



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
FACULTAD DE MEDICINA
SECRETARIA DE SALUD
INSTITUTO NACIONAL DE PEDIATRÍA
CURSO DE ESPECIALIDAD EN PEDIATRÍA MÉDICA**

**APRENDIZAJE AUTOMÁTICO PARA ASISTIR EL DIAGNÓSTICO CLÍNICO
DE LAS INMUNODEFICIENCIAS PRIMARIAS: EL PROYECTO MYCROFT**

TESIS

PARA OBTENER EL TÍTULO DE :

MÉDICO ESPECIALISTA EN PEDIATRÍA MÉDICA

PRESENTA :

DRA. GUADALUPE MALDONADO COLÍN

PROFESOR TITULAR:

DR. ALEJANDRO SERRANO SIERRA

ASESOR DE TESIS:

SAÚL OSWALDO LUGO REYES

CO- TUTORES:

JOSÉ LUIS MORALES

CHIHARU MURATA

México, D. F. 2014



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

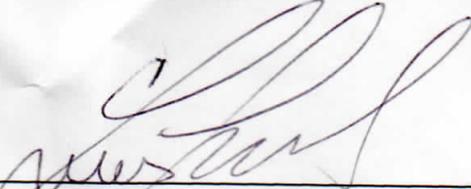
**APRENDIZAJE AUTOMÁTICO PARA ASISTIR EL DIAGNÓSTICO CLÍNICO
DE LAS INMUNODEFICIENCIAS PRIMARIAS: EL PROYECTO MYCROFT**



**DR. ALEJANDRO SERRANO SIERRA
PROFESOR TITULAR DEL CURSO DE
ESPECIALIZACION EN PEDIATRIA**



**DRA. ROSAURA ROSAS VARGAS
DIRECTORA DE ENSEÑANZA**



**DR. LUIS MARTIN GARRIDO GARCIA
JEFE DEL DEPARTAMENTO DE PRE Y POSGRADO**



**SAÚL OSWALDO LUGO REYES
TUTOR DE TESIS**

RESUMEN

Las inmunodeficiencias primarias (IDPs) son un grupo de enfermedades congénitas cuya incidencia global es de 1:10,000, siendo su abordaje complejo y su diagnóstico diferencial amplio. La expectativa de vida sin tratamiento puede ser tan breve como 1 a 2 años, o incapacidad y muerte antes de la adolescencia. Por esto es esencial para el sistema de salud su diagnóstico y tratamiento oportunos. El proyecto Mycroft es un proyecto basado en técnicas de AA que busca asistir en el diagnóstico clínico de las IDPs. Las variables ingresadas incluirán: edad de presentación, aislamientos, sitios de infección, manifestaciones asociadas, complicaciones y resultados de laboratorio. A partir de la base de datos de entrenamiento se creará un algoritmo de toma de decisiones con diferentes metodologías para extraer conocimiento a partir de la información disponible, generando ecuaciones para clasificar correctamente los nuevos casos presentados y para que aprenda con cada utilización incorporando la información de éstos. La meta final es un software con una interface amigable al usuario clínico para obtener un diagnóstico diferencial y un abordaje diagnóstico sugerido.

La inteligencia artificial se define como “el campo de la ciencia y la ingeniería que se ocupa del entendimiento computacional de lo que comúnmente se conoce como comportamiento inteligente, y de la creación de artefactos que exhiban tal comportamiento.”(1)

El Aprendizaje automático (*machine learning*) es una rama de la Inteligencia artificial, que trata de la construcción y estudio de sistemas que puedan aprender de los datos.

En el núcleo del aprendizaje automatizado están la Representación y la Generalización. Se representan los datos ingresados en base a funciones evaluadas en la instancia de datos. La Generalización consiste en que el sistema se desempeñe bien frente a instancias de datos nuevos. Es la habilidad de un algoritmo para desempeñarse con exactitud frente a ejemplos nuevos, luego de

haberse entrenado en un conjunto de datos de aprendizaje. Su objetivo principal es generalizar a partir de la experiencia. Extraer de la distribución de datos de entrenamiento algo, una función que le permita hacer predicciones útiles en casos nuevos.(3)

La Minería de Datos (*Data Mining*) consiste en descubrir patrones en bases de datos, descubrir propiedades previamente ignoradas de los datos. El aprendizaje automático utiliza minería de datos como un paso previo para mejorar la exactitud de predicción, o como “aprendizaje no supervisado”. Cada vez hay más aplicaciones para el AA, y es un campo en constante expansión.

