



FACULTAD DE MEDICINA

ESCUELA DE MEDICINA

Trabajo de titulación previo a la obtención de título de médico.

**Métodos de Correlación Estructura-Actividad Aplicados a Efectos
Adversos de Antibióticos, Cuenca, 2019.**

Autoras:

Jesenia Tatiana Jácome Barrionuevo

Paola Alexandra Guerrero Ramón

Director:

Dr. Johann Radax

Cuenca 2020

RESUMEN

En el siguiente estudio se utilizó el método QSAR para desarrollar un modelo matemático que relacione parámetros estructurales de los antimicrobianos con sus efectos adversos, permitiendo la predicción respectiva en el caso de nuevos fármacos con similares composiciones moleculares.

Para elaborar dicho modelo matemático usamos el método QSAR, el cual agrupa un conjunto de técnicas computacionales relacionadas con diseño y visualización espacial virtual de moléculas, descriptores, bioinformática y estadística. Para el cálculo de los descriptores moleculares está el software Dragon 7, y para el análisis estadístico, el programa Matlab.

Logramos elaborar un modelo satisfactorio con el 67% de precisión en la predicción, el 80% de sensibilidad y el 88% de especificidad.

La importancia de este trabajo yace en la posibilidad de ahorrar tiempo y recursos en el desarrollo de nuevos fármacos y de incrementar la seguridad de estos medicamentos por una preselección molecular acertada.

ABSTRACT

In the following study, the QSAR method was used to develop a mathematical model that relates structural parameters of antimicrobials with their adverse effects, allowing predictions in the case of new drugs with similar molecular compositions. QSAR is a set of computational techniques related to design and virtual spatial visualization of molecules, descriptors, bioinformatics and statistics. For the calculation of the molecular descriptors we used the software Dragon 7, and for statistical analysis the program Matlab. A satisfactory model with 67% precision in prediction, 80% of sensitivity and 88% specificity was achieved. The importance of this study lies in the possibility of saving time and resources in the development of new drugs and increasing drug safety by an adequate molecular pre-selection.

Translated by

Paola Guerrero



PALABRAS CLAVE

QSAR, efectos adversos, predicción, modelos matemáticos, antimicrobianos, moléculas.

LISTA DE ABREVIACIONES

MedDRA: Medical Dictionary for Regulatory Activities (diccionario médico para actividades regulatorias).

QSAR: Quantitative structure-activity relationship (relación cuantitativa estructura – actividad).

FDA: Food and Drug Administration (Administración de alimentos y fármacos)

C.A.S: Chemical Abstracts Service (servicio de resúmenes químicos)

S.M.I.L.E.S.: Simplified Molecular Input Line Entry Specification (especificación de introducción lineal molecular simplificada)

KNN: K-nearest neighbors (K vecinos más cercanos)

N3: N-Nearest Neighbours (N vecinos más cercanos).

SOC: ("System Organ Classes")

N.E.R: No error rate (porcentaje de los clasificados correctos)

INTRODUCCIÓN

El avance de los fármacos creados por las casas farmacéuticas ha conllevado un arriesgado proceso debido a la millonaria inversión, muchas veces sin garantía de que su estudio sea seguro, cumpla los intereses terapéuticos y no sea riesgoso para los pacientes el uso de los nuevos medicamentos. Así mismo existe cierta tranquilidad por los fármacos que han permanecido durante muchos años en el mercado cuyos efectos no han sido preocupantes; sin embargo, no es así sobre los nuevos medicamentos, en donde aún no se sabe bien los efectos que dichos fármacos pueden provocar por lo que hay aún un poco de expectativa. (1).

El adelanto de nuevos fármacos empieza desde las pruebas preclínicas, donde se realizan estudios a nivel molecular, con pruebas en animales; después de esta fase se procede a obtener permisos respectivos para iniciar las investigaciones clínicas. Existe así pues el estudio de fase 1, en donde el objetivo es verificar la farmacología y seguridad del fármaco; posterior a esta fase se encuentra la fase 2 en donde se caracteriza por comprobar la efectividad de los compuestos; como fase 3 que es la final es aquella que se da antes de presentar la solicitud de registro y del permiso de mercantilización del nuevo fármaco. La fase 4, o estudios post marketing se va a realizar después de la aprobación del medicamento (2).

