

Análisis-TMO&AOS

Sebastian Gaviria Sánchez

Estadísticos de resumen

Se tiene gran interés en responder los siguientes cuestionamientos:

- ¿Cuánto se demora en promedio un paciente para completar cada nivel?
- ¿Cuánto se demora en promedio un paciente para terminar el tratamiento?

Para dar respuesta a la primera pregunta se muestra la siguiente tabla

Table 1: Días promedio por nivel del tratamiento

Nivel	Días promedio
TMO NIVEL 1	21.2
TMO NIVEL 2	15.1
TMO NIVEL 3	20.4
TMO NIVEL 4	17.0
TMO NIVEL 5	20.1
TMO NIVEL 6	27.7

Con esta última se tienen las siguientes conclusiones:

- En promedio, los pacientes bajo el tratamiento tardan aproximadamente 21 días en completar el nivel 1 del mismo.
- Así mismo, para el nivel 2 de dicho tratamiento, los pacientes tardan aproximadamente 15 días para completarlo en promedio.
- Similar al nivel 1, en el nivel 3 los pacientes bajo el tratamiento tardan en promedio 20 días en completar el mismo.
- Para el nivel 4, se puede decir que los pacientes en promedio se tardan 17 días en llegar al final de este nivel.
- De manera análoga al nivel 3, en el nivel 5 los pacientes también tardan en promedio 20 días aproximadamente.
- Finalmente y a diferencia de los demás niveles, en el nivel 6 los pacientes suelen tardar más en promedio, con casi 28 días para terminar este nivel.

Ahora, interesa conocer cuánto se demora un paciente en promedio para finalizar el tratamiento en su totalidad. Para esta causa se muestra el siguiente resumen.

Table 2: Días para terminar el tratamiento

Estadístico	Resultado
Mínimo	86.0
Percentil 25	124.2
Mediana	150.5
Promedio	157.8
Percentil 75	181.0
Máximo	328.0

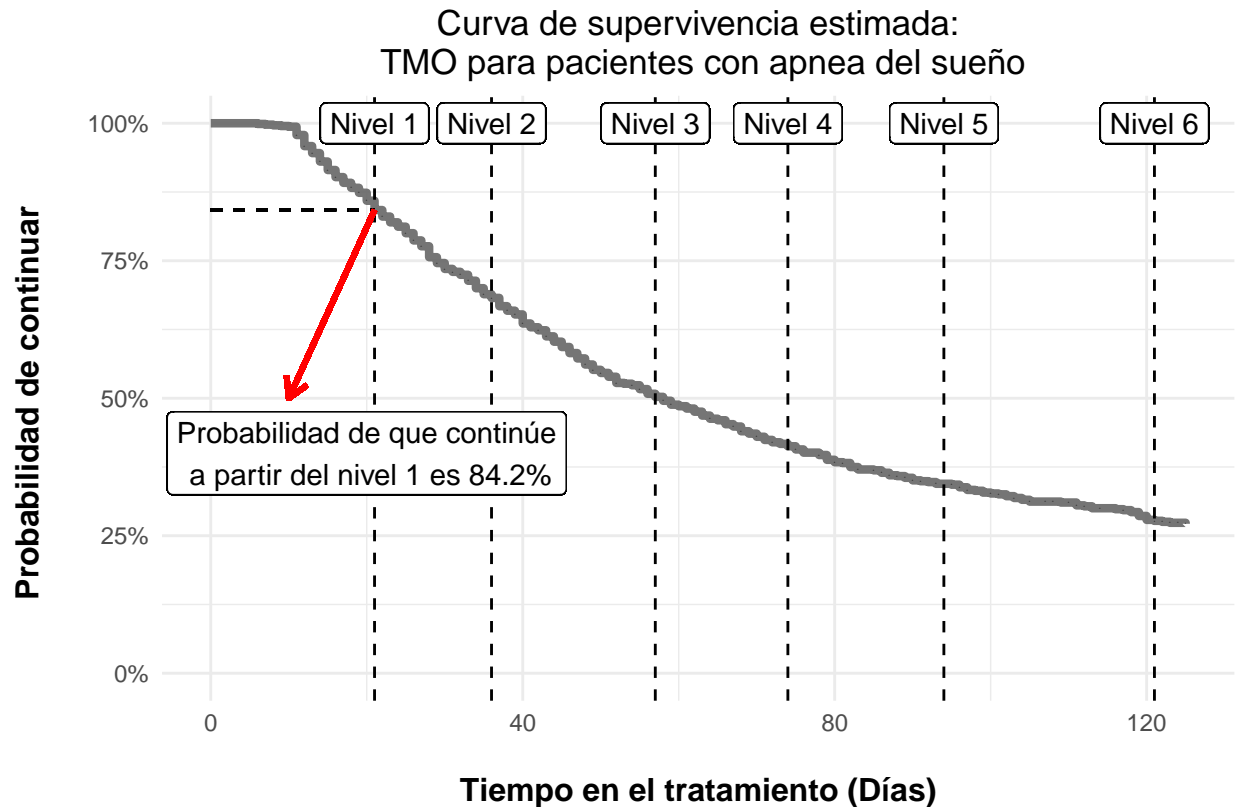
Con este resultado se llega a las siguientes conclusiones:

- El paciente que menos se demoró en terminar el tratamiento lo hizo en 86 días.
- Según el percentil 25, se afirma que el 25% de los pacientes terminan en menos de 125 días.
- Similarmente, según la mediana la mitad de los pacientes termina el tratamiento antes de los 150 días.
- En promedio, los pacientes tardan 158 días en terminar el tratamiento.
- Según el percentil 75, el 75% de los pacientes tarda menos de 181 días en terminar el tratamiento.
- Finalmente, el paciente que más tardó en terminar su tratamiento lo hizo en 328 días.

Curva de supervivencia

En aras de determinar el comportamiento de deserción a lo largo del tratamiento, se desea estimar la curva de supervivencia para observar la probabilidad de que un paciente continúe a partir de un día determinado en el tratamiento o de manera análoga, qué porción de la cohorte inicial va a continuar bajo tratamiento en un momento específico del mismo.

Así pues, se estima la curva de supervivencia $S(t) = P(T > t)$ usando el estimado de Kaplan-Meier. El resultado se ilustra a continuación:



En este gráfico se considera que un paciente se va a demorar por nivel lo obtenido en la sección de estadísticos de resumen, es decir, para el nivel 1 del tratamiento 21 días, para el nivel 2 15 días, para el nivel 3 20 días, para el nivel 4 17 días, para el nivel 5 20 días y para el nivel 6 28 días.

Así entonces, las líneas verticales punteadas demarcan el día esperado para que un paciente termine el nivel correspondiente que se señala en la parte superior.

Dicho esto, se tienen las siguientes conclusiones de interés:

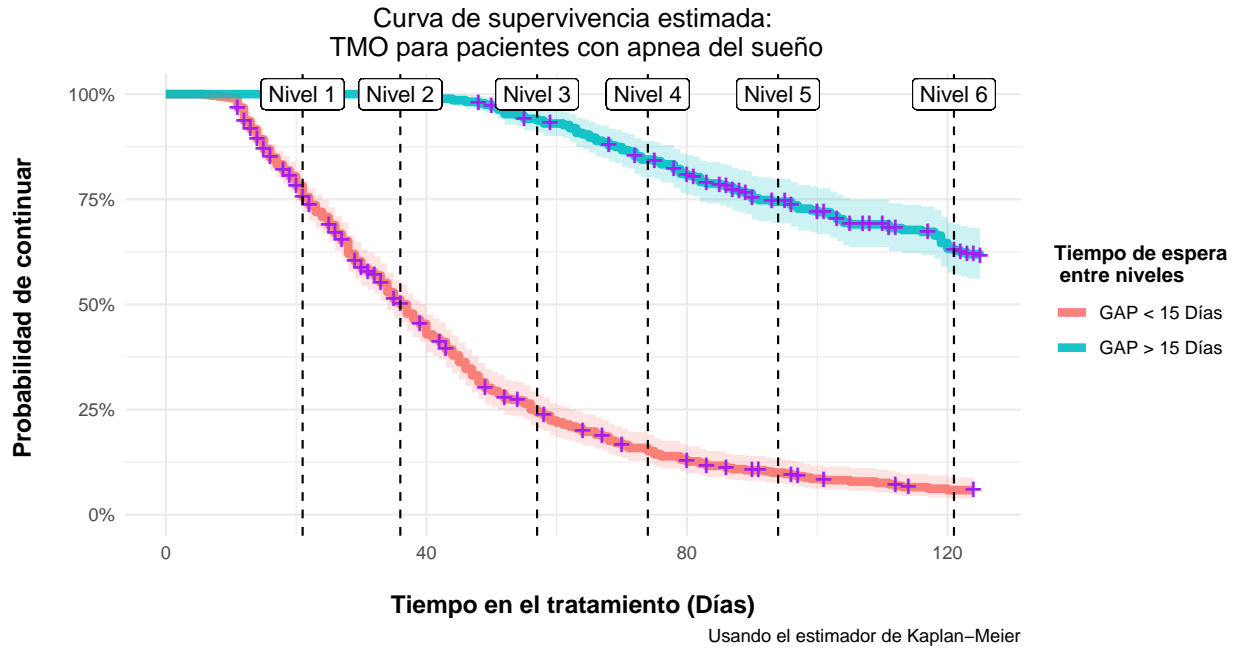
- Un paciente que termine el nivel 1 del tratamiento tendrá una probabilidad del 84.2% de continuar ea nivel 2 o de manera análoga, se espera que el 84.2% de los pacientes que comienzan el tratamiento continúen al nivel 2.
- Similarmente, un paciente que termine el nivel 2 del tratamiento tendrá una probabilidad del 68.2% de continuar al nivel 3 o de manera análoga, se espera que el 68.2% de los pacientes que comienzan el tratamiento continúen al nivel 3.
- De igual forma, un paciente que termine el nivel 3 del tratamiento tendrá una probabilidad del 50.2% de continuar al nivel 4 o de manera análoga, se espera que el 50.2% de los pacientes que comienzan el tratamiento continúen al nivel 4.
- Así mismo, un paciente que termine el nivel 4 del tratamiento tendrá una probabilidad del 41.2% de continuar al nivel 5 o de manera análoga, se espera que el 41.2% de los pacientes que comienzan el tratamiento continúen al nivel 5.
- Igualmente, un paciente que termine el nivel 5 del tratamiento tendrá una probabilidad del 34.4% de continuar al nivel 6 o de manera análoga, se espera que el 34.4% de los pacientes que comienzan el tratamiento continúen al nivel 6.

Diferencia entre grupos de GAP

Una de las dinámicas del tratamiento consiste en programar una cita de control para cada paciente una vez termine uno de los niveles y esté a puertas de comenzar el siguiente; sin embargo estas citas de control en algunas ocasiones suelen programarse en un periodo prolongado, lo que ocasiona una ventana de tiempo considerablemente grande entre niveles del tratamiento. En este caso se considera que una ventana de tiempo mayor a 15 días entre niveles es excesivo e incluso posiblemente perjudicial para los efectos del tratamiento.

Con lo anterior se crea naturalmente el cuestionamiento: ¿Un GAP (ventana) de tiempo excesivo (mayor a quince días) entre niveles ocasiona que los pacientes bajo tratamiento se adhieran menos al mismo?

Para dar respuesta a esta pregunta se clasifican los pacientes de acuerdo a si en algún momento de su tratamiento tuvieron al menos un GAP de tiempo mayor a quince días en algún momento. Habitualmente se crean dos grupos: Aquellos que tuvieron sus citas de control en menos de quince días ($GAP \leq 15$) y aquellos que en algún momento del tratamiento tuvieron un tiempo de espera prolongado entre niveles del mismo ($GAP > 15$). Así pues, para apreciar la adherencia de ambos grupos se grafican sus respectivas curvas de supervivencia como sigue.



Antes de proceder a concluir sobre estas dos curvas, es necesario demostrar estadísticamente si existe una diferencia significativa entre las mismas.

Veamos entonces si existe una diferencia estadísticamente significativa entre las dos curvas de supervivencia ilustradas anteriormente.

Denote $S_1(t)$: Curva de supervivencia para pacientes que no experimentan un GAP grande de tiempo y $S_2(t)$: Curva de supervivencia para pacientes que en algún momento de su tratamiento tuvieron que esperar más de quince días entre niveles.

Así entonces, mediante la prueba Log-Rank se desea juzgar el siguiente juego de hipótesis

$$\begin{cases} H_0 : S_1(t) = S_2(t) \\ H_1 : S_1(t) \neq S_2(t) \end{cases}$$

Con esto, usando un nivel de significancia $\alpha = 0.05$, se obtiene el siguiente resultado

Table 3: Resultados Prueba Log-Rank

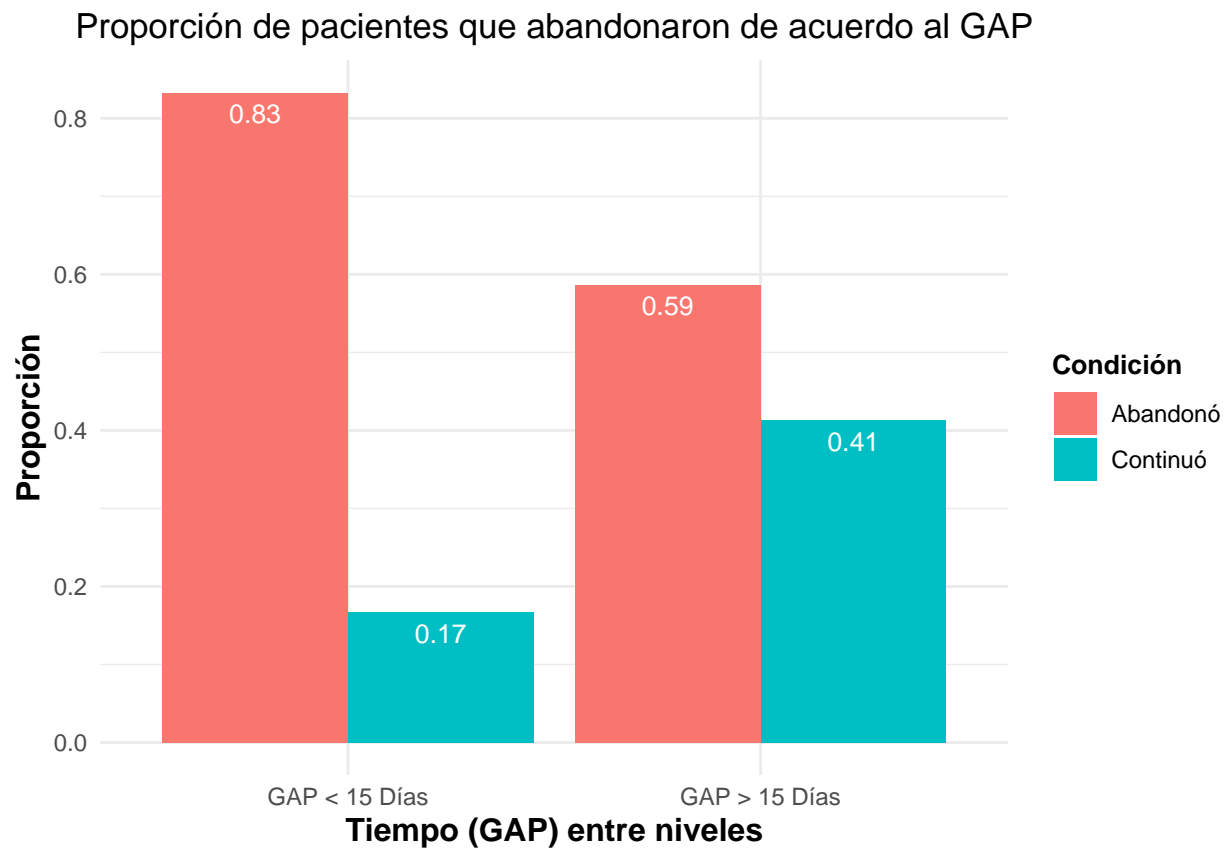
χ^2	Valor P
376.7487	0

Teniendo en cuenta que Valor P $\ll \alpha = 0.05$, se puede rechazar $H_0 : S_1(t) = S_2(t)$ contundentemente, concluyendo así que la supervivencia entre los dos grupos es significativamente diferente.

Así entonces se pueden realizar las siguientes conclusiones.

- La supervivencia de los pacientes que en algún momento de su tratamiento experimentan un GAP de tiempo excesivo parece ser superior a la de los pacientes que no tienen que esperar mucho antes de iniciar su siguiente nivel respectivo, es decir, los pacientes que en algún momento deben esperar mucho por una cita de control parecen adherirse mas al tratamiento, resultado completamente contraintuitivo.
- **Sin embargo, el GAP de tiempo mencionado puede no ser debido exclusivamente al tiempo en que se programa una cita de control. En ocasiones, se puede considerar adecuado que un paciente realice ciertos ejercicios durante un periodo de tiempo antes de comenzar el siguiente nivel, lo cual no es necesariamente negativo, por el contrario puede pensarse que un acompañamiento cercano al paciente lo motiva más y hace que el mismo mejore sus probabilidades de terminar exitosamente el tratamiento.**
- Si bien los pacientes que experimentan un GAP grande de tiempo parecen adherirse más al tratamiento, hay que concluir con cuidado en este aspecto, porque el tiempo prolongado en la programación de estas citas de control pueden estar inflando falsamente la supervivencia de un paciente que más adelante no va a continuar con su tratamiento.

