

*Alonso, María Susana*

## **Ausentismo de pacientes adultos jóvenes: desarrollo de modelos predictivos**

Maestría en Investigación Clínica

*Tesis 2021*

*Cita sugerida:* Alonso MS. Ausentismo de pacientes adultos jóvenes: desarrollo de modelos predictivos [tesis de maestría] [Internet]. [Buenos Aires]: Instituto Universitario Hospital Italiano de Buenos Aires; 2021 [citado AAAA MM DD]. 53 p. Disponible en: <http://trovare.hospitalitaliano.org.ar/descargas/tesisyt/20210521124924/tesis-alonso-susana.pdf>

Este documento integra la colección Tesis y trabajos finales de Trovare Repositorio Institucional del Instituto Universitario Hospital Italiano de Buenos Aires y del Hospital Italiano de Buenos Aires. Su utilización debe ser acompañada por la cita bibliográfica con reconocimiento de la fuente.

Para más información visite el sitio <http://trovare.hospitalitaliano.org.ar/>



Tesis Maestría en Investigación Clínica  
Instituto Universitario Hospital Italiano  
Título: Ausentismo de pacientes adultos jóvenes  
Desarrollo de modelos predictivos

Maestrando: María Susana Alonso<sup>1, 2</sup> Cohorte 2013

[mariasusana.alonso@hospitalitaliano.org.ar](mailto:mariasusana.alonso@hospitalitaliano.org.ar)

Director: Marcelo Risk<sup>2, 3</sup>

[marcelo.risk@hospitalitaliano.org.ar](mailto:marcelo.risk@hospitalitaliano.org.ar)

Director y tutor: Diego Giunta<sup>2,4,5</sup>

[diego.giunta@hospitalitaliano.org.ar](mailto:diego.giunta@hospitalitaliano.org.ar)

1. Instituto de Medicina Traslacional e Ingeniería Biomédica (IMTIB) CONICET
2. Instituto Universitario Hospital Italiano- Hospital Italiano Buenos Aires
3. Director Instituto de Medicina Traslacional e Ingeniería Biomédica (IMTIB) CONICET
4. Área de Investigación en Medicina Interna. Servicio de Clínica Médica. Hospital Italiano de Buenos Aires
5. Área de investigación no Patrocinada. Departamento de Investigación. Hospital Italiano de Buenos Aires

Agradecimientos:

- A los dos directores: el Dr. Marcelo Risk, y el Dr. Diego Giunta, también tutor de la tesis.
- A mi madre y mis seis hijos por todo su apoyo y paciencia.
- A Estanislao por su colaboración.

## Índice de contenidos

<b>Resumen</b>	<b>4</b>
<b>Introducción</b>	<b>5</b>
<b>Materiales y Métodos</b>	<b>8</b>
Diseño	8
Ámbito	8
Variables	9
Consideraciones estadísticas	12
Desarrollo de Modelos predictivos	13
Consideraciones éticas	15
<b>Resultados</b>	<b>16</b>
Análisis descriptivo y tasa de ausentismo	16
Factores asociados a ausentismo	17
Evaluación de modelos de Machine Learning para predicción de ausentismo	17
<b>Discusión</b>	<b>18</b>
Conclusiones	24
<b>Tablas</b>	<b>26</b>
<b>Figuras</b>	<b>37</b>
<b>Referencias bibliográficas</b>	<b>41</b>
<b>Anexo 1. Operacionalización de variables Ausentismo Modelos predictivos</b>	<b>45</b>

## Resumen

**Introducción.** El ausentismo de los pacientes a los turnos médicos programados representa un problema muy importante para los centros de salud, las tasas de inasistencia pueden ir del 10 al 40% aproximadamente, la pérdida económica es significativa y como consecuencia se deteriora la relación médico-paciente. Es primordial conocer los factores asociados al ausentismo para tomar decisiones que puedan solucionar este problema. En este trabajo, desarrollado en un sistema de salud privado, se diseñó una cohorte retrospectiva, se caracterizaron los factores asociados al ausentismo, se estimaron las tasas de ausentismo y se desarrollaron modelos predictivos de Aprendizaje de máquinas.

**Objetivos.** Para las consultas ambulatorias de adultos (edades  $\geq 18$  y  $\leq 45$  años) de Clínica médica/Atención primaria de la salud fueron: 1. Estimar la tasa de ausentismo a consultas médicas clínicas programadas. 2. identificar factores asociados a mayor probabilidad de ausentismo en cada turno. 3. Construir modelos predictivos de ausentismo usando técnicas de aprendizaje de máquinas.

**Materiales y métodos.** Se generó una cohorte retrospectiva de todos los pacientes adultos incluidos (177.105) de 18 a 45 años de edad, que solicitaron al menos un turno programado. Se utilizaron todos los turnos solicitados para estimar las tasas de ausentismo, evaluar factores asociados a ausentismo, generar modelos predictivos de aprendizaje de máquinas para ausentismo. El período del estudio fue desde el 1° de enero de 2015 al 31 de diciembre de 2016 inclusive. Fuente de datos: historia clínica hospitalaria informatizada.

**Resultados.** Se estimó una tasa de ausentismo del 35,6%. El análisis por sexo, para 177.105 turnos, arroja un resultado de 21.461 hombres ausentes, siendo la tasa de ausentismo para este grupo de 32,2 %. Los factores más fuertemente asociados a ausentismo fueron pertenecer a una cobertura médica privada OR 4,912 (IC95% 3,982- 6,059), pertenecer a una cobertura médica privada de alta falta OR 9,339 (IC95% 7,325- 11,906), insuficiencia hepática OR 9,371 (IC95% 1,761- 49,84), la internación actual OR 3,722 (IC95% 2,541- 5,452). La mejor área bajo la curva ROC obtenida fue de 0.694 para el modelo de potenciación extrema del gradiente, las tres variables predictivas de mayor importancia fueron la proporción de faltas de los turnos solicitados en el año previo a la solicitud del turno, la cantidad de problemas activos en la Historia clínica electrónica y el tiempo transcurrido desde la solicitud del turno.

**Conclusiones.** Las tasas de ausentismo detectadas son altas, tanto la tasa global como la de hombres y mujeres. Varios factores se asocian fuertemente al ausentismo, como pertenecer a una cobertura médica privada y la internación al momento del turno. De los modelos predictivos de aprendizaje de máquinas desarrollados, los modelos de potenciación extrema del gradiente y potenciación del gradiente obtuvieron las mejores métricas de calibración.

**Palabras clave:** ausentismo a consultas médicas programadas, atención ambulatoria, tasas de ausentismo, factores asociados al ausentismo, aprendizaje de máquinas, modelos predictivos de ausentismo.

## Introducción

El problema del ausentismo de los pacientes a las citas médicas programadas representa una pérdida económica significativa para los sistemas de salud, aproximadamente las tasas de inasistencia pueden ir desde el 10 al 40%, e inclusive pueden ser más altas [1]. Como consecuencia del impacto negativo generado, los recursos médicos son subutilizados, aumentan los costos de atención médica, disminuye la eficiencia clínica y la productividad. No solamente es un problema económico o de recursos, sino que también se ve afectada la relación médico-paciente tanto en clínica general como en las distintas especialidades médicas [1–3].

El ausentismo o la inasistencia de los pacientes que faltan, sin previo aviso, a los turnos o consultas médicas ambulatorias programadas, se define como la no asistencia por los pacientes a esa consulta sin dar aviso previo, o cuando cancelan los turnos con un tiempo tan reducido que no se puede ubicar a otro paciente privándolo de una consulta [2,4,5]. A pesar de que el problema se trata de solucionar desde tiempo atrás, sigue siendo difícil predecir las ausencias, y los casos en los que el paciente cancele el turno tan cerca del horario establecido que es imposible coordinar las llegadas de otros pacientes cuando ya están programadas [2]. De ahí la importancia de predecir con precisión la probabilidad de que el paciente se ausente.

Hay poco conocimiento en los sistemas de salud sobre los factores asociados a las ausencias de las consultas ambulatorias programadas en cada servicio. Estas son razones por las que se vienen realizando diversos trabajos de investigación para poder caracterizarlos. Por ejemplo, el trabajo de Huang & Hanauer se realizó con la recolección de diez años de datos retrospectivos de un sistema de programación y un sistema de registros de salud electrónicos en una única clínica de pediatría general para desarrollar un modelo predictivo, y Kheirkhah y col estudiaron las tasas de ausentismo en atención primaria y distintas subespecialidades [2,5].

Rebolledo y col encontraron que en Inglaterra el porcentaje de ausentismo llegó a ser de un 11%, en Estados Unidos la tasa de ausentismo del 5 al 34% y en Australia entre el 13 y 23%. También, según cifras oficiales, en el sector de salud pública de Chile la inasistencia de pacientes a consultas médicas de especialistas fue aumentando, en el año 2005 era de alrededor del 13% y en el 2010 cerca del 16%, esto equivale a más de 6,5 millones de consultas de especialistas perdidas durante este período. En ese mismo año, 2010, la tasa llegó al 16,1 % de todas las consultas médicas programadas, considerando solamente el arancel cobrado por el médico, la pérdida económica llegaba a 28 millones de dólares norteamericanos anuales [6]. El alcance de este problema es global, en el continente africano la tasa de ausentismo fue de 43.0%, en Sud América de 27.8%, en Asia 25.1%, América del Norte 23.5%, y en el resto de Europa 19.3% [4]. Un estudio en la Agencia Sanitaria Costa del Sol revela que para una tasa de inasistencia del 13,8%, el impacto económico total fue superior a tres millones de euros, correspondiendo a un 14% del total global de las consultas de la agencia (400.761 turnos) [7]. Otro trabajo de 2010 muestra que de un total de 619 pacientes con cáncer colorrectal el 21% no asistió a uno o más turnos, a pesar de que el seguimiento de rutina tiene un impacto significativo en la supervivencia de estos pacientes [8].

Encontramos diversos estudios que demuestran que la tasa de ausentismo en el grupo de pacientes jóvenes es alta. En 2009 fue publicado un trabajo que reportaba que el ausentismo a citas hospitalarias en todo el Reino Unido cuesta al Servicio Nacional de Salud un estimado en 600 millones de libras esterlinas por año, siendo los hombres jóvenes los más faltadores [9] también en otro trabajo de 2016, entre los predictores de ausentismo fue ser joven de edad [10].

En un estudio de cohorte retrospectivo en nuestro hospital universitario, se estimó una tasa de ausentismo del 27,84% para citas ambulatorias programadas de pacientes mayores de 18 años [11]. Por otro lado, se observó en otro de nuestros estudios, que los pacientes que pierden sus citas programadas son en general jóvenes adultos de 17 a 40 años de edad [12]. Vidal y col encuentran que hay más evidencia de que los pacientes jóvenes pertenezcan al grupo de mayor riesgo de ausencia a las citas programadas en las diferentes áreas de salud, como se ve en el estudio de [13]. Una revisión sistemática sobre las barreras y facilitadores asociados a la asistencia de adultos jóvenes (15-30 años de edad) con enfermedades crónicas como la diabetes mellitus tipo I han demostrado una baja asistencia en este grupo etario [14]. En un estudio de la Agencia Sanitaria Costa del Sol, Málaga, se determinó que el grupo de mayor riesgo de ausentismo en las consultas externas es el de los pacientes jóvenes [7]. En otro estudio realizado en un hospital de cirugía de mano, los pacientes que más se ausentaban eran también los más jóvenes, los pacientes que no asistieron a los turnos y que no avisaron por adelantado para cancelarlos tenían más probabilidades de ser más jóvenes (45 años, DS 16) que los pacientes que asistieron a sus turnos (52 años, DS 17) [3].

Nuestro equipo de trabajo viene generando diferentes aproximaciones para la identificación de los factores asociados al ausentismo y el desarrollo de modelos predictivos para detectar las variables involucradas a este problema [11]. Giunta y col recomiendan que, debido a la existencia de variabilidad en las faltas a las consultas a los turnos programados dentro de la atención médica ambulatoria, se utilicen estimaciones locales para que, al diseñar las intervenciones correspondientes, éstas sean efectivas [11]. Por eso la importancia de identificar los grupos de pacientes que tienen mayor probabilidad de no presentarse a las consultas en clínica médica en nuestro hospital. Es necesario poder predecir cuáles de ellos están en más riesgo de ausentarse para implementar alguna estrategia nueva que pueda disminuir el ausentismo [8]. Briatore y col también recomiendan que, ante la evidencia relacionada con las características de los pacientes y la identificación de turnos con una mayor probabilidad de inasistencia, se deben promover estrategias específicas para reducir la tasa de ausentismo, y también futuras investigaciones sobre este tema [12]. Los modelos predictivos que vienen desarrollándose para predecir cuáles son los factores que se asocian al ausentismo, sirven para orientar el desarrollo de estrategias en la gestión hospitalaria y así lograr reducir la tasa de inasistencia.

Con el progreso reciente en la adquisición de datos digitalizados, el aprendizaje automático o aprendizaje de máquinas (Machine Learning- ML) y la infraestructura informática, las aplicaciones de inteligencia artificial (IA) se están expandiendo a áreas que anteriormente no se pensaba. La IA está cambiando gradualmente la práctica médica [15], permitiendo que se puedan realizar pronósticos más precisos. Para esto se utiliza la estadística por medio del procesamiento de conjuntos de datos masivos (big data). A través de modelos matemáticos en estratos (algoritmos), se pueden detectar patrones no descifrables al corregir los algoritmos (entrenamiento) para aumentar la confianza del modelo predictivo de IA [16]. Unido a la IA está la minería de datos, que es el proceso para descubrir y extraer patrones de una

gran cantidad de datos. Hoy en día hay una necesidad de comprender conjuntos de datos grandes, complejos y enriquecidos con información para todos los campos de la tecnología, los negocios y la ciencia. La industria de la salud en la mayoría de los países evoluciona rápidamente, genera cantidades masivas de datos, como registros médicos electrónicos, informes administrativos y otros. Por medio de la minería de datos, a partir de estos grandes volúmenes de datos, se puede buscar información nueva y valiosa. La minería de datos en el cuidado de la salud se utiliza principalmente para predecir diversas enfermedades, también para ayudar a los médicos en el diagnóstico a la hora de tomar decisiones clínicas. En las ciencias de la salud los modelos predictivos son los más usados. Los modelos predictivos a menudo aplican funciones de aprendizaje supervisado para predecir valores desconocidos o futuros de otras variables de interés [17].

ML podría permitir conocer anticipadamente una probabilidad de ausentismo a un turno programado. A su vez, por medio de la utilización de técnicas de ML se podría conocer anticipadamente el comportamiento de los pacientes. Esta aplicación de la IA, que ha sido ampliamente utilizada por la comunidad científica, permite convertir una gran variedad de bases de datos enormes y heterogéneas en alta calidad de información [18]. A partir de ML, “las fuentes de datos clínicos permiten generar rápidamente modelos de predicción para miles de preguntas clínicas del mismo tipo” [19].

Nelson y col encontraron que el mejor rendimiento, con un área bajo la curva de 0.852 y una precisión promedio de 0.511, fue de los modelos basados en máquinas de aumento de gradiente de alta dimensión [4]. Predecir una enfermedad o un evento en la clínica médica, es cada vez más importante en la medicina actual. Estos modelos de predicción pueden combinar múltiples predictores proporcionando información sobre los efectos relativos de los predictores en el modelo [20]. Es posible encontrar una gran cantidad de información en un amplio campo de factores de predicción débil: los modelos podrían aprovechar esto para lograr un rendimiento mucho mejor de lo que resulta del análisis de cada factor en forma aislada o por combinación lineal. El rendimiento predictivo general puede mejorar sin necesidad de identificar los factores predictivos más fuertes, logrando un modelado que produzca el mejor rendimiento. Se ha visto en algún caso que el mejor rendimiento de un estudio se logró mediante los modelos más complejos, estos reflejan la compleja interacción de factores causales del paciente, ambientales y operativos. Este rendimiento no se logra con los modelos simples, no son adecuados para predecir eventos individuales donde el campo causal es amplio [4]. Dado esto nos propusimos para esta tesis explorar por medio de modelos más complejos no lineales y de alta dimensión los predictores asociados a los pacientes y a los turnos con probabilidad de faltar.

A partir de los resultados de todos los estudios encontrados en la literatura, como también los reportados por Giunta y Briatore (de nuestro equipo de ausentismo de clínica médica), que revelaron que la tasa asociada al ausentismo fue alta, nos preguntamos cómo serían las tasas de ausentismo en el grupo de pacientes jóvenes específicamente de nuestro hospital, para Clínica médica/Atención primaria de la salud, en atención ambulatoria por el servicio de Clínica Médica del Hospital Italiano de Buenos Aires [11,12]. Por lo tanto, siendo esencial buscar soluciones a este problema en el grupo de estos pacientes, trazamos tres objetivos, que fueron, estimar la tasa de ausentismo de la población de pacientes adultos jóvenes de entre 18 y 45 años, detectar los factores asociados al ausentismo, construir y validar modelos predictivos de ausentismo usando técnicas de ML.



## **Materiales y Métodos**

### **Diseño**

Diseñamos una cohorte retrospectiva de todos los pacientes adultos jóvenes que solicitaron al menos un turno programado, los turnos fueron para atención médica ambulatoria en Clínica Médica/Atención primaria de la salud que se solicitaron entre el 1° de enero de 2015 y el 31 de diciembre de 2016 inclusive, del sistema de atención ambulatoria del Hospital Italiano de Buenos Aires (HIBA). Con los turnos solicitados se estimaron las tasas de ausentismo, se evaluaron los factores asociados al ausentismo, se generaron y validaron los modelos predictivos de ausentismo.

