



CENTRO DE  
INVESTIGACIÓN EN  
COMPLEJIDAD SOCIAL



Universidad del Desarrollo  
Facultad de Gobierno

DCCS

DOCTORADO EN  
CIENCIAS DE LA  
COMPLEJIDAD SOCIAL

## Econometría

Trastornos por consumo de sustancias y hospitalizaciones por salud mental en Chile 2010-2022

Amaru Simón Agüero Jiménez

2025-06-21

## Tabla de contenidos

<b>1 Introducción</b>	<b>3</b>
1.1 Relevancia del problema . . . . .	3
1.2 Pregunta de investigación . . . . .	3
1.3 Modelo teórico o mecanismo que explicaría la relación entre las variables que se van a analizar e hipótesis . . . . .	3
<b>2 Metodología</b>	<b>4</b>
2.1 Datos . . . . .	4
2.2 Variables de interés . . . . .	4
2.3 Modelos . . . . .	4
<b>3 Resultados</b>	<b>7</b>
3.1 Estadísticas Descriptivas . . . . .	7
3.2 Modelos de Regresión Logística . . . . .	7
3.2.1 Modelo Base y Progresión . . . . .	7
3.2.2 Predictor tipo de sustancias (Modelo 7) . . . . .	7
3.2.3 Características Demográficas . . . . .	7
3.2.4 Características del Tratamiento . . . . .	7
3.2.5 Historial de Salud Mental . . . . .	8

3.2.6 Retratamientos . . . . .	8
3.2.7 Variables Socioeconómicas . . . . .	8
3.2.8 Interacciones Sexo × Sustancia . . . . .	8
3.3 Modelos Probit . . . . .	8
3.4 Modelos Logit con Ajuste de Firth . . . . .	9
3.5 Comparación de Métodos de Estimación . . . . .	9
3.6 Rendimiento Predictivo . . . . .	9
3.7 Diagnóstico de Modelos . . . . .	9
<b>4 Discusión y Conclusiones</b>	<b>20</b>
4.1 Marco teórico planteado y resultados obtenidos . . . . .	20
4.2 Limitaciones del estudio . . . . .	20
<b>5 Repositorio GitHub</b>	<b>21</b>
<b>Referencias</b>	<b>22</b>

## 1 Introducción

### 1.1 Relevancia del problema

Los trastornos por consumo de sustancias (TCS) constituyen una de las principales causas de carga de enfermedad y mortalidad evitable a escala global. En 2016 se calculó que más de 100 millones de personas sufrían trastorno por consumo de alcohol y decenas de millones presentaban dependencia de opioides, cannabis o cocaína<sup>1</sup>. La frecuente comorbilidad psiquiátrica, depresión, trastornos de ansiedad, psicosis o trastornos de personalidad multiplica la severidad clínica y los costes sociosanitarios<sup>2</sup>. Estudios hospitalarios europeos y norteamericanos muestran que alrededor del 20 % de las admisiones psiquiátricas corresponden a pacientes con diagnóstico dual, fenómeno que favorece re-ingresos y estancias prolongadas<sup>3</sup>.

En Chile, las encuestas nacionales sitúan la prevalencia de abuso o dependencia de sustancias entre el 11 % y el 20 %, una de las más elevadas de Latinoamérica.. Los registros hospitalarios concuerdan con las cifras internacionales: alrededor de una quinta parte de los internados en psiquiatría presenta un TCS como diagnóstico primario o secundario. Esta convergencia evidencia que la hospitalización psiquiátrica es un desenlace clínico crítico en la trayectoria de las adicciones, razón por la cual identificar sus factores determinantes resulta esencial para planificar intervenciones preventivas, asignar recursos y reducir la carga asistencial<sup>2,4-6</sup>.

### 1.2 Pregunta de investigación

¿Qué factores individuales y clínicos se asocian a la hospitalización psiquiátrica dentro del tiempo posterior al alta de un tratamiento por consumo de sustancias, en adultos atendidos en la red chilena entre 2010-2022?

### 1.3 Modelo teórico o mecanismo que explicaría la relación entre las variables que se van a analizar e hipótesis

Se parte de un marco de riesgo acumulativo en el que la probabilidad de hospitalización ( $p$ ) surge de la interacción entre:

- Severidad adictiva: tipo de sustancia principal y número de re-tratamientos, como indicador de recaídas y cronicidad.
- Vulnerabilidad psiquiátrica previa: presencia y número de hospitalizaciones mentales antes del tratamiento por drogas.
- Características del tratamiento: modalidad (ambulatoria vs residencial) y duración total del episodio.
- Variables sociodemográficos: edad (lineal y cuadrática), sexo, estado civil, nivel educacional y región de residencia.

Bajo este esquema se plantean cuatro hipótesis:

$H_1$  – Historial psiquiátrico: cada hospitalización mental previa incrementa significativamente la probabilidad de una nueva hospitalización ( $\uparrow p$ ).

$H_2$  – Re-tratamientos: los pacientes con tres o más reingresos muestran un riesgo superior al de quienes no han recaído.

$H_3$  – Tratamiento residencial: refleja mayor complejidad clínica y eleva la probabilidad de hospitalización frente al tratamiento ambulatorio.

$H_4$  – Sustancias depresoras del SNC: por su perfil farmacológico, conllevan mayor riesgo de complicaciones psiquiátricas agudas.

Este modelo conceptual apoyado por la evidencia internacional sobre patología dual y cursos clínicos severos guiará la exploración empírica de los datos chilenos<sup>5</sup>. Evaluar dichas hipótesis permitirá distinguir los pacientes de alto riesgo y orientar estrategias integradas que reduzcan la probabilidad de hospitalización, mejoren la continuidad asistencial y optimicen el uso de recursos en salud mental tanto en Chile como en otros contextos comparables.

## 2 Metodología

### 2.1 Datos

Este es un estudio de cohorte retrospectiva de pacientes adultos en tratamiento por consumo de sustancias, con datos otorgados por el Servicio Nacional para la Prevención y Rehabilitación del Consumo de Drogas y Alcohol de Chile (SENDA) en convenio con el núcleo milenio de análisis de políticas públicas de drogas (nDP). La cohorte se construyó vinculando los registros administrativos de los pacientes ( $n = 156.771$  episodios de tratamiento entre 97,698 personas en las 16 regiones del país) con los datos de egresos hospitalarios a nivel nacional entre 2010 y 2022 (20.652.003 hospitalizaciones). El registro de pacientes en tratamiento se realizó en una plataforma electrónica denominada SISTRAT, que contenía información sociodemográfica, datos sobre el estado de salud y patrones de consumo de sustancias, entre otras variables, además de información sobre el propio tratamiento (p. ej., fecha de ingreso, egreso, tipo de tratamiento). Los conjuntos de datos de egresos hospitalarios son gestionados por el Departamento de Estadísticas e Información de Salud del Ministerio de Salud e incluyen información sobre todas las hospitalizaciones en centros públicos y privados, así como la causa principal y secundaria de ingreso, la fecha de ingreso y el alta, entre otras variables. Las bases de datos se vincularon de forma determinista mediante un hash de 64 caracteres resultante del cifrado (con un algoritmo SHA-256) del número de identificación único de cada persona.

### 2.2 Variables de interés

Se incluyeron como principal **predictores** múltiples variables relacionadas al **consumo y tratamiento rehabilitador de drogas**. Se consideró la *sustancia principal* por la cual se trató al paciente (alcohol, pasta base de cocaína, cocaína, marihuana, depresores del SNC u otras sustancias; tomando Alcohol como categoría de referencia), el *tipo de plan de tratamiento* (ambulatorio vs. residencial), la duración del tratamiento, el **número de reingresos** a tratamiento (retratamientos, categorizados en 0, 1, 2, 3 o más reingresos) y el **motivo de egreso** del programa (alta terapéutica, derivación, abandono temprano/tardío, etc.). Entre las variables **sociodemográficas** consideradas están la edad (en años) al inicio del tratamiento incorporada también en forma cuadrática (edad<sup>2</sup>) para capturar posibles efectos no lineales, el sexo del paciente, el **estado civil** (por ejemplo, soltero, casado/conviviente, separado, viudo), el **nivel educacional** alcanzado (educación primaria, secundaria, técnica, universitaria) y la situación de **empleo** (categorías como trabajando, desempleado, estudiando, tareas del hogar, jubilado, etc.). Se justó con otras variables del **historial clínico** de salud mental del paciente a través del número de hospitalizaciones psiquiátricas previas al tratamiento de drogas. Todas las variables categóricas se introdujeron mediante codificación *dummy*, definiendo categorías de referencia apropiadas (p. ej., sexo femenino, tratamiento ambulatorio, etc.)

### 2.3 Modelos

El estudio emplea modelos de regresión logística binaria (logit) y probit para analizar los predictores de hospitalización psiquiátrica **posterior** al tratamiento por consumo de sustancias. Se compararon tres métodos de estimación para evaluar la robustez de los resultados. El modelo logit estándar estimó la probabilidad del evento mediante la función de enlace logit:

$$\log \left( \frac{p_i}{1 - p_i} \right) = \mathbf{x}_i^T \boldsymbol{\beta}$$

El modelo probit utilizó la función de distribución acumulada normal estándar:

$$\Phi^{-1}(p_i) = \mathbf{x}_i^T \boldsymbol{\beta}$$

Finalmente, se implementó el método de regresión logística penalizada de Firth para abordar posibles problemas de separación cuasi-completa. Este método maximizó la log-verosimilitud penalizada:

$$\ell^*(\boldsymbol{\beta}) = \ell(\boldsymbol{\beta}) + \frac{1}{2} \log |\mathbf{I}(\boldsymbol{\beta})|$$

donde  $\mathbf{I}(\boldsymbol{\beta})$  fue la matriz de información de Fisher. La penalización de Firth redujo el sesgo en las estimaciones, particularmente relevante en presencia de eventos raros o categorías con pocas observaciones.

En los tres métodos,  $p_i = P(Y = 1 | X)$  es la probabilidad de hospitalización psiquiátrica dada la presencia del vector de predictores  $\mathbf{x}_i^T$ . En ambos casos,  $Y$  es una variable dicotómica indicadora de si el paciente fue hospitalizado en un recinto psiquiátrico tras completar el tratamiento de rehabilitación, y el vector  $\boldsymbol{\beta}$  representan los coeficientes de cada variable explicativa  $\mathbf{x}_i^{T7-9}$ .

