



REVISIÓN

Revisión sistemática de la aplicación de algoritmos de «machine learning» en la esclerosis múltiple

M. Vázquez-Marrufo^{a,*}, E. Sarrias-Arrabal^a, M. García-Torres^b, R. Martín-Clemente^c y G. Izquierdo^d

^a Departamento de Psicología Experimental, Facultad de Psicología, Universidad de Sevilla, Sevilla, España

^b Escuela Politécnica Superior, Universidad Pablo de Olavide, Sevilla, España

^c Departamento de Teoría de la Señal y Comunicaciones, Escuela Técnica Superior de Ingeniería, Universidad de Sevilla, Sevilla, España

^d Unidad de Esclerosis Múltiple, Hospital VITHAS, Sevilla, España

Recibido el 5 de junio de 2020; aceptado el 11 de octubre de 2020

Accesible en línea el 3 de febrero de 2021

PALABRAS CLAVE

Biomarcadores;
Clasificación;
Diagnóstico;
Esclerosis múltiple;
Machine learning;
Predicción

Resumen

Introducción: La aplicación de la inteligencia artificial y en particular de algoritmos de aprendizaje automático o «*machine learning*» (ML) constituye un desafío y al mismo tiempo una gran oportunidad en diversas disciplinas científicas, técnicas y clínicas. Las aplicaciones específicas en el estudio de la esclerosis múltiple (EM) no han sido una excepción mostrando un creciente interés en los últimos años.

Objetivo: Realizar una revisión sistemática de la aplicación de algoritmos de ML en la EM.

Material y métodos: Empleando el motor de búsqueda de libre acceso *PubMed* que accede a la base de datos *MEDLINE*, se seleccionaron aquellos estudios que incluyeran simultáneamente los dos siguientes conceptos de búsqueda: «*machine learning*» y «*multiple sclerosis*». Se rechazaron aquellos estudios que fueran revisiones, estuvieran en otro idioma que no fuera el castellano o el inglés, y aquellos trabajos que tuvieran un carácter técnico y no fueran aplicados para la EM. Se seleccionaron como válidos 76 artículos y fueron rechazados 38.

Conclusiones: Tras la revisión de los estudios seleccionados, se pudo observar que la aplicación del ML en la EM se concentró en cuatro categorías: 1) clasificación de subtipos de pacientes dentro de la enfermedad; 2) diagnóstico del paciente frente a controles sanos u otras enfermedades; 3) predicción de la evolución o de la respuesta a intervenciones terapéuticas y por último 4) otros enfoques. Los resultados hallados hasta la fecha muestran que los diferentes algoritmos de ML pueden ser un gran apoyo para el profesional sanitario tanto en la clínica como en la investigación de la EM.

© 2020 Sociedad Española de Neurología. Publicado por Elsevier España, S.L.U. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

* Autor para correspondencia.

Correo electrónico: marrufo@us.es (M. Vázquez-Marrufo).

KEYWORDS

Biomarkers;
Classification;
Diagnosis;
Multiple sclerosis;
Machine learning;
Prediction

A systematic review of the application of machine-learning algorithms in multiple sclerosis

Abstract

Introduction: The applications of artificial intelligence, and in particular automatic learning or “machine learning” (ML), constitute both a challenge and a great opportunity in numerous scientific, technical, and clinical disciplines. Specific applications in the study of multiple sclerosis (MS) have been no exception, and constitute an area of increasing interest in recent years.

Objective: We present a systematic review of the application of ML algorithms in MS.

Materials and methods: We used the PubMed search engine, which allows free access to the MEDLINE medical database, to identify studies including the keywords “machine learning” and “multiple sclerosis.” We excluded review articles, studies written in languages other than English or Spanish, and studies that were mainly technical and did not specifically apply to MS. The final selection included 76 articles, and 38 were rejected.

Conclusions: After the review process, we established 4 main applications of ML in MS: 1) classifying MS subtypes; 2) distinguishing patients with MS from healthy controls and individuals with other diseases; 3) predicting progression and response to therapeutic interventions; and 4) other applications. Results found to date have shown that ML algorithms may offer great support for health professionals both in clinical settings and in research into MS.

© 2020 Sociedad Española de Neurología. Published by Elsevier España, S.L.U. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introducción

La inteligencia artificial y en particular el aprendizaje automático o «*machine learning*» (ML) son conceptos que están abriéndose paso en los últimos años en diversas disciplinas y la medicina no es una excepción. En la última década se ha producido un incremento significativo de procedimientos y aplicaciones del ML para el estudio de las patologías neurológicas. Esta revisión pretende analizar cuáles han sido las primeras vías de aplicación en una patología de gran impacto sanitario y social como es la esclerosis múltiple (EM)¹. Antes de entrar en materia, es necesario precisar que esta revisión no pretende analizar en profundidad los métodos matemáticos que subyacen a los algoritmos de ML empleados en los diferentes estudios (véase² para una revisión a este respecto). Sin embargo, realizaremos una pequeña introducción sobre diversos conceptos relacionados con el ML ampliamente empleados en la literatura que se analiza en esta revisión. Habitualmente es posible confundir los conceptos de inteligencia artificial y ML. ML es un campo de la inteligencia artificial que se centra en el estudio de algoritmos que aprenden de la experiencia (datos) para mejorar el rendimiento en una tarea específica. Por ejemplo, las tareas de clasificación y regresión permiten hallar patrones, en ocasiones muy complejos, del conjunto de datos de la muestra a estudio. Dentro de esta idea general, existen dos posibles alternativas dentro del ML, los procedimientos supervisados y los no supervisados. En el primer caso, se entrena al algoritmo con un conjunto de datos que han sido etiquetados por el investigador (p. ej. subtipos de pacientes en una patología (esclerosis múltiple remitente-recidivante (EMRR), esclerosis múltiple secundaria progresiva (EMSP), etc.). Este procedimiento tiene un enfoque predictivo y el

algoritmo define los posibles patrones de cada etiqueta a partir de las variables o parte de las variables incluidas en el análisis. Posteriormente, y aquí está la clave del gran interés en este campo, el algoritmo permite identificar a cuál de los patrones previamente definidos (aprendidos por el algoritmo) se ajustan mejor los nuevos datos (un nuevo paciente) que se le proporcionen. En los procedimientos no supervisados, enfocados principalmente en tareas descriptivas, no se produce un etiquetado previo de los datos y el algoritmo busca similitudes entre los datos y define por sí solo, los posibles «clústeres» o grupos presentes en la muestra de datos. Existen múltiples procedimientos en el caso de los métodos supervisados que pueden ser reconocidos por el lector tales como: «*neural networks*», «*random forest*», «*support vector machine*» o en el caso de los métodos no supervisados: «*hierarchical clustering*» o «*self-organizing maps*»³. Una de las grandes ventajas de los algoritmos de ML es que permiten analizar conjuntamente datos cuantitativos (*scores*) y cualitativos (evaluaciones del profesional sanitario) de los pacientes en el proceso de aprendizaje del algoritmo. Otro aspecto clave es si el algoritmo de ML realmente está consiguiendo una eficaz clasificación de los patrones existentes entre los datos. Para ello, se emplean una serie de indicadores como son la precisión, sensibilidad o la especificidad entre otros. Sin embargo, lamentablemente no todos los estudios que se han seleccionado en este estudio muestran siempre todas estas medidas o incluso desarrollan sus propios indicadores, lo que a veces complica la comparación de resultados entre los diferentes estudios. La presente revisión pretende mostrar el panorama general de las aplicaciones del ML en el caso particular de la EM, centrada principalmente en los intereses y retos del profesional sanitario. Además, se incluirán los valores estimados de pre-

cisión de los algoritmos en su capacidad de categorización de los pacientes.

Materiales y métodos

La revisión se realizó empleando el motor de búsqueda «PubMed» de libre acceso que permite consultar la base de datos MEDLINE. La selección de estudios se realizó introduciendo como conceptos de búsqueda «*machine learning*» y «*multiple sclerosis*» unidos por el operador booleano «AND» para limitar la búsqueda a aquellos estudios que contuvieran ambos conceptos.

La búsqueda se realizó sin límite inferior en el año de publicación y con fecha de cierre del 30 de abril de 2020. Se rechazaron aquellos estudios que tuvieran un carácter técnico y sin una aportación específica a la esclerosis múltiple (criterio de exclusión 1); aquellos que estuvieran en otro idioma que no fuera el castellano o el inglés (criterio de exclusión 2); y aquellos trabajos que fueran revisiones (criterio de exclusión 3). Sin embargo, en este último caso, se rastreó entre sus referencias, para localizar estudios que trataran de la aplicación de ML en la EM y que no hubieran sido mostrados por el motor de búsqueda inicialmente.

Las características de esta investigación (criterios de selección de los estudios) establecen a priori un bajo nivel de sesgo, sin embargo algunas consideraciones se tuvieron en cuenta para minimizar esta posible amenaza. Por un lado, no se establecieron limitaciones en cuanto al tipo de revista (aspectos como la posición de cuartil o el factor de impacto) o en los años de publicación. Sí se eliminaron las revisiones como referencias elegibles (para evitar posibles duplicaciones), aunque sí se emplearon para poder rastrear referencias no halladas por el motor de búsqueda. Por último, no se han incluido aquellos trabajos enfocados a la mejora en la exploración de resonancia magnética nuclear (RMN) aunque se ha realizado un comentario específico al respecto (véase más adelante en el apartado de resultados).