¿Realmente sobreviven más los pacientes con GAP grande?



Note que de acuerdo a este gráfico, para los pacientes que no experimentan un tiempo de espera mayor a quince días entre sus niveles, el 83% de los mismos abandonaron el tratamientos. Sin embargo, para aquellos pacientes que en algún momento del tratamiento tuvieron que esperar más de quince días para iniciar un nuevo nivel, esta proporción es menor, pues en este caso, solo el 59% abandonó el tratamiento.

Análisis descriptivo: Pacientes que llegaron a la meta

Se dispone de una base de datos que contiene información relacionada a 126 pacientes que completaron el tratamiento miofuncional orofacial (TMO) en todos sus niveles. En dicha base de datos se encuentra consignada información relacionada a comorbilidades, condiciones oclusales, diversos índices, edad, sexo etc. de cada uno de los pacientes que completaron el tratamiento. La estructura de la base de datos así como algunas variables se ilustran en la siguiente tabla:

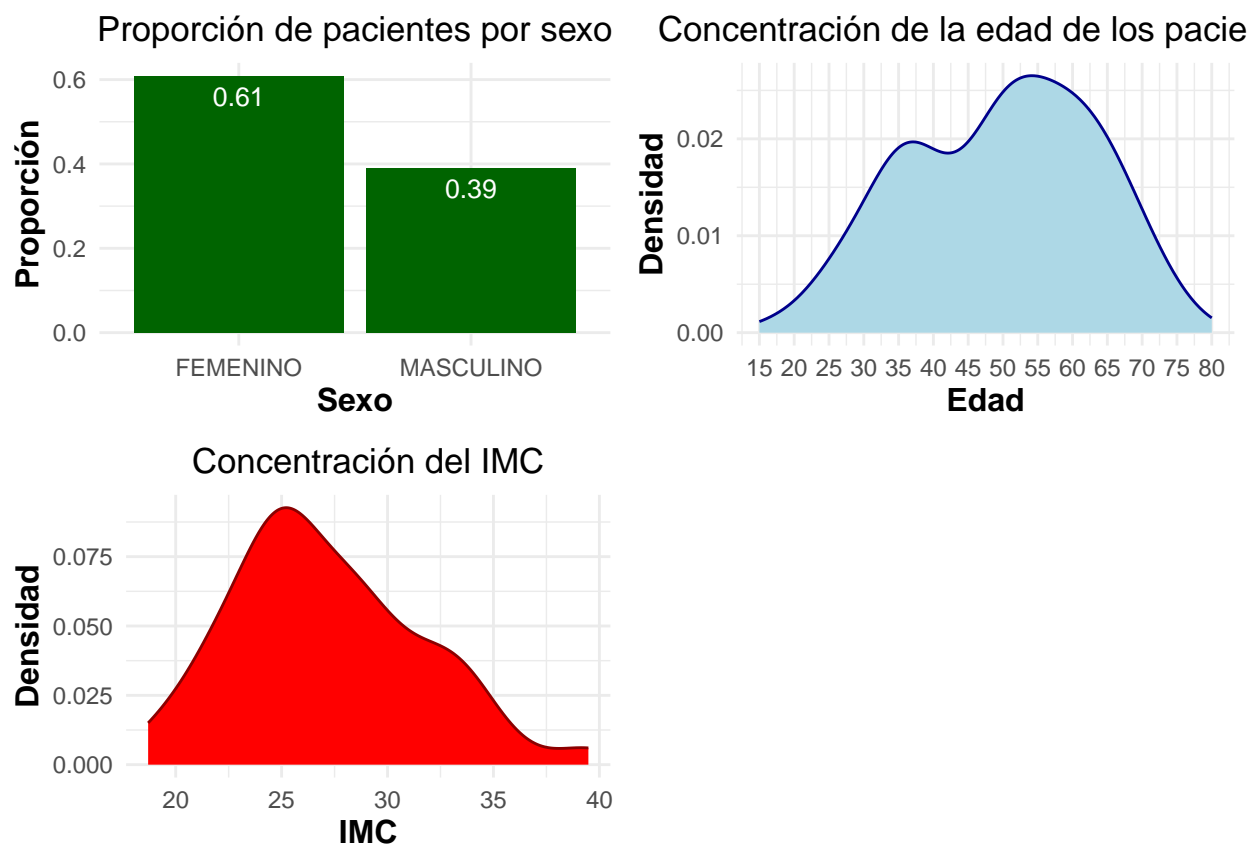
Table 4: Estructura de los datos

ESQUEMA NIVEL 1	EDAD	SEXO	SINTOMAS BASALES	PuntGeneral
TMO NIVEL 1	37	FEMENINO	Cefalea matutina	4.00
TMO NIVEL 1	71	FEMENINO	Ronquido	4.00
TMO NIVEL 1	68	MASCULINO	Nicturia	5.00
TMO NIVEL 1	65	FEMENINO	Cefalea matutina	4.25
TMO NIVEL 1	18	FEMENINO	Cansancio crónico	3.75
TMO NIVEL 1	35	FEMENINO	Cansancio crónico	3.00

En este ejemplo, la variable “PuntGeneral” es una escala continua de 1-5 que indica el nivel de satisfacción/efectividad del tratamiento del paciente, donde 1 indica un paciente que considera inefectivo el tratamiento y 5 indica que el paciente considera el tratamiento absolutamente efectivo.

Variables esquema, sexo, edad e IMC

Así entonces, para dar inicio al análisis descriptivo, se plantean los siguientes gráficos:

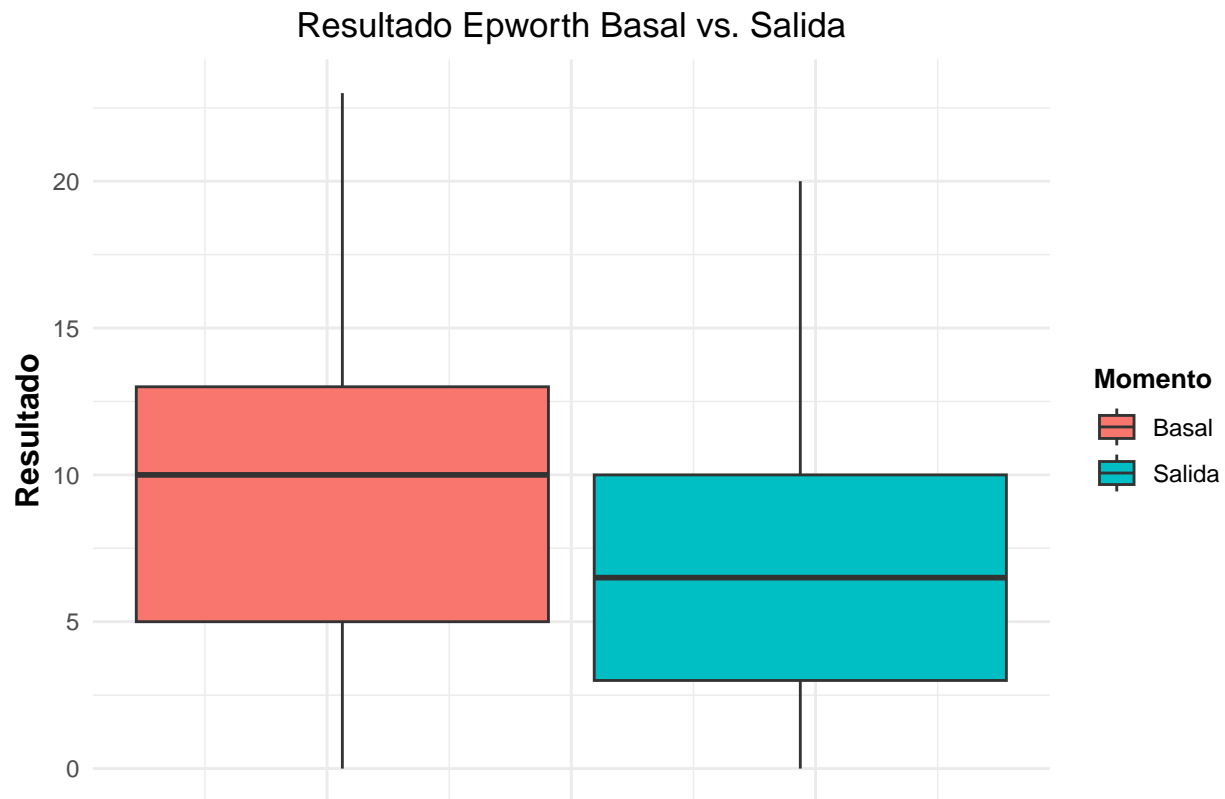


Según estos gráficos se tienen las siguientes apreciaciones:

- El 61% de los pacientes son femeninos, diferencia significativa con la muestra masculina.
- En general, el rango de edad de los pacientes del tratamiento es de los 18 a 75 años, y según el gráfico, hay una concentración destacable de pacientes entre los 30 y 40 años, sin embargo la concentración más apreciable de pacientes se encuentra entre los 50 y 60 años de edad.
- De manera similar, el índice de masa corporal centra usualmente sus valores alrededor del 25, sin embargo se puede notar una pequeña moda alrededor del 32 que podría indicar un porción considerable de pacientes en el rango de la obesidad.

Análisis escala Epworth

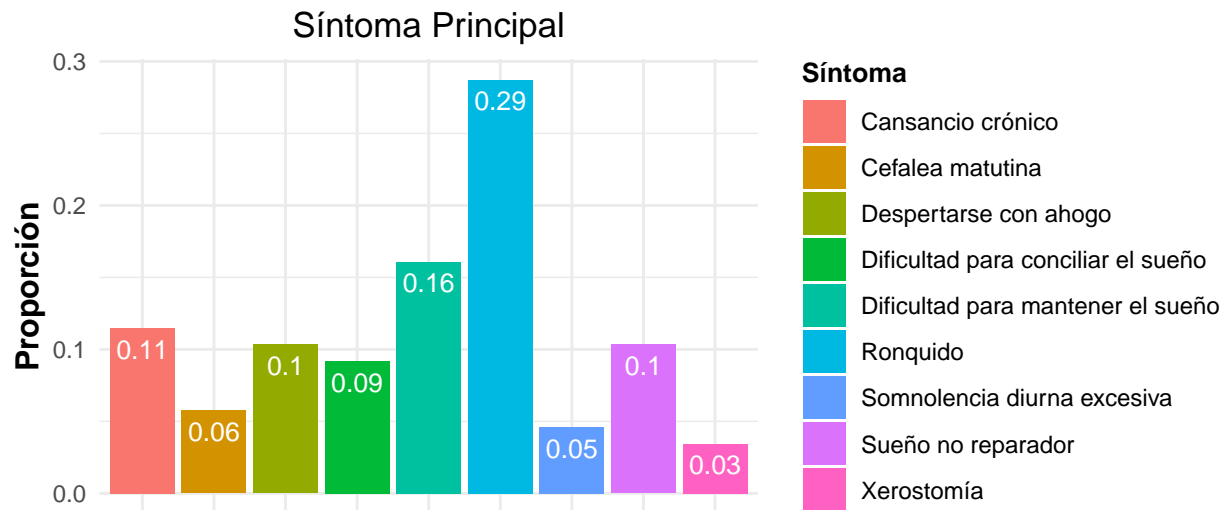
Para lo que respecta a la escala de Epworth, observe el siguiente gráfico:



- De este boxplot se puede apreciar que existe una diferencia aparente en el resultado de la escala de somnolencia entre el momento de entrada y salida de un paciente al tratamiento. En general puede decirse que una vez concluido el tratamiento, los pacientes se autoperciben menos somnolientos respecto al momento de entrada a la terapia miofuncional orofacial.

Síntomas reportados

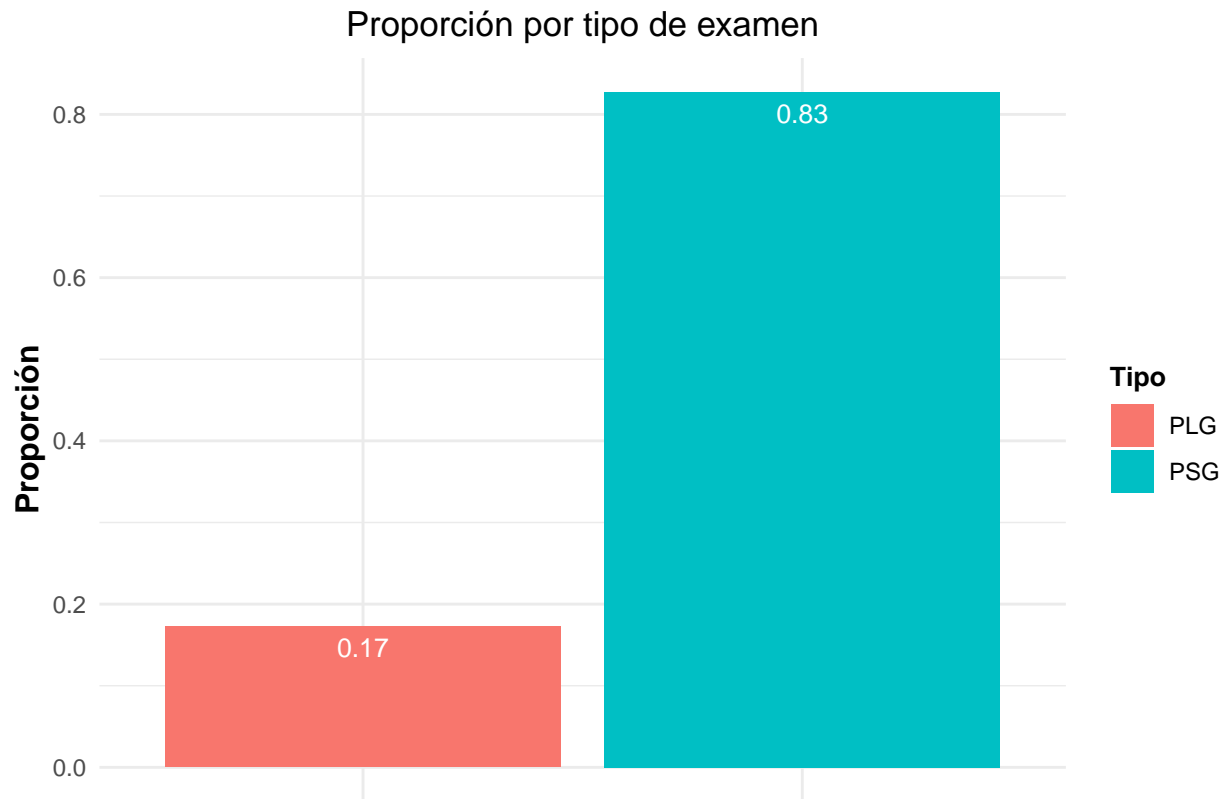
Ahora, en relación a los síntomas que reportan los pacientes al momento de la terapia, observe los siguientes gráficos:



- De este gráfico se aprecia que en este caso, el 29% de los pacientes definieron el ronquido como su síntoma principal. Otro síntoma destacable en este caso es la dificultad para mantener el sueño, puesto el 16% de las veces como síntoma principal.

Tipo de examen de sueño

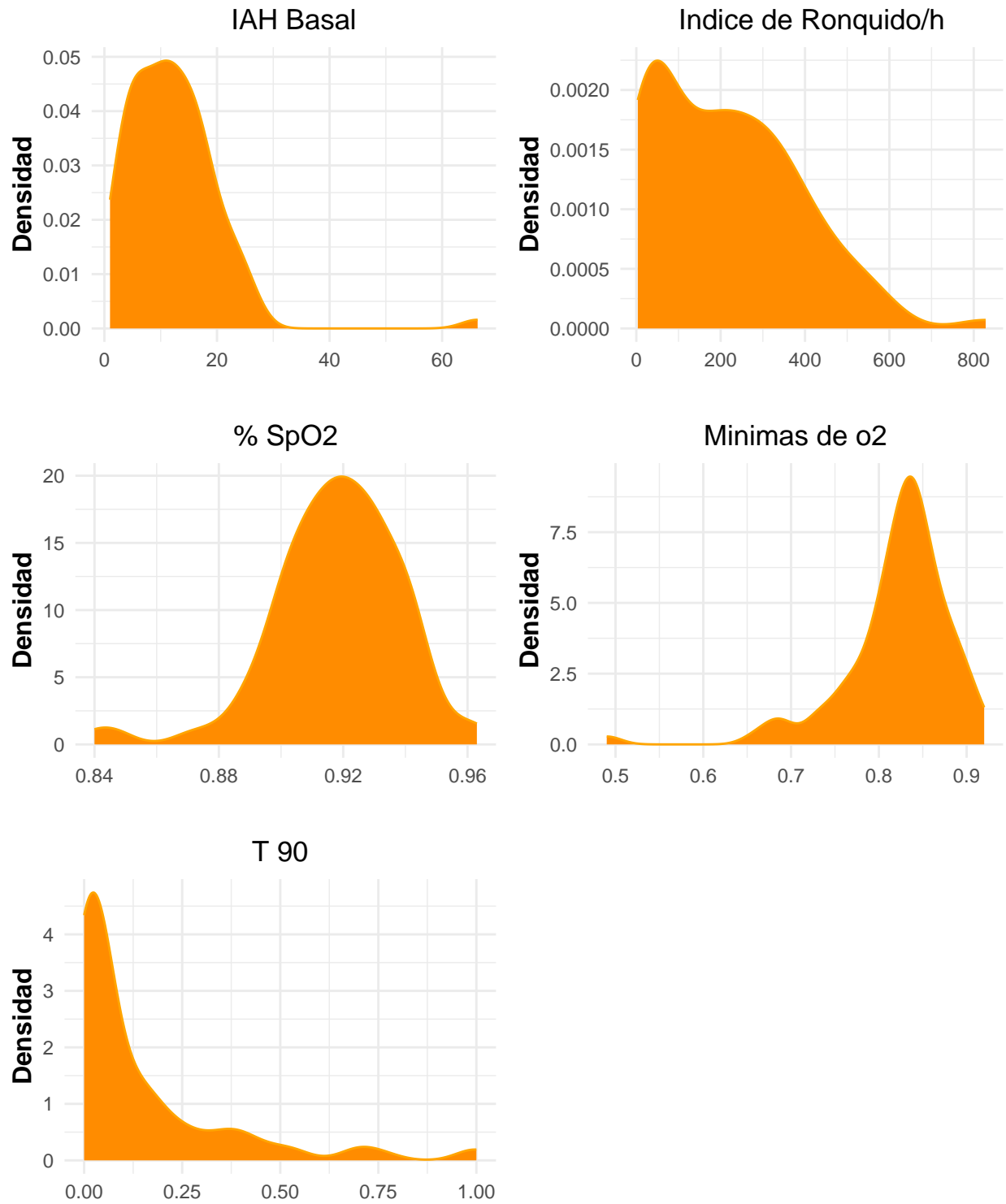
Observe ahora el siguiente gráfico en relación al tipo de examen de sueño



- La mayoría de pacientes que ingresan al tratamiento se les practica una Polisomnografía como examen del sueño, abarcando el 83% de los casos.