### **Ámbito**

El HIBA es un hospital universitario de alta complejidad de la Ciudad Autónoma de Buenos Aires. Funciona como una red integrada de salud. Cuenta con 18 centros de atención ambulatoria, dos hospitales (Hospital Italiano Central y Hospital Agustín A. Roca de San Justo), cuenta aproximadamente con 750 camas entre ambos, y más de 2,5 millones de consultas ambulatorias anuales. En su sistema de atención ambulatoria, se atienden diferentes coberturas de salud, incluyendo el sistema de medicina prepaga del hospital. En este sistema de atención, los especialistas en clínica médica y los médicos de familia cumplen el rol de médicos de cabecera siendo fundamentales en la atención primaria de la salud y el seguimiento ambulatorio de los pacientes. La prepaga del HIBA (Plan de Salud del Hospital Italiano - PS) cuenta con aproximadamente 150.000 afiliados. La atención de los afiliados al PS se realiza en forma exclusiva y completa dentro de este sistema. La prepaga del HIBA (Plan de Salud del Hospital Italiano - PS) cuenta con aproximadamente 150.000 afiliados. La atención de los afiliados al PS es exclusiva dentro de este sistema.

La atención en todos los ámbitos correspondientes al) [21,22]. En este repositorio de datos está contenida toda la información administrativa de los pacientes (como por ejemplo estudios, solicitud de turnos, turnos ambulatorios, internaciones, consumo de fármacos) y toda su información clínica (problemas de salud, diagnósticos clínicos, evoluciones médicas, resultados de estudios entre otros).

Todos los problemas de salud o comorbilidades de los pacientes son codificados automáticamente mediante un servidor de terminología con tesoro local que mapea y codifica la información con el vocabulario controlado SNOMED-CT [23–25]. Debido a que la gestión de turnos se centraliza en un único repositorio de datos y el servidor de terminología permite obtener información del repositorio de datos clínicos, se puede obtener la totalidad de los datos correspondientes a los pacientes y las solicitudes de turnos del sistema de salud completo del HIBA.

### **Población y criterios de selección**

La población blanco del estudio corresponde a los turnos de atención ambulatoria programada por consultorios externos de Clínica médica/Atención primaria de la salud, solicitados por pacientes de 18 a 45 años. Con fecha asignada entre el primero de enero de 2015 y el 31 de diciembre de 2016 inclusive (24 meses en total).

Fueron incluidos los turnos o sobre turnos programados de pacientes adultos jóvenes (edades entre 18 y 45 años) de consultorios externos del sistema de atención ambulatoria del HIBA, correspondiente a Clínica Médica/Atención primaria de la salud (servicio de Clínica Médica, Medicina Familiar). Se excluyeron los turnos de acuerdo a los siguientes criterios: 1. Turnos asignados para procedimientos, evaluaciones prequirúrgicas, y consultas específicas para evaluación previa al ingreso al PS. 2. Los sobretornos espontáneos o sobretornos con nota. 3. Los turnos del médico consigo mismo. 4. Los pacientes con más de 200 turnos en el período de 2 años ya que corresponden a errores y/o pruebas de sistema. 5. Los turnos cancelados o reprogramados. Los sobretornos del primer criterio de exclusión son turnos que no corresponden a turnos para atención médica por consultorios externos. El consultorio de preingreso al PS, no es un consultorio de atención médica, sino una práctica administrativa para los postulantes a afiliarse al Plan de Salud como parte integral del proceso de admisión. Los turnos en los criterios 2 y 3 no son programados y por definición no corresponden a atención programada. No se estudiarán en esta tesis los turnos para prácticas ni procedimientos que tienen una dinámica y características diferentes. La fuente de datos utilizada fue la HCE.

Las unidades de análisis fueron los turnos solicitados tanto en el grupo de turnos de generación, como en el grupo de turnos de validación.

## Variables

### Variable resultado

La variable de resultado fue el ausentismo, evaluado como variable categórica dicotómica. Cuando el paciente se presenta para su atención el día y la hora del turno, se da el presente administrativo en el sistema. Este paso es necesario para la atención médica del paciente y la facturación, por lo cual contiene baja tasa de errores. Cada turno incluido se categorizó como presente o ausente de acuerdo a si el paciente concurrió o no a la consulta según constaba el presente administrativo en el sistema de gestión de turnos.

### Variables explicativas

Las variables explicativas se dividieron en categorías de acuerdo a los conceptos que representan: características basales de los pacientes, características del proceso de solicitud del turno, historial del paciente, características del turno solicitado, clima del día del turno, características clínicas y comorbilidades del paciente. Se presentan en la siguiente tabla.

<b>1. Características basales de los pacientes</b>	Edad Sexo Cobertura médica Cantidad de problemas en la Historia Clínica Electrónica Distancia desde el domicilio al lugar de atención Tiempo desde el domicilio al lugar de atención
<b>2. Características del proceso de solicitud del turno</b>	Método utilizado para la solicitud del turno Sobretorno o turno convencional Tiempo desde la solicitud al turno

<b>3. Historial del paciente</b>	<p>Cantidad de turnos solicitados  Porcentajes de faltas en el último año previo a la solicitud del turno  Internación previa  Internación actual  Fallecimiento previo al turno  Cirugía previa</p>
<b>4. Características del turno solicitado</b>	<p>Más de un turno en el mismo momento  Lugar físico del turno  Turno de primera vez del paciente  Turno de primera vez del médico  Hora del día  Día de la semana  Mes  Horas de alta falta  Días de alta falta  Meses de alta falta  Lugar de alta falta  Porcentaje de faltas del médico a cargo del turno  Presencia de otros turnos y cantidad de ellos solicitados ese mismo día</p>
<b>5. Clima del día del turno</b>	<p>Temperatura máxima, mínima, media  Precipitaciones  Presión barométrica</p>
<b>6. Características clínicas y comorbilidades del paciente</b>	<p>Embarazo  Hipertensión Arterial  Diabetes  Enfermedad Oncológica  Enfermedad Coronaria  Enfermedades Psiquiátricas  Anemia  BMI  Obesidad  Tabaquismo  Alcoholismo  Abuso de sustancias  Score de comorbilidades de Charlson  Linfoma  Leucemia  Accidente cerebro vascular  Demencia  Enfermedad vascular periférica  EPOC  Hemiplejía  Insuficiencia cardíaca congestiva</p>

	Insuficiencia hepática Insuficiencia renal SIDA Insuficiencia hepática
--	---

Características basales de los pacientes que fueron consideradas: la edad, sexo, cobertura médica, ciudad de residencia, distancia del domicilio al lugar de reunión y cantidad de problemas de salud en la HCE. Otras variables demográficas, aunque de interés, no pudieron ser incorporadas ya que no estaban disponibles en las bases de datos secundarias. La cantidad de coberturas de salud posibles (durante el período del estudio fue de 306), se desconoce el perfil de ausentismo de cada una, por esta razón se exploraron los porcentajes de ausentismo para cada cobertura. Se dividieron las coberturas de acuerdo a cuartiles de porcentajes de faltas dentro de cada una durante el período del estudio. El 25% de las coberturas con menor proporción de faltas se incluyó en el primer cuartil y el 25% de las coberturas con mayor porcentaje de faltas en el último cuartil.

Los pacientes del PS y los pacientes privados que pagan la atención, sin ningún tipo de cobertura médica, se estudiaron adicionalmente por separado, ambos grupos tienen un comportamiento diferente. Fueron incluidos como variables diferentes. En este estudio se llamó cobertura privada a quien pagó por su atención, en lugar de que pague una cobertura de salud.

Se incluyeron la distancia y tiempo para llegar desde el domicilio al lugar de atención dentro de esta categoría utilizando los algoritmos de Google Maps para calcular la mínima distancia en kilómetros entre dos direcciones siguiendo el camino transitable más corto en auto y el tiempo estimado para dicho trayecto. En ambos casos se asignaron las coordenadas de latitud y longitud utilizando Google Maps Geocoding API [26]. Se utilizó Google Maps Distance Matrix API para calcular distancia [27].

Las características del proceso de solicitud del turno representan variables de perfiles de pacientes o situaciones diferentes de manera indirecta. Como por ejemplo el método utilizado para la solicitud del turno (telefónico, presencial, a través del portal de salud web POPES, o HCE), o si se trata de un sobretorno programado o turno convencional. A su vez, el tiempo desde la solicitud al turno al momento del turno es un factor habitualmente asociado a ausentismo.

El historial del paciente representa patrones de comportamiento previo al turno asignado. Se incluyeron la cantidad de turnos y el porcentaje de faltas en el año previo a la solicitud del turno. A su vez se incluyeron en esta categoría estar internado al momento del turno, haber

sido internado o intervenido quirúrgicamente en los 30 días previos al turno y el fallecimiento previo. Se consideró el fallecimiento previo a los egresos de una internación que finalizaron como fallecimiento con fecha previa al turno considerado. Si bien el fallecimiento previo parece una obviedad, el sistema gestor de turnos actual no los identifica y son turnos que resultan en faltas con muy alta probabilidad.

Dentro de las características del turno solicitado, se incorporaron el lugar físico del turno, turno de primera vez (primera vez en el HIBA y primera vez con ese médico en particular), hora del día, día de la semana, mes, porcentaje de faltas del médico a cargo del turno, presencia de otros turnos y cantidad solicitados ese mismo día, y si el paciente tenía turnos asignados en el mismo día y hora.

Como el clima es una causa posible de ausentismo, se solicitaron los datos climáticos de Capital Federal al Centro de Información Meteorológica del Servicio Meteorológico Nacional del período evaluado [28] Se incluyeron la temperatura máxima, mínima, media, sensación térmica, precipitaciones y presión barométrica.

Con el fin de simplificar la información y los modelos, se definieron dentro de este grupo, variables indicadoras de alta falta para lugar, médico, hora, día y mes [29]. Se estimó el percentil 75% para las categorías de estas variables: el percentil 75% de porcentaje de faltas por categoría fue 29,6% para las horas del día, de 28,9% para los días de la semana, de 28% para los meses, de 27,9% para los lugares de atención y de 35,7% para los médicos. A partir de la distribución de las frecuencias observadas para cada categoría, se definió a priori un 30% como límite para alta falta. De esta manera, se categorizó como alta falta a todos los lugares de atención con porcentajes de faltas por encima del 30%. Las horas de alta falta fueron 7, 18 y 19 horas; el día de alta falta fue el sábado; y los meses de alta falta fueron junio, julio y diciembre. Otro ejemplo son los médicos, todo medico con porcentaje de faltas mayor a 30%, se clasificó como alta falta.

Se incluyó una categoría con características clínicas y demográficas del paciente de alta frecuencia o relacionadas con las enfermedades crónicas no transmisibles de mayor impacto en la salud: incluyendo hipertensión arterial, diabetes, enfermedad oncológica, enfermedad coronaria. Todas fueron recuperadas a través de los diagnósticos de problemas activos de la HCE codificados según vocabulario controlado SNOMED-CT [23–25,30]. Se incorporaron adicionalmente las enfermedades psiquiátricas debido a que en series anteriores se asocian con el ausentismo. Para representar la carga comórbida del paciente, se calculó el score de comorbilidades de Charlson recuperando los diagnósticos de la misma manera de la HCE.

## **Consideraciones estadísticas**

### **Muestreo y cálculo muestral**

Se llevó a cabo un muestreo aleatorio consecutivo. Por lo que se incluyó la totalidad de los turnos solicitados durante los 24 meses del período del estudio. Los datos para el armado de la base de datos para este estudio se extrajeron de las HCE de los pacientes entre 18 y 45 años, se excluyeron del presente estudio todos los turnos cancelados.

## Análisis estadístico

La unidad de análisis fueron cada uno de los turnos de cada paciente. En el análisis estadístico descriptivo se presentan las variables cuantitativas como media y desvío estándar (DS) o mediana e intervalo intercuartil (IIC) según la distribución observada. Se presentan las variables categóricas como frecuencia relativa en porcentaje y frecuencia absoluta.

Para el cálculo de la tasa de ausentismo se utilizó como numerador la cantidad de turnos ausentes, y como denominador la totalidad de los turnos que cumplieron los criterios de inclusión y exclusión durante el mismo período. Se reportan intervalos de confianza del 95% estimados con la aproximación de Wilson (IC95%) [31].

Para evaluar la asociación univariada entre cada una de las variables independientes con el ausentismo a los turnos, se utilizó un modelo de regresión logística de efectos mixtos, con ordenada al origen aleatorio, considerando el natural agrupamiento de los turnos en cada paciente. Se presentan los Odds Ratio (OR) con sus intervalos de confianza de 95%.

Se consideraron como estadísticamente significativas las probabilidades menores a 0,05. El análisis estadístico descriptivo y la evaluación de cada asociación con el ausentismo se realizó utilizando el software STATA versión 13.1.

## Desarrollo de Modelos predictivos

Se utilizaron técnicas de ML supervisadas para construir los modelos predictivos utilizando el grupo de turnos de generación, y se utilizó el grupo de turnos de validación para validar los modelos. Las técnicas que se utilizaron son RL, GBM, XGBM, RF, DL. Los datos corresponden a turnos con ausentismo, los turnos considerados para el análisis fueron todos los turnos ausentes que corresponden a los criterios de inclusión. En el estudio se incluyeron todos los turnos que tuvo cada paciente, siguiendo metodología utilizada en estudios previos del mismo grupo de investigación. Si bien en el análisis bivariado se utilizó un modelo de regresión que consideró qué turnos estaban agrupados en el mismo paciente, los modelos de aprendizaje de máquinas no consideraron este natural agrupamiento. El entrenamiento y la validación se basaron en registros. Para la generación y validación de los modelos predictivos de ausentismo, se utilizaron los turnos ausentes y presentes (excluyendo los cancelados). Se dividió la muestra aleatoriamente en un grupo de turnos de generación (2/3 118.029 de la muestra turnos) y un grupo de turnos de validación (1/3 59.076 turnos). La división en el grupo de turnos de generación y validación se realizó por turnos y no por sujetos, ya que el objetivo de esta tesis fue generar un modelo que prediga la probabilidad de faltas a un determinado turno (el turno es la unidad de análisis) [32].

Se agrega, a continuación, una breve descripción de cada uno de los métodos utilizados.

**Regresión Logística:** Es un método de análisis predictivo que se utiliza para modelar la probabilidad de un resultado binario. Utiliza una combinación lineal de los diferentes tipos de entradas que pasan a través de la función logística. El desarrollo de predicciones usando regresión logística es fácil de implementar, de interpretar y su uso está muy generalizado en las publicaciones biomédicas [18]. Tiene muchos usos y varias ventajas. Es más rápido y requiere cálculos más estables. La regresión logística con múltiples variables independientes permite realizar una evaluación simultánea de la contribución de muchos predictores de

ausentismo. El modelo de regresión logística permite estimar la probabilidad condicional de presentarse a la cita médica [33].

**Modelo de potenciación del gradiente:** se ha utilizado en regresión y clasificación [18]. El modelo GBM construye un conjunto de árboles de decisión de forma secuencial, este conjunto de árboles de decisión se entrena secuencialmente, y cada árbol de decisión corrige gradualmente los residuos de todos sus árboles precedentes, minimizando las discrepancias entre todos los resultados observados y predichos [34]. GMB Se considera un poderoso instrumento para predecir resultados binarios. Se realiza aplicando de manera iterativa clasificadores simples y se combinan sus soluciones para obtener un mejor resultado de predicción [35].

**Modelo de potenciación extrema del gradiente:** hace predicciones, en promedio, mejores que las de un bosque aleatorio y casi tan buenas como las de las redes neuronales profundas. La mayor fortaleza de XGBM es su velocidad. Mientras que el uso eficiente del bosque aleatorio requiere generar cada árbol en paralelo en un clúster, y las redes neuronales profundas generalmente se ejecutan en GPU, XGBM se puede ejecutar en una sola CPU en menos de un tercio del tiempo que lleva cualquiera de los otros métodos [18]. Se pueden generar muchos modelos a partir de conjuntos de entrenamiento que contienen una gran cantidad de datos y una gran cantidad de descriptores. [36]

**Algoritmo del bosque aleatorio:** es un algoritmo de clasificación de árboles de decisión que producen diferentes modelos (bosque de árboles) al seleccionar un atributo de entrada al azar en cada iteración y aprender si los resultados de la clasificación son mejores o peores. Luego en cada iteración, el atributo se elimina o se incluye según los resultados de las iteraciones anteriores [18].

**Algoritmo de aprendizaje profundo:** El aprendizaje profundo, es un subconjunto del aprendizaje automático. Por su estructura, las redes neuronales, organizadas en múltiples capas, pueden abordar tareas complejas. Un tipo de redes neuronales, para varias situaciones clínicas, demostraron poseer un rendimiento superior en comparación con los algoritmos clásicos de aprendizaje automático, y, en algunos casos, lograron un rendimiento comparable o mejor al de expertos clínicos [37]. Los métodos de Deep learning se han utilizado en el campo de la atención médica. Son útiles para problemas que son difíciles de resolver con los métodos tradicionales. Son óptimos para manejar datos de gran volumen y dimensiones. Sin embargo, Daghistani et al encontraron un solo estudio que utiliza Deep Learning para predecir ausentismo de pacientes a clínicas ambulatorias [18]. M Dashtban y Weizi Li en su trabajo sobre estos métodos, sostienen que son métodos más robustos, que se obtienen buenos resultados en salud, la economía social y la información medioambiental, competentes para identificar patrones y dependencias que no pueden descubrir los seres humanos. Poseen potencial para revelar una estructura desconocida en la predicción del riesgo de ausentismo [38].

