E extension (_2)	Overweight Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,72	0,63
HARMONICS + Descriptive data E flexion (_2) vs E extension (_3)	Overweight Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,69	0,49
HARMONICS + Descriptive data E all positions	Overweight Cormack 4	H2O NAIVE BAYES	219/94	0,75	0,61
HARMONICS + Descriptive data I normal (_1) vs I flexion (_3)	Overweight Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,67	0,49
HARMONICS + Descriptive data I normal (_1) vs I extension (_2)	Overweight Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,65	0,49
HARMONICS + Descriptive data I all positions	Overweight Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,66	0,49
HARMONICS + Descriptive data U normal (_1) vs U flexion (_3)	Overweight Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,66	0,48
HARMONICS + Descriptive data U normal (_1) vs U extension (_2)	Overweight Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,66	0,48
HARMONICS + Descriptive data U flexion (_3) vs U extension (_2)	Overweight Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,68	0,46
HARMONICS + Descriptive data U all positions	Overweight Cormack 4	H2O GRADIENT BOOSTING	219/94	0,68	0,44
HARMONICS+ descriptive data all vocals	Cormack 1 and 4 cases	H2O GRADIENT BOOSTING	100/43	0,84	0,95
HARMONICS + Descriptive data A normal (_1)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,60	0,84

ACCURACY					
DATASET (PARAMETERS COMBINATION)	TYPE OF CASES	BEST MODEL	N PATIENTS	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data A extension (_2)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,87	0,81
HARMONICS + Descriptive data A flexion (_3)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,62	0,81
HARMONICS + Descriptive data E normal (_1)	Cormack 1 and 4 cases	H2O RANDOM FOREST	100/43	0,77	0,88
HARMONICS + Descriptive data E extension (_2)	Cormack 1 and 4 cases	H2O GRADIENT BOOSTING	100/43	0,84	0,86
HARMONICS + Descriptive data E flexion (_3)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,91	0,79
HARMONICS + Descriptive data I normal (_1)	Cormack 1 and 4 cases	H2O RANDOM FOREST	100/43	0,75	0,84
HARMONICS + Descriptive data I extension (_2)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,92	0,77
HARMONICS + Descriptive data I flexion (_3)	Cormack 1 and 4 cases	H2O GRADIENT BOOSTING	100/43	0,82	0,93
HARMONICS + Descriptive data O normal (_1)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,88	0,86
HARMONICS + Descriptive data O extension (_2)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,70	0,79
HARMONICS + Descriptive data O flexion (_3)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,90	0,77
HARMONICS + Descriptive data U normal (_1)	Cormack 1 and 4 cases	H2O GRADIENT BOOSTING	100/43	0,71	0,88
HARMONICS + Descriptive data U extension 2)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,87	0,84
HARMONICS + Descriptive data U Flexion (_3)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,91	0,77
HARMONICS + Descriptive data A normal (_1) vs A flexion (_3)	Cormack 1 and 4 cases	H2O RANDOM FOREST	100/43	0,83	0,81
HARMONICS + Descriptive data A normal (_1) vs A extension (_2)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,88	0,81
HARMONICS + Descriptive data A flexion (_3) vs A extension (_2)	Cormack 1 and 4 cases	H2O GRADIENT BOOSTING	100/43	0,77	0,93
HARMONICS + Descriptive data A all positions	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,83	0,93

ACCURACY					
DATASET (PARAMETERS COMBINATION)	TYPE OF CASES	BEST MODEL	N PATIENTS	TRAINING 70%	TESTING 30%
HARMONICS + Descriptive data E normal (_1) vs E flexion (_3)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,78	0,84
HARMONICS + Descriptive data E normal (_1) vs E extension(_2)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,92	0,77
HARMONICS + Descriptive data E flexion (_2) vs E extension (_3)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,66	0,79
HARMONICS + Descriptive data E all positions	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,95	0,81
HARMONICS + Descriptive data I normal (_1) vs I flexion (_3)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,90	0,79
HARMONICS + Descriptive data I normal (_1) vs I extension(_2)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,84	0,77
HARMONICS + Descriptive data I flexion (_3) vs I extension (_2)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,91	0,77
HARMONICS + Descriptive data I all positions	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,69	0,79
HARMONICS + Descriptive data O normal (_1) vs O flexion (_3)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,72	0,81
HARMONICS + Descriptive data O normal (_1) vs O extension (_2)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,68	0,77
HARMONICS + Descriptive data O flexion(_3) vs O extension (_2)	Cormack 1 and 4 cases	H2O RANDOM FOREST	100/43	0,71	0,72
HARMONICS + Descriptive data O all positions	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,93	0,79
HARMONICS + Descriptive data U normal (_1) vs U flexion (_3)	Cormack 1 and 4 cases	H2O RANDOM FOREST	100/43	0,81	0,84
HARMONICS + Descriptive data U normal (_1) vs U extension (_2)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,86	0,84
HARMONICS + Descriptive data U flexion (_3) vs U extension (_2)	Cormack 1 and 4 cases	H2O RANDOM FOREST	100/43	0,81	0,79
HARMONICS + Descriptive data U all positions	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,93	0,81
ONLY descriptive data		H2O RANDOM FOREST	219/94	0,69	0,61
HARMONICS All vocals		H2O NAIVE BAYES	219/94	0,42	0,50

ACCURACY					
DATASET (PARAMETERS COMBINATION)	TYPE OF CASES	BEST MODEL	N PATIENTS	TRAINING 70%	TESTING 30%
HARMONICS A normal (_1)		H2O NAIVE BAYES	219/94	0,28	0,62
HARMONICS A extension (_2)		H2O GENERALISED LINEAR MODEL	219/94	0,56	0,34
HARMONICS A flexion (_3)		H2O NAIVE BAYES	219/94	0,26	0,70
HARMONICS E normal (_1)		H2O GENERALISED LINEAR MODEL	219/94	0,57	0,66
HARMONICS E extension (_2)		H2O RANDOM FOREST	219/94	0,52	0,55
HARMONICS E flexion(_3)		H2O GRADIENT BOOSTING	219/94	0,76	0,60
HARMONICS I normal (_1)		H2O NAIVE BAYES	219/94	0,57	0,55
HARMONICS I extension (_2)		H2O GRADIENT BOOSTING	219/94	0,58	0,52
HARMONICS I flexion(_3)		H2O NAIVE BAYES	219/94	0,47	0,55
HARMONICS O normal (_1)		H2O GRADIENT BOOSTING	219/94	0,56	0,56
HARMONICS O extension (_2)		H2O GENERALISED LINEAR MODEL	219/94	0,49	0,32
HARMONICS O flexion(_3)		H2O GRADIENT BOOSTING	219/94	0,58	0,49
HARMONICS U normal (_1)		H2O NAIVE BAYES	219/94	0,56	0,56
HARMONICS U extension 2)		H2O NAIVE BAYES	219/94	0,52	0,34
HARMONICS U Flexion (_3)		H2O RANDOM FOREST	219/94	0,37	0,43
HARMONICS A normal (_1) vs A flexion (_3)		H2O GENERALISED LINEAR MODEL	219/94	0,31	0,46
HARMONICS A normal (_1) vs A extension (_2)		H2O NAIVE BAYES	219/94	0,31	0,60
HARMONICS A flexion (_3) vs A extension (_2)		H2O NAIVE BAYES	219/94	0,50	0,69
HARMONICS A all positions		H2O NAIVE BAYES	219/94	0,40	0,72
HARMONICS E normal (_1) vs E flexion (_3)		H2O GENERALISED LINEAR MODEL	219/94	0,38	0,77
HARMONICS E normal (_1) vs E extension(_2)		H2O GENERALISED LINEAR MODEL	219/94	0,60	0,44

ACCURACY					
DATASET (PARAMETERS COMBINATION)	TYPE OF CASES	BEST MODEL	N PATIENTS	TRAINING 70%	TESTING 30%
HARMONICS E flexion (_2) vs E extension (_3)		H2O GRADIENT BOOSTING	219/94	0,66	0,57
HARMONICS E all positions		H2O NAIVE BAYES	219/94	0,65	0,65
HARMONICS I normal (_1) vs I flexion (_3)		H2O NAIVE BAYES	219/94	0,45	0,51
HARMONICS I normal (_1) vs I extension(_2)		H2O NAIVE BAYES	219/94	0,49	0,65
HARMONICS I flexion (_3) vs I extension (_2)		H2O GRADIENT BOOSTING	219/94	0,70	0,47
HARMONICS I all positions		H2O NAIVE BAYES	219/94	0,43	0,61
HARMONICS O normal (_1) vs O flexion (_3)		H2O RANDOM FOREST	219/94	0,57	0,29
HARMONICS O normal (_1) vs O extension (_2)		H2O NAIVE BAYES	219/94	0,49	0,44
HARMONICS O flexion(_3) vs O extension (_2)		H2O GRADIENT BOOSTING	219/94	0,71	0,67
HARMONICS O all positions		H2O RANDOM FOREST	219/94	0,58	0,74
HARMONICS U normal (_1) vs U flexion (_3)		H2O NAIVE BAYES	219/94	0,31	0,52
HARMONICS U normal (_1) vs U extension (_2)		H2O GRADIENT BOOSTING	219/94	0,58	0,40
HARMONICS U flexion (_3) vs U extension (_2)		H2O RANDOM FOREST	219/94	0,28	0,53
HARMONICS U all positions		H2O RANDOM FOREST	219/94	0,24	0,46
Spectral+ Descriptive data all vocals		H2O GRADIENT BOOSTING	219/94	0,66	0,64
SPectral+ Descriptive data A normal		H2O RANDOM FOREST	219/94	0,72	0,66
SPectral+ Descriptive data O extension		H2O RANDOM FOREST	219/94	0,71	0,61
SPectral+ Descriptive data O flexion		H2O NAIVE BAYES	219/94	0,78	0,66
Spectral+ Descriptive data A flexion vs A extension		H2O RANDOM FOREST	219/94	0,66	0,61

ACCURACY					
DATASET (PARAMETERS COMBINATION)	TYPE OF CASES	BEST MODEL	N PATIENTS	TRAINING 70%	TESTING 30%
SPectral+ Descriptive data E flexion vs E extension		H2O NAIVE BAYES	219/94	0,72	0,65
Voice parameters + Descriptive data all vocals		H2O GENERALISED LINEAR MODEL	219/94	0,65	0,45
Voice parameters + Descriptive data A normal		H2O GENERALISED LINEAR MODEL	219/94	0,75	0,65
Voice parameters + Descriptive data O extension		H2O GRADIENT BOOSTING	219/94	0,74	0,64
Voice parameters + Descriptive data O flexion		H2O RANDOM FOREST	219/94	0,74	0,71
Voice parameters + Descriptive data A flexion vs A extension		H2O GENERALISED LINEAR MODEL	219/94	0,74	0,54
Voice parameters + Descriptive data E flexion vs E extension		H2O NAIVE BAYES	219/94	0,78	0,51
Voice parameters + Descriptive data O flexion vs O extension		H2O NAIVE BAYES	219/94	0,81	0,51
Voice parameters + Descriptive data all vocals	Only Cormack 1 and 4	H2O RANDOM FOREST	100/43	0,80	0,86
Voice parameters + Descriptive data A normal	Only Cormack 1 and 4	H2O GENERALISED LINEAR MODEL	100/43	0,79	0,81
Voice parameters + Descriptive data O extension	Only Cormack 1 and 4	H2O GENERALISED LINEAR MODEL	100/43	0,73	0,84
Voice parameters + Descriptive data O flexion	Only Cormack 1 and 4	H2O NAIVE BAYES	100/43	0,93	0,81
Voice parameters + Descriptive data A flexion vs A extension	Only Cormack 1 and 4	H2O GENERALISED LINEAR MODEL	100/43	0,81	0,84
Voice parameters + Descriptive data E flexion vs E extension	Only Cormack 1 and 4	H2O GRADIENT BOOSTING	100/43	0,71	0,91
Voice parameters + Descriptive data O flexion vs O extension	Only Cormack 1 and 4	H2O NAIVE BAYES	100/43	0,93	0,81

11.13. ANNEX 13. Full table 14 Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in both training and testing groups, for precision and recall values for the training group and the testing group respectively.

