UAB
Universitat Autònoma de Barcelona

CLINICAL MONITORING OF MULTIPLE SCLEROSIS PATIENTS USING DIGITAL TECHNOLOGY

Luciana Midaglia Fernández

Universitat Autònoma de Barcelona

Doctoral Program in Medicine

Department of Medicine

Doctoral Thesis

CLINICAL MONITORING OF MULTIPLE SCLEROSIS PATIENTS USING DIGITAL TECHNOLOGY

Author:

Luciana Midaglia Fernández

Supervisors:

Xavier Montalban Gairín

Jaume Sastre Garriga

Tutor:

Vicenç Fonollosa Pla

Barcelona, 2022

Vall d'Hebron Barcelona Hospital Campus | Neurociències

Cemcat Centre d'Esclerosi Múltiple de Catalunya

AGRADECIMIENTOS

En primer lugar me gustaría agradecer a la Dra Delicias Muñoz, quien fue neuróloga adjunta del Servicio de Neurología del Complejo Hospitalario Universitario de Vigo durante mis años de residencia en dicho hospital. Ella me transmitió y contagió el espíritu de trabajar con auténtica pasión. Me enseñó que aunque las condiciones externas no sean las más favorables también se puede desarrollar el espíritu científico que uno lleva dentro y aportar, aunque sea en pequeña escala, un granito más de conocimiento a la ciencia. Gracias por confiar en mí desde un principio y por adentrarme de tu mano en este maravilloso mundo de la esclerosis multiple.

Quiero agradecer a los Dres Xavier Montalban y Mar Tintoré por abrirme las puertas e invitarme a formar parte del Cemcat, uno de los principales centros de referencia en esclerosis multiple a nivel mundial. El trabajar junto a ellos es la culminación de un auténtico sueño cultivado durante muchos años. Es un aprendizaje y crecimiento profesional sin límites. Muchas gracias por haberlo hecho posible.

Gracias al Dr Jaume Sastre-Garriga por el apoyo y seguimiento constante en la elaboración y desarrollo de esta tesis doctoral. Agradecer también al profesor Vicenç Fonollosa por aceptar ser mi tutor, como un gesto de confianza en mí y por supuesto en todo el equipo del Cemcat.

Gracias infinitas a mi familia (mis padres y hermanas) por el amor incondicional de cada día y por los valores inculcados. Por dejarme claro que estuvisteis, estáis y estaréis siempre, en los buenos y malos momentos. Y que aunque nos separe un océano, no hay distancia que pueda contra nuestro amor.

Gracias a mi compañero de vida, Xavier, por acompañarme y apoyarme en mis proyectos profesionales, pero sobre todo por forma parte de mis proyectos vitales.

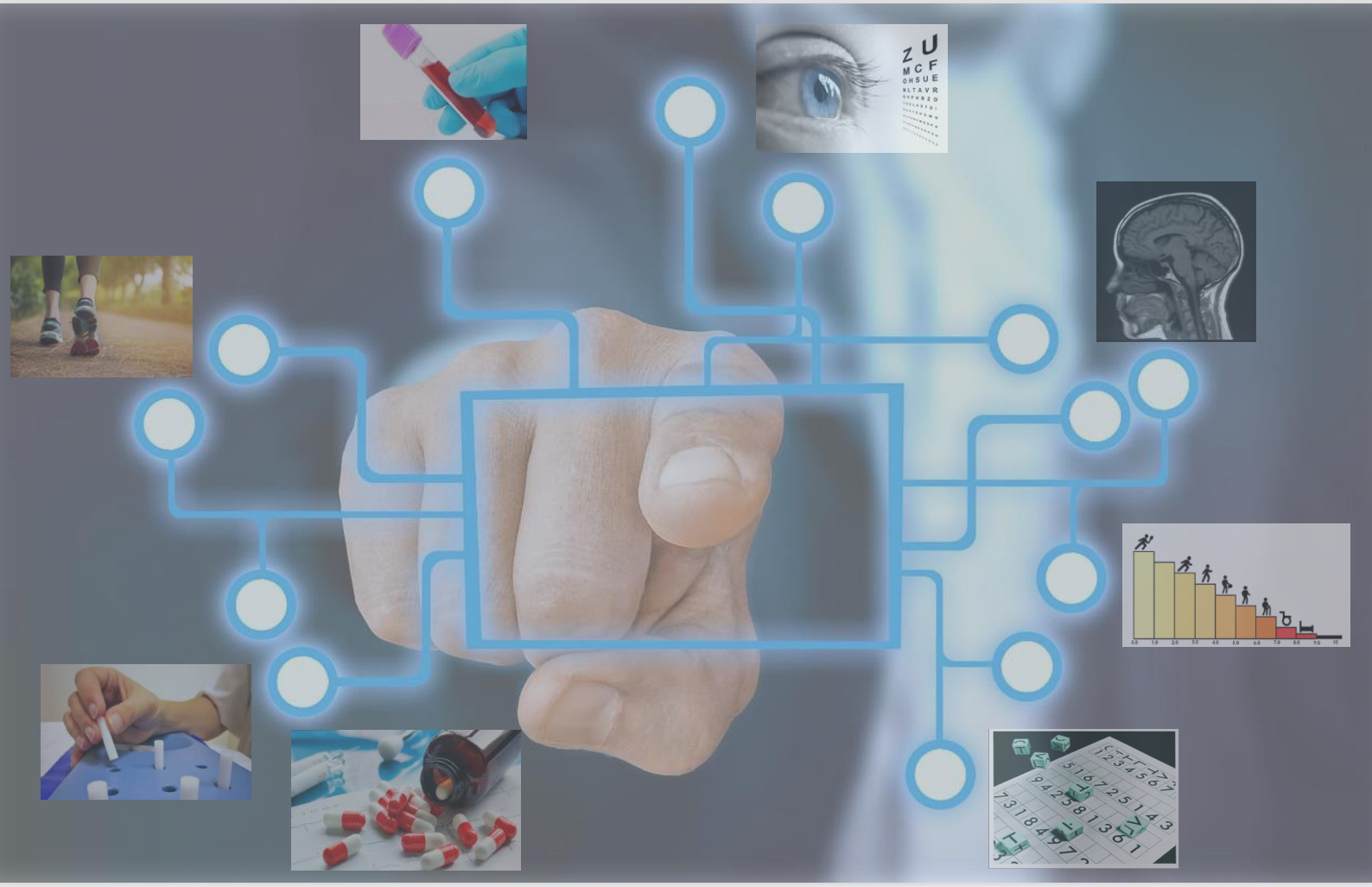
Finalmente agradecer a los pacientes, sus familias, al personal del Cemcat y a la compañía farmacéutica Hoffmann-La Roche Ltd que han apoyado y contribuido al desarrollo de este estudio.

ABBREVIATIONS

antiMOG	anti–Myelin Oligodendrocyte Glycoprotein
App	Application
BBB	Blood-Brain Barrier
BBS	Berg Balance Scale
BMS	Benign multiple sclerosis
CI	Confidence interval
CIS	Clinically isolated syndrome
CNS	Central Nervous System
CSF	Cerebrospinal fluid
CST	Contrast Sensitivity Test
CV	Coefficient of variation
DIS	Dissemination in space
DIT	Dissemination in time
DMD	Disease-modifying drug
DSS	Disability Status Scale
e-health	Elecronic-health
EAE	Experimental autoimmune encephalomyelitis
EBV	Epstein Barr Virus
EBNA-2	Epstein Barr virus nuclear antigen 2
EDSS	Expanded Disability Status Scale
EMA	European Medicines Agency
FAS	Full analysis set
FDR	False discovery rate
FLAIR	Fluid attenuated inversion recovery
FS	Functional Systems
FSMC	Fatigue Scale for Motor and Cognitive Functions
Gd	Gadolinium
GM	Gray matter
HRQoL	Health related QoL
ICC	Intraclass correlation coefficient

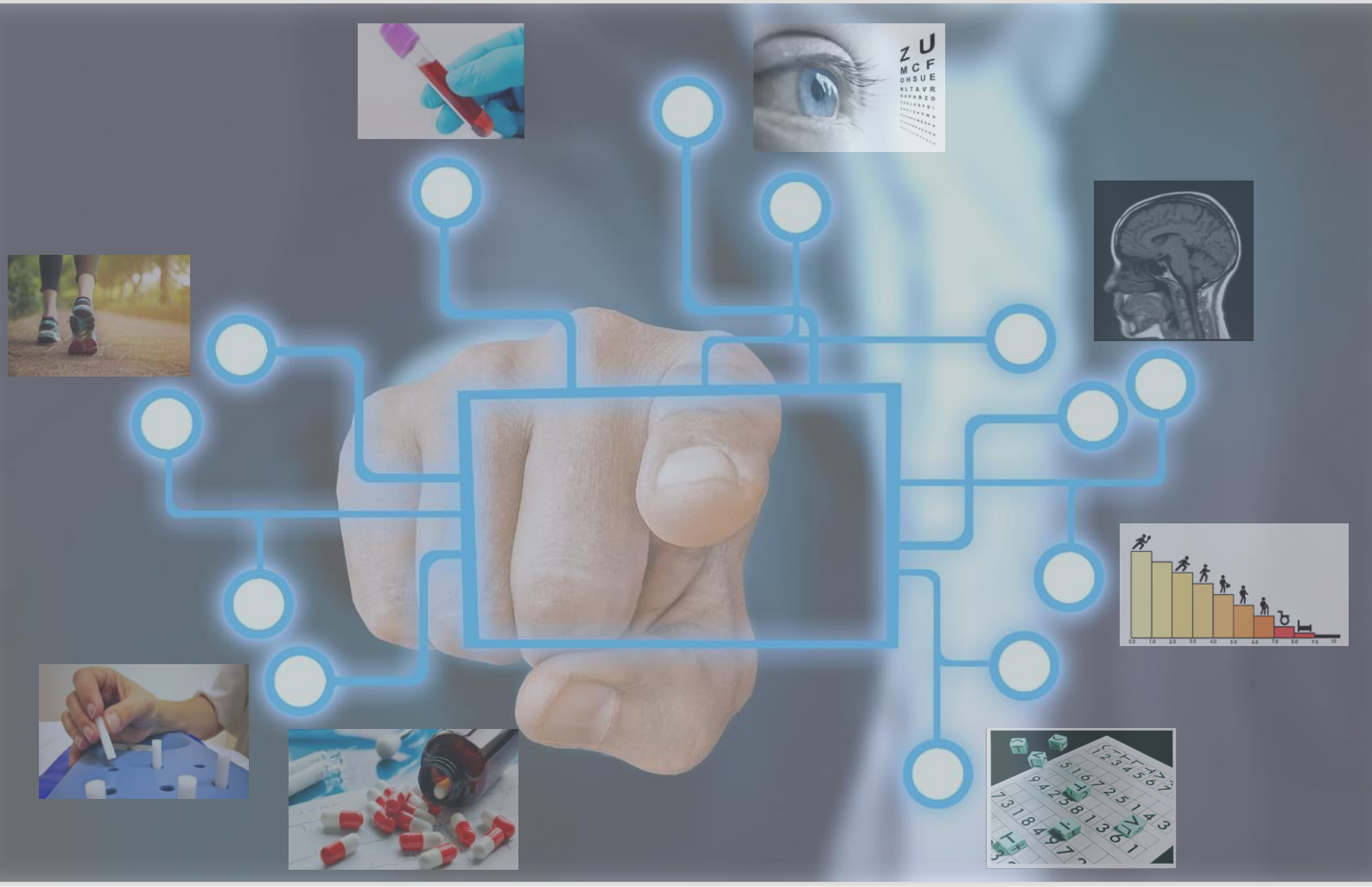
IFN	Interferon
MBP	Myelin basic protein
MDT	Manual Dexterity Test
MHC	Major histocompatibility complex
MMP	Matrix metalloproteinase
MOG	Myelin oligodendrocyte glycoprotein
MRI	Magnetic resonance imaging
MS	Multiple sclerosis
MSFC	Multiple Sclerosis Functional Composite
MSIS-29	Multiple Sclerosis Impact Scale–29
MS PATHS	Multiple Sclerosis Partners Advancing Technology Health Solutions
MSPT	Multiple Sclerosis Performance Test
Neuro-QoL	Quality of life in Neurological disorders questionnaire
NMOSD	Neuromyelitis Optica Spectrum Disorders
NMSS	National Multiple Sclerosis Society
OCB	Oligoclonal bands
PASAT	Paced Auditory Serial Addition test
PDDS	Patient Determined Disease Steps
PHQ-9	Patient Health Questionnaire–9
PIRA	Progression independent of relapse activity
PLP	Proteolipid protein
PML	Progressive multifocal leukoencephalopathy
PPMS	Primary progressive multiple sclerosis
PROs	Patient reported outcomes
PROMs	Patient-reported outcome measures
PST	Processing Speed Test
PwMS	People with multiple sclerosis
QoL	Quality of life
RADAR-CNS	Remote Assessment of Disease and Relapse-Central Nervous System
RAW	Relapse-associated worsening

RIS	Radiologically Isolated Syndrome
RMT	Remote measurement technologies
RP	Relapsing-progressive
RRMS	Relapsing-remitting multiple sclerosis
S1P	Sphingosine-1-phosphate
SBT	Static Balance Test
SDMT	Symbol Digit Modalities Test
SPMS	Secondary progressive multiple sclerosis
T25FW	Timed 25-Foot Walk
TE	Echo time
TI	Inversion time
TR	Repetition time
UTT	U-Turn Test
WST	Walking Speed Test
3D	Three-dimensional
9HPT	Nine-Hole Peg Test
2MWT	Two-Minute Walk Test



INDEX

INDEX



ABSTRACT

ABSTRACT

The present work deals with the search for digital markers for the monitoring of multiple sclerosis (MS).

Background: Current clinical assessments of people with MS (pwMS) are incomplete, episodic and may miss fundamental features of functional fluctuations and changes between visits. Sensor-based monitoring tools may fill a critical gap in MS research and clinical care.

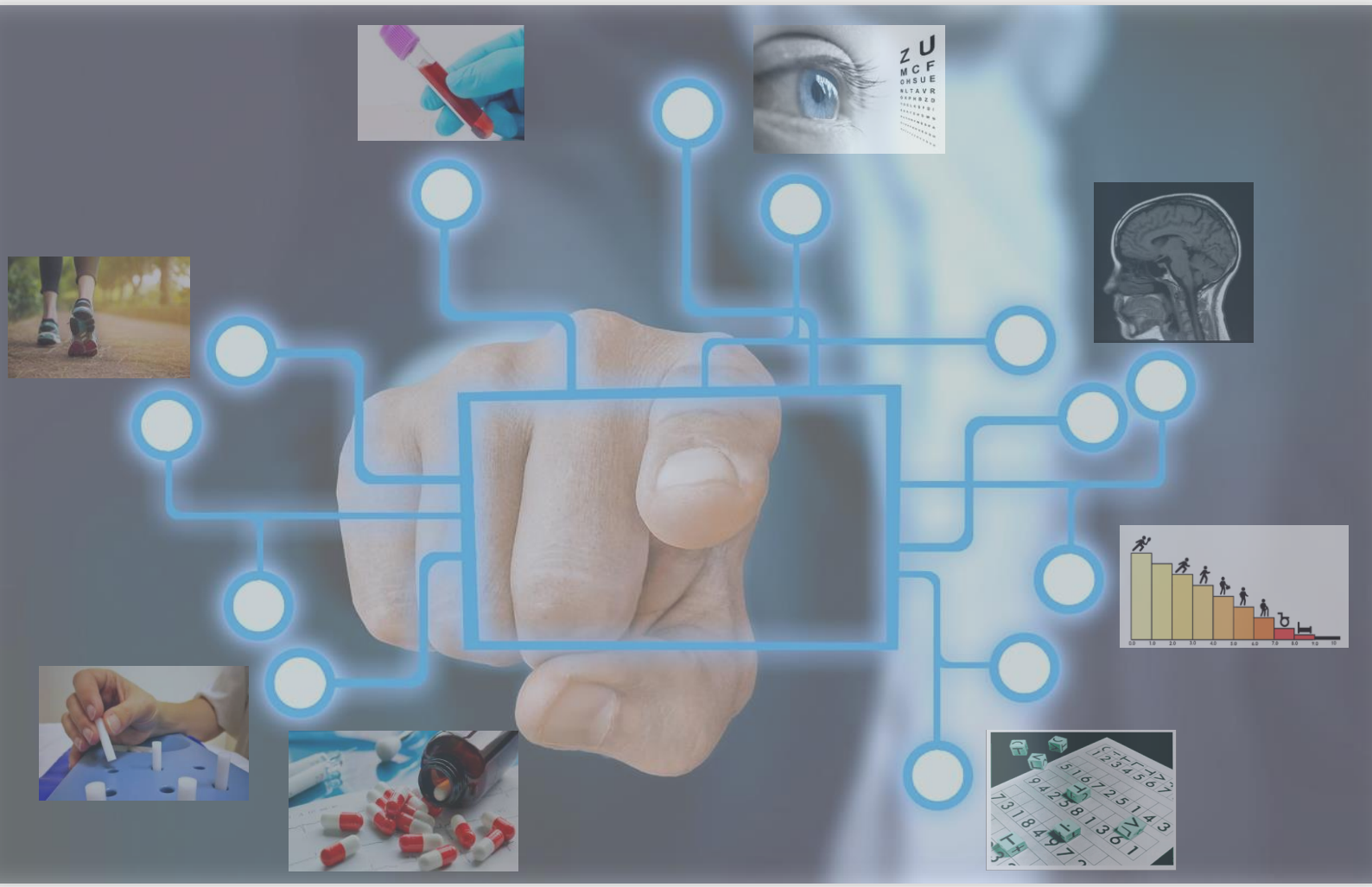
Objectives: We aimed (1) to assess the feasibility of remote active testing and passive monitoring using smartphones and smartwatch technology in pwMS with respect to adherence and satisfaction with the Floodlight MS application (app). (2) To determine the association between exploratory sensor-based outcomes (Floodlight MS app) and conventional MS clinical outcomes. (3) To explore regional neural correlates of digital measures from smartphone sensor-based tests in pwMS.

Methods: We conducted a prospective pilot study including pwMS (aged 20 to 57 years; Expanded Disability Status Scale [EDSS] 0-5.5; n=76) and healthy controls (n=25). All participants performed the Floodlight MS app, comprising active tests (daily, weekly, every two weeks, or on demand) for assessment of cognition (electronic Symbol Digit Modalities Test), upper extremity function (Pinching Test, Draw a Shape Test), and gait and balance (Static Balance Test, U-Turn Test, Walk Test) and passive monitoring (sensor-based gait and mobility) for 24 weeks using a smartphone and smartwatch. (1) The primary analysis assessed adherence (proportion of weeks with at least 3 days of completed testing and 4 hours per day passive monitoring) and questionnaire-based satisfaction. In-clinic assessments including standard clinical measurements and magnetic resonance imaging (MRI) were performed. (2) Intraclass correlation coefficients (ICCs) and age- or sex-adjusted Spearman's rank correlation determined test-retest reliability and correlations with clinical and MRI outcome measures, respectively. (3) Finally, digital measures and standard clinical measures of MS disease state

were correlated against regional and structural MRI outcomes acquired with icobrain ms v.5.0 (icometrix, Leuven, Belgium), including normalized brain volumes for 36 anatomical regions, using univariate Spearman's rank correlation.

Results: (1) PwMS showed 70% (16.68/24 weeks) adherence to active tests and 79% (18.89/24 weeks) to passive monitoring; satisfaction score was on average 73.7 out of 100. Neither adherence nor satisfaction was associated with specific population characteristics. Test-battery assessments had an at least acceptable impact on daily activities in over 80% (61/72) of pwMS. (2) In pwMS, ICCs were moderate-to-good (ICC = 0.61-0.85) across tests. Correlations with domain-specific standard clinical disability measures were significant for all tests in the cognitive (r = 0.82, p < 0.001), upper extremity function (|r|= 0.40/0.64, all p < 0.001), and gait and balance domains (r = -0.25/-0.52, all p<0.05; except for Static Balance Test: r = -0.20, p>0.05). Most tests also correlated with EDSS, 29-item Multiple Sclerosis Impact Scale items or subscales, and normalized brain volume. (3) Worse performance on digital and clinical measures was associated with smaller regional brain volumes and larger ventricular volumes. While digital and clinical measures had many neural correlates in common, some were observed only for digital measures. For example, accuracy over time on the Draw a Shape Test and the time difference between fingers touching the screen on the Pinching Test, but not the Nine-Hole Peg Test (9HPT), correlated with volume of the hippocampus (r = 0.37 [Draw a Shape]/-0.45 [Pinching Test]), thalamus (r = 0.38/-0.41), and pons (r = 0.35/-0.35).

Conclusions: (1) PwMS were engaged and satisfied with the Floodlight MS app. (2) The Floodlight MS app captures reliable and clinically relevant measures of functional impairment in MS. (3) Multiple neural correlates were identified for the digital measures. Some correlates were only observed for the digital but not the standard clinical measures, suggesting that digital measures may yield higher functional specificity.



RESUMEN

RESUMEN

El presente trabajo se basa en la búsqueda de marcadores digitales para el seguimiento de la esclerosis múltiple (EM).

Antecedentes: Las evaluaciones clínicas actuales de personas con EM (pwMS) son incompletas, episódicas y pueden pasar por alto características fundamentales como fluctuaciones y cambios funcionales entre visitas. Las herramientas de monitorización basadas en sensores pueden llenar un vacío crítico en la investigación y la atención clínica de la EM.

Objetivos: Nuestros objetivos eran (1) evaluar la viabilidad de las pruebas activas y la supervisión pasiva de manera remota utilizando teléfonos y relojes inteligentes en pwMS con respecto a la adherencia y la satisfacción con la aplicación (app) Floodlight EM. (2) Determinar la asociación entre los resultados exploratorios basados en sensores (app Floodlight EM) y los resultados obtenidos de la clínica de la EM convencional. (3) Explorar correlatos neuronales regionales de las medidas digitales obtenidas de pruebas basadas en sensores de teléfonos inteligentes en pwMS.

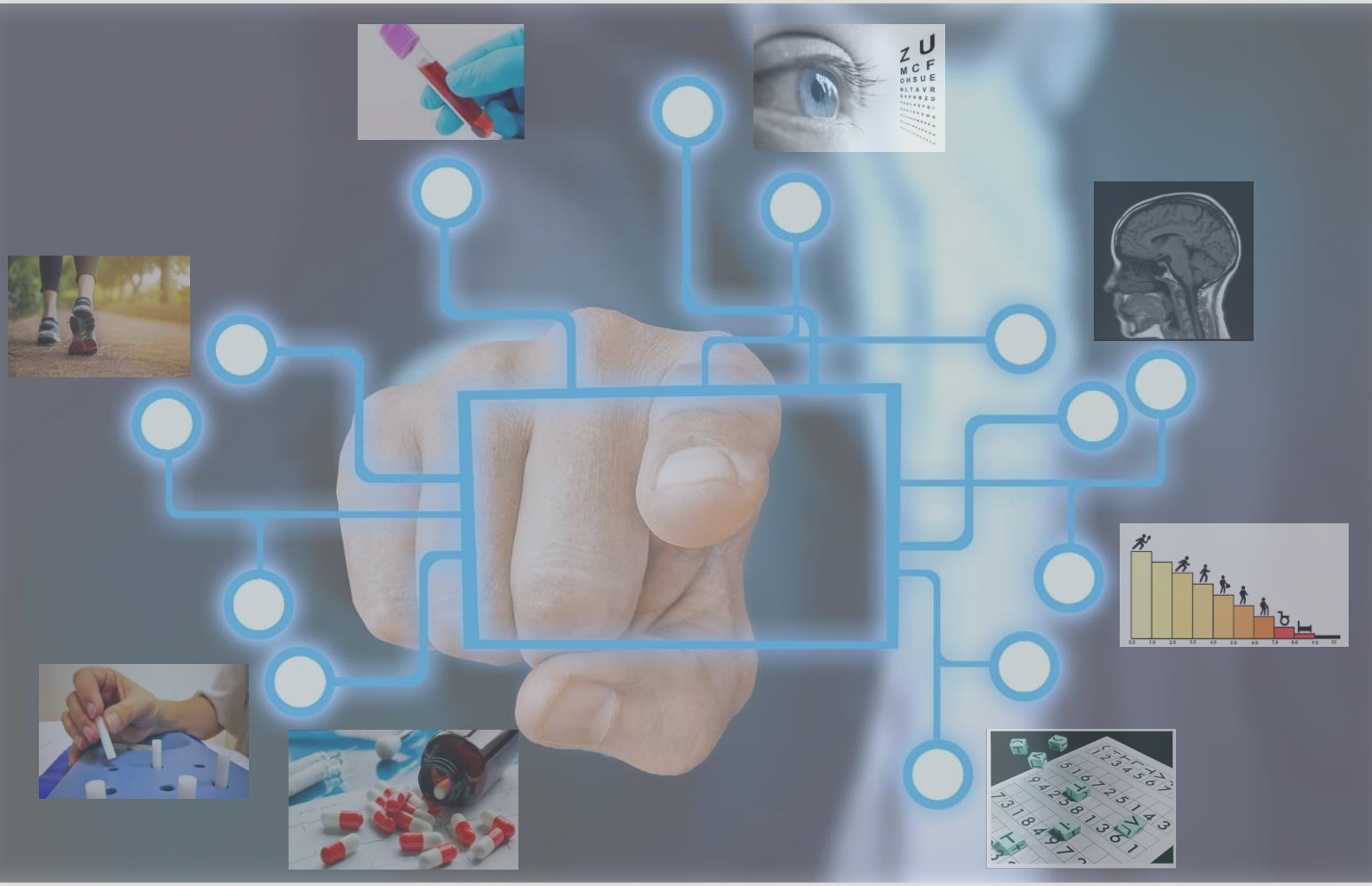
Métodos: Realizamos un estudio piloto prospectivo que incluyó pwMS (edad de 20 a 57 años; Escala de estado de discapacidad expandida [EDSS] 0-5.5; n = 76) y controles sanos (n = 25). Todos los participantes utilizaron la app Floodlight EM, que comprende pruebas activas (diarias, semanales, cada dos semanas o a demanda) diseñadas para evaluar la cognición (prueba electrónica de modalidades de dígitos y símbolos), la función de las extremidades superiores (prueba de pellizco, prueba de dibujar una forma) y la marcha y el equilibrio (prueba de equilibrio estático, prueba de giro en U, prueba de marcha) así como la monitorización pasiva (marcha y movilidad basadas en sensores) durante 24 semanas utilizando un teléfono y reloj inteligentes. (1) El primer análisis evaluó la adherencia (proporción de semanas con al menos 3 días de prueba completa y 4 horas por día de monitorización pasiva) y la satisfacción a través de un cuestionario. A su vez, durante las visitas en el centro se realizaron evaluaciones

clínicas estándar y de resonancia magnética (RM). (2) La fiabilidad test-retest y las correlaciones con las medidas de resultado clínicas y de RM se determinaron mediante los coeficientes de correlación intraclase (ICC) y la correlación de rango de Spearman ajustada por edad o sexo, respectivamente. (3) Finalmente, las medidas digitales y las medidas clínicas estándar del estado de la enfermedad (EM) se correlacionaron con los resultados de RM regional y estructural adquiridos con icobrain ms v.5.0 (icometrix, Lovaina, Bélgica), incluyendo los volúmenes cerebrales normalizados para 36 regiones anatómicas, y utilizando la correlación de rango univariada de Spearman.

Resultados: (1) PwMS mostraron 70% (16,68/24 semanas) de adherencia a las pruebas activas y 79% (18,89/24 semanas) a la monitorización pasiva; la puntuación de satisfacción fue en promedio de 73,7 sobre 100. Ni la adherencia ni la satisfacción se asociaron con características específicas de la población en estudio. Las evaluaciones de la app tuvieron un impacto aceptable en las actividades diarias en más del 80% (61/72) de pwMS. (2) En pwMS, los ICC fueron de moderados a buenos (ICC = 0,61-0,85) en todas las pruebas. Las correlaciones con las medidas estándar de discapacidad clínica específicas para cada dominio fueron significativas para todas las pruebas en los dominios cognitivo ($r = 0,82$, $p < 0,001$), función de las extremidades superiores ($|r| = 0,40/0,64$, todos $p < 0,001$) y marcha y equilibrio ($r = -0,25/-0,52$, todos $p<0,05$; excepto para la prueba de equilibrio estático: $r = -0,20$, $p>0,05$). La mayoría de las pruebas también se correlacionaron con el EDSS, los ítems o subescalas de la escala de impacto de la esclerosis múltiple de 29 ítems y el volumen cerebral normalizado. (3) El peor rendimiento en las medidas digitales y clínicas se asoció con volúmenes cerebrales regionales más pequeños y volúmenes ventriculares más grandes. Si bien las medidas digitales y clínicas tenían muchos correlatos neuronales en común, algunos se observaron solo para las medidas digitales. Por ejemplo, la precisión a lo largo del tiempo en la prueba de dibujar una forma y la diferencia de tiempo entre los toques de la pantalla con los dedos en la prueba de pellizco, pero no en la prueba de las clavijas y nueve agujeros (9HPT), se correlacionaron con el volumen del hipocampo ($r = 0,37$ [prueba de dibujar una

forma]/-0,45 [prueba de pellizco]), tálamo (r = 0,38/-0,41) y protuberancia (r = 0,35/-0,35).

Conclusiones: (1) PwMS estaban comprometidos y satisfechos con la app Floodlight EM. (2) La app Floodlight EM captura medidas confiables y clínicamente relevantes de deterioro funcional en la EM. (3) Múltiples correlatos neuronales fueron identificados para las medidas digitales. Algunas correlaciones solo se observaron para las medidas digitales pero no para las medidas clínicas estándar, lo que sugiere que las medidas digitales pueden proporcionar una mayor especificidad funcional.



INTRODUCTION

1. INTRODUCTION

1.1 INTRODUCTION TO MS

Multiple sclerosis (MS) is an inflammatory-demyelinating and neurodegenerative disease of the central nervous system (CNS) with an autoimmune mechanism, which predominantly affects young adults, being the main non-traumatic cause of disability in this population. It affects more than 2 million people worldwide (at least 400,000 in the United States and 50,000 in Spain)[1-2].