El análisis predictivo y el aprendizaje automático en Medicina.

El proceso de diagnóstico clínico es un proceso cognitivo complejo que involucra entrenamiento, experiencia, reconocimiento de patrones, aplicación de probabilidad condicional, entre otros componentes menos comprendidos.(4) Los médicos veteranos adquieren con el tiempo una intuición que facilita la tarea, producto de la experiencia de haber visto cientos de casos similares a lo largo de su carrera(5). Los clínicos menos experimentados pueden reducir sus errores diagnósticos al hacer consciente y sistemático el proceso (*meta-cognición*), teniendo en cuenta qué tan confiables son sus apoyos diagnósticos (probabilidad pre-prueba, valor predictivo positivo y negativo, razones de verosimilitud) y cuáles son las fuentes más frecuentes de error (fatiga, sesgos, etc.)(6, 7).

El contar con una ayuda electrónica o computacional facilita el proceso diagnóstico, ya que lo hace sistemático y completo, no dependiente de la experiencia o del estado actual del clínico; y elimina la necesidad de una memoria operacional. Bien elaborado, un programa diagnóstico puede ser tan efectivo como un clínico experto, y facilitar grandemente el trabajo de los médicos humanos.(14, 15)

Algunos ejemplos

La Medicina moderna enfrenta el reto de adquirir, analizar y aplicar una gran cantidad de conocimiento necesario para resolver problemas clínicos complejos.(16) Se han desarrollado programas para asistir al clínico en formular un diagnóstico, hacer decisiones terapéuticas, y predecir el pronóstico o resultado final, así como para asistir el aprendizaje médico a partir de la década de 1980 (16, 17). Algunos de los abordajes más utilizados, que han demostrado mayor utilidad incluyen: a) redes neuronales artificiales (ANN)(18-20); b) métodos de razonamiento basado en casos (CBR)(21-23) ; y c) clasificadores bayesianos (BC)(24).

Dietzel y colaboradores en 2011 utilizaron ANN para procesar una base de datos clínica, con el objetivo de clasificar lesiones detectadas mediante mastografía por resonancia artificial (MRM) y predecir malignidad a partir de la arquitectura de la lesión.(20) En un conjunto de datos de 436 lesiones benignas y 648 malignas (clasificadas por histopatología), con 4 abordajes y 18 parámetros a evaluar, encontraron una exactitud diagnóstica elevada (media del área bajo la curva 0.888, $p < 0.001$) y un desempeño robusto (0.885-0.892).

Las ANN han sido utilizadas en el diagnóstico clínico, en el análisis de imágenes radiográficas e histopatológicas, en la interpretación de datos usados en terapia intensiva, entre otros campos de la medicina.¹ Entre los diversos campos de aplicación se encuentra el de la cirugía, algunas otras aplicaciones relevantes incluyen dolor abdominal, glaucoma, lumbalgia y litiasis vesical. Ömer et al. Comprobaron la utilidad de las ANN en el diagnóstico de la apendicitis obteniendo una sensibilidad, una especificidad y un valor predictivo negativo de 100%, 97.2%, 88% y 100% respectivamente. ² demostrando así que esta aplicación es una herramienta diagnóstica especialmente en áreas donde el USG y la TAC no son viables.

En 2011 Tenorio et al utilizaron un sistema de apoyo en decisión clínica para el diagnóstico de la enfermedad celiaca, con apoyo de inteligencia artificial,

encontrando una sensibilidad del 92.9% y una especificidad del 79.2% en comparación con el estándar de referencia, la biopsia.³

El aprendizaje automático también ha sido llevada al campo de la cardiología, Atkov et al. Desarrollaron un proyecto donde el objetivo fue desarrollar redes neuronales artificiales como modelo diagnóstico de las coronariopatías usando factores de riesgo genéticos y no genéticos de estas patologías. Los autores lograron crear modelos con una exactitud del 64-94%, demostrando así la utilidad de estos métodos.⁵

El razonamiento basado en casos (RBC) es un método de abordaje de inteligencia artificial que se basa en experiencias pasadas para resolver problemas actuales. Se utiliza principalmente en ciencias de la salud. La investigación actual está marcada por la riqueza de un amplio rango de aplicaciones y su integración con numerosos abordajes complementarios y tecnología. Las oportunidades abundan, incluyen nuevas contribuciones para la bioinformática, la formalización de razonamiento basado en casos en biomedicina, y por último el desarrollo de mejores sistemas diseñados para abarcar la complejidad de la biomedicina con integración en escenarios clínicos.⁶

Las inmunodeficiencias primarias

Las inmunodeficiencias primarias (IDP) son un grupo de enfermedades congénitas raras, cuya incidencia va de 1:500 a 1:500,000 aproximadamente (la incidencia global aceptada es de 1:10,000), de presentación clínica variable, con más de 170 etiologías genéticas conocidas y al menos 206 fenotipos o enfermedades identificadas.