DATASET (PARAMETERS COMBINATION)	TYPE OF CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				Precision	Recall	Precision	Recall
HARMONICS + Descriptive data All vocals		H2O GENERALISED LINEAR MODEL	219/94	0,40	0,68	0,27	0,89
HARMONICS + Descriptive data A normal (_1)		H2O GENERALISED LINEAR MODEL	219/94	0,42	0,69	0,32	0,79
HARMONICS + Descriptive data A extension (_2)		H2O GRADIENT BOOSTING	219/94	0,62	0,51	0,37	0,74
HARMONICS + Descriptive data A flexion (_3)		H2O GRADIENT BOOSTING	219/94	0,54	0,56	0,33	0,89
HARMONICS + Descriptive data E normal (_1)		H2O GRADIENT BOOSTING	219/94	0,59	0,53	0,33	0,95
HARMONICS + Descriptive data E extension (_2)		H2O NAIVE BAYES	219/94	0,50	0,66	0,31	0,89
HARMONICS + Descriptive data E flexion (_3)		H2O NAIVE BAYES	219/94	0,48	0,66	0,31	0,79
HARMONICS + Descriptive data I normal (_1)		H2O RANDOM FOREST	219/94	0,50	0,86	0,33	0,95
HARMONICS + Descriptive data I extension (_2)		H2O RANDOM FOREST	219/94	0,50	0,79	0,32	0,89
HARMONICS + Descriptive data I flexion (_3)		H2O NAIVE BAYES	219/94	0,47	0,66	0,33	0,79
HARMONICS + Descriptive data O normal (_1)		H2O RANDOM FOREST	219/94	0,53	0,79	0,30	0,95

DATASET (PARAMETERS COMBINATION)	TYPE OF CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				Precision	Recall	Precision	Recall
HARMONICS + Descriptive data O extension (_2)		H2O RANDOM FOREST	219/94	0,51	0,72	0,32	0,84
HARMONICS + Descriptive data O flexion(_3)		H2O NAIVE BAYES	219/94	0,46	0,75	0,29	0,95
HARMONICS + Descriptive data U normal (_1)		H2O RANDOM FOREST	219/94	0,53	0,73	0,35	0,95
HARMONICS + Descriptive data U extension 2)		H2O RANDOM FOREST	219/94	0,57	0,75	0,32	0,95
HARMONICS + Descriptive data U Flexion (_3)		H2O RANDOM FOREST	219/94	0,45	0,78	0,31	0,84
HARMONICS + Descriptive data A normal (_1) vs A flexion (_3)		H2O RANDOM FOREST	219/94	0,44	0,78	0,37	0,74
HARMONICS + Descriptive data A normal (_1) vs A extension (_2)		H2O RANDOM FOREST	219/94	0,64	0,68	0,34	0,79
HARMONICS + Descriptive data A flexion (_3) vs A extension (_2)		H2O RANDOM FOREST	219/94	0,50	0,73	0,33	0,89
HARMONICS + Descriptive data A all positions		H2O RANDOM FOREST	219/94	0,62	0,75	0,31	0,95
HARMONICS + Descriptive data E normal (_1) vs E flexion (_3)		H2O NAIVE BAYES	219/94	0,49	0,69	0,31	0,89
HARMONICS + Descriptive data E normal (_1) vs E extension(_2)		H2O RANDOM FOREST	219/94	0,48	0,85	0,29	0,95

DATASET (PARAMETERS COMBINATION)	TYPE OF CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				Precision	Recall	Precision	Recall
HARMONICS + Descriptive data E flexion (_2) vs E extension (_3)		H2O NAIVE BAYES	219/94	0,58	0,65	0,30	0,84
HARMONICS + Descriptive data E all positions		H2O NAIVE BAYES	219/94	0,47	0,74	0,31	0,84
HARMONICS + Descriptive data I normal (_1) vs I flexion (_3)		H2O GRADIENT BOOSTING	219/94	0,70	0,48	0,36	0,95
HARMONICS + Descriptive data I normal (_1) vs I extension(_2)		H2O GRADIENT BOOSTING	219/94	0,59	0,51	0,36	0,79
HARMONICS + Descriptive data I flexion (_3) vs I extension (_2)		H2O RANDOM FOREST	219/94	0,52	0,79	0,33	0,89
HARMONICS + Descriptive data I all positions		H2O NAIVE BAYES	219/94	0,43	0,72	0,31	0,89
HARMONICS + Descriptive data O normal (_1) vs O flexion (_3)		H2O GRADIENT BOOSTING	219/94	0,50	0,47	0,30	0,95
HARMONICS + Descriptive data O normal (_1) vs O extension (_2)		H2O GENERALISED LINEAR MODEL	219/94	0,37	0,85	0,36	0,74
HARMONICS + Descriptive data O flexion(_3) vs O extension (_2)		H2O GRADIENT BOOSTING	219/94	0,37	0,53	0,28	0,89
HARMONICS + Descriptive data O all positions		H2O GRADIENT BOOSTING	219/94	0,60	0,49	0,31	0,89
HARMONICS + Descriptive data U normal (_1) vs U flexion (_3)		H2O RANDOM FOREST	219/94	0,38	0,78	0,29	0,95

DATASET (PARAMETERS COMBINATION)	TYPE OF CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				Precision	Recall	Precision	Recall
HARMONICS + Descriptive data U normal (_1) vs U extension (_2)		H2O RANDOM FOREST	219/94	0,36	0,88	0,29	0,79
HARMONICS + Descriptive data U flexion (_3) vs U extension (_2)		H2O NAIVE BAYES	219/94	0,42	0,71	0,31	0,58
HARMONICS + Descriptive data U all positions		H2O NAIVE BAYES	219/94	0,44	0,67	0,30	0,74
HARMONICS+ decriptive data all vocals	Overweight Cormack 4	H2O RANDOM FOREST	219/94	0,43	0,65	0,25	0,84
HARMONICS + Descriptive data A normal (_1)	Overweight Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,34	0,75	0,28	0,95
HARMONICS + Descriptive data A extension (_2)	Overweight Cormack 4	H2O GRADIENT BOOSTING	219/94	0,50	0,52	0,30	0,84
HARMONICS + Descriptive data A flexion (_3)	Overweight Cormack 4	H2O RANDOM FOREST	219/94	0,38	0,87	0,25	0,84
HARMONICS + Descriptive data E normal (_1)	Overweight Cormack 4	H2O RANDOM FOREST	219/94	0,39	0,80	0,26	0,84
HARMONICS + Descriptive data I normal (_1)	Overweight Cormack 4	H2O GRADIENT BOOSTING	219/94	0,47	0,55	0,24	0,79
HARMONICS + Descriptive data I extension (_2)	Overweight Cormack 4	H2O GRADIENT BOOSTING	219/94	0,52	0,42	0,23	0,89
HARMONICS + Descriptive data O normal (_1)	Overweight Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,38	0,83	0,29	0,95
HARMONICS + Descriptive data O extension (_2)	Overweight Cormack 4	H2O GRADIENT BOOSTING	219/94	0,48	0,60	0,24	0,95
HARMONICS + Descriptive data O flexion(_3)	Overweight Cormack 4	H2O GRADIENT BOOSTING	219/94	0,70	0,53	0,23	0,79

DATASET (PARAMETERS COMBINATION)	TYPE OF CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				Precision	Recall	Precision	Recall
HARMONICS + Descriptive data U normal (_1)	Overweight Cormack 4	H2O NAIVE BAYES	219/94	0,46	0,70	0,26	0,95
HARMONICS + Descriptive data U extension 2)	Overweight Cormack 4	H2O RANDOM FOREST	219/94	0,52	0,58	0,23	0,95
HARMONICS + Descriptive data U Flexion (_3)	Overweight Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,35	0,87	0,28	1,00
HARMONICS + Descriptive data A normal (_1) vs A flexion (_3)	Overweight Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,45	0,66	0,28	1,00
HARMONICS + Descriptive data A normal (_1) vs A extension (_2)	Overweight Cormack 4	H2O NAIVE BAYES	219/94	0,41	0,77	0,28	0,95
HARMONICS + Descriptive data A flexion (_3) vs A extension (_2)	Overweight Cormack 4	H2O NAIVE BAYES	219/94	0,43	0,62	0,25	0,89
HARMONICS + Descriptive data A all positions	Overweight Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,34	0,81	0,28	1,00
HARMONICS + Descriptive data E normal (_1) vs E flexion (_3)	Overweight Cormack 4	H2O NAIVE BAYES	219/94	0,52	0,59	0,28	0,89
HARMONICS + Descriptive data E normal (_1) vs E extension (_2)	Overweight Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,39	0,67	0,32	0,74
HARMONICS + Descriptive data E flexion (_2) vs E extension (_3)	Overweight Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,37	0,70	0,28	0,95
HARMONICS + Descriptive data E all positions	Overweight Cormack 4	H2O NAIVE BAYES	219/94	0,47	0,69	0,30	0,68

DATASET (PARAMETERS COMBINATION)	TYPE OF CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				Precision	Recall	Precision	Recall
HARMONICS + Descriptive data I normal (_1) vs I flexion (_3)	Overweight Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,36	0,78	0,28	1,00
HARMONICS + Descriptive data I normal (_1) vs I extension (_2)	Overweight Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,36	0,89	0,28	1,00
HARMONICS + Descriptive data I all positions	Overweight Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,36	0,89	0,28	1,00
HARMONICS + Descriptive data U normal (_1) vs U flexion (_3)	Overweight Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,36	0,86	0,28	1,00
HARMONICS + Descriptive data U normal (_1) vs U extension (_2)	Overweight Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,36	0,85	0,28	1,00
HARMONICS + Descriptive data U flexion (_3) vs U extension (_2)	Overweight Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,38	0,79	0,27	1,00
HARMONICS + Descriptive data U all positions	Overweight Cormack 4	H2O GRADIENT BOOSTING	219/94	0,49	0,59	0,25	0,89
HARMONICS+ decriptive data all vocals	Cormack 1 and 4 cases	H2O GRADIENT BOOSTING	100/43	0,53	0,61	1,00	0,60
HARMONICS + Descriptive data A normal (_1)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,26	0,85	0,40	0,80
HARMONICS + Descriptive data A extension (_2)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,65	0,56	0,36	0,80
HARMONICS + Descriptive data A flexion (_3)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,29	0,92	0,36	0,80
HARMONICS + Descriptive data E normal (_1)	Cormack 1 and 4 cases	H2O RANDOM FOREST	100/43	0,47	0,80	0,50	0,80

DATASET (PARAMETERS COMBINATION)	TYPE OF CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				Precision	Recall	Precision	Recall
HARMONICS + Descriptive data E extension (_2)	Cormack 1 and 4 cases	H2O GRADIENT BOOSTING	100/43	0,53	0,50	0,44	0,80
HARMONICS + Descriptive data E flexion(_3)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,77	0,76	0,33	0,80
HARMONICS + Descriptive data I normal (_1)	Cormack 1 and 4 cases	H2O RANDOM FOREST	100/43	0,47	0,84	0,40	0,80
HARMONICS + Descriptive data I extension (_2)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,88	0,68	0,31	0,80
HARMONICS + Descriptive data I flexion(_3)	Cormack 1 and 4 cases	H2O GRADIENT BOOSTING	100/43	0,46	0,55	0,67	0,80
HARMONICS + Descriptive data O normal (_1)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,60	0,69	0,44	0,80
HARMONICS + Descriptive data O extension (_2)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,31	0,95	0,33	0,80
HARMONICS + Descriptive data O flexion(_3)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,76	0,76	0,31	0,80
HARMONICS + Descriptive data U normal (_1)	Cormack 1 and 4 cases	H2O GRADIENT BOOSTING	100/43	0,34	0,56	0,50	0,80
HARMONICS + Descriptive data U extension 2)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,62	0,69	0,40	0,80
HARMONICS + Descriptive data U Flexion (_3)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,77	0,71	0,31	0,80
HARMONICS + Descriptive data A normal (_1) vs A flexion (_3)	Cormack 1 and 4 cases	H2O RANDOM FOREST	100/43	0,59	0,69	0,36	0,80
HARMONICS + Descriptive data A normal (_1) vs A extension (_2)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,69	0,73	0,36	0,80

DATASET (PARAMETERS COMBINATION)	TYPE OF CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				Precision	Recall	Precision	Recall
HARMONICS + Descriptive data A flexion (_3) vs A extension (_2)	Cormack 1 and 4 cases	H2O GRADIENT BOOSTING	100/43	0,67	0,61	0,67	0,80
HARMONICS + Descriptive data A all positions	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,39	0,62	0,75	0,60
HARMONICS + Descriptive data E normal (_1) vs E flexion (_3)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,38	0,62	0,40	0,80
HARMONICS + Descriptive data E normal (_1) vs E extension (_2)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,79	0,74	0,31	0,80
HARMONICS + Descriptive data E flexion (_2) vs E extension (_3)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,33	0,79	0,33	0,80
HARMONICS + Descriptive data E all positions	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,88	0,78	0,36	0,80
HARMONICS + Descriptive data I normal (_1) vs I flexion (_3)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,67	0,69	0,33	0,80
HARMONICS + Descriptive data I normal (_1) vs I extension (_2)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,61	0,82	0,31	0,80
HARMONICS + Descriptive data I flexion (_3) vs I extension (_2)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,79	0,68	0,31	0,80
HARMONICS + Descriptive data I all positions	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,29	0,78	0,33	0,80
HARMONICS + Descriptive data O normal (_1) vs O flexion (_3)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,32	0,79	0,36	0,80