1.1.1 Epidemiology

Age and gender. Being the first demyelinating attack that could lead to the development of MS, a clinically isolated syndrome (CIS) occurs more typically in young adults aged 20-40 years in 70.0% of cases with a mean age of 30 years[3], but suggestive symptoms can present at older and younger ages. In most populations, women are affected more commonly than men, with the female-to-male ratio varying between 1.5:1 and 2.5:1, with a trend toward higher values in the most recent studies, probably related to lifestyle changes and environmental factors. Although relapsing-remitting MS (RRMS) is the most common clinical phenotype, 10-15% of patients have a progressive course from the beginning of the disease, with or without associated relapses. Patients with primary progressive MS (PPMS) are mostly older at onset and the gender difference in this population tends to be smaller[4-8].

Geography. Both incidence and prevalence vary (Fig. 1.1.1A): areas with a high prevalence (60 per 100000 inhabitants or more) include Europe, Southern Canada, Northern United States, New Zealand, and Southeast Australia[9-10]. This could be partly due to racial differences: people of northern European origin appear to be the most susceptible, whereas people of Asian, American Indian or African ancestry have the lowest risk[11-13]. Nevertheless, an increase in prevalence

has been observed in Latin America[11] and a higher incidence of MS was determined in African Americans, with black women having the highest risk of evolving to MS. The reasons behind the increasing incidence are not fully understood, but improvements in hygiene and a decrease in vitamin D are likely contributors[13]. Besides, persons who migrate from a high to a low-risk area after puberty retain their former risk, whilst those who migrate at earlier ages appear to acquire the risk of the new area to which they migrated[14], suggesting that exposure to certain environmental factors at a young age may be key to developing the disease. There has been an attenuation of the north-south gradient after the 1980s, probably due to an increased incidence in lower latitudes[5,15]. Conversely, although prevalence has increased in Australia, New Zealand, Western Europe, and North America, incidence has not in the latter two regions[5], meaning that the observed latitudinal gradient of prevalence could be explained by factors such as survival time or diagnostic accuracy.

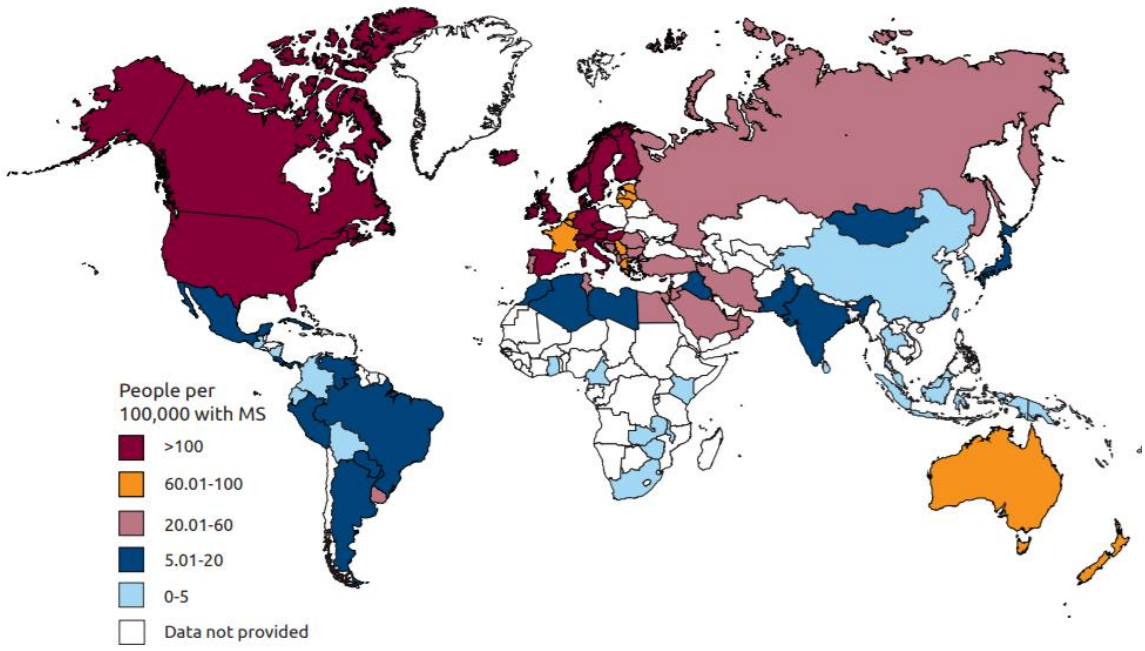


Figure 1.1.1A. Worldwide prevalence of MS. The prevalence of multiple sclerosis (MS) varies between countries. In general, the prevalence of MS is higher in countries of higher latitude and in Western countries. The MS International Federation's Atlas of MS, 2013. © www.atlasofms.org, MSIF 2013.

1.1.2 Risk factors

It is not known whether MS has a single or multiple causes, and rarely (if ever) has a specific etiologic trigger been identified. Nonetheless, various genetic and environmental risk factors have been found[1] (Fig. 1.1.2A).

Genetic risk factors. People with an affected first-degree relative have a 2 to 4% risk of MS (as compared with approximately 0.1% risk in the general population), and concordance in monozygotic twins is 30 to 50%. Different genetic studies, based on available samples from thousands of MS patients and matched controls, have identified more than 200 genetic variants that increase the risk of the disease, of which the most significant continues being the HLA Haplotype DRB1 * 1501 (with an odds ratio of approximately 3). Most risk alleles are associated with genes of the immune pathway, especially with class II alleles of the major histocompatibility complex (MHC). In addition, the risk of MS may vary according to the type of HLA-DRB1 allele, with a protective effect being observed with the HLA-DRB1 * 04, * 07, and * 09 alleles, while a high risk is observed with the DRB1 * 15, * 16, and * 08 alleles. The risk of MS is also associated with other non-MHC susceptibility genes, such as IL2RA, IL7RA and IRF8[16-18]. Epigenetic changes affecting gene expression in response to external influence represent the link between non-genetic and genetic risk factors, which should be extensively studied to improve the knowledge of MS molecular mechanisms[19].

Environmental risk factors

- Vitamin D. There is a close relationship between latitude and exposure to sunlight and therefore with the resulting metabolism of vitamin D[20]. Several studies have shown that taking vitamin D supplements reduces the risk of acquiring MS or that serum levels of 25-hydroxy vitamin D are inversely associated with the risk of MS[21-22].

- Infections. Neurotropic viruses have been studied as possible triggers for MS. To date, there is no specific evidence directly linking viruses to the development of MS, but there is increasing attention on the Epstein Barr Virus (EBV) as a cause

or trigger of MS, observing a relative risk increase of 2.3 times [95% confidence interval (CI) 1.7-3.0] after mononucleosis infectious and elevations of anti-EBV antibody titres before MS onset, especially antibodies against the EBV nuclear antigen 2 (EBNA-2)[23-26]. However, proving a link between EBV and MS is complex since serological evidence of EBV exposure can be found in up to 90% of the general adult population and only about 1 in 500 will develop MS, meaning that other critical factors must influence the development of the disease[27]. In a recent US Army cohort-based study, 801 MS cases and 1566 controls were identified with samples available to assess EBV infection status. At baseline, 35 MS cases and 107 controls were negative for EBV. All but one of these 35 EBV-negative MS cases became infected with EBV during follow-up, and all seroconverted before MS onset. The seroconversion rate among people who developed MS during follow-up (97%) was definitely higher than among people who did not develop MS (57%). The hazard ratio for MS comparing EBV seroconversion versus persistent EBV seronegativity was 32.4 (95% CI 4.3 to 245.3, $p>0.001$)[28]. Another hypothesis proposes that infections during early life could attenuate the immune response to autoantigens and confer some protection against autoimmunity later in life[29-30].

- Gut microbiota. Gut microbiota contributes to preventing colonization of pathogenic microbes, modulating host immune responses, and promoting intestinal barrier integrity[31]. Whenever there is inflammation, it may damage the tight junctions and mucous membranes of the intestine and gastrointestinal mucosa. Toxic by-products in the gastrointestinal tract can be absorbed into the bloodstream and affect many body organs, causing food allergies, inflammation, and autoimmune disorders[32-34]. So, it has been suggested that altered gut microbiota composition might partially be the underlying cause of MS development[35]. Structural changes observed in the gastrointestinal mucosa and an increase in inflammatory T cells in MS patients, may partly explain how the impaired intestinal barrier can lead to the entry of immune cells and inflammatory factors into the brain. Several studies conducted on the role of the intestinal barrier in MS, have demonstrated that probiotics may provide significant

protection by reducing the reactivation of the immune system and inflammatory responses in MS patients[36-37].

- Other potential environmental risk factors. Several studies have shown an association between smoking and an increased risk of CIS and MS[38-39]. The month of birth has been implicated as a risk factor for MS, with a higher probability of developing MS in people born in April and May[40], but it is possible that this effect of birth is a false positive due to confounding factors such as seasonal variation in birth rates, due to excess births in March, April and May[41]. More recent studies suggest that specific diets may act as trigger for MS[42]. Two studies, carried out in vitro and animal models, showed that increased salt concentrations induced a profile of IL-17-producing Th17 cells and led to a worsening of symptoms in mice with experimental autoimmune encephalomyelitis[43-44]. Finally, a higher body mass index during childhood and adolescence has been associated with an increased risk of developing MS, especially in women[45-47]. Other studies suggest that interactions between obesity and HLA-DRB1* 15, as well as past infectious mononucleosis, could trigger the autoimmune response[48-49].

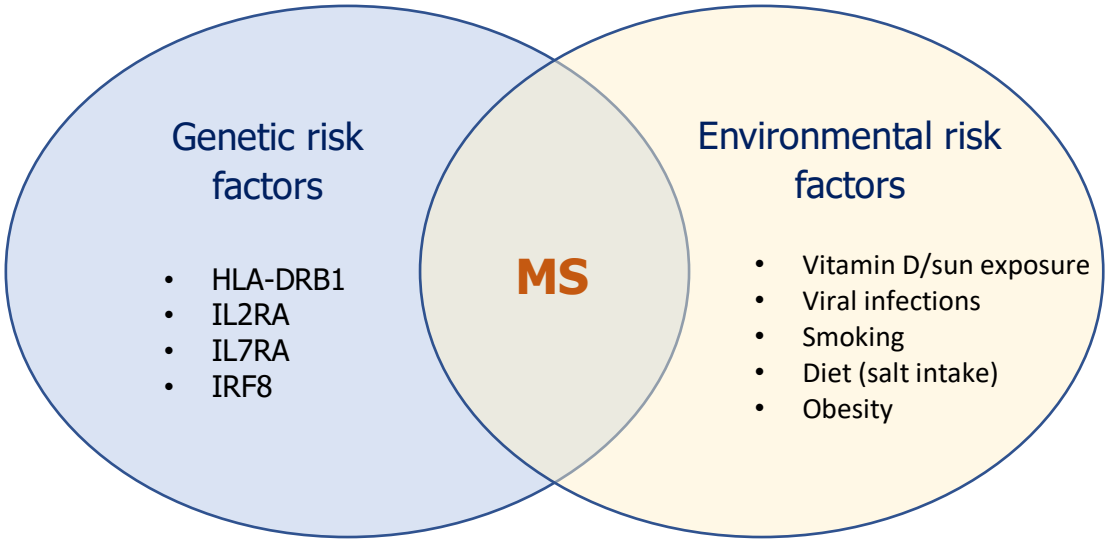


Figure 1.1.2A Main genetic and environmental risk factors for multiple sclerosis (MS).

1.1.3 Pathogeny

The most widely accepted theory is that MS begins as an inflammatory autoimmune disease mediated by autoreactive lymphocytes that cross the blood-brain barrier (BBB) and is dominated by microglial activation and neurodegeneration later in the course of the disease[50-51]. This theory arises from the similarities observed between MS and experimental autoimmune encephalomyelitis (EAE), the animal model of the disease, which is induced by immunizing animals with proteins derived from myelin such as proteolipid protein (PLP), the Myelin oligodendrocyte glycoprotein (MOG) or myelin basic protein (MBP)[52].

Myelin-specific autoreactive T lymphocytes can be found in peripheral blood and in the cerebrospinal fluid (CSF) of both MS patients and healthy controls, although in the former case, myelin-reactive T lymphocytes are activated more frequently and have a memory phenotype compared to the resting, naive phenotype observed in controls[53-54]. It is not yet clear how these T cells are activated in MS, but various processes have been suggested as triggers in genetically susceptible individuals exposed to the correct environmental factors, such as molecular mimicry (cross-reaction with auto-myelin epitopes of T cells generated against non-self-epitopes)[55-56], peripheral constitutive myelin antigens in cervical lymph nodes according to EAE models[57-58], or a deficient immunoregulatory control in which regulatory T lymphocytes fail to suppress effector cells, allowing them to initiate an immune response within the brain[59].

Transmigration through the BBB is mediated by adhesion molecules expressed in endothelial and immune cells, chemokines, and matrix metalloproteinase (MMP)[60]. Once within the CNS, autoreactive CD4 lymphocytes are reactivated by myelin antigens presented, in the context of HLA class II molecules, by antigen-presenting cells such as macrophages and microglia. This reactivation triggers the release of pro-inflammatory cytokines that further disrupt BBB and stimulate chemotaxis, resulting in a larger wave of inflammatory cell recruitment into the CNS[52,61].

After peripheral activation, IL-12 induces the differentiation of naive CD4 + T cells into Th1 cells, producing pro-inflammatory cytokines such as interferon (IFN)γ that activate macrophages[62]. On the other hand, differentiation into Th2 cells induces an anti-inflammatory response through cytokines such as IL-4[63]. A dysregulation in the balance between Th1 and Th2 cytokines has been classically implicated in the core of MS immunopathogenesis, but over the years, other cells subtypes have been discovered that also play a role in autoimmune diseases. Among them, Th17 cells undergo expansion after exposure to IL-23 produced by macrophages and dendritic cells. This cell lineage secretes IL-17, a pro-inflammatory cytokine[62-64], as well as IL-16 and TNF-α[65-66]. IL-17 further promotes the differentiation and activation of B cells and macrophages[67-68] and allows the penetration of most Th17 cells into the brain[69-70]. However, the antigenic specificity of Th17 cells has not been resolved[50].

So, whilst Th17 and Th1 cells promote inflammation, naturally occurring regulatory T cells (CD4 + CD25 + Treg) and Th2 cells have a role in resolving the inflammatory process by inhibiting T cell proliferation through down regulation of MHC-II expression and reducing the release of inflammatory cytokines[71]. Although the number of Treg cells in both blood and CSF appears to be similar between MS patients and healthy controls, there may be defects in the ability of Treg cells to suppress peripheral activation of myelin-specific T cells in MS patients[59].

Classically, MS was defined as a CD4+ cell-mediated disease. There is evidence that other cell lines also contribute to the disease immunopathogenesis. Regarding CD8+ cells, they are prominently found in MS lesions and, in some studies, CD8+ cells outnumber CD4+ cells[72]. They may promote CNS vascular permeability and the number of infiltrating CD8+ cells in MS lesions correlate with axonal damage[73-74]. Additionally, adoptive transfer of activated myelin-specific CD8+ cells can induce EAE, suggesting a role as effector cells in MS pathogenesis[75]. As for B cells, they can act as antigen presenting cells and, once they mature into plasmatic cells, they can produce auto-antibodies that bind to their target cells and activate the complement cascade or induce macrophage

antibody-mediated phagocytosis. B cells can also produce demyelination through secretion of antibodies against oligodendrocytes that can be mediated or not by complement[62,76].

The role of humoral immunity in MS immunopathogenesis is further supported by the persistent intrathecal production of oligoclonal immunoglobulins, a finding that, although not specific for MS, has been part of the diagnostic criteria in one way or another[77-80]. Another finding supporting the role of the humoral immune response is the presence of follicle-like aggregates with germinal centres in the meninges of patients with secondary progressive MS (SPMS), meaning that B-cell responses like proliferation, antigen-driven maturation selection, and differentiation into plasma cells can be maintained locally inside the CNS and may contribute to the pathogenic process[81]. Finally, B cell ablative therapy with an anti-CD20 monoclonal antibody that depletes naïve and memory B cells, reduces clinical relapses and inflammatory activity on brain MRI, probably by decreasing antigen presenting capacity and cytokine production by B cells[82].

Once T and B cells, plasma cells, and macrophages accumulate in the CNS, proinflammatory cytokines increase the immune response by activating microglia. Contact is then established between the activated microglia and components of the oligodendrocyte-myelin unit, which is opsonised with ligands for both microglial Fc and complement receptors, delivering a lethal signal through the surface-bound TNF-α resulting in demyelination and oligodendrocyte loss [83-85]. Additionally, it has also been suggested that astrocytes could play a more active role in facilitating a glutamate-mediated axonal degeneration and oligodendrocyte damage[85-86] (Fig. 1.1.3A).

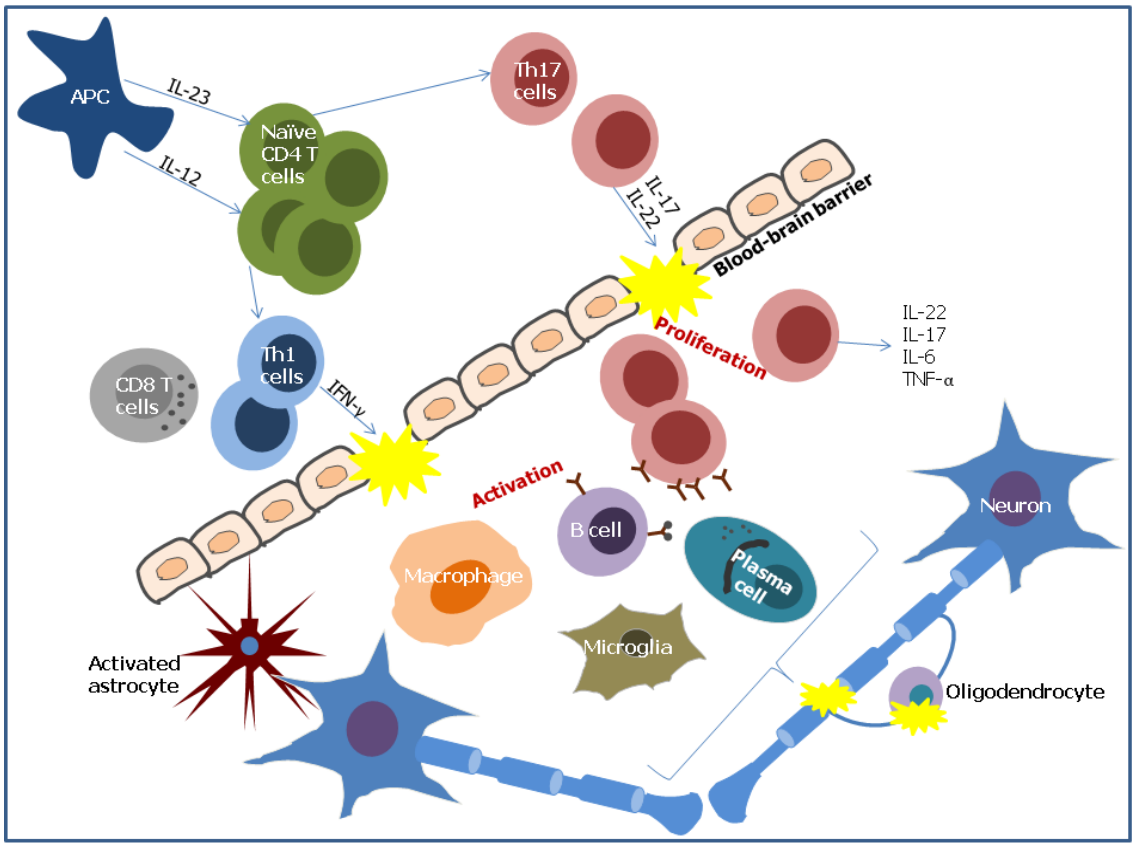


Figure 1.1.3A. Schematic representation of MS immunopathogenesis. APC: antigen presenting cell.

1.1.4 MS clinical manifestations and phenotypes

Clinically MS is characterized by the appearance of episodes of neurological dysfunction, called relapses, with total or partial recovery, which usually last days or weeks. Typical syndromes at onset include monocular visual loss secondary to inflammation of the optic nerve (optic neuritis), limb weakness or sensory loss secondary to spinal inflammation (myelitis), double vision due to brainstem dysfunction, or ataxia due to cerebellar involvement[1,87-88] (Table 1.1.4A). After typically 10 to 20 years from disease onset, a progressive clinical course develops in many of the persons affected, eventually leading to impaired mobility and cognition; approximately 15% of patients have a progressive course from onset[1].

Table 1.1.4A. Typical and atypical clinical presentation of MS

Clinical manifestation	Typical/atypical manifestation	Onset	Involvement	Signs or symptoms	Recovery
Optic neuritis	Typical	Subacute to chronic (hours to days)	Unilateral	Afferent pupillary defect Central visual blurring or scotoma Reduced visual acuity Dyschromatopsia Normal optic disc or optic disc swelling Mild unilateral orbital pain worsened by eye movements	Gradual recovery (2-4 weeks after reaching peak severity)
	Atypical	Acute (seconds to minutes)	Bilateral	Peripheral or altitudinal visual loss Retinal haemorrhages or exudates Severe optic disc swelling No light perception No or severe orbital pain Photophobia	Progressive worsening or no recovery
Brainstem and/or cerebellar syndromes	Typical	Subacute to chronic (hours to days)	Unilateral and localized	Internuclear ophthalmoplegia Multidirectional nystagmus Sixth cranial nerve palsy Ataxia or gait imbalance Vertigo Facial numbness or sensory loss Dysmetria Dysarthria and slurred speech Dysphagia Hearing loss Nausea	Gradual recovery (2-4 weeks after reaching peak severity)
	Atypical		Alternating syndromes	Vascular territories signs	

		Acute (seconds to minutes)		Isolated trigeminal neuralgia Fluctuating ocular or bulbar weakness Fever Meningism	Progressive worsening or no recovery
Myelitis	Typical	Subacute to chronic (hours to days)	Incomplete transverse myelitis Asymmetric involvement	Sensory involvement: paresthesias, Lhermitte sign, impairment of vibration and joint position sense, decreased pain and light touch perception and Uhthoff phenomenon Motor deficits: pyramidal signs, spastic paresis and/or weakness and spasticity Sphincter dysfunction: urinary urgency, hesitancy, urge incontinence, constipation, and faecal incontinence Sexual dysfunction: erectile dysfunction and impotence	Gradual recovery (2-4 weeks after reaching peak severity)
	Atypical	Acute (seconds to minutes)	Complete transverse myelitis Complete Brown-Séquard syndrome Cauda equina syndrome Anterior spinal artery territory lesion Localized or radicular spinal pain	Progressive and symmetrical spastic parapresis Progressive sensory ataxia Sharp level to all sensory modalities Segmental loss of pain and temperature sensation Areflexia and/or spinal shock Acute urinary retention Severe pain	Progressive worsening or no recovery
Cerebral hemispheric syndromes	Typical	Subacute to chronic (hours to days)	Unilateral	Hemisyndrome: hemiparesis and hemisensory deficits Campimetric deficits	Gradual recovery (2-4 weeks after reaching peak severity)
	Atypical	Acute (seconds to minutes)	Bilateral	Encephalopathy Epilepsy Cortical blindness Intracranial hypertension	Progressive worsening or no recovery

MS: multiple sclerosis. Modified from Filippi M, et al[88]

Regarding MS phenotypes, in 1996, in the absence of agreed on biological markers, the Advisory Committee on Clinical Trials of New Agents in MS of the National Multiple Sclerosis Society (NMSS) (USA) reached consensus on definitions and the appropriate terminology to describe clinical outcomes and course patterns in patients with MS. Among several possible clinical patterns commonly used "relapsing-remitting (RR), relapsing-progressive (RP), primary progressive (PP), secondary progressive (SP), benign, and malignant MS", the Committee ultimately decided to retain and redefine the terms RR, PP and SP, and that the term RP MS be dropped, as the term was believed to be vague and overlapped with other disease course subtypes. Besides, the term progressive relapsing (PR) MS emerged as an additional, albeit rare, to reflect its progressive onset and to distinguish it from the term RP. In regard to clinical severity definitions (benign and malignant MS), it was agreed that they should not be the sole determinant of the appropriateness of any available therapeutic measures. It was additionally emphasized that these terms were most useful in the context of research studies and should be used with care in communication with affected individuals, family members, and third-party payers[89].

In October 2012, the Committee met again to review the 1996 MS clinical subtypes descriptions[90]. In this meeting it was concluded that the descriptions of the central MS phenotypes of relapsing and progressive disease should be preserved with some modifications. In this sense, the phenotype PR was removed. The two remaining phenotypes of progressive MS (PP and SP) should be subclassified into progressive disease "with progression" in the event that patients experience clinical evidence of increased and confirmed disability worsening independent of relapses over a defined time interval. Furthermore, in case of associating symptomatic relapses and/or asymptomatic radiological activity (occurrence of contrast enhancing T1 or new or unequivocally enlarging T2 hyperintense lesions), progressive MS should be subclassified as "active" or "non-active" in their absence.

The term CIS was included in the spectrum of MS phenotypes. As for the progressive phenotypes, both for the CIS and RR phenotypes, the "active or non-

active" modifiers should also be applied depending on whether they associate or not the presence of clinical (relapses) and/or subclinical MRI activity (Fig. 1.1.4A and 1.1.4B).

It was also noted that the terms "benign" and "malignant" disease are often misused and should be used with caution.

A clarification of the clinical course descriptors for MS defined in 2013 was recently published, recommending annually evaluations as a minimum time interval for monitoring disease activity and progression. Furthermore, a clarification of the terms *worsening* and *progression* was established. In this sense, it is recommended the term worsening be used to describe any increase in impairment/disability irrespective of whether it has resulted from relapses or increasing disability during the progressive phase of the disease and reserving the term *disease progression* for progressive MS types (PPMS or SPMS) with accrual of disability, regardless of any relapse activity[91].

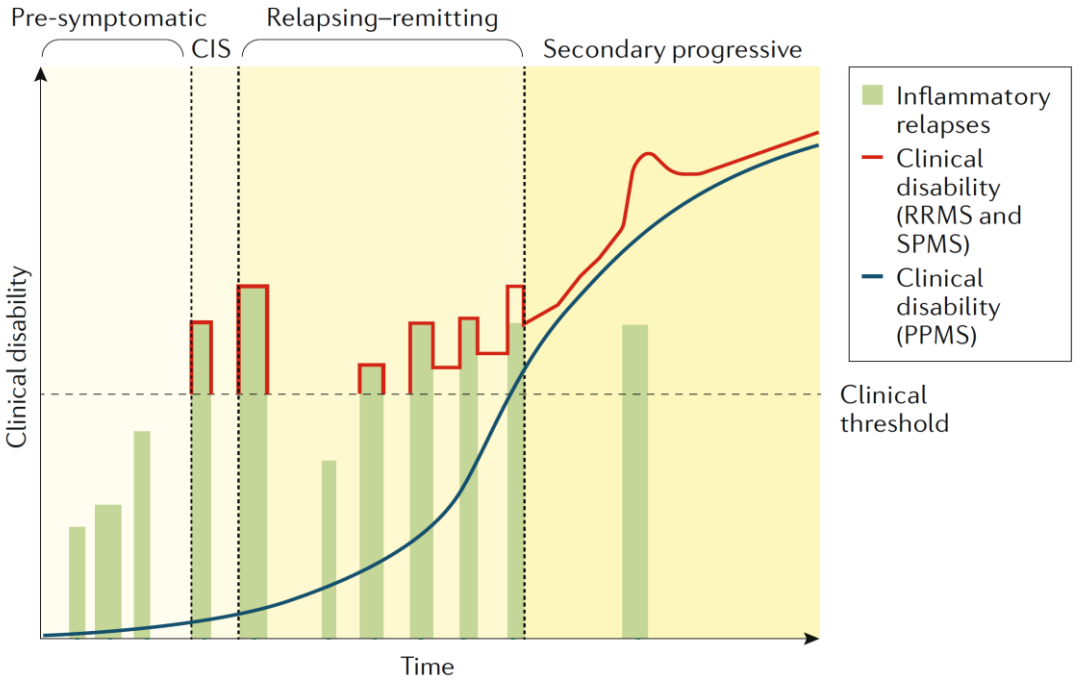


Figure 1.1.4A. Clinical course of MS. The National Multiple Sclerosis Society Advisory Committee on Clinical Trials in multiple sclerosis (MS) defined the following clinical courses of MS: relapsing–remitting MS (RRMS), secondary progressive MS (SPMS), and primary progressive MS (PPMS).

