



UNIVERSIDAD
NACIONAL
DE COLOMBIA

**Modelo de aprendizaje automático para predecir el riesgo de evento adverso
en analgésicos opioides aplicando datos de secuenciación de
última generación (NGS) en una población de pacientes colombianos**

Eliana del Pilar Garzón Venegas

Universidad Nacional de Colombia
Facultad de ingeniería, Departamento de Sistemas e Industrial
Bogotá, Colombia
2022

Modelo de aprendizaje automático para predecir el riesgo de evento adverso en analgésicos opioides aplicando datos de secuenciación de última generación (NGS) en una población de pacientes colombianos

Eliana del Pilar Garzón Venegas

Tesis presentada como requisito parcial para optar al título de:
Magister en Bioinformática

Director:

Ph.D., Fabio Ancízar Aristizábal Gutiérrez

Codirector:

Ph.D., Luis Fernando Niño Vásquez

Línea de Investigación: Farmacogenómica

Grupo de Investigación:

Laboratorio de Investigación en Sistemas Inteligentes – LISI

Universidad Nacional de Colombia

Facultad de ingeniería, Departamento de Sistemas e Industrial

Bogotá, Colombia

2022

“Nunca consideres el estudio como una obligación, sino como una oportunidad para penetrar en el bello y maravilloso mundo del saber.” Albert Einstein.

Dedicatoria

Dedico este trabajo a mis padres, mis hermanos, mi esposo e hijo, quienes han sido un apoyo muy importante en este proceso. A mi hijo Esteban una dedicatoria especial porque se ha convertido en una inspiración y motor de vida.

Agradecimientos

Agradezco a Dios y a la vida por brindarme salud, estabilidad y fortaleza.

A mis padres y hermanos, por su apoyo incondicional.

A mi esposo e hijo, por su comprensión, apoyo, inspiración y tiempo para cumplir mis proyectos académicos.

A mi director y codirector de Tesis, el profesor Fabio Ariztizabal y Luis Fernando Niño, quienes siempre me orientaron en la realización de esta Tesis.

A la Dra Caludia Serrano y el laboratorio Genetix SAS, quienes me brindaron su apoyo con la data empleada en la realización de esta tesis.

A Oscar Escobar, Bioinformático del Laboratorio Genetix, por su apoyo, colaboración y guía.

A David Bustos Cortes, estudiante de pregrado de ingeniería de Sistemas, por sus enseñanzas y tiempo, fue un apoyo muy valioso en la elaboración de los modelos de aprendizaje automático de esta tesis.

A la Universidad Nacional de Colombia y docentes, por brindarme las bases conceptuales y académicas para la realización de este proyecto.

A los integrantes del grupo de investigación LISI (Laboratorio de Investigación en Sistemas Inteligentes) de la Universidad Nacional de Colombia por su guía, ayuda y aportes.

Modelo de aprendizaje automático para predecir el riesgo de evento adverso en analgésicos opioides aplicando datos de secuenciación de última generación (NGS) en una población de pacientes colombianos

Resumen

OBJETIVO: Desarrollar un modelo basado en aprendizaje automático para la predicción del riesgo de evento adverso, a partir del genotipo en farmacogenes asociados con la farmacocinética y farmacodinámica de analgésicos opioides, a partir de datos de secuenciación de última generación (NGS), en una cohorte de pacientes colombianos.

MÉTODOS: Se desarrolló un *pipeline* de anotación de variantes y anotación funcional en 39 genes asociados a la farmacocinética y farmacodinamia de 17 analgésicos opioides de uso común en Colombia a partir de 2080 VCF de exomas provenientes de la secuenciación de nueva generación (NGS). Se realizó un modelo de aprendizaje automático para la clasificación del riesgo efecto adverso utilizando bosques aleatorios, naive Bayes y perceptrón multicapa.

RESULTADOS: El *pipeline* de anotación de variantes y anotación funcional en 39 genes asociados a la farmacocinética y farmacodinamia de 17 analgésicos opioides de uso común en Colombia detecta 9 variantes de riesgo. Los algoritmos de aprendizaje automático se entrenan y evalúan a partir de un dataset compuesto por 1900 variantes genéticas con *score* de riesgo desde 0.5 a 1.5, 50 variantes genéticas con un *score* de 1.5 a 2 y 9 variantes genéticas con *score* de 2.5 a 4, las cuales están asociadas con efecto adverso, se analizan empleando bosques aleatorios, naive bayes y perceptrón multicapa, obteniendo resultados deficientes en la clasificación de la clase 2 y clase 3 debido al desbalance de datos en estas clase, con lo cual se realiza un enriquecimiento del dataset a partir de variantes de la base de datos PharmGKB, ampliando la data de la clase 3. Se evalúan varias combinaciones de clases por medio de bosques aleatorios y perceptrón multicapa, obteniendo los mejores resultados de clasificación, considerado únicamente dos clases, clase 1, suprimiendo valores de *score* de 0.5 y 1 y clase 2 compuesta por la unión de datos con *score* de 2.5 a 5.

CONCLUSIONES: Fue posible desarrollar modelos computacionales de clasificación del riesgo de efecto adverso. Sin embargo, se concluye que el desequilibrio en las clases, genera problemas de clasificación, lo que resulta en una reducción significativa de la sensibilidad y la precisión de los modelos de aprendizaje automático. Para poder generar relaciones entre variantes genéticas y su asociación con la presentación de efecto adverso se hace indispensable considerar variables de diversas ómicas que le den un peso importante a la asociación, acompañado de la información clínica y de seguimiento de los pacientes.

Palabras clave: Analgesicos opioides, evento adverso, farmacogenética, aprendizaje automático.

Machine learning model to predict the risk of adverse event in opioid analgesics using next-generation sequencing (NGS) data in a population of Colombian patients

Abstract

GOAL: To develop a model based on machine learning for prediction of the risk of adverse event, from the genotype in associated pharmacogenes with the pharmacokinetics and pharmacodynamics of opioid analgesics, based on data from next-generation sequencing (NGS), in a cohort of colombian patients.

METHODS: A variant annotation and functional annotation pipeline was developed for 39 genes associated with the pharmacokinetics and pharmacodynamics of 17 commonly used opioid analgesics in Colombia from 2080 VCF exomes from nextgeneration sequencing (NGS). A machine learning model was performed for risk adverse effect classification using random forest, naive Bayes, and multilayer perceptron.

RESULTS: The variant annotation and functional annotation pipeline in 39 genes associated with the pharmacokinetics and pharmacodynamics of 17 commonly used opioid analgesics in Colombia detected 9 risk variants. Machine learning algorithms are trained and evaluated from a dataset composed of 1,900 genetic variants with risk scores from 0.5 to 1.5, 50 genetic variants with scores from 1.5 to 2, and 9 genetic variants with scores from 2.5 to 4. , which are associated with an adverse effect, will be analyzed using occasional forests, naive bayes and multilayer perceptron, obtaining poor results in the classification of class 2 and class 3 due to the imbalance of data in these classes, with which an enrichment is performed. of the dataset from variants of the PharmGKB database, expanding the data of class 3. Various combinations of classes are evaluated by means of random forests and multilayer perceptron, obtaining the best classification results, considering only two classes, class 1, suppressing score values of 0.5 and 1 and class 2 composed of the union of data with scores from 2.5 to 5.

CONCLUSIONS: It was possible to develop computational models for classifying the risk of adverse effects. However, it is concluded that the imbalance in the classes generates classification problems, which results in a significant reduction in the sensitivity and accuracy of the machine learning models. In order to generate relationships between genetic variants and their association with the presentation of adverse effects, it is essential to consider variables of various omics that give significant weight to the association, accompanied by clinical information and patient follow-up.

Keywords: Opioid analgesics, adverse event, pharmacogenetics, machine learning.

Esta tesis de maestría se sustentó el Miercoles 2 de Noviembre de 2022 a las 07:00 am, y fue evaluada por los siguientes jurados:

Clara Isabel Bermudez Santana (Msc - Ph.D.)
Universidad Nacional de Colombia

Juan David Garcia Arteaga (Msc - Ph.D.)
Universidad Nacional de Colombia

Tabla de contenidos

Resumen	6
Abstract	7
Lista de Figuras	9
Lista de tablas	10
Lista de abreviaturas	11
Introducción	1
1. Identificación del problema	2
1.1 Justificación	2
1.2 Objetivo general	4
1.3 Objetivos específicos	4
2. Fármacos opioides	4
3. Efecto adverso en fármacos opioides	8
4. Genes asociados a la farmacodinamia y farmacocinética de los opioides	9
5. Modelos de aprendizaje automático en farmacogenómica	12
6. Metodología	16
7. Resultados	23
8. Discusión	39
9. Conclusiones	42
10. Perspectivas	43
11. Anexos	44
12. Referencias	50

Lista de Figuras

Figura 1. Efectos adversos de los opioides.

Figura 2. Ejemplo de la estructura de un árbol de decisión.

Figura 3. Ejemplo de la estructura de un bosque aleatorio.

Figura 4. Ejemplo de la estructura de una red bayesiana.

Figura 5. Diagrama Pipeline DRAGEN.

Figura 6. Diagrama del pipeline de anotación.

Figura 7. Pipeline para la anotación funcional.

Figura 8. Metodología aprendizaje automático.

Figura 9. Importancia de las variables de entrada usando la librería H2O.

Figura 10. Importancia variables de entrada usando la librería Scikit-learn .

Figura 11. Escala de importancia de las variables de entrada obtenidas usando la librería Scikit-learn.

Figura 12. Cuadrícula de hiperparámetros.

Figura 13. Parámetros de la validación cruzada.

Figura 14. Mejor combinación de hiperparámetros.

Lista de tablas

Tabla 1. Clasificación de los fármacos opioides según su acción en el receptor

Tabla 2. Asociación de variantes genéticas en genes asociados a la farmacocinética y farmacodinamia de los analgésicos opioides con efecto adverso

Tabla 3. Panel de 39 genes asociados a la farmacodinamia y farmacocinética de analgésicos opioides.

Tabla 4. Analgésicos opioides.

Tabla 5. Asignación del score (puntaje) de acuerdo con el riesgo **Tabla 6.**

Tabla de escala de importancia de las variables de entrada usando la librería H2O

Tabla 7. Conjuntos (*datasets*) de entrenamiento y evaluación.

Tabla 8. Parámetros de los modelos de aprendizaje automático

Tabla 9. Resultados - matrices de confusión

Tabla 10. Dataset de entrenamiento y evaluación.

Tabla 11. Resultados - matriz de confusión

Tabla 12. Análisis sin considerar *score* de 0.5 en Clase 1.

Tabla 13. Dataset de entrenamiento y evaluación.

Tabla 14. Análisis eliminando los datos con *scores* de 0.5 y 1 en la clase 1.

Tabla 15. Dataset de entrenamiento y evaluación.

Tabla 16. Análisis eliminando los datos con *scores* de 0.5 y 1 en la Clase 1 y haciendo la unión de la clase 2 y la clase 3, en una sola clase.

Tabla 17. Análisis con solo dos clases.

Lista de abreviaturas

Abreviatura **Término**

NGS Secuenciación de nueva generación

ADME Absorción, Distribución, Metabolismo y Excreción

DMET Metabolismo de fármacos Enzimas y transportadores

G-6-M Glucurónido 6 de morfina G-3-M

Glucurónido 3 de morfina

Introducción

Los opioides son una clase de medicamentos que se unen a los receptores de opiáceos, para tratar el dolor moderado a intenso (1). La respuesta al tratamiento con analgésicos opioides varía de forma importante de unos individuos a otros, variaciones genéticas en las enzimas metabolizadoras de fármacos o los transportadores pueden contribuir a una respuesta variable del tratamiento con opioides y a la presentación de evento adverso al medicamento (2-4).

El uso de opioides no está exento de riesgos. En algunos casos, el uso prolongado puede provocar adicción y pueden estar asociados con múltiples eventos adversos como depresión respiratoria, estreñimiento, náuseas, confusión, somnolencia e incluso la muerte (5-6).

La farmacogenómica es prometedora para aplicaciones como la optimización de medicamentos para pacientes basada en el genotipo. Se centra en la identificación de variantes genéticas que se correlacionan con los efectos de los fármacos en poblaciones, cohortes y pacientes individuales (7).

La secuenciación del exoma por medio de técnicas de nueva generación (NGS) ha proporcionado información relevante de variantes individuales para la genotipificación en farmacogenómica, con un enfoque en los genes que codifican ADME (absorción, distribución, metabolismo, excreción) y DMET (enzimas metabolizantes y transportadores), dando lugar a la medicina de precisión, que tiene como objetivo personalizar la atención médica con decisiones y tratamientos adaptados a cada individuo optimizando el tratamiento y disminuyendo el riesgo de presentación de evento adverso (10-12).

Los métodos de aprendizaje automático, incluidos los métodos de aprendizaje profundo, se han utilizado en farmacogenómica, permitiendo la predicción mecánica de la respuesta a los medicamentos. La genética y la genómica han sido esenciales para la evolución de la medicina de precisión y se adaptan bien a la tecnología de aprendizaje automático, contribuyendo de manera significativa en la comprensión de la enfermedad y el tratamiento (13).

1. Identificación del problema

Los analgésicos opioides se formulan con frecuencia para controlar el dolor agudo y crónico, pero las diferencias genotípicas individuales dificultan predecir y establecer el equilibrio entre la efectividad de los fármacos y la potencial aparición del evento adverso. Para la población colombiana no se ha realizado un estudio que explore la relación del genotipo en farmacogenes con el riesgo de presentar evento adverso basados con la farmacocinética y farmacodinamia en analgésicos opioides por medio de aprendizaje automático. Teniendo en cuenta esto, se vuelven relevantes los resultados que puedan obtenerse de esta investigación ya que brindará información importante al paciente y a los médicos en la selección de intervenciones que logren un efecto terapéutico deseable sin toxicidad, en la población colombiana.

1.1 Justificación

Los analgésicos opioides son un grupo de fármacos utilizados para controlar el dolor crónico, generalmente usados en condiciones médicas como trauma, cirugías, cáncer. La eficacia de los medicamentos opioides está relacionada con su capacidad de actuar sobre los receptores opioides en el sistema nervioso central y en otros tejidos (1). Existen cuatro receptores opioides en el ser humano: mu, delta y kappa, que median los principales efectos farmacológicos de los opioides (2,3).

La respuesta al tratamiento con analgésicos opioides varía de forma importante de unos individuos a otros en función de factores relacionados con el paciente, la enfermedad y su tratamiento (4). Pueden ocasionar eventos adversos como depresión respiratoria, sedación, náuseas, vómitos, estreñimiento, deterioro de la cognición, adicción y, en algunos casos, la muerte. Lo cual hace indispensable optimizar el manejo del dolor, para garantizar la máxima eficacia mientras se previene la toxicidad (5,6).

Los polimorfismos genéticos han sido etiquetados como algunos de los principales responsables de la variación de la respuesta de opioides. El objetivo de la farmacogenética como herramienta clínica en la terapia con opioides es optimizar el alivio del dolor y prevenir los efectos adversos (7).

La farmacogenética es un componente importante de la medicina personalizada la cual ha logrado transformar la predicción de enfermedades, el pronóstico y un diagnóstico individual que permite pensar en intervenciones terapéuticas con alta posibilidad de ser eficaces y previniendo o minimizando riesgos de efectos adversos (8).

La secuenciación del exoma a partir de la secuenciación de última generación (NGS) cubre las regiones codificantes del genoma humano, que permite identificar variantes causales de una enfermedad (9), actualmente tiene una aplicación en medicina personalizada, ya que permite la identificación de variantes genéticas que puede ayudar

a predecir la farmacocinética y farmacodinámica de los medicamentos, lo que puede ayudar a los médicos en la selección de opciones farmacológicas que logren un efecto terapéutico deseable y minimizando riesgos toxicológicos (10,11).

Para los medicamentos con un rango terapéutico estrecho, un pequeño cambio en la actividad funcional puede resultar en un efecto fisiológico drástico que puede conducir a un evento adverso. Las reacciones adversas asociadas a estos medicamentos son una de las principales causas de morbilidad y mortalidad que se pueden evitar fácilmente (12). Con los avances de la medicina personalizada y de precisión y la secuenciación del exoma se pueden predecir los riesgos de enfermedades individuales.

Por otro lado, la inteligencia artificial es un componente clave de la gran revolución basada en los datos que está transformando la medicina; como parte de esto, los algoritmos de aprendizaje automático han aumentado las capacidades predictivas para el riesgo de enfermedades complejas (13). Lo que ayudará a incorporar características genéticas en la atención médica personalizada futura. Los enfoques de aprendizaje automático se han empleado ampliamente en la predicción de riesgos para muchas enfermedades, entre ellas: cáncer; Alzheimer, enfermedad cardiovascular y diabetes (14, 15, 16, 17, 18). En un estudio de diabetes tipo 2 se usó un algoritmo de bosque aleatorio (en inglés, *random forest*) para predecir el riesgo de esta enfermedad (19). La implementación de un bosque aleatorio es muy útil como método de aprendizaje automático para el modelado de riesgo de enfermedades complejas (20, 21, 22).

En un estudio de asociación de los genotipos humanos en los genes TRPV1 y TRPA1 con la sensibilidad a los estímulos de calor y la capsaicina aplicada tópicamente por medio de aprendizaje automático, por medio de algoritmo de bosque aleatorio, se identifica que la complejidad del genotipo es un factor relevante en la modulación de la respuesta individual. Los métodos de aprendizaje de máquina proporcionaron resultados biológicamente importantes, superando a los análisis estadísticos clásicos de genotipo versus asociación de fenotipo (23).

Recientemente, en un estudio muy relevante en Estados Unidos en la estratificación del riesgo de opioides relacionado con la adicción, cuyo objetivo principal era la predicción del riesgo, y el objetivo secundario era la estratificación del riesgo. Se emplearon varias estrategias de aprendizaje automático para identificar cuál método ofrecía la predicción y estratificación de riesgo más precisa. Con los resultados obtenidos, los investigadores concluyeron que su enfoque superó a otros métodos para identificar el riesgo. Los modelos de aprendizaje automático que utilizan datos administrativos, es decir, datos provenientes de historia clínica a lo largo del tiempo, pueden ser una herramienta valiosa y factible para identificar de manera precisa y eficiente a las personas con alto riesgo de sobredosis de opioides (24).

De acuerdo al informe sobre el manejo del abuso de sustancias de la Organización Mundial de la Salud (OMS), se estima que en el mundo mueren cada año 69 000 personas por sobredosis de opioides. Se estima que hay 15 millones de personas en el mundo que presentan adicción a los opioides (25). Para Colombia, en el Estudio Nacional de Consumo de Sustancias Psicoactivas del 2019 se señala que el 1,07 % de

las personas declararon haber usado algún analgésico opioide sin prescripción médica, al menos una vez en su vida. Sobre las intoxicaciones y muertes asociadas a consumo de opioides con y sin usos médicos, los reportes al sistema SIVIGILA indican un comportamiento estable en los últimos tres años (26). Sin embargo, no existe una estadística de evento adverso asociado al consumo de opioides para la población colombiana.