DATASET (PARAMETERS COMBINATION)	TYPE OF CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				Precision	Recall	Precision	Recall
HARMONICS + Descriptive data O normal (_1) vs O extension (_2)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,32	0,84	0,31	0,80
HARMONICS + Descriptive data O flexion(_3) vs O extension (_2)	Cormack 1 and 4 cases	H2O RANDOM FOREST	100/43	0,49	0,88	0,27	0,80
HARMONICS + Descriptive data O all positions	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,75	0,83	0,33	0,80
HARMONICS + Descriptive data U normal (_1) vs U flexion (_3)	Cormack 1 and 4 cases	H2O RANDOM FOREST	100/43	0,52	0,71	0,40	0,80
HARMONICS + Descriptive data U normal (_1) vs U extension (_2)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,59	0,78	0,40	0,80
HARMONICS + Descriptive data U flexion (_3) vs U extension (_2)	Cormack 1 and 4 cases	H2O RANDOM FOREST	100/43	0,40	0,89	0,33	0,80
HARMONICS + Descriptive data U all positions	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,81	0,83	0,36	0,80
ONLY descriptive data		H2O RANDOM FOREST	219/94	0,46	0,67	0,30	0,68
HARMONICS All vocals		H2O NAIVE BAYES	219/94	0,26	0,79	0,27	0,84
HARMONICS A normal (_1)		H2O NAIVE BAYES	219/94	0,21	0,94	0,33	0,84
HARMONICS A extension (_2)		H2O GENERALISED LINEAR MODEL	219/94	0,36	0,79	0,23	0,95
HARMONICS A flexion (_3)		H2O NAIVE BAYES	219/94	0,21	0,98	0,37	0,68
HARMONICS E normal (_1)		H2O GENERALISED LINEAR MODEL	219/94	0,30	0,69	0,34	0,74

DATASET (PARAMETERS COMBINATION)	TYPE OF CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				Precision	Recall	Precision	Recall
HARMONICS E extension (_2)		H2O RANDOM FOREST	219/94	0,38	0,74	0,27	0,74
HARMONICS E flexion(_3)		H2O GRADIENT BOOSTING	219/94	0,46	0,46	0,29	0,68
HARMONICS I normal (_1)		H2O NAIVE BAYES	219/94	0,34	0,72	0,27	0,74
HARMONICS I extension (_2)		H2O GRADIENT BOOSTING	219/94	0,36	0,60	0,26	0,74
HARMONICS I flexion(_3)		H2O NAIVE BAYES	219/94	0,26	0,75	0,29	0,84
HARMONICS O normal (_1)		H2O GRADIENT BOOSTING	219/94	0,28	0,62	0,30	0,89
HARMONICS O extension (_2)		H2O GENERALISED LINEAR MODEL	219/94	0,26	0,81	0,23	1,00
HARMONICS O flexion(_3)		H2O GRADIENT BOOSTING	219/94	0,33	0,61	0,26	0,84
HARMONICS U normal (_1)		H2O NAIVE BAYES	219/94	0,34	0,67	0,30	0,84
HARMONICS U extension 2)		H2O NAIVE BAYES	219/94	0,44	0,76	0,23	0,95
HARMONICS U Flexion (_3)		H2O RANDOM FOREST	219/94	0,26	0,92	0,25	0,95
HARMONICS A normal (_1) vs A flexion (_3)		H2O GENERALISED LINEAR MODEL	219/94	0,23	1,00	0,26	0,89
HARMONICS A normal (_1) vs A extension (_2)		H2O NAIVE BAYES	219/94	0,22	0,94	0,31	0,79
HARMONICS A flexion (_3) vs A extension (_2)		H2O NAIVE BAYES	219/94	0,26	0,71	0,38	0,84
HARMONICS A all positions		H2O NAIVE BAYES	219/94	0,41	0,78	0,39	0,68
HARMONICS E normal (_1) vs E flexion (_3)		H2O GENERALISED LINEAR MODEL	219/94	0,24	0,96	0,44	0,63
HARMONICS E normal (_1) vs E extension(_2)		H2O GENERALISED LINEAR MODEL	219/94	0,30	0,68	0,26	0,95
HARMONICS E flexion (_2) vs E extension (_3)		H2O GRADIENT BOOSTING	219/94	0,37	0,63	0,28	0,68

DATASET (PARAMETERS COMBINATION)	TYPE OF CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				Precision	Recall	Precision	Recall
HARMONICS E all positions		H2O NAIVE BAYES	219/94	0,40	0,55	0,33	0,68
HARMONICS I normal (_1) vs I flexion (_3)		H2O NAIVE BAYES	219/94	0,26	0,88	0,27	0,84
HARMONICS I normal (_1) vs I extension(_2)		H2O NAIVE BAYES	219/94	0,29	0,84	0,33	0,68
HARMONICS I flexion (_3) vs I extension (_2)		H2O GRADIENT BOOSTING	219/94	0,48	0,47	0,25	0,84
HARMONICS I all positions		H2O NAIVE BAYES	219/94	0,26	0,90	0,30	0,74
HARMONICS O normal (_1) vs O flexion (_3)		H2O RANDOM FOREST	219/94	0,33	0,68	0,21	0,95
HARMONICS O normal (_1) vs O extension (_2)		H2O NAIVE BAYES	219/94	0,31	0,88	0,23	0,79
HARMONICS O flexion(_3) vs O extension (_2)		H2O GRADIENT BOOSTING	219/94	0,50	0,47	0,33	0,63
HARMONICS O all positions		H2O RANDOM FOREST	219/94	0,37	0,66	0,38	0,42
HARMONICS U normal (_1) vs U flexion (_3)		H2O NAIVE BAYES	219/94	0,23	0,92	0,29	0,95
HARMONICS U normal (_1) vs U extension (_2)		H2O GRADIENT BOOSTING	219/94	0,30	0,57	0,24	0,89
HARMONICS U flexion (_3) vs U extension (_2)		H2O RANDOM FOREST	219/94	0,22	0,97	0,26	0,74
HARMONICS U all positions		H2O RANDOM FOREST	219/94	0,21	0,21	0,21	0,21
Spectral+ Descriptive data all vocals		H2O GRADIENT BOOSTING	219/94	0,34	0,64	0,32	0,68
SPectral+ Descriptive data A normal		H2O RANDOM FOREST	219/94	0,41	0,77	0,35	0,79
SPectral+ Descriptive data O extension		H2O RANDOM FOREST	219/94	0,41	0,78	0,30	0,74
SPectral+ Descriptive data O flexion		H2O NAIVE BAYES	219/94	0,50	0,57	0,34	0,74
Spectral+ Descriptive data A flexion vs A extension		H2O RANDOM FOREST	219/94	0,40	0,86	0,33	0,89
SPectral+ Descriptive data E flexion vs E extension		H2O NAIVE BAYES	219/94	0,45	0,71	0,33	0,68

DATASET (PARAMETERS COMBINATION)	TYPE OF CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				Precision	Recall	Precision	Recall
Voice parameters + Descriptive data all vocals		H2O GENERALISED LINEAR MODEL	219/94	0,33	0,70	0,27	1,00
Voice parameters + Descriptive data A normal		H2O GENERALISED LINEAR MODEL	219/94	0,44	0,61	0,32	0,63
Voice parameters + Descriptive data O extension		H2O GRADIENT BOOSTING	219/94	0,52	0,51	0,33	0,74
Voice parameters + Descriptive data O flexion		H2O RANDOM FOREST	219/94	0,47	0,83	0,38	0,63
Voice parameters + Descriptive data A flexion vs A extension		H2O GENERALISED LINEAR MODEL	219/94	0,43	0,75	0,29	0,84
Voice parameters + Descriptive data E flexion vs E extension		H2O NAIVE BAYES	219/94	0,47	0,58	0,28	0,89
Voice parameters + Descriptive data O flexion vs O extension		H2O NAIVE BAYES	219/94	0,62	0,47	0,28	0,89
Voice parameters + Descriptive data all vocals	Only Cormack 1 and 4	H2O RANDOM FOREST	100/43	0,68	0,73	0,44	0,80
Voice parameters + Descriptive data A normal	Only Cormack 1 and 4	H2O GENERALISED LINEAR MODEL	100/43	0,35	0,69	0,36	0,80
Voice parameters + Descriptive data O extension	Only Cormack 1 and 4	H2O GENERALISED LINEAR MODEL	100/43	0,37	0,78	0,40	0,80
Voice parameters + Descriptive data O flexion	Only Cormack 1 and 4	H2O NAIVE BAYES	100/43	0,84	0,71	0,36	0,80
Voice parameters + Descriptive data A flexion vs A extension	Only Cormack 1 and 4	H2O GENERALISED LINEAR MODEL	100/43	0,38	0,78	0,40	0,80
Voice parameters + Descriptive data E flexion vs E extension	Only Cormack 1 and 4	H2O GRADIENT BOOSTING	100/43	0,29	0,66	0,57	0,80
Voice parameters + Descriptive data O flexion vs O extension	Only Cormack 1 and 4	H2O NAIVE BAYES	100/43	0,75	0,78	0,36	0,80

11.14. ANNEX 14. Full Table 15. Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in both training and testing groups, for AUC and Pr AUC values for the training group and the testing group respectively (Simplified)

DATASET (PARAMETERS COMBINATION)	CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				AUC	Pr AUC	AUC	Pr AUC
HARMONICS + Descriptive data All vocals		H2O GENERALISED LINEAR MODEL	219/94	0,55	0,33	0,62	0,27
HARMONICS + Descriptive data A normal (_1)		H2O GENERALISED LINEAR MODEL	219/94	0,60	0,32	0,67	0,27
HARMONICS + Descriptive data A extension (_2)		H2O GRADIENT BOOSTING	219/94	0,60	0,39	0,71	0,31
HARMONICS + Descriptive data A flexion (_3)		H2O GRADIENT BOOSTING	219/94	0,61	0,40	0,67	0,27
HARMONICS + Descriptive data E normal (_1)		H2O GRADIENT BOOSTING	219/94	0,61	0,42	0,69	0,30
HARMONICS + Descriptive data E extension (_2)		H2O NAIVE BAYES	219/94	0,62	0,41	0,67	0,26
HARMONICS + Descriptive data E flexion(_3)		H2O NAIVE BAYES	219/94	0,61	0,38	0,65	0,24
HARMONICS + Descriptive data I normal (_1)		H2O RANDOM FOREST	219/94	0,62	0,41	0,71	0,30
HARMONICS + Descriptive data I extension (_2)		H2O RANDOM FOREST	219/94	0,62	0,44	0,70	0,29
HARMONICS + Descriptive data I flexion(_3)		H2O NAIVE BAYES	219/94	0,61	0,39	0,67	0,25
HARMONICS + Descriptive data O normal (_1)		H2O RANDOM FOREST	219/94	0,62	0,45	0,69	0,31
HARMONICS + Descriptive data O extension (_2)		H2O RANDOM FOREST	219/94	0,62	0,38	0,70	0,29
HARMONICS + Descriptive data O flexion(_3)		H2O NAIVE BAYES	219/94	0,62	0,39	0,66	0,25
HARMONICS + Descriptive data U normal (_1)		H2O RANDOM FOREST	219/94	0,62	0,43	0,74	0,33
HARMONICS + Descriptive data U extension 2)		H2O RANDOM FOREST	219/94	0,62	0,41	0,74	0,34
HARMONICS + Descriptive data U Flexion (_3)		H2O RANDOM FOREST	219/94	0,61	0,35	0,68	0,29