Índices y variables numéricas

Se plantean ahora los siguientes gráficos relativos a índices y variables numéricas



De este conjunto de gráficos se aprecia lo siguiente:

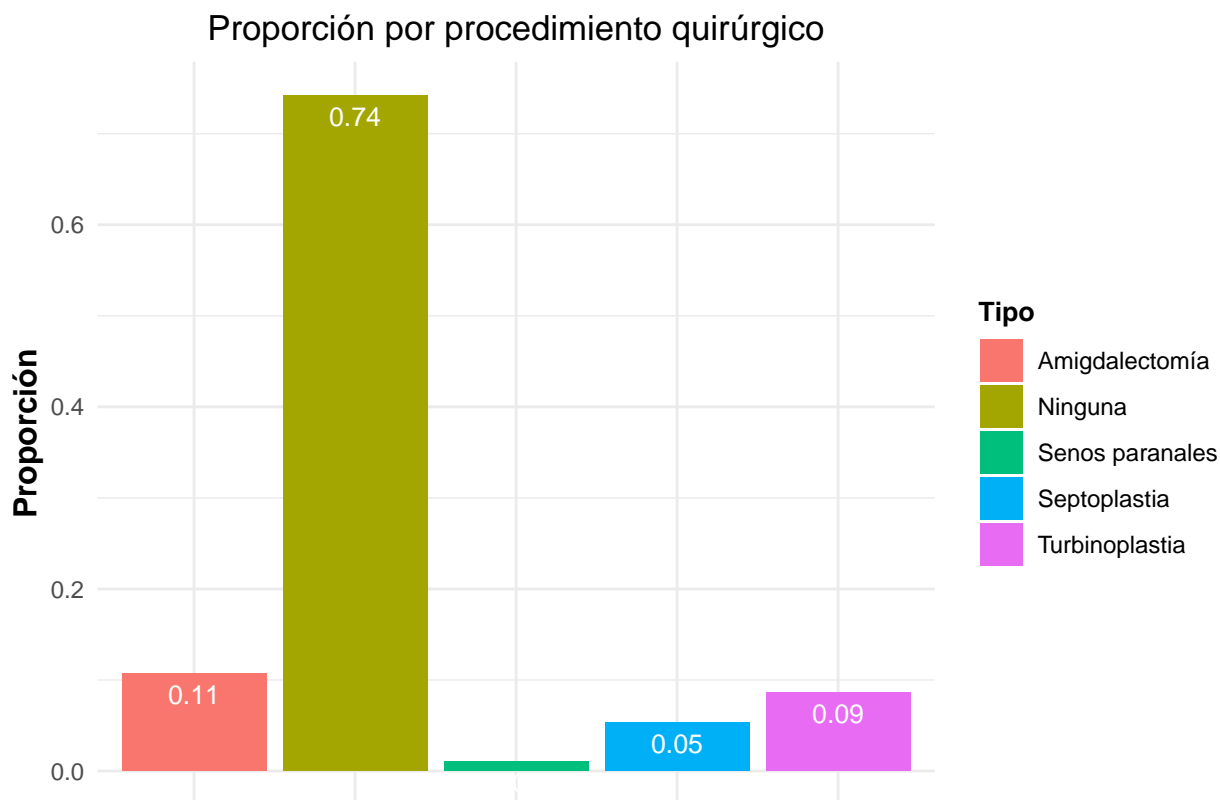
- Los valores más usuales para el índice de apneas hipoapneas de sueño se centran alrededor de 10. La

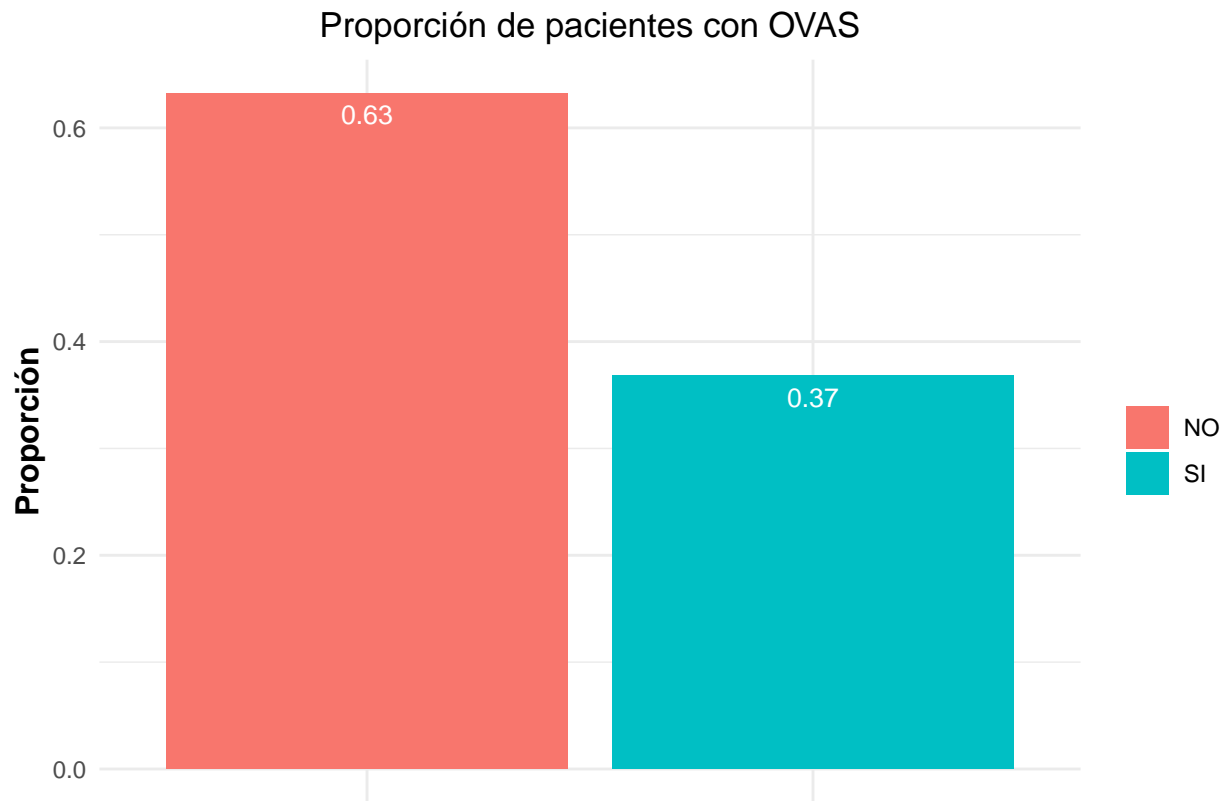
gran mayoría de los pacientes registran este índice entre los valores de 0 y 20; sin embargo existen pacientes que tienen registros más altos llegando al caso extremo de un paciente que registra un índice de 66.3 apneas hipoapneas de sueño.

- La curva de densidad para el índice de ronquido se nota un poco más extendida sobre su rango, donde un poco menos de la mitad de los pacientes registran menos de 100 en este índice de ronquido. Nuevamente se aprecia una cola hacia la derecha de la distribución muestral de esta variable, lo que indica presencia de valores extremos por encima de 500.
- Los porcentajes de SpO2 varían entre 0.84 y 0.96, centrando sus valores (usuales) en el 92%. La mayoría de pacientes en este aspecto registran entre 90% y 94% de saturación promedio por horas de sueño grabadas durante el examen; sin embargo existe una ligera cola a la izquierda de esta distribución que indica valores por debajo del 88% que resultan atípicos según el comportamiento normal de esta variable.
- Una situación similar se aprecia cuando se observa la curva de las mínimas de O₂. En este punto los valores se concentran entre 0.75 y 0.90. Sin embargo y de manera análoga al gráfico anterior, en esta distribución se aprecia una cola a la izquierda lo que indica la presencia de valores extremos y atípicamente pequeños.
- Finalmente, en lo que respecta al porcentaje de tiempo en el que la saturación de oxígeno estuvo por debajo de 90% por el total de horas de sueño grabadas, se puede apreciar que los valores se concentran a la izquierda del 25%, siendo los porcentajes más pequeños los más usuales. Sin embargo se aprecia que existen valores grandes que extienden la distribución hasta un 100%.

Relativo a vías aéreas superiores

En relación a las vías aéreas superiores, se plantean los siguientes gráficos:

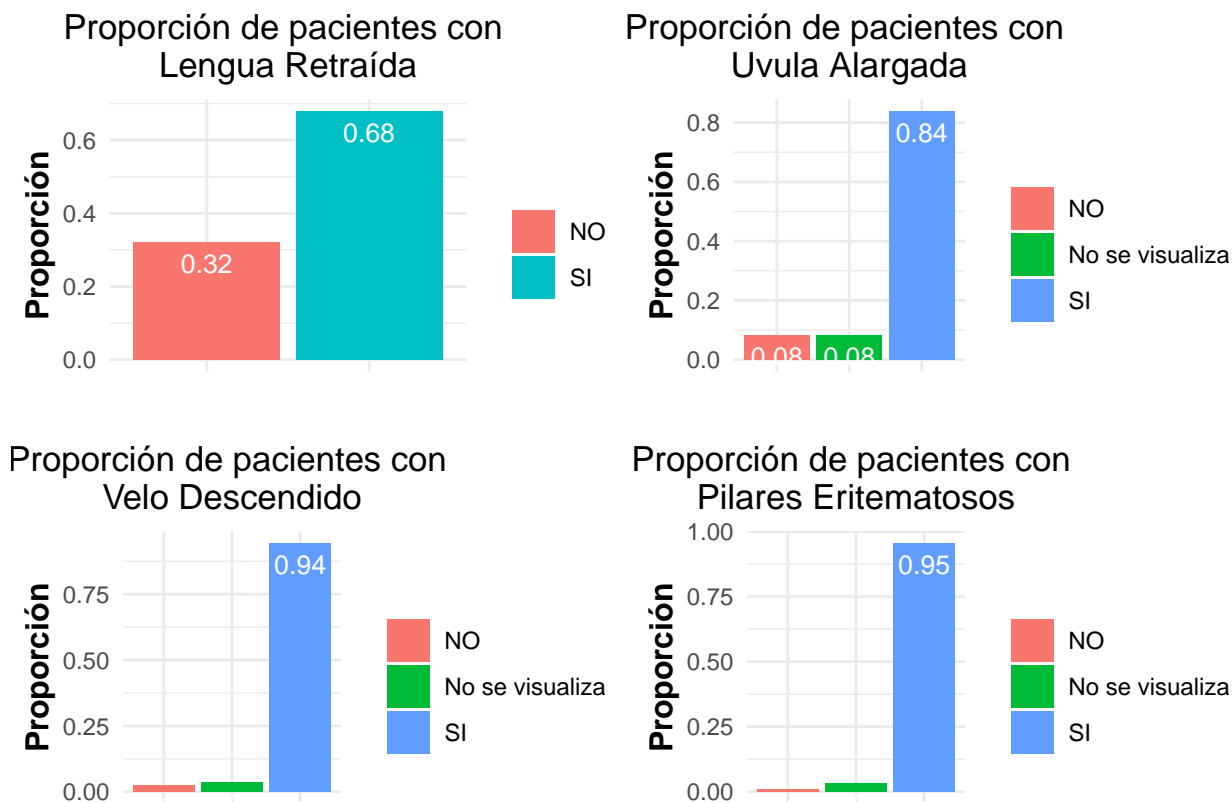




- Del primer gráfico se observa que, el 74% de los pacientes no ha tenido una cirugía en sus vías aéreas superiores. El 11% ha tenido una Amigdalectomía, el 9% una Turbinoplastia y el 5% una Septoplastia.
- Del segundo se concluye que el 63% de los pacientes no poseen obstrucciones de la vía aérea superior.

Condiciones relacionadas a la cavidad oral

En los siguientes gráficos se analizan algunas condiciones relacionadas a la cavidad oral de los pacientes:



- Del primero se nota que el 68% de los pacientes que terminaron su tratamiento poseen la condición de lengua retraída.
- Similarmente, el 84% de los pacientes poseen una uvula alargada, mientras que en un 8% de las personas no se logra concluir sobre esta condición.
- Igualmente, en lo que al velo descendido respecta, se observa que el 94% de los pacientes del tratamiento tienen esta condición.
- De manera casi idéntica, el 95% de los pacientes registra pilares eritematosos.

¿Existe entonces una relación entre las últimas dos condiciones? Véase la siguiente tabla:

Table 5: VELO DESCENDIDO VS PILARES ERITEMATOSOS

	NO	No se visualiza	SI
NO	0	0	2
No se visualiza	0	2	1
SI	1	1	80

De esta tabla se aprecia que prácticamente todos los pacientes que poseen velo descendido, poseen además pilares eritematosos; sin embargo, para probar esta cuestión de manera formal, se plantea el siguiente juego de hipótesis:

$$\begin{cases} H_0 : \text{Las condiciones velo descendido y pilares eritematosos no están relacionadas} \\ H_1 : \text{Las condiciones velo descendido y pilares eritematosos están relacionadas} \end{cases}$$

Para juzgar este juego de hipótesis se plantea el test χ^2 de Pearson para independencia de tablas de contingencia. Los resultados son como sigue:

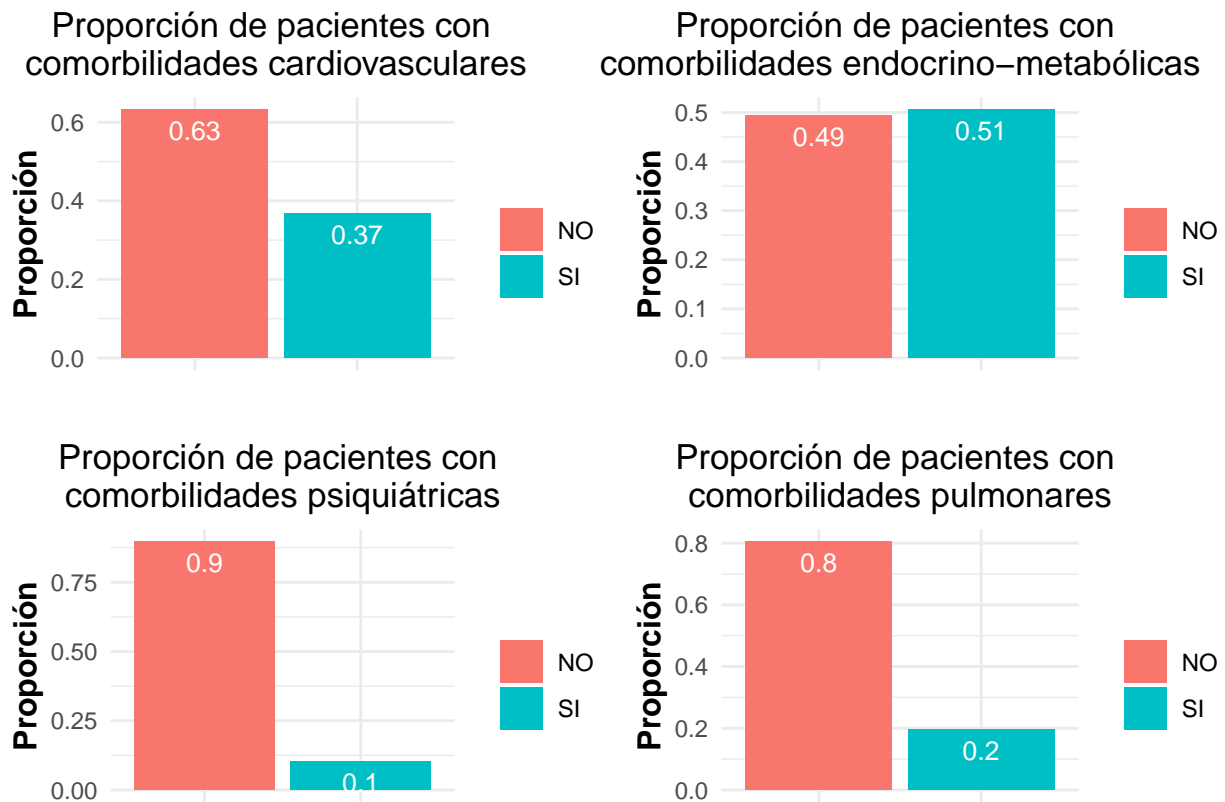
Table 6: Resultados Prueba de Independencia

χ^2	DF	Valor P
37.3373	4	1.5E-7

De acuerdo a este resultado y usando cualquier nivel de significancia (α) usual, se rechaza la hipótesis nula (H_0) concluyendo de este modo que existe una relación estadísticamente significativa entre las condiciones bucales “velo descendido” y “pilares eritematosos” de los pacientes en cuestión.