El **desarrollo de los modelos** siguió un enfoque escalonado. Se comenzó estimando un modelo base con un conjunto inicial de covariables fundamentales, al cual se le fueron incorporando progresivamente bloques adicionales de variables para evaluar su contribución incremental al ajuste. La significancia estadística de cada bloque añadido se evaluó mediante pruebas de **razón de verosimilitud** (LR) para modelos anidados<sup>10</sup>. Para la comparación de modelos se calculó el **criterio de información de Akaike** (AIC) de cada especificación, considerando preferible aquel modelo con menor AIC (indicativo de mejor equilibrio entre buen ajuste y parsimonia). Asimismo, se estimó el **área bajo la curva ROC** (AUC) para cada modelo, como medida de desempeño predictivo (discriminación entre pacientes con y sin hospitalización)<sup>11</sup>. Con objeto de caracterizar integralmente el rendimiento predictivo, se estimaron además las siguientes métricas, definidas a partir de la matriz de confusión (verdaderos positivos TP, falsos positivos FP, verdaderos negativos TN y falsos negativos FN):

- Sensibilidad:

$$\text{Sensibilidad} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

- Especificidad:

$$\text{Especificidad} = \frac{\text{TN}}{\text{TN} + \text{FP}}$$

- Precisión (valor predictivo positivo, PPV):

$$\text{PPV} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

- Valor predictivo negativo (NPV):

$$\text{NPV} = \frac{\text{TN}}{\text{TN} + \text{FN}}$$

- Exactitud global:

$$\text{Exactitud} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}}$$

Estas métricas complementaron la información aportada por la AUC y permiten describir, de forma robusta, la capacidad de los modelos para identificar correctamente los casos de hospitalización psiquiátrica y minimizar los errores de clasificación<sup>12,13</sup>.

Una vez seleccionados los mejores modelos según las métricas anteriores se evaluaron su calibración mediante la comparación entre las probabilidades predichas y las proporciones observadas. Las probabilidades predichas  $\hat{p}_i$  se agruparon en deciles basados en sus cuantiles. Para cada decil  $j$ , se calculó la probabilidad predicha promedio  $\bar{\hat{p}}_j = \frac{1}{n_j} \sum_{i \in G_j} \hat{p}_i$  y la proporción observada  $\bar{y}_j = \frac{1}{n_j} \sum_{i \in G_j} y_i$ . Se construyeron intervalos de confianza del 95% utilizando el error estándar  $SE_j = \sqrt{\frac{\bar{y}_j(1-\bar{y}_j)}{n_j}}$ .

La bondad de ajuste global se evaluó mediante el test de Hosmer-Lemeshow:

$$\chi^2_{HL} = \sum_{j=1}^{10} \frac{(O_j - E_j)^2}{E_j(1 - \bar{\hat{p}}_j)} + \frac{(n_j - O_j - n_j + E_j)^2}{(n_j - E_j)\bar{\hat{p}}_j}$$

donde  $O_j$  representó los eventos observados y  $E_j$  los eventos esperados en cada grupo. Bajo la hipótesis nula de buena calibración, el estadístico siguió una distribución  $\chi^2$  con 8 grados de libertad<sup>14</sup>.

Luego se calcularon los residuos de deviance para identificar patrones sistemáticos y observaciones atípicas:

$$d_i = \text{sign}(y_i - \hat{p}_i) \sqrt{2 \left[ y_i \log \left( \frac{y_i}{\hat{p}_i} \right) + (1 - y_i) \log \left( \frac{1 - y_i}{1 - \hat{p}_i} \right) \right]}$$

Los residuos se graficaron contra los valores ajustados y se aplicó un suavizado loess para detectar desviaciones de la linealidad. Se identificaron como valores atípicos aquellas observaciones con  $|d_i| > 3$ .

Por último se calculó el Factor de Inflación de Varianza Generalizado (GVIF) para evaluar la multicolinealidad entre predictores. Para cada variable o conjunto de variables dummy  $j$ , el GVIF se definió como:

$$\text{GVIF}_j = \det(\mathbf{R}_j) \det(\mathbf{R}_{-j})^{-1}$$

donde  $\mathbf{R}_j$  es la matriz de correlación del subconjunto de variables asociadas con el término  $j$ , y  $\mathbf{R}_{-j}$  es la matriz de correlación de todas las demás variables predictoras. El determinante  $\det(\mathbf{R})$  mide el volumen del elipsoide de correlación; un determinante cercano a cero indica alta multicolinealidad. Dado que el GVIF depende del número de parámetros asociados a cada variable, se aplicó la corrección  $\text{GVIF}^{1/(2 \times df_j)}$ , donde  $df_j$  representó los grados de libertad de la variable  $j$ . Se consideraron problemáticos valores superiores a  $\sqrt{5} \approx 2.24$ , indicando que la varianza del coeficiente se infló más de 5 veces debido a la colinealidad<sup>15</sup>.

Todos los análisis consideraron un nivel de significancia de  $\alpha = 0.05$  e intervalos de confianza con un nivel de confianza de  $1 - \alpha = 0.95$ .

## 3 Resultados

### 3.1 Estadísticas Descriptivas

La Tabla 1 presenta las características de la muestra analítica final de 97,618 pacientes adultos en tratamiento por consumo de sustancias entre 2010 y 2022. De estos, 7,456 (7.6%) experimentaron al menos una hospitalización psiquiátrica posterior al tratamiento. La edad promedio fue de 33 años (DE = 10), con predominio masculino (74%). Las sustancias principales de tratamiento fueron pasta base de cocaína (36.3%), alcohol (34.6%) y cocaína (19.7%). La mayoría recibió tratamiento ambulatorio (89.1%) frente al residencial (10.9%). Un 11% presentó hospitalizaciones psiquiátricas previas al tratamiento, con una media de 0.2 hospitalizaciones previas (DE = 0.9). El 74.5% no tuvo retratamientos, mientras que 16.7% tuvo uno, 5.3% dos, y 3.6% tres o más retratamientos.

### 3.2 Modelos de Regresión Logística

#### 3.2.1 Modelo Base y Progresión

La Tabla 2 presenta los resultados de los modelos logit estándar. El intercepto del Modelo 3 ( $\hat{\beta}_0 = -2.972$ , IC 95%: [-3.092, -2.852]) indica que los log-odds de hospitalización psiquiátrica dado el grupo de referencia (mujeres con tratamiento ambulatorio por alcohol, edad promedio, sin hospitalizaciones previas) son significativamente negativos, correspondiendo a una probabilidad basal de  $\frac{e^{-2.972}}{1+e^{-2.972}} = 0.049$  o 4.9%.

La Tabla 6 muestra las comparaciones entre modelos anidados mediante pruebas de razón de verosimilitud. La inclusión de hospitalizaciones previas (Modelo 2 vs 3) produjo la mayor reducción en deviance ( $\chi^2_{LR} = 2,681.75$ , gl = 1,  $p < 0.001$ ), seguida por la adición de retratamientos (Modelo 3 vs 4:  $\chi^2_{LR} = 3,610.47$ , gl = 3,  $p < 0.001$ ).

#### 3.2.2 Predictor tipo de sustancias (Modelo 7)

En la Tabla 2 se observa que el mayor efecto positivo sobre los log-odds de hospitalización psiquiátrica es el consumo de depresores del SNC, tomando alcohol como referencia ( $\hat{\beta} = 0.463$ , IC 95%: [0.243, 0.683]), lo que representa un odds ratio de  $e^{0.463} = 1.59$ , indicando 1.59 veces más de probabilidad de hospitalización sobre la no hospitalización, comparado con pacientes tratados por alcohol (El coeficiente se mantiene estable en todos los modelos indicando su robustez). La cocaína mostró una asociación negativa no significativa ( $\hat{\beta} = -0.019$ , IC 95%: [-0.154, 0.117]), mientras que la pasta base no mostró diferencias significativas ( $\hat{\beta} = 0.042$ , IC 95%: [-0.068, 0.153]). Estos efectos negativos si son significativos en la interacción con sexo. Esto indica que la probabilidad de la hospitalización psiquiátrica es menor en hombres que consumen pasta base o cocaína que en mujeres como se evalúa más adelante.

#### 3.2.3 Características Demográficas

El sexo masculino se asoció con menores log-odds de hospitalización ( $\hat{\beta} = -0.105$ , IC 95%: [-0.203, -0.006]), representando un odds ratio de 0.90. La edad mostró una relación no lineal significativa: el efecto lineal fue positivo ( $\hat{\beta} = 0.089$ , IC 95%: [0.074, 0.105]), mientras que el término cuadrático fue negativo ( $\hat{\beta} = -0.001$ , IC 95%: [-0.001, -0.001]), sugiriendo que el riesgo de hospitalización aumenta con la edad pero a una tasa decreciente (Tabla 2, Modelo 7).

#### 3.2.4 Características del Tratamiento

Existe una mayor probabilidad de hospitalización psiquiátrica posterior dado un plan residencial versus ambulatorio, ( $\hat{\beta} = 0.279$ , IC 95%: [0.205, 0.353]), con un odds ratio de 1.32. Comparando los modelos en la Tabla 2, este efecto

se redujo considerablemente desde el Modelo 2 ( $\hat{\beta} = 0.763$ ) al incluir hospitalizaciones previas y retratamientos, sugiriendo que parte del efecto inicial se debía a mayor severidad basal en pacientes residenciales (Sesgo por omisión negativo).

### 3.2.5 Historial de Salud Mental

El número de hospitalizaciones psiquiátricas previas fue el predictor más fuerte ( $\hat{\beta} = 0.478$ , IC 95%: [0.455, 0.500]). Hay un aumento de los odds de hospitalización posterior en  $e^{0.478} = 1.61$  veces dado un aumento de eventos de hospitalización psiquiátrica posterior. Este efecto permaneció estable a través de todos los modelos (4-7), indicando robustez ante la inclusión de covariables adicionales (Tabla 2).

### 3.2.6 Retratamientos

Los retratamientos mostraron una relación dosis-respuesta con la hospitalización psiquiátrica. Comparado con ningún retratamiento, un retratamiento aumentó los log-odds en 1.048 (IC 95%: [0.986, 1.110], OR = 2.85), dos retratamientos en 1.597 (IC 95%: [1.514, 1.679], OR = 4.94), y tres o más en 2.277 (IC 95%: [2.191, 2.362], OR = 9.75) (Tabla 2, Modelo 7).