DATASET (PARAMETERS COMBINATION)	CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				AUC	Pr AUC	AUC	Pr AUC
HARMONICS + Descriptive data A normal (_1) vs A flexion (_3)		H2O RANDOM FOREST	219/94	0,62	0,38	0,71	0,30
HARMONICS + Descriptive data A normal (_1) vs A extension (_2)		H2O RANDOM FOREST	219/94	0,61	0,46	0,68	0,26
HARMONICS + Descriptive data A flexion (_3) vs A extension (_2)		H2O RANDOM FOREST	219/94	0,61	0,39	0,66	0,25
HARMONICS + Descriptive data A all positions		H2O RANDOM FOREST	219/94	0,61	0,47	0,69	0,28
HARMONICS + Descriptive data E normal (_1) vs E flexion (_3)		H2O NAIVE BAYES	219/94	0,62	0,41	0,68	0,26
HARMONICS + Descriptive data E normal (_1) vs E extension(_2)		H2O RANDOM FOREST	219/94	0,62	0,43	0,68	0,27
HARMONICS + Descriptive data E flexion (_2) vs E extension (_3)		H2O NAIVE BAYES	219/94	0,62	0,44	0,67	0,28
HARMONICS + Descriptive data E all positions		H2O NAIVE BAYES	219/94	0,62	0,42	0,68	0,26
HARMONICS + Descriptive data I normal (_1) vs I flexion (_3)		H2O GRADIENT BOOSTING	219/94	0,61	0,38	0,72	0,30
HARMONICS + Descriptive data I normal (_1) vs I extension(_2)		H2O GRADIENT BOOSTING	219/94	0,61	0,41	0,69	0,28
HARMONICS + Descriptive data I flexion (_3) vs I extension (_2)		H2O RANDOM FOREST	219/94	0,62	0,42	0,73	0,34
HARMONICS + Descriptive data I all positions		H2O NAIVE BAYES	219/94	0,61	0,38	0,67	0,26
HARMONICS + Descriptive data O normal (_1) vs O flexion (_3)		H2O GRADIENT BOOSTING	219/94	0,59	0,28	0,61	0,24
HARMONICS + Descriptive data O normal (_1) vs O extension (_2)		H2O GENERALISED LINEAR MODEL	219/94	0,61	0,33	0,70	0,29

DATASET (PARAMETERS COMBINATION)	CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				AUC	Pr AUC	AUC	Pr AUC
HARMONICS + Descriptive data O flexion(_3) vs O extension (_2)		H2O GRADIENT BOOSTING	219/94	0,62	0,27	0,63	0,25
HARMONICS + Descriptive data O all positions		H2O GRADIENT BOOSTING	219/94	0,58	0,30	0,70	0,30
HARMONICS + Descriptive data U normal (_1) vs U flexion (_3)		H2O RANDOM FOREST	219/94	0,67	0,31	0,59	0,22
HARMONICS + Descriptive data U normal (_1) vs U extension (_2)		H2O RANDOM FOREST	219/94	0,67	0,28	0,64	0,28
HARMONICS + Descriptive data U flexion (_3) vs U extension (_2)		H2O NAIVE BAYES	219/94	0,70	0,36	0,60	0,24
HARMONICS + Descriptive data U all positions		H2O NAIVE BAYES	219/94	0,68	0,36	0,64	0,26
HARMONICS+ decriptive data all vocals	Overweight Cormack 4	H2O RANDOM FOREST	219/94	0,67	0,34	0,62	0,27
HARMONICS + Descriptive data A normal (_1)	Overweight Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,71	0,33	0,63	0,26
HARMONICS + Descriptive data A extension (_2)	Overweigh Cormack 4	H2O GRADIENT BOOSTING	219/94	0,61	0,29	0,60	0,23
HARMONICS + Descriptive data A flexion (_3)	Overweigh Cormack 4	H2O RANDOM FOREST	219/94	0,73	0,31	0,55	0,20
HARMONICS + Descriptive data E normal (_1)	Overweigh Cormack 4	H2O RANDOM FOREST	219/94	0,73	0,31	0,56	0,21
HARMONICS + Descriptive data I normal (_1)	Overweigh Cormack 4	H2O GRADIENT BOOSTING	219/94	0,67	0,33	0,50	0,20
HARMONICS + Descriptive data I extension (_2)	Overweigh Cormack 4	H2O GRADIENT BOOSTING	219/94	0,63	0,32	0,53	0,20
HARMONICS + Descriptive data O normal (_1)	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,73	0,34	0,64	0,26
HARMONICS + Descriptive data O extension (_2)	Overweigh Cormack 4	H2O GRADIENT BOOSTING	219/94	0,68	0,35	0,53	0,21
HARMONICS + Descriptive data O flexion(_3)	Overweigh Cormack 4	H2O GRADIENT BOOSTING	219/94	0,75	0,43	0,49	0,20
HARMONICS + Descriptive data U normal (_1)	Overweigh Cormack 4	H2O NAIVE BAYES	219/94	0,74	0,37	0,59	0,22

DATASET (PARAMETERS COMBINATION)	CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				AUC	Pr AUC	AUC	Pr AUC
HARMONICS + Descriptive data U extension 2)	Overweigh Cormack 4	H2O RANDOM FOREST	219/94	0,73	0,37	0,54	0,22
HARMONICS + Descriptive data U Flexion (_3)	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,73	0,33	0,64	0,27
HARMONICS + Descriptive data A normal (_1) vs A flexion (_3)	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,69	0,32	0,63	0,27
HARMONICS + Descriptive data A normal (_1) vs A extension (_2)	Overweigh Cormack 4	H2O NAIVE BAYES	219/94	0,74	0,38	0,61	0,23
HARMONICS + Descriptive data A flexion (_3) vs A extension (_2)	Overweigh Cormack 4	H2O NAIVE BAYES	219/94	0,70	0,35	0,58	0,22
HARMONICS + Descriptive data A all positions	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,69	0,33	0,66	0,29
HARMONICS + Descriptive data E normal (_1) vs E flexion (_3)	Overweigh Cormack 4	H2O NAIVE BAYES	219/94	0,74	0,39	0,63	0,24
HARMONICS + Descriptive data E normal (_1) vs E extension(_2)	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,70	0,34	0,70	0,31
HARMONICS + Descriptive data E flexion (_2) vs E extension (_3)	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,70	0,34	0,67	0,29
HARMONICS + Descriptive data E all positions	Overweigh Cormack 4	H2O NAIVE BAYES	219/94	0,74	0,39	0,63	0,24
HARMONICS + Descriptive data I normal (_1) vs I flexion (_3)	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,72	0,33	0,64	0,27
HARMONICS + Descriptive data I normal (_1) vs I extension(_2)	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,70	0,33	0,64	0,27
HARMONICS + Descriptive data I all positions	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,70	0,33	0,64	0,27
HARMONICS + Descriptive data U normal (_1) vs U flexion (_3)	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,73	0,33	0,64	0,28

DATASET (PARAMETERS COMBINATION)	CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				AUC	Pr AUC	AUC	Pr AUC
HARMONICS + Descriptive data U normal (_1) vs U extension (_2)	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,70	0,33	0,64	0,26
HARMONICS + Descriptive data U flexion (_3) vs U extension (_2)	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,73	0,35	0,66	0,28
HARMONICS + Descriptive data U all positions	Overweigh Cormack 4	H2O GRADIENT BOOSTING	219/94	0,60	0,29	0,58	0,22
HARMONICS+ decriptive data all vocals	Cormack 1 and 4 cases	H2O GRADIENT BOOSTING	100/43	0,73	0,15	0,83	0,51
HARMONICS + Descriptive data A normal (_1)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,62	0,16	0,84	0,30
HARMONICS + Descriptive data A extension (_2)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,68	0,20	0,85	0,30
HARMONICS + Descriptive data A flexion (_3)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,67	0,18	0,85	0,29
HARMONICS + Descriptive data E normal (_1)	Cormack 1 and 4 cases	H2O RANDOM FOREST	100/43	0,77	0,16	0,83	0,32
HARMONICS + Descriptive data E extension (_2)	Cormack 1 and 4 cases	H2O GRADIENT BOOSTING	100/43	0,63	0,10	0,88	0,27
HARMONICS + Descriptive data E flexion(_3)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,85	0,23	0,83	0,21
HARMONICS + Descriptive data I normal (_1)	Cormack 1 and 4 cases	H2O RANDOM FOREST	100/43	0,70	0,14	0,85	0,30
HARMONICS + Descriptive data I extension (_2)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,85	0,22	0,79	0,18
HARMONICS + Descriptive data I flexion(_3)	Cormack 1 and 4 cases	H2O GRADIENT BOOSTING	100/43	0,65	0,11	0,94	0,52
HARMONICS + Descriptive data O normal (_1)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,75	0,28	0,86	0,32
HARMONICS + Descriptive data O extension (_2)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,70	0,19	0,82	0,27
HARMONICS + Descriptive data O flexion(_3)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,85	0,22	0,76	0,20
HARMONICS + Descriptive data U normal (_1)	Cormack 1 and 4 cases	H2O GRADIENT BOOSTING	100/43	0,51	0,16	0,83	0,34

DATASET (PARAMETERS COMBINATION)	CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				AUC	Pr AUC	AUC	Pr AUC
HARMONICS + Descriptive data U extension 2)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,78	0,29	0,83	0,28
HARMONICS + Descriptive data U Flexion (_3)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,84	0,22	0,80	0,21
HARMONICS + Descriptive data A normal (_1) vs A flexion (_3)	Cormack 1 and 4 cases	H2O RANDOM FOREST	100/43	0,75	0,19	0,82	0,23
HARMONICS + Descriptive data A normal (_1) vs A extension (_2)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,82	0,23	0,82	0,27
HARMONICS + Descriptive data A flexion (_3) vs A extension (_2)	Cormack 1 and 4 cases	H2O GRADIENT BOOSTING	100/43	0,55	0,09	0,86	0,35
HARMONICS + Descriptive data A all positions	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,64	0,18	0,91	0,40
HARMONICS + Descriptive data E normal (_1) vs E flexion (_3)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,66	0,17	0,84	0,30
HARMONICS + Descriptive data E normal (_1) vs E extension(_2)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,88	0,26	0,81	0,20
HARMONICS + Descriptive data E flexion (_2) vs E extension (_3)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,59	0,16	0,82	0,27
HARMONICS + Descriptive data E all positions	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,85	0,32	0,84	0,23
HARMONICS + Descriptive data I normal (_1) vs I flexion (_3)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,87	0,23	0,81	0,21
HARMONICS + Descriptive data I normal (_1) vs I extension(_2)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,86	0,21	0,79	0,19
HARMONICS + Descriptive data I flexion (_3) vs I extension (_2)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,87	0,23	0,81	0,19
HARMONICS + Descriptive data I all positions	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,68	0,17	0,83	0,27

DATASET (PARAMETERS COMBINATION)	CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				AUC	Pr AUC	AUC	Pr AUC
HARMONICS + Descriptive data O normal (_1) vs O flexion (_3)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,66	0,18	0,82	0,26
HARMONICS + Descriptive data O normal (_1) vs O extension (_2)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,63	0,18	0,83	0,27
HARMONICS + Descriptive data O flexion(_3) vs O extension (_2)	Cormack 1 and 4 cases	H2O RANDOM FOREST	100/43	0,68	0,14	0,73	0,16
HARMONICS + Descriptive data O all positions	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,91	0,29	0,83	0,23
HARMONICS + Descriptive data U normal (_1) vs U flexion (_3)	Cormack 1 and 4 cases	H2O RANDOM FOREST	100/43	0,68	0,13	0,79	0,22
HARMONICS + Descriptive data U normal (_1) vs U extension (_2)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,77	0,29	0,84	0,26
HARMONICS + Descriptive data U flexion (_3) vs U extension (_2)	Cormack 1 and 4 cases	H2O RANDOM FOREST	100/43	0,82	0,25	0,81	0,20
HARMONICS + Descriptive data U all positions	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,90	0,31	0,82	0,22
ONLY descriptive data		H2O RANDOM FOREST	219/94	0,69	0,34	0,56	0,22
HARMONICS All vocals		H2O NAIVE BAYES	219/94	0,50	0,20	0,63	0,29
HARMONICS A normal (_1)		H2O NAIVE BAYES	219/94	0,39	0,16	0,71	0,33
HARMONICS A extension (_2)		H2O GENERALISED LINEAR MODEL	219/94	0,60	0,29	0,47	0,18
HARMONICS A flexion (_3)		H2O NAIVE BAYES	219/94	0,43	0,19	0,63	0,28
HARMONICS E normal (_1)		H2O GENERALISED LINEAR MODEL	219/94	0,55	0,23	0,66	0,27
HARMONICS E extension (_2)		H2O RANDOM FOREST	219/94	0,54	0,24	0,58	0,23
HARMONICS E flexion(_3)		H2O GRADIENT BOOSTING	219/94	0,65	0,33	0,61	0,25
HARMONICS I normal (_1)		H2O NAIVE BAYES	219/94	0,59	0,27	0,65	0,36
HARMONICS I extension (_2)		H2O GRADIENT BOOSTING	219/94	0,50	0,22	0,57	0,23