Mediante la información de secuenciación de última generación (NGS) de una cohorte representativa de pacientes colombianos se desea realizar una asociación de los genotipos en farmacogenes involucrados con la farmacocinética y farmacodinámica de medicamentos opioides y su asociación con el riesgo de presentar evento adverso por medio de aprendizaje automático, la información obtenida proporcionarán una mayor comprensión de estas asociaciones genéticas con el evento adverso y mejorará el uso de la farmacogenética como herramienta de medicina personalizada en la población colombiana.

1.2 Objetivo general

Desarrollar un modelo basado en aprendizaje automático para la predicción del riesgo de evento adverso, a partir del genotipo en farmacogenes asociados con la farmacocinética y farmacodinámica de analgésicos opioides, a partir de datos de secuenciación de última generación (NGS), en una cohorte de pacientes colombianos.

1.3 Objetivos específicos

- Realizar la clasificación funcional *in silico* de SNPs a partir del análisis terciario de la secuenciación de última generación (NGS) en una cohorte de pacientes colombianos.
- Diseñar un modelo de aprendizaje automático para predecir el efecto adverso de analgésicos opioides, a partir de datos de secuenciación de última generación (NGS) en una cohorte de pacientes colombianos.
- Implementar un modelo computacional que permita determinar la correlación del genotipo con el riesgo de presentación de evento adverso en farmacogenes asociados con la respuesta a analgésicos opioides.

2. Fármacos opioides

Los fármacos opioides derivan del opio y se utilizan para el tratamiento del dolor crónico de moderado a intenso. Su acción analgésica se da por la acción combinada de varias estructuras situadas desde la corteza cerebral hasta la médula espinal.

Ejercen su acción a través de la fijación a receptores específicos, que se encuentran en el sistema nervioso, así como en otros tejidos. Existen cuatro receptores opioides en el ser humano: mu (MOR), delta (DOR) y kappa (KOR), que median los principales efectos farmacológicos de los opioides (27,28,29).

2.1 Clasificación de fármacos opioides

Los opioides pueden clasificarse de acuerdo a su afinidad y eficacia en la unión al receptor en agonistas puros, agonistas-antagonistas mixtos, agonistas parciales y antagonistas puros (30), ver **Tabla 1**.

Tabla 1. Clasificación de los fármacos opioides según su acción en el receptor.

AGONISTAS	AGONISTAS-ANTAGONISTAS	AGONISTAS PARCIALES
Dolor moderado	Pentazocina	Buprenorfina
Codeína	Butorfanol	
Dihidrocodeína	Nalbufina	
Dextropropoxifeno	Meptazinol	
Tramadol		
Dolor intenso, vida media corta		
Morfina		
Oxicodona		
Meperidina		
Diacetilmorfina (heroína)		
Fentanilo		
Dolor intenso, vida media larga		
Metadona		

2.1.1 Agonistas puros

Se comportan como agonistas exclusivos sobre receptores mu (μ), mostrando la máxima actividad intrínseca. Pertenecen a este grupo la morfina, la heroína, la petidina, la metadona, el fentanilo y sus derivados (30,31,32). Se clasifican en:

- Opioides mayores: Se caracterizan por no tener techo analgésico, el grado de analgesia crece ilimitadamente con la dosis.
- Opioides menores: Se caracterizan por poseer techo analgésico.

2.1.2 Agonistas / antagonistas mixtos

Son opioides agonistas sobre receptores kappa (κ) pero que ejercen un efecto agonista parcial o antagonista sobre receptores mu (μ). El efecto final que este tipo de fármacos produzcan dependerá del grado de afinidad sobre cada uno de los tipos de receptores. Pertenecen a este grupo la pentazocina, el butorfanol o la nalorfina (30,31,32).

2.1.3 Agonistas parciales

Actúan sobre receptores mu (μ) con actividad intrínseca inferior a la de los agonistas puros. En presencia de un agonista puro, pueden comportarse también como antagonista. Pertenecen a este grupo la buprenorfina (30,31,32).

2.1.4 Antagonistas puros

Tienen afinidad por todos los receptores opiáceos, pero carecen de actividad intrínseca. Compiten con los agonistas por el receptor, desplazándolos y sirven para evitar sus efectos secundarios. Pertenecen a este grupo la naloxona y la naltrexona (30,31,32).

2.1 Farmacocinética

Los opiáceos se absorben en el tracto gastrointestinal, las mucosas nasal y oral, en el pulmón y por inyección subcutánea o intramuscular. También, pueden ser administrados por vía intravenosa, espinal y transdérmica. La concentración del fármaco en los tejidos depende de su grado de perfusión, siendo máxima en el hígado, el riñón, el pulmón y el cerebro. El paso de la barrera hematoencefálica es limitado y atraviesan la barrera placentaria alcanzando la circulación fetal. Se metabolizan principalmente en el hígado y la eliminación tiene lugar, sobre todo, por vía renal (33).

2.2 Farmacodinamia

Los opioides disminuyen la percepción del estímulo doloroso determinando un estado con ausencia de dolor o un dolor muy leve. La valoración del efecto analgésico debe hacerse periódicamente teniendo en cuenta la edad y estado del paciente. La concentración analgésica mínima efectiva (MEAC): es el mínimo nivel plasmático de opioide necesario a partir del cual se puede atenuar el dolor en un paciente. Cuando se alcanza, pequeñas variaciones en la concentración reducen significativamente la

intensidad del dolor, delimitando la denominada "ventana analgésica". La MEAC no es un valor fijo para cada opioide, varía en cada individuo por lo que la analgesia no se relaciona siempre de manera directa con la concentración plasmática de opioide, ni con su farmacocinética (34).

2.3 Principios activos

El principio activo es el componente del medicamento que posee actividad biológica, a continuación, se describen los principios activos de los opioides.

2.4.1 Codeína

La codeína es un agonista de los receptores μ , clasificado como un opioide débil, con pocos efectos secundarios. La codeína, para el efecto analgésico, se comporta como un profármaco que requiere de la presencia de enzimas de Citocromo P-450 para su conversión a morfina que es el principio activo responsable de tal efecto. (35). La variabilidad interindividual de esta enzima en el ser humano compromete la efectividad de la codeína (36).

2.4.2 Tramadol

El Tramadol es un medicamento con múltiples mecanismos de acción, considerado un opioide débil. Es útil en el control del dolor moderado. Es un agonista puro, no selectivo sobre los receptores opioides μ , δ y κ , con mayor afinidad por los receptores μ . Es metabolizado en el hígado por las isoenzimas CYP3A4 y CYP2D6 (40).

2.4.3 Morfina

La Morfina es un alcaloide. Actualmente sigue siendo el analgésico más eficaz para aliviar los dolores agudos. Ejerce su acción sobre receptores opioides específicos del sistema nervioso central y del tubo digestivo (40).

2.4.4 Hidromorfona

La Hidromorfona es un derivado de la morfina con leves cambios en su estructura, por la presencia de un grupo cetónico en el carbono 6. Es un agonista puro, cuyos efectos analgésicos son mediados por receptores μ . Tiene baja afinidad sobre receptores κ . Se utiliza en dolor pos-operatorio y en el tratamiento del dolor crónico por cáncer (40).

2.4.5 Oxycodona

La Oxycodona es un opioide semisintético, agonista puro de receptores μ y κ . Se utiliza para tratar dolor de moderado a severo (49).

2.4.6 Fentanilo

El Fentanilo es de origen sintético, con una estructura química que le provee una elevada liposolubilidad. Se comporta como un agonista de receptores μ y está disponible para la administración por las vías oral transmucosa, parenteral, transdérmica y epidural. Se utiliza para dolores intensos, dolor post operatorio. Puede

usarse para tratar a pacientes que sufren de dolor crónico y presentan tolerancia física a otros opioides (35).

2.4.7 Buprenorfina

La Buprenorfina es un opioide semisintético derivado de la tebaína, con alta liposolubilidad. Su peso molecular y configuración estructural aumentan su penetración tisular por vía transmucosa y transdérmica. Es un analgésico agonista/antagonista opiáceo que se une a los receptores μ y k (40). Se usan para aliviar el dolor fuerte en personas que necesiten medicamento para el dolor todo el tiempo durante un período prolongado.

2.4.8 Naloxona

La Naloxona es un antagonista puro de los opioides. Se utiliza para revertir efectos adversos de los opioides que comprometan la vida del paciente, en especial, depresión respiratoria (36).

La farmacocinética y farmacodinamia de los principios activos de los analgésicos opioides se especifican en el anexo 1.

3. Efecto adverso en fármacos opioides

Los opiáceos, como todos los fármacos, pueden producir diversos efectos adversos (ver **Figura 1**), los cuales se pueden clasificar en:

Efectos gastrointestinales: Son los más frecuentes, interfieren con la funcionalidad, calidad de vida y adherencia terapéutica. Se producen en todo el sistema gastrointestinal por diferentes vías, entre los efectos más comunes, náuseas, vómito y estreñimiento (60).

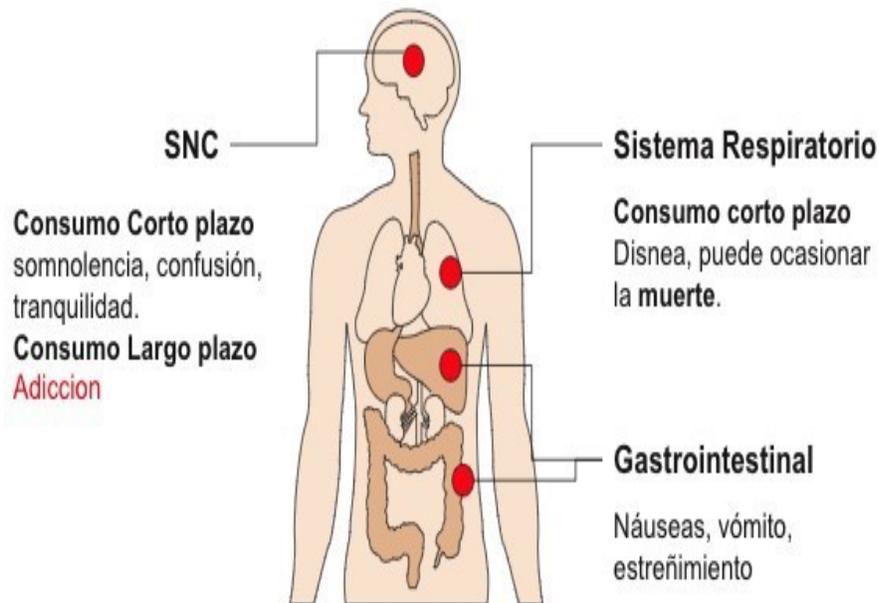
Efectos respiratorios: Los opioides producen diferentes efectos sobre el sistema respiratorio, como la depresión respiratoria, disnea, insuficiencia cardíaca, edema pulmonar cardiogénico y rigidez torácica (60).

Efectos en el sistema nervioso central y el funcionamiento cognitivo: Se asocian con el estado de ánimo, concentración, memoria, actividad psicomotora, pueden poner en riesgo el funcionamiento de los pacientes en cuanto a sus actividades físicas, trabajo, actividades sociales y ocio. Dentro de los efectos más comunes se encuentra la Sedación, delirio y alucinación (60).

Efecto de tolerancia, dependencia y adicción: En el uso prolongado de opioides en el control del dolor crónico maligno y no maligno se generan mecanismos adaptativos, tanto a nivel del receptor como celular, sináptico y en niveles de conexión

responsables del desarrollo de los fenómenos de tolerancia, dependencia y adicción. Además influyen los factores genéticos y psicosociales (61).

Figura 1. Efectos adversos de los opioides.



4. Genes asociados a la farmacodinamia y farmacocinética de los opioides

Las dos características fenotípicas clínicamente relevantes más importantes en los medicamentos opioides son: la dosis del opioide requerida para la analgesia y la incidencia de efectos adversos relacionados con los opioides.

La farmacocinética y farmacodinamia de los fármacos opioides pueden verse alterados por los genes que codifican las proteínas o enzimas involucradas en estas vías. La absorción, distribución, metabolismo y eliminación son procesos farmacocinéticos. Entre estos procesos, el metabolismo de un fármaco es importante, ya que juega un papel importante en la determinación de las concentraciones plasmáticas del fármaco original, así como en la producción de metabolitos activos o inactivos (66). Con respecto a los opioides, la mayoría de esta clase de fármacos son metabolizados por el citocromo P450 hepático. Se ha encontrado que los polimorfismos de un solo nucleótido en los genes que codifican estas enzimas CYP450 están asociados con variaciones en las concentraciones plasmáticas del fármaco original y sus metabolitos. De manera similar, los SNP en los genes que codifican las enzimas metabolizadoras de otros opioides (UGT2B7 para morfina,

CYP2D6 para tramadol y CYP3A4 para fentanilo y sufentanilo) pueden afectar las concentraciones plasmáticas de los respectivos fármacos originales y sus metabolitos, lo que influye en su relación dosisrespuesta (73).

La actividad farmacodinámica de los opioides está determinada tanto por el receptor del fármaco como por el transportador del fármaco. Se ha encontrado que los polimorfismos de un solo nucleótido en los genes que codifican estos receptores y transportadores de fármacos están asociados con variaciones en la eficacia de los opioides, es decir, en la analgesia (73).

Las variaciones genéticas se han asociado con la respuesta a los medicamentos. Se realiza una búsqueda bibliográfica en publicaciones sobre genes asociados a la respuesta de medicamentos analgésicos opioides y su posible asociación con un efecto adverso, descritos en la Tabla 2.

Tabla 2. Asociación de variantes genéticas en genes asociados a la farmacocinética y farmacodinamia de los analgésicos opioides con efecto adverso.

Gen	Ubicación	Medicamento	Variante	Descripción	Efecto	Referencia
ABCC3	17q21.33	Morfina	rs4793665	Metabolismo	Aclaramiento de la morfina	62
ABCB1	7q21.12	morfina	rs1128503	Metabolismo	Disminución de la concentración de la morfina	63
		fentanilo	rs1128503	Dosis	Incremento de dosis	64
		oxicodona	rs1045642	Dosis	Ajuste dosis	65
ADRB2	5q32	fentanilo	rs1042718	toxicidad	Hipotension	66
ALDH5A1	6p22.3	Metadona	rs2760118	Eficacia	Reducción de la actividad enzimática	67
ANKK1	11q23.2	Metadona	rs1800497	Dosis	Inceremento de dosis	68
BDNF	11p14.1	Metadona	rs2030324	Dosis	Ajuste dosis	69
CACNA1E	1q25.3	Fentanilo	rs3845446	Dosis	Ajuste dosis	70
CCL11	17q12	Metadona	rs1129844	Asociado a toxicidad	Insomnio	71
COMT	22q11.21	Morfina	rs4680	Eficacia y toxicidad	Nauseas, vomito, estreñimiento, delirio, somnolencia	72
		Oxicodona	rs4680	Eficacia	Incremento de la respuesta	73
		Metadona	rs933271	Dosis, eficacia	Incremento dosis	74

Gen	Ubicación	Medicamento	Variante	Descripción	Efecto	Referencia
CYP1A2	15q24.1	Metadona	CYP1A2*1B	Metabolismo, dosis	tiempo de aclaramiento medicamento	75
CYP2B6	19q13.2	Tramadol	rs3745274	Metabolismo	Ajuste dosis	76
		Metadona	rs2279343	Dosis	Incremento de dosis	77
CYP2C19	10q23.33	Metadona	CYP2C19*1, CYP2C19*2, CYP2C19*3	Metabolismo	variacion en la farmacocinetica	78
CYP2C9	10q23.33	Metadona	CYP2C9*1, CYP2C9*2, CYP2C9*3	Metabolismo, dosis, eficacia	variacion en la farmacocinetica	79
CYP2D6	22q13.2	Codeina	CYP2D6*2, CYP2D6*2xN	Metabolismo	Efectos adversos, asociados a la muerte	80
		Morfina		Metabolismo	Ajuste dosis	81
		Trmadol	CYP2D6*4, CYP2D6*6, CYP2D6*9	Metabolismo	Ajuste dosis	82
		Oxicodona		Metabolismo	Ajuste dosis	83
CYP3A4	7q22.1	Fentanilo	CYP3A4*36	Dosis, eficacia	Ajuste dosis	84
		Oxicodona	CYP3A4*1, CYP3A4*22	Metabolismo	Disminucion metabolismo	85
CYP3A5	7q22.1	Fentanilo	CYP3A5*1, CYP3A5*3	Metabolismo	Incremento de la concentracion	86
		oxicodona	CYP3A5*1, CYP3A5*3	Metabolismo	Ajuste dosis	87
DBH	9q34.2	naltexona	rs1611115	toxicidad, eficacia	Síndrome de abstinencia	88
DRD2	11q23.2	Metadona	rs6275	eficacia, dosis, toxicidad	Incremento de dosis	89
FAAH	1p33	Morfina	rs324420	toxicidad	prurito, estreñimiento, delirio, insuficiencia respiratoria, nauseas, vomito	90
GAD1	2q31.1	Metadona	rs3762555	Dosis	Requerimineto de mayor dosis	91
GNB3	12p13.31	Metadona	rs5443	toxicidad	Aumento frceuencia cardiaca	92
HCN1	5p12	tramadol	rs79983226	Metabolismo	menor conversión de tramadol a O-desmetiltramadol (M1)	93
IL1B	2q14.1	Morfina	rs1143634	Dosis,eficacia	Aumento de dosis	94
KCNJ6	21q22.13	Fentanilo	rs2835859	Dosis, eficacia	Ajuste dosis	95
		Metadona	rs2070995		Ajuste dosis	96
		Morfina	rs6517442	Eficacia	Incremento de la respuesta	97
MYD88	3p22.2	fentanilo	rs6853	toxicidad	Disfuncion cognitiva	98
NGF	1p13.2	Metadona	rs2239622	Dosis	Aumento de dosis	99
NTRK2	9q21.33	Metadona	rs4358872	Dosis	Aumento de dosis	100
		buprenorfina	rs678849	Eficacia	Incremento de la respuesta	101
		metadona	rs581111	Eficacia	Disminucion de la respuesta del farmaco	102
OPRD1	1p35.3	Oxicodona	rs529520	Dosis	Incremento de dosis	103
		Morfina	rs1051660	Dosis, eficacia	Aumento de dosis	104
OPRK1	8q11.23	Metadona	rs3802281	Dosis	Aumento de dosis, farmacodependencia	105
OPRM1	6q25.2	oxicodona	rs1799971	Eficacia	Disminucion analgesia	106
		Morfina	rs1799971	Toxicidad	prurito, estreñimiento, delirio, insuficiencia respiratoria, nauseas, vomito	107
P2RX7	12q24.31	fentanilo	rs1718125	Dosis	Aumento de dosis	108
RHBDF2	17q25.1	fentanilo, oxicodona, metadona	rs12948783	Eficacia	menor alivio del dolor	109
SLC22A1	6q25.3	tramadol	rs12208357	Metabolismo	Disminucion farmacocinética	110
SLC6A4	17q11.2	Morfina	rs1042173	Dosis	Aumento de dosis	111
SULT1A1	16p11.2	tapentadol	rs28374453	Metabolismo	menor sulfatación de tapentadol	112
SULT1A3	16p11.2	Morfina	rs1293732453	Metabolismo	Ajuste dosis	113
TXNRD2	22q11.21	Metadona	rs933271	Eficacia	menor respuesta al tratamiento	114
UGT1A1	2q37.1	Morfina	UGT1A1*28	toxicidad	Apnea	115
UGT2B7	4q13.2	fentanilo	rs7668282	Eficacia	Analgesia	116
		buprenorfina	rs7662029	Metabolismo	Disminucion de la concentracion	117
		Codeina	rs7439366	Dosis	Ajuste dosis	118

5. Modelos de aprendizaje automático en farmacogenómica

Los modelos de aprendizaje automático se pueden definir como el estudio de algoritmos informáticos que mejoran automáticamente a través de la experiencia (119). Estos métodos permiten obtener resultados satisfactorios en términos de predicción y clasificación.