### 3.2.7 Variables Socioeconómicas

Entre las variables educacionales (referencia: otra/no especificada), ninguna categoría mostró efectos significativos en el Modelo 7. El estado civil casado/conviviente mostró menores log-odds comparado con divorciado/separado ( $\hat{\beta} = -0.126$ , IC 95%: [-0.213, -0.039], OR = 0.88). En ocupación (referencia: desempleado), trabajar se asoció con menores log-odds ( $\hat{\beta} = -0.381$ , IC 95%: [-0.442, -0.319], OR = 0.68), mientras que jubilados/pensionados mostraron mayores log-odds ( $\hat{\beta} = 0.359$ , IC 95%: [0.179, 0.538], OR = 1.43) (Tabla 2).

### 3.2.8 Interacciones Sexo × Sustancia

La interacción entre sexo masculino y pasta base fue significativa y negativa ( $\hat{\beta} = -0.257$ , IC 95%: [-0.385, -0.129]), indicando que el efecto de la pasta base sobre la hospitalización psiquiátrica difiere entre hombres y mujeres. Para hombres tratados por pasta base, el efecto combinado es  $0.042 + (-0.257) = -0.215$ , sugiriendo menor riesgo que las mujeres tratadas por la misma sustancia (Tabla 2, Modelo 7).

## 3.3 Modelos Probit

La Tabla 3 presenta los modelos probit, mostrando patrones similares pero con coeficientes escalados según la función de enlace probit. El intercepto del Modelo 7 ( $\hat{\beta}_0 = -2.747$ ) corresponde a  $\Phi(-2.747) = 0.003$ , donde  $\Phi$  es la función de distribución acumulada normal estándar. Los efectos mantuvieron su dirección y significancia: depresores ( $\hat{\beta} = 0.240$ ,  $p < 0.001$ ), edad ( $\hat{\beta} = 0.040$ ,  $p < 0.001$ ), plan residencial ( $\hat{\beta} = 0.155$ ,  $p < 0.001$ ), y hospitalizaciones previas ( $\hat{\beta} = 0.242$ ,  $p < 0.001$ ). Las comparaciones entre modelos anidados (Tabla 7) mostraron patrones similares a los modelos logit, con la mayor mejora al incluir hospitalizaciones previas ( $\chi^2_{LR} = 2,673.74$ , gl = 1,  $p < 0.001$ ).

### 3.4 Modelos Logit con Ajuste de Firth

La Tabla 4 presenta los resultados del método de Firth, diseñado para abordar problemas de separación. Los coeficientes fueron virtualmente idénticos a los del logit estándar, sugiriendo ausencia de problemas de separación en estos datos. El Modelo 7 de Firth mostró el mismo patrón de efectos significativos, con mínimas diferencias en las estimaciones puntuales. Las comparaciones entre modelos (Tabla 8) replicaron los hallazgos de los otros métodos.

### 3.5 Comparación de Métodos de Estimación

La Tabla 5 y 9 muestra la comparación de AIC y AUC entre los tres métodos de estimación. Los modelos logit estándar y Firth produjeron resultados prácticamente idénticos, con el AIC más bajo en el Modelo 7 (44,540.24 y 44,540.33 respectivamente), mientras que el modelo probit mostró un AIC ligeramente superior (44,623.88). Las métricas de discriminación fueron consistentes entre métodos, con AUC de 0.779 para logit, 0.779 para probit, y 0.779 para Firth en el Modelo 7.

### 3.6 Rendimiento Predictivo

La Tabla 9 compara las métricas de rendimiento para todos los modelos. El Modelo 7 logit mostró: sensibilidad 63.6%, especificidad 79.3%, y exactitud global 78.1%. El valor predictivo positivo fue 20.2% mientras que el valor predictivo negativo alcanzó 96.3%, reflejando la baja prevalencia del evento. Métricas similares se obtuvieron para probit (sensibilidad 64.3%, especificidad 78.6%) y Firth (sensibilidad 63.6%, especificidad 79.2%).

La Figura 1 presenta las curvas ROC para todos los modelos. La progresión del AUC desde el Modelo 3 (0.692) al Modelo 7 (0.779) indica que las hospitalizaciones previas y retratamientos aportan la mayor capacidad discriminativa, mientras que las variables socioeconómicas y regionales contribuyen marginalmente. La comparación entre Modelo 6 y 7 ( $\Delta\text{AUC} = 0.002$ ) sugiere que la complejidad adicional del modelo completo ofrece ganancia predictiva limitada.

### 3.7 Diagnóstico de Modelos

La Figura 2, 3 y 4 presenta los diagnósticos para las tres variantes del Modelo 7. La calibración evaluada mediante el test de Hosmer-Lemeshow reveló falta de ajuste significativa en los tres modelos: logit estándar ( $\chi^2_{HL} = 33.54$ ,  $gl = 8$ ,  $p < 0.001$ ), probit ( $\chi^2_{HL} = 46.71$ ,  $gl = 8$ ,  $p < 0.001$ ), y Firth ( $\chi^2_{HL} = 33.32$ ,  $gl = 8$ ,  $p < 0.001$ ). Los gráficos de calibración mostraron subestimación sistemática en probabilidades altas ( $>0.6$ ) para los tres métodos.

El análisis de multicolinealidad mediante GVIF mostró valores aceptables para todas las variables en los tres modelos. Los valores  $GVIF^{1/(2 \times df)}$  más altos correspondieron a edad<sup>2</sup> (aproximadamente 44.0) debido a su correlación con el término lineal, y región (aproximadamente 38.0) por sus múltiples categorías. Ninguna variable superó el umbral problemático después del ajuste por grados de libertad.

Los residuos de deviance no mostraron patrones sistemáticos contra valores ajustados en ninguno de los tres modelos, aunque se observó mayor variabilidad en probabilidades extremas, consistente con la naturaleza binaria del resultado.

Tabla 1: Estadísticas descriptivas según el estado de hospitalización psiquiátrica

Variable	Level	No Mental Hosp	Mental Hosp	Total
<b>Sample Size</b>		90239	7456	97695
<b>First Treatment Duration (days)</b>		228.3 (198.3, 0.0, 93.0, 173.0, 305.0, 4056.0)	226.3 (220.6, 0.0, 84.0, 162.0, 300.0, 2947.0)	228.1 (200.1, 0.0, 92.0, 172.0, 305.0, 4056.0)
<b>Total Treatment Duration (days)</b>		300.8 (274.9, 0.0, 115.0, 222.0, 397.0, 6168.0)	463.4 (436.0, 0.0, 154.0, 333.0, 625.0, 4258.0)	313.2 (293.5, 0.0, 117.0, 229.0, 411.0, 6168.0)
<b>Number of Previous Mental Hospitalizations</b>		0.1 (0.7, 0.0, 0.0, 0.0, 0.0, 63.0)	1.0 (2.2, 0.0, 0.0, 0.0, 1.0, 31.0)	0.2 (0.9, 0.0, 0.0, 0.0, 0.0, 63.0)
<b>Number of Re-treatments</b>		0.3 (0.8, 0.0, 0.0, 0.0, 0.0, 11.0)	1.1 (1.5, 0.0, 0.0, 1.0, 2.0, 14.0)	0.4 (0.9, 0.0, 0.0, 0.0, 0.0, 14.0)
<b>Total Number of Treatments</b>		1.3 (0.8, 1.0, 1.0, 1.0, 1.0, 12.0)	2.1 (1.5, 1.0, 1.0, 2.0, 3.0, 15.0)	1.4 (0.9, 1.0, 1.0, 1.0, 2.0, 15.0)
<b>Sex</b>	Female	22778 (25.2%)	2627 (35.2%)	25405 (26.0%)
	Male	67461 (74.8%)	4829 (64.8%)	72290 (74.0%)
<b>Principal Substance</b>	Alcohol	31137 (34.5%)	2633 (35.3%)	33770 (34.6%)
	Other substances	400 (0.4%)	52 (0.7%)	452 (0.5%)
	Cocaine	18014 (20.0%)	1190 (16.0%)	19204 (19.7%)
	Depressants	1372 (1.5%)	237 (3.2%)	1609 (1.6%)
	Marijuana	6737 (7.5%)	470 (6.3%)	7207 (7.4%)
	Cocaine paste	32579 (36.1%)	2874 (38.5%)	35453 (36.3%)
<b>Plan Type</b>	Outpatient	81024 (89.8%)	5987 (80.3%)	87011 (89.1%)
	Residential	9215 (10.2%)	1469 (19.7%)	10684 (10.9%)
<b>Previous Mental Hospitalization</b>	No	82324 (91.2%)	4660 (62.5%)	86984 (89.0%)
	Yes	7915 (8.8%)	2796 (37.5%)	10711 (11.0%)
	0	69368 (76.9%)	3375 (45.3%)	72743 (74.5%)
<b>Re-treatments Category</b>	1	14276 (15.8%)	1994 (26.7%)	16270 (16.7%)
	2	4185 (4.6%)	989 (13.3%)	5174 (5.3%)
	3+	2410 (2.7%)	1098 (14.7%)	3508 (3.6%)
<b>Discharge Reason</b>	Late Dropout	31440 (34.8%)	2246 (30.1%)	33686 (34.5%)
	Early Dropout	14023 (15.5%)	1126 (15.1%)	15149 (15.5%)
	Administrative Discharge	7165 (7.9%)	703 (9.4%)	7868 (8.1%)
	Therapeutic Discharge	23731 (26.3%)	1467 (19.7%)	25198 (25.8%)
	Referral	13879 (15.4%)	1914 (25.7%)	15793 (16.2%)
	NA	1 (0.0%)	0 (0.0%)	1 (0.0%)
<b>Marital Status</b>	Divorced / Separated	9359 (10.4%)	982 (13.2%)	10341 (10.6%)
	Married / Partnered	30592 (33.9%)	2158 (28.9%)	32750 (33.5%)
	Other / No answer	185 (0.2%)	18 (0.2%)	203 (0.2%)
	Single	49038 (54.3%)	4194 (56.2%)	53232 (54.5%)
	Widowed	1065 (1.2%)	104 (1.4%)	1169 (1.2%)
<b>Education Level</b>	Primary or less	23823 (26.4%)	1970 (26.4%)	25793 (26.4%)
	Secondary	49018 (54.3%)	3822 (51.3%)	52840 (54.1%)
	University or higher	7329 (8.1%)	739 (9.9%)	8068 (8.3%)
	Unknown	415 (0.5%)	39 (0.5%)	454 (0.5%)
	Technical	9578 (10.6%)	886 (11.9%)	10464 (10.7%)
	NA	76 (0.1%)	0 (0.0%)	76 (0.1%)
<b>Employment Status</b>	Unemployed	28646 (31.7%)	3102 (41.6%)	31748 (32.5%)
	Studying	1306 (1.4%)	86 (1.2%)	1392 (1.4%)
	Other	7890 (8.7%)	775 (10.4%)	8665 (8.9%)
	Retired / Pensioned	1358 (1.5%)	234 (3.1%)	1592 (1.6%)
	Household tasks	5328 (5.9%)	632 (8.5%)	5960 (6.1%)
	Working	45711 (50.7%)	2627 (35.2%)	48338 (49.5%)
<b>Treatment Center Region</b>	De Antofagasta	3574 (4.0%)	207 (2.8%)	3781 (3.9%)
	De Arica Y Parinacota	2096 (2.3%)	261 (3.5%)	2357 (2.4%)
	De Atacama	2459 (2.7%)	108 (1.4%)	2567 (2.6%)
	De Aysen Del General Carlos Ibáñez Del Campo	964 (1.1%)	160 (2.1%)	1124 (1.2%)
	De Coquimbo	3661 (4.1%)	139 (1.9%)	3800 (3.9%)
	De La Araucanía	2555 (2.8%)	246 (3.3%)	2801 (2.9%)
	De Los Lagos	3419 (3.8%)	403 (5.4%)	3822 (3.9%)
	De Los Ríos	1471 (1.6%)	152 (2.0%)	1623 (1.7%)
	De Magallanes Y La Antártica Chilena	1089 (1.2%)	195 (2.6%)	1284 (1.3%)
	De Nuble	687 (0.8%)	83 (1.1%)	770 (0.8%)
	De Tarapaca	2019 (2.2%)	309 (4.1%)	2328 (2.4%)
	De Valparaíso	6681 (7.4%)	866 (11.6%)	7547 (7.7%)
	Del Bío-Bío	6293 (7.0%)	563 (7.6%)	6856 (7.0%)
	Del Libertador General Bernardo O'Higgins	4725 (5.2%)	315 (4.2%)	5040 (5.2%)
	Del Maule	5401 (6.0%)	375 (5.0%)	5776 (5.9%)
	Metropolitana	43144 (47.8%)	3074 (41.2%)	46218 (47.3%)
	NA	1 (0.0%)	0 (0.0%)	1 (0.0%)