DATASET (PARAMETERS COMBINATION)	CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				AUC	Pr AUC	AUC	Pr AUC
HARMONICS I flexion(_3)		H2O NAIVE BAYES	219/94	0,50	0,21	0,65	0,29
HARMONICS O normal (_1)		H2O GRADIENT BOOSTING	219/94	0,55	0,22	0,62	0,23
HARMONICS O extension (_2)		H2O GENERALISED LINEAR MODEL	219/94	0,53	0,24	0,56	0,23
HARMONICS O flexion(_3)		H2O GRADIENT BOOSTING	219/94	0,52	0,23	0,59	0,20
HARMONICS U normal (_1)		H2O NAIVE BAYES	219/94	0,53	0,23	0,64	0,29
HARMONICS U extension 2)		H2O NAIVE BAYES	219/94	0,55	0,27	0,52	0,20
HARMONICS U Flexion (_3)		H2O RANDOM FOREST	219/94	0,44	0,18	0,55	0,21
HARMONICS A normal (_1) vs A flexion (_3)		H2O GENERALISED LINEAR MODEL	219/94	0,45	0,17	0,61	0,24
HARMONICS A normal (_1) vs A extension (_2)		H2O NAIVE BAYES	219/94	0,42	0,17	0,65	0,25
HARMONICS A flexion (_3) vs A extension (_2)		H2O NAIVE BAYES	219/94	0,46	0,20	0,69	0,28
HARMONICS A all positions		H2O NAIVE BAYES	219/94	0,43	0,19	0,64	0,25
HARMONICS E normal (_1) vs E flexion (_3)		H2O GENERALISED LINEAR MODEL	219/94	0,58	0,24	0,70	0,32
HARMONICS E normal (_1) vs E extension(_2)		H2O GENERALISED LINEAR MODEL	219/94	0,54	0,23	0,64	0,28
HARMONICS E flexion (_2) vs E extension (_3)		H2O GRADIENT BOOSTING	219/94	0,61	0,26	0,56	0,22
HARMONICS E all positions		H2O NAIVE BAYES	219/94	0,53	0,25	0,66	0,31
HARMONICS I normal (_1) vs I flexion (_3)		H2O NAIVE BAYES	219/94	0,51	0,20	0,65	0,32
HARMONICS I normal (_1) vs I extension(_2)		H2O NAIVE BAYES	219/94	0,57	0,26	0,66	0,30
HARMONICS I flexion (_3) vs I extension (_2)		H2O GRADIENT BOOSTING	219/94	0,53	0,25	0,59	0,25
HARMONICS I all positions		H2O NAIVE BAYES	219/94	0,56	0,24	0,65	0,29
HARMONICS O normal (_1) vs O flexion (_3)		H2O RANDOM FOREST	219/94	0,55	0,23	0,46	0,24

DATASET (PARAMETERS COMBINATION)	CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				AUC	Pr AUC	AUC	Pr AUC
HARMONICS O normal (_1) vs O extension (_2)		H2O NAIVE BAYES	219/94	0,57	0,25	0,55	0,22
HARMONICS O flexion(_3) vs O extension (_2)		H2O GRADIENT BOOSTING	219/94	0,55	0,24	0,62	0,32
HARMONICS O all positions		H2O RANDOM FOREST	219/94	0,53	0,24	0,58	0,30
HARMONICS U normal (_1) vs U flexion (_3)		H2O NAIVE BAYES	219/94	0,42	0,19	0,67	0,29
HARMONICS U normal (_1) vs U extension (_2)		H2O GRADIENT BOOSTING	219/94	0,48	0,20	0,49	0,18
HARMONICS U flexion (_3) vs U extension (_2)		H2O RANDOM FOREST	219/94	0,48	0,17	0,61	0,25
HARMONICS U all positions		H2O RANDOM FOREST	219/94	0,42	0,17	0,59	0,23
Spectral+ Descriptive data all vocals		H2O GRADIENT BOOSTING	219/94	0,56	0,22	0,60	0,27
SPectral+ Descriptive data A normal		H2O RANDOM FOREST	219/94	0,74	0,33	0,70	0,30
SPectral+ Descriptive data O extension		H2O RANDOM FOREST	219/94	0,73	0,32	0,65	0,26
SPectral+ Descriptive data O flexion		H2O NAIVE BAYES	219/94	0,70	0,31	0,67	0,23
Spectral+ Descriptive data A flexion vs A extension		H2O RANDOM FOREST	219/94	0,68	0,30	0,69	0,29
SPectral+ Descriptive data E flexion vs E extension		H2O NAIVE BAYES	219/94	0,70	0,29	0,69	0,21
Voice parameters + Descriptive data all vocals		H2O GENERALISED LINEAR MODEL	219/94	0,51	0,28	0,61	0,23
Voice parameters + Descriptive data A normal		H2O GENERALISED LINEAR MODEL	219/94	0,71	0,31	0,62	0,24
Voice parameters + Descriptive data O extension		H2O GRADIENT BOOSTING	219/94	0,65	0,35	0,64	0,25
Voice parameters + Descriptive data O flexion		H2O RANDOM FOREST	219/94	0,77	0,38	0,67	0,28
Voice parameters + Descriptive data A flexion vs A extension		H2O GENERALISED LINEAR MODEL	219/94	0,75	0,36	0,64	0,24
Voice parameters + Descriptive data E flexion vs E extension		H2O NAIVE BAYES	219/94	0,67	0,35	0,62	0,23

DATASET (PARAMETERS COMBINATION)	CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				AUC	Pr AUC	AUC	Pr AUC
Voice parameters + Descriptive data O flexion vs O extension		H2O NAIVE BAYES	219/94	0,69	0,37	0,61	0,23
Voice parameters + Descriptive data all vocals	Only Cormack 1 and 4	H2O RANDOM FOREST	100/43	0,76	0,14	0,75	0,25
Voice parameters + Descriptive data A normal	Only Cormack 1 and 4	H2O GENERALISED LINEAR MODEL	100/43	0,69	0,18	0,82	0,26
Voice parameters + Descriptive data O extension	Only Cormack 1 and 4	H2O GENERALISED LINEAR MODEL	100/43	0,68	0,19	0,87	0,32
Voice parameters + Descriptive data O flexion	Only Cormack 1 and 4	H2O NAIVE BAYES	100/43	0,86	0,20	0,84	0,23
Voice parameters + Descriptive data A flexion vs A extension	Only Cormack 1 and 4	H2O GENERALISED LINEAR MODEL	100/43	0,78	0,23	0,86	0,30
Voice parameters + Descriptive data E flexion vs E extension	Only Cormack 1 and 4	H2O GRADIENT BOOSTING	100/43	0,60	0,15	0,86	0,35
Voice parameters + Descriptive data O flexion vs O extension	Only Cormack 1 and 4	H2O NAIVE BAYES	100/43	0,88	0,23	0,84	0,23

11.15. ANNEX 15. Full Table 16. Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in both training and testing groups, for F1 and Log loss values for the training group and the testing group respectively.

DATASET (PARAMETERS COMBINATION)	CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				F1	Log Loss	F1	Log Loss
HARMONICS + Descriptive data All vocals		H2O GENERALISED LINEAR MODEL	219/94	0,48	0,60	0,41	0,54
HARMONICS + Descriptive data A normal (_1)		H2O GENERALISED LINEAR MODEL	219/94	0,50	0,61	0,45	0,49
HARMONICS + Descriptive data A extension (_2)		H2O GRADIENT BOOSTING	219/94	0,52	0,53	0,49	0,48
HARMONICS + Descriptive data A flexion (_3)		H2O GRADIENT BOOSTING	219/94	0,54	0,54	0,49	0,56
HARMONICS + Descriptive data E normal (_1)		H2O GRADIENT BOOSTING	219/94	0,54	0,51	0,49	0,53
HARMONICS + Descriptive data E extension (_2)		H2O NAIVE BAYES	219/94	0,56	0,86	0,46	
HARMONICS + Descriptive data E flexion(_3)		H2O NAIVE BAYES	219/94	0,54	0,88	0,44	0,98
HARMONICS + Descriptive data I normal (_1)		H2O RANDOM FOREST	219/94	0,61	0,51	0,49	0,52
HARMONICS + Descriptive data I extension (_2)		H2O RANDOM FOREST	219/94	0,60	0,55	0,47	0,56
HARMONICS + Descriptive data I flexion(_3)		H2O NAIVE BAYES	219/94	0,54	0,82	0,47	0,91
HARMONICS + Descriptive data O normal (_1)		H2O RANDOM FOREST	219/94	0,62	0,45	0,45	0,48
HARMONICS + Descriptive data O extension (_2)		H2O RANDOM FOREST	219/94	0,57	0,60	0,46	0,47
HARMONICS + Descriptive data O flexion(_3)		H2O NAIVE BAYES	219/94	0,56	0,89	0,44	
HARMONICS + Descriptive data U normal (_1)		H2O RANDOM FOREST	219/94	0,62	0,55	0,51	0,54
HARMONICS + Descriptive data U extension 2)		H2O RANDOM FOREST	219/94	0,61	0,51	0,48	0,52
HARMONICS + Descriptive data U Flexion (_3)		H2O RANDOM FOREST	219/94	0,55	0,49	0,46	0,50

DATASET (PARAMETERS COMBINATION)	CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				F1	Log Loss	F1	Log Loss
HARMONICS + Descriptive data A normal (_1) vs A flexion (_3)		H2O RANDOM FOREST	219/94	0,56	0,52	0,49	0,52
HARMONICS + Descriptive data A normal (_1) vs A extension (_2)		H2O RANDOM FOREST	219/94	0,58	0,75	0,48	0,48
HARMONICS + Descriptive data A flexion (_3) vs A extension (_2)		H2O RANDOM FOREST	219/94	0,58	0,57	0,49	0,58
HARMONICS + Descriptive data A all positions		H2O RANDOM FOREST	219/94	0,62	0,61	0,46	0,49
HARMONICS + Descriptive data E normal (_1) vs E flexion (_3)		H2O NAIVE BAYES	219/94	0,56	0,88	0,47	0,98
HARMONICS + Descriptive data E normal (_1) vs E extension(_2)		H2O RANDOM FOREST	219/94	0,60	0,49	0,44	0,54
HARMONICS + Descriptive data E flexion (_2) vs E extension (_3)		H2O NAIVE BAYES	219/94	0,61	0,80	0,44	
HARMONICS + Descriptive data E all positions		H2O NAIVE BAYES	219/94	0,57	0,89	0,45	
HARMONICS + Descriptive data I normal (_1) vs I flexion (_3)		H2O GRADIENT BOOSTING	219/94	0,53	0,53	0,52	0,52
HARMONICS + Descriptive data I normal (_1) vs I extension(_2)		H2O GRADIENT BOOSTING	219/94	0,54	0,53	0,49	0,51
HARMONICS + Descriptive data I flexion (_3) vs I extension (_2)		H2O RANDOM FOREST	219/94	0,61	0,56	0,49	0,55
HARMONICS + Descriptive data I all positions		H2O NAIVE BAYES	219/94	0,53	0,93	0,46	0,98
HARMONICS + Descriptive data O normal (_1) vs O flexion (_3)		H2O GRADIENT BOOSTING	219/94	0,47	0,58	0,45	0,53
HARMONICS + Descriptive data O normal (_1) vs O extension (_2)		H2O GENERALISED LINEAR MODEL	219/94	0,51	0,59	0,48	0,48
HARMONICS + Descriptive data O flexion(_3) vs O extension (_2)		H2O GRADIENT BOOSTING	219/94	0,42	0,99	0,43	0,76
HARMONICS + Descriptive data O all positions		H2O GRADIENT BOOSTING	219/94	0,41	0,61	0,46	0,47
HARMONICS + Descriptive data U normal (_1) vs U flexion (_3)		H2O RANDOM FOREST	219/94	0,48	0,51	0,44	0,54
HARMONICS + Descriptive data U normal (_1) vs U extension (_2)		H2O RANDOM FOREST	219/94	0,50	0,49	0,43	0,50
HARMONICS + Descriptive data U flexion (_3) vs U extension (_2)		H2O NAIVE BAYES	219/94	0,52	0,81	0,40	0,98
HARMONICS + Descriptive data U all positions		H2O NAIVE BAYES	219/94	0,52	0,88	0,42	