Resultados

Tras la búsqueda, se obtuvieron 107 estudios de los cuales: 31 fueron rechazados en atención al criterio de exclusión 1; ninguno fue retirado por el criterio de exclusión 2; y 7 fueron excluidos por el criterio de exclusión 3 (revisiones). En el caso de los estudios rechazados por el criterio 3, el rastreo posterior de sus propias referencias rindió 7 estudios no hallados en la búsqueda inicial y que por tanto fueron añadidos a la muestra seleccionada para esta revisión. El total de estudios aceptados finalmente en esta revisión fue de 76 y el número de rechazados 38 (fig. 1).

Respecto de las diversas revisiones halladas en la búsqueda, es necesario destacar que muestran ya un cierto grado de madurez del campo del ML en la disciplina de la Neurología. Sin embargo, la mayoría de estas revisiones abordan múltiples patologías neurológicas y no se centran exclusivamente sobre la esclerosis múltiple^{3–7}. Tan solo hemos encontrado dos revisiones enfocadas principalmente en la EM. Una primera revisión que analiza posibles futuros

biomarcadores de la enfermedad determinados mediante espectroscopia de resonancia magnética⁸ y otro estudio que describe diversas aplicaciones digitales en la EM y donde se recoge un apartado de la aplicación de algoritmos de ML en esta enfermedad⁹. Otra categoría que aparece bien diferenciada entre los estudios encontrados en la búsqueda es aquella en la que se investiga sobre la idoneidad de los algoritmos de ML para la realización de análisis automatizados en la exploración RMN. En este sentido, y considerando que la extensión que ocupan las otras categorías que trataremos a continuación ya está muy ajustada con los límites que la «Neurología» sugiere para un artículo de revisión, hemos decidido incluir al final de este párrafo las referencias de estos estudios para que el lector interesado en esta área pueda ampliar la información, pero no entraremos en un análisis en detalle de los mismos^{10–21}. No obstante, podemos indicar que los estudios de ML en estos casos, pretenden cuantificar si los algoritmos pueden realizar una clasificación automática de las lesiones con un nivel adecuado de precisión. Los diversos resultados han mostrado que todavía no existen altos niveles de precisión en su aplicación y la participación del profesional sanitario es necesaria. En particular, algunos estudios apuntan a que el aspecto crítico en la aplicación del ML no depende tanto de los algoritmos empleados como de los parámetros incluidos de la RMN¹¹. Sin duda, futuros estudios serán necesarios para obtener procedimientos válidos en la aplicación clínica del ML en la identificación de lesiones en la RMN.

Los estudios seleccionados pueden ser clasificados en las siguientes cuatro categorías: clasificación de subtipos dentro de la enfermedad; diagnóstico; predicción; y otros enfoques.

Clasificación de subtipos dentro de la enfermedad

A pesar de la corta trayectoria de aplicación de los algoritmos de ML a la EM, es posible observar un claro cambio de tendencia en su breve historia sobre los objetivos perseguidos (fig. 2). En los estudios iniciales, el principal reto era la clasificación de los pacientes dentro de la enfermedad (subtipos). Como puede observarse en la tabla 1, los estudios con este objetivo^{22–33} se presentan de manera más homogénea a lo largo de la pasada década respecto a otras aplicaciones (diagnóstico o predicción) (tablas 2 y 3) que han aumentado considerablemente su número en los últimos años (2017–2020).

Los diversos estudios han permitido clasificar a pacientes en diversas categorías: EMRR; esclerosis múltiple benigna; síndrome clínico aislado (CIS); síndrome radiológico aislado (RIS); EM pediátrica de inicio temprano o tardío; o por características como la carga lesional o el comienzo en la enfermedad; o incluso atendiendo al grado de discapacidad cognitiva. En cuanto a las variables empleadas para la clasificación también encontramos una gran variedad de opciones. Desde diversos parámetros de la RMN relacionados con la sustancia gris y blanca, pasando por medidas clínicas (EDSS o duración de la enfermedad), hasta medidas no necesariamente cuantitativas como por ejemplo notas clínicas del paciente.

El rango de precisión alcanzado entre los diferentes estudios está comprendido desde el 70,6 hasta el 96%. Estos

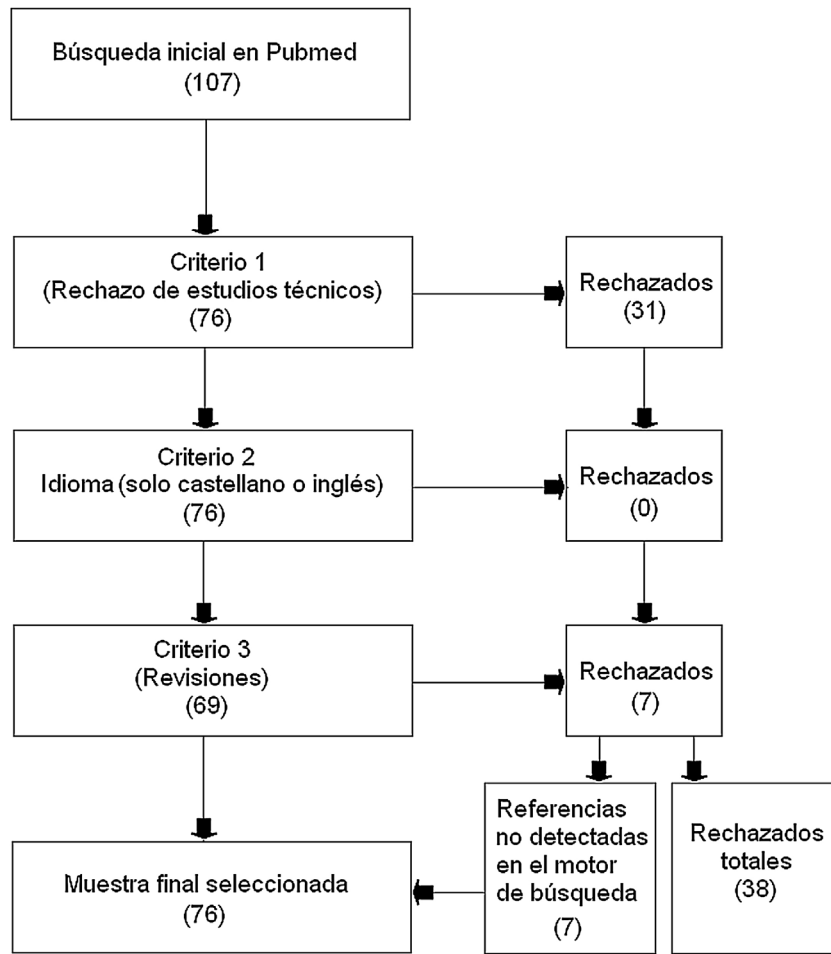


Figura 1 Diagrama de flujo del procedimiento de revisión.

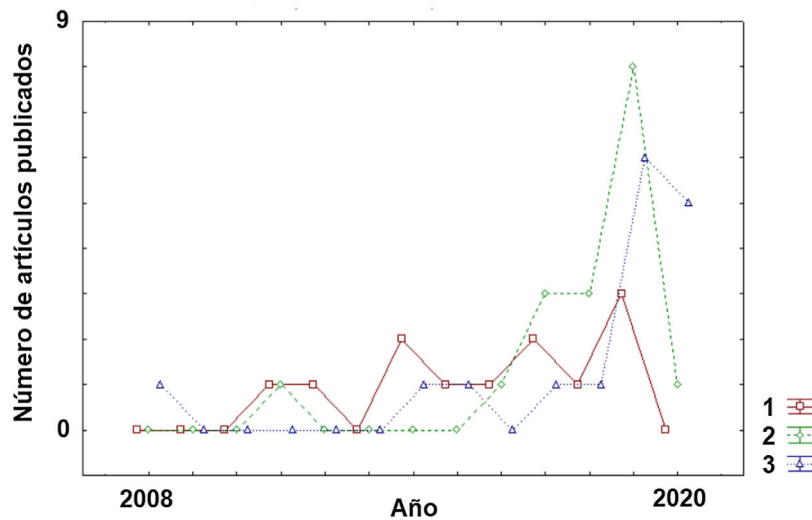


Figura 2 Evolución anual de estudios publicados en PubMed empleando las palabras clave «Machine learning» y «Multiple Sclerosis» y para cada una de las categorías analizadas (1: clasificación; 2: diagnóstico y 3: predicción).

valores varían mucho entre los diferentes estudios sugiriendo que los procedimientos algorítmicos aplicados, las variables empleadas para la clasificación y los subtipos de la enfermedad que se pretenden identificar son determinantes

para el nivel de precisión. Se puede concluir que los algoritmos pueden alcanzar niveles aceptables en la clasificación de los subtipos en pacientes y pueden permitir mejorar la estratificación de los mismos para posteriores estudios de