Comorbilidades

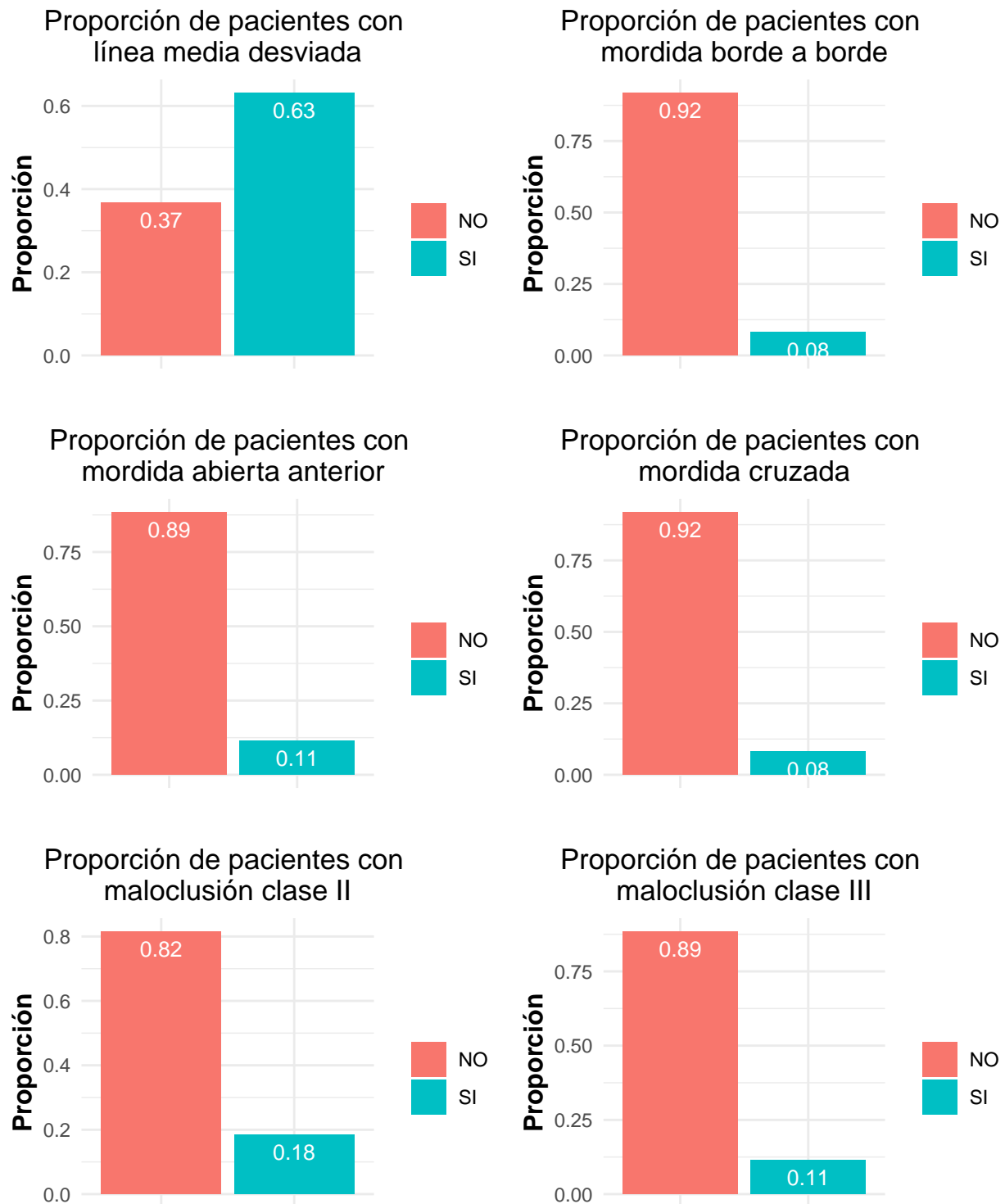
Para tener una perspectiva de las comorbilidades entre los pacientes del tratamiento, se plantean los siguientes gráficos:



- De este conjunto de gráficos se aprecia que, las comorbilidades más usuales entre los pacientes en el tratamiento son las comorbilidades cardiovasculares y las endocrino-metabólicas, esta última presente en casi la mitad de los pacientes.
- La proporción de pacientes con comorbilidades psiquiátricas o pulmonares es baja, con 10% y 20% de los pacientes respectivamente.

Mordida

De manera similar, para tener una perspectiva de las condiciones en la mordida de los pacientes se plantea el siguiente conjunto de gráficos:



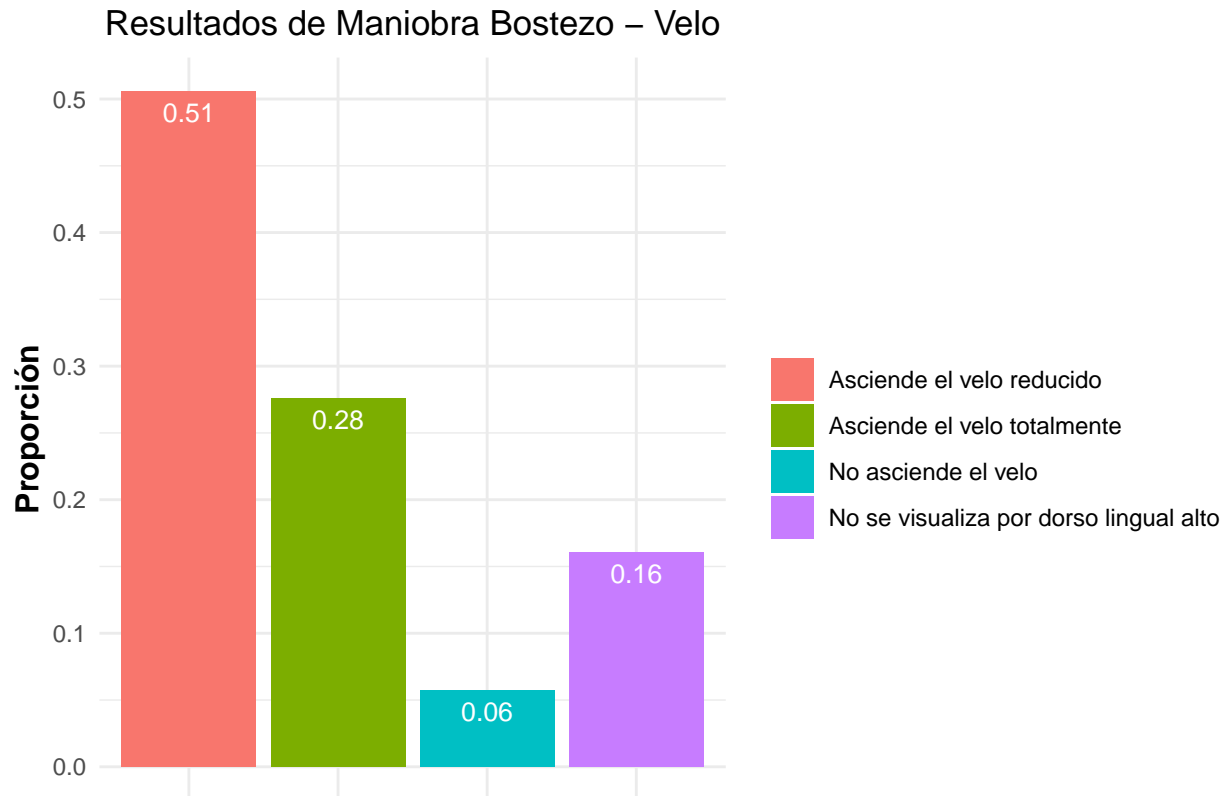
- De las condiciones en la mordida de los pacientes bajo tratamiento, la línea media desviada es la más

frecuente con un 63% de los pacientes que la poseen.

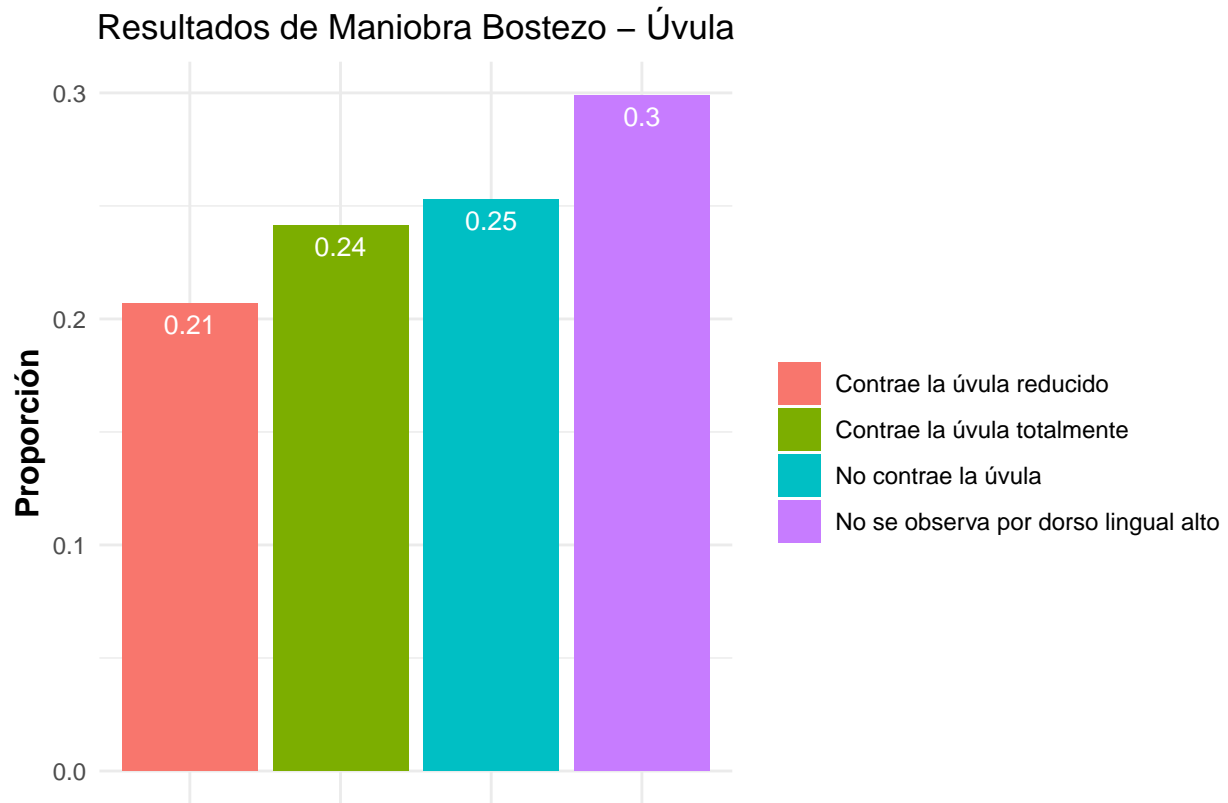
- Condiciones como mordida borde aborde y mordida cruzada son poco usuales entre los pacientes. En todos los casos, menos del 10% de los pacientes tuvieron alguna de estas condiciones.
- El 18% y 11% de los pacientes reportaron maloclusión clase II y maloclusión clase III respectivamente.

Maniobras de bostezo y fonema

Ahora, para visualizar el resultado de las maniobras de bostezo y fonema para los pacientes que completaron el tratamiento, se exponen los siguientes gráficos:

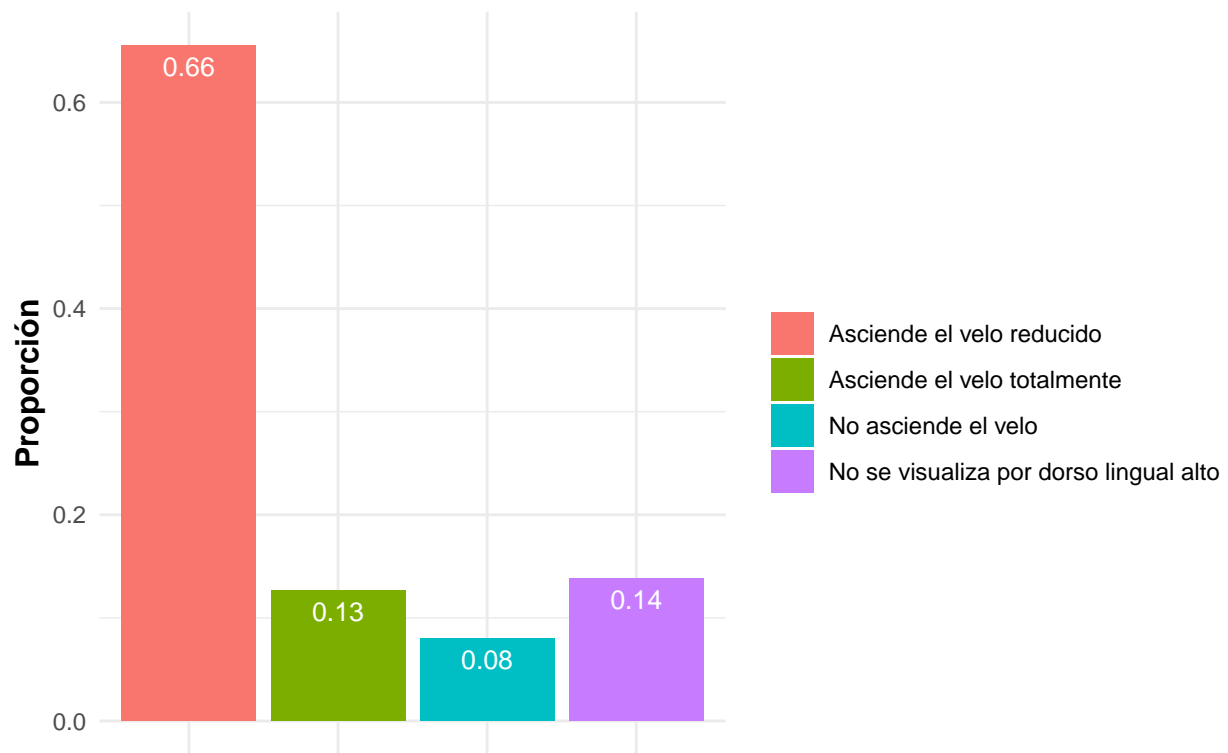


- De este gráfico se puede apreciar que el 79% de los pacientes ascienden el velo en la maniobra de bostezo, con un 51% para aquellos que lo hacen de forma reducida, y un 28% para aquellos que en la maniobra de bostezo ascienden el velo totalmente. Un 16% de los pacientes poseen un dorso lingual lo suficientemente alto para impedir la conclusión en esta maniobra.