Es importante acotar que la presentación de los efectos adversos de los medicamentos se produce en la fase 4 de investigación. Es conocido el “Caso Vioxx” de la casa farmacéutica Merck, cuando después de la comercialización del rofecoxib (nombre comercial Vioxx®), debido a su cardiotoxicidad tuvo que ser retirado del mercado, cuyos juicios penales que fueron más de 17.000 y con una millonaria deuda por más de 4.850 millones de dólares que tuvieron que ser cancelados debido a las demandas impuestas (3).

Es poca la cantidad de personas que saben sobre las serias reacciones que pueden provocar los nuevos medicamentos, es en los Estados Unidos en donde se producen cada año más de 2,74 millones de efectos adversos serios a fármacos. Existe un estimado de que alrededor de 128.000 personas mueren cada año por esta causa, siendo los fármacos la cuarta causa de muerte en conjunto con eventos cerebrovasculares (1,4).

Existe una medida preventiva para evitar problemas debido a la toxicidad de las moléculas, la cual se ha venido realizando bajo investigaciones que es mediante la

metodología QSAR (Quantitative Structure–Activity Relationship)(5–7). Esta metodología nos ha permitido predecir mediante la actividad farmacológica de las sustancias, el efecto esperado o deseado. (8–10).

En el presente trabajo de titulación se utilizó el método QSAR para poder desarrollar un modelo matemático que nos permita predecir la toxicidad en moléculas con actividad antimicrobiana.

MATERIALES Y MÉTODOS

La base de datos de los efectos secundarios de las moléculas estudiadas se obtuvo de SIDER 4.1: Side Effect Resource (11) cuya base de datos contiene miles de moléculas con sus efectos secundarios según los acápites respectivos (antibióticos, antimicóticos, antiparasitarios y antivirales). De esta base de datos extrajimos 149 moléculas, a cada una de las cuales se asoció su código S.M.I.L.E.S que es una forma de representar la molécula como secuencia de letras.

S.M.I.L.E.S. (Simplified Molecular Input Line Entry Specification) es una especificación para delinear sin ambigüedades una estructura molecular utilizando cadenas cortas ASCII (código de caracteres apoyado en el alfabeto latino). Pueden ser extraídas por editores moleculares para su conversión en dibujos bidimensionales o tridimensionales de dichas moléculas (12).

PubChem es una página que contiene una de las colecciones más grande del mundo en relación a la información química con libre acceso, en ella se puede buscar diferentes productos químicos ya sea por su nombre, estructura, forma molecular etc. En esa página podemos encontrar propiedades químicas, información de seguridad, toxicidad, actividades biológicas, etc. (13) por ejemplo: amoxicilina (Ver tabla N° 1)

El número de registro CAS (Chemical Abstracts Service) consiste en una tipificación numérica que se usa para secuencias biológicas, compuestos biológicos, aleaciones y preparados. Asigna estos identificadores a cada compuesto químico que ha sido descrito en la literatura. El CAS también tiene una base de datos de los diferentes compuestos químicos, conocida como registro CAS, compuesta por más de 123 millones de compuestos los cuales están catalogados y numerados. El propósito es permitir realizar