RR MS accounts for ~85% of patients and is characterized by the occurrence of relapses at irregular intervals with complete or incomplete neurological recovery; the average relapse frequency is ~1.1 per year early in the disease course but seems to decrease with advancing disease, increasing neurological dysfunction and age. Most patients with RRMS will develop SP MS, which is characterized by progressive, irreversible disability that occurs independently of the presence of relapses. Conversion to SPMS occurs in ~2–3% of patients per year. Approximately 10–15% of patients present with PPMS, which is characterized by disease progression from the onset, resulting in gradual, progressive, and permanent neurological deficits for >1 year with or without relapses. The term clinically isolated syndrome (CIS) was also included to denote those patients whose first clinical presentation has characteristics of inflammatory demyelination that could be MS but who do not fulfil its diagnostic criteria. Modified from Filippi M, et al[88]

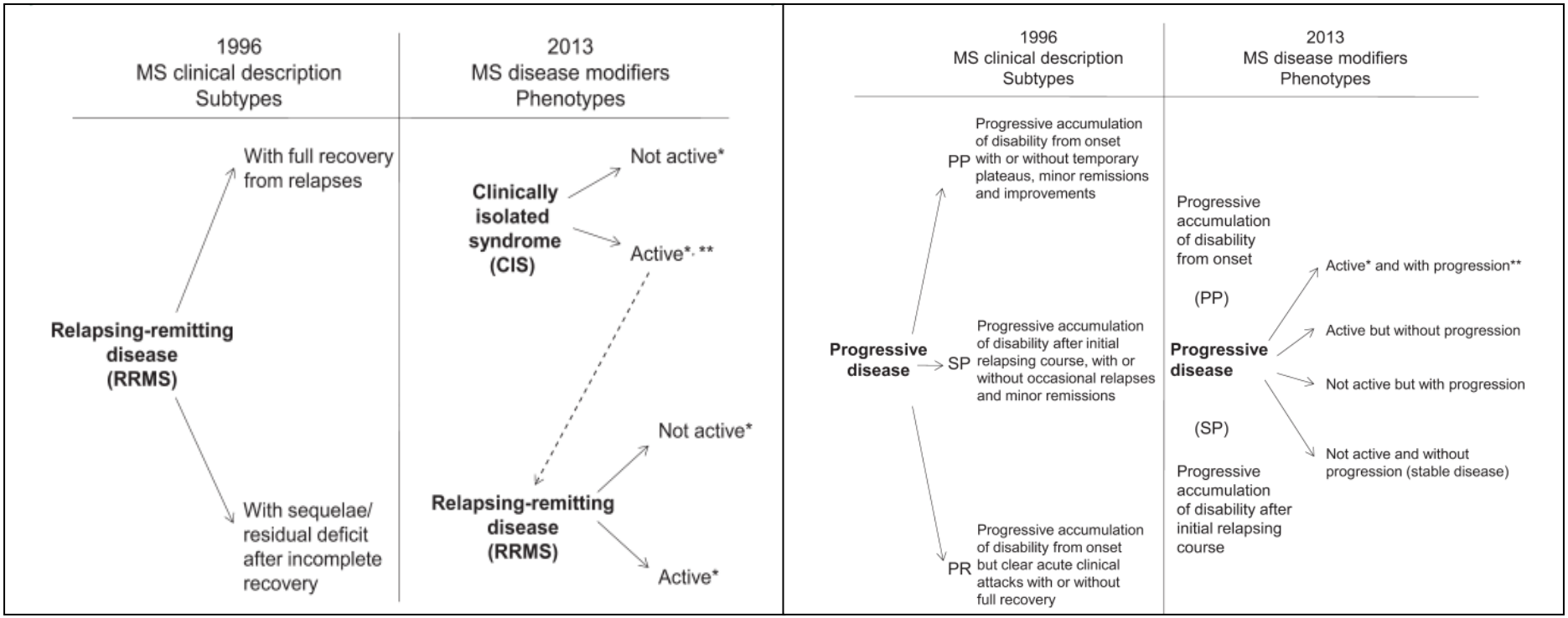


Figure I-4b. The 1996 vs 2013 multiple sclerosis phenotype descriptions for relapsing and progressive disease. *Activity determined by clinical relapses and/or MRI activity (contrast-enhancing lesions; new or unequivocally enlarging T2 lesions, assessed at least annually.*,**CIS, if subsequently clinically active and fulfilling current multiple sclerosis (MS) diagnostic criteria, becomes relapsing-remitting MS (RRMS). **Progression measured by clinical evaluation, assessed at least annually. PP= primary progressive; PR= progressive relapsing; SP= secondary progressive. Modified from Lublin FD, et al[90]

With the widespread use of MRI, a new term "Radiologically Isolated Syndrome" (RIS) emerged due to the identification of lesions suggestive of MS in asymptomatic individuals[92-93]. Not until 2009 was the expression RIS introduced by Okuda et al (Table 1.1.4B), and formally defined for the first time[94]. Several studies suggest that up to 34% of patients with RIS may develop a clinical attack in the next 5 years. Factors such as male sex, young age, and the presence of spinal cord lesions increase the risk of having a first clinical event[95]. However, RIS has not been considered as a MS subtype per se, since clinical symptoms suggestive of demyelinating disease are lacking and MRI findings alone may be non-specific[96].

Table 1.1.4B. Definition of RIS according to Okuda and colleagues.

A. The presence of incidentally identified CNS white matter anomalies meeting the following MRI criteria:

1. Ovoid, well-circumscribed, and homogeneous foci observed with or without involvement of the corpus callosum.
2. T2 hyperintensities measuring ≥ 3 mm and fulfilling Barkhof criteria (at least three out of four) for dissemination in space.
3. Anomalies not following a clear vascular pattern.
4. Structural neuroimaging abnormalities identified not explained by another disease process B. No historical accounts of remitting clinical symptoms consistent with neurological dysfunction

C. The MRI anomalies do not account for clinically apparent impairments in social, occupational, or generalized area of functioning.

D. The MRI anomalies are not due to the direct physiological effects of substances (recreational drug use, toxic exposure) or a medical condition.

E. Exclusion of individuals with MRI phenotypes suggestive of leukoaraiosis or extensive white matter changes lacking clear involvement of the corpus callosum.

F. The CNS MRI anomalies are not better accounted for by another disease process

In recent years, the published MAGNIMS consensus guidelines on MRI criteria for the diagnosis of MS propose that the same MRI criteria used to establish

dissemination in time (DIT) and dissemination in space (DIS) in MS (see below), should also be applied to RIS[97].

1.1.5 Diagnosis

Diagnostic criteria for multiple sclerosis, based primarily on clinical, imaging, and laboratory evidence, have evolved over time.

- **MRI**

To this day, MRI is the test of choice to support the clinical diagnosis of MS[91]. Clinical findings are also considered to determine fulfilment of DIS and DIT[77].

In the context of a first demyelinating event suggestive of MS, MRI diagnostic criteria can be applied[98-100]. With the criteria evolution and emergence of the McDonald criteria, MRI was incorporated as a robust tool to demonstrate DIS and DIT of CNS lesions.

Apart from its use in diagnosis, MRI has also gained a fundamental role in the monitoring of the response to treatments (use in the follow-up of inflammation and / or neurodegeneration) as well as in the early recognition of adverse events associated with treatment (for example, progressive multifocal leukoencephalopathy (PML) and other opportunistic infections)[101].

DIS. The original McDonald criteria as well as their revision in 2005 established the concept of DIS on MRI by using the Barkhof-Tintoré criteria (Table 1.1.5A). At least three of the four criteria must be present to fulfil DIS[102-103]. The clinical concept of DIS was maintained in case of a multifocal attack or according to findings on the neurological examination (ie, optic neuritis plus extensor plantar response). In cases in which MRI findings are not characteristic and do not fulfil the Barkhof-Tintore criteria, authors established an alternative criterion based on presence of at least 2 lesions on T2-weighted images on MRI plus oligoclonal bands (OCB) or increase of IgG in CSF[77-78]. Considerations regarding spinal cord MRI in the diagnosis were made: one spinal cord lesion could be equivalent to an infratentorial lesion to fulfil the Barkhof criteria. The McDonald criteria were

revised again in 2010 to incorporate new evidence and to simplify their application whilst preserving sensitivity and specificity[79]. The criteria proposed by the MAGNIMS group[97] were incorporated in the 2010 revised version of the McDonald criteria[83]. In this proposal, DIS was demonstrated by presence of one or more T2 lesions in at least two of the four typical MS anatomical topographies: periventricular, juxtacortical, infratentorial or spinal cord or by the development of a further clinical attack implicating a different nervous system site. Symptomatic lesions of the brainstem or spinal cord were excluded and did not contribute to lesion count.

Table 1.1.5A. DIS on MRI according to the Barkhof-Tintoré and MAGNIMS criteria

DIS criteria on MRI	
Barkhof-Tintoré	**MAGNIMS**
At least three of the following: 1. One Gd-enhancing lesion or nine T2 hyperintense lesions if there is no Gd-enhancing lesion 2. At least three periventricular lesions 3. At least one juxtacortical lesion 4. At least one infratentorial lesion	At least one T2 lesion in at least two of the following areas: • Periventricular • Juxtacortical • Infratentorial • Spinal cord
Note: In 2001, one spinal cord lesion can be substituted for one brain lesion. In 2005, individual spinal cord lesions can be included in the total lesion count.	Note: In the case of brainstem or spinal cord syndromes, the symptomatic lesions are excluded and do not contribute to lesion count.

The International Panel on Diagnosis of Multiple Sclerosis met for the last time in 2016. The most recent revision of McDonald criteria includes (Table 1.1.5B), based on histopathological studies and in addition to juxtacortical lesions, cortical

lesions as new topography to fulfil MRI criteria for DIS, although it recognized that standard MRI currently has limited ability to detect cortical lesions or distinguish cortical lesions in MS from those with other causes. In addition, care is needed to distinguish potential cortical lesions from neuroimaging artefacts[80]. Additionally, recent studies have shown that inclusion of symptomatic lesions in the MRI determination of DIS or DIT increases diagnostic sensitivity with little or no reduction in specificity[104-105]. Based on these data, the Panel also recommended including symptomatic and asymptomatic MRI lesions in the determination of DIS and DIT in the last McDonald criteria revision. An exception relates to lesions in the optic nerve in a patient presenting with optic neuritis, as the Panel felt evidence was insufficient to support inclusion of the optic nerve as a site to determine DIS in these patients[80].

DIT. In the 2001 and 2005 McDonald criteria, if a new MRI was performed three months after the CIS and a Gd-enhancing lesion in an asymptomatic region was detected, DIT was fulfilled. When analysing T2-weighted images, detection of at least one new lesion, if the baseline scan had been performed at least 30 days after symptom onset, could also be used to demonstrate DIT. Hence, MRI now allowed an earlier diagnosis. As with DIS, later studies proposed some alternatives to establish DIT[98-100]. Since these studies demonstrated that specificity was preserved, the 2010 criteria also incorporated a simplified version of DIT. Simultaneous presence of asymptomatic Gd-enhancing and non-enhancing lesions at any time, or a new T2 or Gd-enhancing lesion on follow-up MRI, irrespective of its timing with reference to a baseline scan, or a second clinical attack fulfilled DIT. Thus, the main advantage of the 2010 criteria was the possibility of establishing DIS and DIT at the time of the CIS[79]. As mentioned above, the latest McDonald revision (2017), further simplified the current radiological criteria by allowing the inclusion of symptomatic Gd-enhancing lesions in the determination of DIT. In turn, when in a first MRI it is not possible to fulfill DIT, but the patient presents with a typical CIS and fulfilment of clinical or MRI criteria for DIS are demonstrated and there is no better explanation for the clinical presentation, the demonstration of CSF-specific oligoclonal bands in

the absence of other CSF findings atypical of MS allows a diagnosis of this disease to be made (Table 1.1.5C)[90].

Table 1.1.5B. 2017 McDonald criteria for demonstration of dissemination in space and time by MRI in a patient with a clinically isolated síndrome
• **DIS** can be demonstrated by one or more T2-hyperintense lesions* that are characteristic of MS in two or more of four areas of the CNS: periventricular, cortical or juxtacortical, and infratentorial brain regions, and the spinal cord
• **DIT** can be demonstrated by the simultaneous presence of gadolinium-enhancing and non-enhancing lesions* at any time or by a new T2-hyperintense or gadolinium-enhancing lesion on follow-up MRI, with reference to a baseline scan, irrespective of the timing of the baseline MRI

MS: multiple sclerosis, DIS: dissemination in space, DIT: dissemination in time, CNS: central nervous system, MRI: magnetic resonance imaging.
*Unlike the 2010 McDonald criteria, no distinction between symptomatic and asymptomatic MRI lesions is required.

Table 1.1.5C. The 2017 McDonald criteria for diagnosis of multiple sclerosis in patients with an attack at onset

	Number of lesions with objective clinical evidence	**Additional data needed for a diagnosis of MS**
≥ 2 clinical attacks	≥ 2	None*
≥ 2 clinical attacks	1 (as well as clear-cut historical evidence of a previous attack involving a lesion in a distinct anatomical location**)	None*
≥ 2 clinical attacks	1	DIS demonstrated by an additional clinical attack implicating a different CNS site or by MRI+
1 clinical attack	≥ 2	DIT demonstrated by an additional clinical attack or by MRI¨ OR demonstration of CSF-specific oligoclonal bands§

1 clinical attack	1	DIS demonstrated by an additional clinical attack implicating a different CNS site or by MRI+ AND DIT demonstrated by an additional clinical attack or by MRI¨ OR demonstration of CSF-specific oligoclonal bands§

MS: multiple sclerosis, DIS: dissemination in space, DIT: dissemination in time, CNS: central nervous system, MRI: magnetic resonance imaging, CSF: cerebrospinal fluid.
*No additional tests are required to demonstrate DIS and DIT. However, unless MRI is not possible, brain MRI should be obtained in all patients in whom the diagnosis of MS is being considered. In addition, spinal cord MRI or CSF examination should be considered in patients with insufficient clinical and MRI evidence supporting MS, with a presentation other than a typical clinically isolated syndrome, or with atypical features. If imaging or other tests (eg, CSF) are undertaken and are negative, caution needs to be taken before making a diagnosis of MS, and alternative diagnoses should be considered. ** Clinical diagnosis based on objective clinical findings for two attacks is most secure. Reasonable historical evidence for one past attack, in the absence of documented objective neurological findings, can include historical events with symptoms and evolution characteristic for a previous inflammatory demyelinating attack; at least one attack, however, must be supported by objective findings. In the absence of residual objective evidence, caution is needed. +The MRI criteria for DIS are described in panel 5. ¨The MRI criteria for DIT are described in panel 5. §The presence of CSF-specific oligoclonal bands does not demonstrate DIT per se but can substitute for the requirement for demonstration of this measure.

Currently, brain and spinal cord volume measures have no role in the MS diagnostic criteria[77] or disease course classification[90], but a body of evidence that these measures are valuable for early evaluation of treatment responses and prediction of disease evolution has been steadily growing alongside improvements in methodology that could facilitate widespread implementation of these measures in clinical practice[106-107].

- **CSF**

General findings. CSF appearance and pressure are usually normal. Total leukocyte count is normal in approximately 60.0% of patients, exceeds 15 cells/µl in less than 5.0% of cases, and rarely exceeds 50 cells/µl (a finding that should be considered a red flag)[108]. Lymphocytes are the predominant cell type. CSF protein levels are usually normal.

IgG OCB. Elevation of CSF immunoglobulin levels relative to other proteins suggests intrathecal synthesis. Such increase is predominantly IgG, although IgM and IgA synthesis can be increased as well. Positive CSF is based on the finding of either OCB different from any such bands in serum or by an increased IgG index. IgG level is usually expressed by use of the IgG index (normal value <0.66-0.90, depending on the laboratory) or by use of formulae for intrathecal fluid synthesis of IgG. Such abnormality can be found in up to 90.0% of patients with MS[109-111]. OCB represent limited classes of antibodies that are depicted as discrete bands on agarose gel. Around 8.0% of CSF samples of patients without MS also contain OCB, and most are the result of chronic CNS infections, viral infections, and neuropathies. Of the five OCB patterns, types 2 (OCB in CSF but not serum) and 3 (OCB in CSF plus additional identical OCB in CSF and serum) indicate intrathecal IgG synthesis. Expert recommendations on evaluation of CSF in patients suspected of having MS indicate that the most informative analysis is qualitative assessment of CSF for OCB, best performed using isoelectric focusing on agarose gel followed by immunodetection by blotting or fixation, as this method achieves the best sensitivity and specificity[112].

- **Multimodal evoked potentials**

Multimodal evoked potentials are the electrical events generated in the CNS by peripheral stimulation. They are used to detect abnormal CNS function that may be clinically undetectable. Detection of such subclinical lesion in a site remote from the region of clinical dysfunction supports the diagnosis of MS and may also help define the anatomical site of the lesion in tracts not easily visualized by imaging (optic nerves, dorsal columns). The most commonly used are somatosensory, visual, and brainstem auditory evoked potentials. Patients with MS have abnormal visual evoked potentials in 85.0% of cases, but ocular or retinal disorders should be excluded. Somatosensory evoked potentials are abnormal in 77.0% of cases, including approximately 50.0% of those who do not have sensory symptoms or signs. Brainstem auditory evoked potentials abnormalities are present in 67.0% of cases. Thus, visual evoked potentials are probably useful, whereas somatosensory evoked potentials are possibly useful,

and there is insufficient evidence to recommend brainstem auditory evoked potentials as a test for diagnostic purposes[113].

1.1.6 Differential diagnosis / MS misdiagnosis

MS has a wide spectrum of clinical manifestations, and the proper interpretation of the radiological findings may pose an actual challenge at the time of diagnosis[114-116]. Neither highly specific nor sensitive biomarker for MS has been identified to date. Different disorders and conditions can mimic the radiological and/or clinical appearance of MS[117]. Non-MS inflammatory demyelinating diseases such as Neuromyelitis Optica Spectrum Disorders (NMOSD) and disorders associated with anti–myelin oligodendrocyte glycoprotein (antiMOG) antibodies and incidental and migraine-related white matter changes are the most common to mimic MS[118]. Other possible sources of misdiagnosis also include infectious, metabolic and vascular diseases[119]. Despite the current well-validated diagnostic criteria for MS[120], the occurrence of misdiagnosis remains a challenge with significant implications for patients, their families and the healthcare system[121]. Current evidence supports the benefits of an early diagnosis and disease-modifying drug (DMD) onset seeking to reduce long-term disability[122]. In given cases, this may lead neurologists to rush into a diagnosis of MS, hence constituting a factor playing a role in mistaken diagnosis[123]. The latest revisions of the McDonald criteria have a strong focus on the role of MRI, increasing its sensitivity though with low impact on specificity for MS diagnosis, making early diagnosis more likely[124]. However, the degree of emphasis allocated to the role of MRI during the diagnostic process can also contribute to misdiagnosis by wrong interpretation of non-specific or non-inflammatory demyelinating lesions. The same is true for the application of the radiological criteria in patients with atypical clinical presentations and also in asymptomatic patients[125]. For this reason, all criteria sets published to date, from Schumacher's to McDonald's, emphasized on the need to rule out any 'other better explanation' for the clinical scenario before making a definitive diagnosis of MS[124-126]. Several studies have described that, of all new patients with a suspicion of or established MS diagnosis

referred to centres specializing in demyelinating disorders, 30%–67% did not have MS eventually[127]. Unfortunately, around 50% of such patients had initiated DMD at the time of determining the misdiagnosis, having exposed to possible secondary emotional and physical damage. Misdiagnosis is also associated with the clinical challenge of having to approach undoing of a diagnosis of MS[128].

In a recent study performed at our Centre[129] the frequency of patients' diagnoses changing from 'established MS' to a final diagnosis of "non-MS" was lower (7.1%) than in previous studies (but only 31.6% of all patients were referred to our centre with a diagnosis of 'established MS'). Despite the low misdiagnosis rate, when we attempt to extrapolate that frequency into the total Spanish population with MS of about 46,000 individuals[1320] (100 per every 100,000 inhabitants), then the number of misdiagnosed patients is estimated to be beyond 3000. Multifocal white matter lesions deemed non-specific and not suggestive of MS were reported in all patients whose diagnosis changed from 'established MS' to "non-MS", pointing to the fact that MRI could be a strong contributor to MS misdiagnosis, especially when not considering clinical aspects or when the accompanying symptoms are atypical. In the cases that MRI findings are inconclusive, searching for spinal cord lesions could be helpful as they either do not occur or will have different characteristics in other diseases[131]. In addition, assessing the presence of the central vein sign can successfully differentiate MS from non-MS white matter lesions with high sensitivity and specificity in selected cases[132-133]. Obtaining a second opinion from an experienced neuroradiologist is also recommended. Finally, a 'wait and see' attitude within a short reassessment interval may be sometimes prudent in patients with atypical clinical presentations or atypical MRI findings, as new data may confirm or rule out MS diagnosis, thus avoiding moving into unnecessary and potentially harmful treatments[134].

1.1.7 Treatment

Currently more than a dozen DMDs are available in order to reduce the frequency of relapses and prevent the accumulation of focal lesions in the CNS (Fig. 1.1.7A)

On the basis of the ability of several of these medications to delay a formal diagnosis of MS after an initial attack, there has been a general move toward early treatment, although, the long-term value of this approach with respect to preventing progressive MS remains uncertain. No available drug has curative properties or reverses established neuronal damage[1].

According to the decrease in the annualized relapses rate in RRMS patients, DMDs have been classified as moderate-intermediate efficacy and high efficacy therapies. Interferons, glatiramer acetate and the newer orals drugs teriflunomide and dimethyl fumarate are usually considered as being of moderate efficacy. Fingolimod, other sphingosine-1-phosphate (S1P) receptor modulators and cladribine are usually considered as intermediate efficacy drugs. Finally, monoclonal antibodies such natalizumab, alemtuzumab and ocrelizumab, together with mitoxantrone (an antineoplastic agent) are usually considered high efficacy therapies[135].

In the absence of previously licensed treatments for PPMS patients, in 2017 ocrelizumab was the first drug approved as a promising step, but the reasons for ocrelizumab's ability to slow progression[136] remain uncertain. More recently, a phase III trial of siponimod versus placebo in SPMS met its primary objective of reducing the risk of disability progression[137]. In this way, albeit slowly, the therapeutic arsenal for progressive MS phenotypes is being enriched.

Another important trend over time has been to escalate treatment with a target of "no evidence of disease activity," as evidenced by the absence of new lesions, relapses, disability progression and, more recently, tissue atrophy[138-139]; however, it is doubtful that MS can be fully arrested with current therapies.

Several multicentres studies compare early intensive treatment with more conventional treatment escalation approaches. If an escalation strategy is chosen, strict monitoring of disease activity with clinical and radiological parameters is mandatory for deciding whether treatment needs to be changed and when best to make this change[140]. With induction therapy, monitoring of treatment response is also important, but the risks associated with most available

high-efficacy treatments must be considered. Patients should be informed of and adhere to strategies proposed by the regulatory agencies to mitigate the risk[141].

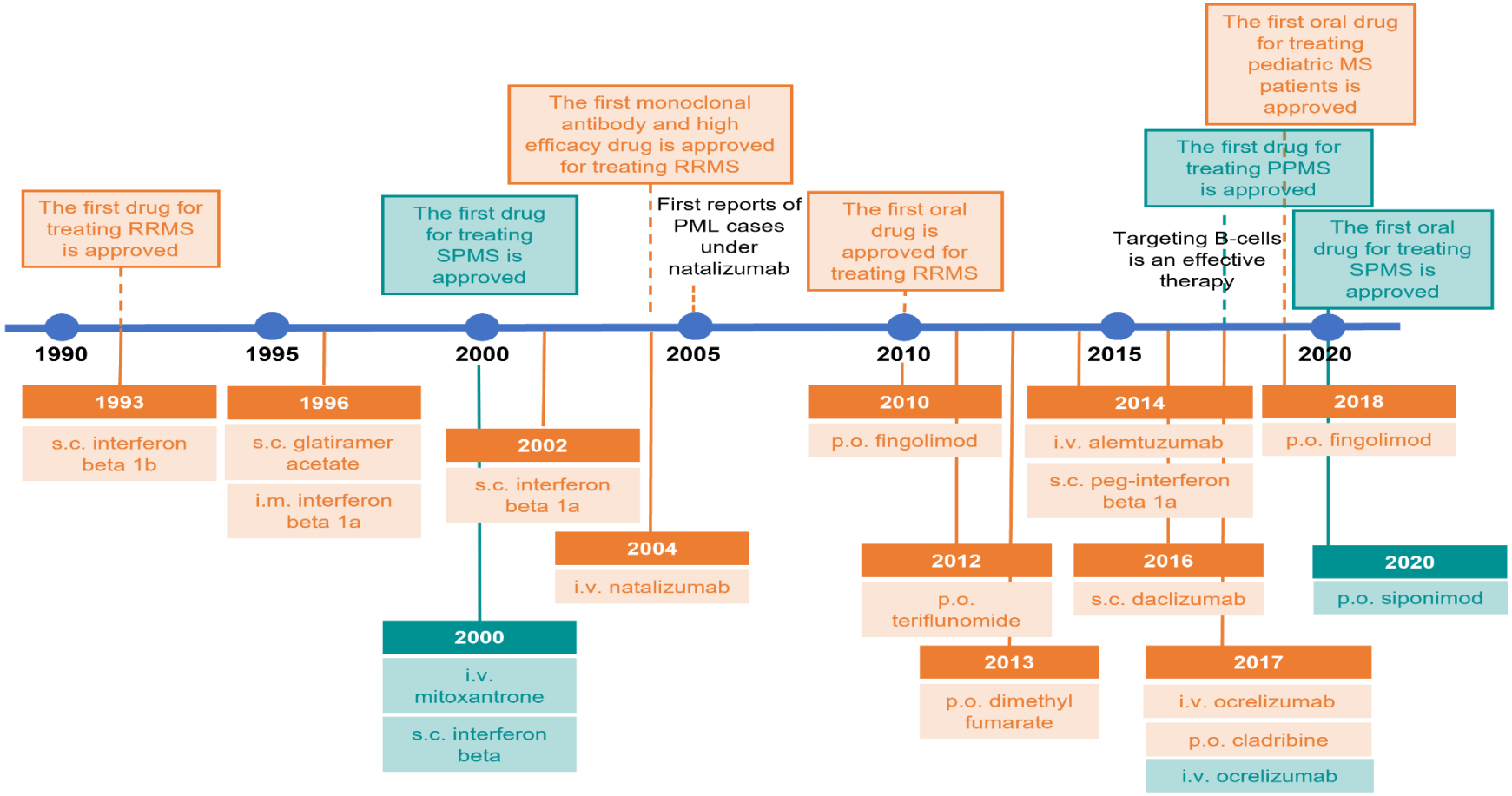


Figure 1.1.7A Timeline of developments in the treatment of multiple sclerosis. Important milestones in the development are shown in green boxes, and drugs approved by the FDA (or the European Medicines Agency for subcutaneous IFNβ for secondary progressive multiple sclerosis (SPMS) and oral cladribine) are shown in orange boxes. MS, multiple sclerosis; PEG, polyethylene glycol; PPMS, primary progressive MS; RRMS, relapsing–remitting multiple sclerosis; SPMS, secondary progressive MS. Modified from Tintoré et al.[142]

1.2 MS PATIENTS MONITORING

Since MS is a highly heterogeneous, multisymptomatic, chronic disease lasting several decades, it offers numerous inter-individual and intra-individual differences, as well as different disease phenotypes evident in different disease stages. Each of these individual differences and disease phenotypes must be addressed when it comes to treating MS, as well as MS-related symptoms (e.g. spasticity, pain, fatigue, cognitive impairment and gait problems). Furthermore, because MS and its symptoms can change over time, it is crucial to detect these changes early in their development. For this purpose, regular neurological examinations and MRI scans, among other potential evaluations, should be performed periodically[143].

1.2.1 Radiological MS monitoring

Radiological activity, in the clinical setting, is primarily defined by the presence of Gd-enhancing T1 or new/enlarging T2 lesions[144]. MRI, in general, cannot only assist in the diagnostic process but is also crucial in regular monitoring to provide information on effectiveness as well as safety of DMDs[101].

Based on the 2021 MAGNIMS–CMSC–NAIMS consensus[144], in patients with CIS in whom the initial brain and spinal cord MRI scans did not show DIS or DIT, according to the 2017 revisions of the McDonald criteria, serial clinical observation and follow-up MRI every 6–12 months are required to identify new disease activity over time. Serial brain MRI studies in individuals with CIS showed accrual of new brain T2-weighted lesions that confirmed DIT and diagnosis of MS in 51% of patients by 6 months and in 74% of patients by 12 months[145].

Before starting or switching a DMD a baseline brain MRI and a new baseline brain MRI, usually at 3–6 months after treatment initiation, should be obtained to avoid misinterpretation of lesions that developed before onset of the therapeutic effect. Then, a yearly brain MRI while the patient is on the DMD is recommended, considering longer intervals in clinically stable patients after the first few years of treatment[144].

Regarding the role of MRI on the safety monitoring of DMDs, it is mainly focused on ruling out or early diagnosing PML, an opportunistic infection with a relatively high incidence in patients treated with natalizumab and with much lower frequency with other MS therapies[146-148]. The imaging findings of patients with early PML and the clinical relevance of brain MRI screening to facilitate early PML diagnosis, leading to a more favourable outcome, have been shown in patients with MS who are treated with natalizumab[149]. But the spectrum of possible safety events is broad and not exclusively restricted to opportunistic infections, and also includes non-infectious CNS comorbidities, such as vascular or neoplastic processes, and atypical demyelinating lesions that potentially mimic MS[150-153].

At the moment there are no MRI sequences used in clinical practice capable of reflecting disease progression as a consequence of the neurodegenerative process of MS, as well as MRI markers for distinguishing PPMS from RRMS and predicting the evolution of RRMS to SPMS. Although non-specific, some spinal cord imaging features (diffuse abnormalities and lesions involving gray matter [GM] and ≥2 white matter columns) are typical of PPMS. In patients with PPMS and those with relapsing MS onset, location of lesions in critical CNS regions (spinal cord, infratentorial regions, and GM) and MRI-detected high inflammatory activity in the first years after diagnosis are risk factors for long-term disability and future progressive disease course[154]. Additionally, measurement of global brain volume and cervical cord area may help better understand the global burden of disease in all clinical MS phenotypes, even in the earliest stages of the condition. While GM volume changes in the brain are more pronounced and clinically relevant than white matter volume changes, their exact relevance in clinical practice is unclear. Although some studies have shown associations between treatment effects on measures of atrophy and disability, the results so far are inconclusive. Several potential sources of substantial error remain, including, but not limited to, differential effects of drugs on brain volume measures, confounding physiological and technical factors and the performance and value of volumetric tools. To make implementation of volume measurements in clinical practice feasible, these potential sources of error need to be accounted

for and appropriately managed, and further research is needed to ensure the accuracy and reliability of the measurements[155-156].

1.2.2 Clinical MS monitoring

1.2.2a What should we monitor?

Clinical MS monitoring mainly relies on the assessment of relapses and disability on a regular basis[157]. Relapses are considered to be the clinical expression of acute inflammatory demyelinating focal lesions disseminated in the CNS, whereas disability progression is considered to reflect the occurrence of axonal loss and neurodegeneration[1].

Relapses are defined as acute/subacute development of new symptoms or worsening of existing symptoms, lasting >24 h, followed by full or partial recovery, in the absence of fever or infection[1]. In relapsing MS patients, relapses may contribute to the accumulation of disability, primarily early in the disease, which is called relapse-associated worsening (RAW)[158]. Different studies demonstrated that higher early relapse frequencies within 2–5 years and shorter first inter-attack intervals herald more rapid deterioration via interaction with the neurodegeneration characterizing secondary progression. They increase the probability of its occurrence, shorten its latency and influence, to a lesser degree, its slope. However, neither the risk of entering the secondary progressive phase nor the latency to onset of progression appear to be related to the total number of attacks during the relapsing-remitting phase or to the rate of relapses experienced during the relapsing-remitting phase after the second year up of the disease. Furthermore, the predictive effect of early relapse rate on disease progression disappears once the progressive course supervenes[159-162].