Las IDP son diversas, con manifestaciones clínicas y severidad variable. Se clasifican en 8 grupos, de acuerdo a la rama de la inmunidad que está afectada: Defectos combinados, Defectos de Anticuerpos, Defectos bien definidos, Defectos

de la fagocitosis, Defectos de la inmunidad innata, Defectos de desregulación inmunológica, Defectos del Complemento, y Enfermedades Autoinflamatorias.

Lo que dificulta el diagnóstico de las IDP es: a) la falta de familiaridad de la mayoría de los clínicos de primer contacto; b) el número creciente, siempre en expansión, de nuevos defectos identificados; c) el vasto número de posibilidades, a la fecha en más de 200, que exige un conocimiento profundo, una gran memoria operacional, y un olfato clínico aguzado. El proceso diagnóstico puede ser complejo aún para quienes se dedican de tiempo completo a investigar inmunodeficiencias, que son muy pocos. En México, un número muy reducido de inmunólogos pediatras se puede considerar experto en el abordaje diagnóstico de las IDP, casi todos ellos concentrados en la Ciudad de México y en el INP.

El proyecto Mycroft

El proyecto Mycroft es un trabajo que pretende asistir en el diagnóstico clínico de las inmunodeficiencias primarias y, eventualmente, para predecir correctamente la probabilidad de encontrar un defecto congénito en niños referidos a la Unidad de Investigación en Inmunodeficiencias por una historia de infecciones recurrentes.

Un sistema basado en técnicas de aprendizaje automático (AA) mejoraría sustancialmente la efectividad en el proceso de realizar un diagnóstico diferencial. El uso de esta clase de sistemas está ampliamente documentado en la literatura especializada. Entre las ventajas aceptadas se encuentran las siguientes: a) reducen la necesidad de experiencia, familiaridad y memoria operacional en el usuario clínico; b) ofrecen un diagnóstico diferencial basado en las variables clínicas observadas en un paciente.

Las variables ingresadas incluirán: edad de presentación, aislamientos, sitios de infección, manifestaciones asociadas, complicaciones y resultados de laboratorio. Y como resultado el sistema ofrecerá una lista de diagnóstico diferencial y un abordaje diagnóstico-terapéutico sugerido.

Además de las ventajas señaladas para los sistemas automatizados (eliminar sesgos de disponibilidad y confirmación, eliminar la necesidad de experiencia y entrenamiento en IDP, prescindir de la memoria operacional del cerebro humano), un sistema basado en aprendizaje automático es capaz de incorporar, previa evaluación, casos nuevos en el modelo de clasificación. El mecanismo consiste en recalcular el modelo de clasificación considerando los casos nuevos y las posibles modificaciones que se hayan hecho en los diagnósticos originales. Esta capacidad es crucial para mantener la funcionalidad y actualidad del sistema considerando los avances y nuevos descubrimientos en el área de IDP.

En la Universidad de Tampere, Finlandia, se está desarrollando un abordaje de IDP asistido mediante un Sistema de conocimiento, una máquina de inferencia, y la interface del usuario, funcional en fase de prototipo, con 708 casos clínicos y 148 atributos o parámetros incluidos. Este grupo cuenta con una amplia experiencia en bioinformática aplicada a IDP (www.bioinf.uta.fi). Su planteamiento es desarrollar un Sistema Experto asociado a CBR (*Nearest Neighbour, leave-one-out strategy*). Resultados preliminares reportan 469 casos correctamente clasificados, una exactitud de 66%, con una mediana de exactitud de clasificación por clase (*classwise*) del 50% (Samarghitean et al, ESID 15th biennial meeting 2012, Florencia, Italia).