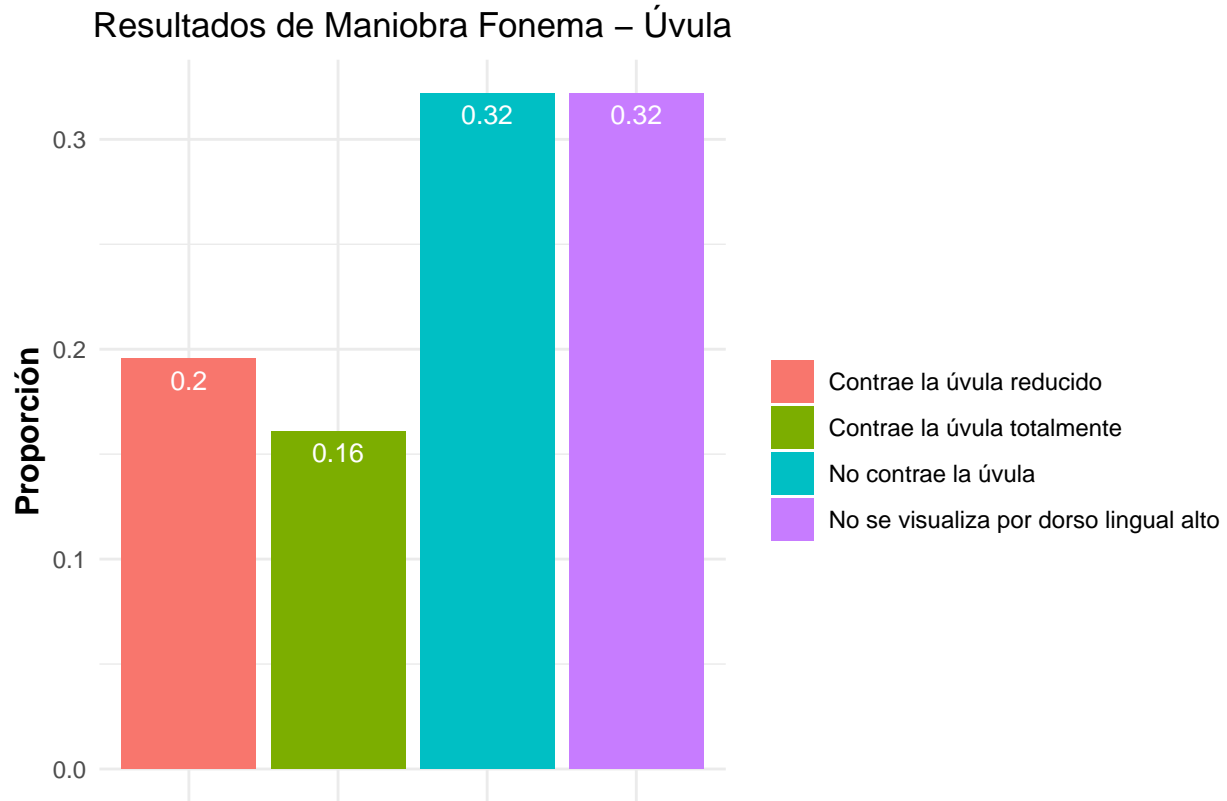


- Continuando con la misma maniobra pero observando la úvula, en el gráfico se logra apreciar que 1/4 de los pacientes no contrae la úvula en la maniobra de bostezo, proporción análoga a aquellos pacientes que si la contraen durante la maniobra. En este caso la mayoría de los pacientes poseen un dorso lingual lo suficientemente alto para que el resultado no sea concluyente, sin embargo la diferencia con los demás grupos es reducida.

Resultados de Maniobra Fonema – Velo



- Pasando a la maniobra de fonema, según el gráfico se puede observar que un 66% de los pacientes asciende el velo de manera reducida en esta prueba. En este caso los pacientes que no ascienden el velo son minoría con un 8% y los demás grupos empatan sus proporciones con cerca de un 14%.



- Finalmente, observando la úvula en la maniobra de fonema, se puede notar que una mayoría de los pacientes no la contraen en este ejercicio con un 32% del total de los participantes. Es destacable en este caso que hay una gran proporción (32%) de pacientes con un dorso lingual tan alto como para no observar el resultado de la práctica.

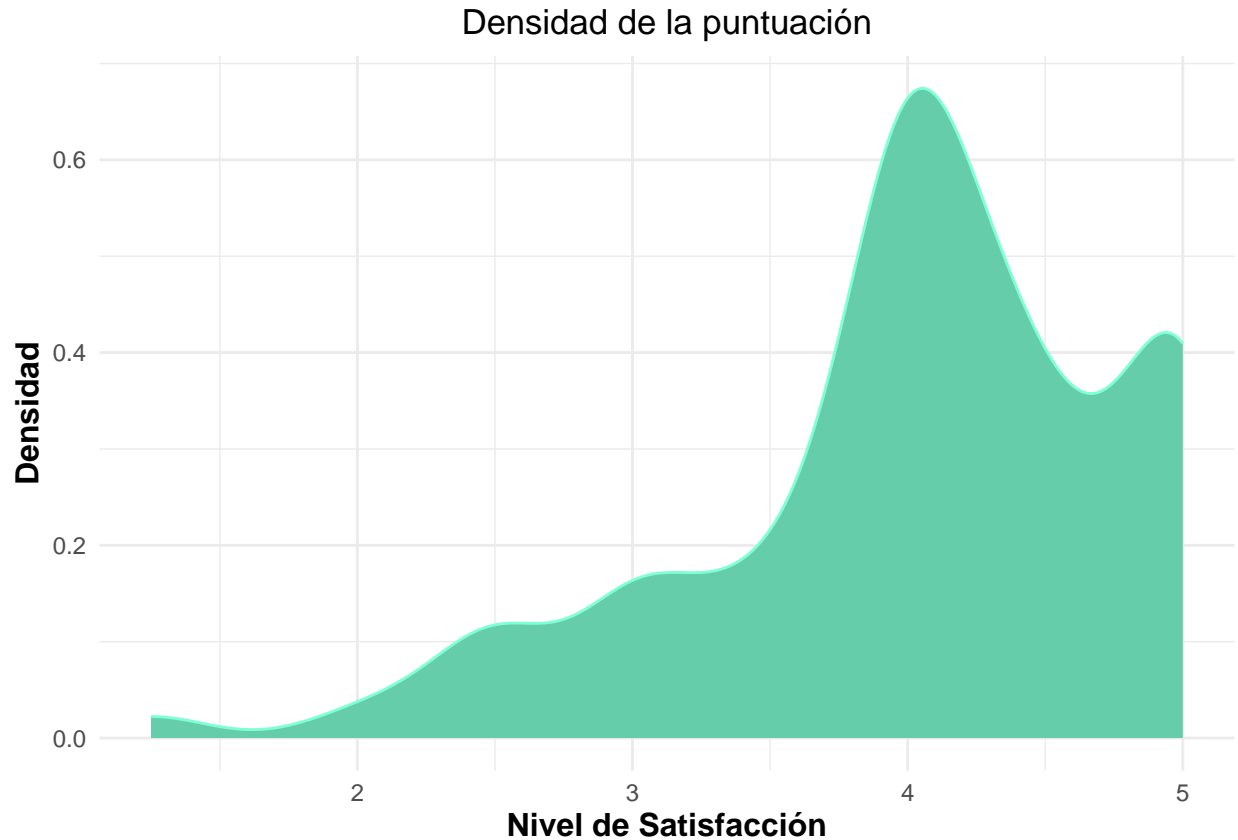
Continuación: Variable respuesta vs. regresores

Como se comentó en el principio del análisis descriptivo, la variable “PuntGeneral” es una escala continua de 1-5 que indica el nivel de satisfacción/efectividad del tratamiento del paciente, donde 1 indica un paciente que considera inefectivo el tratamiento y 5 indica que el paciente considera el tratamiento absolutamente efectivo.

Esta variable configura en sí misma un indicador de desempeño del tratamiento, y es de interés central poder determinar cuáles son los factores que afectan la percepción del paciente en relación a la efectividad del tratamiento. Por esta razón, en aras de una etapa de modelamiento predictivo posterior, se realiza un análisis descriptivo en relación a la puntuación general que asignan los pacientes de acuerdo al éxito del tratamiento.

Distribución de la variable respuesta

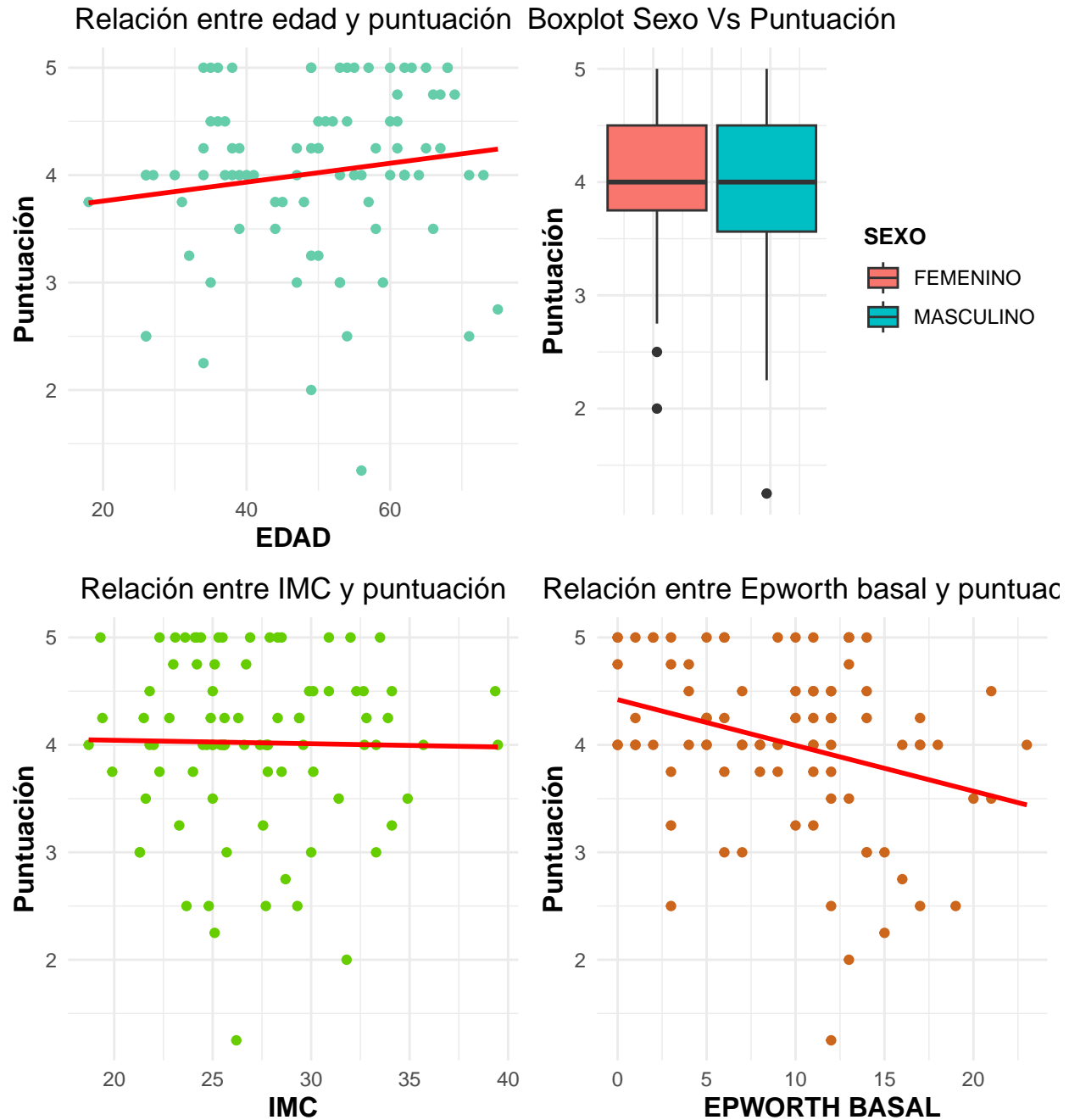
Antes de comenzar con el análisis descriptivo de la respuesta vs las covariables en cuestión, es necesario graficar la curva de densidad de la variable respuesta para determinar así una plausible distribución teórica y para tener una visión general de lo obtenido por los pacientes:



- De este gráfico se puede ver inicialmente que la distribución posee una cola apreciable a la izquierda del 4, adicionalmente se puede ver que la curva es notablemente no-simétrica, por lo cual modelar la respuesta vía un modelo de regresión lineal puede no ser adecuado.
- La mayoría de los puntajes se sitúan en torno al 4, es decir que los pacientes en una gran medida consideran el tratamiento satisfactorio/efectivo. También puede verse que existe una concentración en menor medida cerca al 5, por lo que puede decirse que también hay una porción significativa de los pacientes que consideran el tratamiento absolutamente satisfactorio/efectivo.

Respuesta vs. Edad, Sexo, IMC y Empworth basal

Véase ahora la relación de la puntuación obtenida con las variables mencionadas.

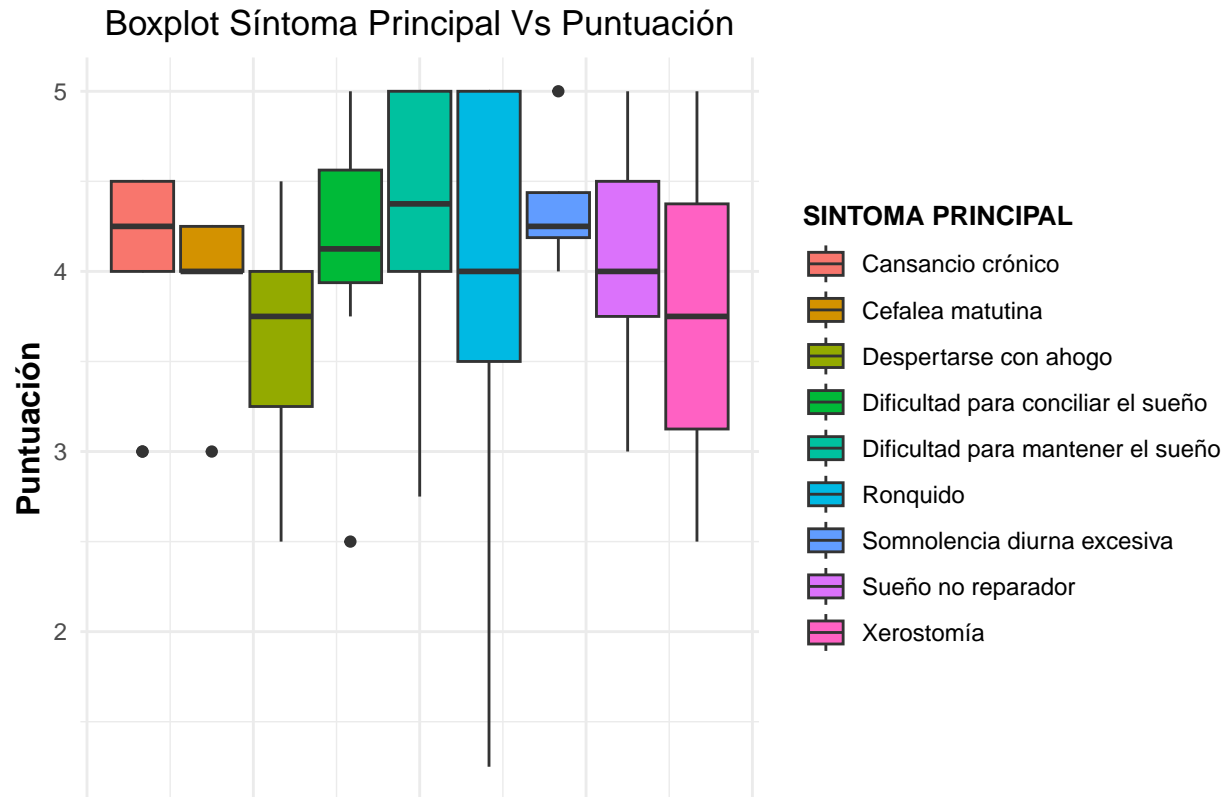


- Del primer gráfico en relación a la edad, se puede notar que aparentemente, los pacientes mayores suelen calificar mejor al tratamiento; sin embargo esta relación es muy débil (la línea roja es casi horizontal) por lo que esta variable puede no ser significativa para el modelo.
- Del gráfico relativo al sexo se puede notar que las dos cajas se traslapan totalmente, lo que indica en este caso que no existe una diferencia significativa entre masculinos y femeninas a la hora de calificar el tratamiento. Nótese que hay dos puntuaciones atípicamente bajas para las mujeres y una para los hombres.

- De manera similar al primer gráfico, en el tercero se puede apreciar que no existe una relación entre el IMC y la puntuación que los pacientes reportan.
- Finalmente, en lo que respecta al resultado del cuestionario epworth se puede notar que sí existe una relación medianamente plausible entre dicha variable y la puntuación de efectividad. Particularmente se nota que los pacientes menos somnolientos suelen calificar mejor al tratamiento.