Se compararon los diferentes modelos por el área bajo la curva ROC (receiver operating characteristic). Se seleccionó el modelo con mayor área bajo la curva ROC. Se presentaron las medidas de discriminación y calibración de los modelos en el grupo de generación y de validación [20].

Para el análisis utilizando las aproximaciones de ML se utilizó el software estadístico R, de código abierto, versión 6.3.1, con el Rstudio versión 1.2.5 [4].

### **Consideraciones éticas**

El estudio fue analizado y aprobado por el Comité de Ética y Protocolos de Investigación (CEPI), del Hospital Italiano de Buenos Aires (HIBA). No se obtuvo consentimiento informado de los pacientes porque el análisis incluyó solamente datos recopilados retrospectivamente que fueron recolectados de las historias clínicas, extraídos de una base anonimizada (Resolución 1480/11 del Ministerio de Salud de la Nación). Se mantuvo el anonimato de los pacientes incluidos en este estudio, manteniendo la confidencialidad de acuerdo a la ley 25.326 de Habeas Data. A cada paciente se le asignó un código de identificación numérico, sólo conocido por el equipo de investigación.



## Resultados

### Análisis descriptivo y tasa de ausentismo

Análisis descriptivo: Se diseñó una cohorte retrospectiva para estudiar las características del ausentismo a turnos programados. Durante el período del estudio de 24 meses, se detectaron 4.839.335 turnos correspondientes a Clínica médica/Atención primaria registrados en el sistema de turnos del HIBA. Los turnos corresponden a 418.981 pacientes, fueron excluidos 4.662.230 turnos. El flujo de turnos incluidos en esta etapa se muestra en la Figura 1. Los turnos incluidos en este estudio fueron 177.105 y corresponden a 68.093 pacientes entre 18 y 45 años. Fueron divididos en 118.029 (66,6 %) turnos en el grupo de generación y 59.076 (33,3%) turnos en el grupo de validación.

De los 177.105 turnos incluidos, 63.063 estuvieron ausentes. La tasa de ausentismo global fue de 35,6% (IC95% 35,4 - 35,8). En los hombres, la tasa de ausentismo fue de 32,2% (21.461/66.615, IC95% 31,8 - 32,6) y en las mujeres fue de 37,1% (40.592/109.480, IC95% 36,79% - 37,36%). Tabla 4.

Las características basales de los pacientes se presentan en la Tabla 1, y las características de los turnos incluidos en la Tabla 2.

Dentro de las características basales de los pacientes, la media de edad fue de 32,2 (7,5) años, el 40,7% (27.289) de los pacientes pertenecen al sexo masculino. Tabla 1.

Dentro de los turnos incluidos, la mediana de edad fue de 33,4 años, con predominio femenino, 62,20 % (108.662), y 37,8% (66.615) de sexo masculino. Tenían una mediana de 13 (7-22) problemas de salud en su HCE y el 97,1% (172.053), tenían un score de comorbilidades de Charlson de 0 puntos. La mediana de tiempo que les lleva llegar al lugar de atención son 12,5 minutos (7,4-20,2). Con respecto a la cobertura de salud, el 61,7% (109.284) correspondían al PS del HIBA, y el 0,5% (923) a turnos privados sin cobertura médica de salud. Tabla 2.

Considerando el historial del paciente al momento del turno, la mediana de los turnos solicitados en el último año era de 3 (1-6), con una mediana de porcentaje de faltas de 20 (0-50). El 0,7% (1.294) había tenido una internación en los 30 días previos al turno, y 0,1% (241) estaban internados al momento del turno. Los pacientes que tuvieron una cirugía previa al turno fueron 0,4% (738). Un solo paciente había fallecido previo al turno solicitado.

Con respecto a las características del proceso de solicitud del turno, el 42,5 % (75.366) de los turnos fueron solicitados telefónicamente y el 21 % (37.262) de manera presencial. Una gran mayoría fueron turnos convencionales 99,7% (176.575) y el 0,3 % (530) sobretornos programados. La mediana de días desde la solicitud al turno es de 15 (7-26). La mayoría de los turnos solicitados corresponden al HIBA central. El turno de primera vez de los pacientes fue del 19,8% (35.047), y los de primera vez con el médico solicitado de 35,82% (905.129).

## Factores asociados a ausentismo

En la Tabla 3 se muestran los factores evaluados y sus OR para ausentismo en los turnos de consultas programadas, con sus intervalos de confianza de 95% y sus p valores. Las asociaciones se exploraron utilizando todos los turnos.

Los factores más fuertemente asociados a ausentismo dentro de las características basales de los pacientes, fueron: pertenecer a una cobertura médica privada OR 4,912 (IC95% 3,982- 6,059), todas las categorías de coberturas agrupadas por cuartiles de porcentajes de faltas con respecto al primer cuartil con asociación más fuerte para el cuarto cuartil OR 9,339 (IC95% 7,325- 11,906), haber solicitado el turno por HC 2,78 (IC95% 2,610- 2,964), la internación al momento del turno 3,72 (IC95% 2,541- 5,452), más de un turno en el mismo momento 2,615 (IC95% 1,271- 5,381), lugar físico del turno San Justo 1,34 (IC95% 1,28 - 1,4), presencia de otros turnos en el mismo día 1,94 (IC95% 1,83- 2,055), la cantidad de turnos solicitados el mismo día 1,859 (IC95% 1,767-1,956), el día sábado 1,58 (IC95% 1,255- 1,988), el mes de diciembre 1,51 (IC95% 1,404- 1,63), los días de alta falta 1,604 (IC95% 1,276-2,015), el lugar de alta falta 1,42 (IC95% 1,366- 1,474), el médico de alta falta 1,417 (IC95% 1,370- 1,465), la insuficiencia hepática 9,37 (IC95% 1,761- 49,84), la insuficiencia renal 1,56 (IC95% 1,055- 2,328), obesidad 1,14 (IC95% 1,06- 1,229), abuso de sustancias 1,47 (IC95% 1,222-1,781).

Los factores protectores más fuertes para ausentismo fueron la hora 14 del día 0,61 (IC95% 0,396-0,948), haber solicitado el turno por la web 0,79 (IC95% 0,765-0,826), tener cobertura médica del PS 0,68 (IC95% 0,658-0,703).

## Evaluación de modelos de Machine Learning para predicción de ausentismo

Todos los modelos de ML fueron estimados utilizando el set de validación y la totalidad de las variables. Las cuatro variables predictoras principales en casi todos los modelos desarrollados fueron la proporción de faltas total de los turnos solicitados en el año previo a la solicitud del turno, la cantidad de problemas activos en la HCE, el tiempo desde la solicitud al turno y la cantidad de turnos solicitados en el año previo al turno, como se muestra en los gráficos de la Figura 2 donde se listan los 10 principales predictores de cada modelo en orden de importancia decreciente.

La comparación entre los diferentes algoritmos estimada con el subgrupo de validación se presenta en la Tabla 5, con sus estadísticos de calibración y discriminación. Se observan mínimas diferencias entre calibración y discriminación de los diferentes modelos.

Los estadísticos de calibración y discriminación para los diferentes algoritmos se presentan en la tabla 4. El modelo que presenta un mejor rendimiento basado en XGBM logró un área bajo la curva de 0.690, para el grupo de generación el valor bajo la curva ROC fue de 0,734. En el análisis de calibración la pendiente de calibración para XGMB es la que mejor se ajusta a la predicha como se muestra en la Figura 3. Otros resultados semejantes a este son los obtenidos por GBM con un área bajo la curva ROC de 0.690, luego le sigue RF con un área bajo la curva ROC de 0.681, los otros dos modelos se alejan un poco de estos resultados.

## Discusión

Durante el desarrollo de esta tesis se estimaron las tasas de ausentismo a las consultas médicas programadas para los turnos de Clínica Médica/Atención Primaria de la Salud, en pacientes jóvenes (18-45 años de edad), y las tasas de ausentismo por sexo, se identificaron los factores asociados a ausentismo y se generaron modelos predictivos de ML para ausentismo. Se utilizaron datos demográficos y datos del sistema de la HCE para desarrollar cinco modelos de aprendizaje de máquinas supervisados para la clasificación de ausentismo de turnos médicos. La muestra corresponde a pacientes jóvenes que solicitaron al menos un turno dentro del periodo del estudio. Como la historia clínica electrónica está orientada a problemas, todos tienen que tener problemas ingresados.

Siendo un problema global las altas tasas de ausentismo, desde tiempo atrás se han desarrollado diferentes trabajos que apuntan a la solución del problema. Por eso en nuestro hospital, el grupo de ausentismo de clínica médica ha desarrollado desde años atrás varios estudios para determinar la tasa de ausentismo a consultas médicas programadas de los pacientes adultos, identificando los factores asociados y generando modelos predictivos de ausentismo. En uno de los estudios, dentro de los tres subgrupos estudiados, las tasas de ausentismo fueron consideradas altas, entre 25% y 34% [11]. La tasa de ausentismo detectada por nuestro estudio con pacientes jóvenes, aunque está dentro del rango encontrado en la literatura, es bastante más alta que las obtenidas en ese estudio realizado en nuestro hospital, esas tasas fueron, la tasa de ausentismo global, que fue de 25,53% para pacientes mayores de 18 años fue de 27,84%, siendo la mediana de edad de 60,4 años y en Clínica médica/Atención primaria de la salud, creemos que parte de esta tasa podría explicarse por las ausencias debidas a los pacientes jóvenes. Probablemente, al excluir al grupo de jóvenes, la tasa en los adultos mayores de 45 años, para la misma área de salud, sea más baja que la que se determinó en ese estudio.

El problema del ausentismo data de hace años, ya Turkcan et al., en 2013 encontraron en la literatura, que los turnos perdidos se asocian a mayores riesgos de los pacientes que padecen ciertas patologías: en cuanto a los controles glucémicos en pacientes diabéticos, mayor riesgo de mortalidad y hospitalización en pacientes en diálisis, el mayor fracaso de las terapias antirretrovirales, provocando que aumenten las tasas de mortalidad en pacientes con VIH y mayor mortalidad en pacientes psiquiátricos [33]. Analizando la literatura más actualizada no pareciera que hayan bajado las tasas globalmente, ya sea por la aplicación de alguna estrategia novedosa u otras razones, sino que se mantienen altas, por lo que siguen siendo una preocupación para todos los centros de salud a lo largo del mundo. El ausentismo a los turnos en las consultas externas de atención especializada es un problema que afecta directamente a la eficiencia del sistema sanitario y su análisis debe ser parte de un control de calidad para la mejora continua del sistema de salud, considerándose uno de los indicadores de la gestión hospitalaria [39].

Nelson y col reconocen que las tasas de ausentismo no han variado en los últimos 10 años, aparentemente no han mejorado debido a dos factores, uno sería que las intervenciones para evitar el ausentismo se aplican de la misma manera a todos los pacientes, como el llamado telefónico como recordatorio, y no sirve para la mayoría de los pacientes, por lo que sería una estrategia ineficiente, el otro factor se refiere a que los sistemas enfocados en intervenciones para predecir la inasistencia de manera individual no son sencillos de diseñar

debido a las múltiples causas probables de inasistencia, como por ejemplo desde las características individuales de los pacientes hasta los fenómenos ambientales [4].

Además del problema de las altas tasas de ausentismo, hay que considerar la variabilidad de las mismas. En la literatura se pueden encontrar numerosos ejemplos de variabilidad de las tasas de ausentismo, tanto globales como también en los diferentes servicios médicos de las instituciones de salud, para distintas edades y poblaciones. Por ejemplo, en 2012, en el General Hospital Chelsea HealthCare Center (MGH Chelsea), se obtuvo la tasa más alta para adultos mayores de 18 años de atención primaria de la salud, que fue del 16% [40]. En el Servicio de Medicina Preventiva del Hospital Clínico Universitario Lozano Blesa de Zaragoza, la tasa de ausentismo fue del 12,5% en las primeras consultas y 11,7% en consultas sucesivas para pacientes externos con diferentes programas de vacunación para grupos de riesgo [41]. La tasa global de ausentismo fue del 12,5% para un estudio en consultas ambulatorias de neurología [42]. Y en otro estudio del King Faisal Specialist Hospital & Research Centre la tasa de ausentismo fue del 11,3% [43]. Turkcan y col encontraron las tasas de ausentismo más altas en 62 artículos publicados en revistas, incluyendo la atención en salud mental y la atención primaria, siendo la media en todos los estudios del 23,8% las tasas fueron de: 24,3% para los estudios asiáticos, 14,9% para los estudios europeos y 27,1% para los estudios norteamericanos. También comentan que es llamativo que algunas clínicas norteamericanas reportan tasas de ausentismo altísimas del 48-64% [33]. En una revisión sistemática de 2018, de resultados obtenidos de 105 estudios de ausentismo, la tasa promedio de ausentismo fue del orden del 23% [44].

Los ejemplos anteriores son ilustrativos en referencia a la variabilidad encontrada en las tasas de ausentismo en distintos contextos como por ejemplo el pediátrico, y también en una clínica de pacientes autistas, donde la tasa de ausentismo fue del 9% y 15% para las consultas iniciales y de seguimiento respectivamente [45]. Las tasas de ausentismo del 20% son comunes en entornos ambulatorios y pueden superar el 50% en otros entornos. Esta situación lo que provoca es que los costos de atención médica aumenten y que los pacientes que faltan a las consultas tengan que buscar una atención de emergencia o que se produzcan ingresos hospitalarios que podrían haberse evitado con la atención ambulatoria adecuada [33]. En un estudio de 2016 desarrollado por Kheirkhah y col, mostraron que la tasa de ausentismo fue alta con un costo económico significativo, siendo la tasa media de ausencias de 10 clínicas estudiadas del 18,8% [5]. Con respecto a nuestro hospital, desconocemos cuál es la variación de las tasas de ausentismo en nuestra población de estudio a lo largo del tiempo. Esperamos también poder investigar cuáles son las tasas para las distintas épocas del año, y la variación de la tasa a lo largo del tiempo de espera a la fecha del turno cuanto el tiempo es largo, se necesita obtener más datos que permitan profundizar en el conocimiento del fenómeno del ausentismo.

El ausentismo implica una carga para el sistema hospitalario, la implementación de sistemas de recordatorios no ha sido eficiente como para producir un impacto en las tasas de ausentismo [5,46]. Bech sostiene que la literatura sobre la inasistencia revela una inmensa necesidad de estudios que aborden el problema del ausentismo aplicando la teoría económica y los estándares de análisis [46]. El alto costo que significa que las tasas de ausentismo sean altas merece que se solucione este problema lo antes posible. La pérdida de recursos económicos trae aparejada también el deterioro de la eficiencia en los recursos humanos de los centros de salud, con todo lo que implica, siendo un problema general el de

las altas tasas de ausentismo en los diferentes centros tanto públicos como privados, se termina afectando en última instancia a la salud pública de la nación. La utilidad de desarrollar modelos predictivos más eficientes permitirá caracterizar con más precisión a los pacientes y turnos de alta falta, siendo esto necesario para llegar a estimar los altos costos asociados a ausentismo aún desconocidos. Así se podrá mejorar el manejo de los recursos hospitalarios tanto humanos como económicos, y de esta manera poder evitar la pérdida de seguimiento de los pacientes con todo lo que ello implica, ya que la pérdida de la posibilidad de su atención afecta la salud de los mismos en distintos grados [11].

La necesidad de solucionar el problema del ausentismo lleva a estimar las tasas en las distintas poblaciones de pacientes, por esa razón decidimos estudiar los pacientes jóvenes, ya que es un rango de edad de los pacientes que muestra una alta falta a las citas programadas [10,13,47,48], y cuyo abordaje no ha sido realizado aún en nuestro hospital. En este estudio la tasa de ausentismo estimada se considera alta, del 35,6%, aproximadamente, las faltas corresponden a 1 de cada 3 turnos, una de las razones que puede explicar que los pacientes jóvenes falten tanto podría estar relacionada a una posible percepción mejor sobre su salud [41]. Se hace imprescindible seguir explorando, dentro de esta población de adultos jóvenes, los factores particulares que puedan estar relacionados al ausentismo, como también caracterizar dentro del grupo a los pacientes con más probabilidad de faltar. Por otro lado, también hay que analizar las causas individuales que los caracterizan, y determinar patrones para un análisis más exhaustivo que podría permitir llegar a lograr buenas predicciones de ausentismo. Por esta razón, y para ir ahondando en el análisis del grupo, también se estimaron las tasas de ausentismo por sexo. Aunque las tasas de hombres y mujeres son altas, en el grupo de las mujeres es más alta que en el de los hombres. Para conocer las diferencias de ausentismo a los turnos entre hombres y mujeres, sería importante en un futuro explorar las causas debidas a esta diferencia y de este modo encontrar qué factores están influyendo en la población de mujeres para que el ausentismo sea tan alto. Hasta el momento no se encontró en la literatura algún estudio que permita explicar el alta falta de las mujeres comparada con la de los hombres, aunque Montesino y col encuentran en la literatura, que la relación entre el sexo y el ausentismo en las consultas no aparece de forma concluyente, si bien, las referencias en cuanto a una mayor frecuencia en varones son superiores [39]. Queda como desafío para nuevos trabajos de investigación sobre el tema.