Esta es la única iniciativa conocida por los autores, aún sin datos publicados, que realiza un abordaje comparable al propuesto. Por lo demás, hay varios sitios web que se ocupan de almacenar, registrar y difundir el conocimiento de las inmunodeficiencias primarias en línea, incluyendo: www.immunodeficiencysearch.com (EUA), <http://www.ncbi.nlm.nih.gov/omim> (EUA, Online Mendelian Inheritance in Man), <http://rapid.rci.riken.jp/RAPID> (Japón), <http://bioinf.uta.fi/idr/index.shtml> (Finlandia); ninguno de los cuales ofrece un diagnóstico clínico asistido mediante aprendizaje automático.

Este trabajo es alentador, porque sugiere un interés compartido y una factibilidad, aunque sus primeros resultados no son aceptables, lo que sugiere que hay campo para crecimiento e investigación. El Sistema Experto utilizado por este grupo, por

definición no es inteligente, y el Razonamiento Basado en Casos empleado como metodología no es un buen ejemplo de aprendizaje automatizado, dado que echa mano de todo el conjunto de datos en cada utilización, y su algoritmo no se modifica; simplemente obtiene un cociente de qué tanto coincide el ejemplo nuevo con todos los ejemplos previamente registrados.

Se puede emplear el conocimiento acumulado en miles de casos de pacientes con inmunodeficiencias primarias publicados en los últimos 50 años, y en registros detallados alrededor del mundo (Australia, Asia, Europa, América), para construir una base de datos con atributos para cada enfermedad, que sirva como entrenamiento a la máquina para aprender a clasificar correctamente nuevos pacientes.

A partir de la base de datos de entrenamiento se crea un algoritmo de toma de decisiones con diferentes metodologías, incluyendo regresión logística multivariante, máquinas de soporte vectorial, vecino más cercano, clasificadores bayesianos ingenuos, y sistema inmune artificial, que extraiga conocimiento a partir de la información disponible, genere ecuaciones para clasificar correctamente los nuevos casos presentados, y que aprenda con cada utilización para incorporar la información de los nuevos casos y volverse más robusto con el tiempo.

El ensamble de varios abordajes o modelos garantiza un mejor resultado final, y el desempeño del nuevo modelo se puede evaluar al ponerlo a prueba con una porción de la base de datos que se reserve para este fin. El resultado del proceso será un diagnóstico diferencial, una lista de mayor a menor probabilidad con las tres Inmunodeficiencias primarias más probables. Junto a cada nombre de la enfermedad se incluirá la etiología genética (*CYBB*, *BTK*, *DOCK8*, etc.) y un abordaje diagnóstico propuesto (ej, "Realizar 123-DHR" o "Detectar expresión de la proteína WASP"). La meta final es una aplicación o software con una interface amigable al usuario clínico, que pueda ingresar sus datos clínicos y de laboratorio, y obtener un diagnóstico diferencial y un abordaje diagnóstico sugerido.

MATERIAL Y METODOS

TIPO DE ESTUDIO: Observacional descriptivo, ambispectivo y transversal.

CATEGORIA: Desarrollo tecnológico e innovación.

METODOS:

I. Conjunto de datos IDP (Base de Conocimiento).

Se revisará los expedientes de pacientes del INP con diagnóstico conocido de IDP para construir una base de datos (Microsoft Excel 2007) con las variables a estudiar (Apéndice A). Aproximadamente 200 atributos, incluyendo por ejemplo: sexo, edad de inicio, infecciones, susceptibilidad, complicaciones, niveles séricos de inmunoglobulinas.

Paralelamente, se construirá una base de datos con información clínica conocida, de laboratorio y genética para las distintas Inmunodeficiencias primarias, con el fin de contar con pacientes “arquetípicos” para cada uno de los genes conocidos.

Las bases de datos se combinarán en una sola y se nutrirán con pacientes propios recién identificados y con pacientes reportados en la literatura médica internacional. Además se buscará acceder a los registros internacionales de pacientes en Europa, Estados Unidos y Latinoamérica, para nutrir y hacer más robusta la base de datos final.

II. Algoritmo de Aprendizaje Automático

Con los datos clínicos, de laboratorio y genéticos se diseñará un algoritmo de aprendizaje automático que otorgue más peso a determinadas variables o atributos. Dicho peso relativo se asignará mediante regresión logística.

Se empleará la plataforma R, una plataforma de programación de acceso libre descargable en línea (www.r-project.org), para desarrollar el modelo de Aprendizaje automático a partir de la base de datos.