Síntoma principal

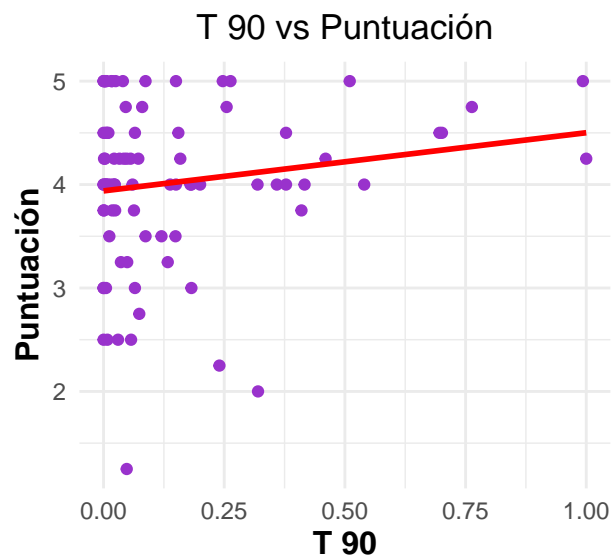
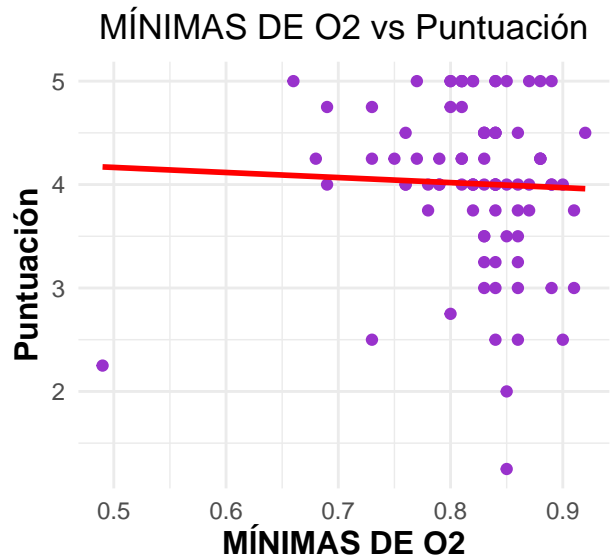
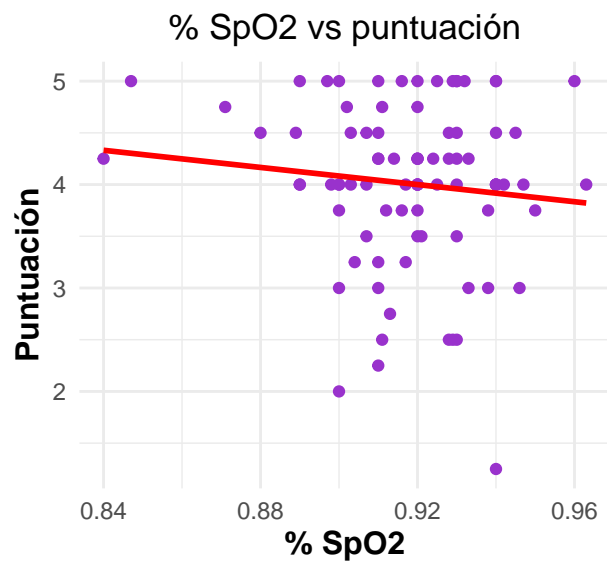
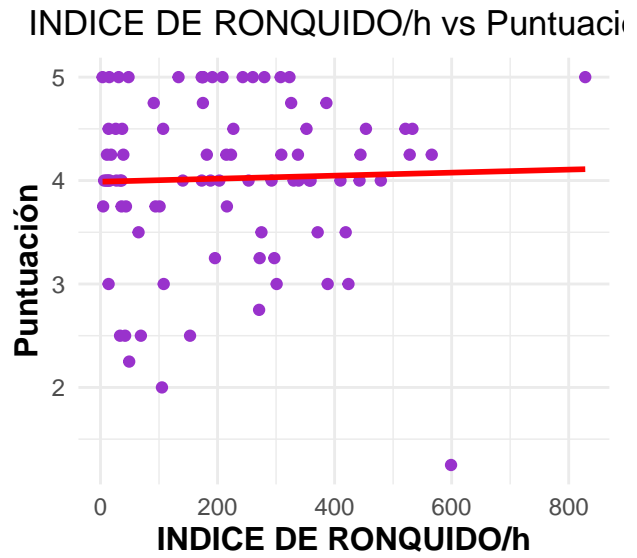
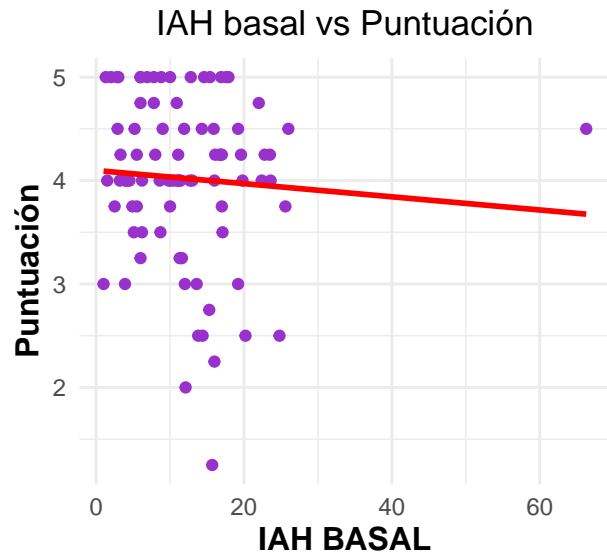
De esta variable surge el cuestionamiento natural: ¿Existe una relación entre el síntoma principal reportado y la puntuación obtenida?



- Se puede apreciar que existe una diferencia significativa entre los paciente que se despiertan con ahogo y los pacientes somnolientos, donde estos últimos suelen calificar mejor los resultados del tratamiento. Este mismo comportamiento es notable con los pacientes que sufren cefalea matutina y cansancio crónico.

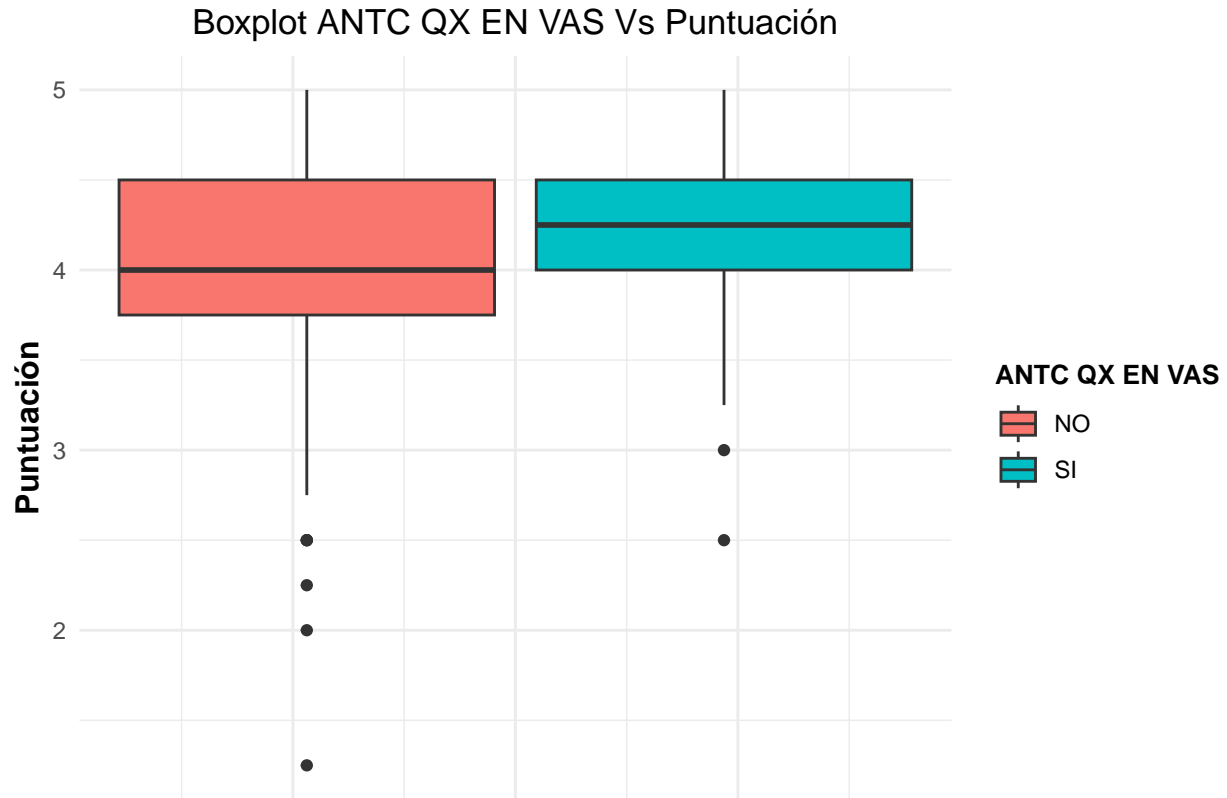
Índices y porcentajes

Ahora se analiza la relación entre algunas variables numéricas y la respuesta en cuestión



- De este grupo de gráficos se puede notar que la variable que tiene una relación apreciable con la respuesta en este caso es T 90, los demás índices y porcentajes no muestran una relación plausible que permita concluir de antemano significancia para explicar el puntaje obtenido.

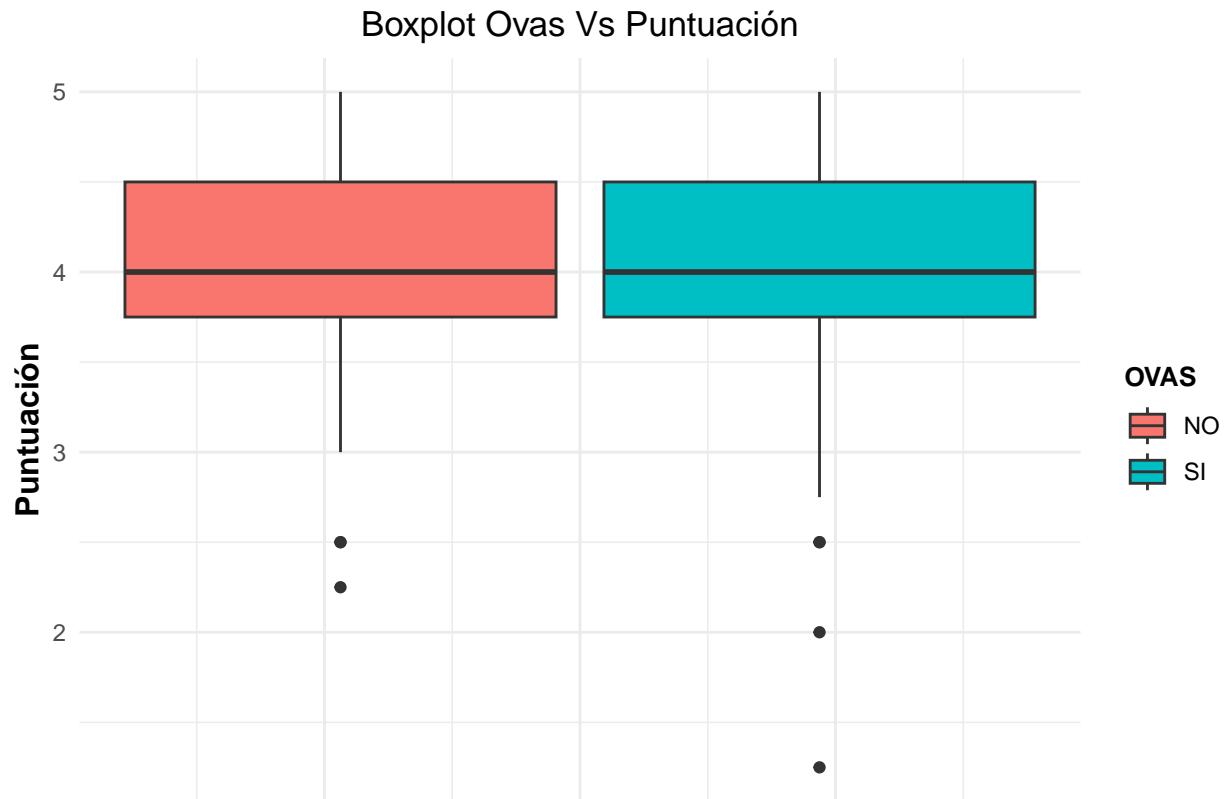
Antecedentes quirurgicos en vías aereas superiores



- Se aprecia que no existe relación aparente entre la percepción de efectividad del tratamiento y los antecedentes quirurgicos en vías aereas superiores.

OVAS

¿Afecta el poseer obstrucciones en la vía aérea superior en la percepción de efectividad del tratamiento?

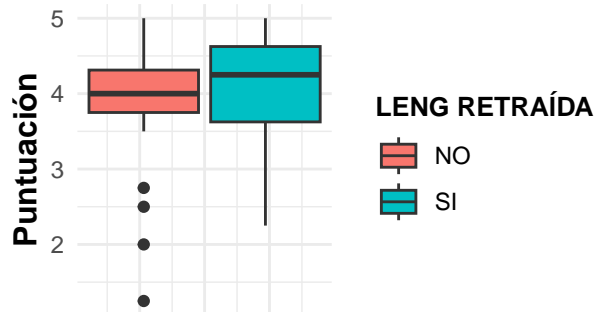


- Similar a lo obtenido en el apartado anterior, el hecho de poseer algún tipo de obstrucción en las vías aéreas superiores aparentemente no afecta el puntaje que se obtiene de efectividad/satisfacción del tratamiento.

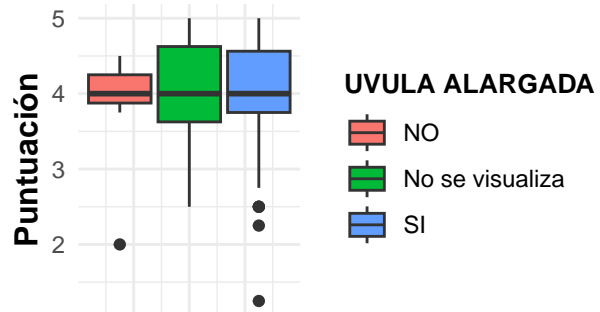
Condiciones de la cavidad oral vs puntuación

Ahora se analizarán las condiciones de la cavidad oral en relación a la variable respuesta en cuestión:

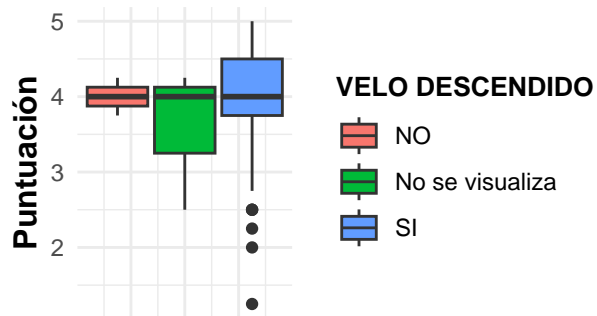
Lengua Retraída Vs Puntuación



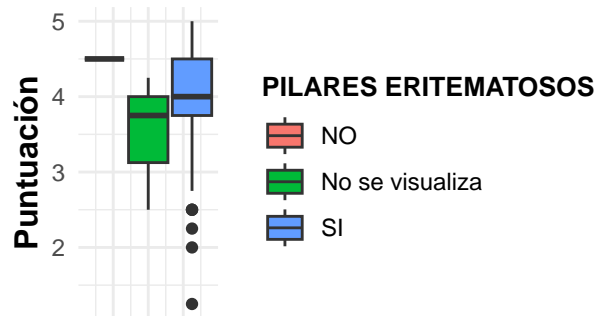
Úvula Alargada Vs Puntuación



Velo Descendido Vs Puntuación



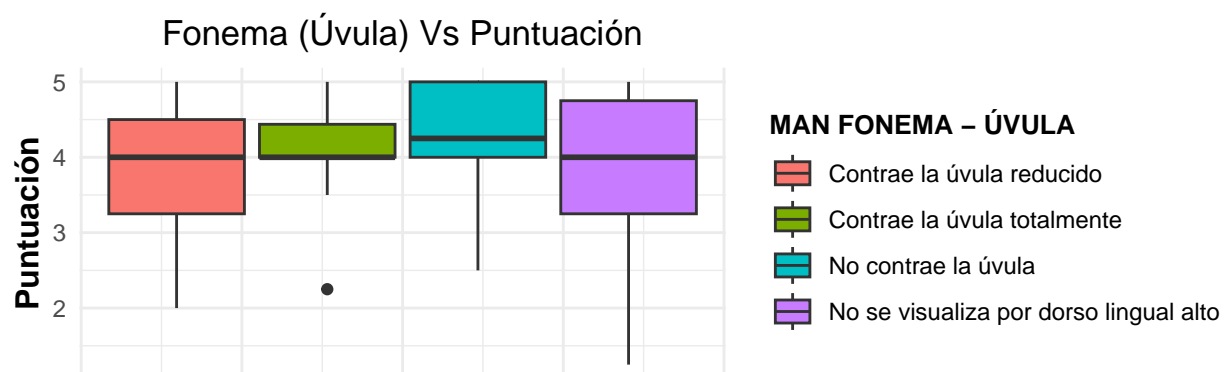
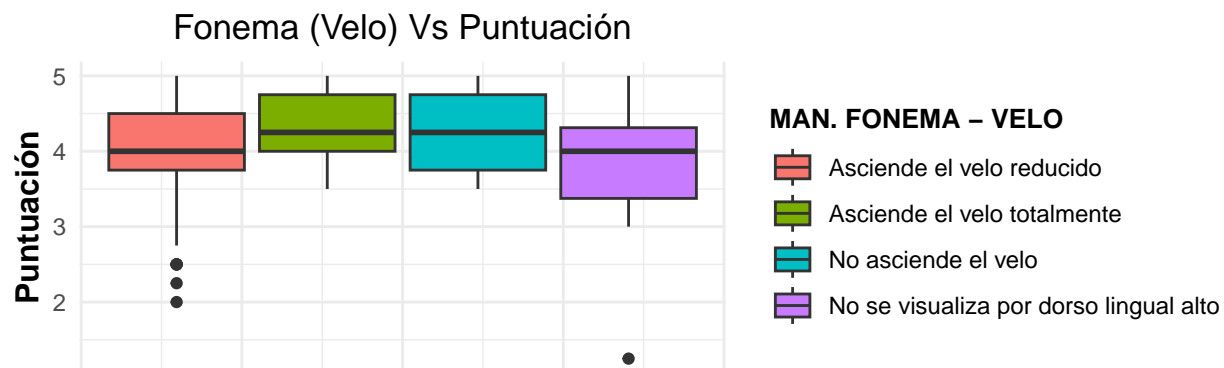
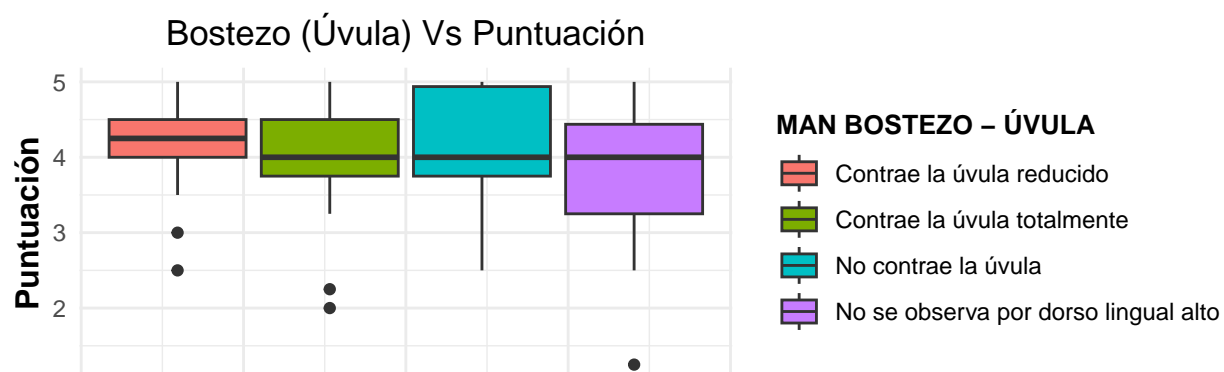
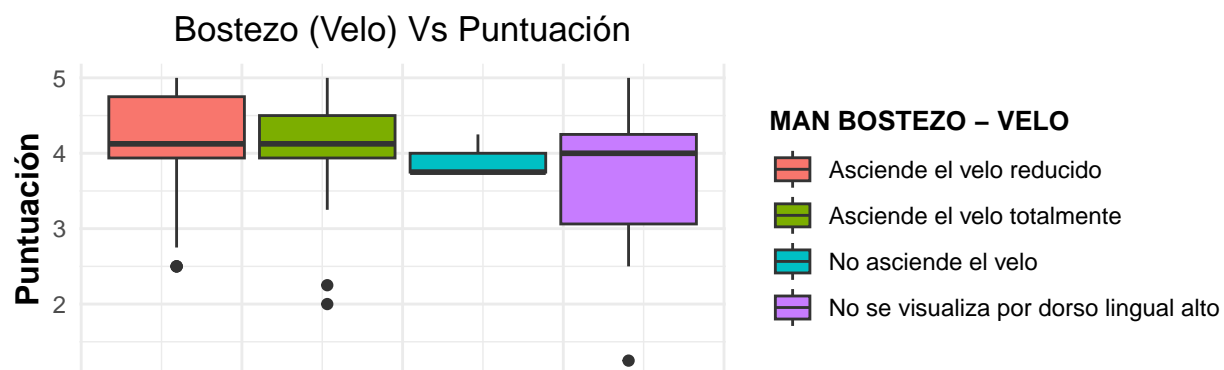
Pilares Eritematosos Vs Puntuación