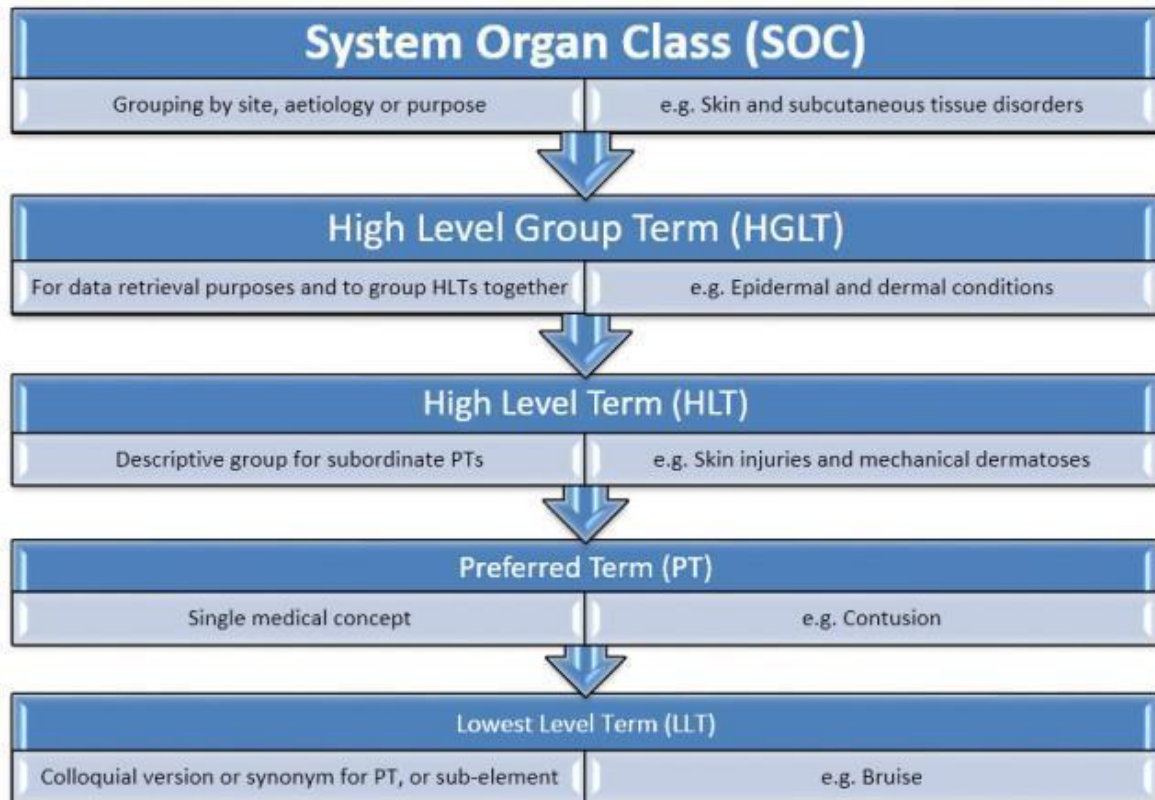
una búsqueda en dicha base de datos la cual sea unificada, actualmente, en su mayoría casi todas dichas moléculas permiten realizar una búsqueda por un número CAS. (14)

Tabla N. °1: Codificación C.A.S. y Canonical S.M.I.L.E.S.

FÁRMACO	C.A.S	S.M.I.L.E. S
Cefepime	88040-23-7	<chem>C[N+]1(CCCC1)CC2=C(N3C(C(C3=O)NC(=O)C(=NOC)C4=CSC(=N4)N)SC2)C(=O)[O-]</chem>
Amoxicillin	26787-78-0	<chem>CC1(C(N2C(S1)C(C2=O)NC(=O)C(C3=CC=C(C=C3)O)N)C(=O)O)C</chem>
Cefuroxime	55268-75-2	<chem>CON=C(C1=CC=CO1)C(=O)NC2C3N(C2=O)C(=C(CS3)COC(=O)N)C(=O)O</chem>

Los efectos adversos de los antimicrobianos dentro de SIDER constan en forma de frecuencias (porcentajes de usuarios afectados) y se describen en el formato MedDRA. De hecho, la razón de seleccionar la base de datos SIDER ha sido, porque ella aplica la nomenclatura internacional MedDRA la cual consiste en una terminología de uso médico internacional que nos ayuda en la recuperación, codificación, análisis y visualización de datos. Es aplicada en todas las fases del desarrollo de los fármacos, libre de maltrato animal, etc. (15)

MedDRA establece una clasificación de los efectos adversos en una jerarquía determinada:



Clase de sistema órgano – Término de grupo de alto nivel – Término de alto nivel –
Término preferido – Término de nivel más bajo)

Usando el software MedDRA Browser se identificaron los SOC para cada efecto secundario registrado.

Una vez realizada la clasificación de los efectos adversos en categorías SOC, se procedió a asignar un factor de ponderación: según la severidad del efecto, 10 para casos severos, 5 para casos medianos y 1 para casos leves. Esto, desde luego, es subjetivo y constituye una debilidad del estudio.

Parte de las moléculas de la base de datos SIDER no contienen información sobre efectos adversos y solo son marcados como p.m. (post mercadeo). Esto sucedió en 40 de las 149 moléculas.

Finalmente se realizó la agrupación de las moléculas y la suma de las frecuencias ponderadas por grupo SOC. Con estos datos sumados se realizó la construcción del modelo.

Tabla N. °2: Principales efectos adversos y el porcentaje según el N. ° de Fármacos.