Tabla 1 Estudios seleccionados de «*machine learning*» en la clasificación de subtipos en la esclerosis múltiple. Clasificación de subtipos de la enfermedad

Referencia	Parámetros	Precisión
González-Rosa et al. (2011)	Puntuaciones de EEG y conductuales (tiempo de reacción) en tareas cognitivas computarizadas	Pacientes EM y controles sanos con distinto nivel de discapacidad cognitiva. (N.D.)
Bendfeldt et al. (2012)	Alteraciones de la sustancia gris en RMN	EM temprana vs. EM tardía (85%) Pacientes con baja vs. alta carga lesional (83%) EM benigna vs. EM no benigna (77%)
Crimi et al. (2014)	Patrones espacio-temporales de lesiones en RMN	Tres tipos de patrones de lesiones fueron encontrados (N.D.)
Weygandt et al. (2014)	RMN (lesiones en sustancia blanca y gris)	Esclerosis múltiple pediátrica (EMP) y sujetos sanos (87,1%) Esclerosis múltiple pediátrica de inicio temprano vs. sujetos sanos y pacientes con esclerosis múltiple pediátrica de inicio tardío (77,3%)
Fiorini et al. (2015)	Múltiples variables que evalúan movilidad, fatiga, capacidad cognitiva, estatus emocional, incontinencia urinaria y calidad de vida	EMRR vs. otras formas de EM (75-78%)
Kocevar et al. (2016)	RMN (imágenes T1 y DTI)	Sujeto sano vs. CIS (91,8%)* CIS vs. EMRR (91,8%)* EMRR vs. EMPP (75,6%)* Comparación entre diversos subtipos CIS-EMRR-EMSP (70,6%)*
Zhong et al. (2017)	RMN (espesor cortical y volumen de sustancia gris subcortical) y conectividad funcional interregional	Sujetos sanos vs. pacientes con discapacidad motora (MI) (88%) Sujetos sanos vs. pacientes sin discapacidad motora (MP) (84%) MI vs. MP (86%)
Ion-Mărgineanu et al. (2017)	Datos demográficos y clínicos (edad, duración de la enfermedad, EDSS), RMN: Carga lesional y ratios de metabolitos (NAA/Cre, NAA/Cho y Cho/Cre).	CIS vs. EMRR (F1-score: 71%)** CIS vs. EMRR + EMSP (F1-score: 72%)** EMRR vs. EMPP (F1-score of 85%)** EMRR vs. EMSP (F1-score of 87%)**
López et al. (2018)	SNP (<i>single nucleotide polymorphism</i>) (genes implicados en procesos inmunes y vías de adhesión celular)	Clústeres de pacientes (96%)
Van Le et al. (2019)	Múltiples variables (medicación administrada, resultados de laboratorio e incluso notas clínicas)	Identificación de pacientes EMRR sin riesgo de convertir a formas progresivas. (Valor predictivo positivo (PPV): 99,1%***)
Mato-Abad et al. (2019)	RMN (espesor cortical, volumen de sustancia gris cortical y subcortical, e integridad de la sustancia blanca)	CIS vs. RIS (76,5%)
González-Campo et al. (2019)	RMN (análisis estructural y funcional de las estructuras cerebrales interoceptivas (ínsula y corteza cingulada anterior)	Pacientes EM con síntomas de fatiga y sin fatiga (AUC > 90%)

AUC: área bajo la curva; CIS: síndrome clínico aislado; EEG: electroencefalografía; EM: esclerosis múltiple; EMPP: esclerosis múltiple primaria progresiva; EMRR: esclerosis múltiple remitente-recidivante; EMSP: esclerosis múltiple secundaria progresiva. N.D.: no descrito; RIS: síndrome radiológico aislado; RMN: resonancia magnética nuclear.

* Todos estos valores corresponden con lo que los autores denominan «*F-measure*» y que incluyen la medida de precisión, pero también lo que es definido como «*recall*» que representa la proporción de verdaderos positivos identificados entre los falsos negativos. Véase la referencia para una definición precisa de la fórmula.

** $F1\text{-score} = 2 \times TP / (2 \times TP + FN + FP)$ (TP: verdadero positivo; FN: falso negativo; FP: falso positivo).

*** $PPV = (\text{verdadero positivo}) / (\text{verdadero positivo} + \text{falso positivo})$.

Tabla 2 Estudios seleccionados de «*machine learning*» en el diagnóstico de la esclerosis múltiple. Diagnóstico

Referencia	Grupos	Marcadores	Precisión
Weygandt et al. (2011)	Pacientes EM vs. controles sanos	RMN (mapeo de patrones de intensidad en áreas lesionadas de la SGAN y SBAN)	Lesiones en sustancia blanca del lóbulo parietal posterior (96%) Patrones en la sustancia gris aparentemente normal (cerebelo) (84%) Patrones en la sustancia blanca aparentemente normal (áreas posteriores del cerebro) (91%)
Ohanian et al. (2016)	Encefalomiелitis miálgica, Síndrome de fatiga crónica y esclerosis múltiple	<i>DePaul Symptom Questionnaire</i> (DSQ) (realizado a través de medios telemáticos)	Las variables inmunitarias fueron las mejores para la categorización (81,2%)
Lötsch et al. (2017)	Pacientes EM vs. controles sanos	Concentraciones en suero de marcadores lipídicos (ceramidas)	94,6%
Ostmeyer et al. (2017)	Pacientes con esclerosis múltiple vs. pacientes con otras enfermedades neurológicas	Marcadores genéticos de linfocitos B	87%
Yoo et al. (2017)	Pacientes EM vs. controles sanos	RMN (mapas de mielina)	87,9%
Zurita et al. (2018)	EMRR vs. controles sanos	RMN (estudio de la conectividad estructural y funcional)	89% (Regiones críticas para la clasificación: córtex occipital derecho, Giro orbital frontal izquierdo, cortezas frontales mediales y giro lingual)
Lötsch et al. (2018)	Pacientes EM vs. controles sanos	Biomarcadores lipídicos en suero	95%
Forbes et al. (2018)	Categorización entre diferentes enfermedades (esclerosis múltiple (EM); enfermedad de Crohn (EC); colitis ulcerosa (CU); artritis reumatoide (AR); controles sanos (CS)	Microbiota intestinal	Múltiples valores de precisión entre las enfermedades: EM vs. EC: 77% EM vs. CU: 73% EM vs. AR: 59% EM vs. CS: 82%
Saccà et al. (2019)	Pacientes EM vs. controles sanos	RMNf (conectividad de redes neuronales en <i>resting-state</i>)	85,7% (red neuronal sensoriomotora)
Neeb et al. (2019)	Pacientes EM vs. controles sanos	RMN	83,7% (RMN libre de artefactos de movimiento) 74,5% (RMN con artefactos de movimiento)
Ahmadi et al. (2019)	Pacientes EM vs. controles sanos	Registro EEG durante tareas atencionales	90-91% según la tarea atencional
Andersen et al. (2019)	Pacientes EM vs. controles sanos	Estudio de 325 metabolitos en suero	12 metabolitos fueron encontrados como marcadores de EM, de los cuales 6 mostraban una AUC >80% (piroglutamato, laurato, acilcarnitina, N-metilmaleimida y fosfatidilcolinas)
Pérez del Palomar et al. (2019)	Pacientes EM vs. controles sanos	TCO	97,24%

Tabla 2 (continuación)

Referencia	Grupos	Marcadores	Precisión
Bang et al. (2019)	Diagnóstico entre diferentes enfermedades (esclerosis múltiple, artritis idiopática juvenil, encefalomielitis/síndrome de fatiga crónica, síndrome de inmunodeficiencia adquirida, infarto y cáncer colorrectal)	Microbiota intestinal	Múltiples valoraciones según las enfermedades comparadas (en todos los casos por encima del 95%)
Azarmi et al. (2019)	Pacientes EM vs. controles sanos	RMNf durante la realización de pruebas cognitivas	95%
Cavaliere et al. (2019)	Pacientes EM vs. controles sanos	TCO	91%
Heming et al. (2020)	Esclerosis múltiple (EM) vs. neurosarcoidosis (NS)	Parámetros de muestra sanguínea y de LCR (Pacientes NS mostraron un incremento de linfocitos T CD4+ y de células plasmáticas en sangre)	Probabilidad de predicción de NS en la muestra 76% y del 99% para los pacientes con EM

AUC: área bajo la curva; EEG: electroencefalografía; EM: esclerosis múltiple; LCR: líquido cefalorraquídeo; RMN: resonancia magnética nuclear; RMNf: resonancia magnética nuclear funcional; SBAN: sustancia blanca aparentemente normal; SGAN: sustancia gris aparentemente normal; TCO: tomografía de coherencia óptica.

respuesta a intervenciones terapéuticas o de progresión de la enfermedad.

Diagnóstico

En el caso de la aplicación de los algoritmos de ML con el objetivo principal del diagnóstico, hallamos un estudio del 2011 y es necesario esperar hasta el 2016 para encontrar nuevos estudios. Pero sin duda, algo que llama la atención es que en el año 2019 se acumularon una gran cantidad de trabajos en esta categoría. Curiosamente, y de modo distinto a lo que ocurrirá en el apartado de «Predicción», en el año 2020 solo se ha publicado un estudio en los primeros 4 meses (tabla 2)^{34–50}. Otro aspecto destacable en esta categoría es que los trabajos se han orientado principalmente para el diagnóstico de la EM frente a controles sanos (13 estudios). En un número menor (5 estudios), la aplicación del algoritmo persiguió la diferenciación entre diferentes entidades patológicas.