Luego de determinar las tasas, se analizaron los factores asociados al ausentismo. Dentro de las características de los pacientes incluidos, el promedio de edad fue de 32,2 años, y los pacientes del sexo masculino fueron el 40,7 %, dentro las características clínicas y comorbilidades del paciente, los pacientes con obesidad fueron el 15,9 % y consumidores de tabaco (tabaquismo) fueron el 14,8 %, y la media del porcentaje de faltas fue de 25. Para los turnos incluidos la mediana de edad fue de 33,4 años, y el 37,8 % corresponde al sexo masculino, los factores que se presentan con mayor frecuencia son, la cobertura médica del plan de salud, el turno convencional, el médico de alta falta, el método telefónico como método utilizado para la solicitud del turno, el tabaquismo, el tiempo desde el domicilio al lugar de atención con una mediana de 12,5, el IBM con una mediana de 24. Entre los factores de riesgo más fuertes para ausentismo encontramos la cobertura médica privada, la cobertura privada de alta falta, la HC como método utilizado para la solicitud del turno, la internación actual, más de un turno en el mismo momento, presencia de otros turnos en el mismo día, cantidad de turnos solicitados el mismo día, la insuficiencia hepática. Los factores con mayor efecto preventivo de faltar fueron, el sexo masculino, la cobertura médica del Plan de Salud,

las 14 hs dentro de las horas del día. El resto de los factores, o no presentan asociación, o esta no es tan fuerte como la de los anteriores.

Según Drewek y col, el tiempo que transcurre hasta el día del turno es un factor de riesgo asociado con las ausencias. Consideran que se necesitan más estudios para identificar e implementar enfoques efectivos que reduzcan el tiempo de espera de los turnos, en ese caso estarían dirigidos especialmente a los nuevos pacientes de las subespecialidades pediátricas [1], sin embargo, en nuestro estudio, el tiempo desde la solicitud del turno no mostró estar asociado al ausentismo. También hay estudios que mostraron que los indicadores más importantes fueron las tasas de ausentismo previas y el número de ingresos hospitalarios [43]. En otro estudio se estimó que los factores más importantes fueron las tasas de ausentismo previas a los turnos y el número de ingresos hospitalarios, cuanto mayor es la tasa de ausentismo con anterioridad, mayores son las posibilidades de que un paciente no se presente en las visitas posteriores [43].

Kheirkhah y col, como comentamos anteriormente, muestran que los costos asociados al ausentismo son muy altos y las consecuencias que provoca esta situación son serias [5], la importancia de desarrollar estrategias que mejoren este problema tiene que ser una de las prioridades de nuestro centro de salud. El problema del ausentismo es global, el desafío de encontrar una respuesta al problema del ausentismo a los turnos programados en los centros de salud lleva años de estudio, actualmente adquirieron gran importancia los modelos predictivos de ML para detectar con gran precisión y rapidez a los pacientes que tienen probabilidad de ausentarse, y poder desarrollar un modelo de predicción eficiente. Desde los años 80 se comenzó a estudiar el ausentismo de pacientes a citas programadas. Se demostró que se pueden predecir con precisión las ausencias mediante árboles de decisión, y detectar factores de ausentismo con regresión logística [43]. En cuanto a los factores predictivos, para Nelson y col, no sirve tomar en cuenta sólo los más genéricos basándose en modelos estadísticos simples lineales y de baja dimensión, sino, que no hay que descartar ninguno cuando se desarrollen los modelos predictivos [4].

La comparación entre los diferentes algoritmos estimada con el subgrupo de validación se presenta en la Tabla 4, con sus estadísticos de calibración y discriminación. Se observan mínimas diferencias entre calibración y discriminación de los diferentes modelos, siendo los algoritmos de XGBM y GBM los que obtuvieron las mayores AUC ROC y las mejores métricas de calibración. Esta estadística, que mide la capacidad del modelo para discriminar entre los pacientes que se ausentaron a los turnos frente a los que asistieron fue, para XGBM, de 0.694, le sigue GBM con un área bajo la curva de 0,690. Los resultados de los modelos predictivos de ML que desarrollamos no llegan a alcanzar el nivel encontrado en otros trabajos con una curva ROC mayor [4,18,38], considerando una excelente discriminación un valor de 0,861 bajo la curva ROC [43]. DL ha demostrado tener menor poder de discriminación, sin embargo, arroja resultados semejantes con respecto a las variables más importantes. En el estudio se incluyeron todos los turnos que tuvo cada paciente, siguiendo metodología utilizada en estudios previos del mismo grupo de investigación. Si bien en el análisis bivariado se utilizó un modelo de regresión que consideró qué turnos estaban agrupados en el mismo paciente, los modelos de aprendizaje de máquinas no consideraron este natural agrupamiento.

El resultado de nuestros modelos nos muestra que las diez variables predictivas de mayor importancia son prácticamente las mismas en los cinco. Las tres con mayor importancia son: la proporción de faltas de los turnos solicitados en el año previo a la solicitud del turno, la cantidad de problemas activos en la HCE del paciente y el tiempo desde la solicitud del turno. El resto de las variables aparecen con menor importancia a estas tres, y con importancia decreciente que varía según el modelo.

Está demostrado que el tiempo hasta la solicitud del turno es un predictor importante, siendo un factor de riesgo asociado a las ausencias [1, 18], inclusive en un estudio en el se evaluaron las tasas de ausentismo en distintos tiempos, aunque no se desarrollaron modelos de ML, concluyeron que las tasas de ausentismo variaban según el tiempo de espera hasta el turno [49], sería interesante considerar este dato para estudiar las tasas de ausentismo en los tiempos relacionado con esta variable. Una variable para considerar es la edad, que aparece entre las 10 de más importancia, habrá que analizar este predictor, ver si en un futuro modelo tiene este comportamiento, porque podría indicar que dentro del rango de edad de los pacientes estudiados pueden existir subgrupos etarios con más probabilidad de faltar. Siendo un estudio de pacientes jóvenes otra variable a tomar en cuenta es el número de personas viviendo en el mismo hogar, considerada en base a la percepción de que las familias con varios hijos a menudo parecen ausentarse a los turnos debido a la complicada logística de manejar las diferentes necesidades de los miembros de la familia, no resultaría difícil de estimar dado que se puede extraer el número de individuos de nuestra HCE que comparten el mismo domicilio del paciente [2]. AlMuhaideb y col, luego de un análisis de la literatura en su estudio sobre predictores de ausentismo utilizando algoritmos de IA [43], muestran que aparecen otros predictores como ser menor de edad, no tener problemas cardíacos, tener depresión y ser drogodependiente, no estar casado y viajar largas distancias. Hay una cantidad de variables para evaluar que no hemos considerado en nuestro estudio, que podrían estar ejerciendo cierta influencia sobre las otras, y que si se incorporan a los modelos tal vez podían obtenerse resultados más precisos. Aparecen tres predictores asociados a las características clínicas y comorbilidades de los pacientes que tienen importancia en nuestros modelos, BMI (Índice de masa corporal), anemia y abuso de sustancias, fueron incorporadas para este estudio, sería interesante incorporar otras características de los pacientes para ser analizadas como posibles factores de riesgo de ausentismo.

Para mejorar nuestros modelos, aunque no es el objetivo de la tesis, sería interesante considerar los logros de distintos autores que encontramos en la literatura, por ejemplo, vemos como Baskaran y col muestran cómo lograron mejorar el desempeño de un modelo de predicción, entre otras cosas agrandaron la base de datos a un conjunto de atributos más completo y también generaron nuevas variables categóricas [50]. También Jothi et al consideran que la precisión de los métodos de ML varía según las características de los conjuntos de datos, como también del tamaño del conjunto de datos entre los conjuntos de entrenamiento y prueba. Los problemas que se presentan con los conjuntos de los datos, también pueden deberse a que los datos disponibles generalmente son escasos. Piensan en que habría que considerar en eliminar de los modelos los predictores que no son importantes [18]. Pero no existe un método de ML adecuado para resolver todos estos problemas. Por otro lado, habría que pensar en la incorporación de variables que no se registran en la HC, resulta interesante que Dashtban y Li hayan desarrollado un modelo basado en factores de salud, medio ambiente y economía social con buenos resultados [38]. Nuestro equipo podría analizar qué mejoras pueden introducirse en los modelos, ya que todavía hay margen para

mejorar la predicción y esto se puede lograr mediante mejores atributos de predicción y / o técnicas mejoradas de aprendizaje automático [50].

Como limitaciones del estudio encontramos que nuestros modelos predictivos de ML no alcanzaron el resultado esperado según otros estudios encontrados en la literatura, en los que el área bajo la curva ROC muestra un mayor rendimiento predictivo, lo que dificulta pensar en cuál modelo elegiríamos para desarrollar predicciones de ausentismo. Es probable que, como los datos se analizaron retrospectivamente, no se cuente con toda la información esperada, dado que no fue posible incluir en el estudio variables nuevas que puedan ser de importancia para la clasificación. Por otro lado si se decide realizar más adelante otros estudios, habría que pensar en la necesidad en ellos otras variables que permitieran mejorar la performance de los mismos, como algunas que aparecen en la literatura, raza, número de miembros de la familia que conviven con el paciente, especialmente si hay niños o adultos mayores que requieren cuidado, situaciones de viudez o pérdida de algún hijo que fueran situaciones cercanas al turno y que podrían asociarse hipotéticamente con ausentismo, otras podrían ser el nivel de estudios, nivel socioeconómico, provincia, localidad, embarazo, cantidad de indicaciones farmacológicas. En el estudio se incluyeron todos los turnos que tuvo cada paciente, siguiendo metodología utilizada en estudios previos del mismo grupo de investigación. No se han identificado las características de los pacientes que tienen mayor probabilidad de ausentarse ni de los turnos con mayor probabilidad de ausentismo. La literatura presenta una serie de estudios realizados en distintos centros y para diferentes especialidades, y con distintos diseños, nuestros resultados no serían extrapolables a otras poblaciones e instituciones, también otras diferencias que encontramos en los trabajos, es que no se utilizan los mismos criterios de inclusión ni las definiciones de las variables [41]. Otras diferencias en las características con nuestro hospital son la gestión de turnos y la población de pacientes, por lo tanto, los resultados no pueden ser generalizados.

Dentro de las fortalezas de nuestro trabajo, encontramos que los datos utilizados en el modelo de predicción son recopilados como parte de las operaciones clínicas de rutina llevadas a cabo en el hospital, por lo que no fue necesario trabajar especialmente para conseguirlos. En el estudio se incluyeron una gran cantidad de turnos y la posibilidad de analizar una cantidad significativa de variables, luego será necesario analizar si se incorporan nuevas variables en el modelo. El hecho de que el período de estudio fue de 24 meses permite que los resultados sean más balanceados o estables, considerando alguna variación de la tasa de ausentismo que pudiera haber ocurrido por temas estacionales, sociales, o de otro tipo. También nuestro equipo ya viene entrenado en este tipo de estudios, y busca constantemente nuevos modelos y estrategias para aportar más información al problema del ausentismo, esto permite que el desarrollo de la tesis haya sido enfocado a un problema relacionado con las necesidades del hospital con una base de conocimiento previo. Consideramos un avance para el equipo el hecho de haber desarrollado modelos predictivos de ML con este estudio.



## Conclusiones

Se alcanzaron los tres objetivos propuestos para este trabajo. Estimamos la tasa de ausentismo y sus factores asociados para la población de adultos jóvenes en los turnos de Clínica médica/Atención primaria de la salud de nuestro hospital y la tasa de ausentismo por sexo para la población de pacientes entre 18 y 45 años del servicio de Medicina ambulatoria/Atención primaria de la salud. Se desarrollaron modelos predictivos de ML para ausentismo. La alta tasa de ausentismo en esta población de pacientes jóvenes requiere atención por parte de nuestra institución, se necesita aplicar estrategias más eficientes para evitar las consecuencias negativas en la atención de los pacientes, En cuanto a la tasa por sexo, se requiere explorar las causas del gran ausentismo en el grupo de mujeres principalmente.

El conocimiento de los factores asociados al ausentismo y la importancia de las variables detectadas en nuestros modelos resultan útiles para la implementación de iniciativas de mejora de la calidad de la atención de los pacientes que permitan reducir la inasistencia y maximizar la productividad en el entorno del consultorio de medicina ambulatoria. El desarrollo de modelos predictivos para los pacientes ambulatorios que se ausentan permite conocer anticipadamente el comportamiento de los pacientes, pueden contribuir a evaluar la relación entre la exposición a factores específicos y el riesgo de ausentarse. Como el ausentismo puede llegar a ser un marcador para la causa de eventos de salud que requieren consultas a guardia, internaciones o incluso que podrían causar la muerte, tiene que ser tratado para disminuir los riesgos que se asocian a tasas altas. Los modelos de riesgo de ausentismo como los que hemos desarrollado se pueden utilizar para mejorar el manejo de recursos de los centros de salud, en particular en este caso de nuestro hospital lo que contribuiría a mejorar el acceso a la atención desarrollando estrategias particulares para este grupo de adultos jóvenes, a partir de los resultados de los modelos desarrollados. Sin embargo, la predicción obtenida no fue la esperada, por lo que sería necesario desarrollar nuevos modelos.

En este estudio, presentamos una evaluación de cinco técnicas de aprendizaje automático para predecir la ausencia de los pacientes utilizando datos médicos de la HCE. Es necesario actualizar los sistemas de las historias clínicas para la recopilación de datos nuevos, que permitirá agilizar cualquier intento de nuevos desarrollos de modelos más eficientes, se necesita mejorar la predicción de las variables, obtener mayor precisión, los modelos se podrán aplicar en las distintas áreas médicas con las correcciones correspondientes.

Los resultados encontrados en la literatura y en nuestro hospital muestran la importancia de desarrollar estrategias eficientes dirigidas a las diferentes poblaciones de pacientes en todas las áreas de salud donde los pacientes sean atendidos luego de obtener un turno programado. El desarrollo de estrategias para evitar el ausentismo de pacientes, basadas en los resultados de los modelos predictivos, implica la necesidad de investigar acerca de la situación socioambiental de los pacientes, desarrollar estrategias de educación para la concientización de las consecuencias de faltar a los turnos programados, investigar con profundidad los turnos que se otorgan con mucho tiempo de anticipación, las causas por las que ciertos médicos son de alta falta (frecuencia de faltas de más del 30%) y seguir especialmente a los pacientes con alta proporción de faltas, también considerar a los que presentan ciertas comorbilidades y abuso de sustancias.

Para el servicio de Clínica médica/ Atención primaria de la salud en pacientes de 18 a 45 años sería conveniente

1. Seguir investigando sobre los factores asociados al ausentismo, detectar otros posibles factores asociados que no hayan aparecido en este trabajo, especialmente en la población de mujeres jóvenes.
2. Estudiar las características particulares de los turnos otorgados con mucho tiempo de anticipación, de los pacientes con probabilidad de faltar y los turnos con mayor probabilidad de ausencias.
3. Desarrollar nuevos modelos de ML que permitan mayor un mayor poder de predicción.
4. Desarrollar nuevas estrategias para disminuir la tasa de ausentismo en esta población de pacientes.
5. Estimar costos asociados al ausentismo.

## Tablas

**Tabla 1.** Características de los pacientes incluidos. Variables cuantitativas: se presentan como media y desvío estándar (DS) o mediana e intervalo intercuartil (IIC) según la distribución observada. Variables categóricas: se presentan como frecuencia relativa en porcentaje y frecuencia absoluta.

	<b>Clínica médica / atención primaria de la salud (177.105)</b>
<b>Características basales de los pacientes</b>	
Edad	32,2 (7,5)
Sexo (masculino)	40,7% (27.289)
<b>Características clínicas y comorbilidades del paciente</b>	
Hipertensión arterial	1,3% (905)
Diabetes	0,9% (634)
Leucemia	0,06% (41)
cáncer	0,62% (425)
linfoma	0,13% (86)
Enfermedad oncológica (f onco)	0,8% (544)
Enfermedad coronaria	0,1% (105)
ACV	0,21% (145)
Enfermedades psiquiátricas	8,2% ( 5.630)
Anemia	5,3% ( 3.610)
BMI	24,1 (21,3-27,7)
Obesidad	15,9% (3.550)
Tabaquismo	14,8% (10.107)
Alcoholismo	0,5% (348)
Abuso de sustancias	0,5% (370)
Score de comorbilidades de Charlson	0 97,6% (66.478)

	1	1,3% (927)
	2	0,9% (611)
	3	0,07% (49)
Demencia		0,01% (6)
Enfermedad vascular		0,01% (6)
EPOC		0,06% (44)
Hemiplejia		0,02% (16)
ICC		0,09% (63)
Insuficiencia hepática		0,01% (8)
Insuficiencia renal		0,15% (105)
SIDA		0% (1)
Úlcera gástrica		0,12% (83)
Cantidad de turnos		2 (1-3)
Cantidad de ausencias		1 (0-1)
Porcentaje de faltas		25 (0-60)

**Tabla 2.** Características de los turnos incluidos. Turnos de Clínica médica / atención primaria de la salud- Grupo de turnos de generación- Grupo de turnos de validación. Variables cuantitativas: se presentan como media y desvío estándar (DS) o mediana e intervalo intercuartil (IIC) según la distribución observada. Variables categóricas: se presentan como frecuencia relativa en porcentaje y frecuencia absoluta.