For clinicians, in the absence of biomarkers and definitive clinical diagnostic criteria, the diagnosis of the conversion from RRMS to SPMS is still challenging and retrospective. It has been seen that up to two-thirds of patients with insidious worsening of disability are still considered by clinicians to have RR MS[163-165]. A

new outcome measure that recently emerged for relapsing MS patients to detect slowly increasing disability is "progression independent of relapse activity (PIRA)", defined as confirmed progression independent of relapses, which may start early, playing a major role as a contributor to the irreversible accumulation of disability throughout the entire course of the disease[158,166].

Dissociation between therapeutic effects on relapse frequency versus effects on progression of disability was shown in different clinical trials[167-170], suggesting that biological mechanisms leading to acute attacks probably differ from those responsible for unremitting disability[171-172].

Based on the above and given the diagnostic and therapeutic limitations in progressive MS, early treatment of patients with RRMS is essential to prevent or delay the characteristics of the initial course, which are associated with poor long-term outcomes[159-162]. As well as the identification of PIRA events early in their development is crucial to optimize the therapeutic strategy[173].

1.2.2b What tools are available for MS monitoring in the clinical setting?

The most important tools used for clinical monitoring of MS patients are the neurological examination and the clinical history[174].

In 1955 Kurztke described "a scale for assessing disability in MS", later known as the Disability Status Scale (DSS)[175]. This scale was used in the first multicentre, randomized, placebo-controlled, double-blind clinical trial to evaluate a therapy against MS. The DSS had 10 values between 0 (normal), extending to state 10 (death from MS). The scale was "intended to measure the maximal function of each patient as limited by neurologic deficits,"' and it was based on neurologic examination. The functional groups, later called Functional Systems (FS), were Pyramidal, Cerebellar, Brain Stem, Sensory, Bowel & Bladder, Visual, Cerebral or Mental, and Other or Miscellaneous Functions. Although the DSS was considered satisfactory in several treatment trials in acute bouts, it was thought that there should be more room for change in studies of chronic MS[176].

Thus, in 1983[177] Kurtzke made some modifications on the scale, then called expanded DSS (EDSS). EDSS provides, for each step from 1 through 9, two steps that together add up to the same step of the original DSS. This division allows a better expression of the neurological examination encoded in the different FS, being able to record greater changes and more accurately reflect the functional status of the patient with MS. In fact, it is fully defined in the lower ranges by the FS grades, while the upper range of the scale (>EDSS 6) measures handicaps of MS patients. The determination of EDSS 4 – 6 is heavily dependent on aspects of walking ability (Table 1.2.2A).

Table 1.2.2A. Definitions for a standardised, quantified neurological examination and assessment of Kurtzke's Expanded Disability Status Scale in Multiple Sclerosis

Expanded Disability Status Scale	
0	Normal neurological exam (all FS grade 0)
1.0	No disability, minimal signs in one FS (one FS grade 1)
1.5	No disability, minimal signs in more than one FS (more than one FS grade 1)
2.0	Minimal disability in one FS (one FS grade 2, others 0 or 1)
2.5	Minimal disability in two FS (two FS grade 2, others 0 or 1)
3.0	Moderate disability in one FS (one FS grade 3, others 0 or 1) though fully ambulatory; or mild disability in three or four FS (three /four FS grade 2, others 0 or 1) though fully ambulatory
3.5	Fully ambulatory but with moderate disability in one FS (one FS grade 3) and mild disability in one or two FS (one/two FS grade 2) and others 0 or 1; or fully ambulatory with two FS grade 3 (others 0 or 1); or fully ambulatory with five FS grade 2 (others 0 or 1)
4.0	Ambulatory without aid or rest for ≥500 meters; up and about some 12 hours a day despite relatively severe disability consisting of one FS grade 4 (others 0 or 1) or combinations of lesser grades exceeding limits of previous steps
4.5	Ambulatory without aid or rest for ≥300 meters; up and about much of the day, characterised by relatively severe disability usually consisting of one FS grade 4 and combination of lesser grades exceeding limits of previous steps
5.0	Ambulatory without aid or rest for ≥200 meters (usual FS equivalents include at least one FS grade 5, or combinations of lesser grades usually exceeding specifications for step 4.5)
5.5	Ambulatory without aid or rest for ≥100 meters
6.0	Unilateral assistance (cane or crutch) required to walk at least 100 meters with or without resting (see chapter 8, Ambulation)
6.5	Constant bilateral assistance (canes or crutches) required to walk at least 20 meters without resting (see chapter 8, Ambulation)
7.0	Unable to walk 5 meters even with aid, essentially restricted to wheelchair; wheels self and transfers alone; up and about in wheelchair some 12 hours a day 7.5 unable to take more than

	a few steps; restricted to wheelchair; may need some help in transferring and in wheeling self
8.0	Essentially restricted to bed or chair or perambulated in wheelchair, but out of bed most of day; retains many self-care functions; generally has effective use of arms
8.5	Essentially restricted to bed much of the day; has some effective use of arm(s); retains some self-care functions
9.0	Helpless bed patient; can communicate and eat
9.5	Totally helpless bed patient; unable to communicate effectively or eat/swallow
10	Death due to MS

Another important instrument for MS patients monitoring is the Multiple Sclerosis Functional Composite (MSFC)[178], which was developed in 1994 by the MS Society's Clinical Assessment Task Force as an additional clinical measure of MS disability progression. The primary goal for creating the MSFC was to improve the standard measure of MS disability for clinical trials and to develop a multidimensional metric of overall MS clinical status more amenable to statistical methodologies[179]. MSFC is a three-part performance scale for evaluating the degree of impairment in MS patients. It includes the assessment of leg function by moving a short walking distance ("Timed 25-Foot Walk", T25FW), the assessment of arm function using breadboard test ("Nine-Hole Peg Test", 9HPT) and an attention/concentration test to assess cognitive functions ("Paced Auditory Serial Addition test", PASAT). Subsequently, a visual pathway test (low contrast letter accuity) was also included and PASAT was replaced by Symbol Digit Modalities Test (SDMT)[180].

Over time, the EDSS became the most popular and widely used instrument to assess the disability status of MS patients, while the MSFC was relegated, exclusively, to clinical trials[181]. But, beyond these scales, and in the absence of biomarkers, there is a well-known tool that, although very old, should never be underestimated: the clinical history. A systematic and oriented anamnesis could be a key anchor in the identification of progressive disease and disease activity[174].

The characteristics, in terms of validity, reliability, sensitivity to change and interpretability, of the EDSS and MSFC scores are described below (Table 1.2.2B).

Table 1.2.2B. Characteristics of EDSS and MSFC

	EDSS	MSFC
Validity	EDSS showed good correlation with the Barthel Index, the London Handicap Scale, the Scripps Neurological Rating Scale, the Functional Independence Measure[182], the physical function ability of the Short Form-36[183-185] and other patient-reported outcomes. Amato and colleagues[186] certify a good validity of the EDSS; Ebers[187] described the EDSS as "well validated".	MSFC showed good correlation with other indicators of disease in MS, including the EDSS, MRI measures[180,188], patient reported health[189-191] and employment[192]. While the overall MSFC scores distinguish across MS disabilities, the PASAT does not do as well[193]. In fact, Brochet et al.[180] found no correlation between PASAT and EDSS.
Reliability	It was shown a moderate inter-rater reliability[194-195], with higher intra-rater agreement. Different studies showed a substantial proportion of disagreement between evaluators when referring to the lowest scores of the scale (1.0-3.5)[196-197]	Inter-rater and intra-rater reliability were shown to be high for MSFC overall score as well as for its component measures[193,198-199].
Sensitivity to change	Some studies showed that the rates of change varied depending on the initial value. Greater rates of change were observed for minor severity of disease, but from an EDSS score of 6, the EDSS showed very little change in sensitivity[200-201]. Different authors found EDSS to be less sensitive to change than the Dmax (the maximum distance that a person can go) and T8 (time to walk 8 minutes)[202], the Rivermead Mobility Index, the Ambulatory Index and the 10 minutes walking time test Vaney[203], and the DSS[201].	Some studies indicate the MSFC to be more sensitive than EDSS to detecting changes in disease[204-205]. Authors also found patients with MSFC deteriorating by at least one standard deviation to have an odds ratio of 2.1 for sustained EDSS deterioration[206-207].
Interpretability	There is no clear recommendation on how to interpret changes in EDSS values. Noseworthy et al.[194] recommended a progression of 1.0 as a meaningful change in clinical trials. Francis et al.[208] suggested a change of 1.5 points to be more appropriate. The guidelines of the European Medicines Agency (EMA)[209] recommend defining treatment success or	A major drawback of the interpretation of the MSFC is its scoring mechanism (z-scores)[210]. To overcome this limitation, Bosma et al.[211] found optimal cut-off values, considering clinically relevant a 20% change for both T25FW and 9HPT. Hoogervorst et al.[189] established that 20% change in all three MSFC measures

	treatment failure of either reaching a certain EDSS score or a sustained change in sufficient volume. A separate consideration of the lower and upper value range of the EDSS is more recently recommended: 1 point on the EDSS scale with a baseline EDSS score less than or equal to 5.5 and 0.5 points in an EDSS score over 5.5.	were meaningful to patient's own perception of health.

1.2.2c Current limitations in MS clinical monitoring

Over time, few changes have taken place in MS clinical monitoring[157]. The neurological exam is arguably the most elegant in all of medicine, yet our standard tools for evaluating MS patients have become antiquated. J. Madison Tyler designed the first reflex hammer in 1881 and Henrich Rumpf published on the use of the tuning fork for vibration sense in 1889[212-213]. In 1983 Kurtzke, using these tools and based on the neurological examination findings, described the EDSS, which continues to be the current gold standard in routine clinical practice to demonstrate the functional status of MS patients[177,181].

Despite EDSS was consolidated as a "gold standard" for disability monitoring in MS patients, it has some disadvantages[157]. For example, the lack of accuracy when evaluating non-physical symptoms of MS such as cognitive impairment, fatigue or depression[214-215]. The isolated combination of the EDSS value and disease duration has led to mistakenly define in many cases, as benign MS, patients with low physical impact of the disease, but with poor quality of life due to a significant burden of invisible symptoms[216-217]. In this sense, patient reported outcomes (PROs) focused on health-related quality of life including symptoms, functions, cognitive status and social aspects among others, used as complementary measures to the EDSS, could contribute to obtaining a more real picture of the patient's functional status[218-219].

Another disadvantage of the EDSS is that time and space requirements limit its use in clinical practice[157]. Regarding this, a recent survey addressed to

neurologists specialized in inflammatory-demyelinating disorders reported that only 25.5% use the EDSS on a regular basis, a frequency that was even lower when they were specifically asked about the evaluation of the patients' walking distance[220].

Furthermore, EDSS only reflects the patient physical status at a specific point in time during infrequent visits (once- or twice-yearly in-clinic visits) that often do not correlate with the real patient performance throughout the whole day, nor throughout the days[221].

Finally, on the one hand, although the overall inter-rater EDSS reliability is good-moderate, different studies showed a substantial proportion of disagreement between evaluators when referring to the lowest scores of the scale (1.0-3.5)[194-195]. On the other hand, EDSS has shown weak sensitivity in detecting small changes in function over time[200-201].

Therefore, despite recent advances in the care of MS patients, the available tools used by neurologists still have several limitations for the clinical monitoring[157]. Issues like the early detection of disease progression[158], the quantification of disability progression/worsening, the characterization of invisible symptoms related to aspects of quality of life, as well as the correct identification of certain possible relapses sometimes misinterpreted as fluctuations or pseudo-relapses, although critical to optimally adapt the therapeutic strategy, currently remain a challenge[222-225].

In this context, recent advances in sensor technology and the computer control industry could provide the MS specialist with the opportunity to measure neurological function in exciting new ways that may enable the capture of subclinical symptoms/progression of the disease and the prevention of adverse outcomes[226].

1.2.3 Patient Reported Outcomes

Different studies showed that invisible symptoms of MS, such as fatigue, depression, and cognitive impairment, may have a greater negative impact than

physical symptoms on the quality of life (QoL) of patients with MS[214-215,218-219]. However, these symptoms are still being underreported in registries based on clinician reported outcomes, which focus on the assessment of relapses, radiological activity, and disability measured by the EDSS[227]. This situation emphasizes the importance of including standardized PROs in clinical practice to better understand the impact of symptom burden in MS and thus have a more realistic picture of the patient's functional status[218-219].

PRO is an umbrella term used to define any information "of a patient's health condition that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else"[228]. PROs commonly include measures of symptoms, functioning (activity limitations), QoL and health related QoL (HRQoL) as well as patient satisfaction, compliance, and treatment preferences[218-219]. PROs measurement instruments, usually standardized questionnaires, are called patient-reported outcome measures (PROMs)[229]. PROMs can be categorized as either generic or targeted. Generic measures include questions that are general enough for use with both healthy and a variety of clinical populations. On the other side, there are PROMS targeted to specific diseases, symptom clusters/ domains, regions or populations[230].

PROs have been mainly used in post-marketing, observational studies, but nowadays they are increasingly used as secondary or tertiary outcomes in MS clinical trials on DMDs and symptomatic treatments, whereas in rehabilitation trials they are used as primary or co-primary outcomes[231].

For their widespread use, PROMs should meet some fundamental characteristics, such as having been validated, being available in different languages, being developed in partnership with patients and with an understanding of what really matters to them and being cost-effective and well-accepted by patients[230-231]. Ideally, PRO data should be linked with other clinical data of the disease (derived from clinical registries) to create integrated datasets that can provide large amounts of information which could serve as key prognostic covariates in the assessment of MS[232-233].

However, beyond what has been described above, many limitations remain to be solved before successfully incorporating the contributions of patients into clinical practice. Most clinicians are not familiar with how to translate the QoL data provided by PROs into a medical action or how to interpret when the score of an evaluated domain represents a clinically relevant condition. In turn, despite their popularity, little is known about the clinical utility of these PROMs in MS clinical monitoring, particularly when evaluating changes in disability and QoL over the longer term[229-231].

Therefore, in order to adequately measure patient health, illness, and healthcare performance, it is mandatory to define a standardized capture of patient clinical data by PROMs to target on precisely defined and measurable aspects of health that can be improved by medical interventions[230-231].

1.3 e-HEALTH

Elecronic-health (e-health) refers to the use of electronic communication tools and computers by patients to monitor or maintain their own health and by doctors to provide better care[234]. It has created opportunities for integration of datasets across centres and countries to generate datasets large and comprehensive enough to characterize less common populations and events and support the development of more meaningfully predictive models of individual prognosis. The latter promise greater personalization of care and more rapid discrimination of the relative benefits and risks of new DMDs[235-236].

More than 95% MS patients have access to a mobile device and most use it routinely[237]. The last decade has seen an explosion in the capability of monitoring individuals via biosensors in smartphones or wearable devices, and the range of parameters which can be measured by such technologies will continue to grow[238-239]. Smartphones can be used in a patient's home environment as often as necessary, being familiar and discreet[239]. Furthermore, most off-the-shelf smartphones contain sensors with the capacity to gather objective data unaffected by inter- and intra-rater variability[221].

Patients' symptoms and physical impairments often go undetected by healthcare providers, particularly in intervals between clinic visits[218-219]. In addition, MS patients are frequently struck by different degrees of cognitive impairment and may forget what they were feeling one or two weeks before the scheduled visits[240-241].

The development of remote measurement technologies (RMT) is an innovation which could, in the foreseeable future, be used to predict and avert negative clinical outcomes by providing real time information on the patient's current clinical state, as well as providing predictive information indicative of a future deterioration[242]. In common with other "big data" projects, RMT has been hailed as a potential "paradigm shift" in the way in which clinical services can be delivered but will require healthcare adaptation for its delivery[234].

Not only objective measurements but also PRO data could be captured with smartphone technology and have the potential to enable more frequent, decentralised, and home-based care to supplement the infrequent in-clinic assessments typically offered to patients. Mutually sharing this information with patients can help focus the clinical conversation or empower shared decision making[232-233].

Today, smartphone applications and biosensors perform a variety of functions. Many of these tools are wellness applications designed to support day-to-day disease management, for example through symptom or medication intake tracking, visit-scheduling, provision of disease education, and connectivity to supportive care facilities or patient social media networks[243-248]. Other smartphone-based solutions enable assessment of functional parameters affected by the disease, such as mobility and cognition, or therapeutic benefit, such as for fatigue or depression[249].

Accelerometers are the most widely used biosensors to measure mobility in MS patients. The reliability and accuracy of step counting by accelerometers have been correlated with step counts in healthy controls and in patients with mild disability, as well as with EDSS between 0 and 6.5[2520]. However, we need long-term studies to determine if accelerometers are useful in detecting increased disability or disease progression over time. In this sense, a pilot study that included 11 MS patients, and a more recent one with 95 patients, showed during the one-year follow-up period that the accelerometer could identify fluctuations and deterioration of functional status of patients earlier and with greater precision than the traditional tests used in clinical practice, including the EDSS, even when these scores remained stable[251-252]. Another study aimed at evaluating early changes in gait through accelerometers and included 102 patients and 22 healthy controls. Participants were asked to perform the T25FW twice at a 'self-selected' speed, followed by twice at a maximum speed. The analysis of subgroups between controls and patients showed significant differences in the mean length of the step, gait speed, support time foot and swing time. The differences between the subgroups according to the EDSS were more pronounced in the

tests with maximum speed gait[253]. Despite its promising benefits, wrist-worn accelerometers have identified limitations, such as that work best in patients with lower disability levels, do not capture well upper limb function, work best for walking as opposed to other forms of activity and can be confounded by activities like washing dishes. Data provided by accelerometers may be enhanced with use of gyroscope or other complementary sensors that provide additional information[226].

Gyroscopes have been used in approximately 20% of research studies on biosensing in MS[254]. They have shown good test–retest reliability[255-256]. They have been used not only to improve the accuracy of accelerometer data, but also to detect falls, digitize the timed up and go test, quantify gait, and measure tremor[255-258].

Other types of sensors that have been used in MS patients include grip sensors, electrodermal sensors (to detect stress and fatigue), and surface or portable electromyograph[259-261]. A number of clinical trials from the past few years include biosensor outcomes, from rehabilitation trials to more standard pharmacological studies, exploring sensor data as secondary or exploratory outcomes (clinicaltrials.gov)[226].

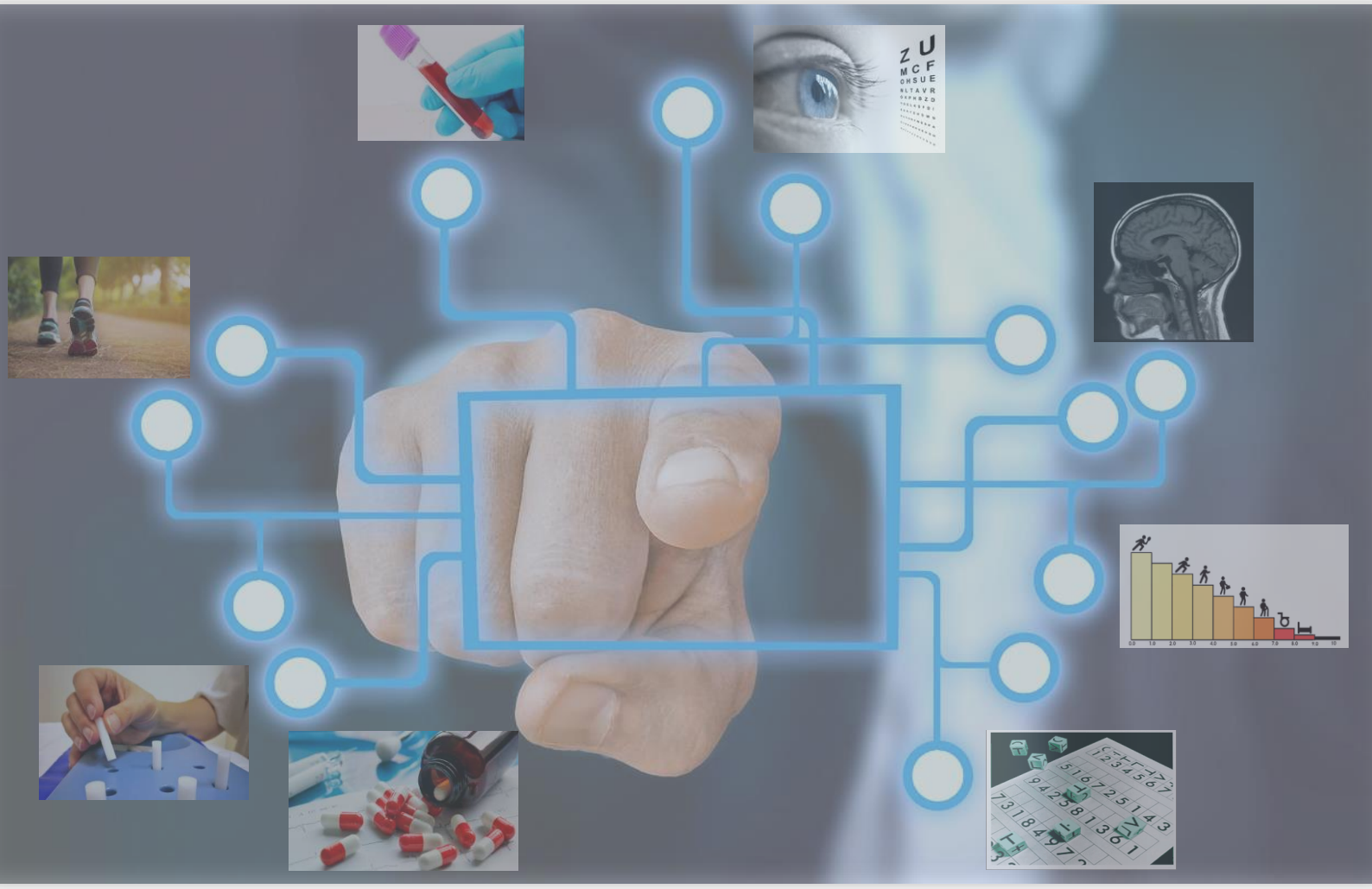
So, in order to overcome some difficulties in clinical practice and minimize the burden on care providers and the healthcare system, digital technology could play a key role by offering an easy way to collect extensive clinical information, including free-living data and PRO focused on non-physical MS symptoms, providing a more complete and real picture of the functional situation of MS patients, ultimately improving the understanding of individual disease trajectories and response to intervention[157]. Despite their promise, however, smartphone-based solutions have not yet been fully integrated into routine medical practice.

The optimal approach to validate new technology for clinical use is a multi-step, rigorous process. First, extraction of sensor data through more standard signal processing must be related to a ground truth about the disease state to ensure

clinical significance. Second, reproducibility and reliability need to be established. Precision of measurement will ultimately determine utility. Finally, sensitivity to change over time must be established. This is perhaps the most challenging to determine. Finding associations in cross section are much easier, but to advance the field of MS therapeutics, in particular for progressive MS, measures sensitive to change over the short term are mandatory. Other factors that will determine the adoption of biosensors into MS clinical research and practice include degree of disruption to patient life or clinical work flow, cost and the ethics of confidentiality, and patient safety[226].

Practical considerations include setting standards for data quality and collection for these devices. Real-time delivery of analyzed data will be more useful for clinicians. To achieve this goal, it will be necessary to provide software (along with hardware) that is commercial grade in terms of reliability and that is either integrated with a sensor device or easily uploaded to a PC or mobile device. Bluetooth or other wireless connectivity to transfer data is essential[226,262].

Multi-sensor systems are likely to rise to the top of the biosensing world as they can provide complementary data for a richer representation of a particular motion or event. The ever-increasing variety of biosensors and the exponential increase in publications make it inevitable that neurologists in the coming decades will see discrete biosensor data leak into their clinics, enforcing the rules of reliability and clinical utility to determine which biosensors rise to the top[226].



HYPOTHESES

2. HYPOTHESES

We started out with the following general hypothesis: digital technology can contribute to clinical MS monitoring providing a more granular and realistic picture of the functional situation of MS patients.

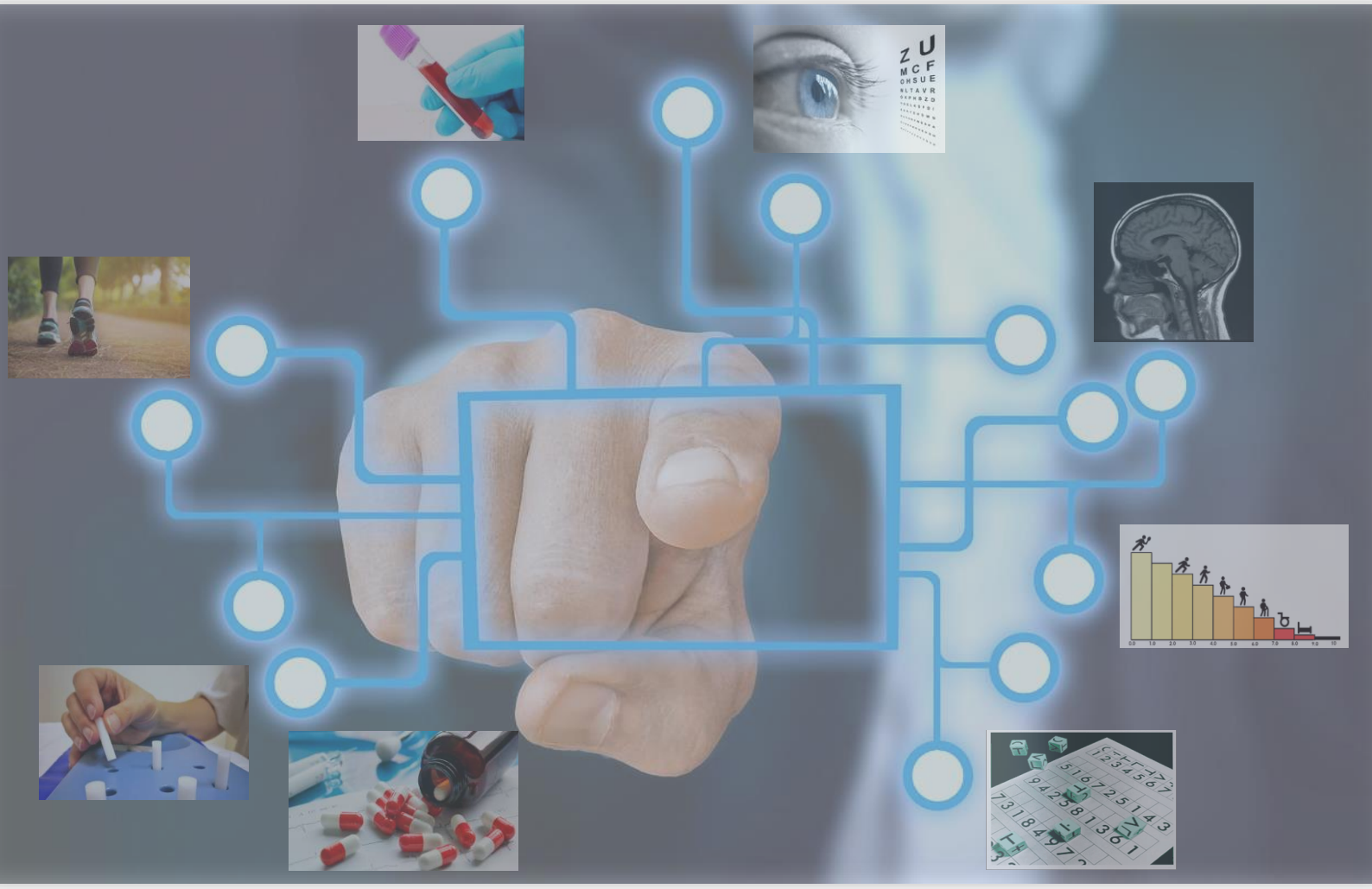
Although the search for digital markers for MS monitoring is a topic of great interest and addressed worldwide, nowadays we do not have any available digital tool in routine clinical practice for this purpose. As noted above, the optimal approach to validating a new technology for clinical use is a rigorous, multi-step process that is often not followed. In this thesis project we focus on the main areas (feasibility and clinical validity) that must be addressed early during the design of a digital tool. In this way, together with Hoffmann-La Roche, we created the Floodlight MS application (app) that contains a series of active tests and passive monitoring to measure, at home and in an unsupervised setting, functional ability in cognitive, upper extremity, and gait and balance domains. Then, we launched the pilot study "Monitoring of Multiple Sclerosis Participants With the Use of Digital Technology (Floodlight study)" to explore its feasibility and clinical validity.

1. We wanted to assess the adherence to smartphone- and smartwatch-based assessments and collect the feedback from the participants. In a recent study, it was reported that both healthy controls (HCs) and people with MS (pwMS) were capable of completing daily tasks on a smartphone for one year. So, the hypothesis is that healthy volunteers and MS patients adhere well to the Floodlight MS app and that its use does not have a major impact on their daily activities.

2. We wanted to determine the association between exploratory sensor-based outcomes derived from the components of the Floodlight MS app and conventional MS clinical outcomes. Smartphone sensor-based,

remotely administered digital tests represent a promising new avenue to capture disability burden with quantitative accuracy. Such tests typically allow the remote capture of ecologically valid measures in a frequent, time- and cost-efficient fashion with minimal patient burden. In addition, they quantify multiple aspects of nervous system function during a given task, in opposition to clinician-administered disability tests that only assess the capacity to complete a task with a performance summary score. In recent years, sensor-based tests have been increasingly used to assess functional ability in MS, as well as in Parkinson's disease and Huntington's disease. The hypothesis is that components of the Floodlight MS app correlate well with clinician-administered assessments tipically use for MS monitoring (such as EDSS, MSFC).

3. We wanted to evaluate the correlation between brain structure at a global and regional level as measured with MRI and digital measures derived from Floodlight MS app. Although there have been efforts to link clinical disability to regional atrophy, the relationship between digital measures of functional ability and MRI outcomes has not been extensively studied. We hypothesize that global and regional atrophy measured by MRI could correlate well with the digital measures captured with Floodlight MS app.