Las principales variables que se han empleado para mejorar nuestra capacidad de diagnóstico proceden de la técnica de RMN en diferentes parámetros de la misma (mapas lesionales, conectividad funcional, etc.). Sin embargo, existen otros marcadores que ofrecen interesantes niveles de precisión como los marcadores serológicos o genéticos, o pruebas como la tomografía de coherencia óptica (TCO) o la electroencefalografía (EEG).

El rango de precisión se presenta desde niveles muy buenos (98,2%) hasta niveles moderados (59%). Curiosamente, ambos resultados se obtuvieron empleando las características de la microbiota intestinal entre diferentes entidades patológicas. En el primer caso, se trataba de distinguir la EM entre otras enfermedades como artritis idiopática o

síndrome de fatiga crónica⁴⁷. En el segundo caso, ese bajo nivel de precisión se produjo a la hora de categorizar entre la EM y la artritis reumatoide⁴⁰.

Predicción

Es conveniente destacar que en este apartado se han incluido trabajos que en el fondo realizan en algunos casos un proceso de clasificación en subtipos dentro de la enfermedad. Es el caso de estudios en los que se trata de discriminar a buenos de malos respondedores en tratamientos farmacológicos. La decisión de ubicar estos estudios en esta categoría radica en que la finalidad última de la aplicación del algoritmo de ML no es solo la clasificación sino la predicción.

Si se revisa la tabla 3^{51–67} se puede observar que se registró un primer estudio en el año 2008 y tuvieron que pasar varios años (hasta el 2014) para encontrar nuevos estudios con el objetivo de la predicción. Sin embargo, este objetivo es el que hoy por hoy más crecimiento ha experimentado en los dos últimos años con 6 artículos tan solo en los primeros 4 meses del año 2020. Las diversas líneas de actuación en este campo comprenden dos objetivos principales. Por un lado, predecir la posible evolución del paciente (p. ej. riesgo de convertirse en otra forma más evolucionada de la enfermedad), y por otro, anticipar la posible respuesta del paciente a una terapia farmacológica para ayudar al profesional en la elección del tratamiento más adecuado. Es sin duda, una aproximación para un tipo de medicina más personalizada que contempla la posible heterogeneidad de los condicionantes específicos de cada paciente. Respecto a cuánto de precisos son los algoritmos de ML en estos dos casos, en la predicción en la progresión de la enfermedad,

Tabla 3 Estudios seleccionados de «*machine learning*» en el apartado de predicción. Predicción

Referencia	Predicción	Marcador	Precisión
Corvol et al. (2008)	Conversión de CIS a EMCD	Expresión génica en células T CD4(+)	92%
Wottschel et al. (2014)	Probabilidad de un segundo brote que lleva a un cambio de diagnóstico (CIS a EMCD)	Combinación de medidas clínicas, sociodemográficas y RMN	Predicción de la presencia de EMCD entre los CIS (71,4%)
Baranzini et al. (2015)	Asociación de un biomarcador con la respuesta al tratamiento con IFN	Triplete de genes (CASP2/IRF4/IRF6) fueron relacionados con una peor respuesta al tratamiento	68%
Zhao et al. (2017)	Clasificación de pacientes que tendrán una mejor o peor progresión en la enfermedad	Combinación de medidas clínicas, sociodemográficas y RMN	86%
Kiiski et al. (2018)	Predicción de la progresión de la discapacidad cognitiva	Múltiples variables: Registro de EEG durante la realización de tareas cognitivas; evaluación neuropsicológica (MACFIMS) y datos sociodemográficos (edad, educación) y clínicos (EDSS)	N.D.
Zhang et al. (2019)	Conversión de CIS a EM	Características de las lesiones en RMN	84,5%.
Fagone et al. (2019)	Predicción de la respuesta de pacientes EMRR al tratamiento con natalizumab antes de iniciar la terapia	Subset de genes de las células T (CD4+)	89,3%
Goyal et al. (2019)	1) Predicción de la progresión en la enfermedad 2) Clasificación de EMRR o no EMRR	1) Nivel en suero de ocho citocinas (IL-1 β , IL-2, IL-4, IL-8, IL-10, IL-13, IFN- γ , y TNF- α) 2) Citocinas, edad, género, duración de la enfermedad, EDSS y «Multiple Sclerosis Severity Score» (MSSS)	1) 91% 2) 70%
Sun et al. (2019)	Diferenciar pacientes con riesgo de caídas	Evaluaciones de posturografía estática y riesgo fisiológico de caída	86%
Law et al. (2019)	Predicción de la progresión de la discapacidad a corto plazo en EMSP	Variables: EDSS, MSFC, volumen de la lesión en T2, duración de la enfermedad, edad y género	AUC: 61,8%
Wottschel et al. (2019)	Predicción de brotes en CIS para convertirse en EM	RMN (carga lesional de sustancia blanca, espesor cortical y volumen de regiones corticales y de sustancia blanca)	Rango entre 73% a 92,9% (entre distintos centros)
Jackson et al. (2020)	Determinación de genes relacionados con la progresión de la discapacidad	113 variantes genéticas relacionados con la severidad de la EM	N.D.
Ebrahimkhani et al. (2020)	Biomarcadores de la actividad de la enfermedad que predicen la respuesta a la terapia con fingolimod	« <i>Circulating exosome microRNA</i> » y actividad de la enfermedad determinada por RMN con contraste de gadolinio	Estratificación mediante micro-RNA permite determinar una respuesta positiva al tratamiento con fingolimod (N.D.)

Tabla 3 (continuación)

Referencia	Predicción	Marcador	Precisión
Ghafouri-Fard et al. (2020)	Predicción de riesgo de padecer EM mediante el contenido genético del paciente	« <i>Single nucleotide polymorphisms</i> » (SNP)	64,73%
Seccia et al. (2020)	Predicción de la conversión de EMRR a EMSP	21 variables (demográficas, clínicas, RMN y tratamientos terapéuticos)	Positive predictive value (PPV): 42%
Yperman et al. (2020)	Predicción de la progresión de la discapacidad empleando potenciales evocados motores (MEPs)	Serie temporales completas de los MEPs en vez de parámetros discretos de los mismos	AUC = 0,75
Jin et al. (2020)	Predicción de la respuesta al IFN (buenos o malos respondedores) según la genética del paciente	41 genes (véase la referencia para una lista completa de los mismos)	80,95%

AUC: área bajo la curva; CIS: síndrome clínico aislado; EDSS: escala expandida de discapacidad (expanded disability status scale); EM: esclerosis múltiple; EMCD: esclerosis múltiple clínicamente definida; EMRR: esclerosis múltiple remitente-recidivante; EMSP: esclerosis múltiple secundaria progresiva; IFN: interferón; MACFIMS: *Minimal Assessment of Cognitive Function in Multiple Sclerosis battery*; N.D.: no descrito; RMN: resonancia magnética nuclear.

los porcentajes de precisión estuvieron en un rango entre 61,8 y 92,9%. En el caso de la precisión en clasificar a los pacientes en función de la respuesta que mostrarán a terapias farmacológicas el rango de precisión es similar con una horquilla entre el 68 y el 89,3%.

Otros enfoques

En este último apartado hemos querido incluir otras líneas de actuación que se están produciendo en la aplicación del ML en la EM. Una de las líneas más activas en este sentido es la que trata de desarrollar nuevos indicadores (*scores*) que integran diferentes variables y consigan mejorar la evaluación de los pacientes. Un primer ejemplo de esta aplicación fue desarrollado por el equipo de Kosa et al.⁶⁸ (2016). Este grupo planteó la necesidad de obtener un indicador más sensible para la evaluación de cambios en la progresión de la enfermedad y así poder anticipar lo antes posible la respuesta a tratamientos farmacológicos en pacientes con la forma progresiva. El nuevo evaluador denominado «CombiWISE» integra valores de varias escalas: «*Expanded disability status scale*» (EDSS), «*Scripps neurological rating scale*», «*25 foot walk test*» y el «*9 hole peg test*». Los resultados mostraron en el seguimiento de 98 pacientes con la forma progresiva que el conjunto de estos indicadores era mejor que las medidas de atrofia en RMN o que la escala EDSS por sí solas. De hecho un punto en la EDSS correspondía con 7,5 puntos en el nuevo indicador aumentando por tanto la resolución en la evaluación de los pacientes. En otro estudio⁶⁹ se desarrolló un indicador denominado «*Brain Age*» para poder determinar si vale como marcador de curso clínico y severidad de la enfermedad. Los resultados mostraron que este indicador mostraba, en promedio, un valor superior en pacientes con EM respecto de sujetos sanos y que correlacionaba con la atrofia cerebral y la carga lesional en sustancia blanca. Desafortunadamente no se encontraron buenas correlaciones entre este indicador y diversas

variables clínicas de los pacientes, lo cual hace dudar de su potencial uso clínico. En otros casos, los algoritmos de ML se han centrado en optimizar las capacidades de evaluación ya existentes. Una línea de actuación muy activa se ha concentrado en la determinación de la discapacidad motora y su relación con diversos parámetros evaluados del paciente. En estos estudios principalmente se suele analizar la marcha del paciente mediante sensores colocados en la piel^{70–73}. Otra de las líneas que acumula más estudios es la que analiza las posibles causas de la enfermedad. En este sentido, se han desarrollado estudios con dos objetivos: describir posibles factores de riesgo^{74,75} e identificar potenciales marcadores de la EM para desarrollar posibles dianas terapéuticas útiles en futuros tratamientos^{76–79}. Por último, quisiéramos destacar una línea de actuación que trata de valorar con diversos enfoques si la aplicación del ML resulta beneficiosa para el profesional sanitario en términos eminentemente prácticos. En este sentido, se ha tratado de valorar si la colaboración entre los profesionales sanitarios y la aplicación de los algoritmos de ML dan mejores resultados en la evaluación de los pacientes que cuando se realiza esta por separado. Para ello se realizó un estudio en el que se trataba de predecir el curso de la enfermedad en función de los historiales de los pacientes con EM⁸⁰. El estudio concluyó que la cooperación entre profesionales humanos y el ML rendía mejores valores que cuando los pronósticos se realizaban por separado.