En farmacogenética, los modelos de aprendizaje automático han mostrado un rendimiento satisfactorio en la predicción de la respuesta a fármacos en varios campos, como el cáncer, la depresión y la terapia anticoagulante. El modelo de bosques aleatorios es la técnica de aprendizaje automático supervisado, aplicada con mayor frecuencia (120).

5.1 Árbol de decisión

Un árbol de decisión es un algoritmo de aprendizaje supervisado no paramétrico, que se utiliza tanto para tareas de clasificación como de regresión. Tiene una estructura de árbol jerárquico, que consta de un nodo raíz, ramas, nodos internos y nodos hoja. El árbol de decisión comienza con un nodo raíz, que no tiene ramas entrantes. Las ramas salientes del nodo raíz alimentan los nodos internos, también conocidos como nodos de decisión (Figura 2). En función de las características disponibles, ambos tipos de nodos realizan evaluaciones para formar subconjuntos homogéneos, que se indican mediante nodos hoja o nodos terminales. Los nodos hoja representan todos los resultados posibles dentro del conjunto de datos (121).

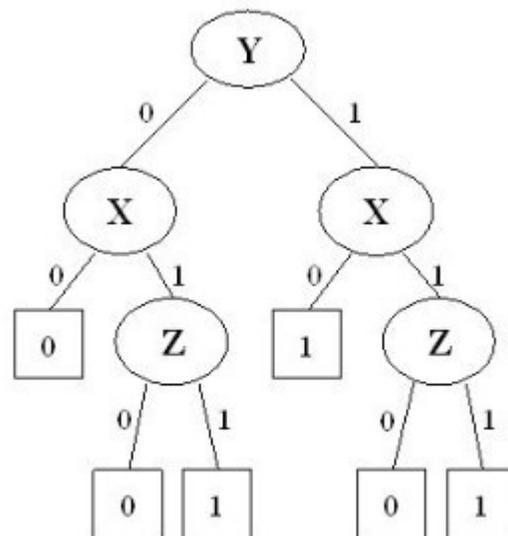


Figura 2. Ejemplo de la estructura de un árbol de decisión.

5.1.2 Parámetros del algoritmo

Durante la fase de entrenamiento el modelo de árboles de decisión permite ajustar los parámetros teniendo en cuenta los datos recibidos. Los hiper-parámetros son parámetros no inferidos por el modelo durante el entrenamiento. El algoritmo va a crear un árbol de decisión tan complejo como sea necesario para adaptarse a los datos de entrenamiento, lo que puede ocasionar sobre-entrenamiento del modelo. Para evitar esto deben aplicarse varios parámetros en la construcción del modelo (121), como se detallan a continuación.

- **Profundidad del árbol (max_depth):** Se define como la profundidad máxima de cada árbol. Un número de profundidad demasiado alto puede provocar un sobreajuste o una variación alta. Un número de profundidad demasiado bajo puede provocar un ajuste insuficiente o aumento del sesgo.

- **Número mínimo de muestras requerido para dividir un nodo (min_samples_split) :** Si un nodo no contiene el número de muestras mínimo especificado, no podrá ser dividido, limitando el sobre-entrenamiento. Un número muy bajo de "división mínima de muestras" puede provocar un ajuste excesivo o una varianza alta. Un valor alto de "división mínima de muestras" puede dar lugar a un ajuste insuficiente o aumento del sesgo.

- **Número mínimo de muestras por hoja (min_samples_left):** Es el número mínimo de muestras necesarias por hoja. Un número de "hojas de muestras mínimas" que es demasiado bajo puede provocar un sobreajuste o una variación alta. Un número de "hoja de muestras mínimas" que es demasiado alto puede conducir a un ajuste insuficiente o aumento del sesgo.

- **Número máximo de hojas (max_leaf_nodes):** Este parámetro permite limitar el número de hojas a crear en el árbol, lo que lleva a modelos más generalizables. Un número de "características máximas" alto se aproxima a la búsqueda exhaustiva de un árbol de decisión sin un conjunto de parámetros. Un número de profundidad máxima bajo selecciona menos características aleatorias para tener en cuenta al elegir la siguiente división en el árbol.

- **Mínimo de impureza para dividir un nodo (min_impurity_split):** si la impureza de un nodo no alcanza el valor mínimo, no se permitirá su división. Un número bajo de "disminución mínima de impurezas" hará que el árbol divida los datos en divisiones "teóricamente" válidas; Un número alto de "disminución mínima de impurezas" hará que el proceso de entrenamiento sea demasiado estricto para elegir la próxima división, lo que generará un alto sesgo.

5.1.3 Medidas para evitar sobre-entrenamiento

El sobreajuste ocurre cuando un modelo estadístico se ajusta mucho a los datos de entrenamiento, lo cual hace que el modelo no responda de manera confiable a valores futuros. Si el sobre-entrenamiento o la complejidad del modelo dan como resultado un ajuste excesivo, entonces una forma de prevenir esto es detener el proceso de entrenamiento antes, también conocido como detención anticipada (*early stopping*), o reducir la complejidad del modelo eliminando las entradas menos relevantes (122). Detectar el sobreajuste es útil, pero es necesario aplicar técnicas que pueden evitarlo:

- Detención temprana (*early stopping*):** este método busca detener el entrenamiento antes de que el modelo comience a aprender el ruido dentro del modelo.

- Entrenamiento con más datos:** expandir el conjunto de entrenamiento para incluir más datos puede aumentar la precisión del modelo al brindar más oportunidades para analizar la relación dominante entre las variables de entrada y salida.

- Aumento de datos:** considerar el aumento del conjunto de datos con datos relevantes.

- Selección de características:** La selección de características es el proceso de identificar las más importantes dentro de los datos de entrenamiento y eliminar las irrelevantes o redundantes.

- Regularización:** si se produce un sobreajuste cuando un modelo es demasiado complejo, debe reducirse el número de parámetros. La regularización aplica una "penalización" a los parámetros con los coeficientes más grandes, lo que posteriormente limita la cantidad de varianza en el modelo.

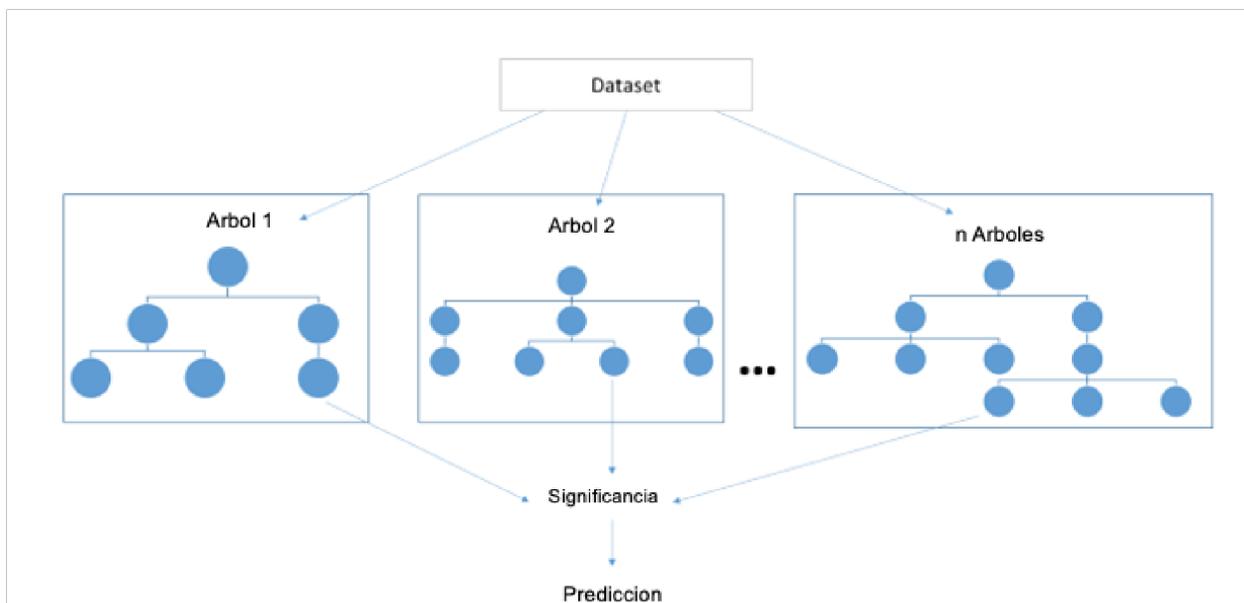
5.2 Bosque aleatorio

Los Bosques Aleatorios (*Random Forest*) están fundamentados en un método de clasificación basado en árboles de decisión. Estos métodos se caracterizan por segmentar el espacio de predicción en un número de regiones. Este algoritmo utiliza muchos modelos de árbol de clasificación o regresión para subconjuntos aleatorios de los datos de entrada y utiliza el resultado combinado (*forest*) para la predicción (ver Figura 3). Es una técnica de clasificación con un poderoso desempeño al usar *bagging* y selección de variables aleatorias para la construcción de cada árbol (123).

Una de las principales ventajas de los bosques aleatorios es su interpretabilidad. Esto se expresa en la posibilidad de calcular fácilmente la importancia de las variables en el modelo. Finalmente, el algoritmo posee una serie de hiperparámetros que pueden ser ajustados:

- Profundidad del árbol (max depth): profundidad máxima de cada árbol.
- Estimadores (n estimators): Se define como la cantidad de árboles.
- Cantidad de variables por árbol (max features): En cada partición (*split*) del árbol, se puede fijar la cantidad de variables a comparar.

Figura 3. Ejemplo de la estructura de un bosque aleatorio.



5.3 Redes bayesianas

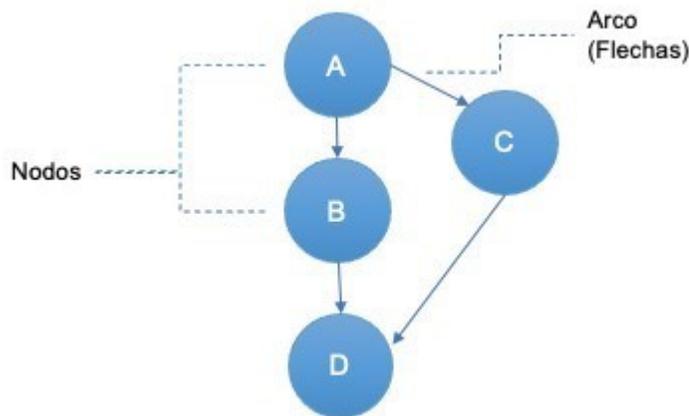
Una red bayesiana es un grafo acíclico dirigido donde los nodos representan a las variables y las flechas representan relaciones de dependencia directa entre las variables. Para cada nodo existe una distribución de probabilidad local que depende del estado de su nodo principal. Una red bayesiana consta de un modelo estructural, cualitativo, que brinda una representación visual de las interacciones entre las variables y de un conjunto de distribuciones locales de probabilidad, cuantitativo, que permite efectuar inferencia probabilística y además mide el impacto de una variable o conjuntos de variables sobre otras. Ambas partes determinan una distribución de probabilidad conjunta de las variables de un determinado problema que se puede expresar de manera compacta

mediante el uso extensivo de la independencia condicional, como se muestra en la siguiente ecuación (124,125,126):

$$P(X_1, X_2, \dots, X_n) = \prod P(X_i | Pa(X_i))$$

$Pa(X_i)$: representa al conjunto de nodos principales de X_i (nodos cuyas flechas apuntan a X_i).

Figura 4. Ejemplo de la estructura de una red bayesiana.



6. Metodología

El presente es un estudio observacional retrospectivo y analítico utilizando la información de pacientes colombianos secuenciados por secuenciación de última generación (NGS) del Laboratorio Genetix SAS.

6.1 Población del estudio

La población de pacientes correspondió a 2080 pacientes que ingresaron al laboratorio Genetix SAS, para el análisis de exoma completo (27.000 genes), para el diagnóstico de diversas enfermedades genéticas. Los pacientes participantes accedieron a compartir su información genética para la elaboración de proyectos de investigación por medio de la aceptación del consentimiento informado.

6.2 Criterios de inclusión

- Ser colombiano
- Pacientes que previamente por medio de consentimiento informado autorizan el uso de sus datos genéticos para investigación.

6.3 Criterios de exclusión

- Tener una nacionalidad diferente a la colombiana
- Pacientes que no autoricen el uso de sus datos genéticos

6.4 Análisis Bioinformático

Los datos fueron obtenidos a través de la secuenciación del exoma completo a partir de secuenciación de nueva generación (NGS), la cual se llevó a cabo a partir del kit TruSeq DNA Exome y secuenciado en la plataforma MiSeq de illumina, y se fundamenta en los siguientes procesos: • **Calidad de los datos**

Se consideran las muestras que cumplan con la calidad de análisis, se tuvieron en cuenta:

1. FASTQ que cumplen la calidad del software FASTQC
2. FASTQ con pesos superiores a 8 Gb
3. Porcentaje (%Q30): Este valor es un promedio de toda la longitud de lectura. Se tienen en cuenta valores superiores de 90% de Q30.
4. Cobertura media superior a 90 x

• Análisis bioinformático primario

El análisis de datos primario consiste en la detección y el análisis de datos sin procesar (análisis de señales), con el objetivo de generar lecturas de secuencias legibles (llamadas de base) y calificar la calidad de la base. Las salidas de este análisis primario son archivos FASTQ, generados por el secuenciador Myseq de Illumina.

• Análisis bioinformático secundario

El análisis de datos secundario incluye la alineación de las lecturas con el genoma de referencia humano y el llamado de variantes, se utiliza el genoma de referencia GRCh37-hg19, y se procesa por medio del *pipeline* generado por DRAGUEN (ver Figura 5).



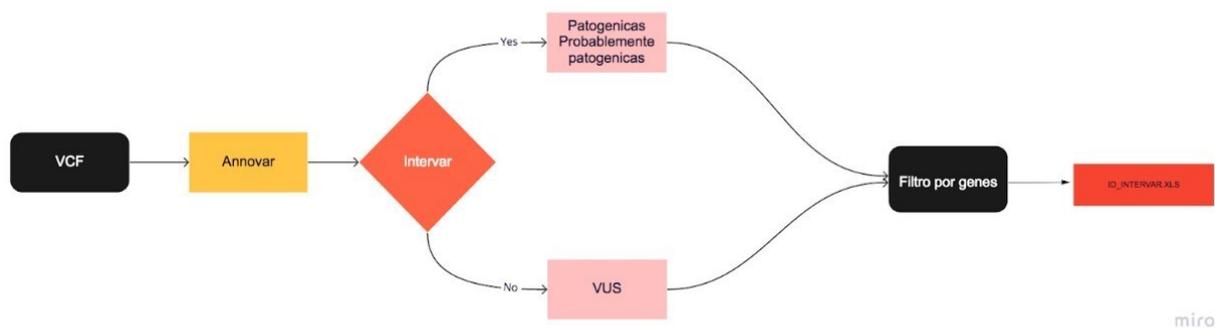
Figura 5. Diagrama del *pipeline* DRAGEN (127).

• Análisis Terciario

El análisis terciario tiene como objetivo interpretar los datos. Se compone por la anotación de variantes, visualización de datos, priorización y filtrado de variantes.

En este análisis se realiza la anotación funcional de 2080 VCF, utilizando la herramienta ANNOVAR e InterVar. Para este proceso se genera un *pipeline* de automatización (Figura 6) en el cual se realiza un filtro de 39 genes (ver Tabla 3) asociados a la farmacodinamia y farmacocinética de 17 analgesios opioides (ver Tabla 4); posteriormente, se filtran las variantes de acuerdo a la clasificación definida en las guías ACMG. Se genera una salida en formato excel, con cuatro hojas de cálculo, definidas como variantes totales, variantes benignas y probablemente benignas, variantes de significado incierto y variantes patogénicas y probablemente patogénicas.

Figura 6. Diagrama del pipeline de anotación.



miro

Tabla 3. Panel de 39 genes asociados a la farmacodinamia y farmacocinética de analgesicos opioides.

PANEL DE GENES		
ABCC3	CYP2D6	NTRK2
ABCB1	CYP3A4	OPRD1
ADRB2	CYP3A5	OPRK1
ALDH5A1	DBH	OPRM1
ANKK1	DRD2	P2RX7
BDNF	FAAH	RHBDF2
CACNA1E	GAD1	SLC22A1
CCL11	GNB3	SLC6A4
COMT	HCN1	SULT1A1
CYP1A2	IL1B	SULT1A3
CYP2B6	KCNJ6	TXNRD2
CYP2C19	MYD88	UGT1A1
CYP2C9	NGF	UGT2B7

Tabla 4. Analgesicos opioides.

Analgesico Opiode
Alfentanilo
Buprenorfina
Codeina
Dextropropoxifeno
Dihidrocodeina
Fentanilo
Hidromorfona
Metadona
Morfina
Naloxona
Naltrexona
Oxicodona
Oximorfona
Remifentanilo
Sufentanilo
Tapentadol
Tramadol

Se seleccionan las variantes de significado incierto, patogénicas y probablemente patogénicas, y se analiza el conjunto de variantes con relación a la frecuencia poblacional, evidencia bibliográfica, predicción *in silico*, conservación evolutiva, bases de datos clínicas y específicas para farmacogenómica (PharmGKB) (ver figura 7), con la finalidad de generar un score (ver Tabla 5) que permite determinar la relación de variante y efecto adverso.