OBJECTIVES

3. OBJECTIVES

The objectives of this thesis are:

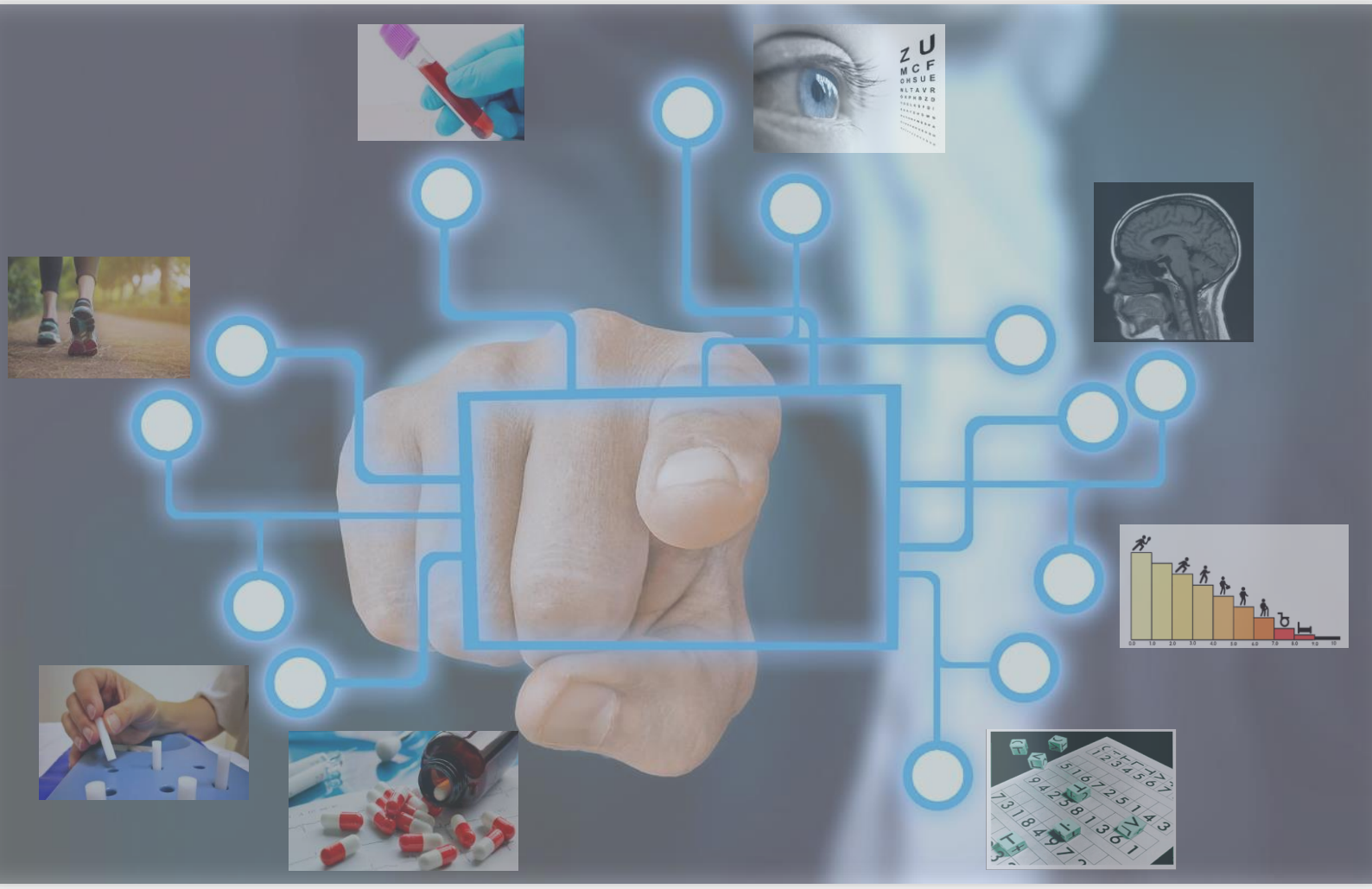
- Main objective

1- To evaluate participants adherence to smartphone- and smartwatch-based assessments integrated in the Floodlight MS app and collect feedback from MS patients and HCs on the smartphone and smartwatch schedule of assessments and its impact on their daily activities using a patient satisfaction questionnaire.

- Secondary objectives

2- To determine the association between exploratory sensor-based outcomes derived from the respective components of the Floodlight MS app and conventional MS clinical outcomes and explore whether the Floodlight MS app test could differentiate between participants with and without MS.

3- To link digital measures of functional disability to specific MRI global and regional atrophy patterns.



MATERIALS AND METHODS

4. MATERIALS AND METHODS

4.1 Trial Design and Participants

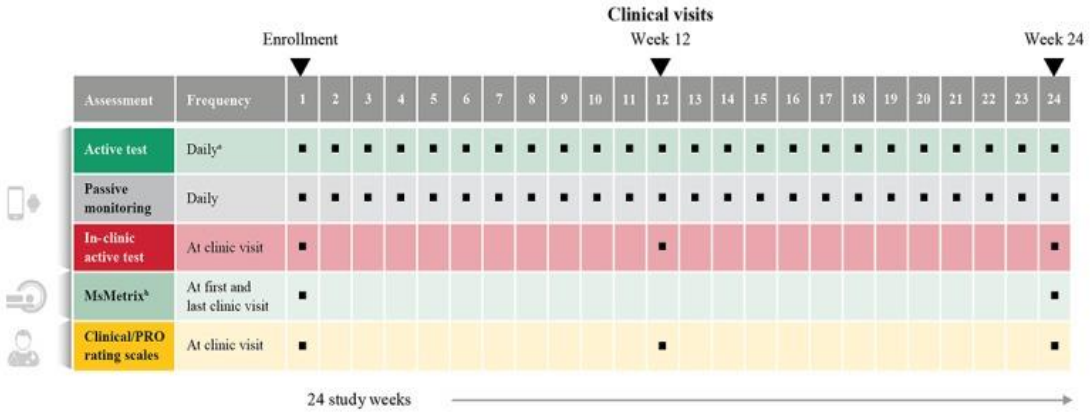
After providing written informed consent, pwMS and HCs, preferentially partners or cohabitants, were evaluated for eligibility for enrollment in the study "Monitoring of Multiple Sclerosis Participants With the Use of Digital Technology (Floodlight MS app)". Eligibility criteria for pwMS included the ability to comply with the study protocol, age 18 to 55 years, diagnosis of MS (2010 revised McDonald criteria, treated or untreated)[79], EDSS score 0 to 5.5 (inclusive), and weight from 99 to 243 lbs (45 to 110 kg). An EDSS score of 5.5 as a maximum limit was meant to ensure any patient with relapsing or progressive MS would not have any significant difficulty in participating in the proposed testing as per study protocol. Further details on eligibility criteria are provided in Appendix 10.1

This study was conducted at two sites in two countries with a total of 76 pwMS and 25 HCs; 60 pwMS and 20 HCs were recruited from our Centre, the Multiple Sclerosis Centre of Catalonia, Vall d'Hebron University Hospital, Barcelona, Spain, and 16 pwMS and 5 HCs were recruited from the University of California, San Francisco, California.

The protocol, informed consent forms, any information given to the participants, and relevant supporting information were reviewed and approved by the institutional review board/ethics committee before the study was initiated. Confidentiality was maintained by assigning each participant enrolled in the study a unique identification number.

4.2 Study Design

The Floodlight study combines continuous sensor data capture with smartphones and smartwatches and standard clinical outcome measures. Eligible pwMS and HCs were enrolled in the study and assessed clinically at the enrollment visit, week 12, and termination visit (week 24). In addition, participants were asked to perform a set of daily active tests and contribute sensor data via passive



monitoring with smartphone and smartwatches over a period of 24 weeks (Fig. 4.2A).

Figure 4.2A. Floodlight study design. PRO: patient-reported outcome. "a" indicates that active tests were administered daily, weekly or every two weeks (see next figure for schedule). "b" indicates that brain magnetic resonance imaging was performed in people with multiple sclerosis.

4.3 In-Clinic Assessments

At each scheduled in-clinic visit (enrollment, week 12, and week 24), the following reference clinical tests were performed for all participants: 9HPT, oral version of SDMT[263-265], T25FW test, Berg Balance Scale (BBS)[266], Fatigue Scale for Motor and Cognitive Functions (FSMC)[267], and Patient Health Questionnaire–9 (PHQ-9)[268]. For pwMS only, disability was measured by EDSS, Patient Determined Disease Steps (PDDS)[269], and Multiple Sclerosis Impact Scale–29 (MSIS-29; version 2)[270-271]. While performing some of the in-clinic tests, pwMS and HCs were asked to carry or wear the smartphone and smartwatch to collect sensor data alongside the in-clinic measures. On the scheduled in-clinic visits, the smartphone and smartwatch Floodlight active tests were performed under investigator supervision. The satisfaction questionnaire (Appendix 10.2) assessed pwMS' and HCs' experience regarding smartphone and smartwatch use and its impact on their daily activities at the week 12 visit and at the study termination/early discontinuation visit. Brain MRI was performed in pwMS at the enrollment visit and at week 24.

4.4 Smartphone and Smartwatch Testing

At the enrollment visit, pwMS and HCs were provided with the Floodlight solution that included a smartphone and smartwatch preconfigured so participants could only run the Floodlight software. A belt bag was also provided for participants to carry their smartphone in an anterior medial position. The smartphone and smartwatch pair contained a preinstalled app that prompted the user to perform various assessments, referred to as active tests. The app also passively recorded sensor data, referred to as passive monitoring. At the enrollment visit, participants received training on the use of the smartphone and smartwatch and were provided with supporting content to help them complete the tests successfully. Participants were instructed to complete the active tests at approximately the same time each day and carry the smartphone and smartwatch throughout the day, recharging the devices overnight. Data transfer from the smartphone and smartwatch is described in Appendix 10.1.

4.4.1 Floodlight Active Tests

PwMS and HCs were asked to perform various active tests (daily, weekly, every two weeks, or on demand) via the smartphone (Fig. 4.4.1A and Table 4.4.1A). These novel active tests were developed to be self-administered by pwMS to capture MS symptoms. A range of clinical and sensor-based assessments were chosen to capture the most prominent symptoms of MS from a broad spectrum of symptoms. PwMS and HCs were required to wear the smartwatch throughout the active tests.

	Active tests									Passive monitoring	
Test type	Experience sampling			Cognition	Hand & arm		Gait & posture			Gait & posture	
Test name	DMQ	ST	MSIS-29	SDMT	Pinching Test	Draw a Shape Test	SBT	5UTT	2MWT	Gait Behaviour	Mobility Pattern
Frequency	Daily	Fortnightly & ad hoc	Fortnightly	Weekly	Daily	Daily	Daily	Daily	Daily	Continuous	Continuous

Figure 4.4.1A. Floodlight active tests and their schedule frequency. DMQ: Daily Mood Question; MSIS-29: Multiple Sclerosis Impact Scale–29; SBT: Static Balance Test; SDMT: Symbol Digit Modalities Test; ST: Symptom Tracker; 2MWT: Two-Minute Walk Test; 5UTT: 5 U-Turn Test.

Table 4.4.1A. Floodlight MS app: active tests

Domain and test	**Short description**
Daily hand motor function tests[a]	
• Draw a Shape Test	The aim of the Draw a ShapeTest is to assess fine finger/manual dexterity while the participants are instructed to hold the mobile device in the untested hand and draw on the smartphone touchscreen six prewritten alternating shapes of increasing complexity (linear, rectangular, circular, sinusoidal, and spiral) with the second finger of the tested hand as fast and as accurately as posible within a maximum time (30 seconds for each of the two attempts per shape).
• Pinching Test	The aim of the Pinching Test is to assess fine pinching/grasping dexterity while the participants are instructed to hold the mobile device in the untested hand and touch the screen with two fingers from the tested hand (thumb + second or thumb + third finger preferred) to squeeze/pinch as many round shapes (tomatoes) as they can during 30 seconds.
Daily gait tests[b]	
• Two-Minute Walk Test (2MWT)	Participants are instructed to walk as fast as long as they can for 2 minutes but walk safely. The 2MWT is a simple test that is required to be performed on an even ground in a place where participants have identified they could walk straight for as far as ≥200 meters without U-turns. Participants are allowed to wear regular footwear and an assistive device and/or orthotic as needed.
• 5 U-Turn Test (5UTT)	The aim of this test is to assess difficulties or unusual patterns in performing U-turns while walking on a short distance at comfortable pace. The 5UTT can be performed indoor or outdoors, on an even ground where participants are instructed to walk safely and perform five successive U-turns going back and forward between two points a few meters apart for 1 minute. Participants are allowed to wear regular footwear and an assistive device and/or orthotic as needed.

• Static Balance Test (SBT)	Participants are asked to stand still unsupported for 30 seconds with relaxed arms straight alongside the body if possible.
Weekly cognitive test • Electronic version of the Symbol Digit Modalities Test (SDMT)	The aim of SDMT testing is to detect impairment of key neurocognitive functions that underlie many substitution tasks.
Patient-reported outcomes (PROs) • Daily Mood Question (DMQ)	This test represents an assessment of participants' perceived overall state by responding daily to the question "How do you feel now?" on a 5-item Likert scale, ranging from excellent to horrible.
• Electronic version of the Multiple Sclerosis Impact Scale-29, version 2 (MSIS-29); pwMS only	This questionnaire measures the physycal and psychological impact of multiple sclerosis.
• Multiple Sclerosis Symptom Tracker (MSST); pwMS only	Patients are asked if they experienced new or significantly worsening symptoms during the last 2 weeks. If yes, onset of the symptoms and the patients' perception to whether they think they experienced a relapse (yes, no or unsure) are recorded.

[a] Tests alternatingly performed with right and left hand; users are instructed on daily alternation.

[b] Recommended position of smartphone in an anterior medial position in the belt bag

4.4.2 Floodlight Passive Monitoring

Passive monitoring collected metrics on gait and mobility throughout the day in a continuous and unobtrusive manner. Participants were instructed to carry their smartphone preferably in an anterior medial position in a belt bag or, alternatively, in their pocket and wear the smartwatch all day as they went about their daily routine until the devices ran out of charge.

4.5 Signal processing

As the smartphone-based tests were performed without supervision by a physician or study coordinator, quality control steps were applied to identify and exclude individual assessments that were performed incorrectly. This ensures the measurements are both reliable and accurate. To exclude such incorrectly performed assessments, quality control flags were defined for each test. In addition, only sufficiently adherent participants, that is, those who contributed at least six individual assessments in the course of the study, were included in the analysis of that particular test. Applying these two quality control steps resulted in the final dataset consisting of valid assessments.

Next, all valid assessments contributed by a participant were aggregated to study test-retest reliability and Spearman's rank correlations in a cross-sectional analysis. As the tests were performed once daily at most, the test-retest analysis was based on the median test performance on the active tests and passive monitoring during 12 two-week windows. Two-week windows were chosen to reduce variability that is independent of general disease status and might be attributed to good or bad days or to differences between weekdays and weekends. For the cross-sectional correlation analysis, the median test performance across the entire study duration as well as the mean of the three in-clinic assessments (mean of two assessments for MRI) were computed.

4.6 Statistical Analyses

The analyses of the first objective of this study were descriptive. Statistical tests were exploratory and conducted at the two-sided 5% significance level without adjustment for multiple comparisons. The analyses were based on all enrolled patients (full analysis set [FAS]). Patients who prematurely withdrew from the

study for any reason were still included in the FAS. Supportive analyses of selected variables were carried out in the per-protocol population, which included pwMS who completed at least 1 week in the study and did not discontinue due to an adverse event or a reason unrelated to the use of the Floodlight solution (Appendix 10.3).

Adherence was evaluated for the following tests and test groups: all Floodlight MS active tests, 2MWT, all active tests except 2MWT, smartphone use, smartwatch use, and for the per-protocol population.

Adherence to active tests was measured as the proportion of study weeks with at least 3 days of completed testing (study co-primary endpoint). Adherence to sensor-based passive monitoring was measured as the proportion of study weeks with at least 3 days of passive monitoring for at least 4 hours per day while the devices were worn by the participant[272] (study co-primary endpoint). Descriptive statistics of calculated adherence were reported for all active tests, all active tests except 2MWT, 2MWT, smartphone use, and smartwatch use. Categorical and numeric variables with fewer than five values were tested for association with adherence using the Kruskal-Wallis test. The association between continuous variables was assessed using the Spearman rank correlation.

Participant complete abandoning of active testing and passive monitoring was also investigated in a time-to-event survival analysis based on the Kaplan-Meier method along the Floodlight study. The abandoning event was defined as the last week in which the participant was adherent according to the definitions above for active tests and passive monitoring. Active tests performed on days of in-clinic visits were not considered in the adherence calculation to focus the abandoning analysis on the remote use. Participants leaving the study before the terminal visit were considered as censored. The impact of different characteristics on adherence was assessed using Cox regression.

A satisfaction score was developed from the satisfaction questionnaire (Appendix 10.2) that sums the individual answers to questions 1-7 and 10-12 rescaled to 0-100 from their original Likert scale. An interquestion correlation analysis was performed to ensure questions are equally correlated and can be combined.

Descriptive statistics of satisfaction score and items are reported, along with covariate analyses of demographics and disease state. Analysis of the change in satisfaction score between week 12 and week 24 are reported using the Wilcoxon signed-rank test.

Patient baseline characteristics incorporated as covariates in the analysis of correlation with Floodlight adherence and satisfaction outcomes were age, gender, body mass index, time since first MS symptom onset, EDSS, T25FW time, 9HPT time, and the oral SDMT correct responses. A descriptive analysis of safety variables, including adverse events and serious adverse events, was carried out in the FAS.

For the second objective analyses, test–retest reliability was evaluated with intraclass correlation coefficients (ICC) separately in pwMS and HC, which considered all consecutive 2-week windows. Generally, at least three valid individual assessments were required for each 2-week period. An exception was made for the e-SDMT to accommodate its weekly testing schedule; only one valid e-SDMT assessment for each 2-week window was considered sufficient. Test-retest reliability was considered as poor (ICC < 0.5), moderate (ICC = 0.5–0.75), good (ICC = 0.75–0.9), or excellent (ICC > 0.9). To examine the agreement of the test features from the Floodlight MS app with standard clinical and MRI measures in pwMS, the test features were correlated against domain-specific clinical measures (oral SDMT, 9HPT, Berg Balance Scale [BBS], T25FW), EDSS, MSIS-29 items or subscales, T2 FLAIR (fluid attenuated inversion recovery) lesion volume, and normalized brain volume using Spearman's rank correlation. In addition, test features from UTT and Walk Test were correlated against corresponding features obtained from Passive Monitoring. The strength of correlation was considered as good-to-excellent (|r|> 0.75), moderate-to-good (|r|= 0.5–0.75), fair (|r|= 0.2–0.49), or not correlated (|r|< 0.25), where |r|represents the absolute value. Partial correlation analyses were performed for the upper extremity function tests and gait assessments to investigate the contribution of each test and test feature in predicting 9HPT times and T25FW times, respectively. All test features from the Pinching Test and the Draw a Shape

Test as well as 9HPT time were included in the upper extremity function model. Each of these features was correlated against each other while controlling for the remaining features in the model. The partial correlation for the gait assessments was run in a similar fashion but included all test features from the UTT, Walk Test, and Passive Monitoring in addition to T25FW time instead. All correlations were adjusted for age and sex with a robust linear model. Statistical significance was set at $p<0.05$ without correction for multiple comparisons.

For the third objective, brain MRI scans were collected from pwMS at baseline and at week 24 with Siemens 3-Tesla scanners (Siemens Healthcare GmbH, Erlangen, Germany) following optimized clinical practice protocols in place at the two investigating centers: (1) Hospital Universitari Vall d'Hebron, Barcelona, three-dimensional (3D) T1-weighted (repetition time [TR]=2,300 ms; echo time [TE]=2.98 ms; inversion time [TI]=900 ms; voxel size=1 × 1 × 1 mm) FLAIR (TR=6,000 ms; TE=394 ms; TI=2,100 ms; voxel size=1 × 1 × 1 mm) sequences acquired in a Siemens Trio Tim; (2) University of California, San Francisco, 3D T1-weighted (TR=2,300 ms; TE=2.98 ms; TI=900 ms; voxel size=1 × 1 × 1 mm) FLAIR (TR=5,000 ms; TE=389 ms; TI=1,800 ms; voxel size=1 × 1 × 1 mm) sequences acquired in a Siemens Skyra. MRI images of the brain were analyzed with icobrain ms v.5.0 (icometrix, Leuven, Belgium), a CE-certified algorithmic solution for MS MRI quantification[273-274], that uses a combination of (1) classical intensity-based modelling and multi-atlas segmentation, and (2) deep learning based on a convolution neural network. Icobrain ms allowed extraction of normalized volumes for 34 anatomical regions, including cerebral gray matter regions (*n*=20), deep gray matter regions (*n*=7), brainstem (*n*=3), cerebellum (*n*=2), lateral ventricles (*n*=1), and cerebral white matter (*n*=1). In addition, icobrain ms allowed calculation of the corpus callosum area (*n*=1) and upper cervical cord area (*n*=1). As the cervical spinal cord was within the field of view on brain scans, the mean upper cervical cord area along C2–C3 levels was computed from the brain images[275-276] using the icobrain ms fully automated pipeline[277-278] based on the Spinal Cord Toolbox[279] (see Table 4.6A for more details). No participants were excluded from the analyses due to low quality MRI data.

First, we correlated digital measures obtained with the Floodlight MS app in pwMS and measures obtained with the standard clinical measures of MS disease state (9HPT, oral SDMT, BBS, T25FW and EDSS) against regional, structural MRI outcomes using univariate Spearman's rank correlation. Since the digital measures may be confounded by age, sex and body mass index, we adjusted for these confounders in both analyses with a robust linear model. Statistical significance was set at $q<0.05$ after applying the false discovery rate (FDR) correction to correct for multiple comparisons. Next, we estimated the variance in the digital measures that can be explained by volumetric MRI data in a multivariate analysis. A Bayesian ridge regression model[280-281] with leave-one-out cross validation was used estimate the R^2 score, i.e., the proportion of the variance in the dependent variable (digital measure or standard clinical measure) that is predictable from the independent variables (structural MRI outcomes). Two models were considered for each digital measure: (1) the univariate Whole Brain model, which included the normalized whole brain volume to estimate the variance in the digital measures and standard clinical measures; (2) the multivariate Parcellation model, which instead included the normalized volume and cross-sectional area measurements of 36 MRI regions. Given the relatively short duration of the study (24 weeks) and the stability of the clinical and MRI measures over the study period, a data aggregation approach was followed to reduce variability and to deal more effectively with missing data. For both analyses, the data were aggregated as follows: for the digital measures, the median across all valid active tests was calculated; for the standard clinical measures and MRI outcomes, the mean of the three clinical visits (baseline, week 12, and week 24) and two MRI scans (baseline and week 24), respectively, were calculated.

Table 4.6A. Outcome variables

Floodlight MS app				
Functional domains	Smartphone-based active test	Functional subdomain	Digital measure	Higher scores indicate performance that is:
Upper extremity function	Draw a Shape Test	Accuracy	Mean trace accuracy	Better
		Swiftness of movement	Mean trace celerity	Better
		Smoothness of movement	CV of linear, angular and radial drawing velocity	Worse
	Pinching Test	Accuracy	Total number of successful pinches	Better
		Responsiveness	Gap time between pinch attempts	Worse
		Pinching asynchrony	Double touch asynchrony of successful pinches	Worse
Cognition	e-SDMT	Information processing speed	Number of correct responses	Better
			Maximum gap time between correct responses	Worse
		Cognitive fatigue	Speed fatigability index of the last 30 seconds	Better
Gait and balance	SBT	Postural control	Sway path	Worse

	UTT	Turning ability (dynamic balance and gait)	Turn speed	Better
	2MWT	Gait pace	Step frequency	Better
		Gait variability	Step frequency variance	Worse
		Gait intensity	Step power	Better

Normalized volumes derived from brain MRI scans[a]

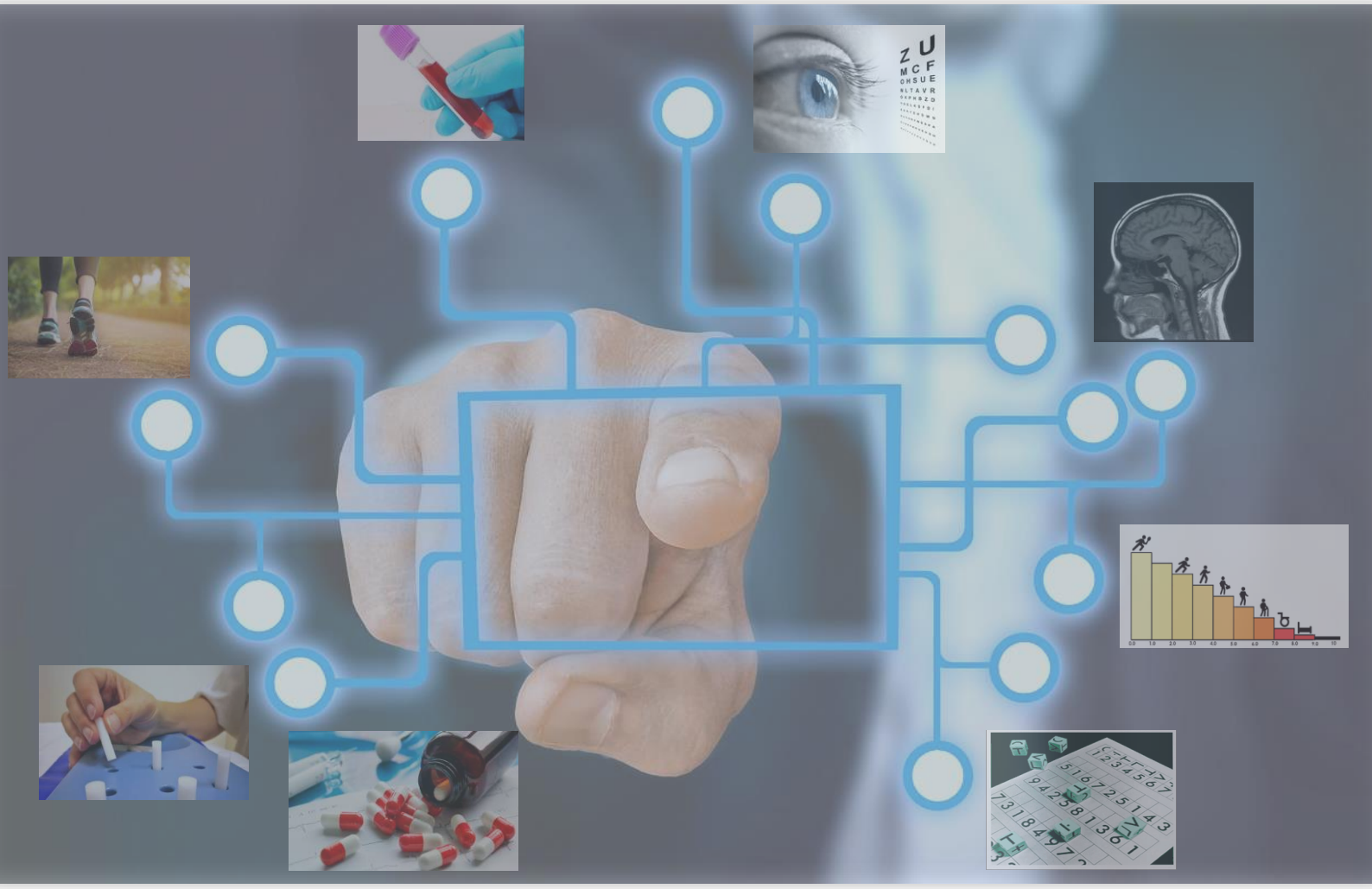
Cerebral cortex gray matter	Deep gray matter	Brainstem and upper cervical cord area	Other
1. Middle frontal	1. Thalamus	**1.** Midbrain	**1.** Lateral ventricles
2. Superior frontal	2. Hippocampus	**2.** Pons	**2.** White matter
3. Lateral and medial orbitofrontal	3. Putamen	**3.** Medulla oblongata	**3.** Corpus callosum area[a]
4. Precentral	4. Globus pallidus	**4.** Upper cervical cord area[a,b]	**4.** Cerebellum - hemispheres
5. Postcentral	5. Caudate nucleus		**5.** Cerebellum - vermis
6. Paracentral	6. Amygdala		
7. Opercularis, triangularis, orbitalis	7. Accumbens		
8. Insula			
9. Precuneus			
10. Cuneus, pericalcarine, lingual			
11. Anterior cingulate			

12. Posterior cingulate
13. Isthmus cingulate
14. Parahippocampus, fusiform, entorhinal
15. Superior parietal
16. Supramarginal
17. Inferior parietal
18. Lateral occipital
19. Inferior and middle temporal
20. Superior and transverse temporal

2MWT, Two-Minute Walk Test; e-SDMT, smartphone-based Symbol Digit Modalities Test; SBT, Static Balance Test; UTT, U-Turn Test.

[a]Volumes were calculated for all regions with the exception of the corpus callosum and the upper cervical cord for which areas were calculated

[b]Upper cervical cord area was determined as mean area along C2 and C3 levels



RESULTS

5. RESULTS

5.1 Baseline Demographics and Characteristics

Participants' baseline demographics for the FAS are described in Table 5.1A. There was an expected greater proportion of females among the pwMS compared with HCs. The majority (69/76, 91%) of pwMS had relapsing-remitting MS (RRMS), with a mild EDSS score (mean 2.4) and presumably "normal" hand/arm function based on an upper limit of normal range defined as the average 9HPT time for HCs plus two standard deviations.

Overall, 92% (70/76) of pwMS and 64% (16/25) of HCs who enrolled in the Floodlight study reached the week 24 visit. Reasons for discontinuation from the study are described in Appendix 10.3.

Table 5.1A. Demographics and clinical characteristics of pwMS and HCs at baseline.