Efecto adverso	Suma ponderada	N.º de Fármacos	
		Sin efecto adverso	Con efecto adverso
Infections and infestations	9429,01	54%	46%
General disorders and administration site conditions	9043,52	50%	50%
Infections and infestations	9029,78	63%	37%
Nervous system disorders	7618,26	41%	59%
General disorders and administration site conditions	5748,13	70%	30%
Blood and lymphatic system disorders	5270,97	63%	37%
Respiratory, thoracic and mediastinal disorders	4251,93	68%	32%
Skin and subcutaneous tissue disorders	2892,16	45%	55%

Empleando el software DRAGON 7, se calculó descriptores moleculares bidimensionales, a partir de los S.M.I.L.E.S., en total 761.

Las actividades adversas reportadas de los fármacos se dividieron en dos clases: inactiva o poco activa (clase 1) con 104 fármacos y activa (clase 2) con 45 fármacos. Esta partición se realizó utilizando como umbral inicial el octogésimo percentil de la distribución de las respuestas.

El grupo de calibración se ha usado para el desarrollo del modelo mediante dos métodos de clasificación:

- 1) Método no paramétrico KNN (16) acoplado a la selección de los descriptores significativos mediante método de los algoritmos genéticos según Leardi (17). Este método es muy simple desde un punto de vista teórico: una muestra es clasificada en función de la clase dominante de los K vecinos más próximos.
- 2) Método no paramétrico N3 (18) aplicado sobre los descriptores seleccionados previamente con KNN (ver punto 1). N3 una evolución del KNN en la cual se considera durante la clasificación, la contribución de todos los objetos en el

espacio multidimensional, ponderados por una contribución de acuerdo a la cercanía del objeto a clasificar. En este caso el parámetro alfa óptimo ha sido de 1.

Los parámetros de calidad utilizados para el modelo de clasificación y de validación son la tasa de aciertos (NER), la sensibilidad (S_n), especificidad (S_p) y precisión (Pr) de las clases (19).

La precisión representa la capacidad del modelo de clasificación de no incluir en una clase determinada elementos de otras clases:

$$Pr_g = \frac{n_{gg}}{n'_g}$$

Donde n'_g es el número total de muestras asignadas a la g-esma clase y n_{gg} es el número de elementos de la clase g asignados correctamente a la clase g

La sensibilidad describe la habilidad del modelo de clasificar los elementos de la clase g en la misma clase g

$$Sn_g = \frac{n_{gg}}{n_g}$$

Donde n_g es el número total de elementos perteneciente a la clase g.

La especificidad caracteriza la habilidad de la clase g de rechazar las muestras de todas las demás clases y se define como:

$$Sp_g = \frac{\sum_{k=1}^G (n'_k - n_{gk})}{n - n_g} \quad \text{for } k \neq g$$

Donde n'_k es el número total de elementos asignados a la k-esma clase

$$n'_k = \sum_{g=1}^G n_{gk}$$

La tasa de no aciertos (N.E.R) es el promedio de las sensibilidades de las clases:

$$NER = \frac{\sum_{g=1}^G S n_g}{G}$$

Donde G es el número total de elementos.

Consensus: para evaluar las capacidades de la integración de los métodos de clasificación utilizados, se han designado las clases obtenidas de los tres métodos de asignación mediante criterio de mayoría de voto (13)

RESULTADOS

Los algoritmos genéticos acoplados con el método KNN permitieron la selección de un modelo óptimo de dos descriptores, los cuales se detallan en la Tabla 1.

Tabla N. ° 3 Detalle de los descriptores moleculares para el modelo KNN obtenido por algoritmos genéticos.

No	Descriptor	Bloque	Descripción
1	nStructures	Constitucional	número de estructuras desconectadas
2	Autovalor de la Matriz de Burden pesada por las masas	Autovalores de la matriz de Burden	Autovalor mas grande

Los algoritmos genéticos acoplados con el método N3 permitieron la selección de un modelo óptimo de siete descriptores, los cuales se detallan en la Tabla 2.

Tabla N. ° 4 Detalle de los descriptores moleculares para el modelo N3 obtenido por algoritmos genéticos.