	<b>Clínica médica / atención primaria de la salud (177.105)</b>	<b>Generación (118.029)</b>	<b>Validación (59.076)</b>
<b>Características basales de los pacientes</b>			
Edad	33,4 (26,7-39,3)	33,4 (26,7-39,4)	33,4 (26,7-39,3)
Sexo (masculino)	37,8% (66.615)	37,7% (44.279)	38,3% (22.336)
Cobertura médica	PSalud Privado	61,7% (109.284) 0,5% (923)	61,7% (72.859) 0,5% (629)
Cuartilos de cobertura médica (cuartilos de faltas)			
Cuartilo inferior o baja falta (1)	63,7% (112.801)	63,7% (75.250)	63,5% (37.551)
Cuartilo (2)	24,5% (43.477)	24,5% (28.956)	24,5% (14.521)
Cuartilo (3)	11,3% (20.013)	11,2% (13.265)	11,4% (6.748)
Cuartilo superior o alta falta (4)	0,4% (814)	0,4% (558)	0,4% (256)
Cantidad de problemas en la HC electrónica	13 (7-22)	13 (7-22)	13 (7-22)
Distancia desde el domicilio al lugar de atención			
Distancia lineal (fórmula)	2,8 (1,4-5,7)	2,8 (1,3-5,7)	2,8 (1,4-5,8)
Distancia (Google Maps)	3,9 (2-7,3)	3,9 (2-7,3)	3,9 (2-7,3)
Tiempo desde el domicilio al lugar de atención	12,5 (7,4-20,2)	12,5 (7,4-20,2)	12,4 (7,4-20,2)
<b>Características del proceso de solicitud de turno</b>			
Método utilizado para la solicitud del turno			
Telefónico	42,5% (75.366)	42,4 (50.053)	42,8 (25.313)
Personal	21% (37.262)	21 (24.853)	21 (12.409)
Web	28,8% (50.983)	28,8% (34.077)	28,6 (16.906)
Por HC	7,6% (13.494)	7,6% (9.046)	7,5% (4.448)
Sobretorno o turno convencional			
Turno convencional	99,7% (176.575)	99,7% (117.681)	99,7% (58.894)
Sobretorno programado	0,3% (530)	0,3% (348)	0,3% (182)
Tiempo desde la solicitud al turno	15 (7-26)	15 (7-26)	15 (7-25,8)

<b>Historial del paciente</b>			
Cantidad de turnos solicitados	3 (1-6)	3 (1-6)	3 (1-6)
Porcentaje de faltas en el último año previo a la solicitud del turno	20 (0-50)	20 (0-50)	20 (0-50)
Internación previa (en 30 días previos a la fecha del turno)	0,7% (1.294)	0,8% (900)	0,6% (394)
Internación actual	0,1% (241)	0,1% (165)	0,1% (76)
Fallecimiento previo al turno	0% (1)	0% (1)	0% (0)
Cirugía previa	0,4% (738)	0,4% (494)	0,4% (244)
<b>Características del turno solicitado</b>			
Más de un turno en el mismo momento	0,04% (74)	0,04% (45)	0,05% (29)
Lugar físico del turno			
Hospital central	36,1% (64.052)	36% (42.591)	36% (21.461)
San Justo	17% (30.137)	17% (20.120)	17% (10.017)
Belg II. Virrey del Pino	4,1% (7.284)		
Flores I - Carabobo	3,8% (6.788)	3,8% (4.526)	3,8% (2.262)
Turno de primera vez			
paciente	19,8% (35.047)	19,6% (23.184)	20% (11.863)
médico	40% (70.825)	39,9% (47.053)	40,2% (23.772)
Hora del día			
9	12,4% (22.029)	12,5% (14.761)	12,3% (7.268)
10	10,4 % (18.444)	10,3% (12.219)	10,5% (6.225)
17	10,1 % (17.912)	10,2% (12.054)	9,9% (5.858)
8	9% (16.067)	9% (10.712)	9% (5.355)
16	8,7% (15.400)	8,7% (10.318)	8,6% (5.082)
Día de la semana			
Lunes	20,7% (36.811)	20,8% (24.549)	20,7% (12.262)
Martes	21,7% (38.545)	21,8% (25.742)	21,6% (12.803)
Miércoles	20% (35.441)	20% (23.639)	20% (11.802)
Jueves	20,5% (36.358)	20,5% (24.214)	20,5% (12.144)
Viernes	16,5% (29.263)	16,4% (19.409)	16,7% (9.854)
Sábado	0,4% (687)	0,4% (476)	0,3% (211)
Mes del turno			
1	7,1% (12.674)	7,1% (8.414)	7,2% (4.260)
2	6,7% (11.880)	6,7% (7.893)	6,7% (3.987)
3	8,7% (15.370)	8,6% (10.182)	8,8% (5.188)
4	9% (15.971)	9% (10.659)	9% (5.312)
5	9% (15.897)	9% (5.312)	9% (15.897)
6	9,4% (16.708)	9,3% (11.064)	9,5% (5.644)
7	8,3% (14.740)	8,3% (9.789)	8,4% (4.951)
8	9% (16.084)	9% (10.692)	9,1% (5.392)

	9	8,9% (15.882)	9% (10.680)	8,8% (5.202)
	10	8,2% (14.609)	8,2% (9.665)	8,3 % (4.944)
	11	8,3% (14.698)	8,3 % (9.854)	8,2% (4.844)
	12	7,1% (12.592)	7,2% (8.413)	7% (4.179)
Horas de alta falta		12,9 % (22.974)	12,8% (15.154)	13,2% (7.820)
Días de alta falta		0,4% (687)	0,4% (476)	0,3% (211)
Meses de alta falta		24,8% (44.040)	24,8% (29.266)	25% (14.774)
Lugar de alta falta		36,2% (140.911)	20,5% (24.153)	20,4% (12.041)
Médico de alta falta		53,2% (123.875)	30,1% (35.535)	30% (17.695)
Cantidad de turnos solicitados el mismo día		1 (1-1)	1 (1-1)	1 (1-1)
Proporción de faltas del médico a cargo del turno		27,8 (7,3)	27,8 (7,3)	27,7 (7,2)
Presencia de otros turnos y cantidad de ellos solicitados ese mismo día		6,3% (11.177)	6,3% (7.492)	6,2% (3.685)
<b>Clima del día del turno</b>				
Temperatura máxima, mínima y media				
	Máxima	22,5 (6,4)	22,6 (6,4)	22,5 (6,3)
	Mínima	13,6 (5,6)	13,7 (5,6)	13,6 (5,6)
	Media	17,7 (5,8)	17,8 (5,9)	17,7 (5,8)
Precipitaciones		0 (0-0,3)	0 (0-0,3)	0 (0-0,3)
Sensación térmica	máxima	21,8 (16,6-28,1)	21 (16,6-28,1)	21,8 (16,6-28,1)
	mínima	13,5 (6)	13,5 (6)	13,5 (6)
Presión barométrica		1013.4 (5,8)	1013.4 (5,9)	1013.4 (5,8)
<b>Características clínicas y comorbilidades del paciente</b>				
Score de comorbilidades de Charlson				
	0	97,1% (172.053)	97,2% (114.660)	97% (57.393)
	1	1,7% (3.078)	1,8% (2.069)	1,7% (1.079)
	2	1% (1.765)	1% (1.156)	1% (609)
	3	0% (724)	0,1% (80)	0% (44)
Embarazo		0,7% (1.323)	0,8 (881)	0,75 (442)
Hipertensión arterial		1,6 % (2.858)	1,6% (1.884)	1,6% (974)

Diabetes mellitus	1,2% (2.086)	1,2% (1.411)	1,1% (675)
Enfermedad oncológica	0,9% (1.586)	0,9% (1.052)	0,9% (534)
Enfermedad coronaria	0,2% (329)	0,2% (225)	0,2% (104)
Enfermedades psiquiátricas	11% (19.430)	11% (12.964)	11% (6.466)
Anemia	7% (12.493)	7% (8.290)	7,1% (4.203)
BMI	24 (21,3-27,7)	24 (21,3-27,7)	24 (21,3-27,7)
Obesidad	16% (10.340)	16% (6.883)	16% (3.457)
Tabaquismo	16,7% (29.590)	16,7% (19.678)	16,7% (9.912)
Alcoholismo	0,6% (1.130)	0,6% (740)	0,6% (390)
Abuso de sustancias	0,7% (1.207)	0,7% (815)	0,7% (1207)
Linfoma	0,1% (223)	0,1% (150)	0,1% (73)
Leucemia	0% (90)	0,1% (62)	0% (28)
Accidente cerebro vascular	0% (36)	0% (19)	0% (17)
Demencia	0% (36)	0% (19)	0% (17)
Enfermedad vascular periférica	0% (13)	0% (11)	0% (2)
EPOC	0% (142)	0,1% (97)	0% (45)
Hemiplejía	0% (35)	0% (22)	0% (13)
Insuficiencia cardíaca congestiva (icc)	0,1% (220)	0,1% (146)	0,1% (74)
Insuficiencia hepática	0% (13)	0% (9)	0% (4)
Insuficiencia renal (irc)	0,1% (266)	0,1% (170)	0,1% (96)
SIDA	0% (3)	0% (3)	0% (0)
Úlcera gástrica	0,1 (288)	0,2% (201)	0,1% (87)



**Tabla 3.** Factores asociados al ausentismo en turnos de consultas programadas. Se presentan los OR crudos de cada variable para ausentismo, sus intervalos de confianza de 95% y los p valores.

	<b>OR (IC 95%)</b>
<b>Características basales de los pacientes</b>	
Edad	0,987 (IC95% 0,985-0,989) p <0,001
Sexo (masculino)	0,78 (IC95% 0,754- 0,806) p <0,001
Cobertura médica	
Plan de Salud	0,68 (IC95% 0,658-0,703) p <0,001
Privado	4,91 (IC95% 3,982- 6,059) p <0,001
Cuartilos de cobertura médica (cat ref 1er cuartil)	
Cuartilo inferior o baja falta 1	1 (ref)
Cuartilo 2	1,32 (IC95% 1,278-1,377) p <0,001
Cuartilo 3	1,77 (IC95% 1,690-1,869) p <0,001
Cuartilo superior o alta falta 4	9,34 (IC95% 7,325- 11,906) p <0,001
Cantidad de problemas en la HC electrónica	1 (IC95% 1,001- 1,003) p <0,001
Distancia desde el domicilio al lugar de atención	
Distancia lineal (fórmula)	1,02 (IC95% 1,014- 1,025) p <0,001
Distancia (Google Maps)	1,01 (IC95% 1,009- 1,017) p <0,001
<b>Características del proceso de solicitud de turno</b>	
Método utilizado para la solicitud del turno (cat ref Telefónico)	
Telefónico	1 (ref)
Personal	1,18 (IC95% 1,141- 1,232) p <0,001
Web	0,79 (IC95% 0,765-0,826) p <0,001
Por HC	2,78 (IC95% 2,610- 2,964) p <0,001
Sobretornos con respecto a turnos convencionales	1,06 (IC95% 0,805-1,400) p <0,670
Tiempo desde la solicitud al turno	1,01 (IC95% 1.008 -1.010) p <0,001
<b>Historial del paciente</b>	
Cantidad de turnos solicitados	1,04 (IC95% 1,037-1,045) p <0,001
Porcentaje de faltas en el último año previo a la solicitud del turno	1,01 (IC95% 1,010- 1,011) p <0,001
Internación previa (30 días previos a la fecha del turno)	1,08 (IC95% 0,928- 1,273) p <0,299
Internación actual	3,72 (IC95% 2,541- 5,452) p < 0,001

Fallecimiento previo al turno	1
Cirugía previa	0,870 (IC95% 0,694-1,092) p < 0,232
<b>Características del turno solicitado</b>	
Más de un turno en el mismo momento	2,615 (IC95% 1,271- 5,381) p < 0,009
Lugar físico del turno (agrupado) (cat ref Central) San Justo Belg II Virrey del Pino	1 (ref) 1,34 (IC95% 1,28 -1,4) p <0,001 0,89 (IC95% 0,862-0,924) p <0,001
Turno de primera vez paciente (primer turno en el hospital) médico (primera vez con ese médico)	0,97 (IC95% 0,938-1) p <0,135 1,01 (IC95 0,984- 1,045) p <0,342
Hora del día (cat ref 7l)	1 (ref) 7 8 0,70 (IC95% 0,453-1,082) p <0,109 9 0,71 (IC95% 0,461-1,101) p <0,128 10 0,72 (IC95% 0,469-1,120) p <0,148 11 0,72 (IC95% 0,470-1,124) p <0,152 12 0,67 (IC95% 0,434-1,040) p <0,075 13 0,67 (IC95% 0,431-1,033) p <0,070 14 0,61 (IC95% 0,396-0,948) p <0,028 15 0,65 (IC95% 0,419-1) p <0,052 16 0,70 (IC95% 0,456-1,091) p <0,118 17 0,72 (IC95% 0,468-1,117) p <0,145 18 0,65 (IC95% 0,424-1,013) p <0,057 19 0,65 (IC95% 0,423-1,018) p <0,060
Día de la semana (cat ref lunes)	1 (ref) Lunes Martes 0,97 (IC95% 0,930-1,016) p <0,221 Miércoles 0,97 (IC95% 0,931- 1,020) p <0,271 Jueves 0,94 (IC95% 0,901-0,986) p <0,011 Viernes 1,05 (IC95% 1,004-1,104) p <0,033 Sábado 1,58 (IC95% 1,255-1,988) p <0,001
Mes (cat ref enero)	1 (ref) 1 2 1,09 (IC95% 1,014-1,181) p <0,020 3 1,12 (IC95% 1,042-1,202) p <0,002 4 1,23 (IC95% 1,150-1,324) p <0,001 5 1,21 (IC95% 1,129-1,300) p <0,001 6 1,28 (IC95% 1,194-1,374) p <0,001 7 1,29 (IC95% 1,202-1,388) p <0,001 8 1,18 (IC95% 1,105-1,274) p <0,001 9 1,10 (IC95% 1,031-1,188) p < 0,005 10 1,20 (IC95% 1,119-1,294) p <0,001 11 1,25 (IC95% 1,166-1,346) p <0,001 12 1,51 (IC95% 1,404- 1,63) p <0,001

Horas de alta falta		0,947 (IC95% 0,906-0,990) p< 0,017
Días de alta falta		1,604 (IC95% 1,276-2,015) p< 0,001
Meses de alta falta		1,161 (IC95% 1,125-1,2) p <0,001
Lugar de alta falta		1,42 (IC95% 1,366- 1,474) p< 0,001
Porcentaje de faltas del médico a cargo del turno		1,034 (IC95% 1,031-1,036) p <0,001
Médico de alta falta		1,417 (IC95% 1,370- 1,465) p <0,001
Presencia de otros turnos en el mismo día		1,94 (IC95% 1,83- 2,055) p <0,001
Cantidad de turnos solicitados el mismo día		1,859 (IC95% 1,767-1,956) p <0,001
<b>Clima del día del turno</b>		
Temperatura máxima, mínima y media	Máxima Mínima Media	0,99 (IC95% 0,994-0,999) p <0,007 0,99 (IC95% 0,995- 1) p <0,161 0,99 (IC95% 0,995- 0,999) p <0,032
Precipitaciones		1 (IC95% 1- 1) p <0,005
Sensación térmica		0,998 (IC95% .995- 1) p < 0,111
Presión barométrica		0,99 (IC95% 0,993-0,998) p <0,003
<b>Características clínicas y comorbilidades del paciente</b>		
Embarazo		1,02 (IC95% 0,860-1,205) p <0,830
Hipertensión arterial		0,95 (IC95% 0,838-1,085) p <0,476
Diabetes		0,98 (IC95% 0,842-1,143) p < 0,813
Enfermedad oncológica		0,97 (IC95% 0,818-1,164) p <0,792
Enfermedad coronaria		1,07(IC95 % 0,716- 1,613) p <0,728
Enfermedades psiquiátricas		1,05 (IC95% 0,997-1,108) p <0,061
Anemia		0,92 (IC95% 0,869-0,99) p <0,024
BMI		1 (IC95% 0,999-1,003) p <0,173
Obesidad		1,14 (IC95% 1,06- 1,229) p <0,001
Tabaquismo		1,04 (IC95% 0,996- 1,085) p <0,07
Alcoholismo		1,20 (IC95% 0,985-1,465) p <0,069
Abuso de sustancias		1,47 (IC95% 1,222-1,781) p<0,001
Score de comorbilidades de Charlson (por cada		1,03 (IC95% 0,975- 1,101) p < 0,243

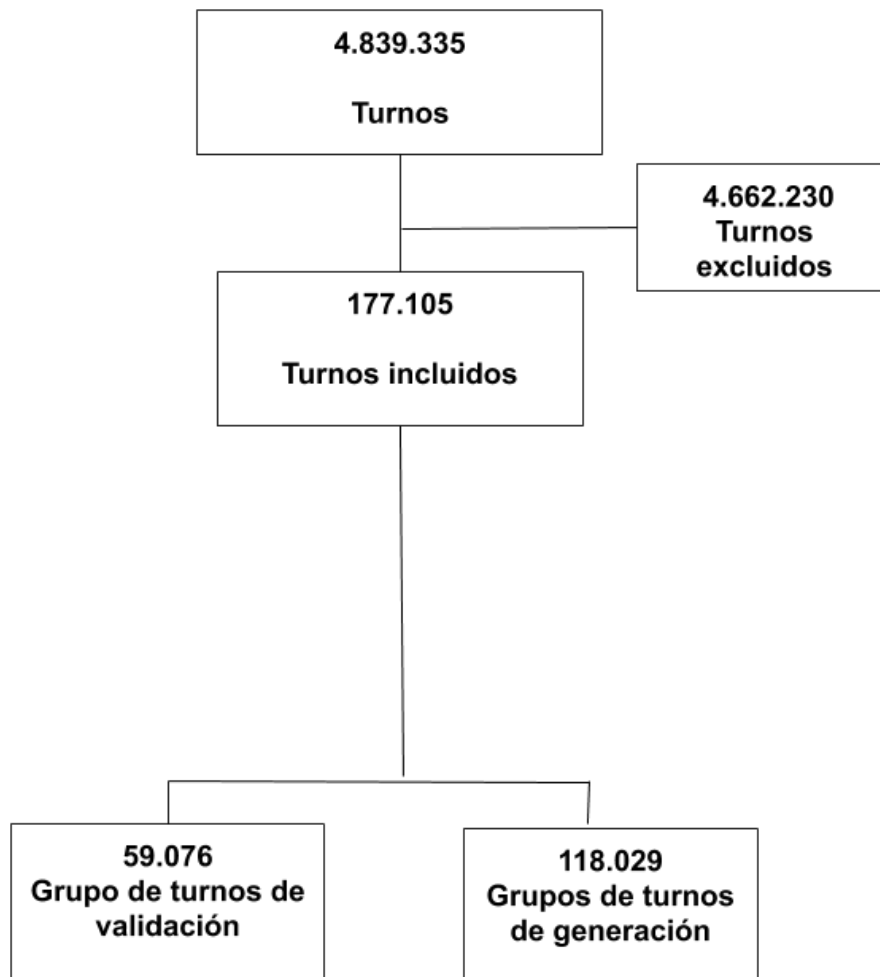
aumento en una unidad del Score)	
Linfoma	1,17 (IC95% 0,761-1,819) p <0,463
Leucemia	0,76 (IC95% 0,43- 1,353) p <0,356
Cáncer (charlson onco)	0,97(IC95% 0,801- 1,193) p <0,826
Enfermedad cerebro vascular	1,13 (IC95% 0,81-1,582) p <0,467
Demencia	0,70 (IC95% 0,278- 1,78) p <0,458
Enfermedad vascular	2,74 (IC95% 0,765- 9,8) p <0,121
EPOC	1,65 (IC95% 1- 2,738) p <0,05
Hemiplegia	2,5 (IC95% 0,958- 6,51) p <0,06
Insuficiencia cardíaca congestiva (icc)	1,79 (IC95% 1,184- 2,711) p <0,006
Insuficiencia hepática	9,37 (IC95% 1,761- 49,84) p <0,009
Insuficiencia renal (irc)	1,56 (IC95% 1,055- 2,328) p <0,026
SIDA	0,98 (IC95% 0,963-0,995) p <0,014
Úlcera gástrica	0,73 (IC95% 0,447-1,215) p <0,233

**Tabla 4.** Sumario de análisis de calibración y discriminación, donde se observan mínimas diferencias entre calibración y discriminación de los diferentes modelos, siendo los algoritmos de XGBM y GBM los que obtuvieron las mayores AUC ROC y las mejores métricas de calibración.