Parameter	**pwMS (n=76)**	**HCs (n=25)**
Age (years), mean (SD)	39.5 (7.9)	34.9 (9.3)
Female, n (%)	53 (70)	7 (28)
MS phenotype, n (%)		
- Primary progressive MS	3 (4)	-[a]
- Secondary progressive MS	4 (5)	-
- Relapsing-remitting MS	69 (91)	-
Time since MS symptom onset (years), mean (SD)	11.3 (7.0)[b]	-
Proportion of pwMS with ≥1 relapse in the past year, n (%)	18 (24)	-
EDSS, mean (SD)	2.4 (1.4)	-
Proportion of pwMS with ≥1 T1 Gd[c]-enhancing lesions, n (%)	2 (3)[d]	-
Total FLAIR[e] lesion volume (mL), mean (SD)	6.3 (7.5)[f]	-
9-Hole Peg Test (seconds), mean (SD)		
- Dominant hand	22.1 (4.6)[g]	18.9 (2.1)
- Non-dominant hand	22.8 (24.9)[h]	19.5 (2.0)
Timed 25-Foot Walk Test (seconds), mean (SD)	6.0 (2.1)[b]	5.0 (1.0)
Symbol Digit Modalities Test (correct responses), mean (SD)	53.8 (11.8)[b]	63.8 (10.0)
Berg Balance Scale, mean (SD)	52.5 (5.7)[i]	56.0 (0)[j]
Patient Determined Disease Steps, mean (SD)	1.5 (1.6)	-
Fatigue Scale for Motor and Cognitive Functions (total score), mean (SD)	59.1 (22.7)[g]	25.5 (6.0)
Patient Health Questionnaire-9, mean (SD)	8.3 (6.1)[k]	2.4 (2.9)[l]
Participants with any previous medications, n (%)	46 (61)	6 (24)
Previous disease modifying treatment[m], n (%)		
- Daclizumab (Zinbryta)	0 (0)	-

- Glatiramer Acetate (Copaxone)	12 (16)	-
- Glatiramer Acetate (Glatopa)	1 (1)	-
- IFN[n] B-1ª IM[o] (Avonex)	4 (5)	-
- IFN B 1ª SC[p] (Rebif)	5 (7)	-
- IFN B 1b SC (Betaseron/Betaferon)	6 (8)	-
- IFN B 1b SC (Extavia)	1 (1)	-
- Pegylated IFN 1ª (Plegridy)	2 (3)	-
- Dimethyl fumarate (Tecfidera)	9 (12)	-
- Fingolimod (Gilenya)	9 (12)	-
- Teriflunomide (Aubagio)	3 (4)	-
- Alemtuzumab (Lemtrada)	2 (3)	-
- Mitoxantrone (Novantrone)	1(1)	-
- Natalizumab (Tysabri)	19 (25)	-
- Other[q]	5 (7)	-

pwMS: people with multiple sclerosis, HCs healthy controls

[a]Not applicable. [b] n=75. [c]Gd: gadolinium. [d] n=68. [e]FLAIR: fluid-attenuated inversion recovery. [f] n=70. [g] n=73. [h] n=74. [i] n=71. [j] n=22. [k] n=60. [l] n=20. [m]Total baseline disease-modifying treatment history. [n]IFN: interferon. [o]IM: intramuscular. [p]SC: subcutaneous. [q]Hidroferol; Radiance study (RPC1063 versus IFN β-1a); Rituximab (Rituxan)

5.2 Adherence

Over a period of 18 months (November 2016-April 2018), more than 6 terabytes of raw data were collected from 76 pwMS and 25 HCs. Participants performed 67,544 active tests, of which 9,787 were the 2MWT, and recorded 200,171 hours of passive monitoring, of which 113,165 hours were captured with the smartwatch. Over 24 weeks, most participants performed 5 to 7 active tests per week, including the 2MWT (Fig. 5.2A). Adherence of pwMS to completing active tests and passive monitoring was good and remained stable over time after week 6 (Fig. 5.2B, C). Even in the last week of the 24-week study, participants completed all active tests on average 4 out of 7 days per week, and recorded at least 4 hours of data via passive monitoring on average 4 out of 7 days per week. The lowest average adherence over 24 weeks was observed for active tests including the 2MWT and the 2MWT only, with participants showing highest

average adherence for passive monitoring (Fig. 5.2D). A total of 70% (16.68/24 weeks) of participants were adherent to all active tests, 75% (17.95/24 weeks) to all active tests except 2MWT, 71% (17.13/24 weeks) to 2MWT, 79% (18.89/24 weeks) to smartphone- or smartwatch-based passive monitoring, 66% (15.74/24 weeks) to smartphone-based passive monitoring, and 74% (17.69/24 weeks) to smartwatch-based passive monitoring.

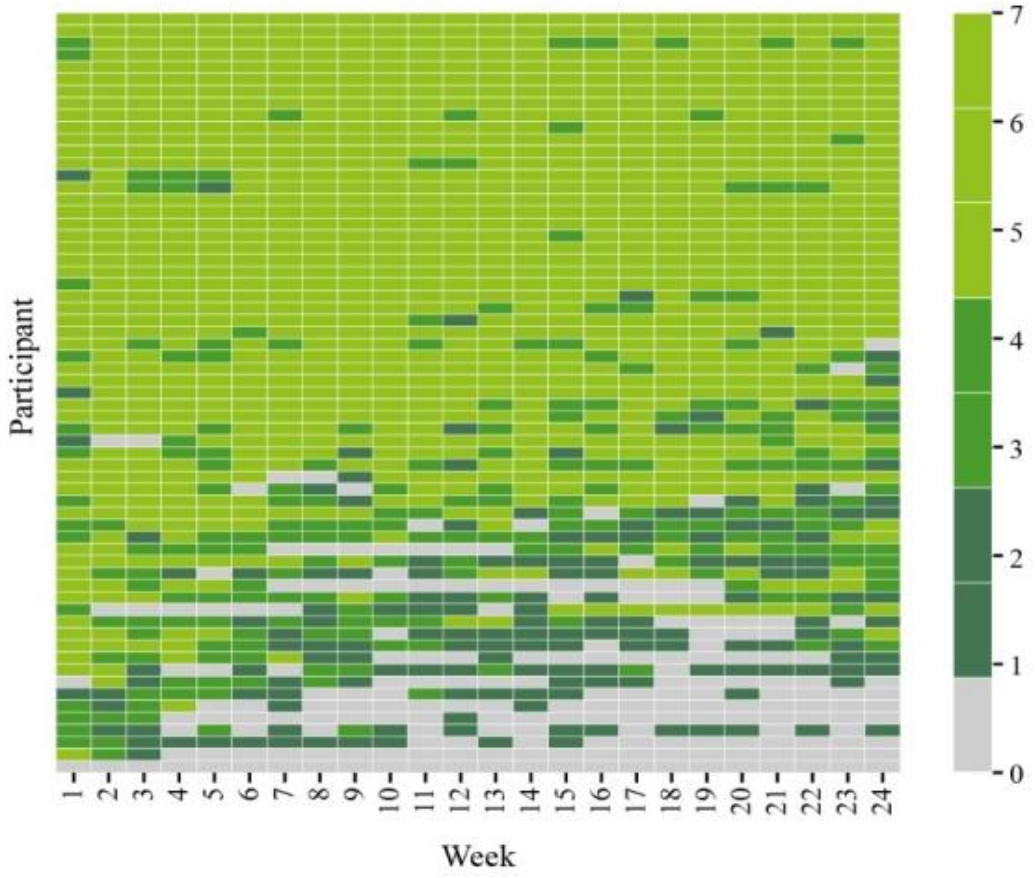


Figure 5.2A. Adherence of pwMS to active tests for individual participants: number of performed active tests per week [level of activity (light green: high; dark green/grey: low] over individual study weeks [columns]).

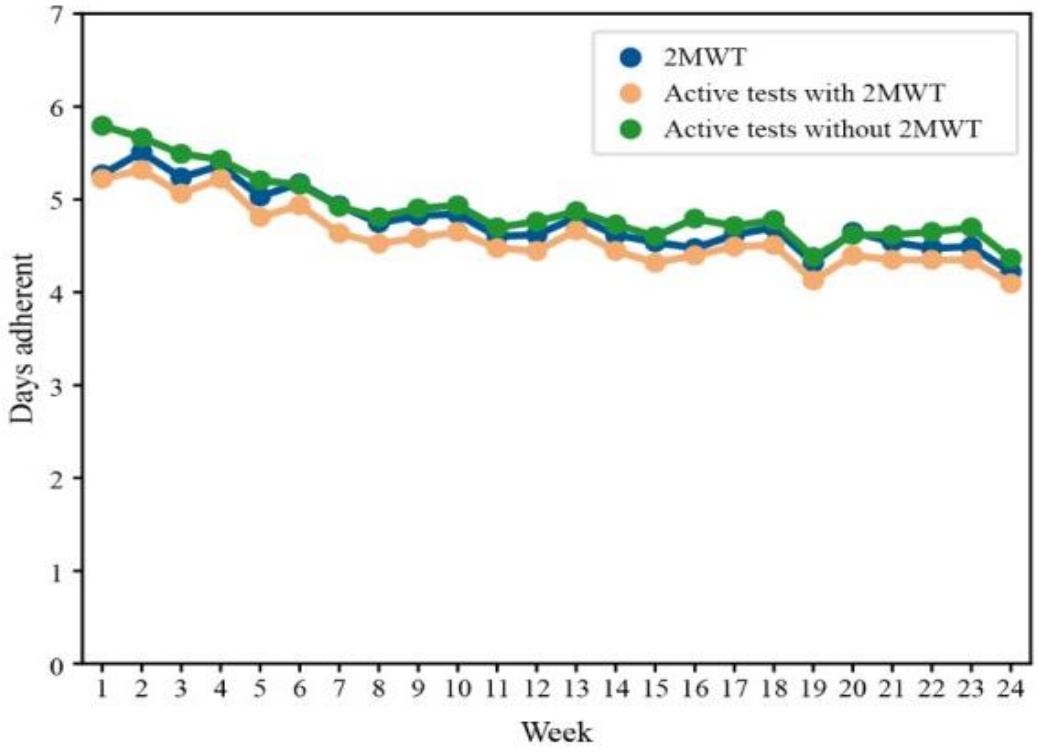


Figure 5.2B. Adherence of pwMS to active tests. 2MWT: Two-Minute Walk Test.

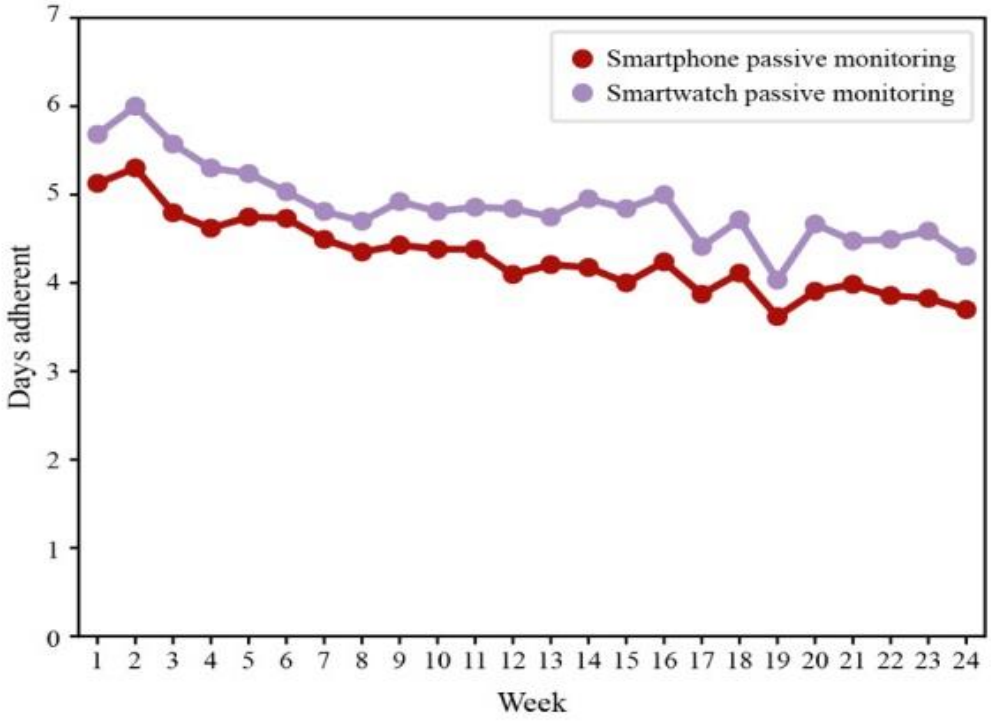


Figure 5.2C. Adherence of pwMS to smartphone and smartwatch passive monitoring. Days with more than 4 hours of passive monitoring on a device are considered as adherent.

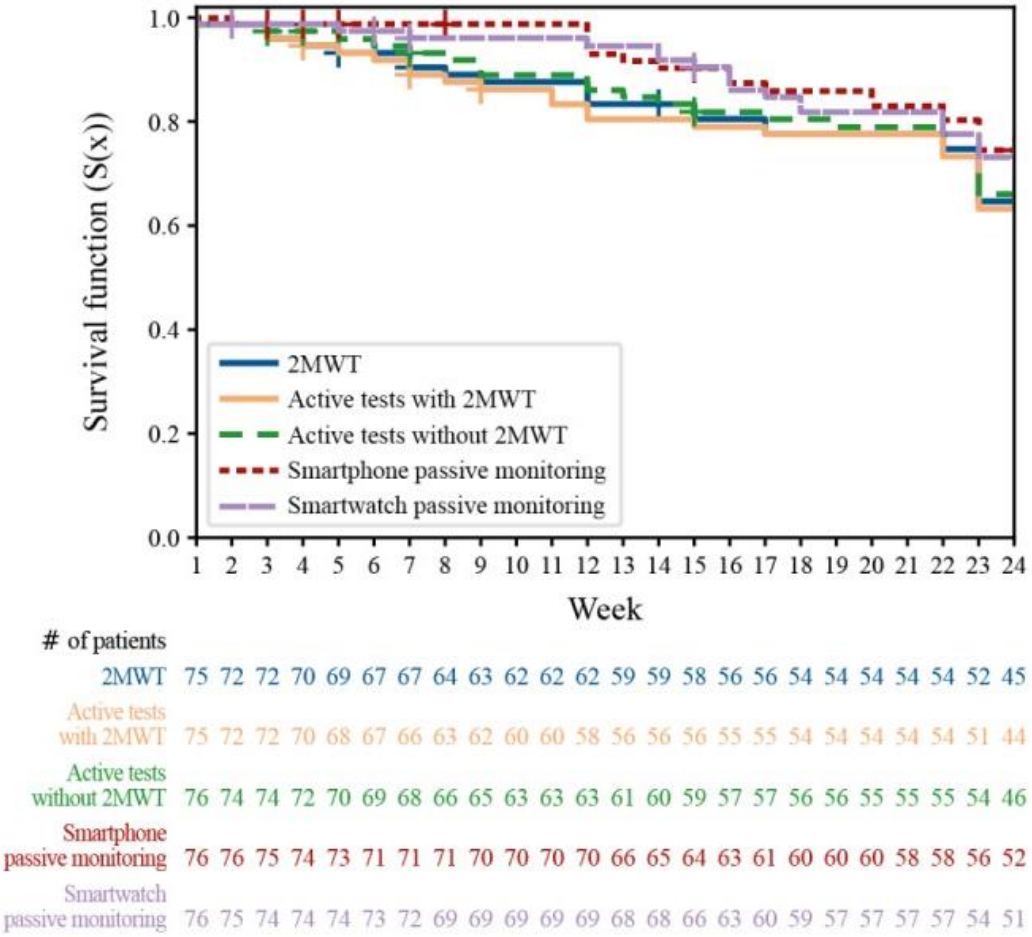


Figure 5.2D. Adherence of pwMS to active tests and passive monitoring. The results of the time-to-event survival analysis based on the Kaplan–Meier method along the Floodlight study. The abandoning event was defined as the last week in which the participant was adherent according to the definitions for active tests and passive monitoring. Active tests performed on days of in-clinic visits were not considered in the adherence calculation, to focus the abandoning analysis on the remote use. Participants leaving the study before the terminal visit were considered as censored. 2MWT: Two-Minute Walk Test.

Correlation was explored between adherence measures and pwMS population characteristics. Only disease duration showed significant small negative correlation with measures of adherence (Spearman rank correlations; 2MWT adherence: –0.42, p<.001; smartphone passive monitoring adherence: –0.29, p=.02; all active tests except 2MWT adherence: –0.37, p=.003; and smartwatch passive monitoring adherence: –0.27, p=.04), indicating that disease severity and demographics did not appear to play a significant role in adherence.

5.3 Patient Satisfaction

The average overall satisfaction score among pwMS who completed the study at week 12 (n=64) was 74.1 out of a possible 100 and remained stable at week 24 (study termination/early discontinuation visit [n=68]) with 73.7 out of 100 (Wilcoxon signed-rank test p=.71). There was one significant association between overall satisfaction score and gender (p=.04). Individual questions from the satisfaction questionnaire (Appendix 10.2) were analyzed for their association with pwMS population characteristics, described in Appendix 10.1.

Implications of the use of the Floodlight MS app in pwMS were assessed from individual questions from the patient satisfaction questionnaire. When asked to rate the impact of the smartphone, smartwatch, and active tests on daily living, more than 80% (61/72) of pwMS perceived the Floodlight MS app to have at least an acceptable impact on daily activities (Fig. 5.3A). Nearly 50% (32/71) of participants had no issue with any of the active tests, and only one-third would prefer to avoid the 2MWT, most likely due to increased burden from execution—for example, having to find a place to perform the test or not wanting to go outside in bad weather (Fig. 5.3B). Without providing any data feedback to the pwMS throughout the study, more than 60% (46/72) of participants would have liked to continue using Floodlight MS app "to understand my MS better and improve my disease management" (Figure 5.3C). Approximately 90% (65/72) of pwMS indicated their interest to see the results of the tests, which will be addressed in later Roche-sponsored studies using Floodlight MS app (CONSONANCE [NCT03523858] and Floodlight Open [floodlightopen.com]; Fig. 5.3D). Analysis of patient responses to the satisfaction questionnaire is described in Appendix 10.1.

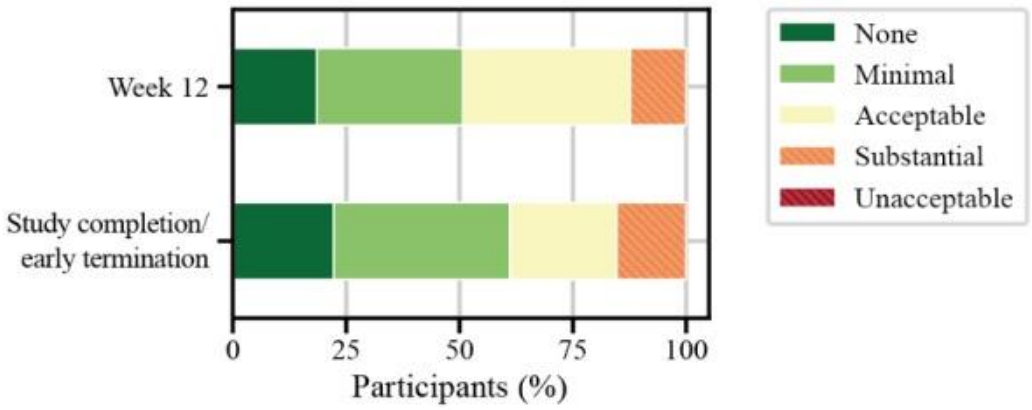


Figure 5.3A. Implications of Floodlight MS app in people with multiple sclerosis for "impact on daily activities" from the patient satisfaction questionnaire.

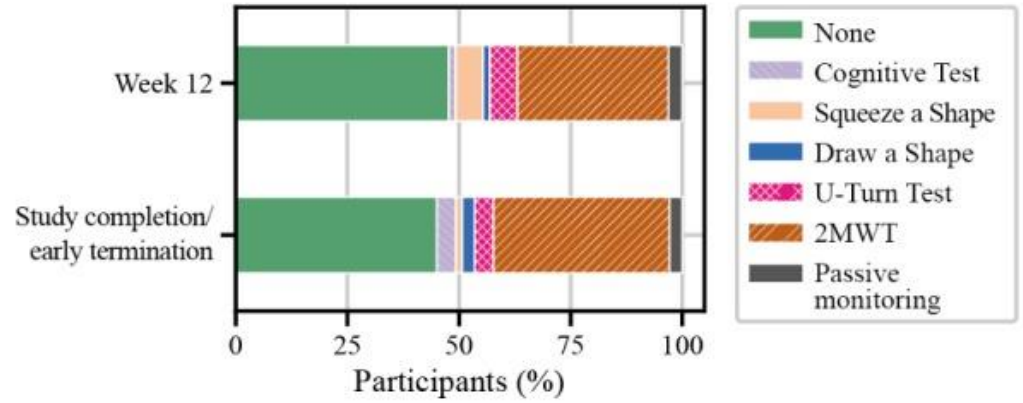


Figure 5.3B. Implications of Floodlight MS app in people with multiple sclerosis for "avoiding one component of FLOODLIGHT" from the patient satisfaction questionnaire. 2MWT: Two-Minute Walk Test.

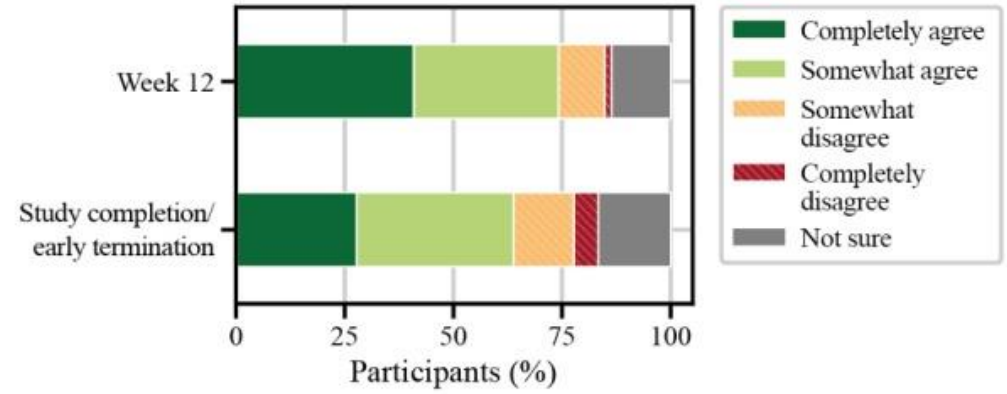


Figure 5.3C. Implications of Floodlight MS app in people with multiple sclerosis for "desire to continue using the Floodlight MS app" from the patient satisfaction questionnaire.

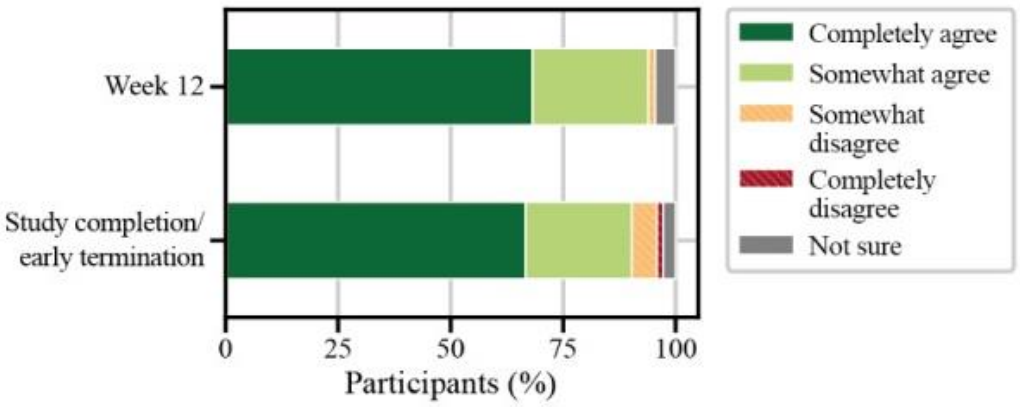


Figure 5.3D. Implications of Floodlight MS app in people with multiple sclerosis for "prefer to see results immediately to monitor" from the patient satisfaction questionnaire.

5.4 Test-retest reliability

Test-retest reliability was assessed in pwMS and HCs with valid assessments in all consecutive 2-week windows (range: 32-46 pwMS and 8-11 HCs). In pwMS, ICCs were moderate or good (Table 5.4A), suggesting that reliable data can be captured with the Floodlight MS app. In HCs, where the group sizes were lower, ICCs were mostly poor to good.

Table 5.4A. Test-retest reliability in pwMS and age- and sex-adjusted Spearman´s rank correlation analysis of the Floodlight app in pwMS.

Domain	Test		Feature	Test–retest reliability ICC (2,1) (95% CI)		Spearman's rank correlations for PwMS				
				HC	PwMS	Domain-specific standard clinical measure	EDSS	MSIS-29 subscale/ items[a]	T2 FLAIR lesion volume (mL)	Normalized brain volume (mL)
Cognition		e-SDMT	Number of correct responses (*n*)	0.55 (0.34–0.80)[b]	0.85 (0.76–0.91)[c]	Oral SDMT	0.82***	−0.43***	−0.52***	−0.42***	0.54***
Upper extremity function		Pinching Test	Double touch asynchrony (s)	0.72 (0.53–0.90)[d]	0.71 (0.62–0.80)[c]	9HPT	0.64***	0.30*	0.35**	0.17	−0.26*
		Pinching Test	Number of successful pinches (*n*)	0.81 (0.65–0.94)[d]	0.72 (0.61–0.82)[c]	9HPT	−0.52***	−0.26*	−0.33**	−0.12	0.32**
		Draw a Shape Test	Overall mean trace accuracy	0.53 (0.32–0.79)[b]	0.85 (0.79–0.90)[e]	9HPT	−0.48***	−0.40***	−0.40***	−0.26*	0.33**
		Draw a Shape Test	Overall mean trace celerity (s^{-1})	0.45 (0.25–0.73)[b]	0.81 (0.73–0.87)[e]	9HPT	−0.40***	−0.08	0.03	−0.26*	0.24*
Gait and balance		SBT	Sway path (m/s^2)	0.40 (0.20–0.73)[f]	0.71 (0.61–0.80)[g]	BBS	−0.20	0.24*	0.31**	0.21	−0.05
		UTT	Turn speed (rad/s)	0.45 (0.24–0.75)[d]	0.83 (0.76–0.89)[h]	T25FW	−0.52***	−0.45***	−0.39***	−0.13	0.27*
		Walk Test	Step power	0.85 (0.70–0.95)[f]	0.78 (0.70–0.86)[i]	T25FW	−0.31**	−0.28*	−0.24*	0.04	−0.02
		Passive Monitoring	Turn speed (rad/s)	0.66 (0.42–0.89)[j]	0.72 (0.61–0.82)[k]	T25FW	−0.25*	−0.27*	−0.12	−0.11	0.14
		Passive Monitoring	Step power	0.63 (0.39–0.88)[j]	0.61 (0.48–0.74)[k]	T25FW	−0.33**	−0.19	−0.22	0.09	0.00

PwMS: people with multiple sclerosis; HCs: healthy controls; PoC: Proof of Concept; ICC: intraclass correlation coefficient; CI: confidence interval; EDSS: Expanded Disability Status Scale; MSIS-29: 29-item Multiple Sclerosis Impact Scale; FLAIR: fluid-attenuated inversion recovery; e-SDMT: electronic Symbol Digit Modalities Test; 9HPT: Nine-Hole Peg Test; SBT: Static Balance Test; BBS: Berg Balance Scale; UTT, U-Turn Test; T25FW Timed 25-Foot Walk. Colored background indicates significant correlation in expected direction. aThe e-SDMT was correlated against the psychological subscale, the Pinching and Draw a Shape Tests against the arm-related items (items 2, 6 and 15), and all other tests against the physical subscale. bn = 11. cn = 46. dn = 10. en = 44. f n = 9. gn = 42. hn = 41. i n = 39. J n = 8. kn =32. *p < 0.05; **p < 0.01; ***p < 0.001.

5.5 Correlation of Floodlight MS app with clinical and MRI measures

The age- and sex-adjusted Spearman's rank correlation analysis in pwMS is summarized in Fig. 5.5A. All statistically significant correlations were in the expected direction. Thus, increasing levels of MS-related disability were associated with worse performance on the Floodlight MS app. Overall, strongest correlations of test features were observed with the respective domain-specific standard clinical disability measures. These correlations were good-to-excellent in the cognitive domain ($r = 0.82$) and fair or moderate-to-good in the upper extremity function domain ($|r| = 0.40–0.64$) and gait and balance domain ($r = −0.25$ to $−0.52$, all $p<0.05$). Only the SBT did not correlate with its domain-specific standard clinical measure, the BBS ($r = −0.20$, $p>0.05$). Most test features also correlated with EDSS (all $p<0.05$ except for Draw a Shape Test overall mean trace celerity and Passive Monitoring step power) and their respective MSIS-29 subscale or items (all $p<0.05$ except for Draw a Shape Test overall mean trace celerity, Passive Monitoring turn speed, and Passive Monitoring step power). Normalized brain volume correlated significantly with test features across all domains with the strongest association found with e-SDMT ($r = 0.54$, $p<0.001$). Similar results were obtained with unadjusted measures.