Figura 7. Pipeline para la anotación funcional.

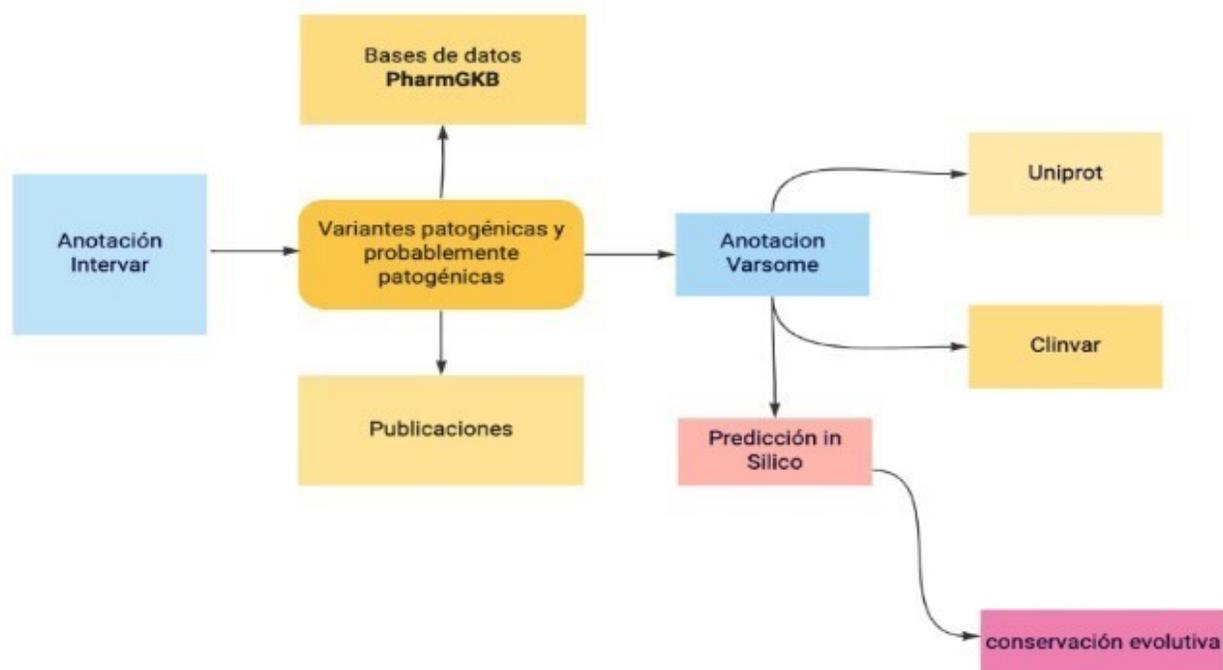


Tabla 5. Asignación del score (puntaje) de acuerdo con el riesgo.

Score	Clasificación de riesgo
0 a 1.5	Moderado
2	Medio
3 a 5	Alto

6.5 Modelo de aprendizaje automático

La metodología de la generación del modelo de aprendizaje automático se genera siguiendo las siguientes etapas (figura 8):

- **Entendimiento del contexto biológico:**

Dentro de esta etapa se define el problema y los objetivos del modelo. Los polimorfismos genéticos han sido asociados con la variación de la respuesta de los analgésicos opioides. El objetivo de desarrollar un modelo basado en aprendizaje automático es lograr la predicción del riesgo de evento adverso, a partir del genotipo en genes asociados con la farmacocinética y farmacodinámica de analgésicos opioides.

- **Entendimiento de los datos:**

Dentro de esta etapa se definen las estrategias para definir el problema, para lo cual se realiza la anotación funcional de las variantes identificadas.

- **Preparación del conjunto de datos (*data set*):**

En esta etapa se eliminan las variantes de mala calidad, con profundidad menor a 20 X, variantes clasificadas como benignas y probablemente benignas. Se realiza la selección de las variantes de significado incierto, patogénicas y probablemente patogénicas de acuerdo a la evidencia de puntaje (*score*) y asociación a efecto adverso.

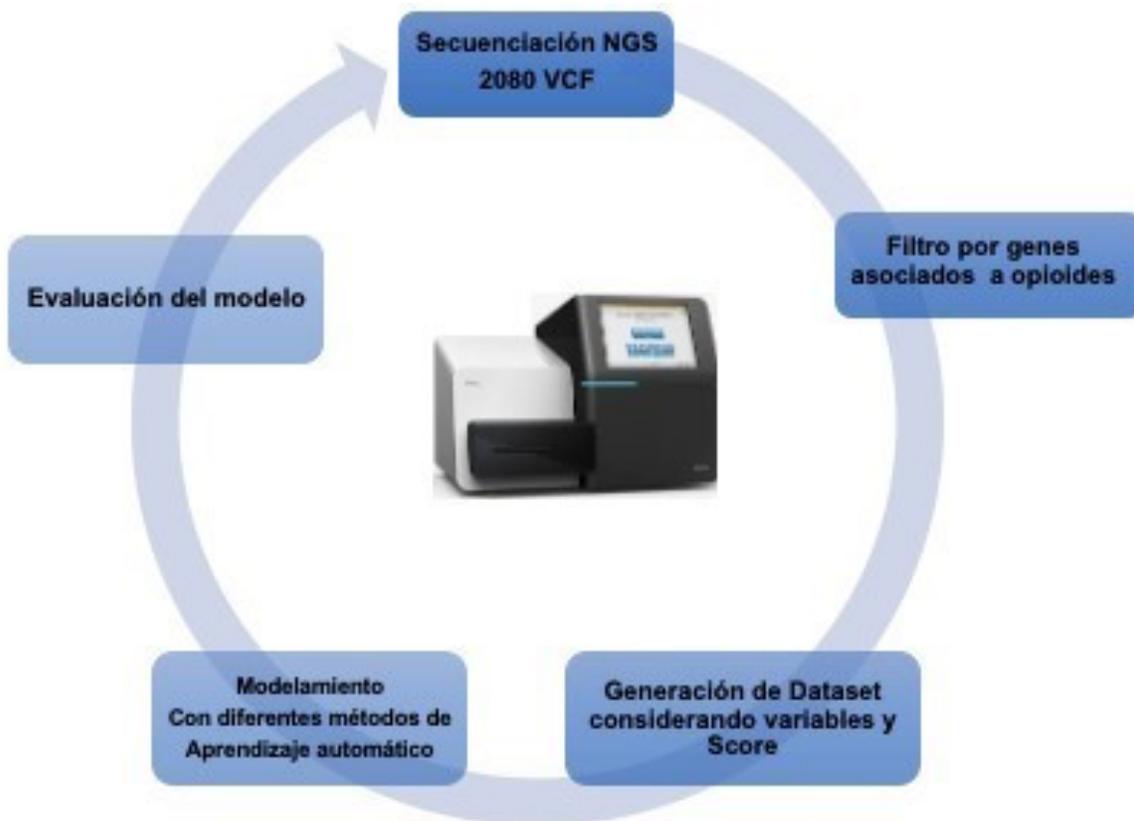
- **Modelación:**

Se seleccionan los modelos de aprendizaje automático, bosques aleatorios y perceptrón multicapa.

- **Evaluación del modelo:**

Se realiza el análisis de las matrices de confusión y accuracy de los diferentes métodos.

Figura 8. Metodología de aprendizaje automático.



7. Resultados

7.1 Resultados modelos de aprendizaje automático

Las variables que se consideran para el modelo de clasificación son:

Input:

- cromosoma
- posición cromosómica
- alelo wild type
- alelo alterno
- Gen
- Conservación

Output:

- score

7.1.2 Definición de las variables de entrada

A continuación se realiza la definición de las variables empleadas como entrada (*input*) en los modelos de aprendizaje automático desde el contexto biológico:

- Cromosoma: Estructura que se encuentra en el interior del núcleo de una célula el cual está compuesto por proteínas y ADN. Por lo habitual, una célula normal humana contiene 23 pares de cromosomas.
- Posición cromosómica: Indica la posición dentro del genoma humano, en la cual se encuentre una variante genética
- Alelo wild type: Alelo “normal” en un locus
- Alelo Alterno: Cambio de nucleótido wild type en un locus
- Gen: Unidad funcional y física de la herencia. Los genes son segmentos de ADN; la mayoría de los genes contienen información para generar una proteína específica.
- Conservación evolutiva: Son secuencias biológicas conservadas en ácidos nucleicos, proteínas, dentro de múltiples especies de organismos.

7.1.3 Efecto e importancia de cada variable de entrada (*input*) en la clasificación de los modelos de aprendizaje automático

El efecto de cada una de las variables de entrada (*input*) en la clasificación del sexto modelo de aprendizaje automático fueron calculados por permutación, la importancia de características de permutación (PFI) permite interpretar las predicciones del modelo de aprendizaje proporcionando la contribución relativa que cada característica aporta a la predicción (128).

Esto fue realizado usando la librería H2O, y se obtuvo que la variable con mayor importancia es el cromosoma, seguida por la posición, el gen, el alelo de referencia y el alelo alternativo (ver figura 9), la cuales fueron calculadas con la sumatoria de permutaciones (ver Tabla 6).

Figura 9. Importancia de las variables de entrada usando la librería H2O.

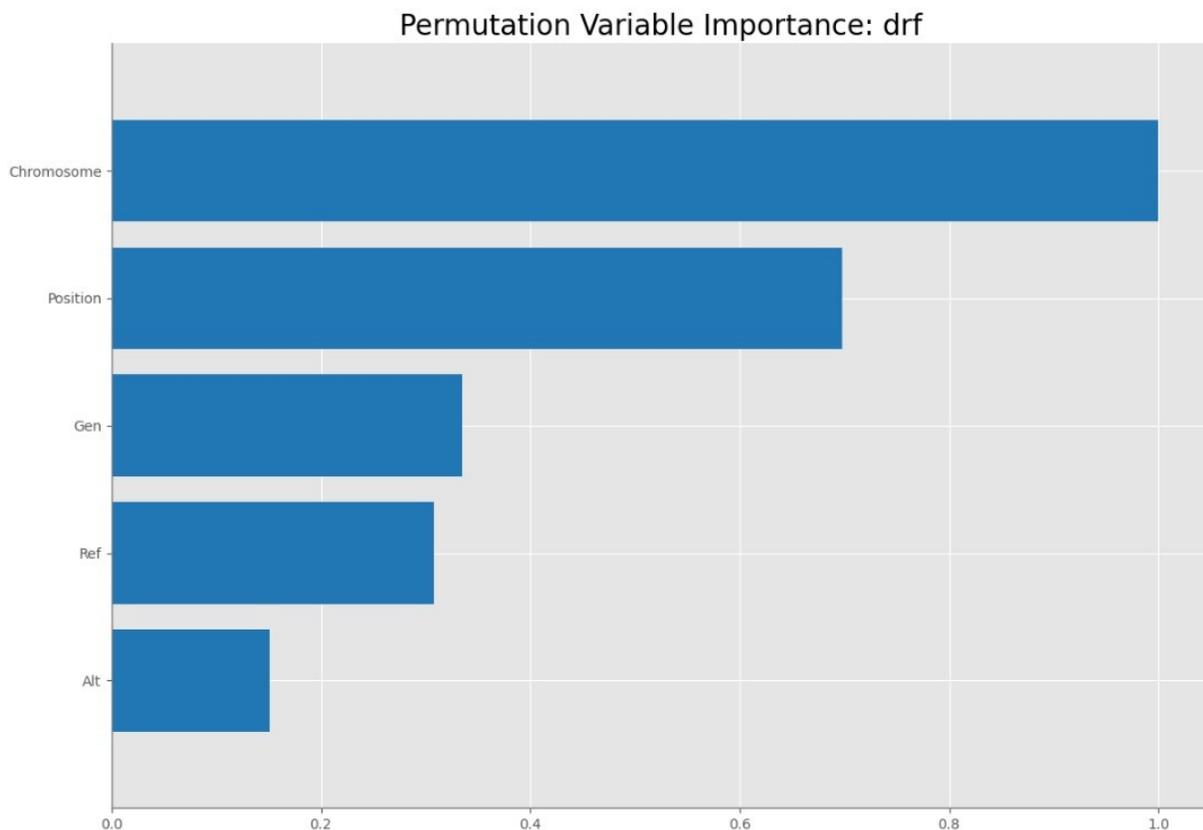


Tabla 6. Tabla de escala de importancia de las variables de entrada usando la librería H2O

	Variable	Relative Importance	Scaled Importance	Percentage
0	Chromosome	0.075130	1.000000	0.401395
1	Position	0.052435	0.697924	0.280143
2	Gen	0.025185	0.335226	0.134558
3	Ref	0.023101	0.307488	0.123424
4	Alt	0.011320	0.150674	0.060480

Por otro lado, la información calculada acerca de la importancia de las variables en el modelo de bosques aleatorios usando scikit-learn se presenta en la figura 10. En este ranking de las 6 variables más relevantes se encuentra en primer lugar la posición, posteriormente la conservación, con amplia ventaja de las variables gen, alelo de referencia y alelo alterno (ver **Tabla 7**).

Figura 10. Importancia de las variables de entrada usando la librería Scikit-learn .

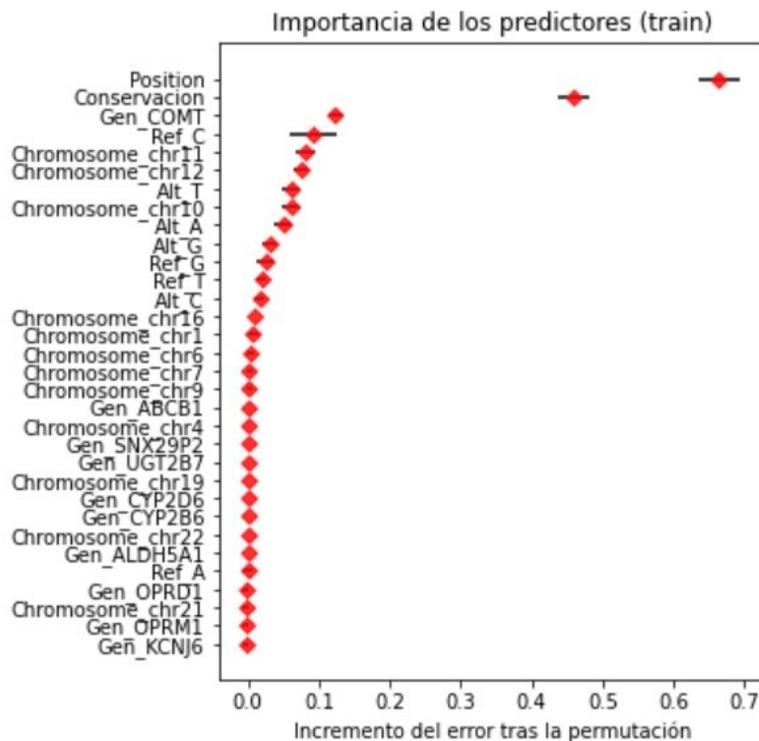


Figura 11. Escala de importancia de las variables de entrada obtenida usando la librería Scikitlearn

	importances_mean	importances_std	feature
0	0.664074	0.029494	Position
1	0.458803	0.022795	Conservacion
13	0.123499	0.009267	Gen_COMT
4	0.092334	0.031847	Ref_C
30	0.081957	0.013212	Chromosome_chr11
31	0.074673	0.010535	Chromosome_chr12
9	0.060945	0.012404	Alt_T
19	0.060932	0.012747	Chromosome_chr10
7	0.049940	0.012016	Alt_A
5	0.030212	0.010228	Alt_G
10	0.024960	0.013043	Ref_G
14	0.019934	0.006608	Ref_T
25	0.016394	0.006636	Alt_C
23	0.009264	0.000000	Chromosome_chr16
15	0.007354	0.006841	Chromosome_chr1
22	0.003705	0.004538	Chromosome_chr6
8	0.001853	0.003705	Chromosome_chr7
18	0.001853	0.003705	Chromosome_chr9
21	0.000000	0.000000	Gen_ABCB1
28	0.000000	0.000000	Gen_SNX29P2
24	0.000000	0.000000	Chromosome_chr22
6	0.000000	0.000000	Gen_UGT2B7
20	0.000000	0.000000	Gen_CYP2D6
17	0.000000	0.000000	Gen_ALDH5A1
3	0.000000	0.000000	Chromosome_chr4
12	0.000000	0.000000	Chromosome_chr19
11	0.000000	0.000000	Gen_CYP2B6
16	-0.000063	0.005959	Ref_A
26	-0.001916	0.003831	Gen_OPRD1
27	-0.001916	0.003831	Chromosome_chr21
29	-0.001916	0.003831	Gen KCNJ6

Con base en los resultados sobre la importancia de las variables con las dos librerías empleadas la variable cromosoma y posición fueron las más relevantes. Desde el contexto biológico, estas dos variables indican el cromosoma humano en el cual se encuentra el gen con la variante de interés, la coordenada indica la posición dentro del cromosoma, en la cual se está dando la variante genética, y, de esta manera, es posible ubicar el gen con el cual es posible generar asociaciones con la farmacocinética y farmacodinamia en los analgésicos opioides.

7.1.4 Hiperparámetros

Para buscar determinar los valores óptimos, se probaron diferentes posibilidades:

1. Se escogió un conjunto de valores para los hiperparámetros.
 - *Grid search* (búsqueda por cuadrícula): se realizó una búsqueda exhaustiva sobre un conjunto de valores previamente definidos.
 - *Random search* (búsqueda aleatoria): se evaluaron valores aleatorios dentro de unos límites definidos.
2. Para cada valor (en realidad, combinación de valores), se entrenaron los modelos y se estimó su error.
3. Finalmente, se ajustaron de nuevo los modelos, con todos los datos de entrenamiento y con los mejores hiperparámetros encontrados (ver Figura 12).

7.1.4.1 *Grid search* (búsqueda por cuadrícula)

El algoritmo Grid Search funciona probando todas las combinaciones posibles de parámetros, que se desean probar, en el modelo. Cada uno de esos parámetros se prueba en una serie de iteraciones de validación cruzada.

Se presentan los valores de la cuadrícula de hiperparámetros generada para el sexto modelo de aprendizaje automático que presenta los mejores resultados de entrenamiento.

De acuerdo a los resultados del árbol de selección, se obtienen 32 características que se pasan al modelo Scikit-learn, en el cual se establece la cuadrícula de hiperparametros. Esta cuadrícula de hiperparametros se aplica usando validación cruzada.

A continuación se presenta la cuadrícula de hiperparámetros generada (ver figura 12); también, se establecen los parámetros de la validación cruzada (ver figura 13) y se presentan los resultados de la mejor combinación encontrada (ver figura 14).