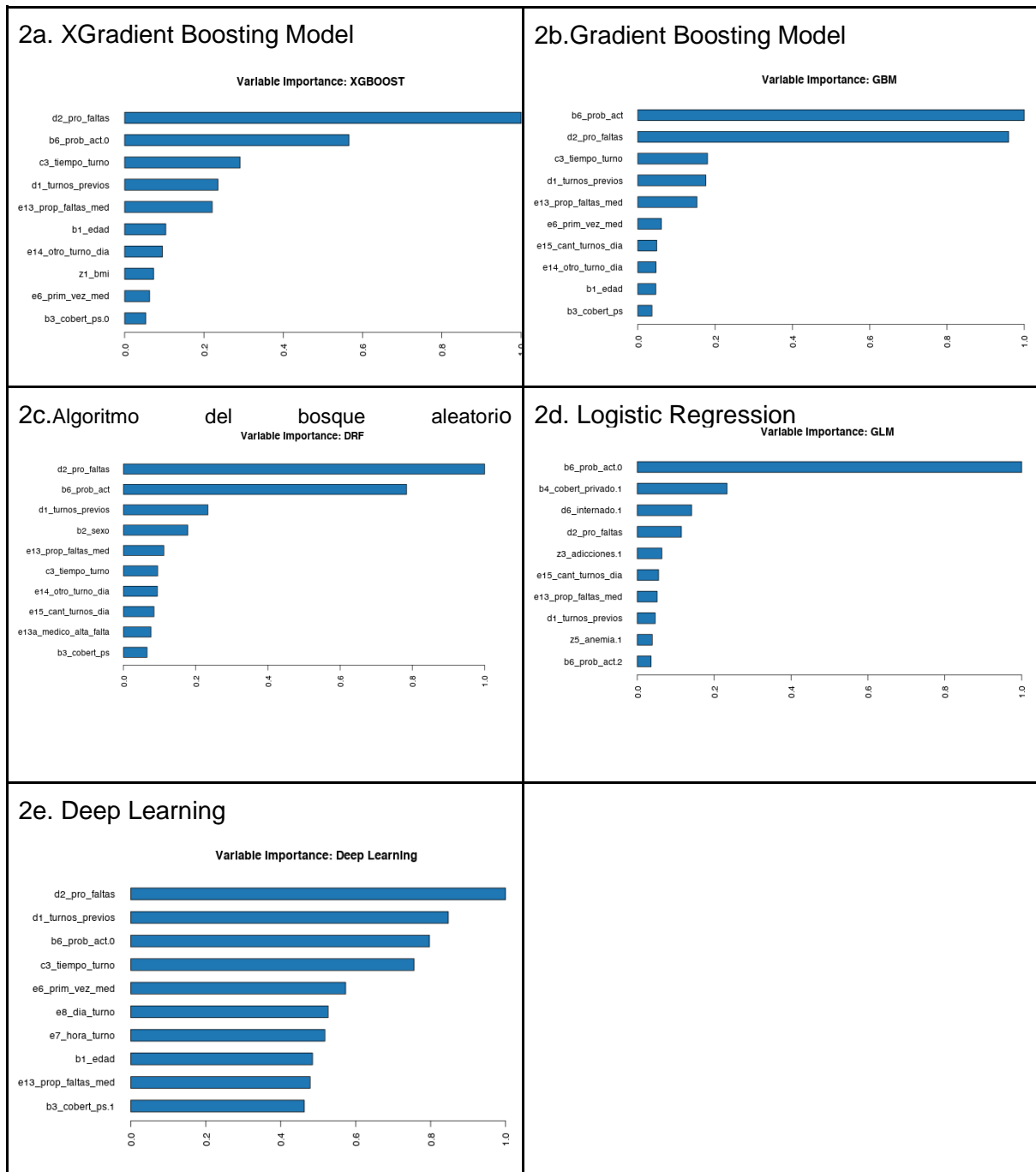
<b>Modelo</b>	<b>Calibración (ordenada al origen)</b>	<b>Calibración (pendiente de calibración)</b>	<b>Discriminación (Área bajo la curva ROC) con el entrenamiento</b>	<b>Discriminación (Área bajo la curva ROC) con la validación</b>
Extreme Gradient Boosting	0,013	0,979	0.734	0,694
Gradient Boosting Model	-0,002	1,021	0,728	0,690
Random Forest	-0,149	1,433	0.678	0,681
Logistic Regression	-0,002	1,021	0,675	0,673
Deep Learning	0,102	0,634	0.647	0,614

## Figuras

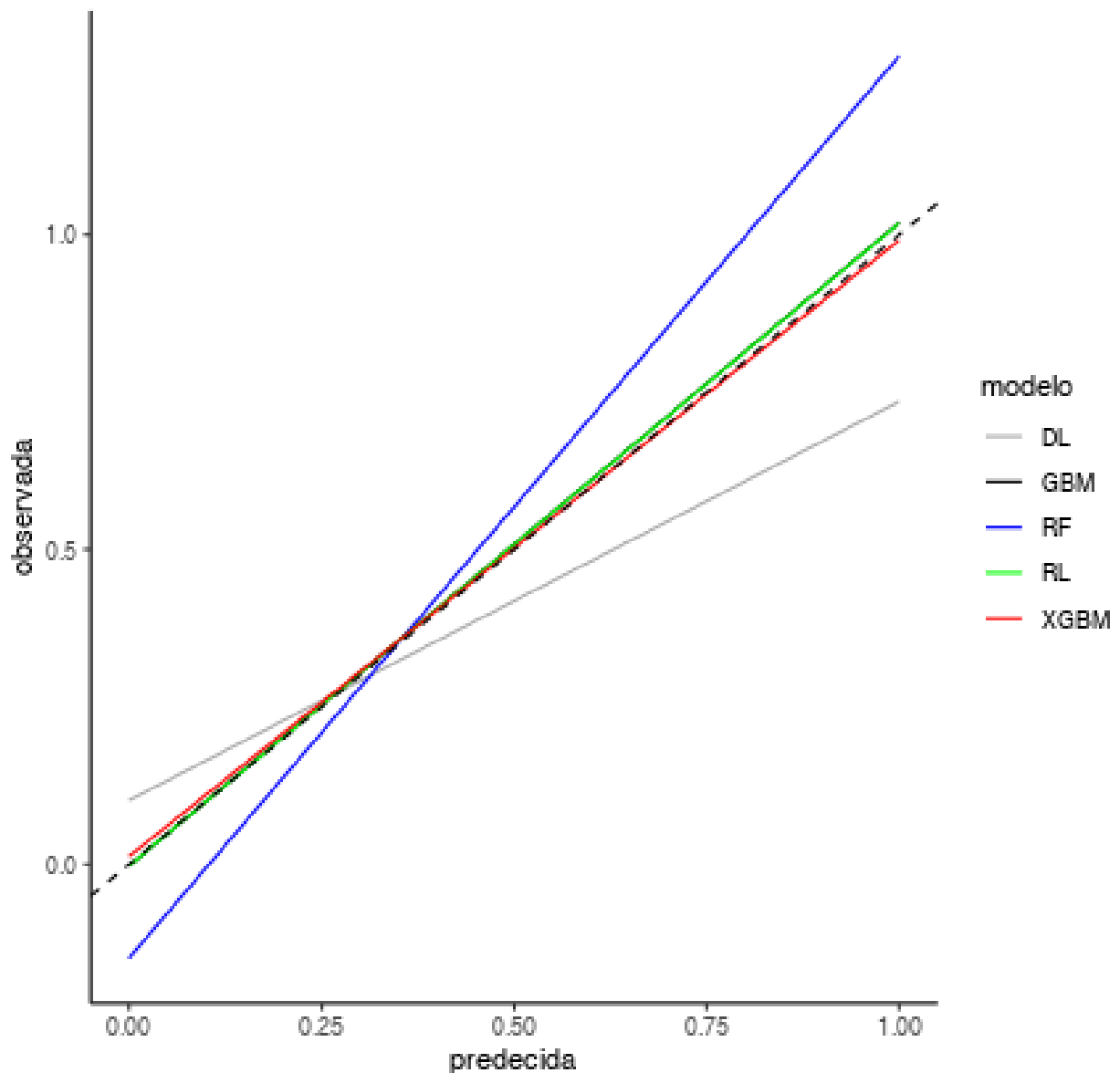
Figura 1. Flujo de turnos incluidos.



**Figura 2. Descripción de la importancia de las variables.** Clasificación de la importancia de las características de los factores en los modelos de aprendizaje automático desarrollados. Se registraron las 10 primeras características en orden de importancia. Gráficos a. Modelo de potenciación extrema del gradiente (XGradient Boosting Model), b. Modelo de potenciación del gradiente (Gradient Boosting Model). c. Algoritmo del bosque aleatorio (Random Forest), d. Regresión Logística (Logistic Regression), e. Algoritmo de aprendizaje profundo (Deep Learning).

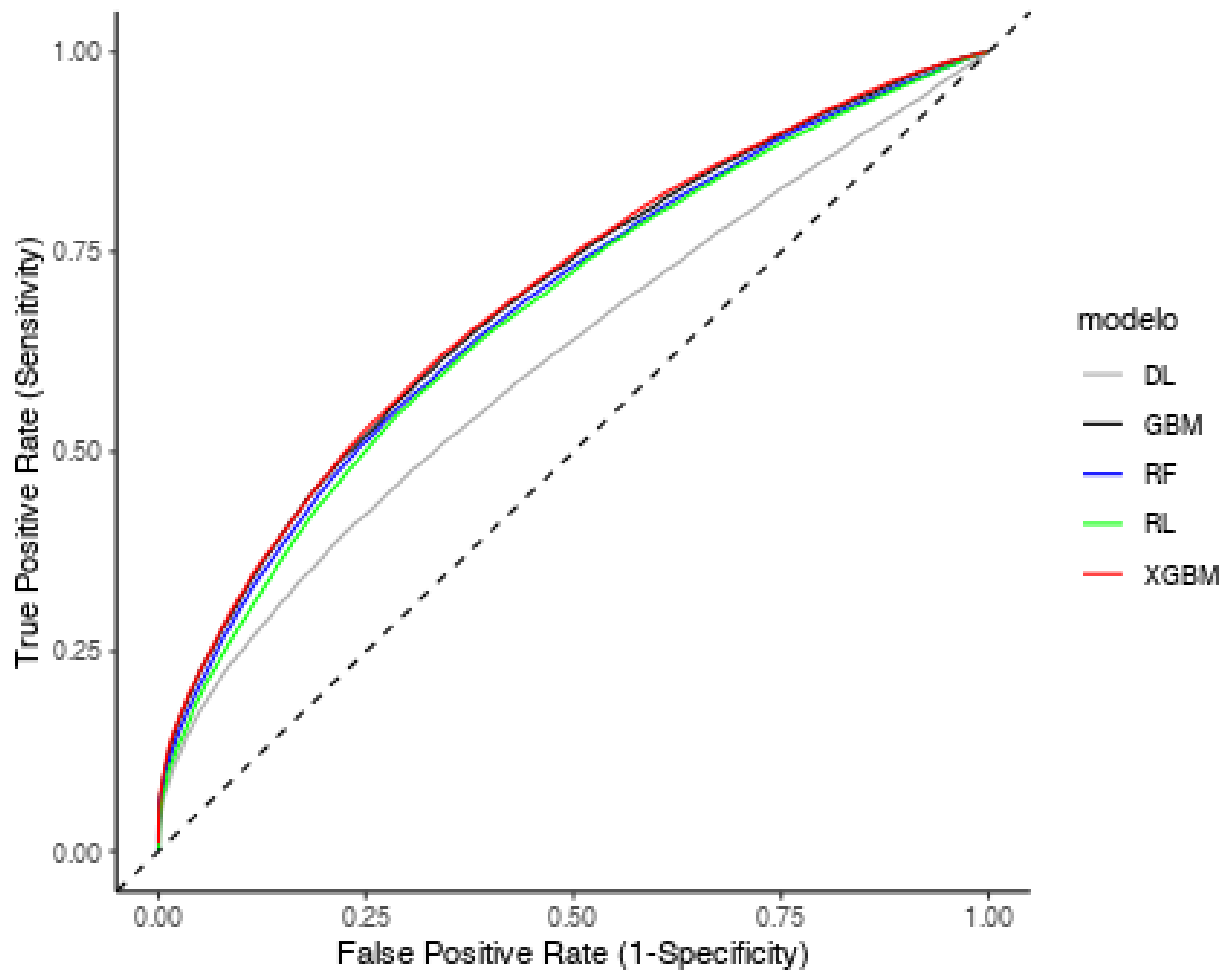


**Figura 3. Análisis de calibración.** La recta de calibración mide el rendimiento de clasificación, cuánto mejor modelo es capaz de clasificar las ausencias, en el caso del mejor modelo la recta de calibración debería tener pendiente 1 y ordenada al origen 0 (línea punteada). Los algoritmos de XGBM y GBM obtuvieron las mejores métricas de calibración.





**Figura. 4. Validación de los modelos predictivos.** Análisis de discriminación: curva ROC de los modelos de ML desarrollados. Mide el rendimiento de clasificación, cuánto modelo es capaz de clasificar las ausencias.



## Referencias bibliográficas

1. Drewek R, Mirea L, David Adelson P. Lead Time to Appointment and No-Show Rates for New and Follow-up Patients in an Ambulatory Clinic. *The Health Care Manager*. 2017. pp. 4–9. doi:10.1097/hcm.000000000000148
2. Huang Y, Hanauer DA. Patient no-show predictive model development using multiple data sources for an effective overbooking approach. *Appl Clin Inform*. 2014;5: 836–860.
3. Menendez ME, Ring D. Factors Associated with Non-Attendance at a Hand Surgery Appointment. *HAND*. 2015. pp. 221–226. doi:10.1007/s11552-014-9685-z
4. Nelson A, Herron D, Rees G, Nachev P. Predicting scheduled hospital attendance with artificial intelligence. *NPJ Digit Med*. 2019;2: 26.
5. Kheirkhah P, Feng Q, Travis LM, Tavakoli-Tabasi S, Sharafkhaneh A. Prevalence, predictors and economic consequences of no-shows. *BMC Health Serv Res*. 2016;16: 13.
6. Rebolledo EAS, De la Cruz Medías R, Silva GB. Nonattendance to medical specialists' appointments and its relation to regional environmental and socioeconomic indicators in the Chilean public health system. *Medwave*. 2014. pp. e6023–e6023. doi:10.5867/medwave.2014.09.6023
7. Determinants and economic cost of patient absenteeism in outpatient departments of the Costa del Sol Health Agency. *Anales del Sistema Sanitario de Navarra*. 2015. doi:10.23938/assn.0072
8. Kosmider S, Shedda S, Jones IT, McLaughlin S, Gibbs P. Predictors of clinic non-attendance: opportunities to improve patient outcomes in colorectal cancer. *Intern Med J*. 2010;40: 757–763.
9. Missed appointments are costing the UK NHS over £600 million per year. *PharmacoEconomics & Outcomes News*. 2009. pp. 11–11. doi:10.2165/00151234-200905850-00035
10. Eid WE, Shehata SF, Cole DA, Doerman KL. PREDICTORS OF NONATTENDANCE AT AN ENDOCRINOLOGY OUTPATIENT CLINIC. *Endocr Pract*. 2016;22: 983–989.
11. Giunta DH, Serena MA. Nonattendance rates of scheduled outpatient appointments in a university general hospital. *The International Journal of Health Planning and Management*. 2019. pp. 1377–1385. doi:10.1002/hpm.2797
12. Briatore A, Giunta D, Baum A, Luna D, Waisman G, de Quiros FGB. Factors associated with nonattendance at clinical medicine scheduled outpatient appointments in a university general hospital. *Patient Preference and Adherence*. 2013. p. 1163. doi:10.2147/ppa.s51841
13. Vidal P, Ramón JM, Goday A, Parri A, Crous X, Trillo L, et al. Lack of adherence to follow-up visits after bariatric surgery: reasons and outcome. *Obes Surg*. 2014;24: 179–183.
14. Hynes L, Byrne M, Dinneen SF, McGuire BE, O'Donnell M, Mc Sharry J. Barriers and facilitators associated with attendance at hospital diabetes clinics among young adults (15-30 years) with type 1 diabetes mellitus: a systematic review. *Pediatr Diabetes*. 2016;17: 509–518.