Además, los algoritmos de ML pueden ayudar al profesional sanitario en reducir la contribución de factores que pueden estar confundiendo para realizar una correcta clasificación. Por ejemplo, un estudio ha mostrado cómo puede reducirse, sin necesidad de la intervención por parte del profesional sanitario, la contribución de factores como la edad que pueden deshacer la homogeneidad de los grupos (clústeres) descritos por el algoritmo de ML⁸¹. En otros casos, la aplicación de ML proporciona una mejora en la realización de la prueba reduciendo el tiempo de operación como en el caso de la TCO, donde se consigue una exploración de

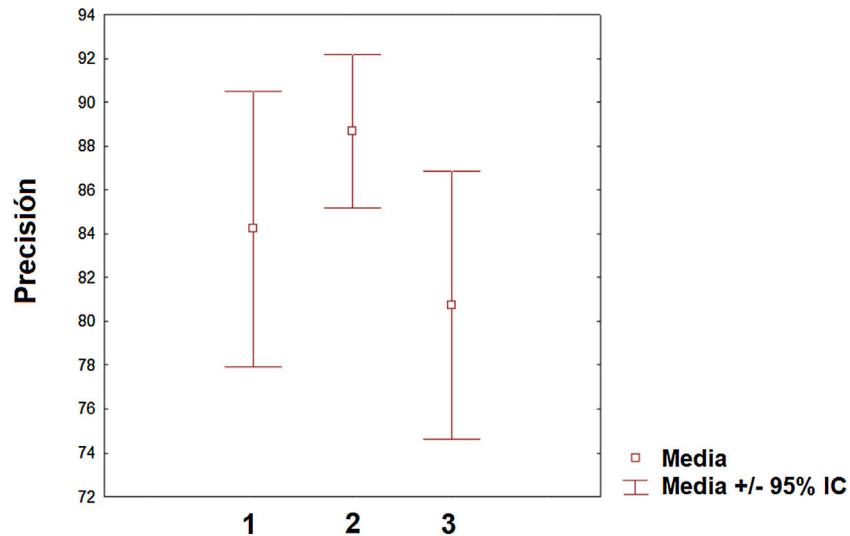


Figura 3 Nivel medio de precisión e intervalos de confianza (95%) en cada una de las categorías analizadas (1: clasificación; 2: diagnóstico y 3: predicción).

las superficies de la retina y del edema macular con una precisión mejor que la habitualmente lograda con métodos más costosos en términos de tiempo, y realizada en tan solo 10 segundos⁸². Otra aplicación en la línea de apoyar al profesional sanitario, en este caso al neuropsicólogo, es la de automatizar parcialmente la evaluación cognitiva en los pacientes⁸³. Durante la progresión de la enfermedad, el paciente puede ir experimentando deterioro cognitivo y uno de los test aplicados para su evaluación es el «*Brief Visuospatial Memory Test-Revised*» (BVMT-R), en el que el paciente tiene que reconocer figuras geométricas y dibujarlas. Tras la digitalización del test y un entrenamiento del algoritmo, los resultados mostraron una precisión del 80% en la evaluación del test considerando la evaluación del profesional como *gold-standard*. Los autores concluyeron que la primera evaluación sigue siendo una tarea para el profesional sanitario pero futuras mejoras en el aprendizaje de los algoritmos puedan permitir que las reevaluaciones en estudios longitudinales sean desarrolladas de manera fiable por los algoritmos de ML.

Discusión

Las aplicaciones más intuitivas de los algoritmos de ML en la EM son sin duda el diagnóstico (o clasificación si hablamos de subtipos dentro de la enfermedad) y la predicción de la evolución o respuesta a las posibles estrategias terapéuticas. Sin embargo, las aplicaciones posibles de los algoritmos de ML son enormes y recientes estudios muestran nuevas vías de investigación que se abren cara al futuro. En el caso particular de la EM, el análisis de todas las ramas «ómicas» con ingentes cantidades de datos pueden ser abordadas con gran potencia de análisis, algo que hasta hace unos años resultaba muy complicado realizar con las técnicas convencionales de análisis estadístico. Particularmente interesante es la aplicación de estos algoritmos en el estudio de posibles

biomarcadores identificativos de la enfermedad que nos permitan comprender la etiología precisa de la misma y trazar posibles vías terapéuticas en el futuro.

Pero sin duda, el campo con una mayor proyección para su aplicación clínica es el de la predicción de la respuesta ante los diversos tratamientos farmacológicos que existen y se generarán en los próximos años. De hecho, la predicción es la aplicación que ha aumentado más en el número de publicaciones en los dos últimos años. Sin embargo, todavía existe un desequilibrio de los trabajos aplicados para predecir la progresión de la enfermedad (línea mayoritaria) de aquellos que se orientan hacia la predicción de la respuesta terapéutica (línea minoritaria). Sin duda, en los próximos años, esta última línea será líder en la aplicación del ML en la EM así como en otras patologías neurológicas. Respecto de la precisión obtenida entre las diferentes categorías (clasificación, diagnóstico y predicción), se puede decir que el valor medio es similar aunque el intervalo de confianza sea algo distinto entre ellas (fig. 3). Es necesario indicar que la comparación entre las diversas categorías debe considerarse con precaución dado que las medidas de precisión, el número de variables incluidas y los procedimientos de ML empleados difieren de unos estudios a otros. Sin embargo, es posible concluir que existen niveles relevantes de precisión en los primeros estudios realizados en este campo. Dado el carácter de entrenamiento que requieren estos algoritmos de ML, es necesario que las instituciones sanitarias hagan un esfuerzo por cooperar en repositorios de datos para poder incrementar la capacidad de los diversos algoritmos en la descripción de los posibles patrones dentro de la enfermedad. En el caso del contenido genético, este enfoque es particularmente interesante dado que esto permitirá una caracterización más precisa del paciente y personalizar aún más si cabe en el futuro los tratamientos. Sin duda, nos encontramos en el camino hacia una medicina personalizada sugerida desde múltiples estudios⁸⁴⁻⁸⁶. Es esperable que, en los próximos años, aquellos sistemas sanitarios que hagan un esfuerzo en la implementación de este tipo de

herramientas de la inteligencia artificial liderarán parte de los avances más relevantes en la práctica clínica. Entre las limitaciones de este estudio debemos destacar algunas consideraciones. El motor de búsqueda *PubMed* puede cometer errores en su rastreo de los artículos publicados hasta la fecha. Sin embargo, el análisis de los apartados de bibliografía de los diferentes artículos analizados minimiza la posibilidad de que un número elevado de artículos relacionados con el ML y la EM no hayan sido incluidos en la revisión. Otro aspecto que puede tenerse en cuenta es que existen otros conceptos como «*deep learning*» que pueden ser de interés en su aplicación en la EM y están relacionados con el ML. En futuras revisiones sobre el ML y su aplicación sobre la EM u otras patologías, deberá contemplarse la posibilidad de ampliar las palabras clave a otras posibles acepciones como «*deep learning*» o «*network representation learning*».

Conclusión

El presente trabajo de revisión ha permitido observar que la aplicación de los algoritmos de ML en la EM está siendo un campo muy activo mostrando un crecimiento exponencial en los dos últimos años. Las principales áreas de aplicación son la clasificación de diferentes subtipos dentro de la enfermedad, el diagnóstico frente a sujetos sanos y otras enfermedades y la predicción de la evolución del paciente o la respuesta a un tratamiento terapéutico. Sin embargo, es posible encontrar también estudios que tratan de optimizar técnicas o buscar nuevos indicadores para mejorar la evaluación clínica en términos de tiempo, coste y bienestar para el paciente. En general, lo observado en los diferentes estudios es que la aplicación de este tipo de algoritmos ML obtiene valores de precisión en algunos casos excelentes (>90%), y puede ser un importante apoyo para el profesional sanitario en el diagnóstico y pronóstico dentro de la EM.