Note:  
For numeric variables: mean (sd, min, q1, median, q3, max). For categorical variables: n (%)

Tabla 2: Modelos de regresión logística (Logit): Predictores de la hospitalización psiquiátrica posterior al tratamiento rehabilitador de drogas

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Intercept	-2.756*** [-2.871, -2.641]	-2.831*** [-2.946, -2.715]	-2.972*** [-3.092, -2.852]	-3.593*** [-3.721, -3.464]	-3.549*** [-3.943, -3.156]	-3.800*** [-4.225, -3.375]	-5.395*** [-5.923, -4.866]
Substance: Other substances	0.454** [0.160, 0.747]	0.422** [0.128, 0.717]	0.283 [0.030, 0.596]	0.354* [0.031, 0.677]	0.336* [0.013, 0.659]	0.271 [-0.051, 0.593]	0.026 [-0.536, 0.587]
Substance: Cocaine	-0.136*** [-0.210, -0.062]	-0.166*** [-0.240, -0.092]	-0.086* [-0.162, -0.010]	-0.191*** [-0.269, -0.113]	-0.188*** [-0.267, -0.110]	-0.120** [-0.201, -0.038]	-0.019 [-0.154, 0.117]
Substance: Depressants	0.569*** [0.424, 0.715]	0.544*** [0.398, 0.690]	0.458*** [0.304, 0.612]	0.485*** [0.325, 0.645]	0.463*** [0.303, 0.623]	0.435*** [0.273, 0.596]	0.463*** [0.243, 0.683]
Substance: Marijuana	-0.038 [-0.143, 0.068]	-0.025 [-0.130, 0.081]	-0.023 [-0.133, 0.086]	-0.013 [-0.125, 0.099]	-0.019 [-0.131, 0.093]	0.013 [-0.101, 0.127]	-0.093 [-0.301, 0.116]
Substance: Cocaine paste	0.150** [0.091, 0.208]	0.007 [-0.053, 0.068]	0.101** [0.039, 0.163]	-0.149*** [-0.214, -0.084]	-0.138*** [-0.203, -0.073]	-0.105** [-0.175, -0.036]	0.042 [-0.068, 0.153]
Age	0.013*** [0.011, 0.016]	0.014*** [0.011, 0.016]	0.011*** [0.009, 0.013]	0.013*** [0.011, 0.016]	0.015*** [0.013, 0.018]	0.015*** [0.012, 0.018]	0.089*** [0.074, 0.105]
Sex: Male	-0.470*** [-0.521, -0.420]	-0.447*** [-0.498, -0.396]	-0.382*** [-0.434, -0.329]	-0.245*** [-0.300, -0.191]	-0.251*** [-0.305, -0.196]	-0.215*** [-0.275, -0.155]	-0.105* [-0.203, -0.006]
Plan: Residential	0.763*** [0.699, 0.827]	0.527*** [0.459, 0.594]	0.421*** [0.351, 0.490]	0.402*** [0.332, 0.472]	0.284*** [0.210, 0.358]	0.279*** [0.205, 0.353]	
Nº previous mental hosp.		0.520*** [0.498, 0.542]	0.524*** [0.502, 0.546]	0.517*** [0.495, 0.540]	0.483*** [0.460, 0.505]	0.478*** [0.455, 0.500]	
Retreatments: 1			1.063*** [1.002, 1.124]	1.063*** [1.002, 1.124]	1.059*** [0.998, 1.121]	1.048*** [0.986, 1.110]	
Retreatments: 2			1.619*** [1.538, 1.701]	1.619*** [1.537, 1.700]	1.615*** [1.533, 1.697]	1.597*** [1.514, 1.679]	
Retreatments: 3+			2.291*** [2.206, 2.375]	2.288*** [2.203, 2.372]	2.302*** [2.216, 2.387]	2.277*** [2.191, 2.362]	
Education: Primary				-0.078	-0.091	-0.066	
Education: Secondary				-0.439, 0.284]	-0.454, 0.272]	-0.430, 0.297]	
Education: University				-0.109	-0.111	-0.097	
Education: Higher technical				-0.469, 0.251]	-0.473, 0.251]	-0.459, 0.266]	
Marital status: Married/Partnered				-0.306, 0.429]	-0.321, 0.419]	-0.322, 0.419]	
Marital status: Other/No answer				0.015	0.021	0.016	
Marital status: Single				-0.351, 0.380]	-0.347, 0.389]	-0.352, 0.385]	
Marital status: Widowed				-0.153***	-0.123**	-0.126**	
Occupation: Studying				-0.239, -0.067]	-0.210, -0.036]	-0.213, -0.039]	
Occupation: Other				0.085	-0.001	0.029	
Occupation: Retired/Pensioned				-0.443, 0.613]	-0.539, 0.536]	-0.508, 0.567]	
Occupation: Household tasks				0.015	0.003	0.016	
Occupation: Working				-0.070, 0.100]	-0.083, 0.089]	-0.070, 0.101]	
Age <sup>2</sup>				-0.186	-0.244*	-0.113	
Interaction: Male × Other substances				-0.415, 0.042]	-0.475, -0.013]	-0.345, 0.120]	
Interaction: Male × Cocaine					-0.251*	-0.129	
Interaction: Male × Depressants					-0.488, -0.014]	-0.367, 0.110]	
Interaction: Male × Marijuana					-0.139**	-0.119*	
Interaction: Male × Cocaine paste					-0.230, -0.047]	-0.211, -0.027]	
<b>Model statistics</b>					0.128	0.359***	
N observations	97,618	97,618	97,618	97,618	97,618	97,618	97,618
Log Likelihood	-26023.76	-25776.08	-24435.20	-22629.97	-22602.34	-22290.06	-22223.12
Deviance	52047.51	51552.15	48870.41	45259.94	45204.68	44580.12	44446.24
AIC	52063.51	51570.15	48890.41	45285.94	45246.68	44662.12	44540.24
BIC	52139.42	51655.55	48985.29	45409.29	45445.94	45051.16	44986.21
Pseudo R <sup>2</sup> (McFadden)	0.012	0.021	0.072	0.141	0.142	0.154	0.156
Pseudo R <sup>2</sup> (Nagelkerke)	0.016	0.028	0.092	0.176	0.177	0.191	0.194

Note:

\*\*\* p < 0.001, \*\* p < 0.01, \* p < 0.05

95% Confidence intervals in brackets

Reference categories: Alcohol (substance), Female (sex), Outpatient (plan), etc.