No.	Descriptor	Bloque	Descripción
1	nR04	Basic descriptors	number of 4-membered rings
2	MATS5v	2D autocorrelations	Moran autocorrelation of lag 5 weighted by van der Waals volumen
3	SM3_D/Dt	Distance/detour matrix (D/Dt)	spectral moment of order 3 from distance/detour matrix
4	MATS2e	Moran autocorrelations	Moran autocorrelation of lag 2 weighted by Sanderson electronegativity
5	CATS2D_05_DN	Basic descriptors	CATS2D Donor-Negative at lag 05
6	CATS2D_03_PL	Basic descriptors	CATS2D Positive-Lipophilic at lag 03
7	B04[O-O]	Binary Atom Pairs of order 4	Presence/absence of O - O at topological distance 4

Los parámetros de calidad de los dos modelos sea en clasificación que en predicción se reportan en la tabla N. ° 3.

Tabla N. ° 5. Resumen de los parámetros de calidad de los modelos obtenidos.

Modelo	NER	Clase 1			Clase 2		
		Precisión	Sensibilidad	Especificidad	Precisión	Sensibilidad	Especificidad
KNN clasificación	0.70	0.88	0.93	0.48	0.63	0.48	0.93
KNN predicción	0.62	0.85	0.8	0.44	0.36	0.44	0.80
N3 clasificación	0.67	0.89	0.68	0.67	0.34	0.67	0.68
N3 predicción	0.40	0.74	0.57	0.22	0.12	0.22	0.57

De la tabla N. ° 5 se nota, en general, que el mejor modelo es el de KNN, sin embargo, en predicción la sensibilidad de la clase 2 es mejor con el método N3.

Por esta razón se ha procedido a realizar un “consensus” entre las respuestas originales, las calculadas por KNN y las calculadas por N3 y recalculando con KNN las clasificaciones y predicciones.

Los resultados se presentan en la tabla N. ° 6

Tabla N. ° 6 Resumen de los parámetros de calidad del modelo de “consensus” obtenido

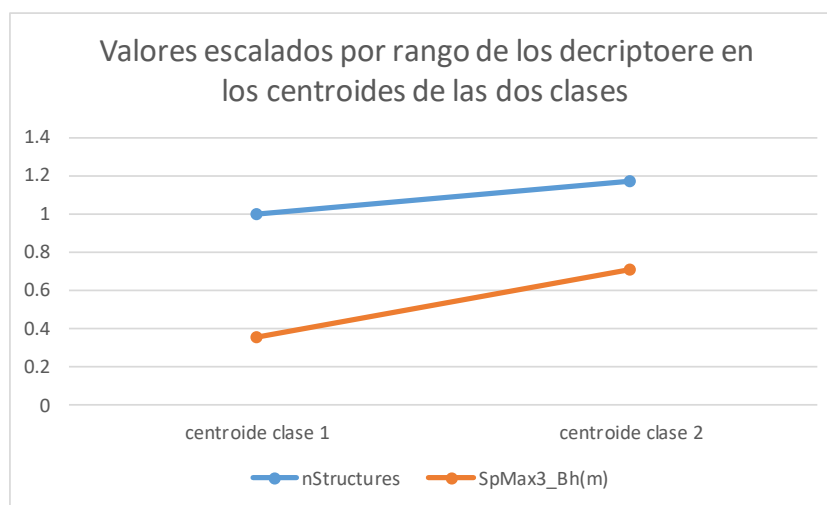
Modelo	NER	Clase 1			Clase 2		
		Precisión	Sensibilidad	Especificidad	Precisión	Sensibilidad	Especificidad
KNN clasificación consensus	0.77	0.92	0.93	0.61	0.65	0.93	0.61
KNN predicción consensus	0.84	0.94	0.88	0.80	0.67	0.80	0.88

De la tabla N. ° 6 se desprende que el modelo de consensus es muy bueno para predecir efectos adversos mientras que en cada uno de los otros se tiene siempre una clase no bien clasificada y predicha.

DISCUSIÓN

Para comprender la importancia de los descriptores en la predicción de la clase de pertenencia, se han calculado los valores centrales de los mismos (Estudentizados) en cada clase.

Figura N. ° 1: gráfico de los centroides de cada descriptor en cada clase.



De la gráfica se nota inmediatamente que los fármacos de la clase 1 (menos efectos adversos) tienen medianamente valores de los descriptores nStructure y SpMax3_B(m) más altos que los fármacos de la clase más activa.