Financiación

Este trabajo ha sido subvencionado por el proyecto PSI2016-78133-P del Ministerio de Ciencia, Innovación y Universidades (España).

Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

Bibliografía

- Vázquez-Marrufo M, Borges M, González-Rosa JJ. *Neuropsicología del deterioro cognitivo en EM*. Madrid: Editorial Síntesis; 2018.
- Nagy Z. *Early detection of neurological disorders using machine learning systems*. Birmingham UK: Packt Publishing Ltd; 2018.
- Stafford IS, Kellerman M, Mossotto E, Beattie RM, MacArthur BD, Ennis S. A systematic review of the applications of artificial intelligence and machine learning in autoimmune diseases. *NPJ Digit Med*. 2020;3:30, <http://dx.doi.org/10.1038/s41746-020-0229-3>.
- Oxtoby NP, Alexander DC. EuroPOND consortium Imaging plus X: Multimodal models of neurodegenerative disease. *Curr Opin Neurol*. 2017;30:371–9, <http://dx.doi.org/10.1097/WCO.0000000000000460>.
- Mateos-Pérez JM, Dadar M, Lacalle-Aurioles M, Iturria-Medina Y, Zeighami Y, Evans AC. Structural neuroimaging as clinical predictor: A review of machine learning applications. *Neuroimage Clin*. 2018;20:506–22, <http://dx.doi.org/10.1016/j.nicl.2018.08.019>.
- Sakai K, Yamada K. Machine learning studies on major brain diseases: 5-year trends of 2014–2018. *Jpn J Radiol*. 2019;37:34–72, <http://dx.doi.org/10.1007/s11604-018-0794-4>.
- Raghavendra U, Acharva UR, Adeli H. Artificial Intelligence Techniques for automated diagnosis of neurological disorders. *Eur Neurol*. 2019;82(1–3):41–64, <http://dx.doi.org/10.1159/000504292>.
- Swanberg KM, Landheer K, Pitt D, Juchem C. Quantifying the metabolic signature of multiple sclerosis by in vivo proton magnetic resonance spectroscopy: Current challenges and future outlook in the translation from proton signal to diagnostic biomarker. *Front Neurol*. 2019;10:1173, <http://dx.doi.org/10.3389/fneur.2019.01173>.
- Matthews PM, Block VJ, Leocani L. E-health and multiple sclerosis. *Curr Opin Neurol*. 2020;33:271–6, <http://dx.doi.org/10.1097/WCO.0000000000000823>.
- Yamamoto D, Arimura H, Kakeda S, Magome T, Yamashita Y, Toyofuku F, et al. Computer-aided detection of multiple sclerosis lesions in brain magnetic resonance images: False positive reduction scheme consisted of ruled-based, level set method, and support vector machine. *Comput Med Imaging Graph*. 2010;34:404–13, <http://dx.doi.org/10.1016/j.compmedimag.2010.02.001>.
- Sweeney EM, Vogelstein JT, Cuzzocreo JL, Calabresi PA, Reich DS, Crainiceanu CM, et al. A comparison of supervised machine learning algorithms and feature vectors for MS lesion segmentation using multimodal structural MRI. *Plos One*. 2014;9:e95753, <http://dx.doi.org/10.1371/journal.pone.0095753>.
- Cabezas M, Oliver A, Valverde S, Beltran B, Freixenet J, Vilanova JC, et al. BOOST: A supervised approach for multiple sclerosis lesion segmentation. *J Neurosci Methods*. 2014;237:108–17, <http://dx.doi.org/10.1016/j.jneumeth.2014.08.024>.
- Deshpande H, Maurel P, Barillot C. Classification of multiple sclerosis lesions using adaptive dictionary learning. *Comput Med Imaging Graph*. 2015;47(Pt 1):2–10, <http://dx.doi.org/10.1016/j.compmedimag.2015.05.003>.
- Roure E, Olver A, Cabezas M, Valverde S, Pareto D, Vilanova JC, et al. A toolbox for multiple sclerosis lesion segmentation. *Neuroradiology*. 2015;57:1031–43, <http://dx.doi.org/10.1007/s00234-015-1552-2>.
- Fartaria MJ, Bonnier G, Roche A, Kober T, Meuli R, Rotzinger D, et al. Automated detection of white matter and cortical lesions in early stages of multiple sclerosis. *J Magn Reson Imaging*. 2016;43:1445–54, <http://dx.doi.org/10.1002/jmri.25095>.
- Brosch T, Tang LY, Youngin Y, Li DK, Traboulsee A, Tam R. Deep 3D convolutional encoder networks with shortcuts for multiscale feature integration applied to multiple sclerosis lesion segmentation. *IEEE Trans Med Imaging*. 2016;35:1229–39, <http://dx.doi.org/10.1109/TMI.2016.2528821>.
- Salem M, Cabezas M, Valverde S, Pareto D, Oliver A, Salvi J, et al. A supervised framework with intensity subtraction and deformation field features for the detection of new T2-w lesions in multiple sclerosis. *Neuroimage Clin*. 2017;17:607–15, <http://dx.doi.org/10.1016/j.nicl.2017.11.015>.
- Commowick O, Istace A, Kain M, Laurent B, Leray F, Simon M, et al. Objective evaluation of multiple