Models 6-7 additionally adjusted for Region

Model 7 includes Sex × Substance interaction and Age<sup>2</sup>

Tabla 3: Modelos de regresión Probit: Predictores de la hospitalización psiquiátrica posterior al tratamiento rehabilitador de drogas

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Intercept	-1.563*** [-1.620, -1.506]	-1.600*** [-1.658, -1.543]	-1.670*** [-1.729, -1.611]	-1.958*** [-2.021, -1.895]	-1.917*** [-2.113, -1.721]	-2.040*** [-2.252, -1.828]	-2.747*** [-3.006, -2.488]
Substance: Other substances	0.239** [0.086, 0.393]	0.218** [0.064, 0.373]	0.145 [0.017, 0.306]	0.174* [0.006, 0.341]	0.163 [0.004, 0.331]	0.131 [0.037, 0.299]	-0.009 [0.297, 0.278]
Substance: Cocaine	-0.067*** [-0.103, -0.031]	-0.081*** [-0.117, -0.045]	-0.041* [-0.078, -0.005]	-0.095*** [-0.134, -0.057]	-0.093*** [-0.132, -0.055]	-0.061** [-0.101, -0.021]	-0.013 [0.082, 0.056]
Substance: Depressants	0.304*** [0.226, 0.383]	0.291*** [0.212, 0.370]	0.246** [0.165, 0.328]	0.254*** [0.170, 0.339]	0.245*** [0.160, 0.330]	0.230*** [0.144, 0.315]	0.240*** [0.123, 0.356]
Substance: Marijuana	-0.016 [-0.067, 0.035]	-0.010 [-0.061, 0.042]	-0.014 [-0.067, 0.039]	-0.006 [-0.061, 0.048]	-0.010 [-0.065, 0.045]	0.006 [-0.050, 0.062]	-0.052 [0.156, 0.051]
Substance: Cocaine paste	0.073** [0.043, 0.102]	0.007 [-0.023, 0.037]	0.050** [0.020, 0.081]	-0.067*** [-0.099, -0.035]	-0.062*** [-0.094, -0.029]	-0.046** [-0.080, -0.011]	0.028 [0.029, 0.085]
Age	0.007*** [0.006, 0.008]	0.007*** [0.006, 0.008]	0.006*** [0.004, 0.007]	0.007*** [0.005, 0.008]	0.008*** [0.006, 0.009]	0.007*** [0.006, 0.009]	0.040*** [0.033, 0.048]
Sex: Male	-0.234*** [-0.259, -0.209]	-0.223*** [-0.248, -0.197]	-0.194*** [-0.220, -0.167]	-0.127*** [-0.154, -0.099]	-0.129*** [-0.157, -0.101]	-0.112*** [-0.142, -0.081]	-0.058* [-0.108, -0.009]
Plan: Residential	0.388*** [0.354, 0.421]	0.273*** [0.238, 0.308]	0.226*** [0.189, 0.262]	0.217*** [0.180, 0.254]	0.158*** [0.119, 0.196]	0.155*** [0.116, 0.194]	
Nº previous mental hosp.		0.260*** [0.248, 0.272]	0.261*** [0.249, 0.273]	0.258*** [0.246, 0.270]	0.243*** [0.231, 0.255]	0.242*** [0.229, 0.254]	
Retreatments: 1			0.517*** [0.487, 0.548]	0.518*** [0.487, 0.548]	0.517*** [0.486, 0.548]	0.512*** [0.481, 0.543]	
Retreatments: 2			0.819*** [0.775, 0.863]	0.819*** [0.775, 0.863]	0.816*** [0.771, 0.860]	0.808*** [0.763, 0.852]	
Retreatments: 3+			1.209*** [1.161, 1.256]	1.207*** [1.159, 1.255]	1.212*** [1.164, 1.261]	1.201*** [1.152, 1.249]	
Education: Primary				-0.053 [-0.233, 0.127]	-0.056 [-0.238, 0.126]	-0.043 [-0.226, 0.139]	
Education: Secondary				-0.072 [-0.251, 0.107]	-0.068 [-0.249, 0.113]	-0.060 [-0.242, 0.121]	
Education: University				0.017 [-0.167, 0.200]	0.014 [-0.171, 0.200]	0.016 [-0.170, 0.201]	
Education: Higher technical				-0.011 [-0.193, 0.172]	-0.002 [-0.186, 0.182]	-0.002 [-0.187, 0.182]	
Marital status: Married/Partnered				-0.078*** [-0.122, -0.035]	-0.065** [-0.109, -0.022]	-0.068** [-0.112, -0.024]	
Marital status: Other/No answer				0.006 [-0.264, 0.277]	-0.040 [-0.315, 0.235]	-0.022 [-0.297, 0.252]	
Marital status: Single				0.005 [-0.038, 0.048]	-0.004 [-0.047, 0.040]	0.000 [-0.043, 0.044]	
Marital status: Widowed				-0.095 [-0.211, 0.020]	-0.126* [-0.243, -0.009]	-0.065 [-0.183, 0.052]	
Occupation: Studying					-0.128* [-0.243, -0.012]	-0.074 [-0.191, 0.042]	
Occupation: Other					-0.074** [-0.121, -0.028]	-0.064** [-0.111, -0.018]	
Occupation: Retired/Pensioned					0.067 [-0.024, 0.157]	0.170*** [0.076, 0.265]	
Occupation: Household tasks					-0.044 [-0.100, 0.012]	-0.047 [-0.103, 0.010]	
Occupation: Working					-0.178*** [-0.209, -0.148]	-0.186*** [-0.217, -0.155]	
Age <sup>2</sup>					-0.000*** [-0.000, -0.000]		
Interaction: Male × Other substances					0.257 [-0.098, 0.612]		
Interaction: Male × Cocaine					-0.078 [-0.159, 0.003]		
Interaction: Male × Depressants					0.061 [-0.112, 0.234]		
Interaction: Male × Marijuana					0.096 [-0.025, 0.216]		
Interaction: Male × Cocaine paste					-0.126*** [-0.191, -0.060]		
<b>Model statistics</b>							
N observations	97,618	97,618	97,618	97,618	97,618	97,618	97,618
Log Likelihood	-26021.79	-25776.43	-24439.56	-22653.35	-22625.26	-22325.50	-22264.94
Deviance	52043.57	51552.86	48879.12	45306.71	45250.52	44651.00	44529.88
AIC	52059.57	51570.86	48899.12	45332.71	45292.52	44733.00	44623.88
BIC	52135.48	51656.26	48994.01	45456.06	45491.78	45122.04	45069.85
Pseudo R <sup>2</sup> (McFadden)	0.012	0.021	0.072	0.140	0.141	0.152	0.155
Pseudo R <sup>2</sup> (Nagelkerke)	0.016	0.028	0.092	0.174	0.176	0.189	0.192

Note:

\*\*\* p < 0.001, \*\* p < 0.01, \* p < 0.05

95% Confidence intervals in brackets

Reference categories: Alcohol (substance), Female (sex), Outpatient (plan), etc.

Models 6-7 additionally adjusted for Region

Model 7 includes Sex × Substance interaction and Age<sup>2</sup>

Tabla 4: Modelos de regresión logística con ajuste (Firth): Predictores de hospitalización psiquiátrica posterior al tratamiento rehabilitador de drogas

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Intercept	-2.756*** [-2.870, -2.641]	-2.830*** [-2.946, -2.715]	-2.971*** [-3.091, -2.851]	-3.592*** [-3.720, -3.463]	-3.536*** [-3.928, -3.144]	-3.784*** [-4.208, -3.361]	-5.374*** [-5.901, -4.847]
Substance: Other substances	0.462** [0.169, 0.755]	0.430** [0.137, 0.724]	0.292 [0.020, 0.604]	0.362* [0.040, 0.684]	0.344* [0.022, 0.666]	0.279 [-0.043, 0.600]	0.052 [-0.504, 0.608]
Substance: Cocaine	-0.136*** [-0.210, -0.062]	-0.166*** [-0.240, -0.092]	-0.086* [-0.162, -0.010]	-0.191*** [-0.269, -0.113]	-0.188*** [-0.266, -0.110]	-0.120** [-0.201, -0.038]	-0.018 [-0.154, 0.117]
Substance: Depressants	0.571*** [0.425, 0.716]	0.545*** [0.400, 0.691]	0.460*** [0.306, 0.613]	0.486*** [0.327, 0.646]	0.465*** [0.304, 0.625]	0.436*** [0.274, 0.597]	0.465*** [0.245, 0.685]
Substance: Marijuana	-0.037 [-0.143, 0.069]	-0.024 [-0.130, 0.082]	-0.022 [-0.132, 0.087]	-0.012 [-0.124, 0.100]	-0.018 [-0.130, 0.094]	0.014 [-0.100, 0.127]	-0.090 [-0.298, 0.119]
Substance: Cocaine paste	0.150** [0.091, 0.208]	0.007 [-0.053, 0.068]	0.101** [0.038, 0.163]	-0.149*** [-0.214, -0.084]	-0.138*** [-0.203, -0.073]	-0.105** [-0.175, -0.036]	0.042 [-0.068, 0.153]
Age	0.013*** [0.011, 0.016]	0.014*** [0.011, 0.016]	0.011*** [0.009, 0.013]	0.013*** [0.011, 0.016]	0.015*** [0.013, 0.018]	0.015*** [0.012, 0.018]	0.089*** [0.074, 0.104]
Sex: Male	-0.470*** [-0.521, -0.420]	-0.447*** [-0.498, -0.397]	-0.382*** [-0.434, -0.329]	-0.245*** [-0.300, -0.191]	-0.251*** [-0.305, -0.196]	-0.215*** [-0.275, -0.155]	-0.105* [-0.203, -0.007]
Plan: Residential	0.763*** [0.699, 0.827]	0.527*** [0.459, 0.594]	0.421*** [0.351, 0.490]	0.402*** [0.332, 0.472]	0.284*** [0.210, 0.358]	0.279*** [0.205, 0.353]	
Nº previous mental hosp.		0.520*** [0.498, 0.542]	0.524*** [0.501, 0.546]	0.517*** [0.494, 0.539]	0.482*** [0.460, 0.504]	0.477*** [0.454, 0.499]	
Retreatments: 1			1.063*** [1.002, 1.124]	1.063*** [1.002, 1.124]	1.059*** [0.997, 1.120]	1.047*** [0.986, 1.109]	
Retreatments: 2			1.619*** [1.538, 1.701]	1.619*** [1.537, 1.700]	1.614*** [1.532, 1.697]	1.596*** [1.513, 1.678]	
Retreatments: 3+			2.290*** [2.206, 2.374]	2.287*** [2.203, 2.371]	2.300*** [2.215, 2.385]	2.275*** [2.189, 2.360]	
Education: Primary			-0.089	-0.103	-0.078		
Education: Secondary			-0.449, 0.270	-0.464, 0.259	-0.440, 0.284		
Education: University			-0.121	-0.122	-0.108		
Education: Higher technical			-0.479, 0.237	-0.482, 0.238	-0.469, 0.252		
Marital status: Married/Partnered			0.050	0.038	0.038		
Marital status: Other/No answer			-0.316, 0.416	-0.330, 0.406	-0.331, 0.406		
Marital status: Single			0.003	0.010	0.005		
Marital status: Widowed			-0.361, 0.367	-0.356, 0.376	-0.361, 0.372		
Occupation: Studying			-0.153*** [-0.239, -0.067]	-0.123** [-0.210, -0.036]	-0.126** [-0.213, -0.039]		
Occupation: Other			0.109	0.023	0.054		
Occupation: Retired/Pensioned			-0.414, 0.632	-0.509, 0.555	-0.478, 0.586		
Occupation: Household tasks			0.015	0.003	0.015		
Occupation: Working			-0.070, 0.100	-0.083, 0.088	-0.070, 0.101		
Age <sup>2</sup>			-0.182	-0.240*	-0.109		
Interaction: Male × Other substances			-0.411, 0.046	-0.471, -0.010	-0.341, 0.123		
Interaction: Male × Cocaine			-0.246*	-0.124			
Interaction: Male × Depressants			-0.483, -0.010	-0.362, 0.114			
Interaction: Male × Marijuana			-0.138**	-0.118*			
Interaction: Male × Cocaine paste			-0.230, -0.047	-0.210, -0.027			
<b>Model statistics</b>			0.130	0.360***			
N observations	97,618	97,618	97,618	97,618	97,618	97,618	97,618
Log Likelihood	-26023.76	-25776.08	-24435.22	-22629.97	-22602.35	-22290.09	-22223.16
Deviance	52047.52	51552.16	48870.44	45259.94	45204.70	44580.18	44446.33
AIC	52063.52	51570.16	48890.44	45285.94	45246.70	44662.18	44540.33
BIC	52139.43	51655.56	48985.33	45409.30	45445.97	45051.23	44986.30
Pseudo R <sup>2</sup> (McFadden)	0.012	0.021	0.072	0.141	0.142	0.154	0.156
Pseudo R <sup>2</sup> (Nagelkerke)	0.016	0.028	0.092	0.176	0.177	0.191	0.194