15. Yu K-H, Beam AL, Kohane IS. Artificial intelligence in healthcare. *Nat Biomed Eng.* 2018;2: 719–731.
16. Miller DD, Brown EW. Artificial Intelligence in Medical Practice: The Question to the Answer? *Am J Med.* 2018;131: 129–133.
17. Jothi N, Rashid NA, Husain W. Data Mining in Healthcare – A Review. *Procedia Computer Science.* 2015. pp. 306–313. doi:10.1016/j.procs.2015.12.145
18. Daghistani T, AlGhamdi H, Alshammari R, AlHazme RH. Predictors of Outpatients' No-Show: Big Data Analytics using Apache Spark. doi:10.21203/rs.3.rs-33216/v1
19. Chen JH, Asch SM. Machine Learning and Prediction in Medicine — Beyond the Peak of Inflated Expectations. *New England Journal of Medicine.* 2017. pp. 2507–2509. doi:10.1056/nejmp1702071
20. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *European Heart Journal.* 2014. pp. 1925–1931. doi:10.1093/eurheartj/ehu207
21. Website. [cited 18 Jan 2021]. Available: Daniel L, Paula O, Alejandro LO, Eduardo R, Federico P, Adrián G, et al. Implementación de una Historia Clínica Electrónica Ambulatoria: El Proyecto Itálica. 6to Simposio de Informática en Salud - 32 JAIIO 2003. 2003. Available: <http://dx.doi.org/>
22. Plazzotta F, Luna D, de Quirós FGB. Sistemas de información en salud: integrando datos clínicos en diferentes escenarios y usuarios. *Revista Peruana de Medicina Experimental y Salud Pública.* 2015. p. 343. doi:10.17843/rpmesp.2015.322.1630
23. SNOMED Home page. In: SNOMED [Internet]. [cited 13 Jun 2020]. Available: <http://www.snomed.org/snomed-ct>
24. Franco M, Mv GB, Otero C, Landoni MC, Benitez S, Borbolla D, et al. Problem Oriented Medical Record: Characterizing the Use of the Problem List at Hospital Italiano De Buenos Aires. *Stud Health Technol Inform.* 2015;216. Available: <https://pubmed.ncbi.nlm.nih.gov/26262179/>
25. Luna D, Franco M, Plaza C, Otero C, Wassermann S, Gambarte ML, et al. Accuracy of an electronic problem list from primary care providers and specialists. *Stud Health Technol Inform.* 2013;192: 417–421.
26. Geocoding Web Content. Google Maps Mashups with Google Mapplets. pp. 59–85. doi:10.1007/978-1-4302-0996-6\_3
27. Get Started | Distance Matrix API | Google Developers. In: Google Developers [Internet]. [cited 13 Jun 2020]. Available: <https://developers.google.com/maps/documentation/distance-matrix/start?hl=es>
28. Servicio Meteorologico Nacional. In: Servicio Meteorologico Nacional [Internet]. [cited 13 Jun 2020]. Available: <https://www.smn.gob.ar/>
29. Devasahay SR, Karpagam S, Ma NL. Predicting appointment misses in hospitals using data analytics. *Mhealth.* 2017;3: 12.
30. Plazzotta F, Otero C, Luna D, de Quiros FG. Natural Language Processing and Inference Rules as Strategies for Updating Problem List in an Electronic Health Record. *Stud Health Technol Inform.* 2013;192. Available:

<https://pubmed.ncbi.nlm.nih.gov/23920937/>

31. Wilson EB. Probable Inference, the Law of Succession, and Statistical Inference. *Journal of the American Statistical Association*. 1927. pp. 209–212. doi:10.1080/01621459.1927.10502953
32. Chaibub Neto E, Pratap A, Perumal TM, Tummalacherla M, Snyder P, Bot BM, et al. Detecting the impact of subject characteristics on machine learning-based diagnostic applications. *NPJ Digit Med*. 2019;2: 99.
33. Turkcan A, Nuti L, DeLaurentis P-C, Tian Z, Daggy J, Zhang L, et al. No-Show Modeling for Adult Ambulatory Clinics. *Handbook of Healthcare Operations Management*. 2013. pp. 251–288. doi:10.1007/978-1-4614-5885-2\_10
34. Kong G, Lin K, Hu Y. Using machine learning methods to predict in-hospital mortality of sepsis patients in the ICU. *BMC Med Inform Decis Mak*. 2020;20: 251.
35. Mayr A, Binder H, Gefeller O, Schmid M. The evolution of boosting algorithms. From machine learning to statistical modelling. *Methods Inf Med*. 2014;53: 419–427.
36. Extreme Gradient Boosting as a Method for Quantitative StructureActivity Relationships. doi:10.1021/acs.jcim.6b00591.s005
37. Chassagnon G, Vakalopoulou M, Paragios N, Revel M-P. Deep learning: definition and perspectives for thoracic imaging. *Eur Radiol*. 2020;30: 2021–2030.
38. Dashtban M, Li W. Deep Learning for Predicting Non-attendance in Hospital Outpatient Appointments. *Proceedings of the 52nd Hawaii International Conference on System Sciences*. 2019. doi:10.24251/hicss.2019.451
39. Montesino-Semper M, Rodrigo Rincón I. Factores determinantes y coste económico del absentismo de pacientes en consultas externas. *An Sist Sanit Navar*. 2016;39: 335–336.
40. Kaplan-Lewis E, Percac-Lima S. No-Show to Primary Care Appointments. *Journal of Primary Care & Community Health*. 2013. pp. 251–255. doi:10.1177/2150131913498513
41. Hernández-García I, Chaure-Pardos A, Moliner-Lahoz J, Prieto-Andrés P, Mareca-Doñate R, Giménez-Júlvez T, et al. Absentismo, y factores asociados, en las citas programadas de una consulta externa de Medicina Preventiva. *Journal of Healthcare Quality Research*. 2018. pp. 82–87. doi:10.1016/j.cali.2017.12.006
42. Do DH, Siegler JE. Diagnoses and other predictors of patient absenteeism in an outpatient neurology clinic. *Neurology: Clinical Practice*. 2018. pp. 318–326. doi:10.1212/cpj.0000000000000488
43. AlMuhaideb S, Alswailem O, Alsubaie N, Ferwana I, Alnajem A. Prediction of hospital no-show appointments through artificial intelligence algorithms. *Annals of Saudi Medicine*. 2019. pp. 373–381. doi:10.5144/0256-4947.2019.373
44. Dantas LF, Fleck JL, Cyrino Oliveira FL, Hamacher S. No-shows in appointment scheduling - a systematic literature review. *Health Policy*. 2018;122: 412–421.
45. Kalb LG, Freedman B, Foster C, Menon D, Landa R, Kishfy L, et al. Determinants of appointment absenteeism at an outpatient pediatric autism clinic. *J Dev Behav Pediatr*. 2012;33: 685–697.

46. Bech M. The economics of non-attendance and the expected effect of charging a fine on non-attendeers. *Health Policy*. 2005;74: 181–191.
47. Cashman SB, Savageau JA, Lemay CA, Ferguson W. Patient Health Status and Appointment Keeping in an Urban Community Health Center. *Journal of Health Care for the Poor and Underserved*. 2004. pp. 474–488. doi:10.1353/hpu.2004.0037
48. Mesa MLJ, Jabalera Mesa ML, Morales Asencio JM, Rivas Ruiz F. Factores determinantes y coste económico del absentismo de pacientes en consultas externas de la Agencia Sanitaria Costa del Sol. *Anales del Sistema Sanitario de Navarra*. 2015. pp. 235–245. doi:10.4321/s1137-66272015000200007
49. DuMontier C, Rindfleisch K, Pruszynski J, Frey JJ 3rd. A multi-method intervention to reduce no-shows in an urban residency clinic. *Fam Med*. 2013;45: 634–641.
50. Baskaran V, Guergachi A, Bali RK, Naguib RNG. Predicting breast screening attendance using machine learning techniques. *IEEE Trans Inf Technol Biomed*. 2011;15: 251–259.

## Anexo 1. Operacionalización de variables Ausentismo Modelos predictivos

<b>Sintaxis</b>	<b>Código</b> (módulo+número de variable). <b>Nombre de variable</b> (nombre de variable en base de datos): Tipo de variable/ <b>Rol en estudio</b> . Si se trata de un campo condicional aclarar condicional a qué variable. Descripción, definición operativa, procedimiento de medición. Validación de medición si aplica. Citas o referencias bibliográficas. Marco temporal. Unidad/Rangos posibles/Codificación (si aplica). Fuente de origen de la información. Unidad de análisis. <b>Captura Primaria/No va en CRF</b>
-----------------	---

### a. Variables de resultado

a3. **Ausentismo** (a3\_ausent): **Variable de Resultado**. Categórica Nominal Dicotómica. Cada turno programado será categorizado de acuerdo a la variable Presencia, Ausencia o Cancelación, como ausentes de acuerdo a si el paciente concurre o no a la consulta. No = 0; Sí = 1. Calculada

### b. Características basales de los pacientes

b1. **Edad** (b1\_edad): Cuantitativa continua, dos decimales. Calculada con la fecha de turno solicitado y la fecha de nacimiento, en años Fecha de turno – Fecha de nacimiento. Calculada.

b1a. **Edad mayor a 65 años** (b1a\_edad\_mayor): Categórica nominal dicotómica. Calculada utilizando la variable b1\_edad según sea mayor o igual a 65 años (definición de anciano [1]). Menor a 65 años = 0, 65 años o mayor = 1. Calculada.

b2. **Sexo** (b2\_sexo): Categórica nominal dicotómica. Según el sexo biológico original del paciente. Femenino = 0, Masculino = 1. Padrón.

b3. **Cobertura Médica Plan de Salud** (b3\_cobert\_ps): Variable explicativa. Categórica Nominal Dicotómica. Turno a cargo del Plan de Salud del Hospital Italiano. Generada con la variable cobertura médica x4. Cualquier otra cobertura médica = 0; PS = 1. Calculada

b4. **Cobertura Médica Privado** (b4\_cobert\_privado): Variable explicativa. Categórica Nominal Dicotómica. Turno a cargo del paciente. Generada con la variable cobertura médica x4. Cualquier otra cobertura médica = 0; Privado = 1. Calculada.

b5. **Cobertura médica agrupada** (b5\_cobert\_agrupada): Variable explicativa. Categórica Nominal. Calculada en base a cuartiles de proporción de faltas observadas dentro de cada cobertura médica durante el periodo del estudio. Cuartilo inferior o baja falta = 1; 2C = 2; 3C = 3; cuartilo superior o alta falta 4C = 4. Calculada.

b6. **Cantidad de problemas activos en la HCE** (b6\_prob\_act): Variable explicativa. Cuantitativa Discreta. Cantidad total de problemas activos en la Historia Clínica Electrónica del paciente a la fecha actual. HCE.

**b7a. Distancia desde el domicilio al lugar de atención (b7a\_dist\_formula):** Variable explicativa. Cuantitativa continua. La distancia entre el domicilio del paciente (A) y el lugar físico de atención del turno solicitado (B). Se calculará la distancia en kilómetros entre ambas coordenadas utilizando la fórmula:  $Distancia = 6371 * \text{ArcoCoseno}(\text{Coseno}(\text{Latitud A}) * \text{Coseno}(\text{Latitud B}) * \text{Coseno}(\text{Longitud B} - \text{Longitud A}) + \text{Seno}(\text{Latitud A}) * \text{Seno}(\text{Latitud B}))$ , utilizando las variables x6, x7, x9 y x10 [2].  
 SELECT [3a distancia lineal reducida]![x7\_lat\_centro] AS LatLugar, [3a distancia lineal reducida]![x6\_long\_centro] AS LongLugar, [3a distancia lineal reducida]![x10\_lat\_domicilio] AS LatPac, [3a distancia lineal reducida]![x9\_long\_domicilio] AS LongPac, ([LatLugar]\*3.141592654)/180 AS RLatLugar, ([LongLugar]\*3.141592654)/180 AS RLongLugar, ([LatPac]\*3.141592654)/180 AS RLatPac, ([LongPac]\*3.141592654)/180 AS RLongPac, (Cos([RLatLugar])\*Cos([RLatPac])\*Cos([RLongPac]-[RLongLugar]))+(Sin([RLatLugar])\*Sin([RLatPac])) AS F, (2\*Atn((Sqr(1-[f]\*[f]))/(1+[f]))) \* 6371 AS b7a\_dist\_formula FROM [3a distancia lineal reducida]; La distancia calculada con esta fórmula es una aproximación ya que estima la distancia lineal entre dos puntos en la superficie curva de la tierra a nivel promedio del mar [3–5]. Expresada en kilómetros. Repositorio de datos.

**b7b. Distancia desde el domicilio al lugar de atención (b7b\_dist\_google):** Variable explicativa. Cuantitativa continua. La distancia entre el domicilio del paciente y el lugar físico de atención del turno solicitado calculada use utilizaron los algoritmos de GoogleMaps para calcular la mínima distancia en kilómetros entre dos direcciones siguiendo el camino transitable más corto en auto. Se utilizó Google Maps Geocoding API para buscar latitud y longitud y Google Maps Distance Matrix API [6] para calcular distancia [7]). Se utilizaron las variables x6, x7, x9 y x10. Expresada en kilómetros. Repositorio de datos.

**b7c. Tiempo desde el domicilio al lugar de atención (b7c\_tiempo\_google):** Variable explicativa. Cuantitativa continua. El tiempo en minutos correspondiente a la distancia entre el domicilio del paciente y el lugar físico de atención del turno solicitado calculada use utilizaron los algoritmos de GoogleMaps para calcular la mínima distancia en kilómetros entre dos direcciones siguiendo el camino transitable más corto en auto. Se utilizó Google Maps Geocoding API para buscar latitud y longitud y Google Maps Distance Matrix API [6] para calcular distancia [8]). Se utilizaron las variables x6, x7, x9 y x10. Expresada en minutos. Repositorio de datos.

### c. Características del proceso de solicitud del turno

c1. **Método utilizado para la solicitud del turno (Tipo de reserva)** (c1\_sol\_turno): Variable explicativa. Categórica Nominal. Opción utilizada por el paciente para obtener el turno solicitado utilizando los distintos componentes del sistema de gestión de turnos del HI: telefónico, en persona, a través del portal de salud o asignado por el médico a través de la HCE. T - Telefónica = 1; P - Personal = 2; W - Web = 3; A - Por HC = 4; C - Ausentismo médico = 5. Sistema de gestión de turnos

c2. **Tipo de turno** (c2\_tipo\_turno): Variable explicativa. Categórica Nominal Dicotómica. Categorización de tipo de turno solicitado. Se excluyeron los sobretornos espontáneos y los sobretornos con nota ya que en estos casos no existe solicitud. Corresponde a Sobretorno o turno convencional. TC - Turno convencional = 1; SP - Sobretorno programado = 2. Sistema de gestión de turnos.

c3. **Tiempo desde la solicitud al turno** (c3\_tiempo\_turno). Variable explicativa. Cuantitativa continua. Calculada con la siguiente fórmula:  $x2$ . Fecha de Turno -  $x1$ . Fecha de solicitud de turno . Expresada en días. Calculada. Turno Solicitado

### d. Historial del paciente

d1. **Cantidad de turnos solicitados** (d1\_turnos\_previos): Variable explicativa. Cuantitativa discreta. Cantidad total de turnos solicitados en el año previo al turno solicitado. Sistema de gestión de turnos. Turno Solicitado.

d2. **Proporción de faltas** (d2\_pro\_faltas): Variable explicativa. Cuantitativa continua. Proporción de faltas total de los turnos solicitados en el año previo a la solicitud del turno. Considerando la totalidad de faltas sobre los turnos presentes más los ausentes. Expresada en proporción. Sistema de gestión de turnos. Turno Solicitado/Paciente.

d6. **Internación al momento del turno** (d6\_internado): Variable explicativa. Categórica nominal dicotómica. Internación en uno de los dos hospitales HIBA Central o San Justo al momento del turno o guardia. No se consideraron los episodios A (ambulatorio), D (internación domiciliaria), F (seguimiento domiciliario), T (geriátrico). No = 0; Sí = 1. Bases de internaciones.

d7. **Internación previa al momento del turno** (d7\_int\_previa): Variable explicativa. Categórica nominal dicotómica. Internación en uno de los dos hospitales HIBA Central o San Justo con alta dentro de los 30 días del turno programado (episodios H). No se consideran las evaluaciones en guardia, demanda espontánea, ni domiciliarias. No = 0; Sí = 1. Bases de internaciones.

d8. **Muerte al momento del turno** (d8\_muerte\_previa): Variable explicativa. Categórica nominal dicotómica. Pacientes con internación previa con tipo de egreso defunción. No = 0; Sí = 1. Sistema de gestión de turnos. Bases de internaciones.

d9. **Cirugía Previa** (d9\_qx\_previa): Variable explicativa. Categórica Nominal Dicotómica. Según tenga paso por alguno de los quirófanos del sistema de salud como lugar físico en los últimos 30 días al turno solicitado. No = 0; Sí = 1. Historia clínica electrónica. Paciente



## **e. Características del turno solicitado**

**e0. Más de un turno en el mismo momento** (e0\_tiempo\_exacto): Variable explicativa. Categórica Nominal. El paciente tiene dos o más turnos asignados en el mismo exacto momento. No = 0; Sí = 1. Sistema de gestión de turnos. Turno Solicitado

**e1. Servicio** (e1\_serv): Variable explicativa. Categórica Nominal. Servicio responsable del turno solicitado. Sin codificar. Sistema de gestión de turnos.