- sclerosis lesion segmentation using a data management and processing infrastructure. *Sci Rep.* 2018;8:13650, <http://dx.doi.org/10.1038/s41598-018-31911-7>.
19. Kanber B, Nachev P, Barkhof F, Calvi A, Cardoso J, Cortese R, et al. High-dimensional detection of imaging response to treatment in multiple sclerosis. *NPJ Digit Med.* 2019;2:49, <http://dx.doi.org/10.1038/s41746-019-0144-7>.
 20. Eitel F, Soehler E, Bellmann-Strobl J, Brandt AU, Ruprecht K, Giess RM, et al. Uncovering convolutional neural network decisions for diagnosis multiple sclerosis on conventional MRI using layer-wise relevance propagation. *Neuroimage Clin.* 2019;24:102003, <http://dx.doi.org/10.1016/j.nicl.2019.102003>.
 21. Rachmadi MF, Valdés-Hernández MDC, Li H, Guerrero R, Meijboom R, Wiseman S, et al. Limited one-time sampling irregularity map (LOTS-IM) for automatic unsupervised assessment of white matter hyperintensities and multiple sclerosis lesions in structural brain magnetic resonance images. *Comput Med Imaging Graph.* 2020;79:101685, <http://dx.doi.org/10.1016/j.compmedimag.2019.101685>.
 22. González-Rosa JJ, Vázquez-Marrufo M, Vaquero E, Duque P, Borges M, Gómez-González CM, et al. Cluster análisis of behavioural and event-related potentials during a contingent negative variation paradigm in remitting-relapsing and benign forms of multiple sclerosis. *BMC Neurol.* 2011;11:64, <http://dx.doi.org/10.1186/1471-2377-11-64>.
 23. Bendfeldt K, Klöppel S, Nichols TE, Smieskova R, Kuster P, Traud S, et al. Multivariate pattern classification of gray matter pathology in multiple sclerosis. *Neuroimage.* 2012;60:400–8, <http://dx.doi.org/10.1016/j.neuroimage.2011.12.070>.
 24. Crimi A, Commowick O, Maarouf A, Ferré JC, Bannier E, Tourbah A, et al. Predictive value of imaging markers at multiple sclerosis disease onset based on gadolinium and USPIO-enhanced MRI and machine learning. *PLoS One.* 2014;9:e93024, <http://dx.doi.org/10.1371/journal.pone.0093024>.
 25. Weygandt M, Hummel HM, Schregel K, Ritter K, Allefeld C, Domes E, et al. MRI-based diagnostic biomarkers for early onset pediatric multiple sclerosis. *Neuroimage Clin.* 2014;7:400–8, <http://dx.doi.org/10.1016/j.nicl.2014.06.015>.
 26. Fiorini S, Verri A, Tacchino A, Ponzio M, Bricchetto G, Barla A. A machine learning pipeline for multiple sclerosis course detection from clinical scales and patient reported outcomes. *Conf Proc IEEE Eng Med Biol Soc.* 2015:4443–6, <http://dx.doi.org/10.1109/EMBC.2015.7319381>.
 27. Kocivar G, Stamile C, Hannoun S, Cotton F, Vukusic S, Durand-Dubief F, et al. Graph theory-based brain connectivity for automatic classification of multiple sclerosis clinical courses. *Front Neurosci.* 2016;10:478, <http://dx.doi.org/10.3389/fnins.2016.00478>.
 28. Zhong J, Chen DQ, Nantes JC, Holmes SA, Hodaje M, Koski L. Combined structural and functional patterns discriminating upper limb motor disability in multiple sclerosis using multivariate approaches. *Brain Imaging Behav.* 2017;11:754–68, <http://dx.doi.org/10.1007/s11682-016-9551-4>.
 29. Ion-Mărgineanu A, Kocivar G, Stamile C, Sima DM, Durand-Dubief F, Van Huffel S, et al. Machine learning approach for classifying multiple sclerosis courses by combining clinical data with lesion loads and magnetic resonance metabolic features. *Front Neurosci.* 2017;11:398, <http://dx.doi.org/10.3389/fnins.2017.00398>.
 30. Lopez C, Tucker S, Salameh T, Tucker C. An unsupervised machine learning method for discovering patient clusters based on genetic signatures. *J Biomed Inform.* 2018;85:30–9, <http://dx.doi.org/10.1016/j.jbi.2018.07.004>.
 31. Van Le H, le Truong CT, Kamau AWC, Holmén J, Fillmore C, Kobayashi MG, et al. Identifying patients with relapsing-remitting multiple sclerosis using algorithms applied to US integrated delivery network healthcare data. *Value Health.* 2019;22:77–84, <http://dx.doi.org/10.1016/j.jval.2018.06.014>.
 32. Mato-Abad V, Labiano-Fontcuberta A, Rodríguez-Yáñez S, García-Vázquez R, Munteanu CR, Andrade-Garda J, et al. Classification of radiologically isolated syndrome and clinically isolated syndrome with machine-learning techniques. *Eur J Neurolog.* 2019;26:1000–5, <http://dx.doi.org/10.1111/ene.13923>.
 33. González-Campo C, Salamone PC, Rodríguez-Arriagada N, Richter F, Herrera E, Bruno D, et al. Fatigue in multiple sclerosis is associated with multimodal interoceptive abnormalities. *Mult Scler.* 2019, <http://dx.doi.org/10.1177/1352458519888881>.
 34. Weygandt M, Hackmack K, Pfüller C, Bellmann-Strobl J, Paul F, Zipp F, et al. MRI pattern recognition in multiple sclerosis normal-appearing brain areas. *PLoS One.* 2011;6:e21138, <http://dx.doi.org/10.1371/journal.pone.0021138>.
 35. Ohanian D, Brown A, Sunquist M, Furst J, Nicholson L, Klebek L, et al. Identifying key symptoms differentiating myalgic encephalomyelitis and chronic fatigue syndrome from multiple sclerosis. *Neurology.* 2016;4:41–5.
 36. Lötsch J, Thrun M, Lerch F, Brunkhorst R, Schiffmann S, Thomas D, et al. Machine-learned data structures of lipid marker serum concentrations in multiple sclerosis patients differ from those in healthy subjects. *Int J Mol Sci.* 2017;18:1217, <http://dx.doi.org/10.3390/ijms18061217>.
 37. Ostmeier J, Christley S, Rounds WH, Toby I, Greenberg BM, Monson NL, et al. Statistical classifiers for diagnosis disease from immune repertoires: A case study using multiple sclerosis. *BMC Bioinformatics.* 2017;18:401, <http://dx.doi.org/10.1186/s12859-017-1814-6>.
 38. Yoo Y, Tang LYW, Brosch T, Li DKB, Kolind S, Vava-sour I, et al. Deep learning of joint myelin and T1w MRI features in normal-appearing brain tissue to distinguish between multiple sclerosis patients and healthy controls. *Neuroimage Clin.* 2017;17:169–78, <http://dx.doi.org/10.1016/j.nicl.2017.10.015>.
 39. Zurita M, Montalba C, Labbé T, Cruz JP, Dalboni da Rocha J, Tejos C, et al. Characterization of relapsing-remitting multiple sclerosis patients using support vector machine classifications of functional and diffusion MRI data. *Neuroimage Clin.* 2018;20:724–30, <http://dx.doi.org/10.1016/j.nicl.2018.09.002>.
 40. Lötsch J, Schiffmann S, Schmitz K, Brunkhorst R, Lerch F, Ferreiros N, et al. Machine-learning based lipid mediator serum concentration patterns allow identification of multiple sclerosis patients with high accuracy. *Sci Rep.* 2018;8:14884, <http://dx.doi.org/10.1038/s41598-018-33077-8>.
 41. Forbes JD, Chen CY, Knox NC, Marrie RA, El-Gabalawy H, de Kievit T, et al. A comparative study of the gut microbiota in immune-mediated inflammatory diseases—does a common dysbiosis exist? *Microbiome.* 2018;6:221, <http://dx.doi.org/10.1186/s40168-018-0603-4>.
 42. Saccà V, Sarica A, Novellino F, Barone S, Tallarico T, Filippelli E, et al. Evaluation of machine learning algorithms performance for the prediction of early multiple sclerosis from resting-state fMRI connectivity data. *Brain Imaging Behav.* 2019;13:1103–14, <http://dx.doi.org/10.1007/s11682-018-9926-9>.
 43. Neeb H, Schenk J. Multivariate prediction of multiple sclerosis using robust quantitative MR-based image metrics. *Z Me Phys.* 2019;29:262–71, <http://dx.doi.org/10.1016/j.zemedi.2018.10.004>.
 44. Ahmadi A, Davoudi S, Daliri MR. Computer aided diagnosis system for multiple sclerosis disease based on phase to amplitude coupling in covert visual attention. *Comput Methods Programs Biomed.* 2019;169:9–18, <http://dx.doi.org/10.1016/j.cmpb.2018.11.006>.
 45. Andersen SL, Briggs FBS, Winnike JH, Natanson Y, Maichle S, Knagge KJ, et al. Metabolome-based signature of disease