Figura 12. Cuadrícula de hiperparámetros

```
# Crear Grid con hiperparámetros
param_grid = {'n_estimators' : [150],
              #'max_features' : [3, 4, 7, 9, 15, 20, 30, 40, 45, 50, 60, 70],
              'max_features' : [3, 4, 7, 9, 15, 20, 32],
              'max_depth' : [None, 3, 5, 7, 8, 10, 20, 30, 40, 45],
              'criterion' : ['gini', 'entropy']}
)
```

Figura 13. Parámetros de la validación cruzada

```
# Búsqueda por grid search con validación cruzada
grid = GridSearchCV(
    estimator = RandomForestClassifier(random_state = 123),
    param_grid = param_grid,
    scoring = 'accuracy',
    n_jobs = multiprocessing.cpu_count() - 1,
    cv = RepeatedKFold(n_splits=5, n_repeats=3, random_state=123),
    refit = True, # Deja el modelo con los mejores hiperparámetros
    verbose = 0,
    return_train_score = True
)
```

Figura 14. Mejor combinación de hiperparámetros

	param_criterion	param_max_depth	param_max_features	param_n_estimators	mean_test_score	std_test_score	mean_train_score	std_train_score
118	entropy	20	32	150	0.914001	0.018910	0.999003	0.000814
139	entropy	45	32	150	0.913557	0.018384	0.999003	0.000814
76	entropy	None	32	150	0.913557	0.018384	0.999003	0.000814
125	entropy	30	32	150	0.913557	0.018384	0.999003	0.000814

7.2.1 Primer análisis

En el primer análisis se consideran los modelos de aprendizaje automático de árboles aleatorios, perceptrón multicapa y *naive bayes*. El conjunto de datos se compone de las clases:

Clase 1: Variantes genéticas con *score* de 0 a 1.5

Clase 2: Variantes genéticas con *score* de 1.5 a 2

Clase 3: Variantes genéticas con *score* de 3 a 5

El conjunto de datos inicial se divide en un conjunto de datos de entrenamiento y otro de evaluación como se especifica en la Tabla 7.

Tabla 7. Conjuntos (*datasets*) de entrenamiento y evaluación.

ENTRENAMIENTO		EVALUACION	
70%	Clase 1: 1900 variantes genéticas	30%	Clase 1: 815 variantes genéticas
	Clase 2: 50 variantes genéticas		Clase 2: 21 variantes genéticas
	Clase 3: 9 variantes genéticas		Clase 3: 4 variantes genéticas

Los parámetros utilizados para los diferentes modelos son detallados en la Tabla 8.

Tabla 8. Parámetros de los modelos de aprendizaje automático

MODELO	LIBRERÍA	METODOS
Bosques aleatorios		Estimadores:150 Max features:70, Max depth: 45, Criterio: entropía
Perceptron multicapa	Sklearn	Dos capas ocultas, tasa aprendizaje: 0.05, con 100 neuronas cada una
	Tensorflow	Dos capas ocultas con 200 neuronas cada una, funcion de activacion leaky -relu
Naive Bayes	Sklearn	Modelo BernoulliNB del paquete Naive bayes

Los resultados de clasificación que se obtienen con los tres modelos de aprendizaje automático son deficientes, esto se explica por el desbalance de clases. Se tiene una sobrerrepresentación de variantes en la clase 1, que se definen como las variantes de *score* de 0 a 1.5, que no se asocian al efecto adverso en la farmacocinética y farmacodinamia de los analgésicos opioides (ver Tabla 8).

Tabla 8. Resultados - matrices de confusión

MODELO	MATRIZ DE CONFUSION				RESULTADO
Bosques alatorios	Clase	C1	C2	C3	No clasifica adecuadamente las variantes de clase 2 y 3
	C1	817	0	0	
	C2	19	2	0	
	C3	2	0	0	
Perceptron multicapa	Clase	C1	C2	C3	No clasifica adecuadamente las variantes de clase 2 y 3
	C1	806	6	3	
	C2	18	3	0	
	C3	4	0	0	
Naive Bayes	Clase	C1	C2	C3	No clasifica adecuadamente las variantes de clase 2 y 3
	C1	805	10	0	
	C2	17	4	0	
	C3	4	0	0	

Teniendo en cuenta los resultados en las matrices de confusión (ver Tabla 8), y el desbalance de clases, se decide ajustar el conjunto de datos incluyendo 196 variantes asociadas con el efecto adverso en los analgésicos opioides seleccionados para esta investigación. Estas variantes son polimorfismos asociados con efectos adversos, tomados de las anotaciones clínicas de la base de datos PharmGKB (<https://www.pharmgkb.org/whatIsPharmgkb/clinicalAnnotations> y <https://www.pharmgkb.org/whatIsPharmgkb/variantAnnotations>).

7.2.2 Segundo análisis

En el segundo análisis el conjunto de datos se enriquece con 196 variantes asociadas con el efecto adverso en los analgésicos opioides, aumentando el número de variantes de la clase 3.

Se utiliza el modelo de aprendizaje automático de árboles aleatorios. El conjunto de datos se compone de las tres clases así:

Clase 1: Variantes genéticas con *score* de 0 a 1.5

Clase 2: Variantes genéticas con *score* de 1.5 a 2

Clase 3: Variantes genéticas con *score* de 3 a 5

El conjunto de datos inicial se divide en conjunto de entrenamiento y conjunto de evaluación (ver Tabla 9).

Tabla 9. Dataset de entrenamiento y evaluación.

ENTRENAMIENTO		EVALUACIÓN	
70%	Clase 1: 1909 variantes genéticas	30%	Clase 1: 806 variantes genéticas
	Clase 2: 51 variantes genéticas		Clase 2: 20 variantes genéticas
	Clase 3: 136 variantes genéticas		Clase 3: 73 variantes genéticas

Note que los resultados de clasificación que se obtienen siguen aún siendo deficientes (ver Tabla 10).

Tabla 10. Resultados - matriz de confusión

MODELO	MATRIZ DE CONFUSION	RESULTADO																
Bosques alatorios	<table border="1"> <thead> <tr> <th>Clase</th> <th>C1</th> <th>C2</th> <th>C3</th> </tr> </thead> <tbody> <tr> <th>C1</th> <td>799</td> <td>0</td> <td>7</td> </tr> <tr> <th>C2</th> <td>18</td> <td>1</td> <td>1</td> </tr> <tr> <th>C3</th> <td>51</td> <td>0</td> <td>22</td> </tr> </tbody> </table>	Clase	C1	C2	C3	C1	799	0	7	C2	18	1	1	C3	51	0	22	No clasifica adecuadamente las variantes de clase 2 y 3
Clase	C1	C2	C3															
C1	799	0	7															
C2	18	1	1															
C3	51	0	22															

7.2.3 Tercer análisis

En el tercer análisis se utilizó el conjunto de datos enriquecido con 196 variantes asociadas con el efecto adverso en los analgésicos opioides, aumentando el número de variantes de la clase 3.

Se utilizan los modelos de aprendizaje automático de árboles aleatorios y perceptrón multicapa. El conjunto de datos se compone de las clases:

Clase 1: Variantes genéticas con *score* de 0 a 1.5

Clase 2: Variantes genéticas con *score* de 2

Clase 3: Variantes genéticas con score de 3 a 5

El conjunto de datos inicial se divide en conjunto de entrenamiento y conjunto de evaluación (Tabla 11).

Tabla 11. Dataset de entrenamiento y evaluación.

ENTRENAMIENTO		EVALUACIÓN	
70%	Clase 1: 626 variantes genéticas	30%	Clase 1: 240 variantes genéticas
	Clase 2: 50 variantes genéticas		Clase 2: 21 variantes genéticas
	Clase 3: 146 variantes genéticas		Clase 3: 63 variantes genéticas

Teniendo en cuenta los resultados del segundo análisis, se decide suprimir los datos con score de 0.5 de la clase 1, obteniendo los resultados descritos en la Tabla 12. En donde se evidencia una mejoría en la clasificación para la clase 1 y 3, clasificando mal la clase 2. Siguen existiendo variantes para las tres clases que no se clasifican de manera adecuada.

Tabla 12. Análisis sin considerar score de 0.5 en Clase 1.

Modelo	Matriz de confusión				Resultado
Bosques aleatorios	Clase	C1	C2	C3	No se clasifican adecuadamente las variantes clase 2, mejora la clasificación de clase 3
	C1	217	9	14	
	C2	12	6	3	
	C3	11	0	51	
Perceptron multicapa	Clase	C1	C2	C3	No se clasifican adecuadamente las variantes clase 2, mejora la clasificación de clase 3
	C1	219	8	13	
	C2	14	3	4	
	C3	16	2	44	

7.2.4 Cuarto análisis

En el cuarto análisis se utiliza el conjunto de datos enriquecido con 196 variantes asociadas con el efecto adverso en los analgésicos opioides, aumentando el número de variantes de la clase 3.

Se utilizan los modelos de aprendizaje automático de árboles aleatorios y perceptrón multicapa. El conjunto de datos se compone de las clases:

Clase 1: Variantes genéticas con *score* de 1.5

Clase 2: Variantes genéticas con *score* de 2

Clase 3: Variantes genéticas con *score* de 3 a 5

El conjunto de datos inicial se divide en conjunto de entrenamiento y conjunto de evaluación (ver Tabla 13).

Tabla 13. Dataset de entrenamiento y evaluación.

ENTRENAMIENTO		EVALUACIÓN	
70%	Clase 1: 535 variantes genéticas	30%	Clase 1: 186 variantes genéticas
	Clase 2: 46 variantes genéticas		Clase 2: 25 variantes genéticas
	Clase 3: 140 variantes genéticas		Clase 3: 69 variantes genéticas

En el cuarto análisis se suprimen los datos con *scores* de 0.5 y 1 de la clase 1, obteniendo los resultados descritos en la Tabla 14. En donde se evidencia una mejoría en la clasificación de las clases 1 y 3, clasificando mal la clase 2. Sin embargo, siguen existiendo variantes para las tres clases que no se clasifican de manera adecuada.

Tabla 14. Análisis eliminando los datos con *scores* de 0.5 y 1 en la clase 1.

Modelo	Matriz de confusión				Resultado
Bosques aleatorios	Clase	C1	C2	C3	No se clasifican adecuadamente las variantes clase 2, mejora la clasificación de clase 3
	C1	170	2	14	
	C2	12	7	6	
	C3	6	1	62	
Perceptron multicapa	Clase	C1	C2	C3	No se clasifican adecuadamente las variantes clase 2, mejora la clasificación de clase 3
	C1	171	0	15	
	C2	15	4	6	
	C3	13	3	53	

7.2.5 Quinto análisis

En el quinto análisis se utiliza el conjunto de datos enriquecido con 196 variantes asociadas con el efecto adverso en los analgésicos opioides, aumentando el número de variantes de la clase 3.

Se utilizan los modelos de aprendizaje automático de árboles aleatorios y perceptrón multicapa. El conjunto de datos se compone de las clases:

Clase 1: Variantes genéticas con *score* de 1.5, suprimiendo los datos con *score* de 0.5 y 1

Clase 2: Unión de clases de variantes genéticas con *score* de 2 al 5

El conjunto de datos inicial se divide en conjunto de entrenamiento y conjunto de evaluación (ver Tabla 15).

Tabla 15. Dataset de entrenamiento y evaluación.

ENTRENAMIENTO		EVALUACIÓN	
70%	Clase 1: 555 variantes genéticas	30%	Clase 1: 240 variantes genéticas
	Clase 2: 197 variantes genéticas		Clase 2: 83 variantes genéticas

En el quinto análisis, teniendo en cuenta los resultados del tercer y cuarto análisis, donde se obtienen mejores resultados de clasificación de las variantes, se decide eliminar de la clase 1 los datos con los scores de 0.5 y 1. Se toma la decisión de unir las clases 2 y 3 en una sola clase que comprende los scores del 2 al 5. Obteniendo los resultados presentados en la Tabla 16. En donde se evidencia una mejoría en la clasificación de las clases, sin embargo, siguen existiendo variantes que no se clasifican de manera adecuada.

Tabla 16. Análisis eliminando los datos con scores de 0.5 y 1 en la Clase 1 y haciendo la unión de la clase 2 y la clase 3, en una sola clase.

MODELO	MATRIZ DE CONFUSION	RESULTADO									
Bosques alatorios	<table border="1"> <thead> <tr> <th>Clase</th> <th>C1</th> <th>C2</th> </tr> </thead> <tbody> <tr> <td>C1</td> <td>165</td> <td>21</td> </tr> <tr> <td>C2</td> <td>15</td> <td>79</td> </tr> </tbody> </table>	Clase	C1	C2	C1	165	21	C2	15	79	Se obtienen mejores clasificaciones para la clase 1 y 2
Clase	C1	C2									
C1	165	21									
C2	15	79									
Perceptron multicapa	<table border="1"> <thead> <tr> <th>Clase</th> <th>C1</th> <th>C2</th> </tr> </thead> <tbody> <tr> <td>C1</td> <td>214</td> <td>26</td> </tr> <tr> <td>C2</td> <td>30</td> <td>53</td> </tr> </tbody> </table>	Clase	C1	C2	C1	214	26	C2	30	53	Se obtienen mejores clasificaciones para la clase 1 y 2
Clase	C1	C2									
C1	214	26									
C2	30	53									

7.2.6 Sexto análisis

En el sexto análisis se utilizó el conjunto de datos enriquecido con 196 variantes asociadas con el efecto adverso en los analgésicos opioides, aumentando el número de variantes de la clase 3.

Se utiliza el modelo de aprendizaje automático de árboles aleatorios. El conjunto de datos se compone de las clases:

Clase 1: Variantes genéticas con *score* de 1 a 2

Clase 2: Variantes genéticas con *score* superior a 2

El conjunto de datos inicial se divide en conjunto de entrenamiento y conjunto de evaluación (ver tabla 15).

En el sexto análisis, teniendo en cuenta los resultados del quinto análisis, donde se obtienen mejores resultados de clasificación de las variantes, se decide generar el modelo con dos clases, la clase 1, compuesto por variantes genéticas con *score* de 1.5

a 2 y la clase 2 compuesto con variantes genéticas con score superior a 2. En donde se evidencia una mejoría en la clasificación, con respecto a los cinco modelos realizados (tabla 17). El modelo tiene una precisión del 80%, una exhaustividad del 82%, un F1 de 0.80 y una exactitud del 92%.

Tabla 17. Análisis con solo dos clases.

MODELO				RESULTADO
Bosques alatorios	Clase	C1	C2	Se obtienen mejores clasificaciones para la clase 1 y 2, con menor proporción de variantes mal clasificadas
	C1	249	12	
	C2	11	51	

8. Discusión

En el presente estudio se tienen en cuenta variantes genéticas asociadas a la respuesta farmacocinética y farmacodinámica en la acción de 17 analgésicos opioides formulados en la práctica clínica en Colombia.

La variación genética no representa en la mayoría de la población de estudio un riesgo asociado con un efecto adverso, por lo tanto, se tiene un desequilibrio entre el pequeño conjunto disponible de variantes genéticas asociadas a un efecto adverso y el conjunto de variantes genéticas, mucho más grande de no perjudiciales.

Entre de las variables de entrada del modelo de aprendizaje automático, la variable posición cromosómica presenta valores numéricos altos, que ocasionaban problemas en la clasificación, para lo cual se aplicó un escalamiento, empleando `MinMaxScaler`, para convertir estos valores altos en un rango entre 0 y 1.

En los modelos de bosques aleatorios con un número alto de árboles y de características (el hiperparámetro `max_feature`) se favorece el sobreajuste (*overfitting*). Por lo cual, se recomienda considerar experimentos adicionales con otros valores de estos hiperparámetros que permitan evitar el sobreajuste. Porque el modelo puede ajustarse mucho a los datos y no se identifican patrones generales, ocasionando que el modelo únicamente pueda producir resultados singulares y con la imposibilidad de comprender nuevos datos de entrada.

Los métodos de aprendizaje automático para clasificar variantes en datos con un desequilibrio en las clases, enfrentan problemas de clasificación, lo que resulta en una reducción significativa de la sensibilidad y la precisión (130).

Teniendo en cuenta el desequilibrio en las clases, se realizan varios experimentos con diversos modelos de aprendizaje automático. El modelo inicial con bosques aleatorios, perceptrón multicapa y naive bayes, logra clasificar muy bien la clase 1 que corresponde a las variantes de bajo riesgo, pero presenta resultados muy deficientes al clasificar tanto la clase 2, que representa variantes con significado incierto en el riesgo, como la clase 3 que corresponde a las variantes con alto riesgo de presentación de efecto adverso.

Teniendo en cuenta los resultados iniciales y el sobremuestreo de las mismas sin obtener buenos resultados en la clasificación, para el segundo análisis se enriquece el conjunto de datos inicial incluyendo 196 variantes asociadas con el efecto adverso en los analgésicos opioides seleccionados, enriqueciendo la clase 3, este conjunto de datos se emplea desde el segundo hasta el sexto análisis.

En el segundo análisis se emplea un modelo de bosques aleatorios, en el cual se obtienen mejores resultados de clasificación, para la clase 1, sin embargo, la clase 2 y clase 3, no se clasifican adecuadamente.

Estos resultados con una mejor clasificación en las clases, motiva la aproximación del tercer análisis, en el cual empleamos los modelos de bosques aleatorios y perceptrón

multicapa, eliminando los datos con el score de 0.5 de la clase 1, para eliminar variantes en regiones cromosómicas repetitivas que puedan generar ruido en los datos, los resultados demuestran una mejor clasificación de la clase 1 y la clase 3, sin embargo, la clase dos no se clasifica adecuadamente, en los dos modelos empleados.

Estos resultados impulsan el análisis con los modelos de bosques aleatorios y perceptrón multicapa en un cuarto análisis eliminando los datos con el score de 0.5 y 1 de la clase 1, obteniendo resultados similares, mejora la clasificación en clase 1 y clase 3, pero presenta una clasificación deficiente para los datos de la clase 2.

En el quinto análisis, teniendo en cuenta los resultados del tercer y cuarto análisis, donde se obtienen mejores resultados de clasificación de las variantes, se decide eliminar de la clase 1 los datos con los scores de 0.5 y 1. Se toma la decisión de unir las clases 2 y 3 en una sola clase que comprende los scores del 2 al 5. evidenciando una mejoría en la clasificación de las clases, sin embargo, siguen existiendo variantes que no se clasifican de manera adecuada.

En el sexto análisis, teniendo en cuenta los resultados del quinto análisis, donde se obtienen mejores resultados de clasificación de las variantes, se decide generar el modelo con dos clases, la clase 1, compuesto por variantes genéticas con score de 1.5 a 2, que representaría variantes genéticas sin riesgo y la clase 2 compuesto con variantes genéticas con score superior a 2, las cuales representan un riesgo alto. obteniendo mejores resultados de clasificación de la clase 1 y la clase 2, teniendo una proporción más baja de variantes mal clasificadas para las dos clases, con los dos modelos empleados, representando el modelo más óptimo de clasificación.