DATASET (PARAMETERS COMBINATION)	CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				F1	Log Loss	F1	Log Loss
HARMONICS+ decriptive data all vocals	Overweigh Cormack 4	H2O RANDOM FOREST	219/94	0,51	0,54	0,39	0,49
HARMONICS + Descriptive data A normal (_1)	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,47	0,58	0,43	0,50
HARMONICS + Descriptive data A extension (_2)	Overweigh Cormack 4	H2O GRADIENT BOOSTING	219/94	0,48	0,63	0,44	0,61
HARMONICS + Descriptive data A flexion (_3)	Overweigh Cormack 4	H2O RANDOM FOREST	219/94	0,52	0,52	0,39	0,51
HARMONICS + Descriptive data E normal (_1)	Overweigh Cormack 4	H2O RANDOM FOREST	219/94	0,52	0,48	0,40	0,52
HARMONICS + Descriptive data E extension (_2)	Overweigh Cormack 4		219/94				
HARMONICS + Descriptive data E flexion(_3)	Overweigh Cormack 4		219/94				
HARMONICS + Descriptive data I normal (_1)	Overweigh Cormack 4	H2O GRADIENT BOOSTING	219/94	0,48	0,66	0,37	0,80
HARMONICS + Descriptive data I extension (_2)	Overweigh Cormack 4	H2O GRADIENT BOOSTING	219/94	0,45		0,36	
HARMONICS + Descriptive data I flexion(_3)	Overweigh Cormack 4		219/94				
HARMONICS + Descriptive data O normal (_1)	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,51	0,56	0,44	0,50
HARMONICS + Descriptive data O extension (_2)	Overweigh Cormack 4	H2O GRADIENT BOOSTING	219/94	0,49	0,91	0,38	
HARMONICS + Descriptive data O flexion(_3)	Overweigh Cormack 4	H2O GRADIENT BOOSTING	219/94	0,58		0,36	
HARMONICS + Descriptive data U normal (_1)	Overweigh Cormack 4	H2O NAIVE BAYES	219/94	0,52	0,72	0,41	
HARMONICS + Descriptive data U extension 2)	Overweigh Cormack 4	H2O RANDOM FOREST	219/94	0,52	0,46	0,36	0,55
HARMONICS + Descriptive data U Flexion (_3)	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,50	0,58	0,44	0,50
HARMONICS + Descriptive data A normal (_1) vs A flexion (_3)	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,45	0,58	0,44	0,50
HARMONICS + Descriptive data A normal (_1) vs A extension (_2)	Overweigh Cormack 4	H2O NAIVE BAYES	219/94	0,51	0,71	0,43	0,99

DATASET (PARAMETERS COMBINATION)	CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				F1	Log Loss	F1	Log Loss
HARMONICS + Descriptive data A flexion (_3) vs A extension (_2)	Overweigh Cormack 4	H2O NAIVE BAYES	219/94	0,50	0,67	0,40	0,98
HARMONICS + Descriptive data A all positions	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,47	0,58	0,44	0,49
HARMONICS + Descriptive data E normal (_1) vs E flexion (_3)	Overweigh Cormack 4	H2O NAIVE BAYES	219/94	0,54	0,73	0,43	0,98
HARMONICS + Descriptive data E normal (_1) vs E extension(_2)	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,48	0,57	0,44	0,47
HARMONICS + Descriptive data E flexion (_2) vs E extension (_3)	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,48	0,56	0,43	0,48
HARMONICS + Descriptive data E all positions	Overweigh Cormack 4	H2O NAIVE BAYES	219/94	0,55	0,75	0,41	0,94
HARMONICS + Descriptive data I normal (_1) vs I flexion (_3)	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,49	0,58	0,44	0,50
HARMONICS + Descriptive data I normal (_1) vs I extension(_2)	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,51	0,58	0,44	0,50
HARMONICS + Descriptive data I flexion (_3) vs I extension (_2)	Overweigh Cormack 4		219/94				
HARMONICS + Descriptive data I all positions	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,51	0,58	0,44	0,50
HARMONICS + Descriptive data O normal (_1) vs O flexion (_3)	Overweigh Cormack 4		219/94				
HARMONICS + Descriptive data O normal (_1) vs O extension (_2)	Overweigh Cormack 4		219/94				
HARMONICS + Descriptive data O flexion(_3) vs O extension (_2)	Overweigh Cormack 4		219/94				
HARMONICS + Descriptive data O all positions	Overweigh Cormack 4		219/94				
HARMONICS + Descriptive data U normal (_1) vs U flexion (_3)	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,50	0,58	0,44	0,50
HARMONICS + Descriptive data U normal (_1) vs U extension (_2)	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,50	0,58	0,44	0,50

DATASET (PARAMETERS COMBINATION)	CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				F1	Log Loss	F1	Log Loss
HARMONICS + Descriptive data U flexion (_3) vs U extension (_2)	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	0,50	0,57	0,43	0,49
HARMONICS + Descriptive data U all positions	Overweigh Cormack 4	H2O GRADIENT BOOSTING	219/94	0,43		0,39	
HARMONICS+ decriptive data all vocals	Cormack 1 and 4 cases	H2O GRADIENT BOOSTING	100/43	0,55		0,75	0,45
HARMONICS + Descriptive data A normal (_1)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,38	0,53	0,53	0,39
HARMONICS + Descriptive data A extension (_2)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,50	0,52	0,50	0,38
HARMONICS + Descriptive data A flexion (_3)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,42	0,52	0,50	0,37
HARMONICS + Descriptive data E normal (_1)	Cormack 1 and 4 cases	H2O RANDOM FOREST	100/43	0,56	0,39	0,62	0,29
HARMONICS + Descriptive data E extension (_2)	Cormack 1 and 4 cases	H2O GRADIENT BOOSTING	100/43	0,43	0,57	0,57	0,30
HARMONICS + Descriptive data E flexion(_3)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,69	0,78	0,47	
HARMONICS + Descriptive data I normal (_1)	Cormack 1 and 4 cases	H2O RANDOM FOREST	100/43	0,56	0,43	0,53	0,29
HARMONICS + Descriptive data I extension (_2)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,70	0,72	0,44	
HARMONICS + Descriptive data I flexion(_3)	Cormack 1 and 4 cases	H2O GRADIENT BOOSTING	100/43	0,48		0,73	0,89
HARMONICS + Descriptive data O normal (_1)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,57	0,42	0,57	0,40
HARMONICS + Descriptive data O extension (_2)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,45	0,43	0,47	0,43

DATASET (PARAMETERS COMBINATION)	CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				F1	Log Loss	F1	Log Loss
HARMONICS + Descriptive data O flexion(_3)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,68	0,76	0,44	
HARMONICS + Descriptive data U normal (_1)	Cormack 1 and 4 cases	H2O GRADIENT BOOSTING	100/43	0,34	0,49	0,62	0,30
HARMONICS + Descriptive data U extension 2)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,57	0,40	0,53	0,46
HARMONICS + Descriptive data U Flexion (_3)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,67	0,68	0,44	
HARMONICS + Descriptive data A normal (_1) vs A flexion (_3)	Cormack 1 and 4 cases	H2O RANDOM FOREST	100/43	0,61	0,43	0,50	0,28
HARMONICS + Descriptive data A normal (_1) vs A extension (_2)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,64	0,63	0,50	
HARMONICS + Descriptive data A flexion (_3) vs A extension (_2)	Cormack 1 and 4 cases	H2O GRADIENT BOOSTING	100/43	0,49	0,48	0,73	0,28
HARMONICS + Descriptive data A all positions	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,44	0,50	0,67	0,24
HARMONICS + Descriptive data E normal (_1) vs E flexion (_3)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,40	0,46	0,53	0,36
HARMONICS + Descriptive data E normal (_1) vs E extension(_2)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,74	0,71	0,44	
HARMONICS + Descriptive data E flexion (_2) vs E extension (_3)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,38	0,48	0,47	0,38
HARMONICS + Descriptive data E all positions	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,80	0,63	0,50	
HARMONICS + Descriptive data I normal (_1) vs I flexion (_3)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,68	0,78	0,47	

DATASET (PARAMETERS COMBINATION)	CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				F1	Log Loss	F1	Log Loss
HARMONICS + Descriptive data I normal (_1) vs I extension(_2)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,66	0,79	0,44	
HARMONICS + Descriptive data I flexion (_3) vs I extension (_2)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,67	0,75	0,44	
HARMONICS + Descriptive data I all positions	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,39	0,47	0,47	0,47
HARMONICS + Descriptive data O normal (_1) vs O flexion (_3)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,43	0,46	0,50	0,39
HARMONICS + Descriptive data O normal (_1) vs O extension (_2)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,42	0,45	0,44	0,39
HARMONICS + Descriptive data O flexion(_3) vs O extension (_2)	Cormack 1 and 4 cases	H2O RANDOM FOREST	100/43	0,56	0,44	0,40	0,33
HARMONICS + Descriptive data O all positions	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,77	0,69	0,47	
HARMONICS + Descriptive data U normal (_1) vs U flexion (_3)	Cormack 1 and 4 cases	H2O RANDOM FOREST	100/43	0,54	0,39	0,53	0,35
HARMONICS + Descriptive data U normal (_1) vs U extension (_2)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	0,57	0,40	0,53	0,40
HARMONICS + Descriptive data U flexion (_3) vs U extension (_2)	Cormack 1 and 4 cases	H2O RANDOM FOREST	100/43	0,54	0,33	0,47	0,32
HARMONICS + Descriptive data U all positions	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	0,78	0,47	0,50	
ONLY descriptive data		H2O RANDOM FOREST	219/94	0,50	0,60	0,41	0,52
HARMONICS All vocals		H2O NAIVE BAYES	219/94	0,36	0,56	0,41	
HARMONICS A normal (_1)		H2O NAIVE BAYES	219/94	0,35	0,73	0,47	0,48

DATASET (PARAMETERS COMBINATION)	CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				F1	Log Loss	F1	Log Loss
HARMONICS A extension (_2)		H2O GENERALISED LINEAR MODEL	219/94	0,46	0,68	0,37	0,55
HARMONICS A flexion (_3)		H2O NAIVE BAYES	219/94	0,35	0,60	0,48	0,52
HARMONICS E normal (_1)		H2O GENERALISED LINEAR MODEL	219/94	0,40	0,72	0,47	0,49
HARMONICS E extension (_2)		H2O RANDOM FOREST	219/94	0,40	0,57	0,40	0,53
HARMONICS E flexion(_3)		H2O GRADIENT BOOSTING	219/94	0,44		0,41	
HARMONICS I normal (_1)		H2O NAIVE BAYES	219/94	0,42	0,65	0,40	0,47
HARMONICS I extension (_2)		H2O GRADIENT BOOSTING	219/94	0,38	0,67	0,38	0,54
HARMONICS I flexion(_3)		H2O NAIVE BAYES	219/94	0,37	0,73	0,43	0,51
HARMONICS O normal (_1)		H2O GRADIENT BOOSTING	219/94	0,37	0,66	0,45	0,50
HARMONICS O extension (_2)		H2O GENERALISED LINEAR MODEL	219/94	0,39	0,68	0,37	0,51
HARMONICS O flexion(_3)		H2O GRADIENT BOOSTING	219/94	0,38	0,64	0,40	0,50
HARMONICS U normal (_1)		H2O NAIVE BAYES	219/94	0,39	0,62	0,44	0,87
HARMONICS U extension 2)		H2O NAIVE BAYES	219/94	0,42	0,68	0,37	0,89
HARMONICS U Flexion (_3)		H2O RANDOM FOREST	219/94	0,39	0,71	0,40	0,50
HARMONICS A normal (_1) vs A flexion (_3)		H2O GENERALISED LINEAR MODEL	219/94	0,37	0,71	0,40	0,50
HARMONICS A normal (_1) vs A extension (_2)		H2O NAIVE BAYES	219/94	0,35	0,73	0,44	0,54
HARMONICS A flexion (_3) vs A extension (_2)		H2O NAIVE BAYES	219/94	0,36	0,81	0,52	0,53
HARMONICS A all positions		H2O NAIVE BAYES	219/94	0,35	0,78	0,50	0,61