Este resultado puede ser interpretado de la siguiente manera: las moléculas de la clase menos activa (clase 1) presentan un número inferior de grupos (estructuras) no conectados por caminos que las de la clase más activa (clase 2) indicando una menor conexión entre grupos de átomos en la molécula que las de la clase 2.

Además, el descriptor SpMax3_B(m) indica que las moléculas de la clase 1 son menos grandes que las de la clase 2.

En síntesis, la mayor frecuencia de las denuncias y gravedad de efectos secundarios aparece conectada por la complejidad molecular del fármaco.

La importancia de este trabajo está en el hecho que para predecir la clase de pertenencia de nuevos fármacos es suficiente ponerlo en el test de validación y aplicar la predicción utilizando como base el training set actual.

Dentro de las fortalezas de nuestro estudio se encuentran:

- No se necesita personas de prueba
- Fácil de aplicar
- Es un método nuevo

Debilidades con el estudio:

- No es fácil de encontrar estudios comparativos
- La ponderación es muy subjetiva
- Puede existir un sesgo debido a que la base de datos está conformada por datos reportados.

CONCLUSIONES

La metodología QSAR permite predecir, con un grado de certeza satisfactoria, los posibles efectos secundarios que una molécula puede tener, usando como base para la predicción la estructura molecular. Este hecho puede ahorrar costos millonarios en el desarrollo farmacológico, acortar el tiempo en la preselección de las moléculas; se considera que solo 250 de cada 10.000 moléculas de la etapa de investigación básica se aprueban y se utilizan en la fase de investigación preclínica (20). Seguramente contribuirá

a una mayor seguridad farmacológica, pues en muchas ocasiones los efectos secundarios recién se hacen manifiestos en la fase del mercadeo del fármaco. Un notable ejemplo es el caso de Vioxx (rofecoxib) que causó decenas de miles de muertes (el número exacto hasta hoy en día permanece desconocido) y juicios billonarios (sin mencionar la tragedia humana que ha provocado).

Los sistemas de predicción son definidos como técnicas que permiten construir y estudiar nuevas suposiciones, a través del machine learning, llamado también aprendizaje automático. Es un campo magnífico, rama de la inteligencia artificial, la cual permite desarrollar técnicas para que las computadoras aprendan. De manera más objetiva, los científicos buscan algoritmos para convertir ciertos datos en programas de ordenador, sin la necesidad de escribirlos, éstos a su vez son capaces de permitir comportamientos y deducciones para un conjunto más grande de datos. Este método está también relacionado con el reconocimiento de patrones, por lo que puede ser visto como un experimento de automatizar partes del procedimiento científico mediante ciertos métodos matemáticos, por lo tanto se ha convertido en un proceso de inducción del conocimiento. (21)

AGRADECIMIENTO

Agradecemos a nuestros padres y demás familiares por su apoyo absoluto durante nuestra carrera universitaria, a nuestros docentes por todos los conocimientos adquiridos y a nuestro tutor de tesis por su ayuda desinteresada en la realización de nuestro proyecto.

BIBLIOGRAFIA

1. Light DW. [En línea]. 27 junio 2014. New Prescription Drugs: A Major Health Risk With Few Offsetting Advantages [citado 4 noviembre 2019]. Disponible: <https://ethics.harvard.edu/blog/new-prescription-drugs-major-health-risk-few-offsetting-advantages>.
2. Lipsky MS, Sharp LK. From Idea to Market: The Drug Approval Process. *J Am Board Fam Pract*. 2001;14:362-7.
3. Berenson A. Merck Agrees to Settle Vioxx Suits for \$4.85 Billion. *The New York Times* [En línea]. 9 noviembre 2007 [citado 31 marzo 2019]. Disponible: <https://www.nytimes.com/2007/11/09/business/09merck.html>
4. Schroeder MO. *US News & World Report* [En línea]. 27 septiembre 2016. Death By Prescription [citado 4 noviembre 2019]. Disponible: <https://health.usnews.com/health-news/patient-advice/articles/2016-09-27/the-danger-in-taking-prescribed-medications>
5. Roy K, Kar S, Das RN. *Understanding the Basics of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment*. Londres: Academic Press; 2015. DOI: 10.1016/C2014-0-00286-9
6. Kubinyi H. QSAR and 3D QSAR in drug design Part 1: methodology. *Drug Discovery Today*. 1997;2(11):457-67. DOI: 10.1016/S1359-6446(97)01079-9
7. Kubinyi H. QSAR and 3D QSAR in drug design Part 2: applications and problems. *Drug Discovery Today*. 1997;2(12):538-46. DOI: 10.1016/S1359-6446(97)01084-2
8. Du Q-S, Huang R-B, Chou K-C. Recent Advances in QSAR and Their Applications in Predicting the Activities of Chemical Molecules, Peptides and Proteins for Drug Design. *Current Protein & Peptide Science*. 2008;9(3):248-59. DOI: 10.2174/138920308784534005