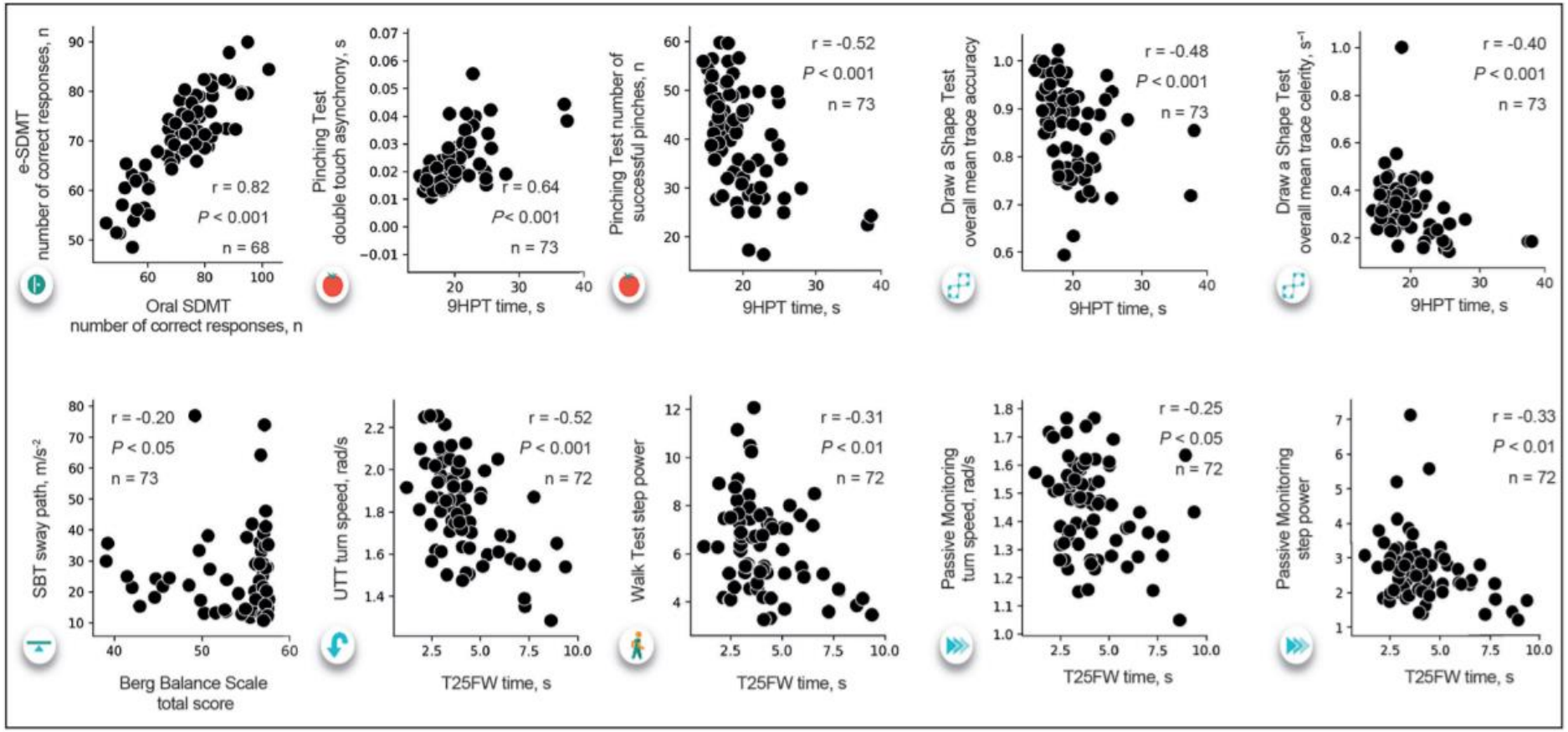


Figure 5.5A. Age- and sex-adjusted Spearman's rank correlations between active tests and passive monitoring (vertical axis) and their respective domain-specific standard clinical measures (horizontal axis). e-SDMT: electronic Symbol Digit Modalities Test; SDMT: Symbol Digit Modalities Test; 9HPT: Nine-Hole Peg Test; SBT: Static Balance Test; UTT: U-Turn Test; T25FW: Timed 25-Foot Walk.

Next, we assessed correlations between active gait tests and passive monitoring. UTT turn speed showed moderate-to-good correlation with Passive Monitoring turn speed ($r = 0.43$, $p<0.001$). Stronger, good-to-excellent positive correlations were observed between Walk Test step power and Passive Monitoring step power ($r = 0.76$, $p<0.001$; Fig. 5.5B). The partial correlation analysis revealed that both the Pinching Test (double touch asynchrony: partial $r = 0.37$, $p<0.001$) and Draw a Shape Test (overall mean trace accuracy: partial $r = -0.40$, $p<0.001$; overall mean trace celerity: partial $r = -0.30$, $p<0.01$) contain independent information in predicting 9HPT time. Similarly, the UTT carries unique information in predicting T25FW time when correcting for the other gait features (partial $r = -0.31$, $p<0.01$). Comparing the performance of the e-SDMT and oral SDMT in pwMS and HCs showed that both HCs and pwMS achieved on average 4.94 and 3.81 fewer correct responses on the e-SDMT than on the oral SDMT, respectively (paired t-test: $p<0.05$ and $p<0.001$, respectively).

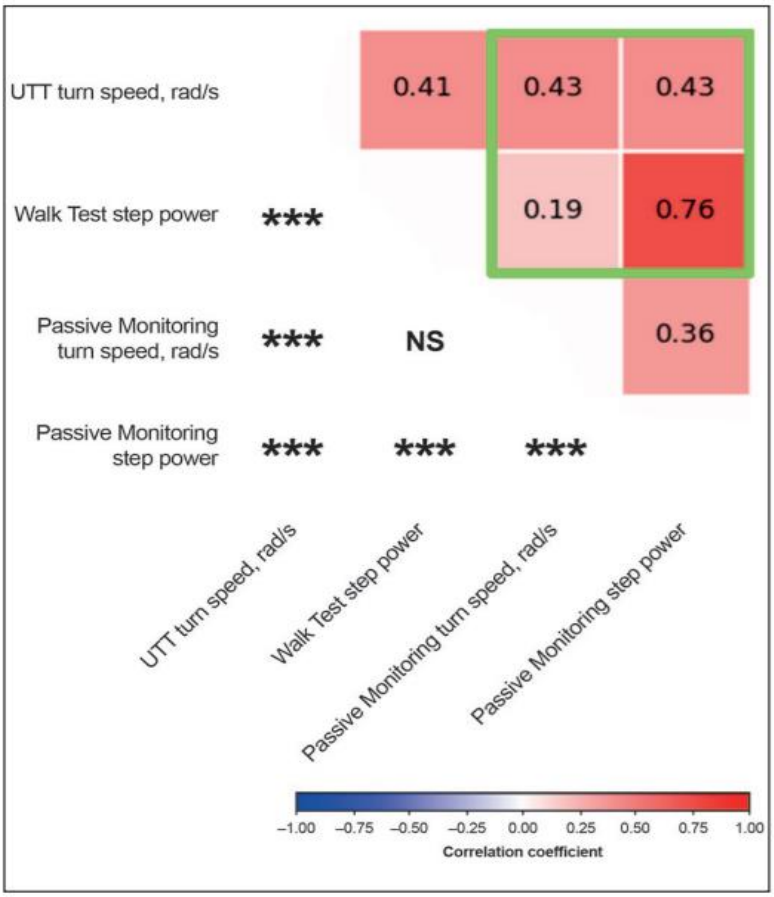


Figure 5.5B. Age- and sex-adjusted Spearman's rank correlations between passive monitoring and active gait features. UTT: U-Turn Test; NS: not significant. ***p < 0.001.

5.6 Correlations of the digital measures and the standard in-clinic assessments with regional, structural MRI outcomes

The study enrolled 76 pwMS, of which 62 (82%) pwMS were included in the analyses. Of the 14 excluded pwMS, 12 were excluded for poor adherence to the active tests and two for missing MRI scans. Of those pwMS included, 68% were female, mean age was 39.7 years (SD: 7.5), mean EDSS at baseline was 2.5 (SD: 1.4; range: 0.0-5.5), and 89% were diagnosed with RRMS.

Correlations of global and regional MRI atrophy patterns with the digital measures and the standard clinical measures are summarized in Fig. 5.6A. Brain maps highlighting the statistically significant correlations after FDR correction are provided in Fig. 5.6B. Higher EDSS scores correlated with smaller volumes of the cerebral white matter (r=− 0.42, q<0.05), lateral occipital lobe (r= −0.43, q<0.05), insula (r= −0.37, q<0.05), putamen (r= −0.36, q<0.05), and globus pallidus (r= −0.38, q<0.05).

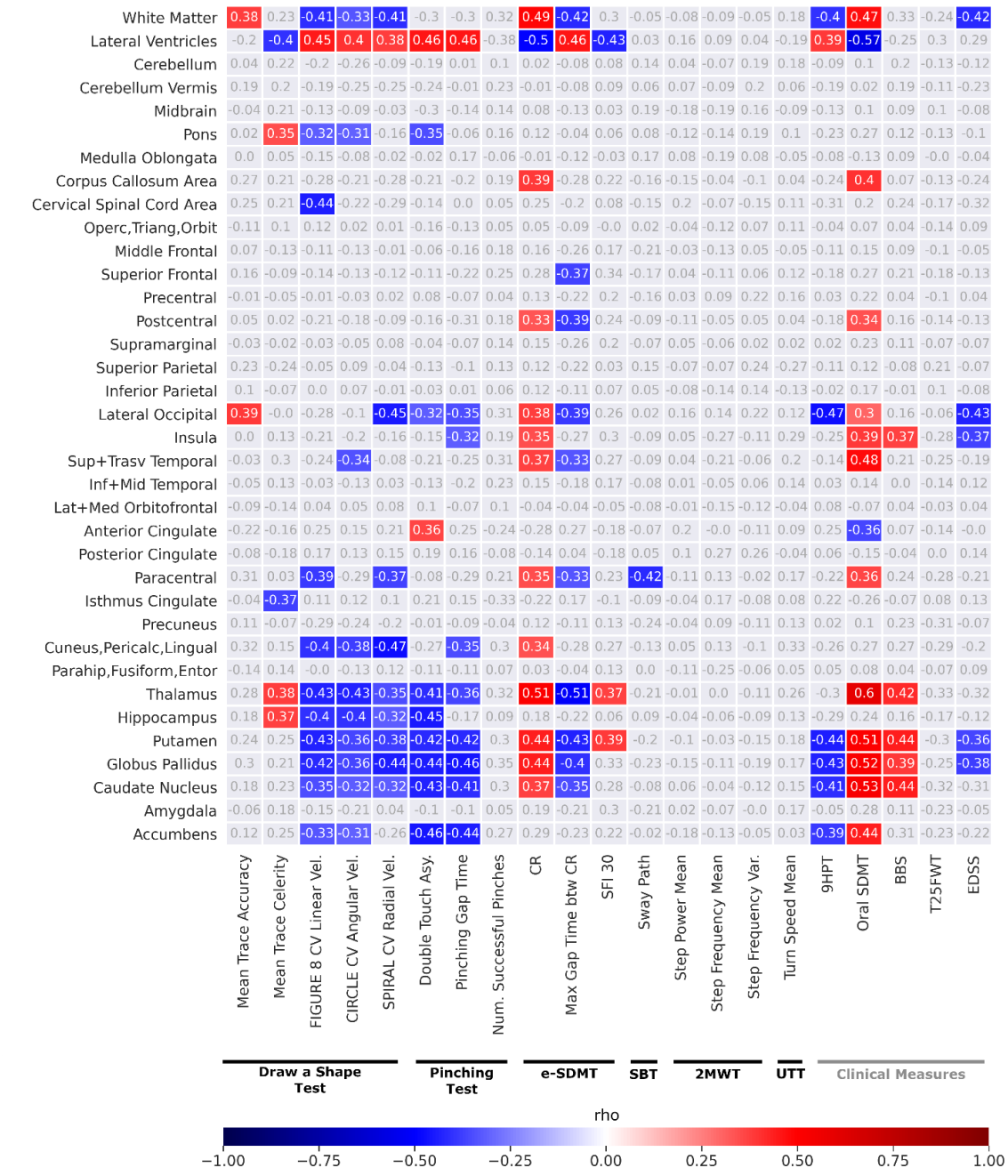


Figure 5.6A. Spearman's rank correlation analysis of digital measures and standard in-clinic assessments with global and regional MRI outcomes. Statistically significant positive and negative correlations are highlighted in red and blue, respectively. Higher values equate to better performance on the oral SDMT and BBS, as well as for digital measures assessing trace accuracy and trace celerity on the Draw a Shape Test; number of pinches on the Pinching Test; number of correct responses and SFI 30 on the e-SDMT; mean step power and mean step frequency on the 2MWT; and mean turn speed on the UTT. In contrast, higher values equate to worse performance on the EDSS, 9HPT and T25FW, as well as for digital measures assessing CV linear, angular and radial velocity on the Draw a Shape Test; double touch asynchrony and pinching gap time on the Pinching Test; max gap time between correct responses on the e-SDMT; sway path on the SBT; and step frequency variance on the 2MWT. 2MWT, Two-Minute Walk Test; 9HPT, Nine-Hole Peg Test; Asy., asynchrony; BBS, Berg

Balance Scale; CR, correct responses; CV, coefficient of variation; EDSS, Expanded Disability Status Scale; Entor, entorhinal cortex; e-SDMT, smartphone-based Symbol Digit Modalities Test; Lat, lateral; Med, medial; Mid, middle; Operc, opercularis; Orbit, orbitalis; Parahip, parahippocampus; SBT; Static Balance Test; SDMT, Symbol Digit Modalities Test; SFI, speed fatigability index; Sup, superior; T25FW, Timed 25-Foot Walk; Trasv, transverse; Triang, triangularis; UTT, U-Turn Test; Vel., drawing velocity.

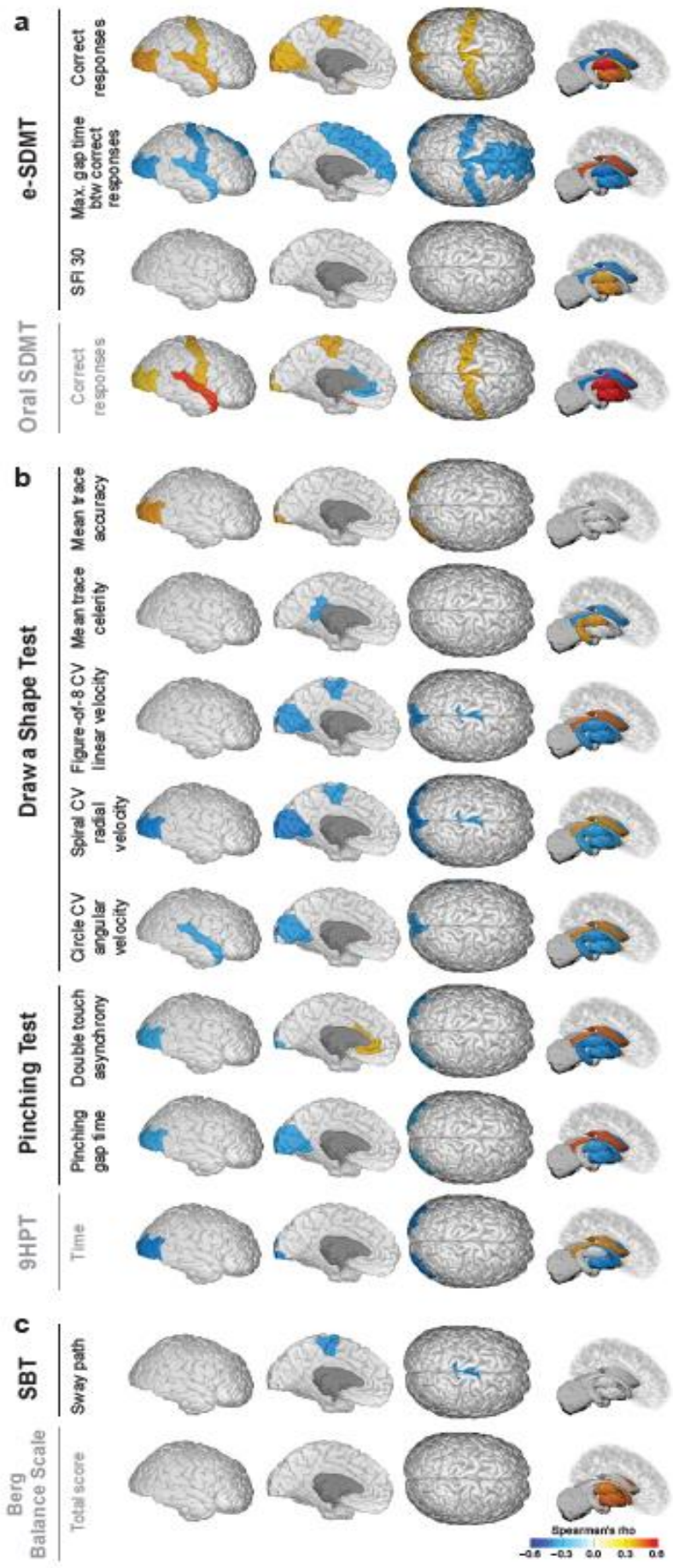


Figure 5.6B. Statistically significant Spearman's rank correlations after FDR correction and adjustment for age, sex, and BMI
Spearman's rank correlations between regional brain volume measured by MRI and measures of **a** the cognitive domain, **b** upper extremity function domain, and **c** balance domain are shown for four planes (from left to right: cortical outer, inner, top, and subcortical). Higher values equate to better performance for digital measures assessing number of correct responses and SFI 30 on the e-SDMT; trace accuracy and trace celerity on the Draw a Shape Test; and double touch asynchrony on the Pinching Test. In contrast, higher values equate to worse performance for digital measures assessing max. gap time between correct responses on the e-SDMT; CV linear, angular, and radial velocity on the Draw a Shape Test; pinching gap time on the Pinching Test; and sway path on the SBT.

9HPT Nine-Hole Peg Test, *BMI* body mass index, *btw* between, *CV* coefficient of variation, *e-SDMT* smartphone-based electronic Symbol Digit Modalities Test, *FDR* false discovery rate, *SBT* Static Balance Test, *SDMT* Symbol Digit Modalities Test, *SFI* speed fatigability index

The e-SDMT and oral SDMT showed a similar correlation pattern with the regional structural MRI outcomes (Fig. 5.6A and Fig. 5.6Ba). On both, a lower number of correct responses was significantly associated with smaller volumes and areas of the cerebral white matter (r=0.49 [e-SDMT]/r= 0.47 [oral SDMT]), corpus callosum (r= 0.39/0.40), postcentral gyrus (r= 0.33/0.34), lateral occipital lobe (r= 0.38/0.30), insula (r= 0.35/0.39), superior and transverse temporal gyrus (r= 0.37/0.48), paracentral gyrus (r= 0.35/0.36), and deep gray matter structures such as the thalamus, putamen, globus pallidus, and caudate nucleus (r= 0.37–0.51/0.51–0.60; all q<0.05). A lower number of correct responses on either test was also associated with a larger volume of the lateral ventricles (r= −0.50/−0.57; both q<0.05). Two correlations were only observed on the e-SDMT, but not on the oral SDMT. This included the correlation between the e-SDMT number of correct responses and the volume of the cuneus, pericalcarine cortex, lingual gyrus (r= 0.34; q<0.05), and between the e-SDMT maximum gap duration between correct responses and the volume of the superior frontal lobe (r= −0.37; q<0.05).

Likewise, for each regional MRI outcome that showed significant correlation with the 9HPT there was at least one Draw a Shape Test measure that was significantly correlated with the same MRI outcome (Fig. 5.6A and Fig. 5.6Bb). Worse performance on either test was associated with smaller volumes of the cerebral white matter ($|r|$ = 0.33–0.41 for all measures with q<0.05); visual areas (lateral occipital lobe [r= −0.32 to −0.47 for all measures with q<0.05]); and deep gray nuclei such as the putamen, globus pallidus, caudate nucleus, and accumbens ($|r|$ =0.31–0.44 for all measures with q<0.05); as well as a larger volume of the lateral ventricles ($|r|$ = 0.38–0.45 for all measures with q<0.05). However, some differences were noted across the individual Draw a Shape Test measures. Only trace celerity and velocity-based measures of round shapes correlated with deep gray matter volume ($|r|$ = 0.31–0.44 for all measures with q<0.05), including the thalamus ($|r|$ = 0.35–0.43 for all measures with q<0.05) and hippocampus ($|r|$ = 0.32–0.40 for all measures with q<0.05). Of note, no association was observed between 9HPT and either the thalamus or hippocampus

(both $q \geq 0.05$). Other correlations observed only with the Draw a Shape Test, but not the 9HPT, include those between the velocity-based measures of round shapes and the volume of the cuneus, pericalcarine cortex, lingual gyrus ($r = -0.38$ to -0.47 for all measures with $q < 0.05$); and between variability of linear drawing velocity on the figure-of-8 and the cervical spinal cord area ($r = -0.44$; $q < 0.05$). Additionally, better performance on overall mean trace celerity, but not the 9HPT, correlated with a smaller volume of the isthmus cingulate ($r = -0.37$; $q < 0.05$).

Similarly, for each correlation observed with the 9HPT, at least one Pinching Test measure was significantly associated with the same MRI outcome, except for the cerebral white matter (Fig. 5.6A and Fig. 5.6Bb). Some correlations, however, were observed only with the Pinching Test. This includes correlations between longer double touch asynchrony and smaller volumes of the pons ($r = -0.35$; $q < 0.05$), thalamus ($r = -0.41$; $q < 0.05$), hippocampus ($r = -0.45$; $q < 0.05$), and interestingly with a larger volume of the anterior cingulate cortex ($r = 0.36$, $q < 0.05$). Additionally, longer gap duration between pinches was also associated with a smaller volume of the thalamus ($r = -0.36$; $q < 0.05$), as well as being specifically associated with a smaller volume of the insula ($r = -0.32$; $q < 0.05$), and the cuneus, pericalcarine cortex, and lingual gyrus ($r = -0.35$; $q < 0.05$).

On the SBT, larger sway path correlated with a smaller volume of the paracentral lobule ($r = -0.42$, $q < 0.05$; Fig. 5.6A and Fig. 5.6Bc). In contrast, a lower total score on the BBS (worse ability to balance) correlated mostly with smaller volumes of deep gray matter structures ($r = 0.39-0.44$ for correlations with $q < 0.05$; Fig. 5.6A and Fig. 5.6Bc).

In contrast, no correlations with any of the regional MRI outcomes was found for any of the gait assessments, neither for the digital measures derived from the 2MWT or UTT nor for the T25FW (Fig. 5.6A; all $q \geq 0.05$).

Next, the variance observed in the digital measures that can be explained by structural MRI outcomes was estimated using a Bayesian ridge regression model. Up to a third of the variance observed in the digital measures (R^2≤34%) could be explained when using either normalized total brain volume as individual predictor ("Whole Brain" model) or volumetric data from the 36 individual MRI regions as multiple predictors ("Parcellation" model). In the upper extremity function domain, however, comparable or higher R^2 values were obtained with the Parcellation model compared with the Whole Brain model. This suggests that using multiple predictors, thus multiple brain regions, can explain as much or more of the observed variance in the digital measures (Fig. 5.6C). The increase in R^2 was most noticeable for mean trace celerity (R^2=0.23 vs. – 0.01), spiral coefficient of variation (CV) linear drawing velocity (R^2=0.22 vs. 0.02), figure-of-8 CV linear drawing velocity (R^2=0.19 vs. 0.12), and double touch asynchrony (R^2=0.16 vs. 0.06, respectively, for the Parcellation vs. Whole Brain model). This was also reflected by a less uniform distribution of the individual Bayesian ridge regression coefficients across the 36 MRI regions (Fig. 5.6D). By comparison, the R^2 score was more comparable across the two models for the standard clinical measure 9HPT (R^2=0.28 vs. 0.25) (Fig. 5.6C).

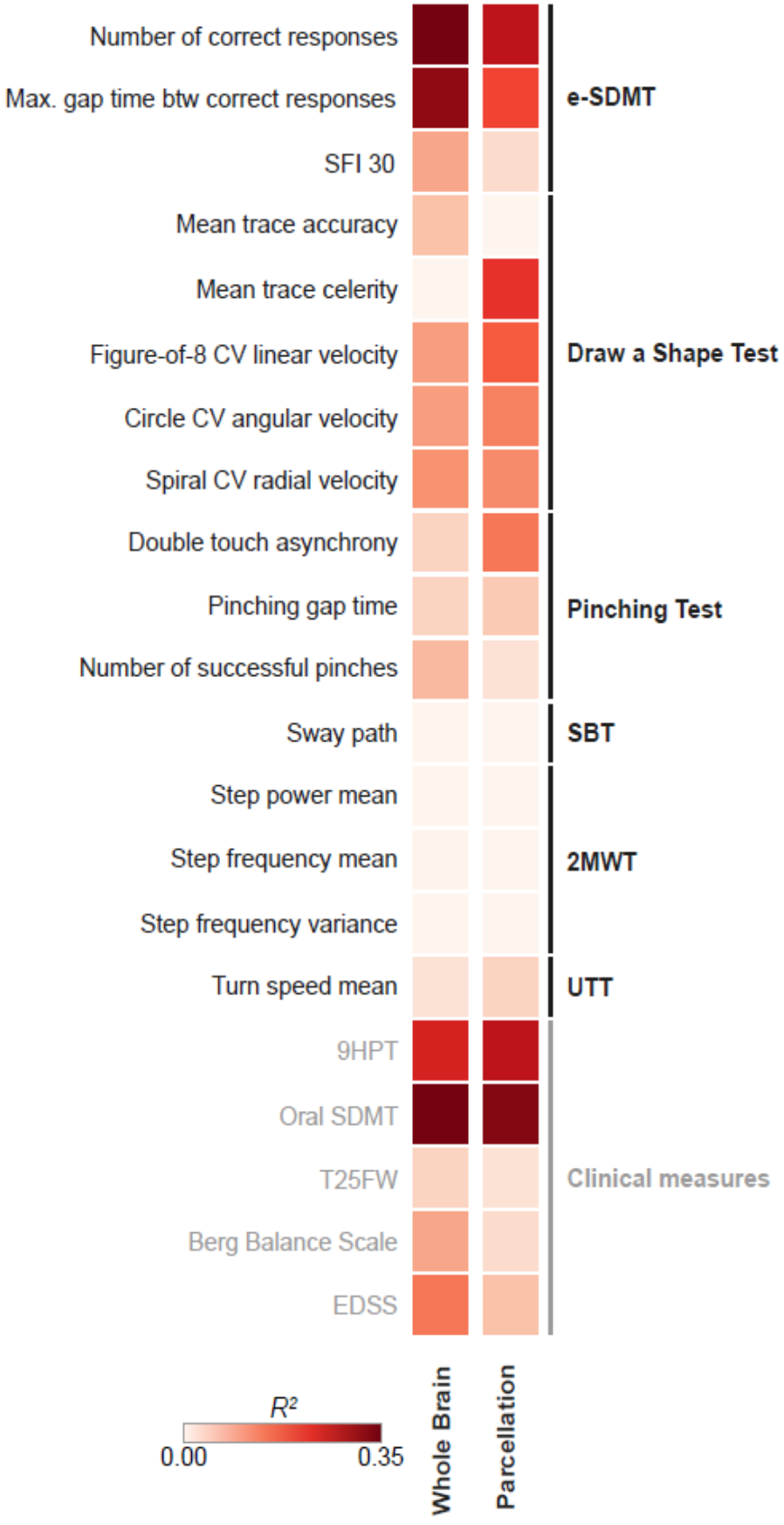


Figure 5.6C. Bayesian ridge regression model with leave-one-out cross validation for estimating the variance (R^2) in the digital measures and standard in-clinic assessments that can be explained by volumetric MRI
Two models were applied. The first model ("Whole Brain") included normalized brain volume and the three demographic variables: age, sex, and body mass index (top row). The second model ("Parcellation") included all 36 regional MRI regions and the same three demographic variables (bottom row). Compared with the standard in-clinic assessments, the digital measures tend to have a larger R^2 score in the Parcellation model vs. in the Whole Brain model, which may reflect higher functional specificity. This is particularly evident on the Draw a Shape Test (mean trace celerity and spiral CV radial velocity) and on the Pinching Test (double touch asynchrony).