Estos resultados permiten analizar que al eliminar variantes con score de 0.5 del dataset, se generan mejores resultados en los modelos empleados, porque les permite a los modelos diferenciar entre las variantes pertenecientes a las clases 2 y 3. En adición el hecho de tener variantes genéticas en las mismas posiciones cromosómicas genera inconvenientes en el entrenamiento del modelo. Se conjetura que para poder resolver este último problema se requeriría ampliar la información teniendo en cuenta otras ómicas que permitan diferenciar las variantes genéticas.

Para obtener mejores resultados en los modelos de aprendizaje automático para la clasificación del riesgo de efecto adverso en analgésicos opioides en la población colombiana se proponen varias aproximaciones:

1. Considerar conjuntos de datos con enriquecimiento de variantes genéticas de riesgo.
2. Implementar medidas para evitar el sobreentrenamiento, aplicando técnicas adicionales como cross-validation, regularización, early stopping, pruning, bayesian priors en los parámetros.

3. Emplear modelos adicionales enfocados al desequilibrio de clases en datos genómicos humanos, como los son HyperSMURF (131)
4. Emplear técnicas técnica de sobremuestreo de minorías sintéticas (SMOTE) (132)
5. Utilizar más variables, relacionadas con el contexto omico en el cual se puedan generar relaciones de causalidad entre la presentación del efecto adverso.
6. Emplear población previamente descrita, con asociaciones de presentación de efecto adverso, para aumentar la probabilidad de hallazgos de variantes genéticas de importancia
7. Disponer de la información clínica completa de los pacientes participantes para poder generar relaciones de causalidad y seguimiento.

9. Conclusiones

En este trabajo se realizó el análisis de un conjunto de datos de exomas de pacientes colombianos secuenciados por NGS aplicando métodos de inteligencia computacional con el fin de clasificar el riesgo de efecto adverso a analgésicos opioides. Los datos incluían variantes genéticas, cromosoma, posición cromosómica, alelo *wild type*, alelo alterno, predicción funcional y *score*.

Se identifican variantes de riesgo en genes asociados a la farmacodinamia y farmacocinética de analgésicos opioides.

Se concluye que el uso de técnicas de inteligencia computacional y bioinformática es posible realizar la clasificación del riesgo de efecto adverso relacionado con variantes genéticas en genes relacionados a la respuesta farmacológica de los analgésicos opioides.

Fue posible desarrollar modelos computacionales de clasificación del riesgo de efecto adverso. Sin embargo, se concluye que el desequilibrio en las clases y el sobreentrenamiento, genera problemas de clasificación, lo que resulta en una reducción significativa de la sensibilidad y la precisión de los modelos de aprendizaje automático. Para poder generar relaciones entre variantes genéticas y su asociación con la presentación de efecto adverso se hace indispensable considerar variables de diversas ómicas que le den un peso importante a la asociación, acompañado de la información clínica y de seguimiento de los pacientes.

10. Perspectivas

Se sugiere realizar análisis similares al llevado a cabo en este trabajo sobre conjuntos con un mayor número de datos e incluir dentro de la población de estudio pacientes que hayan presentado efecto adverso relacionado al consumo de analgésicos opioides.

Se sugiere realizar experimentación adicional y emplear técnicas que eviten el sobreentrenamiento, empleando detención temprana (*early stopping*), entrenamiento del modelo con más datos, considerar el aumento de datos con datos relevantes, realizar una selección de características adecuado y emplear regularización.

Se sugiere tener conjuntos de datos más homogéneos, emplear modelos adicionales enfocados al desequilibrio de clases en datos genómicos humanos y emplear técnicas de sobremuestreo de minorías sintéticas, para asegurar mayor sensibilidad y precisión de los modelos de aprendizaje automático.

Además, se sugiere la inclusión de otras variables desde diferentes ómicas, y tener en cuenta análisis de causalidad haplotípica en la presentación del efecto adverso del fármaco.

11. Anexos

1. Farmacocinética y farmacodinamia de los analgésicos opioides

1.1 Codeína

Farmacocinética

- **Absorción:** Este fármaco se absorbe bien por la vía oral. La vida media de la codeína en plasma es de 2.5 a 3.5 horas (36).
- **Distribución:** El volumen de distribución es de 2.2 lt/kg (37).
- **Biotransformación y excreción:** Para que la codeína pueda ejercer su efecto analgésico debe ser convertida en morfina. El 10% de la dosis administrada se convierte en morfina; este será el metabolito responsable del efecto analgésico. Una vez que se realiza esta reacción de biotransformación tanto el metabolito como la codeína libre se excretan por vía renal aproximadamente un 90% en 24 horas (38).

Farmacodinamia

Una vez que la codeína, en la fracción señalada, se convierte a morfina, se produciría interacción con los receptores μ y se presentará el efecto analgésico. La codeína es un analgésico débil cuando se usa solo, pero este efecto mejora cuando se utiliza con analgésicos no opioides, tales como paracetamol, diclofenaco e ibuprofeno (39).

1.2 Tramadol

Farmacocinética

- **Absorción:** Este fármaco se absorbe por la vía oral. La vida media de la codeína en plasma es de 2 horas (41).
- **Distribución:** El volumen de distribución es de 2.7 lt/kg (42).
- **Biotransformación y excreción:** Sufre reacciones de biotransformación y se excreta por vía renal. Luego de la administración oral de Tramadol, el 90% se excreta por riñón y 10% por las heces (43).

Farmacodinamia

El Tramadol es un fármaco con múltiples mecanismos de acción que contribuyen a su efecto analgésico. La analgesia inducida por Tramadol no es revertida en su totalidad por Naloxona (44).

1.3 Morfina

Farmacocinética

- **Absorción:** Este fármaco se absorbe por la vía oral, realiza un primer paso hepático, por lo que su disponibilidad es del 20-30%. Se absorbe en el intestino delgado, lo cual permite su administración en forma de tabletas y soluciones. Las vías subcutánea, intravenosa, intratecal, epidural, intra-articular y rectal,

constituyen otras opciones para la administración de este medicamento. Dada su naturaleza hidrofílica, la morfina no puede ser administrada por las vías nasal, sublingual, transdérmica (45,46).

- **Distribución:** El volumen de distribución de este medicamento es de 3.3 +/- 0.9 lt /kg, lo cual implica que se distribuye ampliamente en los diferentes tejidos corporales. Debido al metabolismo de la morfina y a su vida media de 2-4 horas, luego de una dosis única, a las 24 horas no se observan concentraciones plasmáticas significativas, pero debido a su volumen de distribución accede al SNC y Sistema Nervioso Periférico (47).
- **Biotransformación:** La morfina se biotransforma por reacciones de glucuronización al glucurónido 6 de morfina (G-6-M), metabolito con la actividad analgésica y al glucurónido 3 de morfina (G-3-M); ambos logran atravesar la barrera hematoencefálica. El G-3-M posee muy poca o ninguna actividad analgésica; sin embargo, se señala como el metabolito responsable de los efectos de toxicidad en el SNC que incluyen excitación, alucinaciones y fenómenos de hiperalgesia (44).
- **Excreción:** La vía renal es la principal ruta de eliminación de morfina y de sus metabolitos, por lo cual no debe administrarse en pacientes con falla renal. Los metabolitos conjugados están expuestos a recirculación enterohepática, por lo cual una pequeña fracción del fármaco puede aparecer en heces (46).

Farmacodinamia

La interacción de morfina con los receptores μ genera una cadena de señalización que incluye más de un tipo de proteínas G y más de un efector, en la asta dorsal y la sustancia gris periacueductal a partir de la cual, la interacción de morfina con este receptor activa vías descendentes que impactan en la asta dorsal en el sitio mencionado. La interacción con receptores μ disminuye la motilidad del intestino e incrementa el tono de los esfínteres. De la interacción con estos receptores se produce un efecto analgésico en el dolor agudo y crónico, en especial en dolor nociceptivo, y se considera menos eficaz en el manejo del dolor neuropático (46).

1.4 Hidromorfona

Farmacocinética

- **Absorción:** Es un fármaco altamente liposoluble que se absorbe por vía oral, parenteral y rectal. La infusión subcutánea continua es posible con hidromorfona por su solubilidad, buena absorción y porque no irrita los tejidos (36). Su biodisponibilidad por vía oral es de 30-40% y su relación oral/parenteral es de 5:1, con un efecto máximo a los 30-60 minutos y una duración de 4-5 horas (37).
- **Distribución:** El volumen de distribución de este medicamento es cercano a los 2.90 lt/kg, lo cual implica que su distribución excede al agua corporal total

y le permite alcanzar sitios distantes desde el lugar de administración. Su vida media es de aproximadamente 2-3 horas (37). • **Biotransformación y excreción**

- La hidromorfona se biotransforma por la vía de conjugación con glucurónidos y genera varios metabolitos activos. Entre ellos, el glucuronido³ de hidromorfona (H3G) que tiene propiedades excitatorias en el SNC (48). • **Farmacodinamia**
- La Hidromorfona es un agonista de los receptores μ , muchas de sus acciones son comparables a las de morfina. Las dosis deben individualizarse y titularse. Es eficaz en dolor agudo y crónico oncológico y no oncológico, por vía intravenosa y subcutánea, así como en Analgesia Controlada por el Paciente. Su unión a proteínas plasmáticas es limitada (<30%) (48).

1.5 Oxidodona Farmacocinética

- **Absorción:** La oxidodona se absorbe por vía oral, tiene un primer paso hepático y su biodisponibilidad es 60-87% (50).
- **Distribución:** Su volumen de distribución es de 2 lts/kg. Su vida media es de 2.6 horas, por lo que para alcanzar concentraciones estables se requiere que transcurran 12 horas luego de iniciado el tratamiento (51).
- **Biotransformación y excreción:** La vida media plasmática es de 3-5 horas y alcanza niveles estables en 24-36 horas. Se biotransforma a nivel hepático por reacciones de Odesmetilación que generan metabolitos inactivos que son posteriormente eliminados por la vía renal; parte del compuesto original es eliminado por esta misma vía. Se metaboliza en el hígado por el sistema enzimático citocromo P-450 (CYP3A4 y CYP2D6) a noroxidodona (analgésico débil a concentraciones elevadas) y oximorfona (analgésico potente a concentraciones bajas), que no contribuyen al efecto farmacológico (51).

Farmacodinamia

La Oxidodona se considera un agonista de receptores μ , por ello es capaz de inducir un efecto analgésico significativo, cuya intensidad se relaciona con la dosis (51).

1.6 Fentanilo

Farmacocinética

- **Absorción:** La estructura química del fentanilo, como derivado de fenilpiperidinas, con dos anillos aromáticos, hace de este compuesto una

sustancia de gran solubilidad en lípidos, lo que le permite atravesar las mucosas, membranas celulares y la piel. Esto hace que fentanilo sea útil en el alivio del dolor interrecurrente. La concentración plasmática máxima para analgesia se alcanza en aproximadamente 5 minutos después de su administración intravenosa (52).

- **Distribución:** El volumen de distribución del fentanilo es de 4 lt/kg, lo cual implica que el producto tiene una distribución que sobrepasa los compartimentos acuosos del cuerpo. Su vida media es de 3.7 horas aproximadamente, lo que implica que luego de la administración de una sola dosis por vía parenteral, el producto desaparecerá del organismo en 18 horas. (37).
- **Biotransformación y excreción:** El Fentanilo tiene un primer paso hepático, vía citocromo P450 3A4 para formar norfentanil, metabolito inactivo; vía hidrólisis genera dos metabolitos adicionales. Se elimina por vía renal y no hay transformación a nivel de la piel. Sus concentraciones a nivel central disminuyen, luego de una sola dosis, como consecuencia de la influencia del fenómeno de redistribución hacia depósitos grasos (53). **Farmacodinamia**

El fentanilo es un agonista de los receptores μ con efectos analgésicos y sedantes. Comparado con morfina tiene un efecto analgésico cien veces más potente (52).

1.7 Buprenorfina

Farmacocinética

- **Absorción:** La absorción es buena por sus diferentes vías de administración, con una biodisponibilidad oral baja y lenta disociación, que se traduce en una duración de acción de 6 horas. La liposolubilidad le permite una absorción excelente por la barrera cutánea (54).
- **Distribución:** Se distribuye ampliamente hacia la circulación plasmática, uniéndose en un 96% a las proteínas del plasma (55).
- **Biotransformación y excreción:** Se metaboliza por vía hepática, se transforma principalmente a buprenorfina glucurónido y, en forma parcial a norbuprenorfina, cuya difusión en el cerebro es muy baja. La excreción biliar es su vía de eliminación principal, independientemente de la vía de administración. Dos tercios del fármaco se eliminan inalterados por las heces y un tercio se elimina sin cambios o de-alquilado a través del aparato urinario; posee circulación enterohepática. La farmacocinética de la buprenorfina no se altera con la edad ni en pacientes con disfunción renal (55,56). **Farmacodinamia**

La buprenorfina se clasifica como un agonista parcial de los receptores opioides μ , posee actividad antagonista sobre los receptores κ y δ . La acción analgésica eficaz de la buprenorfina se alcanza con una ocupación de

receptores relativamente baja, aproximadamente 5-10%, y se ha podido comprobar que debido a este hecho, el grado de analgesia no se correlaciona estrechamente con las concentraciones plasmáticas del fármaco (55,56).

1.8 Naloxona

Farmacocinética

Absorción: Se absorbe con facilidad por el tubo digestivo, la naloxona se metaboliza casi por completo en el hígado, antes de llegar a la circulación general. Su inicio de acción es de 1 a 2 minutos después de su administración (57,58).

Distribución: El efecto máximo por cualquiera de las vías de administración se observa entre los 5 y 15 minutos y tiene un volumen de distribución de 1,8 lt/kg (36).

Biotransformación y excreción: La mayor parte de la naloxona se biotransforma en su primer paso hepático. Es hidrolizada a N-alilo-14-hidroxi7, 8-dihidromorfina y noroximorfona (7, 8-dihidro-14-hidroxinormorfina). Los tres compuestos se someten a glucoronidación, lo que contribuye al alto efecto de primer paso y a su corta duración de acción. La formación de naloxona3glucurónido, que es el principal metabolito que se encuentra en la orina, es mediada por UGT2B7 (59).

Farmacodinamia

Inhíbe competitivamente los opioides en los sitios receptores μ , δ y κ . Naloxona revierte y previene los efectos adversos de los opioides: depresión respiratoria, sedación, hipotensión arterial sistémica, analgesia y espasmo de vías biliares. Para antagonizar los efectos de depresión respiratoria por buprenorfina se requieren dosis más altas de naloxona. Dosis de 1 mg IV de naloxona bloquean por completo los efectos de 25 mg de heroína (36). 2.

2. Resultados de Genotipificación

Los resultados obtenidos en la genotipificación se presentan en tabla 18.

Tabla 18. Resultados de genotipificación.

Genes	HGVS	Clasificación	Medicamento	Tipo de anotación	Número de pacientes		Descripción
<u>SULT1A3</u>	NM_177552.4:c.43G>A(p.Val15Met)	VUS	Tapentadol	Metabolism/pk	1 (74382)	Farmacocinética	Menor tasa de sulfatación
		VUS	O-desmethyltramadol	Metabolism/pk			
<u>SULT1A3</u>	NM_177552.4:c.25C>T(p.Arg9Cys)	VUS	Morfina	Metabolism/pk	5 (63350, 70762, 72953,79163,79200)		
		VUS	Tapentadol	Metabolism/pk			
		VUS	O-desmetiltramadol	Metabolism/pk			
		VUS	Tramadol	Metabolism/pk			
<u>SULT1A1</u>	NM_177530.2:c.110G>A(p.Arg37Gln)	VUS	Tapentadol	Metabolism/pk	1 (59639)		
<u>SULT1A3</u>	NM_177552.4:c.302C>T(p.Pro101Leu)	VUS	Paracetamol	Metabolismo/pk	2 (55577, 76835)		
		VUS	Morfina	Metabolismo/pk			
		VUS	Tapentadol	Metabolismo/pk			
		VUS	O-desmetiltramadol	Metabolismo/pk			

Genes	HGVS	Clasificación	Medicamento	Tipo de anotación	Número de pacientes		Descripción
SLC22A1	NM_003057.3:c.1201G>A(p.Gly401Ser)	VUS	Tramadol	Metabolismo/pk	47	Farmacocinética	Concentraciones plasmáticas aumentadas, se asocia con una dosis reducida de tramadol en personas con dolor
		VUS	Morfina	Metabolismo/pk			Disminución del aclaramiento de morfina en niños
		VUS	o-desmethyltramadol	Metabolismo/pk			Mayores concentraciones de o-desmetiltramadol
<u>CYP2C19</u>	NM_000769.4:c.395G>A(p.Arg132Gln)	VUS	Metadona	Metabolismo/pk	1 (74330)		Aclaramiento reducido
<u>CYP2D6</u>	NM_000106.6:c.1228G>A(p.Glu410Lys)	VUS	codeína	Metabolismo/pk	14	Mayor depuración de codeína	