Note:

\*\*\* p < 0.001, \*\* p < 0.01, \* p < 0.05

95% Confidence intervals in brackets

Reference categories: Alcohol (substance), Female (sex), Outpatient (plan), etc.

Models 6-7 additionally adjusted for Region

Model 7 includes Sex × Substance interaction and Age<sup>2</sup>

Tabla 5: Model Comparison - AIC and AUC across Model Types

Model	Logit AIC	Probit AIC	Firth AIC	Logit AUC	Probit AUC	Firth AUC
<b>Model 1</b>	52063.51	52059.57	52063.52	0.5835	0.5838	0.5835
<b>Model 2</b>	51570.15	51570.86	51570.16	0.6074	0.6075	0.6074
<b>Model 3</b>	48890.41	48899.12	48890.44	0.6923	0.6920	0.6923
<b>Model 4</b>	45285.94	45332.71	45285.94	0.7672	0.7672	0.7672
<b>Model 5</b>	45246.68	45292.52	45246.70	0.7680	0.7681	0.7680
<b>Model 6</b>	44662.12	44733.00	44662.18	0.7768	0.7770	0.7768
<b>Model 7</b>	44540.24	44623.88	44540.33	0.7787	0.7789	0.7786

Tabla 6: Comparación de modelos de regresión logística anidados

Model Comparison	Df	LR Chi <sup>2</sup>	P-value
<b>Null vs Model 1</b>	7	634.227	< 0.001***
<b>Model 1 vs Model 2</b>	1	495.357	< 0.001***
<b>Model 2 vs Model 3</b>	1	2681.749	< 0.001***
<b>Model 3 vs Model 4</b>	3	3610.469	< 0.001***
<b>Model 4 vs Model 5</b>	8	55.261	< 0.001***
<b>Model 5 vs Model 6</b>	20	624.556	< 0.001***
<b>Model 6 vs Model 7</b>	6	133.880	< 0.001***

Note:

\*\*\* p < 0.001, \*\* p < 0.01, \* p < 0.05

Likelihood ratio tests comparing nested models

Tabla 7: Comparaciones de modelos de regresión Probit anidados

Model Comparison	Df	LR Chi <sup>2</sup>	P-value
<b>Null vs Model 1</b>	7	638.166	< 0.001***
<b>Model 1 vs Model 2</b>	1	490.708	< 0.001***
<b>Model 2 vs Model 3</b>	1	2673.743	< 0.001***
<b>Model 3 vs Model 4</b>	3	3572.411	< 0.001***
<b>Model 4 vs Model 5</b>	8	56.195	< 0.001***
<b>Model 5 vs Model 6</b>	20	599.518	< 0.001***
<b>Model 6 vs Model 7</b>	6	121.118	< 0.001***

Note:

\*\*\* p < 0.001, \*\* p < 0.01, \* p < 0.05

Likelihood ratio tests comparing nested models

Tabla 8: Comparaciones de modelos de regresión logística con ajuste de Firth anidados

Model Comparison	Df	LR Chi <sup>2</sup>	P-value
<b>Null vs Model 1</b>	7	634.222	< 0.001***
<b>Model 1 vs Model 2</b>	1	495.356	< 0.001***
<b>Model 2 vs Model 3</b>	1	2681.722	< 0.001***
<b>Model 3 vs Model 4</b>	3	3610.494	< 0.001***
<b>Model 4 vs Model 5</b>	8	55.241	< 0.001***
<b>Model 5 vs Model 6</b>	20	624.519	< 0.001***
<b>Model 6 vs Model 7</b>	6	133.856	< 0.001***

Note:

\*\*\* p < 0.001, \*\* p < 0.01, \* p < 0.05

Likelihood ratio tests comparing nested models

Tabla 9: Performance Metrics for All Models

Model	AUC	Sens	Spec	Acc	Prec	NPV
<b>Standard Logit Models</b>						
Logit Model 1	0.584	0.402	0.714	0.690	0.104	0.935
Logit Model 2	0.607	0.435	0.728	0.706	0.117	0.940
Logit Model 3	0.692	0.561	0.740	0.726	0.151	0.953
Logit Model 4	0.767	0.714	0.700	0.701	0.165	0.967
Logit Model 5	0.768	0.672	0.742	0.737	0.177	0.965
Logit Model 6	0.777	0.655	0.771	0.762	0.191	0.964
Logit Model 7	0.779	0.636	0.793	0.781	0.202	0.963
<b>Probit Models</b>						
Probit Model 1	0.584	0.402	0.715	0.691	0.104	0.935
Probit Model 2	0.607	0.430	0.732	0.709	0.117	0.940
Probit Model 3	0.692	0.560	0.740	0.726	0.151	0.953
Probit Model 4	0.767	0.683	0.732	0.728	0.174	0.965
Probit Model 5	0.768	0.679	0.737	0.733	0.176	0.965
Probit Model 6	0.777	0.656	0.771	0.762	0.191	0.964
Probit Model 7	0.779	0.643	0.786	0.775	0.199	0.964
<b>Firth Logit Models</b>						
Firth Model 1	0.584	0.402	0.714	0.690	0.104	0.935
Firth Model 2	0.607	0.435	0.728	0.705	0.117	0.940
Firth Model 3	0.692	0.561	0.740	0.726	0.151	0.953
Firth Model 4	0.767	0.714	0.700	0.701	0.165	0.967
Firth Model 5	0.768	0.672	0.742	0.737	0.177	0.965
Firth Model 6	0.777	0.656	0.770	0.761	0.191	0.964
Firth Model 7	0.779	0.636	0.792	0.780	0.202	0.963

Notes:

AUC = Area Under ROC Curve; Sens = Sensitivity; Spec = Specificity; Acc = Accuracy

Prec = Precision (PPV); NPV = Negative Predictive Value

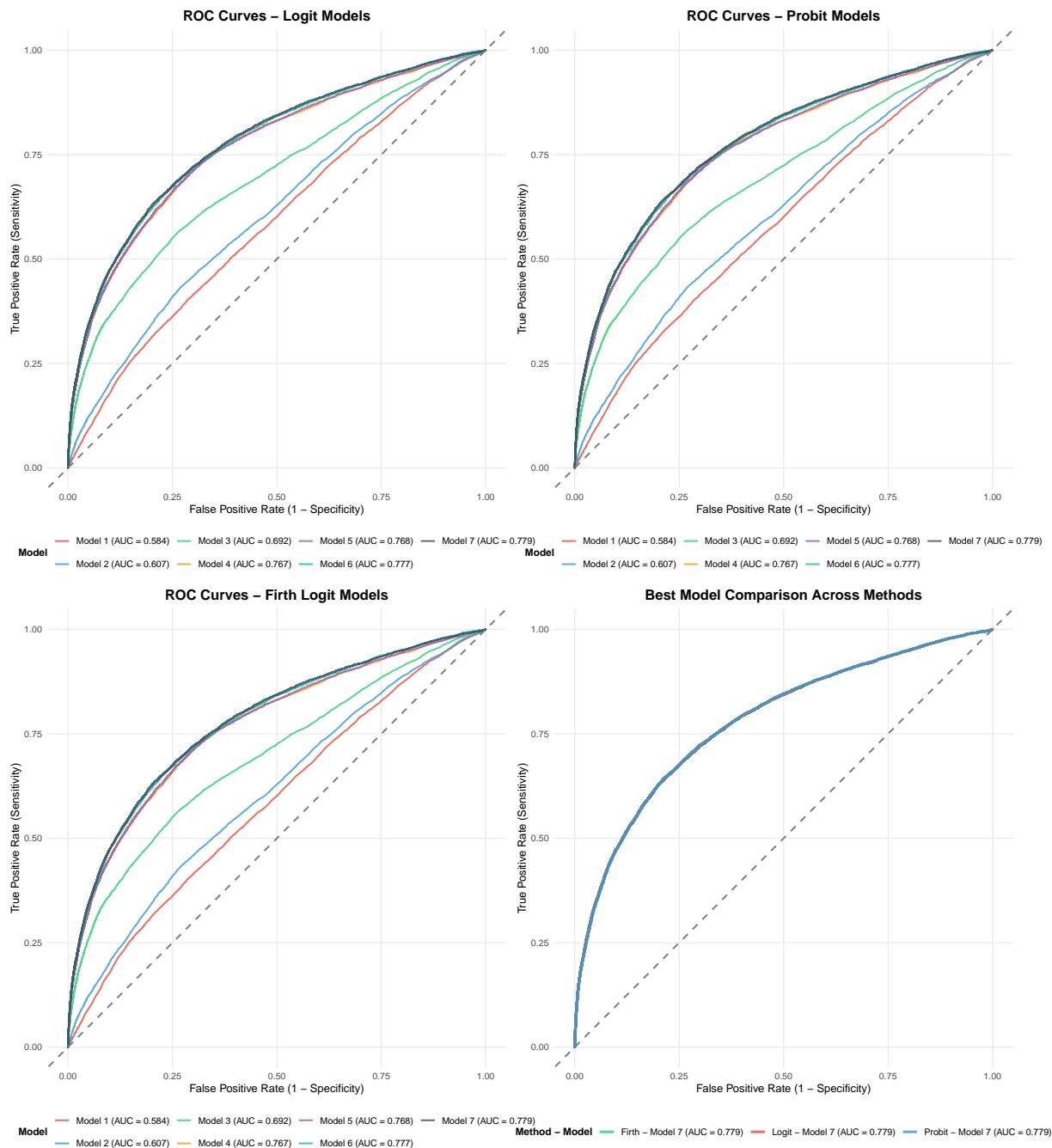


Figura 1: Curvas ROC (Característica Operativa del Receptor) que comparan el rendimiento predictivo de siete modelos utilizando tres métodos de estimación diferentes: regresión logística estándar (logit), regresión probit y regresión logística penalizada de Firth. El cuarto panel muestra una comparación del modelo con mejor rendimiento de cada método. Cada curva muestra el equilibrio entre sensibilidad y especificidad con los valores correspondientes del Área Bajo la Curva (AUC) para la evaluación del modelo.