9. Ghose AK, Viswanadhan VN, editores. Combinatorial library design and evaluation: principles, software tools, and applications in drug discovery. New York: M. Dekker; 2001.
10. Verma J, Khedkar V, Coutinho E. 3D-QSAR in Drug Design - A Review. Current Topics in Medicinal Chemistry. 2010;10(1):95-115. DOI: 10.2174/156802610790232260
11. Kuhn M, Letunic I, Jensen LJ, Bork P. The SIDER database of drugs and side effects. [Internet]. SIDER 4.1: Side Effect Resource. 2015. [citado el 29 de octubre de 2020]. Oct 21. Disponible en: <http://sideeffects.embl.de/>
12. Anderson E. Veith G.D. Weininger D. SMILES: A line notation and computerized interpreter for chemical structures. 1987. Report No. EPA/600/M-87/021. U.S. EPA, Environmental Research Laboratory-Duluth, Duluth, MN 55804
13. National Center of Biotechnology Information. What is PubChem?. [Internet]. PubChem. 2018. [Citado el 29 de octubre de 2020]. Disponible en: <https://pubchem.ncbi.nlm.nih.gov/>
14. Número de registro CAS. [Internet]. Wikipedia, La enciclopedia libre. 2020. [Actualizado el 27 de octubre de 2020]; [Citado el 04 de noviembre de 2020]. Disponible en: https://es.wikipedia.org/wiki/N%C3%BAmero_de_registro_CAS.
15. Whetzel PL, Noy NF, Shah NH, Alexander PR, Nyulas C, Tudorache T, Musen MA. Medical Dictionary for Regulatory Activities Terminology (MedDRA). BioPortal. 2014 [Actualizado el 24 de septiembre de 2020]; [Citado el 04 de noviembre de 2020]. Disponible en: https://bioportal.bioontology.org/ontologies/MEDDRA?fbclid=IwAR3Q9Y5Y5I_aINfBUvEt-NDCPOK7iiNhS0I8UqvVuVNx7jH928xdsxBDSHk
16. Cover T, Hart P. Nearest Neighbor Pattern Classification. In: IEEE TRANSACTIONS ON INFORMATION THEORY [Internet]. Stanford; 2012. [Citado el 29 de octubre de 2020]. Available from: <https://sci-hub.do/10.1109/TIT.1967.1053964>
17. Leardi R. Lupiañez González A. Genetic algorithms applied to feature selection in PLS regression: how and when to use them. Chemometrics and Intelligent Laboratory Systems. 1998.
18. Todeschini R. Ballabio D. Cassotti M. Consonni V. N3 and BNN: Two new

similarity based classification methods in comparison with other classifier. [Internet]. J. Chem. Inf. Model. (2015) [Citado el 29 de octubre de 2020]. Disponible en: <https://doi.org/10.1021/acs.jcim.5b00326>

19. Ballabio D, Consonni V, Todeschini R. The Kohonen and CP-ANN toolbox: a collection of MATLAB modules for Self Organizing Maps and Counter Propagation Artificial Neural Networks. Chemometrics and Intelligent Laboratory Systems. 2009.

20. MSD Salud. Proceso de investigación, desarrollo y aprobación de un fármaco. [Internet]. MSD Salud. 2019. [Citado el 4 de noviembre de 2020]. Disponible en: <https://www.msdsalud.es/informacion-practica/proceso-investigacion-farmaco.html>

21. Aprendizaje automático. [Internet]. Wikipedia, La enciclopedia libre. 2020. [Actualizado el 15 de septiembre de 2020]; [Citado el 04 de noviembre de 2020]. Disponible en: https://es.wikipedia.org/wiki/Aprendizaje_autom%C3%A1tico