DATASET (PARAMETERS COMBINATION)	CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				F1	Log Loss	F1	Log Loss
HARMONICS E normal (_1) vs E flexion (_3)		H2O GENERALISED LINEAR MODEL	219/94	0,39	0,69	0,52	0,50
HARMONICS E normal (_1) vs E extension(_2)		H2O GENERALISED LINEAR MODEL	219/94	0,41	0,68	0,40	0,54
HARMONICS E flexion (_2) vs E extension (_3)		H2O GRADIENT BOOSTING	219/94	0,44	0,68	0,39	0,67
HARMONICS E all positions		H2O NAIVE BAYES	219/94	0,39		0,44	0,63
HARMONICS I normal (_1) vs I flexion (_3)		H2O NAIVE BAYES	219/94	0,40	0,86	0,41	0,52
HARMONICS I normal (_1) vs I extension(_2)		H2O NAIVE BAYES	219/94	0,42	0,77	0,44	0,62
HARMONICS I flexion (_3) vs I extension (_2)		H2O GRADIENT BOOSTING	219/94	0,40	0,63	0,39	0,50
HARMONICS I all positions		H2O NAIVE BAYES	219/94	0,39	0,91	0,43	0,64
HARMONICS O normal (_1) vs O flexion (_3)		H2O RANDOM FOREST	219/94	0,40	0,66	0,35	0,51
HARMONICS O normal (_1) vs O extension (_2)		H2O NAIVE BAYES	219/94	0,44	0,89	0,36	
HARMONICS O flexion(_3) vs O extension (_2)		H2O GRADIENT BOOSTING	219/94	0,39	0,91	0,44	0,70
HARMONICS O all positions		H2O RANDOM FOREST	219/94	0,41	0,55	0,40	0,50
HARMONICS U normal (_1) vs U flexion (_3)		H2O NAIVE BAYES	219/94	0,36	0,74	0,44	0,63
HARMONICS U normal (_1) vs U extension (_2)		H2O GRADIENT BOOSTING	219/94	0,36	0,69	0,38	0,56
HARMONICS U flexion (_3) vs U extension (_2)		H2O RANDOM FOREST	219/94	0,35	0,69	0,39	0,50
HARMONICS U all positions		H2O RANDOM FOREST	219/94	0,35	0,68	0,39	0,50
Spectral+ Descriptive data all vocals		H2O GRADIENT BOOSTING	219/94	0,42	0,57	0,43	0,51
SPectral+ Descriptive data A normal		H2O RANDOM FOREST	219/94	0,53	0,48	0,48	0,46
SPectral+ Descriptive data O extension		H2O RANDOM FOREST	219/94	0,53	0,63	0,43	0,49

DATASET (PARAMETERS COMBINATION)	CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				F1	Log Loss	F1	Log Loss
SPectral+ Descriptive data O flexion		H2O NAIVE BAYES	219/94	0,52	0,34	0,47	
Spectral+ Descriptive data A flexion vs A extension		H2O RANDOM FOREST	219/94	0,52	0,54	0,48	0,46
SPectral+ Descriptive data E flexion vs E extension		H2O NAIVE BAYES	219/94	0,52	0,34	0,44	
Voice parameters + Descriptive data all vocals		H2O GENERALISED LINEAR MODEL	219/94	0,44	0,65	0,42	0,51
Voice parameters + Descriptive data A normal		H2O GENERALISED LINEAR MODEL	219/94	0,49	0,61	0,42	0,53
Voice parameters + Descriptive data O extension		H2O GRADIENT BOOSTING	219/94	0,45	0,34	0,45	
Voice parameters + Descriptive data O flexion		H2O RANDOM FOREST	219/94	0,58	0,53	0,47	0,48
Voice parameters + Descriptive data A flexion vs A extension		H2O GENERALISED LINEAR MODEL	219/94	0,54	0,60	0,43	0,53
Voice parameters + Descriptive data E flexion vs E extension		H2O NAIVE BAYES	219/94	0,51	0,84	0,43	0,98
Voice parameters + Descriptive data O flexion vs O extension		H2O NAIVE BAYES	219/94	0,50	0,88	0,43	
Voice parameters + Descriptive data all vocals	Only Cormack 1 and 4	H2O RANDOM FOREST	100/43	0,58	0,35	0,57	0,33
Voice parameters + Descriptive data A normal	Only Cormack 1 and 4	H2O GENERALISED LINEAR MODEL	100/43	0,44	0,50	0,50	0,40
Voice parameters + Descriptive data O extension	Only Cormack 1 and 4	H2O GENERALISED LINEAR MODEL	100/43	0,47	0,50	0,53	0,43
Voice parameters + Descriptive data O flexion	Only Cormack 1 and 4	H2O NAIVE BAYES	100/43	0,71	0,86	0,50	
Voice parameters + Descriptive data A flexion vs A extension	Only Cormack 1 and 4	H2O GENERALISED LINEAR MODEL	100/43	0,49	0,48	0,53	0,40
Voice parameters + Descriptive data E flexion vs E extension	Only Cormack 1 and 4	H2O GRADIENT BOOSTING	100/43	0,35	0,34	0,67	0,53

DATASET (PARAMETERS COMBINATION)	CASES	BEST MODEL	N PATIENTS	TRAINING 70%		TESTING 30%	
				F1	Log Loss	F1	Log Loss
Voice parameters + Descriptive data O flexion vs O extension	Only Cormack 1 and 4	H2O NAIVE BAYES	100/43	0,76	0,88	0,50	

11.16 ANNEX 16. Full Table 17. Performance of classification models: Combination parameters and group of analyzed patients, type of classification algorithm used, number of participants (N) used in both training and testing groups, for true negatives, false negatives, false positives, true positive values for the training group and the testing group respectively.

DATASET (PARAMETERS COMBINATION)	CASES	BEST MODEL	N PATIENTS	TESTING 30%			
				TN	FN	FP	TP
HARMONICS + Descriptive data All vocals		H2O GENERALISED LINEAR MODEL	219/94	29	2	46	17
HARMONICS + Descriptive data A normal (_1)		H2O GENERALISED LINEAR MODEL	219/94	43	4	32	15
HARMONICS + Descriptive data A extension (_2)		H2O GRADIENT BOOSTING	219/94	51	5	24	14
HARMONICS + Descriptive data A flexion (_3)		H2O GRADIENT BOOSTING	219/94	41	2	34	17
HARMONICS + Descriptive data E normal (_1)		H2O GRADIENT BOOSTING	219/94	38	1	37	18
HARMONICS + Descriptive data E extension (_2)		H2O NAIVE BAYES	219/94	37	2	38	17
HARMONICS + Descriptive data E flexion(_3)		H2O NAIVE BAYES	219/94	41	4	34	15
HARMONICS + Descriptive data I normal (_1)		H2O RANDOM FOREST	219/94	39	1	36	18
HARMONICS + Descriptive data I extension (_2)		H2O RANDOM FOREST	219/94	39	2	36	17
HARMONICS + Descriptive data I flexion(_3)		H2O NAIVE BAYES	219/94	45	4	30	15
HARMONICS + Descriptive data O normal (_1)		H2O RANDOM FOREST	219/94	32	1	43	18
HARMONICS + Descriptive data O extension (_2)		H2O RANDOM FOREST	219/94	41	3	34	16
HARMONICS + Descriptive data O flexion(_3)		H2O NAIVE BAYES	219/94	31	1	44	18
HARMONICS + Descriptive data U normal (_1)		H2O RANDOM FOREST	219/94	42	1	33	18
HARMONICS + Descriptive data U extension 2)		H2O RANDOM FOREST	219/94	37	1	38	18
HARMONICS + Descriptive data U Flexion (_3)		H2O RANDOM FOREST	219/94	40	3	35	16

DATASET (PARAMETERS COMBINATION)	CASES	BEST MODEL	N PATIENTS	TESTING 30%			
				TN	FN	FP	TP
HARMONICS + Descriptive data A normal (_1) vs A flexion (_3)		H2O RANDOM FOREST	219/94	51	5	24	14
HARMONICS + Descriptive data A normal (_1) vs A extension (_2)		H2O RANDOM FOREST	219/94	46	4	29	15
HARMONICS + Descriptive data A flexion (_3) vs A extension (_2)		H2O RANDOM FOREST	219/94	41	2	34	17
HARMONICS + Descriptive data A all positions		H2O RANDOM FOREST	219/94	34	1	41	18
HARMONICS + Descriptive data E normal (_1) vs E flexion (_3)		H2O NAIVE BAYES	219/94	38	2	37	17
HARMONICS + Descriptive data E normal (_1) vs E extension(_2)		H2O RANDOM FOREST	219/94	30	1	45	18
HARMONICS + Descriptive data E flexion (_2) vs E extension (_3)		H2O NAIVE BAYES	219/94	38	3	37	16
HARMONICS + Descriptive data E all positions		H2O NAIVE BAYES	219/94	39	3	36	16
HARMONICS + Descriptive data I normal (_1) vs I flexion (_3)		H2O GRADIENT BOOSTING	219/94	43	1	32	18
HARMONICS + Descriptive data I normal (_1) vs I extension(_2)		H2O GRADIENT BOOSTING	219/94	48	4	27	15
HARMONICS + Descriptive data I flexion (_3) vs I extension (_2)		H2O RANDOM FOREST	219/94	41	2	34	17
HARMONICS + Descriptive data I all positions		H2O NAIVE BAYES	219/94	37	2	38	17
HARMONICS + Descriptive data O normal (_1) vs O flexion (_3)		H2O GRADIENT BOOSTING	219/94	32	1	43	18
HARMONICS + Descriptive data O normal (_1) vs O extension (_2)		H2O GENERALISED LINEAR MODEL	219/94	50	5	25	14

DATASET (PARAMETERS COMBINATION)	CASES	BEST MODEL	N PATIENTS	TESTING 30%			
				TN	FN	FP	TP
HARMONICS + Descriptive data O flexion(_3) vs O extension (_2)		H2O GRADIENT BOOSTING	219/94	32	2	43	17
HARMONICS + Descriptive data O all positions		H2O GRADIENT BOOSTING	219/94	37	2	38	17
HARMONICS + Descriptive data U normal (_1) vs U flexion (_3)		H2O RANDOM FOREST	219/94	30	1	45	18
HARMONICS + Descriptive data U normal (_1) vs U extension (_2)		H2O RANDOM FOREST	219/94	39	4	36	15
HARMONICS + Descriptive data U flexion (_3) vs U extension (_2)		H2O NAIVE BAYES	219/94	50	8	25	11
HARMONICS + Descriptive data U all positions		H2O NAIVE BAYES	219/94	42	5	33	14
HARMONICS+ decriptive data all vocals	Overweigh Cormack 4	H2O RANDOM FOREST	219/94	28	3	47	16
HARMONICS + Descriptive data A normal (_1)	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	29	1	46	18
HARMONICS + Descriptive data A extension (_2)	Overweigh Cormack 4	H2O GRADIENT BOOSTING	219/94	37	3	38	16
HARMONICS + Descriptive data A flexion (_3)	Overweigh Cormack 4	H2O RANDOM FOREST	219/94	27	3	48	16
HARMONICS + Descriptive data E normal (_1)	Overweigh Cormack 4	H2O RANDOM FOREST	219/94	30	3	45	16
HARMONICS + Descriptive data I normal (_1)	Overweigh Cormack 4	H2O GRADIENT BOOSTING	219/94	28	4	47	15
HARMONICS + Descriptive data I extension (_2)	Overweigh Cormack 4	H2O GRADIENT BOOSTING	219/94	17	2	58	17
HARMONICS + Descriptive data O normal (_1)	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	31	1	44	18
HARMONICS + Descriptive data O extension (_2)	Overweigh Cormack 4	H2O GRADIENT BOOSTING	219/94	18	1	57	18
HARMONICS + Descriptive data O flexion(_3)	Overweigh Cormack 4	H2O GRADIENT BOOSTING	219/94	26	4	49	15
HARMONICS + Descriptive data U normal (_1)	Overweigh Cormack 4	H2O NAIVE BAYES	219/94	25	1	50	18
HARMONICS + Descriptive data U extension 2)	Overweigh Cormack 4	H2O RANDOM FOREST	219/94	13	1	62	18