12. Referencias

1. Simon, E. J., Hiller, J. M. & Edelman, I. 1973. Stereospecific binding of the potent narcotic analgesic [3H] etorphine to rat-brain homogenate. *Proceedings of the National Academy of Sciences*, 70, 1947-1949
2. Kieffer, B. L. 1995. Recent advances in molecular recognition and signal transduction of active peptides: receptors for opioid peptides. *Cellular and molecular neurobiology*, 15, 615- 635
3. Chen, Y., Mestek, A., Liu, J., Hurley, J. A. & Yu, L. 1993. Molecular cloning and functional expression of a mu-opioid receptor from rat brain. *Molecular Pharmacology*, 44, 8-12.
4. Isaza, C., Sepúlveda-Arias, J. C. & Henao, J. 2009. Pharmacogenomics in medicine. *Colombia Médica*, 40, 327-346
5. Madadi P, Ross CJ, Hayden MR, Carleton BC, Gaedigk A, Leeder JS, et al. Pharmacogenetics of neonatal opioid toxicity following maternal use of codeine during breastfeeding: a case-control study. *Clinical pharmacology and therapeutics*. 2009;85(1):31-5. doi: 10.1038/clpt.2008.157. PubMed PMID: 18719619
6. Koren G, Cairns J, Chitayat D, Gaedigk A, Leeder SJ. Pharmacogenetics of morphine poisoning in a breastfed neonate of a codeine-prescribed mother. *Lancet*. 2006;368(9536):704. doi: 10.1016/S0140- 6736(06)69255-6. PubMed PMID: 16920476
7. Owusu Obeng, A., Hamadeh, I., & Smith, M. (2017). Review of opioid pharmacogenetics and considerations for pain management. *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy*, 37(9), 1105-1121
8. Laksman, Z., and Detsky, A. S. (2011). Personalized medicine: understanding probabilities and managing expectations. *J. Gen. Intern. Med.* 26, 204–206. doi: 10.1007/s116060101515-6
9. Mueller, J. J., Schlappé, B. A., Kumar, R., Olvera, N., Dao, F., Abu-Rustum, N., et al. (2018). Massively parallel sequencing analysis of mucinous ovarian carcinomas: genomic profiling and differential diagnoses. *Gynecol. Oncol.* 150, 127–135. doi: 10.1016/j.ygyno.2018.05.008
10. Evans, W. E., and Relling, M. V. (2004). Moving towards individualized medicine with pharmacogenomics. *Nature* 429, 464–468. doi: 10.1038/nature02626
11. Feero, W. G., Guttmacher, A. E., and Collins, F. S. (2008). The genome gets personal – Almost. *JAMA* 299, 1351–1352. doi: 10.1001/jama.299.11.1351
12. Ann K Daly, Pharmacogenetics: a general review on progress to date, *British Medical Bulletin*, Volume 124, Issue 1, December 2017, Pages 65–79
13. Ho, D. S. W., Schierding, W., Wake, M., Saffery, R., & O’Sullivan, J. (2019). Machine Learning SNP Based Prediction for Precision Medicine. *Frontiers in Genetics*, 10.
14. Capriotti, E., Calabrese, R., and Casadio, R. (2006). Predicting the insurgence of human

- genetic diseases associated to single point protein mutations with support vector machines and evolutionary information. *Bioinformatics* 22, 2729–2734. doi: 10.1093/bioinformatics/btl423
15. Cruz, J. A., and Wishart, D. S. (2006). Applications of machine learning in cancer prediction and prognosis. *Cancer Inform.* 2, 59–77. doi: 10.1177/117693510600200030
16. Palaniappan, S., and Awang, R. (2008). “Intelligent heart disease prediction system using data mining techniques,” in *Proceedings of the 2008 IEEE/ACS Int. Conf. Comput. Syst. Appl.* (Doha), 108–115. doi: 10.1109/AICCSA.2008.4493524
17. Yu, W. (2010). Application of support vector machine modeling for prediction of common diseases: the case of diabetes and pre-diabetes. *BMC Med. Inform. Decis. Mak.* 10:16. doi: 10.1186/1472-6947-10-16
18. Zhang, D., and Shen, D. (2012). Multi-modal multi-task learning for joint prediction of multiple regression and classification variables in Alzheimer’s disease. *Neuroimage* 59, 895–907. doi: 10.1016/j.neuroimage.2011.09.069
19. López, B., Torrent-Fontbona, F., Viñas, R., and Fernández-Real, J. M. (2017). Single nucleotide polymorphism relevance learning with random forests for type 2 diabetes risk prediction. *Artif. Intell. Med.* 85, 43–49. doi: 10.1016/j.artmed.2017.09.005
20. Boulesteix, A. L., Janitza, S., Kruppa, J., and König, I. R. (2012). Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* 2, 493–507. doi: 10.1002/widm.1072
21. Chen, X., and Ishwaran, H. (2012). Random forests for genomic data analysis. *Genomics* 99, 323–329. doi: 10.1016/j.ygeno.2012.04.003
22. Austin, P. C., Tu, J. V., Ho, J. E., Levy, D., and Lee, D. S. (2013). Using methods from the data-mining and machine-learning literature for disease classification and prediction: a case study examining classification of heart failure subtypes. *J. Clin. Epidemiol.* 66, 398–407. doi: 10.1016/j.jclinepi.2012.11.008
23. Kringel, D., Geisslinger, G., Resch, E., Oertel, B. G., Thrun, M. C., Heinemann, S., & Lötsch, J. (2018). Machine-learned analysis of the association of next-generation sequencing–based human TRPV1 and TRPA1 genotypes with the sensitivity to heat stimuli and topically applied capsaicin. *Pain*, 159(7), 1366
24. Lo-Ciganic, W. H., Huang, J. L., Zhang, H. H., Weiss, J. C., Wu, Y., Kwoh, C. K., ... & Kuza, C. C. (2019). Evaluation of Machine-Learning Algorithms for Predicting Opioid Overdose Risk Among Medicare Beneficiaries With Opioid Prescriptions. *JAMA network open*, 2(3), e190968-e190968
25. Organización Mundial de la salud (OMS), Manejo del abuso de sustancias. Recuperado de https://www.who.int/substance_abuse/information-sheet/es/

26. Observatorio de Drogas de Colombia (O.D.C), Alerta informativa acerca del uso de sustancias y preparados con actividad psicoactiva utilizadas para el manejo del dolor (opioides), 08-12-019. Recuperado de <http://www.odc.gov.co/Portals/1/SAT/Alerta%20Opioides.pdf>
27. EVANS, C. J., KEITH, D., MORRISON, H., MAGENDZO, K. & EDWARDS, R. H. 1992. Cloning of a delta opioid receptor by functional expression. SCIENCE-NEW YORK THEN WASHINGTON-, 258, 1952-1952.
28. KIEFFER, B. L. 1995. Recent advances in molecular recognition and signal transduction of active peptides: receptors for opioid peptides. Cellular and molecular neurobiology, 15, 615- 635.
29. CHEN, Y., MESTEK, A., LIU, J., HURLEY, J. A. & YU, L. 1993. Molecular cloning and functional expression of a mu-opioid receptor from rat brain. Molecular Pharmacology, 44, 8-12.
30. Lorenzo P, Moreno A, Leza JC, Lizasoain I, Moro MA, Portolés A. Velazquez. Farmacología Básica y Clínica. 19ª ed. Editorial Médica Panamericana; 2018.
31. Iórez J. Fármacos analgésicos opioides. En: Flórez J, Armijo JA, Mediavilla A, editores. Farmacología humana. 4ª ed. Barcelona: Masson; 2003. p. 461-478
32. Rang HP, Dale MM, Ritter JM, Moore PK. editores. Farmacología. 5ª ed. Madrid: Elsevier; 2004. p. 562-583
33. Cherny NI. Opioid analgesics. Drugs 1996; 51: 713-737
34. Mendoza Patiño, N. *Farmacología médica* (3ª edición). Editorial Médica Panamericana.
35. Colleau, S. Evaluación de la escalera analgésica de la Organización Mundial de la Salud en su 20º aniversario. Cancer Pain Release. 2006;19:2. 2. Lacy Ch, Armstrong L, Goldman M, Lance L. Drug Information Handbook. Lexi Comp 15th Ed. p. 410-412.
36. Gutstein HB, Akil H. Opioid Analgesics. En: Brunton L, Lazo J, Parker K. Goodman and Gilman's The Pharmacological Basis of Therapeutics. 2006.
37. Thummel K. Shen D, Isoherranen, Smith H. Appendix II Design and Optimization of Dosage Regimens: Pharmacokinetic data. In: Brunton L, Lazo J, Parker K: Goodman and Gilman's The Pharmacological Basis of Therapeutics. 11th ed. 2006. p. 1.813
38. Vascello L and Maquillan R. Opioid Analgesics and Routes of Administration. De Leon Casasola O. Cancer Pain. Pharmacological Interventional and Palliative care Approaches. Elsevier inc. 2006:171-193.
39. Smith HS. Variations in opioid responsiveness. Pain Physician. 2008 Mar-Apr; 11(2):237-48.
40. Colleau, S M. Cancer Press Release 2006; Vol 19: p. 2: Evaluación de la escalera analgésica de la Organización Mundial de la Salud en su 20º aniversario.
41. Leppert W, Luczak J. The role of tramadol in cancer pain treatment-a review. Support Care Cancer. 2005 Jan;13(1):5-17. Epub 2004 Nov 18
42. Grond S, Sablotzki A. Clinical pharmacology of tramadol. Clin Pharmacokinet. 2004;43(13):879-923

43. Leppert W, Luczak J. The role of tramadol in cancer pain treatment-a review. *Support Care Cancer*. 2005 Jan;13(1):5-17. Epub 2004 Nov 18.
44. Vascello L, Maquillan R. Opioid Analgesics and Routes of Administration. De Leon Casasola O. *Cancer Pain. Pharmacological Interventional and Palliative Care Approaches*. Elsevier inc. 2006:171-193.
45. Lacy Ch, Armstrong L, Goldman M, Lance L. *Drug Information Handbook*. Lexi Comp 15th Ed. Page 1171-1175.
46. Gutstein HB, Akil H. Opioid Analgesics, Chap 21. In: Brunton L; Lazo J; Parker K: Goodman and Gilman's *The Pharmacological Basis of Therapeutics*. 11th ed. 2006.
47. Thummel K, Shen D, Isoherranen, Smith H. Appendix II Design and Optimization of Dosage Regimens: Pharmacokinetic data. In: Brunton L; Lazo J; Parker K: Goodman and Gilman's. *The Pharmacological Basis of Therapeutics*. 11th ed. 2006. p.1851.
48. Schumacher M, Basbaun A, Way W. in: Katzung B. *Basic and clinical pharmacology*. 10th Ed, Mc Graw-Hill Companies The; 2006. Chap 31
49. Ross FB, Smith MT. The intrinsic antinociceptive effects of oxycodone appear to be kappa-opioid receptor mediated. *Pain*. 1997 Nov;73(2):151-7.
50. Benziger DP, Kaiko RF, Miotto JB, Fitzmartin RD, Reder RF, Chasin M. Differential effects of food and the bioavailability of controlled-release oxycodone tablets and immediate-release oxycodone solution. *J Pharm Sci*. 1996 Apr;85(4):407-10.
51. Kaiko RF, Benziger DP, Fitzmartin RD, Burke BE, Reder RF, Goldenheim PD. Pharmacokinetic-pharmacodynamic relationships of controlled-release oxycodone. *Clin Pharmacol Ther*. 1996 Jan;59:52-61.
52. Darwish M, Kirby M, Robertson P Jr, Tracewell W, Jiang JG. Absolute and relative bioavailability of fentanyl buccal tablet and oral transmucosal fentanyl citrate. *J Clin Pharmacol*. 2007 Mar;47(3):343-50.
53. Lacy Ch, Armstrong L, Goldman M, Lance L. *Drug Information Handbook*. Lexi Comp 15th ed: p. 693-699
54. Huang P, Kehner GB, Cowan A, Liu-Chen LY. Comparison of pharmacological activities of buprenorphine and norbuprenorphine: norbuprenorphine is a potent opioid agonist. *J Pharmacol Exp Ther*. 2001 May;297(2):688-95.
55. Böger RH. Renal impairment: a challenge for opioid treatment? The role of buprenorphine. *Palliat Med*. 2006;20 Suppl 1:s17-23.
56. Zaki PA, Keith DE Jr, Brine GA, Carroll FI, Evans CJ. Ligand-induced changes in surface mu-opioid receptor number: relationship to G protein activation? *J Pharmacol Exp Ther*. 2000 Mar;292(3):1127-34
57. McNicol ED, Boyce D, Schumann R, Carr DB. Mu-opioid antagonists for opioid-induced bowel dysfunction. *Cochrane Database Syst Rev*. 2008 Apr 16;(2):CD006332.

58. Nebreda C, Bruno U. Manual de fármacos utilizados en el tratamiento del dolor crónico. IAHPC 2001;42-43
59. Trescot AM, Datta S, Lee M, Hansen H. Opioid pharmacology. *Pain Physician*. 2008 Mar;11(2 Suppl):S133-53.
60. Pappagallo M, Stewart W, Woods M. Constipation symptoms in long-term users of opioid analgesic therapy. Poster-Abstracts. American Pain Society Annual Meeting, Fort Lauderdale, Florida, October 21-24, 1999.
61. Homann NG, Olofsson O, Salen B, Wickstrom L. Prevalence of abuse and dependency in chronic pain patients. *Int J Addict*. 1995 Jun;30(8):919-27.
62. Hahn, D., Fukuda, T., Euteneuer, J. C., Mizuno, T., Vinks, A. A., Sadhasivam, S., & Emoto, C. (2020). Influence of MRP3 genetics and hepatic expression ontogeny for morphine disposition in neonatal and pediatric patients. *The Journal of Clinical Pharmacology*, 60(8), 992-998.
63. Lam J, Woodall KL, Solbeck P, Ross CJ, Carleton BC, Hayden MR, Koren G, Madadi P. Codeine-related deaths: The role of pharmacogenetics and drug interactions. *Forensic Sci Int*. 2014 Jun;239:50-6. doi: 10.1016/j.forsciint.2014.03.018. Epub 2014 Mar 26. PMID: 24747667.
64. Zhang J, Zhang L, Zhao X, Shen S, Luo X, Zhang Y. Association between MDR₁/CYP3A4/OPRM₁ gene polymorphisms and the post-caesarean fentanyl analgesic effect on Chinese women. *Gene*. 2018 Jun 30;661:78-84. doi: 10.1016/j.gene.2018.03.081. Epub 2018 Mar 27. PMID: 29601950.
65. Naito T, Takashina Y, Yamamoto K, Tashiro M, Ohnishi K, Kagawa Y, Kawakami J. CYP3A5*3 affects plasma disposition of noroxycodone and dose escalation in cancer patients receiving oxycodone. *J Clin Pharmacol*. 2011 Nov;51(11):1529-38. doi: 10.1177/0091270010388033. Epub 2011 Jan 5. PMID: 21209234.
66. Saiz-Rodríguez M, Ochoa D, Herrador C, Belmonte C, Román M, Alday E, Koller D, Zubiaur P, Mejía G, Hernández-Martínez M, Abad-Santos F. Polymorphisms associated with fentanyl pharmacokinetics, pharmacodynamics and adverse effects. *Basic Clin Pharmacol Toxicol*. 2019 Mar;124(3):321-329. doi: 10.1111/bcpt.13141. Epub 2018 Oct 30. PMID: 30281924.
67. Fonseca F, Gratacòs M, Escaramís G, De Cid R, Martín-Santos R, Farré M, Estivill X, Torrens M. ALDH5A1 variability in opioid dependent patients could influence response to methadone treatment. *Eur Neuropsychopharmacol*. 2014 Mar;24(3):4204. doi: 10.1016/j.euroneuro.2013.10.003. Epub 2013 Oct 18. PMID: 24230997.
68. Hung CC, Chiou MH, Huang BH, Hsieh YW, Hsieh TJ, Huang CL, Lane HY. Impact of genetic polymorphisms in ABCB1, CYP2B6, OPRM1, ANKK1 and DRD2 genes on methadone therapy in Han Chinese patients. *Pharmacogenomics*. 2011

Nov;12(11):1525-33. doi: 10.2217/pgs.11.96. Epub 2011 Sep 8. PMID: 21902500.

69. de Cid R, Fonseca F, Gratacòs M, Gutierrez F, Martín-Santos R, Estivill X, Torrens M. BDNF variability in opioid addicts and response to methadone treatment: preliminary findings. *Genes Brain Behav.* 2008 Jul;7(5):515-22. doi: 10.1111/j.1601183X.2007.00386.x. PMID: 18182069.
70. Yoshida K, Nishizawa D, Ichinomiya T, Ichinohe T, Hayashida M, Fukuda K, Ikeda K. Prediction formulas for individual opioid analgesic requirements based on genetic polymorphism analyses. *PLoS One.* 2015 Jan 23;10(1):e0116885. doi: 10.1371/journal.pone.0116885. PMID: 25615449; PMCID: PMC4304713.
71. Kuo HW, Liu TH, Tsou HH, Hsu YT, Wang SC, Fang CP, Liu CC, Chen ACH, Liu YL. Inflammatory chemokine eotaxin-1 is correlated with age in heroin dependent patients under methadone maintenance therapy. *Drug Alcohol Depend.* 2018 Feb 1;183:19-24. doi: 10.1016/j.drugalcdep.2017.10.014. Epub 2017 Nov 28. PMID: 29222992.
72. Cargnin S, Magnani F, Viana M, Tassorelli C, Mittino D, Cantello R, Sances G, Nappi G, Canonico PL, Genazzani AA, Raffaelli W, Terrazzino S. An opposite-direction modulation of the COMT Val158Met polymorphism on the clinical response to intrathecal morphine and triptans. *J Pain.* 2013 Oct;14(10):1097-106. doi: 10.1016/j.jpain.2013.04.006. Epub 2013 Jun 15. PMID: 23773341.
73. Jones JD, Mumtaz M, Manubay JM, Mogali S, Sherwin E, Martinez S, Comer SD. Assessing the contribution of opioid- and dopamine-related genetic polymorphisms to the abuse liability of oxycodone. *Pharmacol Biochem Behav.* 2019 Nov;186:172778. doi: 10.1016/j.pbb.2019.172778. Epub 2019 Sep 4. PMID: 31493434; PMCID: PMC6801039.
74. Chang HW, Ho WC, Huang CL, Wang RY. Precision therapeutic opioid dosing implications from genetic biomarkers and craving score. *Medicine (Baltimore).* 2020 May 29;99(22):e20429. doi: 10.1097/MD.00000000000020429. PMID: 32481444.
75. Crettol S, Déglon JJ, Besson J, Croquette-Krokar M, Hämmig R, Gothuey I, Monnat M, Eap CB. ABCB1 and cytochrome P450 genotypes and phenotypes: influence on methadone plasma levels and response to treatment. *Clin Pharmacol Ther.* 2006 Dec;80(6):668-81. doi: 10.1016/j.clpt.2006.09.012. PMID: 17178267.
76. Saiz-Rodríguez M, Ochoa D, Román M, Zubiaur P, Koller D, Mejía G, AbadSantos F. Involvement of CYP2D6 and CYP2B6 on tramadol pharmacokinetics. *Pharmacogenomics.* 2020 Jul;21(10):663-675. doi: 10.2217/pgs-2020-0026. Epub 2020 Jun 15. PMID: 32538291

77. Chen YJ, Lu JT, Huang CW, Wu WH, Lee KF, Liu HT, Shih-Hsin Wu L. Pharmacogenetic study of methadone treatment for heroin addiction: associations between drug-metabolizing gene polymorphisms and treatment efficacy. *Pharmacogenet Genomics*. 2022 Jan 1;32(1):31-38. doi: 10.1097/FPC.0000000000000450. PMID: 34380995.
78. Kringen MK, Chalabianloo F, Bernard JP, Bramness JG, Molden E, Høiseth G. Combined Effect of CYP2B6 Genotype and Other Candidate Genes on a Steady-State Serum Concentration of Methadone in Opioid Maintenance Treatment. *Ther Drug Monit*. 2017 Oct;39(5):550-555. doi: 10.1097/FTD.0000000000000437. PMID: 28723731.
79. Fonseca F, de la Torre R, Díaz L, Pastor A, Cuyàs E, Pizarro N, Khymenets O, Farré M, Torrens M. Contribution of cytochrome P450 and ABCB1 genetic variability on methadone pharmacokinetics, dose requirements, and response. *PLoS One*. 2011 May 12;6(5):e19527. doi: 10.1371/journal.pone.0019527. PMID: 21589866; PMCID: PMC3093392.
80. Koren G, Cairns J, Chitayat D, Gaedigk A, Leeder SJ. Pharmacogenetics of morphine poisoning in a breastfed neonate of a codeine-prescribed mother. *Lancet*. 2006 Aug 19;368(9536):704. doi: 10.1016/S0140-6736(06)69255-6. PMID: 16920476.
81. Badaoui S, Hopkins AM, Rodrigues AD, Miners JO, Sorich MJ, Rowland A. Application of Model Informed Precision Dosing to Address the Impact of Pregnancy Stage and CYP2D6 Phenotype on Foetal Morphine Exposure. *AAPS J*. 2021 Jan 6;23(1):15. doi: 10.1208/s12248-020-00541-1. PMID: 33404848.
82. Halling J, Weihe P, Brosen K. CYP2D6 polymorphism in relation to tramadol metabolism: a study of faroese patients. *Ther Drug Monit*. 2008 Jun;30(3):271-5. doi: 10.1097/FTD.0b013e3181666b2f. PMID: 18520597.
83. Samer CF, Daali Y, Wagner M, Hopfgartner G, Eap CB, Rebsamen MC, Rossier MF, Hochstrasser D, Dayer P, Desmeules JA. The effects of CYP2D6 and CYP3A activities on the pharmacokinetics of immediate release oxycodone. *Br J Pharmacol*. 2010 Jun;160(4):907-18. doi: 10.1111/j.1476-5381.2010.00673.x. PMID: 20590587; PMCID: PMC2935997.
84. Dong ZL, Li H, Chen QX, Hu Y, Wu SJ, Tang LY, Gong WY, Xie GH, Fang XM. Effect of CYP3A4*1G on the fentanyl consumption for intravenous patient-controlled analgesia after total abdominal hysterectomy in Chinese Han population. *J Clin Pharm Ther*. 2012 Apr;37(2):153-6. doi: 10.1111/j.1365-2710.2011.01268.x. Epub 2011 Apr 29. PMID: 21535061.
85. Zhu GD, Whitley P, LaRue L, Adkins B, Dawson E, Huskey A, Capparelli EV, Del Tredici AL. Impact of genetic variation in CYP2C19, CYP2D6, and CYP3A4 on oxycodone and its metabolites in a large database of clinical urine drug tests.