**e2. Servicio agrupado** (e2\_serv\_agrup): Variable explicativa. Categórica Nominal. Calculada en base a la variable Servicio según agrupación en Clínica Médica/atención primaria de la salud (servicio de Clínica Médica), Subespecialidades Clínicas (Alergia, Cardiología, Gastroenterología, Hepatología, Hipertensión arterial, Nefrología, Neumonología, Neurología, Oncología), Especialidades quirúrgicas (Cirugía General, Ginecología, Neurocirugía, Obstetricia, Oftalmología, Otorrinolaringología, Traumatología y ortopedia, Urología). Clínica Médica/atención primaria de la salud = 1; Subespecialidades Clínicas = 2; Especialidades quirúrgicas = 3. Calculada.

**e3. Lugar físico del turno** (e3\_lugar\_turno): Variable explicativa. Categórica Nominal. Lugar donde se realizará la atención correspondiente al turno solicitado. Sistema de gestión de turnos. Turno Solicitado

**e4. Lugar físico del turno agrupado** (e4\_lugar\_turno\_agrup): Variable explicativa. Categórica Nominal. Lugar agrupado donde se realizará la atención correspondiente al turno solicitado en las siguientes categorías: hospital central, hospital San Justo y centros de atención periférica incluyendo consultorios extramuros. Generada utilizando la variable Lugar físico del turno. Central = 1; San Justo = 2; Otro = 3. Sistema de gestión de turnos. Turno Solicitado

**e5. Turno de primera vez Paciente** (e5\_prim\_vez\_pte): Variable explicativa. Categórica Nominal Dicotómica. Pacientes que carecen de turnos previos en el historial de atención del sistema de salud del hospital italiano. No = 0; Sí = 1. Sistema de gestión de turnos. Turno Solicitado

**e6. Turno de primera vez Médico** (e6\_prim\_vez\_med): Variable explicativa. Categórica Nominal Dicotómica. Pacientes que carecen de turnos previos con el médico responsable del turno solicitado en el historial de atención del sistema de salud del hospital italiano. No = 0; Sí = 1. Sistema de gestión de turnos. Turno Solicitado

**e7. Hora del día del turno** (e7\_hora\_turno): Variable explicativa. Categórica Nominal. Corresponde a la hora del día del turno solicitado. Codificado de 7 a 20 hs. Sistema de gestión de turnos. Turno Solicitado

**e8. Día de la semana del turno** (e8\_dia\_turno): Variable explicativa. Categórica Nominal. Corresponde al día de la semana del turno solicitado. Domingo = 1; Lunes = 2; ... Sábado = 7. Sistema de gestión de turnos. Turno Solicitado

**e9. Mes del turno** (e9\_mes\_turno): Variable explicativa. Categórica Nominal. Corresponde al mes del turno solicitado. Codificado de 1 a 12. Sistema de gestión de turnos. Turno Solicitado.

e10. **Horas de alta falta** (e10\_hora\_alta\_falta): Variable explicativa. Categórica Nominal. Se categorizará como Alta falta las horas del día que se encuentren en una frecuencia mayor al 30% de faltas (correspondientes a 7, 18 y 19 horas). Baja falta = 0; Alta falta = 1. Calculada. Turno Solicitado

e11. **Días de alta falta** (e11\_dias\_alta\_falta): Variable explicativa. Categórica Nominal. Se categorizará como Alta falta los días que se encuentren en una frecuencia mayor al 30% de faltas (día 7 - sábado). Baja falta = 0; Alta falta = 1. Calculada. Turno Solicitado

e12. **Meses de alta falta** (e12\_meses\_alta\_falta): Variable explicativa. Categórica Nominal. Se categorizará como Alta falta los meses que se encuentren en una frecuencia mayor a 28% de faltas (12, 7 y 6). Baja falta = 0; Alta falta = 1. Calculada. Turno Solicitado

e12a. **Lugar de alta falta** (e12a\_lugar\_alta\_falta): Variable explicativa. Categórica Nominal. Se categorizará como alta falta los lugares que se encuentren en una frecuencia mayor a 30%. Baja falta = 0; Alta falta = 1. Calculada. Turno Solicitado

e13. **Proporción de faltas del médico a cargo del turno** (e13\_prop\_faltas\_med): Variable explicativa. Cuantitativa continua. Proporción de ausentismo del médico en el año previo al turno solicitado. Expresado en proporción. Sistema de gestión de turnos. Médico

e13a. **Médico de alta falta** (e13a\_medico\_alta\_falta): Variable explicativa. Categórica Nominal. Se categorizará como Alta falta los médicos que tengan una frecuencia de faltas por encima del 30%. Baja falta = 0; Alta falta = 1. Calculada. Turno Solicitado

e14. **Presencia de otros turnos ese mismo día** (e14\_otro\_turno\_dia): Variable explicativa. Categórica Nominal Dicotómica. Presencia de otros turnos en el día del turno solicitado. No = 0; Sí = 1. Sistema de gestión de turnos. Turno Solicitado

e14a. **Presencia de otros turnos a esa misma hora y día** (e14\_misma\_hd): Variable explicativa. Categórica Nominal Dicotómica. Presencia de otros turnos en el mismo horario y día del turno solicitado. No = 0; Sí = 1. Sistema de gestión de turnos. Turno Solicitado

e15. **Cantidad turnos solicitados ese mismo día** (e15\_cant\_turnos\_dia): Variable explicativa. Cuantitativa discreta. Cantidad de turnos solicitados en el día del turno solicitado, incluyendo turnos para realización de estudios programados con turno. Sistema de gestión de turnos. Turno Solicitado.

e16. **Temperatura máxima** (e16\_temp\_max): Variable explicativa. Temperatura máxima registrada en el día del turno solicitado correspondiente a Capital federal. Expresada en grados centígrados. Centro de Información Meteorológica del Servicio Meteorológico Nacional (<http://www.smn.gov.ar/>). Día del Turno Solicitado

e17. **Temperatura mínima** (e17\_temp\_min): Variable explicativa. Temperatura mínima registrada en el día del turno solicitado correspondiente a Capital federal. Expresada en grados centígrados. Centro de Información Meteorológica del Servicio Meteorológico Nacional (<http://www.smn.gov.ar/>). Día del Turno Solicitado.

e18. **Temperatura media** (e18\_temp\_med): Variable explicativa. Temperatura promedio registrada en el día del turno solicitado correspondiente a Capital federal. Expresada en

grados centígrados. Centro de Información Meteorológica del Servicio Meteorológico Nacional (<http://www.smn.gov.ar/>). Día del Turno Solicitado

e19. **Precipitaciones** (e19\_precip): Variable explicativa. Cantidad total de precipitaciones registrada en el día del turno solicitado correspondiente a Capital federal. Expresada en milímetros de lluvia (1 mm de lluvia = 1 litro por metro de superficie). Centro de Información Meteorológica del Servicio Meteorológico Nacional (<http://www.smn.gov.ar/>). Día del Turno Solicitado

e20. **Sensación térmica** (e20\_sens\_termica): Variable explicativa. Presión barométrica registrada en el día del turno solicitado correspondiente a Capital federal. Centro de Información Meteorológica del Servicio Meteorológico Nacional (<http://www.smn.gov.ar/>). Día del Turno Solicitado

e21. **Presión barométrica** (e21\_pres\_barom): Variable explicativa. Presión barométrica registrada en el día del turno solicitado correspondiente a Capital federal. Expresada en hectopascales. Centro de Información Meteorológica del Servicio Meteorológico Nacional (<http://www.smn.gov.ar/>). Día del Turno Solicitado

## f. Características clínicas y comorbilidades del paciente

f1. **Hipertensión arterial** (f1\_hta): Categórica nominal dicotómica. Según conste como problema activo, comorbilidades, diagnósticos, principales o secundarios en la Historia clínica electrónica del paciente según subset definido de términos SNOMED codificados. No=0, Si=1. HCE

f2. **Diabetes mellitus** (f2\_dbt): Categórica nominal dicotómico. Según conste como problema activo, comorbilidades, diagnósticos, principales o secundarios en la Historia clínica electrónica del paciente según subset definido de términos SNOMED codificados. No=0, Si=1. HCE

f3. **Enfermedad oncológica** (f3\_enf\_onco): Categórica nominal dicotómica. Según conste como problema activo, comorbilidades, diagnósticos, principales o secundarios en la Historia clínica electrónica del paciente según subset definido de términos SNOMED codificados. No=0, Si=1. HCE

f4. **Síndrome coronario** (f4\_scoronario): Categórica nominal dicotómica. Según conste como problema activo, comorbilidades, diagnósticos, principales o secundarios en la Historia clínica electrónica del paciente según subset definido de términos SNOMED codificados. No=0, Si=1. HCE

f5. **Enfermedad psiquiátrica** (f5\_enf\_psiq): Categórica nominal dicotómica. Según conste como problema activo, comorbilidades, diagnósticos, principales o secundarios en la Historia clínica electrónica del paciente según subset definido de términos SNOMED codificados. No=0, Si=1. HCE

f6. **Score de comorbilidades de Charlson** (f6\_charlson): Categórica ordinal. Charlson Comorbidity Index, es un sistema de evaluación de la esperanza de vida a los diez años, en dependencia de la edad en que se evalúa, y de las comorbilidades del sujeto. Además de la

edad, consta de 19 ítems, que si están presentes, se ha comprobado que influyen de una forma concreta en la esperanza de vida del sujeto: Infarto de miocardio, insuficiencia cardiaca congestiva, enfermedad vascular periférica, enfermedad cerebrovascular, demencia, EPOC, patología del tejido conectivo, enfermedad ulcerosa, patología hepática ligera, patología hepática moderada/grave, diabetes, diabetes con lesión orgánica, hemiplejía, patología renal moderada o grave, neoplasia, leucemias, linfoma maligno, metástasis óseas, SIDA. <http://www.samiuc.es/index.php/calculadores-medicos/calculadores-de-evaluadores-pronosticos/indice-de-comorbilidad-de-charlson-cci.html>. Todos los componentes provienen de problemas activo, comorbilidades, diagnósticos, principales o secundarios en la Historia clínica electrónica del paciente según subset definido de términos SNOMED codificados. HCE

Insuficiencia cardiaca congestiva (f6\_charlson\_ICC)

enfermedad vascular periférica (f6\_charlson\_enf\_vasc\_p)

enfermedad cerebrovascular (f6\_charlson\_ACVAIT)

demencia (f6\_charlson\_demencia)

EPOC (f6\_charlson\_EPOC)

enfermedad ulcerosa (f6\_charlson\_Ulcera\_gast)

patología hepática ligera/moderada/grave (f6\_charlson\_Insuf\_Hepat)

hemiplejía (f6\_charlson\_hemiplejia)

patología renal moderada o grave (f6\_charlson\_IRC)

neoplasia (f6\_charlson\_cancer)

leucemias (f6\_charlson\_leucemia)

linfoma (f6\_charlson\_linfoma)

SIDA (f6\_charlson\_SIDA)

## **z. Variables agregadas (categóricas, si/no)**

**z1\_BMI BMI $\geq$ 30 BMI, kg (peso)/ m<sup>2</sup> (altura)**

**z1a\_obesidad BMI $\geq$ 30**

**z2\_TBQ**

**z3\_adicciones**

**z4\_alcoholismo**

**z5\_anemia**

**z6\_provincia**

**z7\_localidad**

**z8\_embarazo**

**x. Variables administrativas (necesarias para construir las variables de análisis)**

x1. **ID** (x1\_id): Numérica. Identificador único de personas según aparece en padrón hospitalario, Indicador de clúster. Identificador del paciente (acepta duplicados). Padron/Repositorio de datos.fsexo

x2. **Fecha de turno** (x2\_fecha\_turno): Fecha larga. Fecha del turno solicitado para atención médica por ambulatorio por consultorios externos. Sistema de gestión de turnos.

x3. **Fecha de Nacimiento** (x3\_fecha\_nac): Fecha corta. Fecha corta. Según el DNI. Padrón.

x4. **Cobertura médica** (x4\_cob\_med): Variable explicativa. Categórica Nominal. Cobertura médica del paciente, definida como cualquier prestador o sistema de salud responsable del costo económico del turno solicitado. Sin codificar. Sistema de gestión de turnos.

x6. **Longitud Centro del turno** (x6\_long\_centro): Construcción de otras variables. Cuantitativa continua. Las coordenadas se asignarán utilizando el software Mapas HI (desarrollo del Departamento de Información Hospitalaria del Hospital Italiano de Buenos Aires) usando la variable x5. En radianes. Calculada.

x7. **Latitud Centro del Turno** (x7\_lat\_centro): Construcción de otras variables. Cuantitativa continua. Las coordenadas se asignarán utilizando el software Mapas HI (desarrollo del Departamento de Información Hospitalaria del Hospital Italiano de Buenos Aires) usando la variable x5. En radianes. Calculada.

x9. **Longitud Domicilio Paciente** (x9\_long\_domicilio): Construcción de otras variables. Cuantitativa continua. Las coordenadas se asignarán utilizando el software Mapas HI (desarrollo del Departamento de Información Hospitalaria del Hospital Italiano de Buenos Aires) usando la variable x8. Para los pacientes del interior se utilizarán las coordenadas del centro de la provincia. En radianes. Calculada.

x10. **Latitud Domicilio Paciente** (x10\_lat\_domicilio): Construcción de otras variables. Cuantitativa continua. Las coordenadas se asignarán utilizando el software Mapas HI (desarrollo del Departamento de Información Hospitalaria del Hospital Italiano de Buenos Aires) usando la variable x8. Para los pacientes del interior se utilizarán las coordenadas del centro de la provincia. En radianes. Calculada.

x11. **Fecha de solicitud de turno** (x11\_fecha\_solicitud): Fecha larga. Fecha de solicitud del turno para atención médica por ambulatorio por consultorios externos. Sistema de gestión de turnos.

x12. **ID Médico a cargo del turno solicitado** (x12\_medico\_turno): Categórica Nominal. Médico responsable de la atención del paciente en el turno solicitado. Codificado de 1 en

adelante por orden de frecuencia de consultas en periodo del estudio. Sistema de gestión de turnos. Turno Solicitado

x13. **Aleatorio** (x13\_aleatorio): Número aleatorio para dividir la muestra en cohorte de generación (2/3) y cohorte de validación (1/3). Número aleatorio

x14. **Generación/Validación** (x14\_gen\_val): En base al campo x13\_aleatorio, cohorte de generación = 1; cohorte de validación = 2. Número aleatorio

## Referencias Bibliográficas

1. Tercera edad - Wikipedia, la enciclopedia libre. [cited 24 Jun 2017]. Available: [https://es.wikipedia.org/wiki/Tercera\\_edad](https://es.wikipedia.org/wiki/Tercera_edad)
2. Distance Calculation Algorithms. In: Australian Government. Geoscience Australia [Internet]. c\=AU\;o\=Australian Government\;ou\=Geoscience Australia; 15 May 2014 [cited 19 Jun 2017]. Available: <http://www.ga.gov.au/scientific-topics/positioning-navigation/geodesy/geodetic-techniques/distance-calculation-algorithms>
3. Geographical distance - Wikipedia. [cited 2 Jun 2017]. Available: [https://en.wikipedia.org/wiki/Geographical\\_distance](https://en.wikipedia.org/wiki/Geographical_distance)
4. Geodesics on an ellipsoid - Wikipedia. [cited 2 Jun 2017]. Available: [https://en.wikipedia.org/wiki/Geodesics\\_on\\_an\\_ellipsoid](https://en.wikipedia.org/wiki/Geodesics_on_an_ellipsoid)
5. Distance Calculation Algorithms. In: Australia Government. Geoscience Australia [Internet]. 15 May 2014 [cited 19 Jun 2017]. Available: <http://www.ga.gov.au/scientific-topics/positioning-navigation/geodesy/geodetic-techniques/distance-calculation-algorithms>
6. Primeros pasos | Google Maps Geocoding API | Google Developers. In: Google Developers [Internet]. [cited 13 Jun 2017]. Available: <https://developers.google.com/maps/documentation/geocoding/start?hl=es>
7. Google Maps Distance Matrix API | Google Developers. In: Google Developers [Internet]. [cited 13 Jun 2017]. Available: <https://developers.google.com/maps/documentation/distance-matrix/?hl=es>
8. Google Maps Distance Matrix API | Google Developers. In: Google Developers [Internet]. [cited 13 Jun 2017]. Available: <https://developers.google.com/maps/documentation/distance-matrix/?hl=es>