DATASET (PARAMETERS COMBINATION)	CASES	BEST MODEL	N PATIENTS	TESTING 30%			
				TN	FN	FP	TP
HARMONICS + Descriptive data U Flexion (_3)	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	26	0	49	19
HARMONICS + Descriptive data A normal (_1) vs A flexion (_3)	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	26	0	49	19
HARMONICS + Descriptive data A normal (_1) vs A extension (_2)	Overweigh Cormack 4	H2O NAIVE BAYES	219/94	28	1	47	18
HARMONICS + Descriptive data A flexion (_3) vs A extension (_2)	Overweigh Cormack 4	H2O NAIVE BAYES	219/94	25	2	50	17
HARMONICS + Descriptive data A all positions	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	26	0	49	19
HARMONICS + Descriptive data E normal (_1) vs E flexion (_3)	Overweigh Cormack 4	H2O NAIVE BAYES	219/94	31	2	44	17
HARMONICS + Descriptive data E normal (_1) vs E extension(_2)	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	45	5	30	14
HARMONICS + Descriptive data E flexion (_2) vs E extension (_3)	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	28	1	47	18
HARMONICS + Descriptive data E all positions	Overweigh Cormack 4	H2O NAIVE BAYES	219/94	44	6	31	13
HARMONICS + Descriptive data I normal (_1) vs I flexion (_3)	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	27	0	48	19
HARMONICS + Descriptive data I normal (_1) vs I extension(_2)	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	27	0	48	19
HARMONICS + Descriptive data I all positions	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	27	0	48	19
HARMONICS + Descriptive data U normal (_1) vs U flexion (_3)	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	26	0	49	19
HARMONICS + Descriptive data U normal (_1) vs U extension (_2)	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	26	0	49	19
HARMONICS + Descriptive data U flexion (_3) vs U extension (_2)	Overweigh Cormack 4	H2O GENERALISED LINEAR MODEL	219/94	24	0	51	19

DATASET (PARAMETERS COMBINATION)	CASES	BEST MODEL	N PATIENTS	TESTING 30%			
				TN	FN	FP	TP
HARMONICS + Descriptive data U all positions	Overweigh Cormack 4	H2O GRADIENT BOOSTING	219/94	24	2	51	17
HARMONICS+ decriptive data all vocals	Cormack 1 and 4 cases	H2O GRADIENT BOOSTING	100/43	38	2	0	3
HARMONICS + Descriptive data A normal (_1)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	32	1	6	4
HARMONICS + Descriptive data A extension (_2)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	31	1	7	4
HARMONICS + Descriptive data A flexion (_3)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	31	1	7	4
HARMONICS + Descriptive data E normal (_1)	Cormack 1 and 4 cases	H2O RANDOM FOREST	100/43	34	1	4	4
HARMONICS + Descriptive data E extension (_2)	Cormack 1 and 4 cases	H2O GRADIENT BOOSTING	100/43	33	1	5	4
HARMONICS + Descriptive data E flexion(_3)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	30	1	8	4
HARMONICS + Descriptive data I normal (_1)	Cormack 1 and 4 cases	H2O RANDOM FOREST	100/43	32	1	6	4
HARMONICS + Descriptive data I extension (_2)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	29	1	9	4
HARMONICS + Descriptive data I flexion(_3)	Cormack 1 and 4 cases	H2O GRADIENT BOOSTING	100/43	36	1	2	4
HARMONICS + Descriptive data O normal (_1)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	33	1	5	4
HARMONICS + Descriptive data O extension (_2)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	30	1	8	4
HARMONICS + Descriptive data O flexion(_3)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	29	1	9	4
HARMONICS + Descriptive data U normal (_1)	Cormack 1 and 4 cases	H2O GRADIENT BOOSTING	100/43	34	1	4	4
HARMONICS + Descriptive data U extension 2)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	32	1	6	4
HARMONICS + Descriptive data U Flexion (_3)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	29	1	9	4
HARMONICS + Descriptive data A normal (_1) vs A flexion (_3)	Cormack 1 and 4 cases	H2O RANDOM FOREST	100/43	31	1	7	4
HARMONICS + Descriptive data A normal (_1) vs A extension (_2)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	31	1	7	4

DATASET (PARAMETERS COMBINATION)	CASES	BEST MODEL	N PATIENTS	TESTING 30%			
				TN	FN	FP	TP
HARMONICS + Descriptive data A flexion (_3) vs A extension (_2)	Cormack 1 and 4 cases	H2O GRADIENT BOOSTING	100/43	36	1	2	4
HARMONICS + Descriptive data A all positions	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	37	2	1	3
HARMONICS + Descriptive data E normal (_1) vs E flexion (_3)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	32	1	6	4
HARMONICS + Descriptive data E normal (_1) vs E extension(_2)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	29	1	9	4
HARMONICS + Descriptive data E flexion (_2) vs E extension (_3)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	30	1	8	4
HARMONICS + Descriptive data E all positions	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	31	1	7	4
HARMONICS + Descriptive data I normal (_1) vs I flexion (_3)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	30	1	8	4
HARMONICS + Descriptive data I normal (_1) vs I extension(_2)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	29	1	9	4
HARMONICS + Descriptive data I flexion (_3) vs I extension (_2)	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	29	1	9	4
HARMONICS + Descriptive data I all positions	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	30	1	8	4
HARMONICS + Descriptive data O normal (_1) vs O flexion (_3)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	31	1	7	4
HARMONICS + Descriptive data O normal (_1) vs O extension (_2)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	29	1	9	4
HARMONICS + Descriptive data O flexion(_3) vs O extension (_2)	Cormack 1 and 4 cases	H2O RANDOM FOREST	100/43	27	1	11	4
HARMONICS + Descriptive data O all positions	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	30	1	8	4
HARMONICS + Descriptive data U normal (_1) vs U flexion (_3)	Cormack 1 and 4 cases	H2O RANDOM FOREST	100/43	32	1	6	4

DATASET (PARAMETERS COMBINATION)	CASES	BEST MODEL	N PATIENTS	TESTING 30%			
				TN	FN	FP	TP
HARMONICS + Descriptive data U normal (_1) vs U extension (_2)	Cormack 1 and 4 cases	H2O GENERALISED LINEAR MODEL	100/43	32	1	6	4
HARMONICS + Descriptive data U flexion (_3) vs U extension (_2)	Cormack 1 and 4 cases	H2O RANDOM FOREST	100/43	30	1	8	4
HARMONICS + Descriptive data U all positions	Cormack 1 and 4 cases	H2O NAIVE BAYES	100/43	31	1	7	4
ONLY descriptive data		H2O RANDOM FOREST	219/94	44	6	31	13
HARMONICS All vocals		H2O NAIVE BAYES	219/94	31	3	44	16
HARMONICS A normal (_1)		H2O NAIVE BAYES	219/94	42	3	33	16
HARMONICS A extension (_2)		H2O GENERALISED LINEAR MODEL	219/94	14	1	61	18
HARMONICS A flexion (_3)		H2O NAIVE BAYES	219/94	53	6	22	13
HARMONICS E normal (_1)		H2O GENERALISED LINEAR MODEL	219/94	48	5	27	14
HARMONICS E extension (_2)		H2O RANDOM FOREST	219/94	38	5	37	14
HARMONICS E flexion(_3)		H2O GRADIENT BOOSTING	219/94	43	6	32	13
HARMONICS I normal (_1)		H2O NAIVE BAYES	219/94	38	5	37	14
HARMONICS I extension (_2)		H2O GRADIENT BOOSTING	219/94	35	5	40	14
HARMONICS I flexion(_3)		H2O NAIVE BAYES	219/94	36	3	39	16
HARMONICS O normal (_1)		H2O GRADIENT BOOSTING	219/94	36	2	39	17
HARMONICS O extension (_2)		H2O GENERALISED LINEAR MODEL	219/94	11	0	64	19
HARMONICS O flexion(_3)		H2O GRADIENT BOOSTING	219/94	30	3	45	16
HARMONICS U normal (_1)		H2O NAIVE BAYES	219/94	37	3	38	16
HARMONICS U extension 2)		H2O NAIVE BAYES	219/94	14	1	61	18
HARMONICS U Flexion (_3)		H2O RANDOM FOREST	219/94	22	1	53	18
HARMONICS A normal (_1) vs A flexion (_3)		H2O GENERALISED LINEAR MODEL	219/94	26	2	49	17
HARMONICS A normal (_1) vs A extension (_2)		H2O NAIVE BAYES	219/94	41	4	34	15

DATASET (PARAMETERS COMBINATION)	CASES	BEST MODEL	N PATIENTS	TESTING 30%			
				TN	FN	FP	TP
HARMONICS A flexion (_3) vs A extension (_2)		H2O NAIVE BAYES	219/94	49	3	26	16
HARMONICS A all positions		H2O NAIVE BAYES	219/94	55	6	20	13
HARMONICS E normal (_1) vs E flexion (_3)		H2O GENERALISED LINEAR MODEL	219/94	60	7	15	12
HARMONICS E normal (_1) vs E extension(_2)		H2O GENERALISED LINEAR MODEL	219/94	23	1	52	18
HARMONICS E flexion (_2) vs E extension (_3)		H2O GRADIENT BOOSTING	219/94	41	6	34	13
HARMONICS E all positions		H2O NAIVE BAYES	219/94	48	6	27	13
HARMONICS I normal (_1) vs I flexion (_3)		H2O NAIVE BAYES	219/94	32	3	43	16
HARMONICS I normal (_1) vs I extension(_2)		H2O NAIVE BAYES	219/94	48	6	27	13
HARMONICS I flexion (_3) vs I extension (_2)		H2O GRADIENT BOOSTING	219/94	28	3	47	16
HARMONICS I all positions		H2O NAIVE BAYES	219/94	43	5	32	14
HARMONICS O normal (_1) vs O flexion (_3)		H2O RANDOM FOREST	219/94	9	1	66	18
HARMONICS O normal (_1) vs O extension (_2)		H2O NAIVE BAYES	219/94	26	4	49	15
HARMONICS O flexion(_3) vs O extension (_2)		H2O GRADIENT BOOSTING	219/94	51	7	24	12
HARMONICS O all positions		H2O RANDOM FOREST	219/94	62	11	13	8
HARMONICS U normal (_1) vs U flexion (_3)		H2O NAIVE BAYES	219/94	31	1	44	18
HARMONICS U normal (_1) vs U extension (_2)		H2O GRADIENT BOOSTING	219/94	21	2	54	17
HARMONICS U flexion (_3) vs U extension (_2)		H2O RANDOM FOREST	219/94	36	5	39	14
HARMONICS U all positions		H2O RANDOM FOREST	219/94	27	3	48	16
Spectral+ Descriptive data all vocals		H2O GRADIENT BOOSTING	219/94	47	6	28	13
SPectral+ Descriptive data A normal		H2O RANDOM FOREST	219/94	47	4	28	15
SPectral+ Descriptive data O extension		H2O RANDOM FOREST	219/94	43	5	32	14
SPectral+ Descriptive data O flexion		H2O NAIVE BAYES	219/94	48	5	27	14

DATASET (PARAMETERS COMBINATION)	CASES	BEST MODEL	N PATIENTS	TESTING 30%			
				TN	FN	FP	TP
Spectral+ Descriptive data A flexion vs A extension		H2O RANDOM FOREST	219/94	40	2	35	17
SPectral+ Descriptive data E flexion vs E extension		H2O NAIVE BAYES	219/94	48	6	27	13
Voice parameters + Descriptive data all vocals		H2O GENERALISED LINEAR MODEL	219/94	23	0	52	19
Voice parameters + Descriptive data A normal		H2O GENERALISED LINEAR MODEL	219/94	49	7	26	12
Voice parameters + Descriptive data O extension		H2O GRADIENT BOOSTING	219/94	46	5	29	14
Voice parameters + Descriptive data O flexion		H2O RANDOM FOREST	219/94	55	7	20	12
Voice parameters + Descriptive data A flexion vs A extension		H2O GENERALISED LINEAR MODEL	219/94	35	3	40	16
Voice parameters + Descriptive data E flexion vs E extension		H2O NAIVE BAYES	219/94	31	2	44	17
Voice parameters + Descriptive data O flexion vs O extension		H2O NAIVE BAYES	219/94	31	2	44	17
Voice parameters + Descriptive data all vocals	Only Cormack 1 and 4	H2O RANDOM FOREST	100/43	33	1	5	4
Voice parameters + Descriptive data A normal	Only Cormack 1 and 4	H2O GENERALISED LINEAR MODEL	100/43	31	1	7	4
Voice parameters + Descriptive data O extension	Only Cormack 1 and 4	H2O GENERALISED LINEAR MODEL	100/43	32	1	6	4
Voice parameters + Descriptive data O flexion	Only Cormack 1 and 4	H2O NAIVE BAYES	100/43	31	1	7	4
Voice parameters + Descriptive data A flexion vs A extension	Only Cormack 1 and 4	H2O GENERALISED LINEAR MODEL	100/43	32	1	6	4
Voice parameters + Descriptive data E flexion vs E extension	Only Cormack 1 and 4	H2O GRADIENT BOOSTING	100/43	35	1	3	4
Voice parameters + Descriptive data O flexion vs O extension	Only Cormack 1 and 4	H2O NAIVE BAYES	100/43	31	1	7	4

10.17. ANNEX 17. KNIME workflow