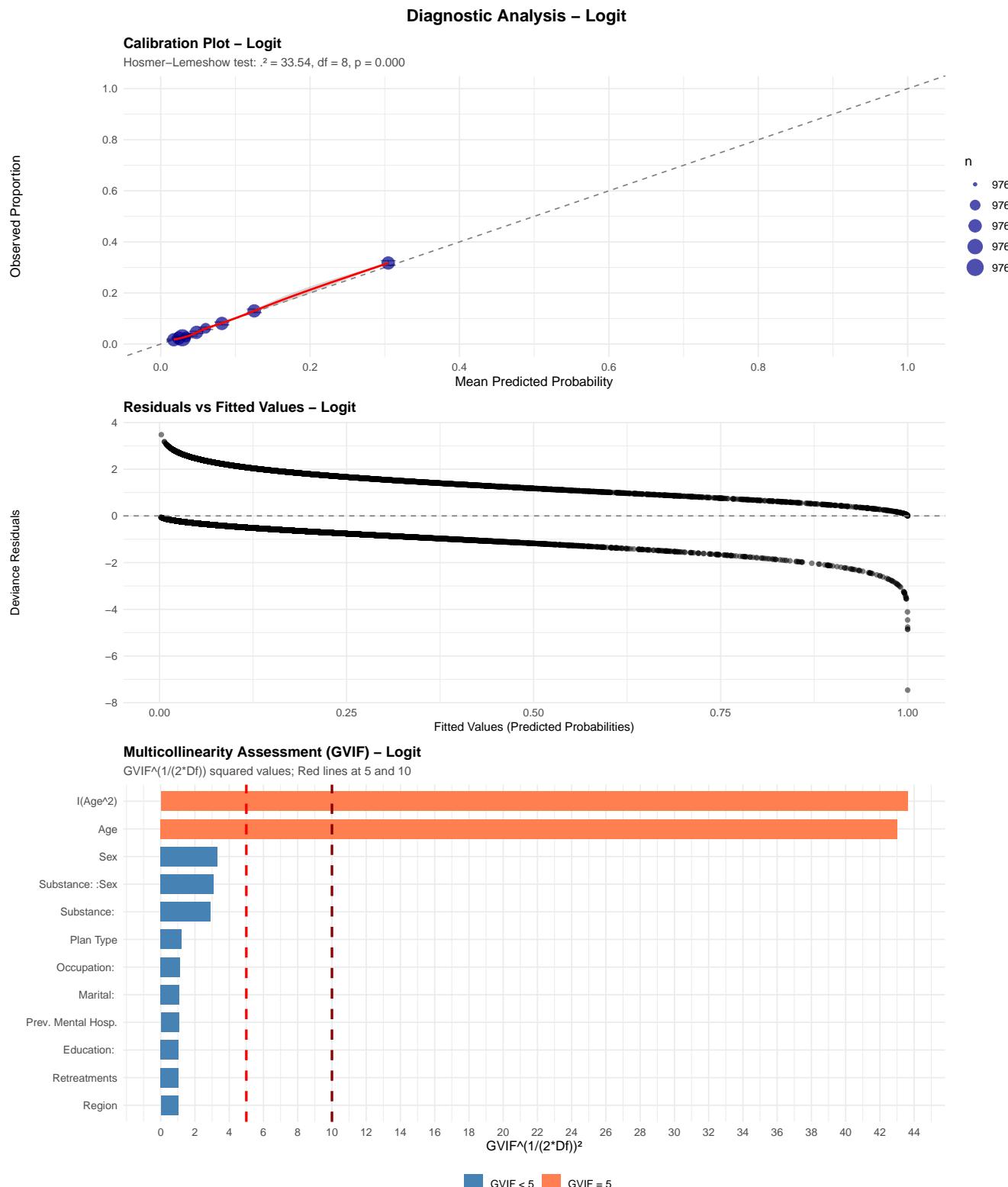


Figura 2: Diagnóstico para las variantes del Modelo 7 que muestra: (1) Gráficos de calibración con la prueba de Hosmer-Lemeshow, (2) Residuos de desviación versus valores ajustados y (3) Evaluación de multicolinealidad utilizando el Factor de Inflación de Varianza Generalizado (GVIF)

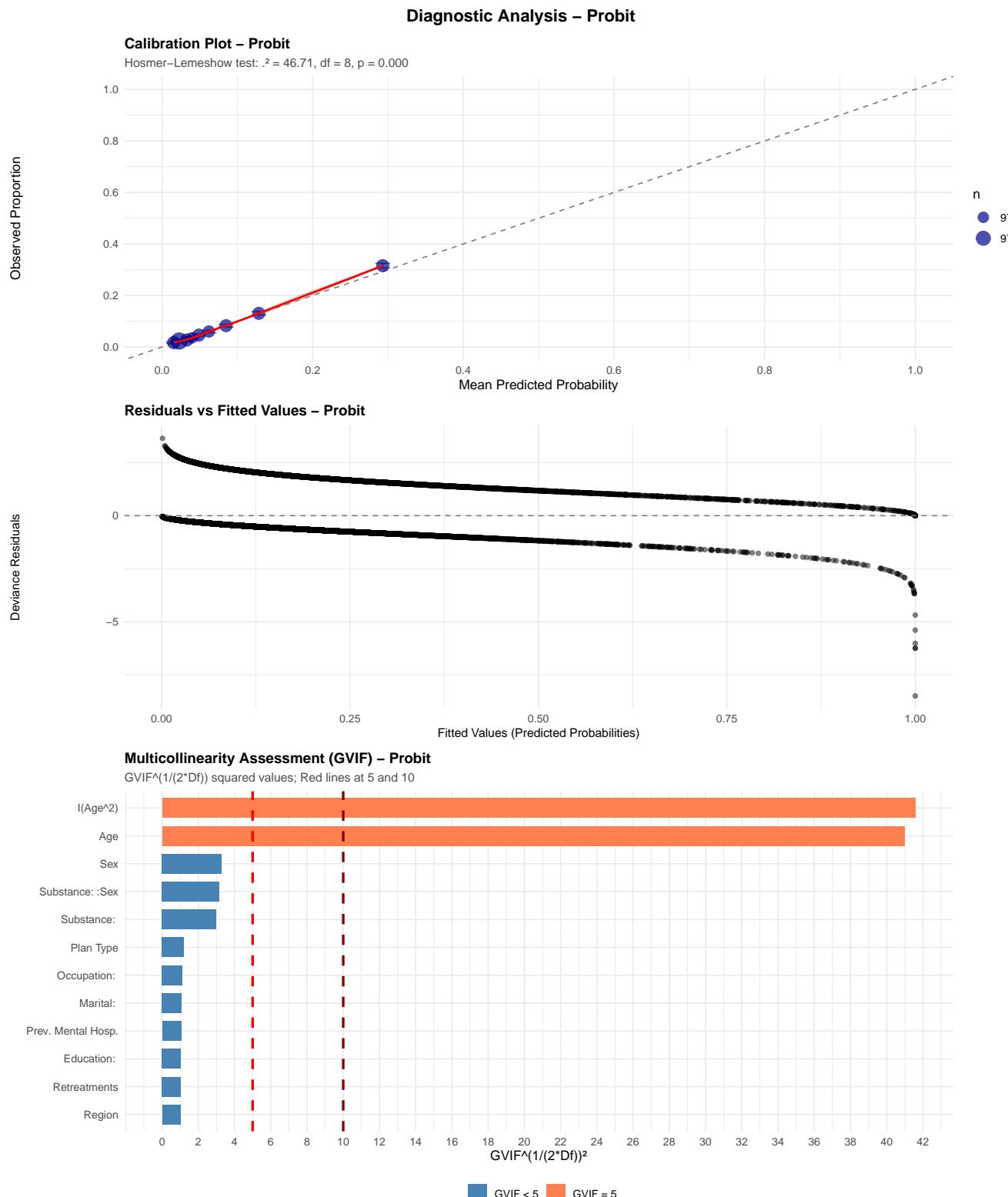


Figura 3: Diagnóstico para las variantes del Modelo 7 que muestra: (1) Gráficos de calibración con la prueba de Hosmer-Lemeshow, (2) Residuos de desviación versus valores ajustados y (3) Evaluación de multicolinealidad utilizando el Factor de Inflación de Varianza Generalizado (GVIF)