Pharmacogenomics J. 2022 Feb;22(1):25-32. doi: 10.1038/s41397-021-00253-5. Epub 2021 Sep 3. PMID: 34480108.

86.Tanaka N, Naito T, Yagi T, Doi M, Sato S, Kawakami J. Impact of CYP3A5*3 on plasma exposure and urinary excretion of fentanyl and norfentanyl in the early postsurgical

period. Ther Drug Monit. 2014 Jun;36(3):345-52. doi: 10.1097/FTD.000000000000029. PMID: 24365989.

87.Naito T, Takashina Y, Yamamoto K, Tashiro M, Ohnishi K, Kagawa Y, Kawakami J. CYP3A5*3 affects plasma disposition of noroxycodone and dose escalation in cancer patients receiving oxycodone. J Clin Pharmacol. 2011 Nov;51(11):1529-38. doi: 10.1177/0091270010388033. Epub 2011 Jan 5. PMID: 21209234.

88.Arias AJ, Gelernter J, Gueorguieva R, Ralevski E, Petrakis IL. Pharmacogenetics of naltrexone and disulfiram in alcohol dependent, dually diagnosed veterans. Am J Addict. 2014 May-Jun;23(3):288-93. doi: 10.1111/j.1521-0391.2014.12102.x. PMID: 24724887; PMCID: PMC4600600.

89.Doehring A, Hentig Nv, Graff J, Salamat S, Schmidt M, Geisslinger G, Harder S, Lötsch J. Genetic variants altering dopamine D2 receptor expression or function modulate the risk of opiate addiction and the dosage requirements of methadone substitution. Pharmacogenet Genomics. 2009 Jun;19(6):407-14. doi: 10.1097/FPC.0b013e328320a3fd. PMID: 19373123.

90.Chidambaran V, Pilipenko V, Spruance K, Venkatasubramanian R, Niu J, Fukuda T, Mizuno T, Zhang K, Kaufman K, Vinks AA, Martin LJ, Sadhasivam S. Fatty acid amide hydrolase-morphine interaction influences ventilatory response to hypercapnia and postoperative opioid outcomes in children. Pharmacogenomics. 2017 Jan;18(2):143-156. doi: 10.2217/pgs-2016-0147. Epub 2016 Dec 15. PMID: 27977335; PMCID: PMC5558540.

91.Shi Y, Li Y, Zhang J, Xiao Y, Yan P, Zhu Y. GAD1 but not GAD2 polymorphisms are associated with heroin addiction phenotypes. Neurosci Lett. 2020 Jan 19;717:134704. doi: 10.1016/j.neulet.2019.134704. Epub 2019 Dec 19. PMID: 31866536.

92.Hung CC, Chiou MH, Huang BH, Hsieh YW, Hsieh TJ, Huang CL, Lane HY. Impact of genetic polymorphisms in ABCB1, CYP2B6, OPRM1, ANKK1 and DRD2 genes on methadone therapy in Han Chinese patients. Pharmacogenomics. 2011 Nov;12(11):1525-33. doi: 10.2217/pgs.11.96. Epub 2011 Sep 8. PMID: 21902500.

93. Wendt FR, Rahikainen AL, King JL, Sajantila A, Budowle B. A genome-wide association study of tramadol metabolism from post-mortem samples. Pharmacogenomics J. 2020 Feb;20(1):94-103. doi: 10.1038/s41397-019-0088-y. Epub 2019 Apr 11. PMID: 30971809.

94.Somogyi AA, Sia AT, Tan EC, Collier JK, Hutchinson MR, Barratt DT. Ethnicitydependent influence of innate immune genetic markers on morphine PCA

requirements and adverse effects in postoperative pain. *Pain*. 2016 Nov;157(11):2458-2466. doi: 10.1097/j.pain.0000000000000661. PMID: 27649267.

95. Yoshida K, Nishizawa D, Ichinomiya T, Ichinohe T, Hayashida M, Fukuda K, Ikeda K. Prediction formulas for individual opioid analgesic requirements based on genetic polymorphism analyses. *PLoS One*. 2015 Jan 23;10(1):e0116885. doi: 10.1371/journal.pone.0116885. PMID: 25615449; PMCID: PMC4304713.

96. Lötsch J, Prüss H, Veh RW, Doehring A. A KCNJ6 (Kir3.2, GIRK2) gene polymorphism modulates opioid effects on analgesia and addiction but not on pupil size. *Pharmacogenet Genomics*. 2010 May;20(5):291-7. doi: 10.1097/FPC.0b013e3283386bda. PMID: 20220551.

97. Elens L, Norman E, Matic M, Rane A, Fellman V, van Schaik RH. Genetic Predisposition to Poor Opioid Response in Preterm Infants: Impact of KCNJ6 and COMT Polymorphisms on Pain Relief After Endotracheal Intubation. *Ther Drug Monit*. 2016 Aug;38(4):525-33. doi: 10.1097/FTD.0000000000000301. PMID: 27027462.

98. Barratt DT, Klepstad P, Dale O, Kaasa S, Somogyi AA. Innate Immune Signalling Genetics of Pain, Cognitive Dysfunction and Sickness Symptoms in Cancer Pain Patients Treated with Transdermal Fentanyl. *PLoS One*. 2015 Sep 2;10(9):e0137179. doi: 10.1371/journal.pone.0137179. PMID: 26332828; PMCID: PMC4557995.

99. Levran O, Peles E, Hamon S, Randesi M, Zhao C, Zhang B, Adelson M, Kreek MJ. Nerve growth factor β polypeptide (NGFB) genetic variability: association with the methadone dose required for effective maintenance treatment. *Pharmacogenomics J*. 2012 Aug;12(4):319-27. doi: 10.1038/tpj.2011.6. Epub 2011 Mar 1. PMID: 21358750; PMCID: PMC3130093.

100. Levran O, Peles E, Randesi M, Shu X, Ott J, Shen PH, Adelson M, Kreek MJ. Association of genetic variation in pharmacodynamic factors with methadone dose required for effective treatment of opioid addiction. *Pharmacogenomics*. 2013 May;14(7):755-68. doi: 10.2217/pgs.13.58. PMID: 23651024; PMCID: PMC3867202.

101. Kranzler HR, Lynch KG, Crist RC, Hartwell E, Le Moigne A, Laffont CM, Andorn AC. A Delta-Opioid Receptor Gene Polymorphism Moderates the Therapeutic Response to Extended-Release Buprenorphine in Opioid Use Disorder. *Int J Neuropsychopharmacol*. 2021 Feb 15;24(2):89-96. doi: 10.1093/ijnp/pyaa069. PMID: 32920647; PMCID: PMC7883889.

102. Jones JD, Mumtaz M, Manubay JM, Mogali S, Sherwin E, Martinez S, Comer SD. Assessing the contribution of opioid- and dopamine-related genetic polymorphisms to the abuse liability of oxycodone. *Pharmacol Biochem Behav*. 2019 Nov;186:172778. doi: 10.1016/j.pbb.2019.172778. Epub 2019 Sep 4. PMID: 31493434; PMCID: PMC6801039.

103. Luo R, Li X, Qin S, Luo Z, Luo X, Hu P, Liu Z, Feng X, Li X. Impact of SNP-SNP interaction among ABCB1, ARRB2, DRD1 and OPRD1 on methadone dosage

requirement in Han Chinese patients. *Pharmacogenomics*. 2017 Dec;18(18):16591670. doi: 10.2217/pgs-2017-0072. Epub 2017 Nov 27. PMID: 29173032.

104.Chatti I, Woillard JB, Mili A, Creveaux I, Ben Charfeddine I, Feki J, Langlais S, Ben Fatma L, Saad A, Gribaa M, Libert F. Genetic Analysis of Mu and Kappa Opioid Receptor and COMT Enzyme in Cancer Pain Tunisian Patients Under Opioid Treatment. *Iran J Public Health*. 2017 Dec;46(12):1704-1711. PMID: 29259946; PMCID: PMC5734971.

105. Zhang Q, Shi M, Tang H, Zhong H, Lu X. κ Opioid Receptor 1 Single Nucleotide Polymorphisms were Associated with the Methadone Dosage. *Genet Test Mol Biomarkers*. 2020 Jan;24(1):17-23. doi: 10.1089/gtmb.2019.0159. PMID: 31940240.

106.Pu J, Wang N, Huang ZK, He XY, Yuan HB. Correlation between gene polymorphism and opioid efficacy in patients with gastric or intestinal cancer. *Eur Rev Med Pharmacol Sci*. 2019 Nov;23(21):9393-9410. doi: 10.26355/eurrev_201911_19432. PMID: 31773688.

107.Pettini E, Micaglio M, Bitossi U, De Gaudio AR, Degl'Innocenti DR, Tofani L, Limatola V, Adembri C, Di Filippo A. Influence of OPRM1 Polymorphism on Postoperative Pain After Intrathecal Morphine Administration in Italian Patients Undergoing Elective Cesarean Section. *Clin J Pain*. 2018 Feb;34(2):178-181. doi: 10.1097/AJP.0000000000000520. PMID: 28591085.

108.Ma J, Li W, Chai Q, Tan X, Zhang K. Correlation of P2RX7 gene rs1718125 polymorphism with postoperative fentanyl analgesia in patients with lung cancer. *Medicine (Baltimore)*. 2019 Feb;98(7):e14445. doi: 10.1097/MD.00000000000014445. PMID: 30762755; PMCID: PMC6408006.

109.Galvan A, Skorpen F, Klepstad P, Knudsen AK, Fladvad T, Falvella FS, Pigni A, Brunelli C, Caraceni A, Kaasa S, Dragani TA. Multiple Loci modulate opioid therapy response for cancer pain. *Clin Cancer Res*. 2011 Jul 1;17(13):4581-7. doi: 10.1158/1078-0432.CCR-10-3028. Epub 2011 May 27. PMID: 21622719.

110.Tzvetkov MV, Saadatmand AR, Lötsch J, Tegeder I, Stingl JC, Brockmöller J. Genetically polymorphic OCT1: another piece in the puzzle of the variable pharmacokinetics and pharmacodynamics of the opioidergic drug tramadol. *Clin Pharmacol Ther*. 2011 Jul;90(1):143-50. doi: 10.1038/clpt.2011.56. Epub 2011 May 11. PMID: 21562485.

111.Zhu XL, Han X, Xin XF, Song HJ, Guo H. Correlations of analgesic dosage of morphine with SLC6A4 gene polymorphisms in patients with lung cancer. *Eur Rev Med Pharmacol Sci*. 2020 May;24(9):5046-5052. doi: 10.26355/eurrev_202005_21197. PMID: 32432768.

112.Rasool MI, Bairam AF, Gohal SA, El Daibani AA, Alherz FA, Abunnaja MS, Alatwi ES, Kurogi K, Liu MC. Effects of the human SULT1A1 polymorphisms on the sulfation

of acetaminophen, O-desmethylnaproxen, and tapentadol. *Pharmacol Rep*. 2019 Apr;71(2):257-265. doi: 10.1016/j.pharep.2018.12.001. Epub 2018 Dec 10. PMID: 30822619; PMCID: PMC6498441.

113. Bairam AF, Rasool MI, Alherz FA, Abunnaja MS, El Daibani AA, Kurogi K, Liu MC. Effects of human SULT1A3/SULT1A4 genetic polymorphisms on the sulfation of acetaminophen and opioid drugs by the cytosolic sulfotransferase SULT1A3. *Arch Biochem Biophys*. 2018 Jun 15;648:44-52. doi: 10.1016/j.abb.2018.04.019. Epub 2018 Apr 26. PMID: 29705271; PMCID: PMC5969526.

114. Duan L, Li X, Yan J, Chen Y, Luo R, Zhang Q, Feng X, Li X. Association of COMT Gene Polymorphisms with Response to Methadone Maintenance Treatment Among Chinese Opioid-Dependent Patients. *Genet Test Mol Biomarkers*. 2020 Jun;24(6):364369. doi: 10.1089/gtmb.2019.0275. Epub 2020 May 14. PMID: 32407152.

115. Holthe M, Klepstad P, Zahlsen K, Borchgrevink PC, Hagen L, Dale O, Kaasa S, Krokan HE, Skorpen F. Morphine glucuronide-to-morphine plasma ratios are unaffected by the UGT2B7 H268Y and UGT1A1*28 polymorphisms in cancer patients on chronic morphine therapy. *Eur J Clin Pharmacol*. 2002 Aug;58(5):353-6. doi: 10.1007/s00228-002-0490-1. Epub 2002 Jul 13. PMID: 12185559.

116. Muraoka W, Nishizawa D, Fukuda K, Kasai S, Hasegawa J, Wajima K, Nakagawa T, Ikeda K. Association between UGT2B7 gene polymorphisms and fentanyl sensitivity in patients undergoing painful orthognathic surgery. *Mol Pain*. 2016 JanDec;12:1744806916683182. doi: 10.1177/1744806916683182. PMID: 28256933; PMCID: PMC5521342.

117. Kaya-Akyüzlü D, Özkan-Kotiloğlu S, Bal C, Yalçın-Şahiner Ş, Avcıoğlu G, Danışman M. Effects of UGT2B7 rs7662029 and rs7439366 polymorphisms on sublingual buprenorphine metabolism in heroin addicts: An improved PCR-RFLP assay for the detection of rs7662029 polymorphism. *Environ Toxicol Pharmacol*. 2022 Aug;94:103902. doi: 10.1016/j.etap.2022.103902. Epub 2022 Jun 10. PMID: 35697190.

118. Baber M, Chaudhry S, Kelly L, Ross C, Carleton B, Berger H, Koren G. The pharmacogenetics of codeine pain relief in the postpartum period. *Pharmacogenomics J*. 2015 Oct;15(5):430-5. doi: 10.1038/tpj.2015.3. Epub 2015 Mar 10. PMID: 25752520.

119. Mitchell T.M. *Machine Learning*. McGraw-hill; New York, NY, USA: 1997.

120. Cilluffo G, Fasola S, Ferrante G, Malizia V, Montalbano L, La Grutta S. Machine Learning: An Overview and Applications in Pharmacogenetics. *Genes (Basel)*. 2021 Sep 26;12(10):1511. doi: 10.3390/genes12101511. PMID: 34680905; PMCID: PMC8535911.

121. Breiman, L. Random forests. *Mach. learning* 45, 5–32 (2001).

122. Christian, Brian; Griffiths, Tom (April 2017), "Chapter 7: Overfitting", Algorithms To Live By: The computer science of human decisions, William Collins, pp. 149–168, ISBN 978-0-00-754799-9
123. Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot y E. Duchesnay: Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.
124. Friedman N., Geiger D., Goldszmidt M. Bayesian networks classifiers. *Machine Learning*, 29:131–163, 1997.
125. Friedman N., Linial M., Nachman I, Pe'er D. Using Bayesian Networks to Analyze Expression Data. *Journal of Computational Biology*, 7:601– 620, 2000.
126. Needham C., Bradford J., Bulpitt A., Westhead D. A primer on learning in Bayesian networks for computational biology. *PLoS Comput Biol*, 3:e129, 2007.
127. .Editado de:
<https://www.illumina.com/products/bytype/informaticsproducts/dragen-bio-it-platform.html>
128. Breiman, Leo. "Random Forests." *Machine Learning* 45 (1). Springer: 5-32 (2001).
129. Galar M, Fernandez A, Barrenechea E, Bustince H, Herrera F. Una revisión de conjuntos para el problema del desequilibrio de clases: enfoques basados en embolsado, refuerzo e híbrido. *sist. Hombre, Cibern. Parte C Aplic. Rev. IEEE Trans.* 2012; 42 :463–484. doi: 10.1109/TSMCC.2011.2161285.
130. Schubach M, Re M, Robinson PN, Valentini G. Imbalance-Aware Machine Learning for Predicting Rare and Common Disease-Associated Non-Coding Variants.
131. *Sci Rep.* 2017 Jun 7;7(1):2959. doi: 10.1038/s41598-017-03011-5. PMID: 28592878; PMCID: PMC5462751.
132. Chawla, N. V., Bowyer, K. W., Hall, L. O. & Kegelmeyer, W. P. SMOTE: Synthetic Minority Over-sampling Technique. *J. Artif. Intell. Res.* 321–357 (2002).