- De este conjunto de gráficos se logra apreciar que las condiciones de lengua retraída y úvula alargada no influyen en el resultado de puntaje; sin embargo, condiciones como velo descendido y pilares eritematosos sí parecen ser significativos para explicar la variable respuesta.
- Es importante tener en cuenta lo obtenido en el análisis descriptivo marginal respectivo a estas variables, pues allí se encontró una relación directa en las últimas dos variables, lo cual podría indicar adecuado incluir un término de interacción en el futuro modelo predictivo.

Maniobras de bostezo y fonema vs puntuación

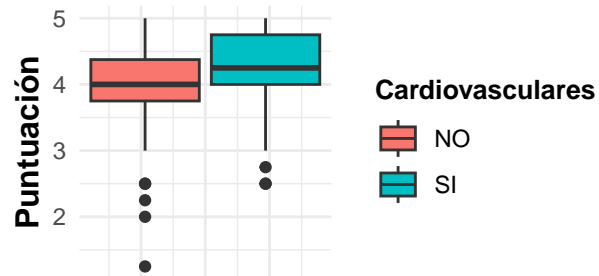
Ahora, para visualizar el resultado de las maniobras de bostezo y fonema para los pacientes que completaron el tratamiento en relación al resultado que reportan, se exponen los siguientes gráficos:



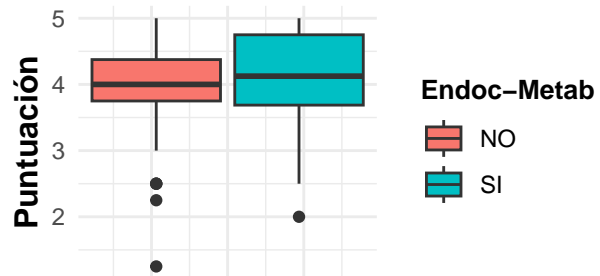
- De este último grupo se nota que en cada gráfico todas las cajas se traslapan unas a otras, por lo cual se puede concluir que estas variables no ayudan a explicar el puntaje obtenido con cada paciente y que en la propuesta de un modelo es probable que no muestren significancia suficiente para ser incluidas en el mismo.

Variables relativas a comorbilidades

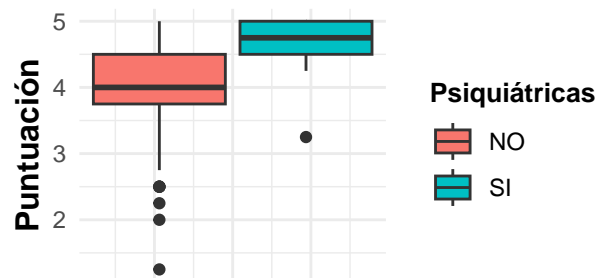
**Comorb Cardiovasculares
Vs Puntuación**



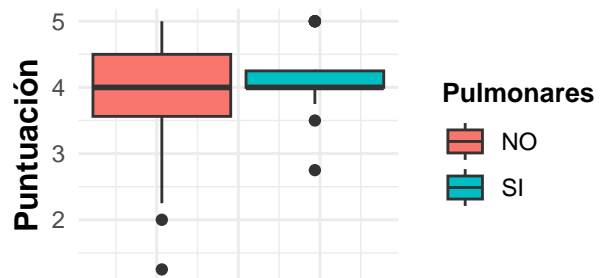
**Comorbilidades Endocrino–Metabólicas
Vs Puntuación**



**Comorbilidades Psiquiátricas
Vs Puntuación**



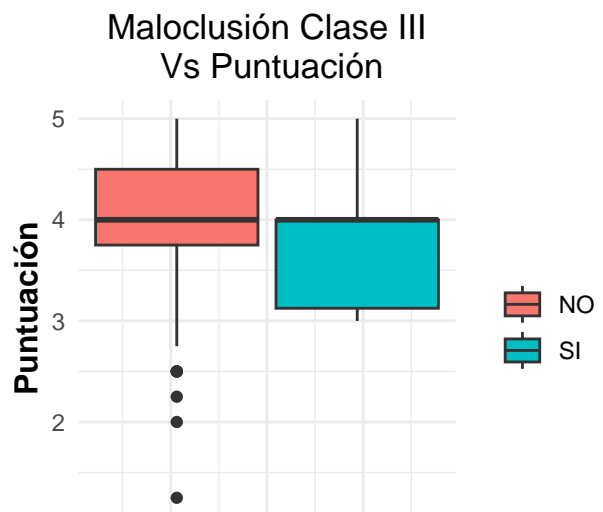
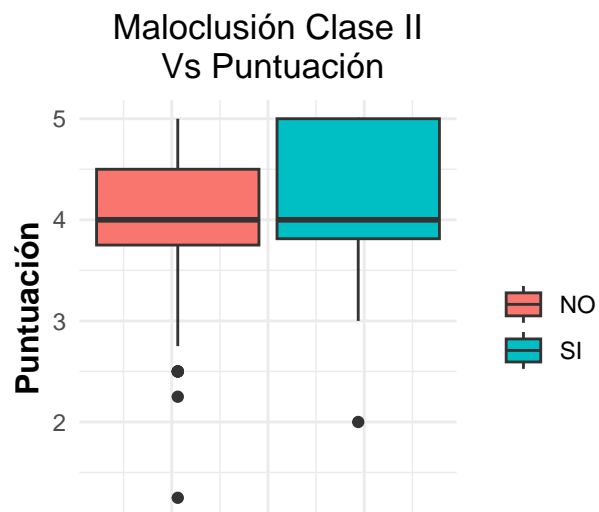
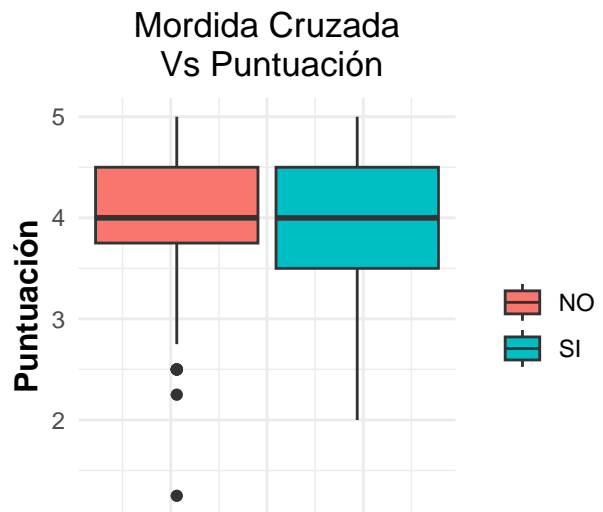
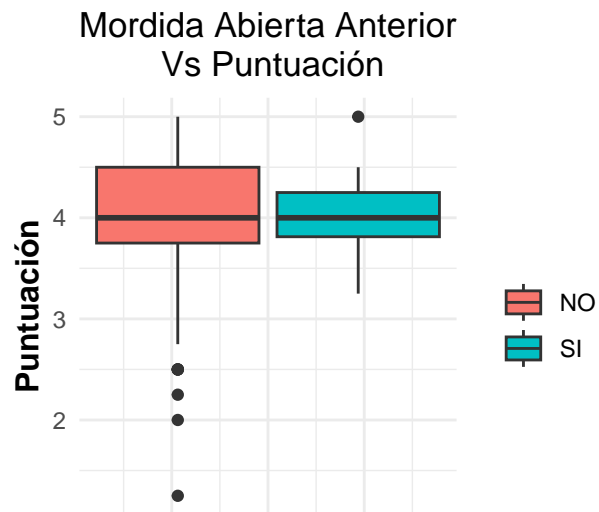
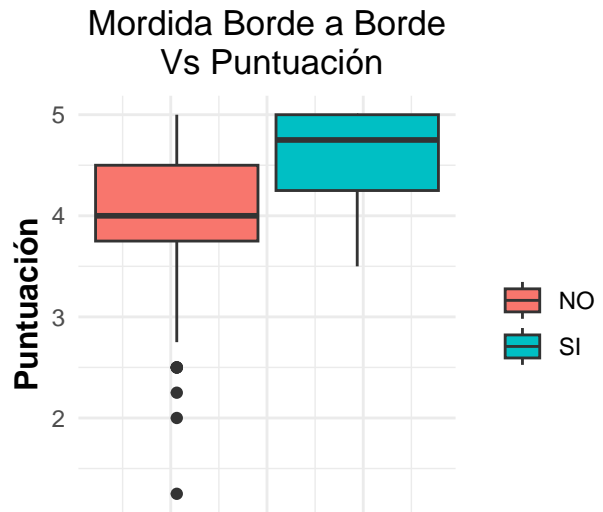
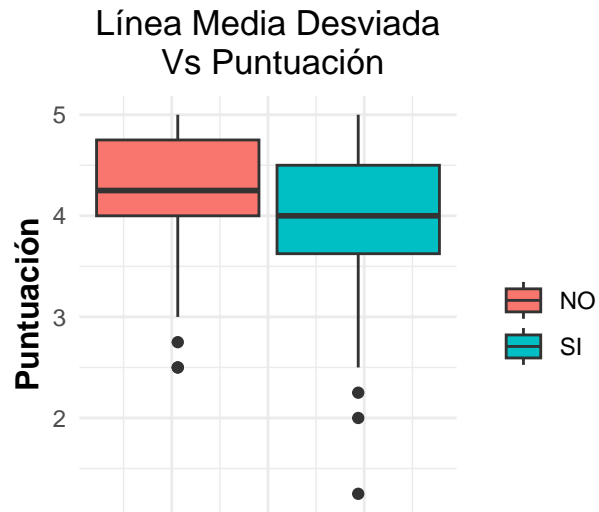
**Comorbilidades Pulmonares
Vs Puntuación**



- En lo que respecta a las comorbilidades, se puede apreciar que existe una relación evidente entre la puntuación que dan los pacientes y si los mismos poseen algún tipo de comorbilidad psiquiátrica, donde aquellos que sí poseen alguna, perciben el tratamiento significativamente más efectivo. Esta relación también es plausible para aquellos pacientes con comorbilidades cardiovasculares, sin embargo en este punto la relación es más débil.

Mordida vs Respuesta

Finalmente, en lo que respecta a la mordida relacionado a la efectividad del tratamiento, se tienen los siguientes gráficos:



- Con estos últimos se puede apreciar que prácticamente en todas las condiciones de la mordida, las cajas se traslapan lo suficiente para descartar una relación significativa con la variable respuesta; sin embargo, en condiciones como mordida borde a borde y maloclusión clase III se puede pensar en una relación débil plausible. En ese sentido sería interesante incluir este grupo de covariables al modelo como una interacción.

Modelación de la variable de puntuación general

Dado que la distribución de la variable respuesta es altamente no simétrica y con una forma desconocida, además de que algunas relaciones con las covariables son débiles, se considera adecuado optar por modelos predictivos robustos propios del machine learning (un enfoque analítico predictivo) que no posean supuestos distribucionales y donde prime el poder predictivo y la minimización del error en las predicciones.

En ese sentido, se tendrán las siguientes consideraciones:

- Para evaluar el desempeño del modelo se utilizará el RMSE (Raíz del error cuadrático medio).

$$RMSE = \sqrt{\sum_{i=1}^n \frac{(y_i - \hat{y}_i)^2}{n}}$$

Con $n = 87$ el número de observaciones en la base de datos, y_i el valor real de la respuesta y \hat{y}_i el valor predicho por el modelo.

- Con esta métrica se escogerá el modelo más preciso.
- Para prevenir el overfitting (sobreajuste) se optará por usar el método cross validation partiendo el conjunto de datos en proporción 80% para entrenamiento y 20% para prueba. Si el RMSE de ambos conjuntos difiere en más de un 15% se considerará que el modelo sobreajusta los datos y será descartado.
- Para el ajuste de los modelos se seleccionarán las variables que poseen una relación destacable con la respuesta: Edad, epworth basal, síntoma principal, T90, comorbilidades cardiovasculares y psiquiátricas, pilares eritematosos, mordida borde a borde y maloclusión clase III.

Modelos a considerar

Para cada tipo de modelo se van a considerar 8 conjuntos de regresores, donde cada fórmula se encuentra anidada en un modelo más específico (ver la siguiente tabla). Para la elección de las fórmulas se tuvieron en cuenta primero las variables que en el análisis descriptivo resultaron más significativas, sucesivamente se fueron agregando variables con las relaciones más plausibles en orden de significancia.

Table 7: Fórmulas consideradas

Modelo	Fórmula
1	PuntGeneral ~ EPWORTH BASAL + SINTOMA PRINCIPAL
2	PuntGeneral ~ EPWORTH BASAL + SINTOMA PRINCIPAL + CO_Psiquiatricas
3	PuntGeneral ~ EPWORTH BASAL + SINTOMA PRINCIPAL + CO_Psiquiatricas + T 90
4	PuntGeneral ~ EPWORTH BASAL + SINTOMA PRINCIPAL + CO_Psiquiatricas + T 90 + EDAD
5	PuntGeneral ~ EPWORTH BASAL + SINTOMA PRINCIPAL + CO_Psiquiatricas + T 90 + EDAD + Maloclusion_clase_III
6	PuntGeneral ~ EPWORTH BASAL + SINTOMA PRINCIPAL + CO_Psiquiatricas + T 90 + EDAD + Maloclusion_clase_III + CO_Cardiovasculares
7	PuntGeneral ~ EPWORTH BASAL + SINTOMA PRINCIPAL + CO_Psiquiatricas + T 90 + EDAD + Maloclusion_clase_III + CO_Cardiovasculares + PILARES ERITEMATOSOS
8	PuntGeneral ~ EPWORTH BASAL + SINTOMA PRINCIPAL + CO_Psiquiatricas + T 90 + EDAD + Maloclusion_clase_III + CO_Cardiovasculares + PILARES ERITEMATOSOS + Mordida_borde_a_borde

Para cada propuesta de modelo, se ajustarán las ocho fórmulas y se elegirá al mejor modelo objetivamente, se considerará a ese como el mejor de su clase y se pondrá a competir con las demás propuestas.

Modelo de regresión lineal múltiple

Se opta por esta clase de modelo como una base o referencia para propuestas futuras. En este caso se aborda este modelo desde una perspectiva analítica predictiva mas no inferencial, por lo cual se obvian los supuestos del modelo y no se realiza análisis de significancia y e inferencia sobre los demás coeficientes.

Los resultados son los siguientes

Table 8: Resultados Modelo Lineal Múltiple

Modelo	RMSE_TRAIN	RMSE_TEST	ErrorRelativo
Fórmula 1	0.7401688	0.6598902	0.11
Fórmula 2	0.7200783	0.7041234	0.02
Fórmula 3	0.7148951	0.6316063	0.12
Fórmula 4	0.7140792	0.6305526	0.12
Fórmula 5	0.7089893	0.6234212	0.12
Fórmula 6	0.7078337	0.6297233	0.11
Fórmula 7	0.7000616	0.6223355	0.11
Fórmula 8	0.6764499	0.6321611	0.07

De este resultado es notable que ninguno de los ocho modelos propuestos con regresión lineal múltiple sobreajusta los datos, es decir, el modelo tiene un buen poder predictivo con datos que no conoce y posee la capacidad de generalizar las predicciones. El mejor modelo encontrado es el que contiene todas las variables significativas, con un error medio de 0.68 en el conjunto de datos de entrenamiento, y un error medio de 0.63 en el conjunto de datos de prueba.