Se comparará el desempeño de al menos tres métodos y sus combinaciones: CBR (NN), MSV, y RLR. La herramienta diagnóstica final se construirá con la combinación de los dos métodos más útiles.

III. DIAGNOSTICO (OUTPUT)

El resultado del proceso será un diagnóstico diferencial, una lista de mayor a menor probabilidad con las tres Inmunodeficiencias primarias más probables. Junto a cada nombre de la enfermedad se incluirá la etiología genética (*CYBB*, *BTK*, *DOCK8*, etc.) y un abordaje diagnóstico propuesto (ej, “Realizar 123-DHR” o “Detectar expresión de la proteína WASP”).

La meta final es una aplicación o software con una interface amigable al usuario clínico, que pueda ingresar sus datos clínicos y de laboratorio, y obtener un diagnóstico diferencial y un abordaje diagnóstico sugerido.

Inicialmente se hará un estudio piloto con una base de datos existente que incluye los pacientes con IDP conocidos de la Unidad de Investigación en Inmunodeficiencias. En etapas intermedias se desarrollarán prototipos de software, que se pondrán a prueba con datos de pacientes conocidos y pacientes de nuevo ingreso, para evaluar su desempeño. Después de desarrollado el sistema, y antes de proponerlo como herramienta diagnóstica, se someterá a un estudio de utilidad diagnóstica con pacientes nuevos (materia de un segundo proyecto).

ANALISIS ESTADISTICO

Se describirá la forma de distribución de las variables cuantitativas, y Estadística descriptiva por medio de los parámetros de Medidas de Tendencia central y Dispersión adecuados para su forma de distribución. La compatibilidad con el modelo normal se evaluará por medio de la prueba de Bondad de ajuste de Shapiro-Wilk. Con las variables categóricas se determinará las frecuencias absolutas y relativas.

Para la construcción de Modelos, los casos clínicos conocidos con mutaciones identificadas se considerarán como diagnóstico definitivo y se utilizarán como estándar de referencia.

La utilidad diagnóstica de cada uno de los modelos se determinará por medio de la Sensibilidad, Especificidad, Valor predictivo positivo, Valor predictivo negativo, y Razón de Verosimilitud.

Para la comparación entre los modelos se evaluará su desempeño con los mismos parámetros de utilidad diagnóstica.

Para todas las pruebas estadísticas la Significancia se reconoce al nivel de $\alpha < 0.05$. Para la construcción del modelo logístico y discriminante se filtran las variables al nivel de $\alpha < 0.2$, y los estimadores puntuales de parámetros de interés se reportarán combinando con su intervalo de confianza (IC) de 95%. Los análisis estadísticos y construcción de algoritmo se realizarán mediante el uso de R (distribución libre) y JMP10 (SAS Institute, Inc.)

REFERENCIAS

1. Shapiro SC. Artificial Intelligence. 2nd ed. Shapiro SC, editor. New York: Wiley; 1992.
2. Mitchell T. Machine learning: McGraw Hill; 1997.
3. Machine learning. en.wikipedia.org: Wikipedia; [cited 2013 Feb. 22]; Available from: http://en.wikipedia.org/wiki/Machine_learning.
4. Segal M. How doctors think, and how software can help avoid cognitive errors in diagnosis. Acta Paediatr. 2007;96(12):1720-2. Epub 2007/09/14.
5. Groves M, O'Rourke P, Alexander H. The clinical reasoning characteristics of diagnostic experts. Medical teacher. 2003;25(3):308-13. Epub 2003/07/26.
6. Sherbino J, Dore KL, Siu E, Norman GR. The effectiveness of cognitive forcing strategies to decrease diagnostic error: an exploratory study. Teaching and learning in medicine. 2011;23(1):78-84. Epub 2011/01/18.
7. Gill CJ, Sabin L, Schmid CH. Why clinicians are natural bayesians. BMJ. 2005;330(7499):1080-3. Epub 2005/05/10.
8. Ebell MH, Smith MA, Barry HC, Ives K, Carey M. The rational clinical examination. Does this patient have strep throat? JAMA : the journal of the American Medical Association. 2000;284(22):2912-8. Epub 2001/01/09.
9. Pines JM. Profiles in patient safety: confirmation bias in emergency medicine. Academic emergency medicine : official journal of the Society for Academic Emergency Medicine. 2006;13(1):90-4. Epub 2005/12/21.
10. Croskerry P. The importance of cognitive errors in diagnosis and strategies to minimize them. Academic medicine : journal of the Association of American Medical Colleges. 2003;78(8):775-80. Epub 2003/08/14.
11. Croskerry P. Cognitive forcing strategies in clinical decisionmaking. Annals of emergency medicine. 2003;41(1):110-20. Epub 2003/01/07.
12. Groves M, O'Rourke P, Alexander H. Clinical reasoning: the relative contribution of identification, interpretation and hypothesis errors to misdiagnosis. Medical teacher. 2003;25(6):621-5. Epub 2004/09/17.
13. Leblanc VR, Brooks LR, Norman GR. Believing is seeing: the influence of a diagnostic hypothesis on the interpretation of clinical features. Academic medicine :