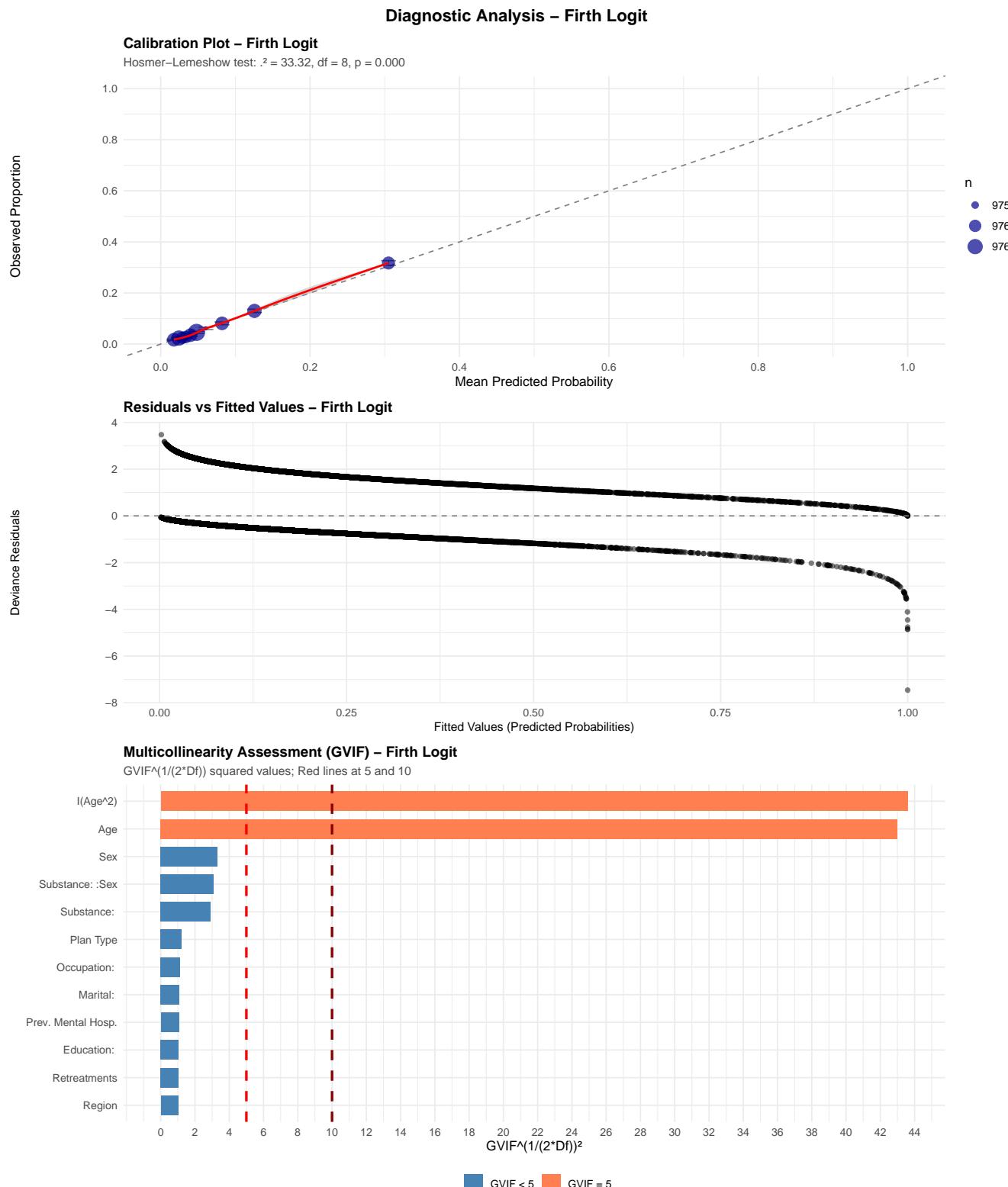


Figura 4: Diagnóstico para las variantes del Modelo 7 que muestra: (1) Gráficos de calibración con la prueba de Hosmer-Lemeshow, (2) Residuos de desviación versus valores ajustados y (3) Evaluación de multicolinealidad utilizando el Factor de Inflación de Varianza Generalizado (GVIF)

## 4 Discusión y Conclusiones

### 4.1 Marco teórico planteado y resultados obtenidos

El modelo teórico asumió un riesgo acumulativo en que la hospitalización psiquiátrica depende de la severidad de la adicción, la vulnerabilidad psiquiátrica previa, las características del tratamiento y factores sociodemográficos. Nuestros resultados respaldan este marco y las hipótesis H1–H4 planteadas. El historial psiquiátrico (H1) surgió como predictor dominante: cada hospitalización mental previa aumentó significativamente la probabilidad de otra hospitalización subsecuente. Asimismo, los retratamientos (H2) mostraron un marcado efecto dosis-respuesta: pacientes con  $\geq 3$  reingresos a tratamiento presentaron casi 10 veces mayor odds de hospitalización que aquellos sin recaídas. El tipo de tratamiento (H3) también fue significativo: la atención residencial se asoció con mayor riesgo relativo a la ambulatoria, aunque parte de esta brecha se redujo al controlar la mayor severidad de los casos derivados a residencias (sesgo de selección). En cuanto al tipo de sustancia (H4), el consumo de depresores del SNC tuvo el efecto más marcado, aumentando 60% la odds de hospitalización comparado con alcohol, en línea con su mayor potencial de complicaciones psiquiátricas agudas. En conjunto, estos hallazgos sustentan el modelo de riesgo acumulativo: la combinación de vulnerabilidad psiquiátrica previa y alta severidad adictiva incrementa sustancialmente el riesgo de descompensación mental, concordando con evidencia internacional en patología dual que vincula la comorbilidad adictiva-psiquiátrica a mayores tasas de rehospitalización.

### 4.2 Limitaciones del estudio

Este estudio observacional presenta varias limitaciones. Primero, la naturaleza no experimental impide inferir causalidad sólida, dado que pueden existir endogeneidades por variables omitidas; por ejemplo, el efecto aparente del tratamiento residencial podría reflejar en parte la mayor gravedad clínica de sus pacientes. Segundo, el uso de registros administrativos conlleva posibles sesgos de medición: las recaídas podrían estar subreportadas (al considerar solo reingresos formales) y solo se registran hospitalizaciones en el sistema de salud, omitiendo episodios no atendidos o diagnósticos no consignados. Tercero, el modelo predictivo mostró limitaciones de ajuste y poder explicativo: la prueba de Hosmer-Lemeshow resultó significativa (falta de calibración) y la capacidad discriminativa fue moderada (AUC ~0.78), lo que sugiere influencia de factores no observados. Finalmente, para afianzar las relaciones causales, se recomienda emplear diseños cuasiexperimentales y datos longitudinales de seguimiento. Dichos enfoques permitirían controlar mejor la endogeneidad, examinar la secuencia temporal entre recaídas y hospitalizaciones, y aportar evidencia más sólida para guiar intervenciones que reduzcan el riesgo de rehospitalización.

## 5 Repositorio GitHub

Este código se replica automáticamente con datos **simulados** debido al tamaño de los datos originales. El código completo de este análisis, incluyendo los modelos estadísticos, las visualizaciones y los diagnósticos, está disponible en el siguiente [repositorio GitHub](#). De igual manera se puede acceder con el siguiente código QR.



El informe .pdf se encuentra en [esta dirección](#). De igual manera se puede acceder con el siguiente código QR.



## Referencias

1. Volkow, N. D., & Blanco, C. (2023). Substance use disorders: a comprehensive update of classification, epidemiology, neurobiology, clinical aspects, treatment and prevention. *World Psychiatry*, 22(2), 203-229. <https://doi.org/10.1002/wps.21073>
2. Connery, H. S., McHugh, R. K., Reilly, M., Shin, S., & Greenfield, S. F. (2020). Substance Use Disorders in Global Mental Health Delivery: Epidemiology, Treatment Gap, and Implementation of Evidence-Based Treatments. *Harvard Review of Psychiatry*, 28(5), 316-327. <https://doi.org/10.1097/HRP.0000000000000271>
3. Gómez-Sánchez-Lafuente, C., Guzman-Parra, J., Suarez-Perez, J., Bordallo-Aragon, A., Rodriguez-de-Fonseca, F., & Mayoral-Cleries, F. (2022). Trends in Psychiatric Hospitalizations of Patients With Dual Diagnosis in Spain. *Journal of Dual Diagnosis*, 18(2), 92-100. <https://doi.org/10.1080/15504263.2022.2053770>
4. Saxena, S., Thornicroft, G., Knapp, J., & Whiteford, M. (2007). Resources for mental health: scarcity, inequity, and inefficiency. *World Psychiatry*, 6(1), 1-10. <https://doi.org/10.1002/wps.21073>
5. Gómez-Sánchez-Lafuente, C., Guzman-Parra, J., Suarez-Perez, J., Mayoral-Cleries, F., & Rodriguez-de-Fonseca, F. (2016). Dual Diagnosis in Spain: Prevalence, Sociodemographic Profile, and Psychiatric Comorbidity in a Sample of Patients Admitted to Psychiatric Inpatient Units. *Journal of Dual Diagnosis*, 12(3-4), 249-258. <https://doi.org/10.1080/15504263.2016.1220207>
6. Rojas, G., Gaete, M., Guajardo, M., Martínez, M., Martínez, M., Fritsch, R., & Araya, R. (2002). Prevalencia de trastornos psiquiátricos en pacientes hospitalizados en un hospital general. *Revista Médica de Chile*, 130(6), 689-696. <https://doi.org/10.4067/S0034-98872002000600008>
7. Cox, D. R. (1958). The regression analysis of binary sequences (with discussion). *Journal of the Royal Statistical Society: Series B (Methodological)*, 20(2), 215-242. <https://doi.org/10.1111/j.2517-6161.1958.tb00292.x>
8. Bliss, C. I. (1934). The method of probits. *Science*, 79(2037), 38-39. <https://doi.org/10.1126/science.79.2037.38>
9. Firth, D. (1993). Bias reduction of maximum likelihood estimates. *Biometrika*, 80(1), 27-38. <https://doi.org/10.1093/biomet/80.1.27>
10. Neyman, J., & Pearson, E. S. (1933). IX. On the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society of London. Series A*, 231(694-706), 289-337. <https://doi.org/10.1098/rsta.1933.0009>
11. Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. En B. N. Petrov & F. Csáki (Eds.), *Proceedings of the 2nd International Symposium on Information Theory* (pp. 267-281). Akadémiai Kiadó.
12. Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143(1), 29-36. <https://doi.org/10.1148/radiology.143.1.7063747>
13. DeLong, E. R., DeLong, D. M., & Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*, 44(3), 837-845.
14. Hosmer, D. W., & Lemeshow, S. (1980). Goodness of fit tests for the multiple logistic regression model. *Communications in Statistics - Theory and Methods*, 9(10), 1043-1069. <https://doi.org/10.1080/03610928008827941>
15. Fox, J., & Monette, G. (1992). Generalized collinearity diagnostics. *Journal of the American Statistical Association*, 87(417), 178-183. <https://doi.org/10.1080/01621459.1992.10475190>