K-Nearest Neighbors Regression

El algoritmo de K vecinos más cercanos para regresión es la extensión del algoritmo de KNN pero en este caso para problemas predictivos con respuesta continua. El algoritmo puede explicarse de manera intuitiva:

- Considere una nube de puntos con los valores de un conjunto de regresores.
- Considere una nueva observación.

- Ubique esa observación en su punto respectivo de la nube de puntos.
- Tome los valores de los k vecinos más cercanos a esa nueva observación.
- Promedie dichos valores y esa será la predicción de la respuesta para la nueva observación.

En este caso particular, el hiperparámetro k se seleccionará mediante el procedimiento LOOCV (leave one out cross validation) sobre el conjunto de entrenamiento. Este procedimiento se hará por cada fórmula.

El resultado se expone como sigue:

Table 9: Resultados Modelo KNN-REG

Modelo	RMSE_TRAIN	RMSE_TEST	ErrorRelativo	BestK
Fórmula 1	0.7626886	0.6663650	0.13	9
Fórmula 2	0.7531820	0.7124703	0.05	9
Fórmula 3	0.7328659	0.7452646	0.02	9
Fórmula 4	0.7599399	0.6965120	0.08	12
Fórmula 5	0.7576740	0.6965120	0.08	12
Fórmula 6	0.7566308	0.6965120	0.08	12
Fórmula 7	0.7568675	0.6965120	0.08	12
Fórmula 8	0.7569104	0.6911396	0.09	12

En este caso la columna “BestK” hace referencia al hiperparámetro K que mejor desempeño predictivo muestra para cada fórmula.

De este último resultado se alcanza a apreciar que en ningún caso el modelo knn-reg sobreajusta los datos y se muestra con un poder predictivo destacable. Sin embargo, este modelo predictivo no logra mejorar los resultados obtenidos por el modelo lineal tradicional en train y test.

Bagging: Random Forest

Random Forest o bosque aleatorio es un algoritmo que extiende el uso de los árboles de decisión.

El principio de esta clase de modelo consiste en ajustar una cantidad n determinada de árboles de decisión, luego se realiza la predicción con cada árbol y se escoge la predicción por mayoría. El funcionamiento simplificado se ilustra así

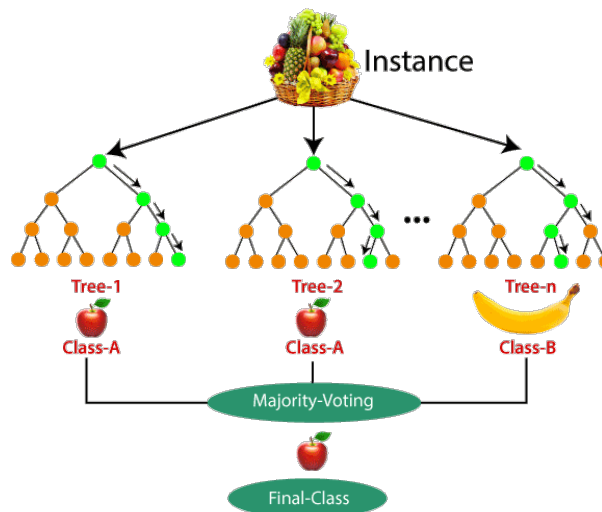


Figure 1: Idea de Random Forest

Table 10: Resultados Modelo Random Forest

Modelo	RMSE_TRAIN	RMSE_TEST	ErrorRelativo
Fórmula 1	0.7019808	0.5941025	0.15
Fórmula 2	0.6847275	0.6278135	0.08
Fórmula 3	0.6347424	0.6060039	0.05
Fórmula 4	0.5789674	0.6385503	0.10
Fórmula 5	0.5797235	0.6226391	0.07
Fórmula 6	0.5703882	0.6158217	0.08
Fórmula 7	0.5923958	0.6003055	0.01
Fórmula 8	0.5973146	0.6080586	0.02

Antes de concluir se debe conocer que el número de árboles empleado es, por defecto, 500.

A partir de este resultado se puede observar que el error relativo del modelo random forest con el primer conjunto de variables independientes es del 15% aproximadamente, por lo cual puede pensarse en overfitting; sin embargo, esto no es así porque el modelo está prediciendo mejor en el conjunto de prueba que en el de entrenamiento, por lo cual no incurre en overfitting. Sin embargo, se debe resaltar que el modelo respectivo a la fórmula 6 reporta un excelente desempeño tanto en train como en test, sin incurrir en overfitting, por lo cual este es un gran candidato a modelo final.

Boosting: Stochastic Gradient Boosting

El boosting como el bagging es un modelo de ensamble, es decir, se compone por varios modelos más básicos que crean un modelo “fuerte” con altas capacidades predictivas. En el método anterior (Random Forest) se vió que este es, de manera resumida, el ensamble en paralelo de varios árboles de decisión básicos con los cuales se puede obtener mediante votación por mayoría, una predicción bastante certera. En este caso también se ensamblan varios árboles de decisión o “modelos débiles” pero en forma secuencial, de manera tal que el siguiente árbol ajuste lo que el anterior no pudo.

Dicho esto, los resultados que se obtienen con el modelo Stochastic Gradient Boosting son

Table 11: Resultados Modelo Stochastic Gradient Boosting

Modelo	RMSE_TRAIN	RMSE_TEST	ErrorRelativo
Fórmula 1	0.7596955	0.6616794	0.13
Fórmula 2	0.7601588	0.7042014	0.07
Fórmula 3	0.7457091	0.7261496	0.03
Fórmula 4	0.7342954	0.7760863	0.06
Fórmula 5	0.6783445	0.7246467	0.07
Fórmula 6	0.7212374	0.7262767	0.01
Fórmula 7	0.7276326	0.7785293	0.07
Fórmula 8	0.7261153	0.7748918	0.07

Dada la complejidad de este algoritmo, se dió vía libre para que el algoritmo computacional desarrollado en la librería caret hiciera el ajuste correspondiente de los hiperparámetros y así evitar errores e inestabilidades numéricas.

Si bien el modelo boosting por descenso del gradiente estocástico no sobreajusta los datos en ningún conjunto de variables independientes, los resultados obtenidos son cuantitativamente inferiores a los obtenidos por el modelo de Random Forest, por lo cual podría considerarse a priori que este tipo de modelo o bien no es adecuado para nuestro conjunto de datos o tal vez requiere de un desarrollo más extenso.

Comparación de los modelos obtenidos

Una vez planteadas las cuatro propuestas de modelos (Regresión lineal múltiple, KNN-REG, Random Forest y Stochastic Gradient Boosting), se selecciona la fórmula que menor rmse reporte en cada conjunto de datos (entrenamiento y prueba) para cada uno de los mismos. En ese caso, se obtienen los “finalistas”

Table 12: Mejores modelos de cada tipo

Modelo	Fórmula	Rmse Train	Rmse Test	Error Relativo
Modelo Lineal	Fórmula 8	0.6764499	0.6321611	0.07
Knn-Reg	Fórmula 1	0.7626886	0.6663650	0.13
Random Forest	Fórmula 6	0.5703882	0.6158217	0.08
Stoch Grad Boost	Fórmula 1	0.7596955	0.6616794	0.13

De esta tabla se puede apreciar que tanto en train como en test, el modelo que más se destaca en términos de rmse es el modelo Random Forest ¿Pero cómo se interpreta el rmse?: En promedio el modelo random forest se equivoca en 0.62 puntos de efectividad/satisfacción del tratamiento, se espera que para un nuevo paciente que ingresa al tratamiento, el modelo acierte su satisfacción futura con 0.62 unidades de error.

Veamos el resultado anterior de manera gráfica



Como es notable en el gráfico, tanto en el conjunto de datos de entrenamiento como en el conjunto de datos de prueba, el modelo “RF” (Random Forest) con la fórmula 6 (ver tabla 7) es el que mejores resultados demuestra en términos de error, por lo cuál este es el modelo final para explicar la respuesta “PuntGeneral”.

Anotaciones sobre el resultado

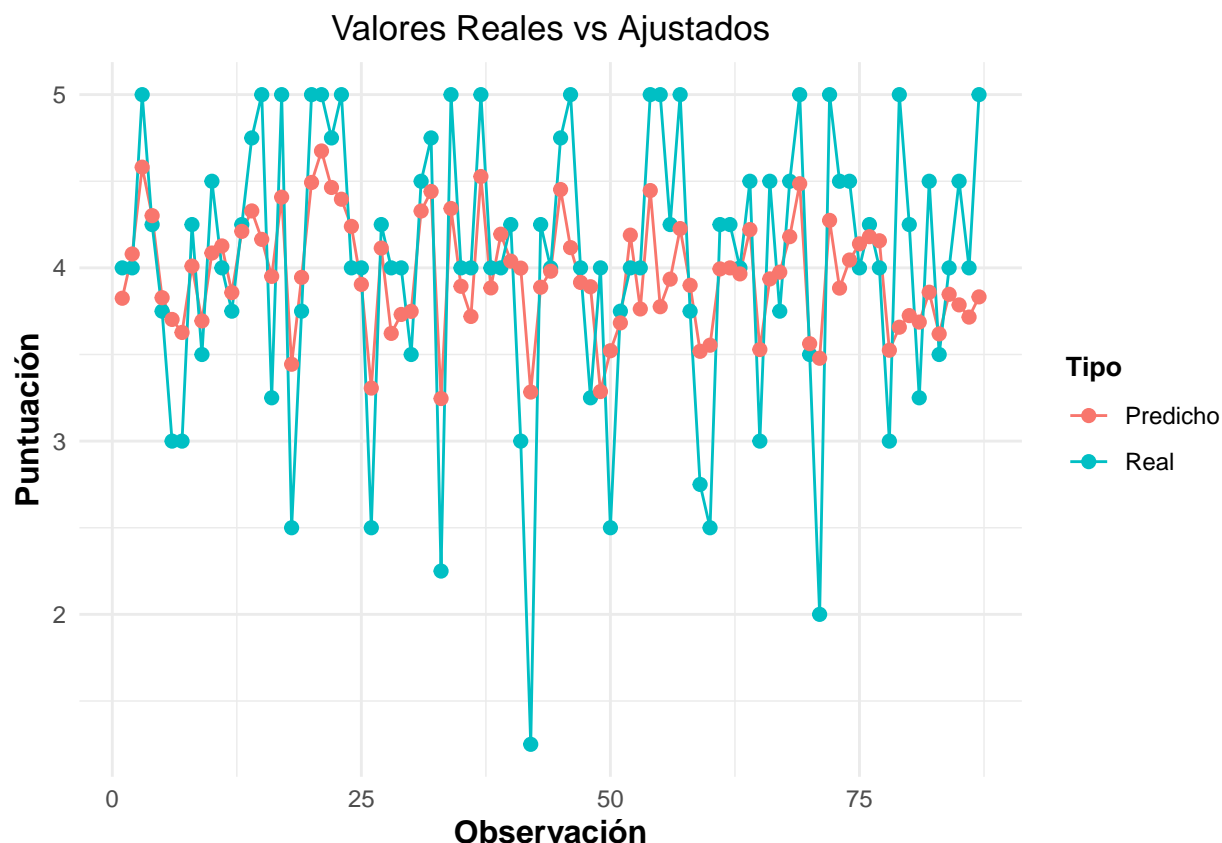
Como se comentó en la sección anterior, el mejor modelo obtenido mediante la metodología propuesta es un modelo propio del paradigma bagging, conocido como modelo Random Forest o modelo de bosques aleatorios. En este caso el modelo se entrenó con 500 árboles aleatorios usando el principio explicado en la respectiva sección. Este modelo ajustó mejor los datos usando la fórmula 6 que se muestra a continuación:

$$\begin{aligned} PuntGeneral \sim EPWORTHBASAL + SINTOMAPRINCIPAL + COPsiquiatricas + T90 \\ + EDAD + MaloclusionClaseIII + COCardiovasculares \end{aligned}$$

Es decir, las covariables que resultaron al menos mínimamente significativas en el análisis descriptivo que mejor explican la puntuación que dan los pacientes al tratamiento son las 7 comprendidas en la anterior fórmula y en ese sentido, serán las que mejor determinen la percepción de efectividad del mismo.

Visualizando el ajuste

Para visualizar el buen ajuste del modelo obtenido se plantea el siguiente gráfico



En este gráfico se puede observar que aunque el modelo parece conservador, realmente las desviaciones corresponden a lo reportado según el RMSE, el cuál es relativamente bajo según la desviación natural de los datos.

Es muy destacable de este modelo que logra capturar fielmente la tendencia de los pacientes, es decir, sube cuando debe subir y baja cuando debe bajar, lo que indica que el modelo está replicando el comportamiento natural de los datos y posee la capacidad de perfilar los pacientes de acuerdo a las variables involucradas en el mismo.

Resultado final: Perfilamiento de los pacientes

Como es de interés primario, ahora el análisis se hará de acuerdo a lo que el modelo dicte como el perfil más exitoso en el tratamiento y lo que prediga como el perfil de paciente que percibe peores resultados. Para esa causa se hará una predicción para todas las combinaciones posibles de las variables escogidas en la fórmula 6 (ver tabla 7) que se encuentren en la base de datos. Posteriormente se reportarán los mejores y peores puntajes de acuerdo al modelo obtenido y se harán las conclusiones respectivas. El resultados se muestra en la siguiente tabla.

Table 13: Perfiles de los pacientes

EPWORTH BASAL	SÍNTOMA PRINCIPAL	PSIQ	T90	EDAD	MALOCCLUSIÓN	CARDIO	PUNTAJE
6	Dificultad para mantener el sueño	SI	0.00	55	NO	NO	4.67
0	Ronquido	SI	0.26	68	NO	SI	4.58
3	Dificultad para mantener el sueño	NO	0.00	62	NO	NO	4.53
9	Somnolencia diurna excesiva	NO	0.01	63	NO	NO	4.49
13	Dificultad para mantener el sueño	SI	0.00	49	NO	NO	4.49
3	Despertarse con ahogo	NO	0.03	26	NO	NO	3.44
19	Xerostomía	NO	0.06	26	NO	NO	3.31
16	Ronquido	NO	0.14	56	NO	NO	3.29
12	Ronquido	NO	0.05	56	NO	NO	3.28
15	Ronquido	NO	0.24	34	NO	NO	3.25

- En esta tabla se ilustran los perfiles de los cinco pacientes que más éxito reportan según el modelo y los cinco perfiles de pacientes que menos satisfacción/efectividad pueden tener en el tratamiento.
- En primera instancia se puede observar que los pacientes más exitosos suelen tener un puntaje de somnolencia bajo en comparación a los menos exitosos.
- En segundo lugar, se puede notar que un paciente con dificultad para mantener el sueño suele percibir el tratamiento más efectivo. En contraparte se aprecia que los roncadores son un perfil típico de un paciente que no suele considerarse satisfecho con el tratamiento.
- Como se vió en el análisis descriptivo, los pacientes psiquiátricos suelen percibir el tratamiento más exitoso que los que no poseen comorbilidades psiquiátricas. Esto es apreciable porque de los cinco perfiles con el puntaje más alto, 3 de ellos involucran el hecho de tener una comorbilidad psiquiátrica.
- En lo que respecta a la variable T90, se puede notar que un paciente con valores muy bajos puede considerar mejor el tratamiento, mientras que aquellos que poseen valores medios/altos en este aspecto calificarán peor al mismo.
- Nótese que los pacientes más mayores suelen considerar al tratamiento como más satisfactorio/efectivo, lo cuál implica directamente que los pacientes relativamente jóvenes suelen considerar menos exitoso el tratamiento.
- En la variable maloclusión clase III se aprecia que tanto los perfiles más exitosos como los menos exitosos no poseen esta condición, sin embargo el análisis descriptivo dicta que aquellos que poseen esta condición suelen calificar menos satisfactorio el tratamiento. Si el modelo indica que esta variable mejora las predicciones, entonces esta consideración es totalmente válida aún no siendo apreciable en la tabla.
- De manera similar ocurre con la variable que indica alguna clase de comorbilidad cardiovascular, si bien no es tan apreciable en la tabla, sí se alcanza a notar que aquellos pacientes que sí tienen comorbilidades cardiovasculares suelen calificar mejor al tratamiento.

- Finalmente, se concluye que se deben “gatillar” o “motivar” a los pacientes con los perfiles mencionados: Epworth basal alto, síntomas relacionados al ronquido, sin comorbilidades psiquiátricas, con valores T90 relativamente altos, jóvenes, que posean maloclusión clase III y que no posean comorbilidades cardiovasculares. Mientras que de aquellos con un epworth basal bajo, con síntomas relacionados a la dificultad para mantener el sueño, con comorbilidades psiquiátricas, con valores T90 bajos, en edades mayores, sin maloclusión clase III y con comorbilidades cardiovasculares se puede esperar de antemano un resultado efectivo/satisfactorio del tratamiento miofuncional orofacial.