journal of the Association of American Medical Colleges. 2002;77(10 Suppl):S67-9.
Epub 2002/10/16.

14. Baxt WG, Shofer FS, Sites FD, Hollander JE. A neural network aid for the early diagnosis of cardiac ischemia in patients presenting to the emergency department with chest pain. *Annals of emergency medicine*. 2002;40(6):575-83.
Epub 2002/11/26.

15. Prabhudesai SG, Gould S, Rekhraj S, Tekkis PP, Glazer G, Ziprin P. Artificial neural networks: useful aid in diagnosing acute appendicitis. *World journal of surgery*. 2008;32(2):305-9; discussion 10-1. Epub 2007/11/29.

16. Ramesh AN, Kambhampati C, Monson JR, Drew PJ. Artificial intelligence in medicine. *Annals of the Royal College of Surgeons of England*. 2004;86(5):334-8.
Epub 2004/08/31.

17. Patel VL, Shortliffe EH, Stefanelli M, Szolovits P, Berthold MR, Bellazzi R, et al. The coming of age of artificial intelligence in medicine. *Artificial intelligence in medicine*. 2009;46(1):5-17. Epub 2008/09/16.

18. Atkov OY, Gorokhova SG, Sboev AG, Generozov EV, Muraseyeva EV, Moroshkina SY, et al. Coronary heart disease diagnosis by artificial neural networks including genetic polymorphisms and clinical parameters. *Journal of cardiology*. 2012;59(2):190-4. Epub 2012/01/06.

19. Yoldas O, Tez M, Karaca T. Artificial neural networks in the diagnosis of acute appendicitis. *The American journal of emergency medicine*. 2011. Epub 2011/09/13.

20. Dietzel M, Baltzer PA, Dietzel A, Zoubi R, Groschel T, Burmeister HP, et al. Artificial Neural Networks for differential diagnosis of breast lesions in MR-Mammography: A systematic approach addressing the influence of network architecture on diagnostic performance using a large clinical database. *European journal of radiology*. 2011. Epub 2011/04/05.

21. Chuang CL. Case-based reasoning support for liver disease diagnosis. *Artificial intelligence in medicine*. 2011;53(1):15-23. Epub 2011/07/16.

22. Park YJ, Chun SH, Kim BC. Cost-sensitive case-based reasoning using a genetic algorithm: application to medical diagnosis. *Artificial intelligence in medicine*. 2011;51(2):133-45. Epub 2011/01/11.
23. Huang ML, Hung YH, Lee WM, Li RK, Wang TH. Usage of case-based reasoning, neural network and adaptive neuro-fuzzy inference system classification techniques in breast cancer dataset classification diagnosis. *Journal of medical systems*. 2012;36(2):407-14. Epub 2010/08/13.
24. Tenorio JM, Hummel AD, Cohrs FM, Sdepanian VL, Pisa IT, de Fatima Marin H. Artificial intelligence techniques applied to the development of a decision-support system for diagnosing celiac disease. *International journal of medical informatics*. 2011;80(11):793-802. Epub 2011/09/16.
25. Casanova JL, Abel L. Primary immunodeficiencies: a field in its infancy. *Science*. 2007;317(5838):617-9. Epub 2007/08/04.
26. Notarangelo LD, Casanova JL. Primary immunodeficiencies: increasing market share. *Current opinion in immunology*. 2009;21(5):461-5. Epub 2009/10/03.
27. Casanova JL, Fieschi C, Zhang SY, Abel L. Revisiting human primary immunodeficiencies. *Journal of internal medicine*. 2008;264(2):115-27. Epub 2008/06/12.