- pathology in MS. *Mult Scler Relat Disord*. 2019;31:12–21, <http://dx.doi.org/10.1016/j.msard.2019.03.006>.
46. Pérez del Palomar A, Cegoñino J, Montolio A, Orduna E, Vilades E, Sebastián B, et al. Swept source optical coherence tomography to early detect multiple sclerosis disease The use of machine learning techniques. *PLoS One*. 2019;14:e0216410, <http://dx.doi.org/10.1371/journal.pone.0216410>.
 47. Bang S, Yoo D, Kim SJ, Jhang S, Cho S, Kim H. Establishment and evaluation of prediction model for multiple disease classification based on gut microbial data. *Sci Rep*. 2019;9:10189, <http://dx.doi.org/10.1038/s41598-019-46249-x>.
 48. Azami F, Miri Ashtiani SN, Shalhaf A, Behnam H, Daliri MR. Granger causality analysis in combination with directed network measures for classification of MS patients and healthy controls using task-related fMRI. *Comput Biol Med*. 2019;115:103495, <http://dx.doi.org/10.1016/j.combiomed.2019.103495>.
 49. Cavaliere C, Vilades E, Alonso-Rodríguez MC, Rodrigo MJ, Pablo LE, Miguel JM, et al. Computer-aided diagnosis of multiple sclerosis using a support vector machine and optical coherence tomography features. *Sensors (Basel)*. 2019;19:5323, <http://dx.doi.org/10.3390/s19235323>.
 50. Heming M, Lohmann L, Schulte-Mecklenbeck A, Brix T, Gross CC, Wiendl H, et al. Leukocyte profiles in blood and CSF distinguish neurosarcoidosis from multiple sclerosis. *J Neuroimmunol*. 2020;341:577171, <http://dx.doi.org/10.1016/j.jneuroim.2020.577171>.
 51. Corvol JC, Pelletier D, Henry RG, Caillier SJ, Wang J, Pappas D, et al. Abrogation of T cell quiescence characterizes patients at high risk for multiple sclerosis after the initial neurological event. *Proc Natl Acad Sci USA*. 2008;105:11839–44, <http://dx.doi.org/10.1073/pnas.0805065105>.
 52. Wottschel V, Alexander DC, Kwok PP, Chard DT, Stromillo ML, De Stefano N, et al. Predicting outcome in clinically isolated syndrome using machine learning. *Neuroimage Clin*. 2014;7:281–7, <http://dx.doi.org/10.1016/j.nicl.2014.11.021>.
 53. Baranzini SE, Madiredy LR, Cromer A, D'Antonio M, Lehr L, Beelke M, et al. Prognostic biomarkers of IFN β therapy in multiple sclerosis patients. *Mult Scler*. 2015;21:894–904, <http://dx.doi.org/10.1177/1352458514555786>.
 54. Zhao Y, Healy BC, Rotstein D, Guttmann CR, Bakshi R, Weiner HL, et al. Exploration of machine learning techniques in predicting multiple sclerosis disease course. *PLoS One*. 2017;12:e0174866, <http://dx.doi.org/10.1371/journal.pone.0174866>.
 55. Kiiski H, Jollans L, Donnchadha SÓ, Nolan H, Lonergan R, Kelly S, et al. Machine learning EEG to predict cognitive functioning and processing speed over a 2-year period in multiple sclerosis patients and controls. *Brain Topogr*. 2018;31:346–63, <http://dx.doi.org/10.1007/s10548-018-0620-4>.
 56. Zhang H, Alberts E, Pongratz V, Mühlau M, Zimmer C, Wiestler B, et al. Predicting conversion from clinically isolated syndrome to multiple sclerosis- An imaging-based machine learning approach. *Neuroimage Clin*. 2019;21:101593, <http://dx.doi.org/10.1016/j.nicl.2018.11.003>.
 57. Gafone P, Mazzone E, Mammata S, Di Marco R, Spinascia F, Basile MS, et al. Identification of CD4+ T cell biomarkers for predicting the response of patients with relapsing-remitting multiple sclerosis to natalizumab treatment. *Mol Med Rep*. 2019;20:678–84, <http://dx.doi.org/10.3892/mmr.2019.10283>.
 58. Goyal M, Khanna D, Rana PS, Khaibullin T, Martynova E, Rizyanov AA, et al. Computational intelligence technique for prediction of multiple sclerosis based on serum cytokines. *Fron Neurol*. 2019;10:781, <http://dx.doi.org/10.3389/fneur.2019.00781>.
 59. Sun R, Hsieh KL, Sosnoff JJ. Fall risk prediction in multiple sclerosis using postural sway measures: A machine learning approach. *Sci Rep*. 2019;9:16154, <http://dx.doi.org/10.1038/s41598-019-52697-2>.
 60. Law MT, Traboulee AL, Li DK, Carruthers RL, Freedman MS, Kolind SH, et al. Machine learning in secondary progressive multiple sclerosis: An improved predictive model for short-term disability progression. *Mult Scler J Exp Transl Clin*. 2019;5, <http://dx.doi.org/10.1177/2055217319885983>, 2055217319885983.
 61. Wottschel V, Chard DT, Enzinger C, Filippi M, Frederiksen JL, Gasperini C, et al. SVM recursive feature elimination analyses of structural brain MRI predicts near-term relapses in patients with clinically isolated syndromes suggestive of multiple sclerosis. *Neuroimage Clin*. 2019;24:102011, <http://dx.doi.org/10.1016/j.nicl.2019.102011>.
 62. Jackson KC, Sun K, Barbour C, Hernández D, Kosa P, Tanigawa M, et al. Genetic model of MS severity predicts future accumulation of disability. *Ann Hum Genet*. 2020;84:1–10, <http://dx.doi.org/10.1111/ahg.12342>.
 63. Ebrahimkhani S, Beadnall HN, Wang C, Suter CM, Barnett MH, Buckland ME, et al. Serum exosome microRNAs predict multiple sclerosis disease activity after fingolimod treatment. *Mol Neurobiol*. 2020;57:1245–58, <http://dx.doi.org/10.1007/s12035-019-01792-6>.
 64. Ghafouri-Fard S, Taheri M, Omrani MD, Daaee A, Mohammad-Rahimi H. Application of artificial neural network for prediction of risk of multiple sclerosis based on single nucleotide polymorphism genotypes. *J MOL Neurosci*. 2020, <http://dx.doi.org/10.1007/s12031-020-01514-x>.
 65. Seccia R, Gammelli D, Dominici F, Romano S, Landi AC, Salvetti M, et al. Considering patient clinical history impacts performance of machine learning models in predicting course of multiple sclerosis. *PLoS One*. 2020;15:e0230219, <http://dx.doi.org/10.1371/journal.pone.0230219>.
 66. Yperman J, Becker T, Valkenburg D, Popescu V, Hellings N, Wiimeersch BV, et al. Machine learning analysis of motor evoked potential time series to predict disability progression in multiple sclerosis. *BMC Neurol*. 2020;20:105, <http://dx.doi.org/10.1186/s12883-020-01672-w>.
 67. Jin T, Wang C, Tian S. Feature selection based on differentially correlated gene pairs reveals the mechanism of IFN β therapy for multiple sclerosis. *PeerJ*. 2020;8:e8812, <http://dx.doi.org/10.7717/peerj.8812>.
 68. Kosa P, Ghazali D, Tanigawa M, Barbour C, Cortese I, Kelley W, et al. Development of a sensitive outcome for economical drug screening for progressive multiple sclerosis treatment. *Fron Neurol*. 2016;7:131, <http://dx.doi.org/10.3389/fneur.2016.00131>.
 69. Høgestøl EA, Kaufmann T, Nygaard GO, Beyer MK, Sowa P, Nordvik JE, et al. Cross-sectional and longitudinal MRI brain scans reveal accelerated brain aging in multiple sclerosis. *Fron Neurol*. 2019;10:450.
 70. McGinnis RS, Mahadevan N, Moon Y, Seagers K, Sheth N, Wright JA Jr, et al. A machine learning approach for gait speed estimation using skin-mounted wearable sensors: From healthy controls to individuals with multiple sclerosis. *PLoS One*. 2017;12:e0178366, <http://dx.doi.org/10.1371/journal.pone.0178366>.
 71. Chitnis T, Glanz BI, Gonzalez C, Healy BC, Saraceno TJ, Sattarnehzad N, et al. Quantifying neurologic disease using biosensor measurements in-clinic and in free-living settings in multiple sclerosis. *NPJ Digit Med*. 2019;2:123, <http://dx.doi.org/10.1038/s41746-019-0197-7>.
 72. Flachenecker F, Gaßner H, Hannik J, Lee DH, Flache-necker P, Winkler J, et al. Objective sensor-based gait measures reflect motor impairment in multiple sclerosis patients: Reliability and clinical validation of a wearable sensor device. *Mult Scler Relat Disord*. 2019;39:101903, <http://dx.doi.org/10.1016/j.msard.2019.101903>.
 73. Vandermeeren S, Bruneel H, Steendam H. Feature selection for machine learning based step length estimation algorithms. *Sensors (Basel)*. 2020;20:778, <http://dx.doi.org/10.3390/s20030778>.

74. Mowry EM, Hedström AK, Gianfrancesco MA, Shao X, Schaefer CA, Shen L, et al. Incorporating machine learning approaches to assess putative environmental risk factors for multiple sclerosis. *Mult Scler Relat Disord.* 2018;24:135–41, <http://dx.doi.org/10.1016/j.msard.2018.06.009>.
75. Briggs FBS, Yu JC, Davis MF, Jiangyang J, Fu S, Parrotta E, et al. Multiple sclerosis risk factors contribute to onset heterogeneity. *Mult Scler Relat Disord.* 2019;28:11–6, <http://dx.doi.org/10.1016/j.msard.2018.12.007>.
76. Fluzino T, Simão ANC, de Carvalho Jennings Pereira WL, Alfieri DF, Oliveira SR, Kallaur AP, et al. Disability in multiple sclerosis is associated with age and inflammatory, metabolic and oxidative/nitrosative stress biomarkers: Results of multivariate and machine learning procedures. *Metab Brain Dis.* 2019;34:1401–13, <http://dx.doi.org/10.1007/s11011-019-00456-7>.
77. Karmonik C, Boone T, Khavari R. Data-driven machine-learning quantifies differences in the voiding initiation network in neurogenic voiding dysfunction in women with multiple sclerosis. *Neurourol J.* 2019;23:195–204, <http://dx.doi.org/10.5213/inj.1938058.029>.
78. De Andrés-Galiana EJ, Bea G, Fernández-Martínez JL, Saligan LN. Analysis of defective pathways and drug repositioning in multiple sclerosis via machine learning approaches. *Comput Biol Med.* 2019;115:103492, <http://dx.doi.org/10.1016/j.compbiomed.2019.103492>.
79. Mezzaroba L, Simão ANC, Oliveira SR, Fluzino T, Alfieri DF, de Carvalho Jennings Pereira WL, et al. Antioxidant and anti-inflammatory diagnostic biomarkers in multiple sclerosis: A machine learning study. *Mol Neurobiol.* 2020;57:2167–78, <http://dx.doi.org/10.1007/s12035-019-01856-7>.
80. Tacchella A, Romano S, Ferraldeschi M, Salvetti M, Zaccaria A, Crisanti A, et al. Collaboration between a human group and artificial intelligence can improve prediction of multiple sclerosis course: A proof-of-principle study. *F1000Res.* 2017;6:2172, <http://dx.doi.org/10.12688/f1000research.13114.2>.
81. Liu J, Brodley CE, Healy BC, Chitnis T. Removing confounding factors via constraint-based clustering: An application to finding homogeneous groups of multiple sclerosis patients. *Artif Intell Med.* 2015;65:79–88, <http://dx.doi.org/10.1016/j.artmed.2015.06.004>.
82. He Y, Carass A, Liu Y, Jedynek BM, Solomon SD, Saidha S, et al. Deep learning based topology guaranteed surface and MME segmentation of multiple sclerosis subjects from retinal OCT. *Biomed Opt Express.* 2019;10:5042–58, <http://dx.doi.org/10.1364/BOE.10.005042>.
83. Birchmeier ME, Studer T. Automated rating of multiple sclerosis test results using a convolutional neural network. *Stud Health Technol Inform.* 2019;259:105–8.
84. Kalincik T, Manouchehrinia A, Sobisek L, Jokubaitis V, Spelman T, Horakova D, et al. Towards personalized therapy for multiple sclerosis: prediction of individual treatment response. *Brain.* 2017;140:2426–43, <http://dx.doi.org/10.1093/brain/awx185>.
85. Baecher-Allan C, Kaskow BJ, Weiner HL. Multiple sclerosis: Mechanisms and immunotherapy. *Neuron.* 2018;97:742–68, <http://dx.doi.org/10.1016/j.neuron.2018.01.021>.
86. Ho D, Quake SR, McCabe ERB, Chng WJ, Chow EK, Ding X, et al. A enabling technologies for personalized and precision medicine. *Trends Biotechnol.* 2020;38:497–518, <http://dx.doi.org/10.1016/j.tibtech.2019.12.021>.