2MWT Two-Minute Walk Test, *9HPT* Nine-Hole Peg Test, *btw* between, *CV* coefficient of variation, *EDSS* Expanded Disability Status Scale, *e-SDMT* smartphone-based electronic Symbol Digit Modalities Test, *SBT* Static Balance Test, *SDMT* Symbol Digit Modalities Test, *SFI* speed fatigability index, *T25FW* Timed 25-Foot Walk, *UTT* U-Turn Test

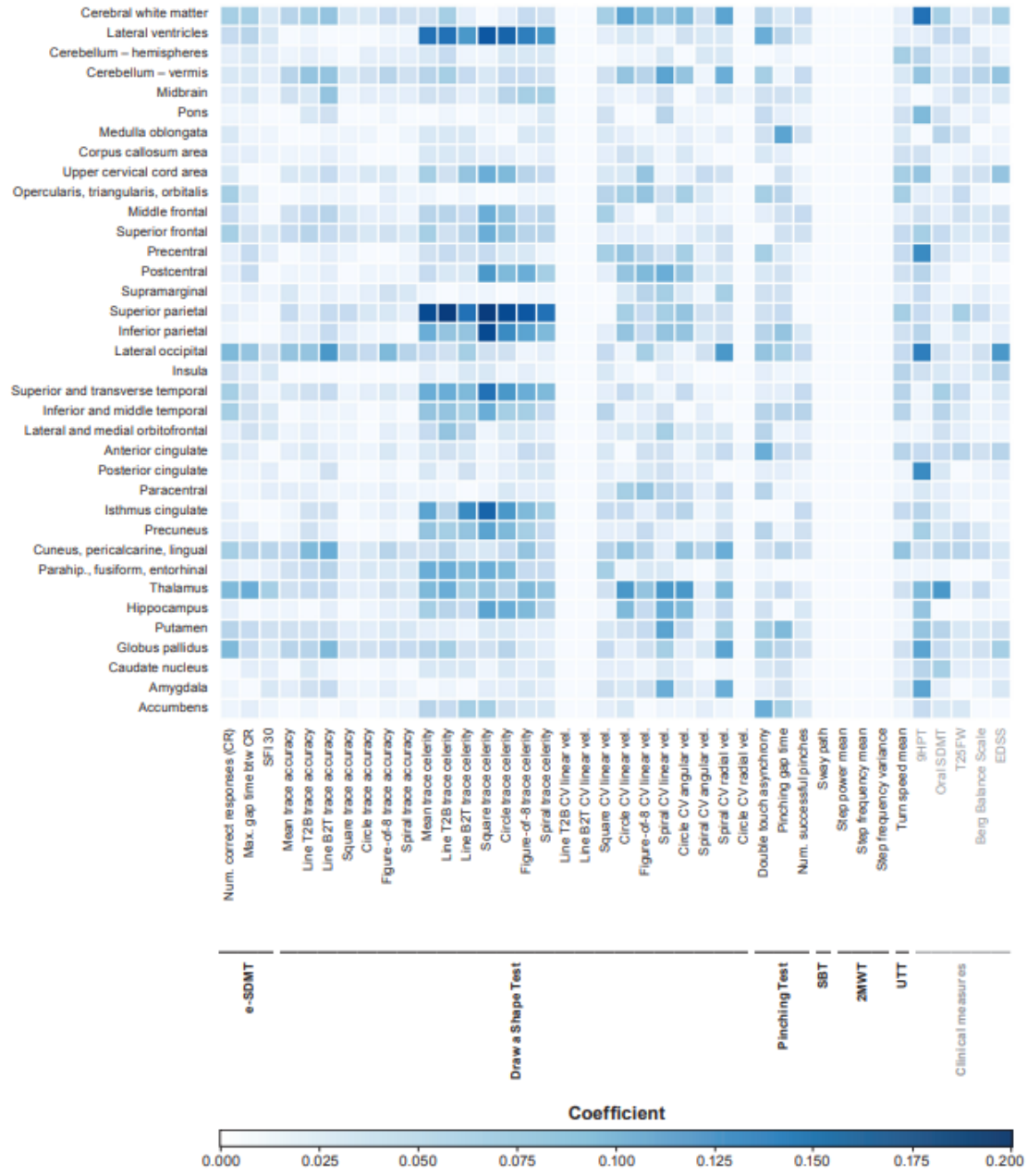
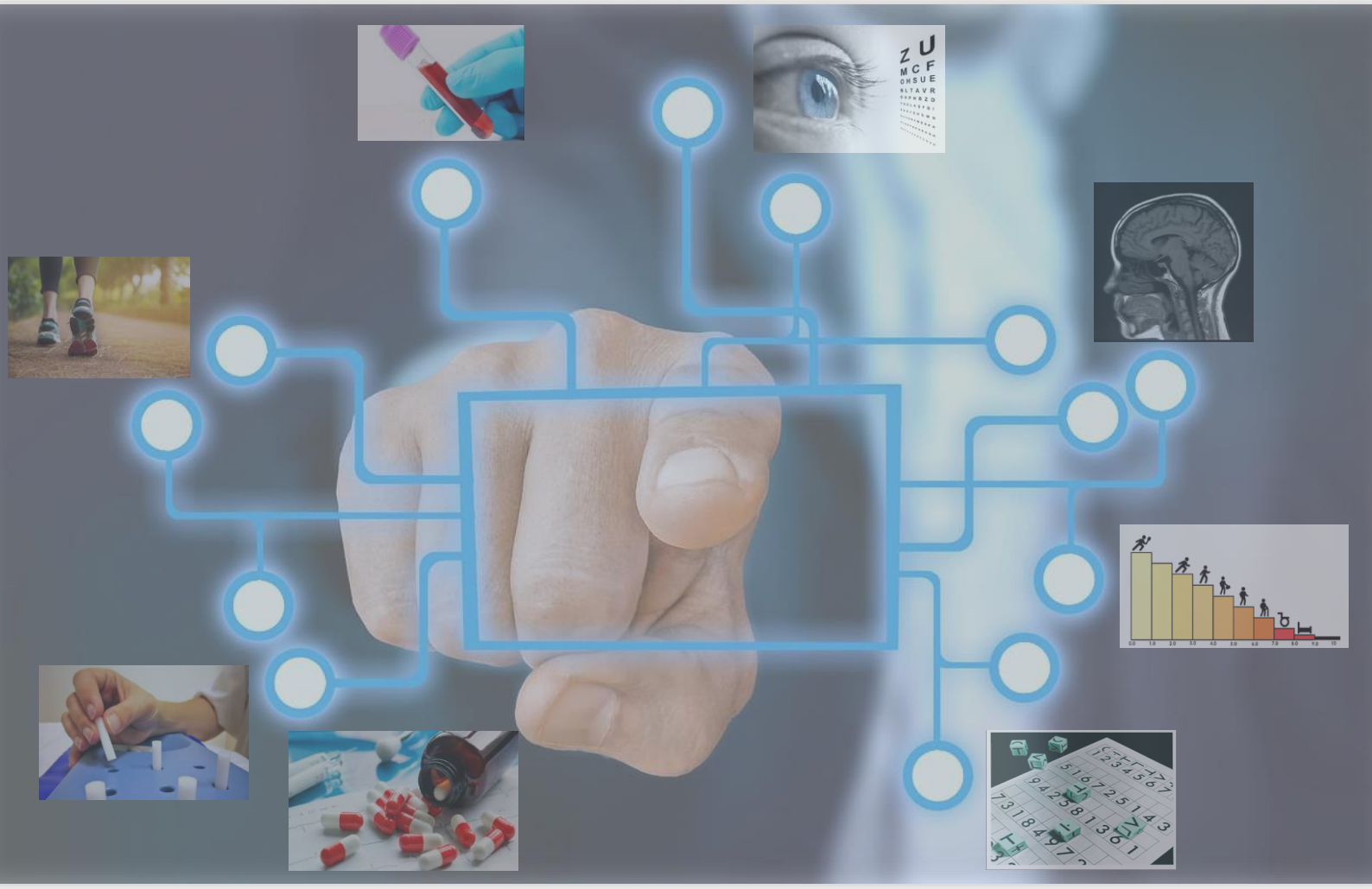


Figure 5.6D. Bayesian ridge regression coefficients across the 36 regional MRI outcomes for digital measures and standard in-clinic assessments Higher values mean better performance on the oral SDMT and Berg Balance Scale, as well as for digital measures assessing number of correct responses and SFI 30 on the e-SDMT; trace accuracy and trace celerity on the Draw a Shape Test; number of pinches on the Pinching Test; mean step power and mean step frequency on the 2MWT; and mean turn speed on the UTT. In contrast, higher values equate to worse performance on the EDSS, 9HPT, and T25FW, as well as for digital measures assessing max. gap time between correct responses on the e-SDMT; CV linear, angular, and radial velocity on the Draw a Shape Test; double touch asynchrony and pinching gap time on the Pinching Test; sway path on the SBT; and step frequency variance on the 2MWT. 2MWT Two-Minute Walk Test, 9HPT Nine-Hole Peg Test, B2T bottom to top, btw between, CR correct responses, CV coefficient of variation, EDSS Expanded Disability Status Scale, e-SDMT smartphone-based electronic Symbol Digit Modalities Test, num. number of, parahip. parahippocampus, SBT Static Balance Test, SDMT Symbol Digit Modalities Test, SFI speed fatigability index, T2B top to bottom, T25FW Timed 25-Foot Walk, UTT U-Turn Test, vel. drawing velocity



DISCUSSION

6. DISCUSSION

Floodlight study demonstrates that the use of smartphones and smartwatches for remote daily active testing and continuous passive monitoring is feasible over 6 months and provides further support to earlier studies, which have shown that HCs and pwMS were capable of completing daily tasks on a smartphone[282]. This study provides further evidence for the use of digital technology, including smartphones, for data collection. Other studies in MS have used smartphone apps to (1) assess steps when walking on a treadmill[283]; (2) assess pain, fatigue, anxiety, and QoL[284]; and (3) assess the feasibility of gathering passive and active performance data[282].

In this protocol, the Floodlight solution collected metrics on cognition, mood, upper extremity function, and gait and posture by instructing participants to perform a set of daily active tests, which should take approximately 5 minutes in total to complete and capture activity data via passive monitoring over a period of 24 weeks. A previous study has shown that 51% of participants (22/38 of pwMS and 17/38 of HCs) completed 12 months of daily data collection, where participants were prompted to complete one assigned test[282]. In the context of the Floodlight study, we observed that overall adherence to active tests was 70% (16.68/24 weeks), which appears to be higher than the adherence of participants to 12 months of daily data collection (39/76, 51%) from Bove et al[282]. However, comparisons between the studies are limited, as the study design and burden of testing are different—for example, the app from Bove et al contained 19 different tests, of which participants were prompted to complete one each day. As the Floodlight MS app was integrated into standalone devices in this study, deployment of the app on participants' own mobile devices may increase adherence because it removes the need to carry a separate, dedicated device and decreases burden on the individual.

A recent study assessing the feasibility of the MS TeleCoach, a novel intervention offering telemonitoring of fatigue and telecoaching of physical activity in people with MS, showed that participants were highly engaged, with 76% (57/75) of

participants completing the study, and 91% (21/23) of a subset of completers showing a median of quite satisfied in the patient satisfaction questionnaire[284]. During the 12-week study period, use of the MS TeleCoach improved fatigue levels in pwMS with moderate to severe fatigue, suggesting that implementation of digital technologies can enhance patient performance. Together with the data presented here, these results indicate that the use of consumer devices by pwMS for sensor data capture fulfills the prerequisites of pwMS satisfaction and acceptable adherence to daily active tests and passive monitoring for potential integration in long-term clinical trials and treatment monitoring.

Test-retest reliability was consistent with ICCs reported for standard clinical measures in pwMS[286]. As anticipated, statistically significant correlations were observed between test features from the Floodlight MS app and related standard clinical and MRI measures. Correlations of similar strength between comparable smartphone sensor-based remote monitoring tests and clinician-administered tests have been reported in the cognitive[287-288], upper extremity function[289], and gait domains[290], despite differences in test design and features. For example, MSCopilot[291] is a mobile app designed for self-assessment of pwMS, which also integrates the four clinician-administered traditional tests belonging to the MSFC: gait, manual dexterity, cognition and visual pathway tests, in digital version. A French, multicentre, open-label, randomized and controlled study included 116 MS patients and 69 HCs with the aim of validating the MSCopilot tool in comparison with the MSFC and the EDSS. Both the mobile app tests and the traditional MSFC tests were carried out in the clinical setting. All participants performed MSCopilot and MSFC tests at day 0. To assess reproducibility, 46 pwMS performed the same tests at day 30 ± 3. The study concluded that MSCopilot was not inferior to the MSFC (area under the curve, 0.92 and 0.89, respectively; p=0.3) when evaluating disability in patients with RR or progressive MS. Thus, MSCopilot composite scores were as highly correlated to the EDSS ($|r| = 0.65$, $p<0.01$) as their MSFC counterparts, confirming the well-known correlation of the MSFC with the EDSS[291].

The active tests and passive monitoring at Floodlight study also showed mostly fair correlations with EDSS and respective MSIS-29 subscales and items, indicating that measurements obtained with the Floodlight MS app agree both with the overall level of MS-related disability and participant's perception of the impact of their disease.

In this sense, Elevate MS study[292], which included pwMS (n = 495) and HCs (n= 134) from the United States, also aimed at evaluating the feasibility and utility of capturing health data related to real-world MS using the 'elevateMS' app during 12 weeks to investigate associations between self-reported MS severity and sensor-based active functional test measurements, adding the assessment of local meteorological conditions impact on the burden of disease. This app included the quality of life in neurological disorders questionnaire (Neuro-QoL), the self-reported physical disability scale (PDDS) and daily health records (symptoms, potential triggers, mobility, pain). Active functional tests consisted of finger tapping, walking, balance, the cognitive test "The voice-controlled Digit Symbol Substitution Test" and the finger-nose test. Local weather data was collected each time the participants completed an active task. The symptoms most described by the patients were fatigue (62.6%), weakness (44.8%), memory / attention problems (42.2%) and difficulty walking (41.4%), while than the triggers found with greater frequency were high ambient temperature (52.3%), stress (50.5%) and sleeping late at nights (44.6%). The PDDS reported by patients and Neuro-QoL scores were significantly associated with the results of the functional tests. Finally, the local temperature was significantly associated with the performance of the active tests of the participants.

Regarding MRI parameters, Floodlight study did not show any strong correlation between T2 FLAIR lesion volume with either active tests or passive monitoring. This is not surprising given the clinic-radiological paradox,[293] which describes the mismatch between the white matter lesion volume and the clinical outcomes in MS, and the subsequent poor cross-sectional correlation between T2-weighted imaging and MS disability measures in a relatively mild MS population[294]. Normalized brain volume correlated with test features from all assessed domains.

This is in line with previously reported correlations between normalized brain volume and measures of MS-related disability[295].

Not surprisingly, the e-SDMT most closely resembles its domain-specific standard clinical measure. The Spearman rho of 0.82 is comparable to the previously reported correlation between other smartphone-based versions of the SDMT and the pen-and-paper version of the SDMT (r= 0.71–0.85)[287-288]. In our study, we noted that the e-SDMT scores tended to be lower than the oral SDMT scores in pwMS and HCs. This is likely due to the longer time required to select the correct response on a smartphone display compared with saying the correct response out loud. Another possible reason is that the e-SDMT displays only one symbol at a time. The oral SDMT, on the contrary, provides participants the entire symbol sequence printed on a sheet of paper[264], thus allowing them to work ahead and use their working memory to a greater extent. This difference also makes the oral SDMT more dependent on eye tracking than the e-SDMT. Given the different concepts assessed by the Floodlight MS app versus clinical measures, 1:1 correlation was not necessarily expected. For example, mean trace celerity did not correlate with EDSS or the arm-related MSIS-29 items. This is likely because these clinical measures do not capture the time component as overall mean trace accuracy correlated significantly with both clinical measures.

The partial correlation analysis presented here revealed that test features from both the Pinching Test and the Draw a Shape Test independently correlate with 9HPT time. This supports the concept that specific sensor-based test features can capture performance outcome information currently not recorded with commonly used in-clinic assessments. This exemplifies the potential of sensor data to characterize functional impairment beyond a single summary score that is typically recorded for in-clinic performance outcome measures. Future work should explore the use of this technology in broader clinical apps and focus on establishing the clinical relevance for the additional information it can provide. It is possible that richer information can be extracted by incorporating additional test features. Initial results on a more comprehensive multidimensional feature

space have been previously reported for the Draw a Shape Test[296] and Walk Test[297-298].

In addition to the active tests, the Floodlight MS app also assesses gait in a free-living situation, or in daily life, through passive monitoring. It has been suggested that signs of gait alteration may be more pronounced during daily life than in conventional in-clinic metrics[299], thereby highlighting the importance of capturing out of clinic performance through passive monitoring. As such, passive monitoring may improve the translation of clinical findings to meaningful care as it informs on the patients' true abilities during daily life activities[221]. A recent study demonstrated the feasibility of passively monitoring gait in a free-living setting, as well as a good correlation between different structured testing measures and free-living features with traditional scales such as MSFC and EDSS in pwMS, using three biosensors connected to the wrist, ankle and sternum[239].

The unique digital measures captured in the Floodlight MS app demonstrated robust correlations with normalized regional brain volumes and areas, and upper cervical cord area. While these correlations overlapped with the correlations observed for standard clinical measures, several of the correlations were specific to the digital measures. These results suggest that the digital measures may yield higher functional specificity and provide novel information on functional ability. Of note, these correlations were observed in an early, relapsing cohort of pwMS, suggesting that the digital measures can capture silent pathology[300-301].

Digital and standard clinical measures were associated with many of the same neural correlates, indicating that both capture similar aspect of functional ability. As expected, both the e-SDMT and oral SDMT correlated with regions implicated in cognitive decline (ventricular expansion and smaller volume of the insula and cerebral white matter), information processing speed (deep gray matter), and working memory (superior temporal gyrus)[302-308]. Correlations were also observed with the volume of the primary somatosensory cortex (postcentral gyrus) and visual processing areas (lateral occipital lobe)[303,309]. Similarly, the

Draw a Shape Test, Pinching Test, and 9HPT were associated with areas of visual processing (lateral occipital lobe, visual deep gray matter) and cognitive decline (ventricular expansion, cerebral white matter)[303-305]. For the SBT, postural control (sway path) correlated inversely with the paracentral lobule volume. This is consistent with findings from a functional imaging study, which suggested a role of the paracentral lobule in proprioceptive processing[310]. In contrast, the BBS correlated more strongly with the deep gray matter volume. The different tasks involved in the two tests, the distinctive sensorimotor aspects they capture and the scoring - the SBT measures the total sway path (a measure of static balance), while the BBS provides an overall score describing both static and dynamic balance - may explain these differences. The mild level of MS-related impairment together with the ceiling effect observed on the BBS may have contributed to the lack of correlations with the cerebellar volume. The former may also explain the lack of correlations found for the sensor-based walking tests (UTT, 2MWT) or the T25FW[311]. With greater levels of impairment, correlations between the T25FW and thalamic volumes or upper cervical cord area could be expected[312-313].

Several correlations were only observed for the digital measures, but not for the standard clinical measures. This was most evident in the upper extremity function domain. Temporal and spatiotemporal measures derived from the Draw a Shape Test such as the overall mean trace celerity and variability of drawing velocity while drawing round shapes specifically correlated with thalamic volume, which has been previously shown to be involved in information processing[314]. These measures also specifically correlated with the volume of the pons. Brainstem atrophy is reported in early stages of MS and a higher pontine lesion load has been associated with upper extremity tremor[315]. In addition, the variability of linear drawing velocity on the figure-of-8 specifically correlated with the upper cervical cord area. This region has been previously linked to upper extremity dysfunction, particularly in more advanced or progressive disease where atrophy is more pronounced[313,316]. The fact that we observed this correlation in a mildly impaired cohort highlights the potential higher sensitivity of Floodlight digital measures. Correlations specific to the digital measures were also observed with

the Pinching Test. Double touch asynchrony, that corresponds to an asynchronous contact of the two fingers with the touchscreen while pinching a tomato, was associated with the volume of anterior cingulate cortex. No association was observed with the 9HPT time. This region is known to selectively modulate motor areas during visually coordinated tasks[317-318]. This result is not surprising considering that double touch asynchrony was specifically developed to assess the ability to perform finger coordination tasks.

The notion that digital measures may offer higher functional specificity is also supported by the explained variance analysis. A low R^2 value in the Whole Brain model but a high R^2 value in the Parcellation model indicates that certain regions contribute more than other regions to the variance observed in the digital or standard clinical measures. This can be seen in the upper extremity function domain, where selected measures derived from the Draw a Shape and Pinching Tests showed a larger increase in R^2 when switching from the Whole Brain to the Parcellation model compared with the 9HPT. This suggests that these digital measures have higher functional specificity than the 9HPT. A high R^2 value in both models, on the other hand, indicates that the different MRI regions contribute equally to the observed variance. The cognitive domain with both the e-SDMT and oral SDMT is a good example of this, with both showing comparable functional specificity.

Study limitations

As this study remains a pilot investigation designed to collect first experiences from continuous sensor data capture, the main limitation is the small sample size and short duration of follow-up. Ongoing Floodlight studies (CONSONANCE and Floodlight Open) will collect longer term data on smartphone-based sensor data capture in a larger number of participants from a broader disability spectrum. Additionally, whether physical and cognitive limitations in people with SPMS and people with PP MS differentially impacts adherence compared with people with RRMS cannot be gleaned from the current data set due to the low numbers of advanced patients enrolled in the study; however, this important question

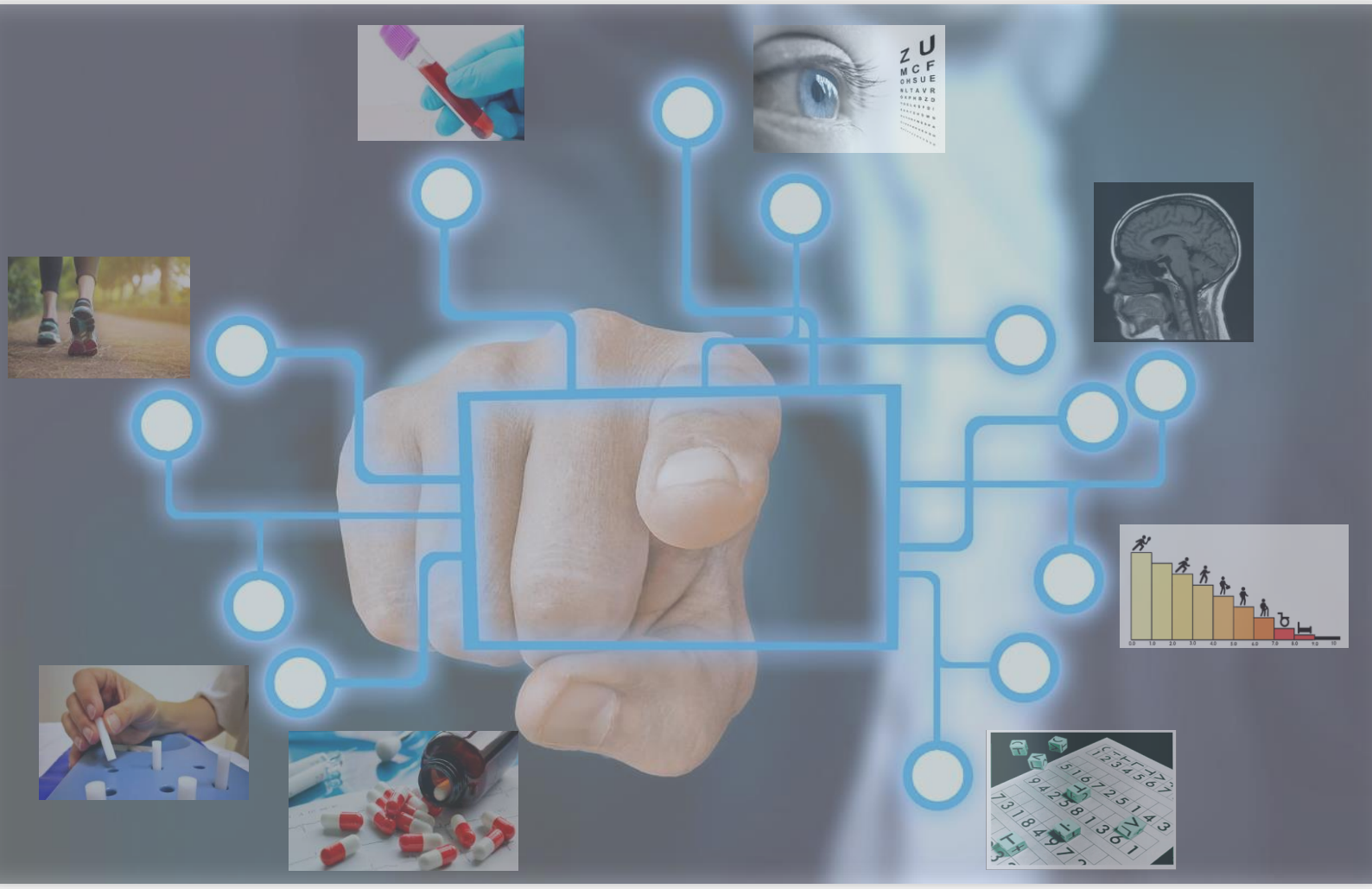
warrants future research exploring remote monitoring in patients with more advanced MS. The importance of continuous monitoring in RRMS should also not be overlooked, as the sensitivity of this novel approach aiming at detecting progression in a real-world setting may provide an earlier window into disease progression outside of the clinic.

Regarding the correlation analysis, most enrolled pwMS had mild disease with limited MS-related disability; the mean EDSS at baseline was 2.4. The current analysis assessed the performance characteristics of the Floodlight MS app in pwMS with EDSS scores in the range of 0.0–5.5. However, it has been previously shown that wearable sensors might not accurately capture step detection at slow walking speeds, particularly at EDSS score of 6.5. This feature should be considered when assessing any wearable monitoring technology. In addition, the analyses presented here were cross-sectional. Due to the relative short duration of the study (24 weeks), a longitudinal analysis on change in functional ability, disease progression, and relapses was not possible. Future studies will lend greater clarity into the use in a broader patient population, including people with more advanced disease, and the test performance over time.

Furthermore, test-retest reliability analysis was conducted in 2-week windows, in which no disease progression was assumed, as each assessment was done at most once per day. Same-day test–retest reliability analysis will be addressed in future work using data from subsequent studies. Further work will also be needed on the development of domain-specific and overall MS outcome measures based on digital health technology.

Regarding the study assessing the regional neural correlates of the digital measures, MRI data was only collected for pwMS. As a result, we were unable to disentangle physiological from pathological effects. In addition, the relatively short study duration of 24 weeks did not allow us to perform any longitudinal analyses or assess the relationship between the digital measures and disease progression. The pwMS enrolled in this study had mostly mild disease, and many had a 9HPT time within the range of healthy individuals or little-to-no gait and balance impairment as measured by the T25FW and BBS. This may have

weakened the correlations, in particular between the digital measures of gait and balance and regional brain volume[319]. In addition, structural, or even functional, cortical reorganization, which helps to maintain functional ability in early stages of the disease despite the structural damage to the brain,[320-322] may have also contributed to the weakened correlations. Nonetheless, we observed correlations of clinical relevance. With longer disease duration, the strength of the observed correlations is expected to increase. Using data from other ongoing and forthcoming studies (CONSONANCE, NCT03523858; Floodlight™ MS – TONiC, ISRCTN11088592) in future work, we will explore both structural cortical and subcortical networks[323] and longitudinal changes in MRI and digital measures to better characterize the utility of sensor-based tests as prognostic biomarkers. Such biomarkers could be used for early identification of patients with silent progression at risk of future disability accrual, and optimization of individual treatment strategies.



CONCLUSIONS

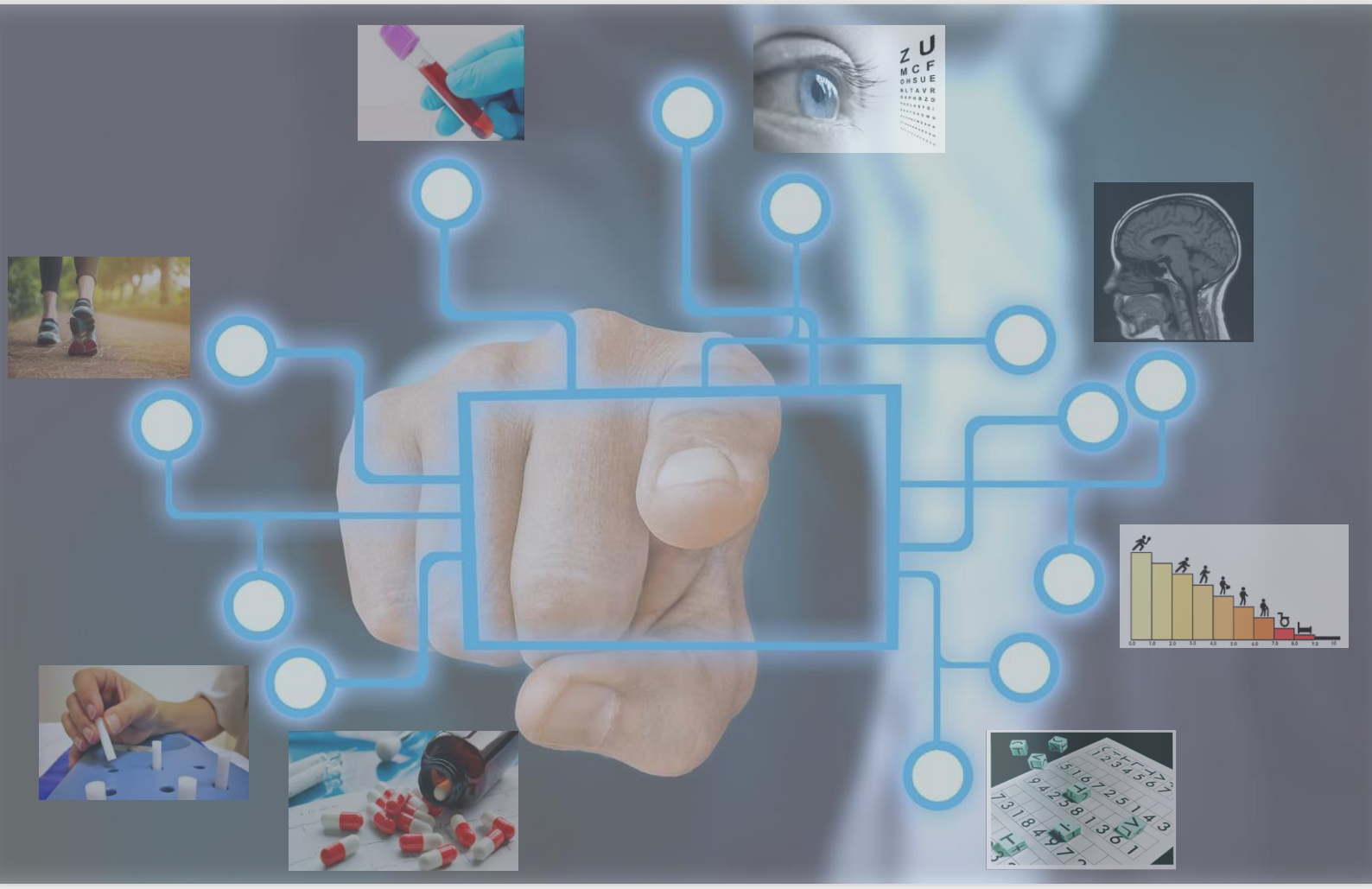
7. CONCLUSIONS

1- Our study showed that pwMS are highly engaged with performing active tests and capturing continuous data via passive monitoring and are satisfied with the Floodlight MS app. Neither satisfaction nor adherence showed strong correlation with study population characteristics. More than 60% (46/72) of people with multiple sclerosis indicated their interest to continue to use Floodlight MS app, and approximately 90% (65/72) wanted to see the results of their tests in real time as biofeedback, which was implemented in future studies using the Floodlight solution.

2- The study also demonstrated that the Floodlight MS app provides reliable measures that align with standard clinical and MRI measures used to quantify MS functional impairment and overall disability. Test–retest reliability was moderate-to-good, and significant correlations in the expected direction were observed between the test features from the Floodlight MS app and standard clinical and MRI measures. While active tests were conducted daily or weekly, passive monitoring permitted the continuous assessment of gait during daily life activities. The higher temporal resolution and multidimensional feature space of functional data collected by this platform hold the potential to capture subtle, potentially disease-related information which are not readily discriminated by clinician administered assessments. It also has the potential to improve and standardize assessment of MS disease over time, provide pwMS and health care professionals in both specialty and primary care environments a better understanding of disease progression, change the way MS is monitored in clinical trials and daily practice, and ultimately improve patient care.

3- Finally, we observed digital measures obtained with the Floodlight MS app correlated with normalized volumes and areas of distinct anatomical regions.

While many of the correlations were also observed with standard clinical measures, some were only observed with the digital measures. In addition, the explained variance analysis suggests a higher functional specificity for digital measures, in particular in the upper extremity function domain. These results indicate that digital measures, by leveraging sensor technology, can probe multiple different neurologic domains rather than just providing an overall assessment of functional ability. Thus, digital measures have the potential to complement standard clinical measures by providing a more detailed picture of MS and a more accessible assessment of functional ability. In future, identification of digital measures of disease progression associated with gray matter pathology could help to identify prognostic biomarkers and individualized therapies with increased efficacy for patients with a predisposition to develop more severe cortical pathology and associated clinical deterioration.



FUTURE RESEARCH

8. FUTURE RESEARCH

The feasibility and clinical validity of the Floodlight app use has been verified throughout the development of this doctoral thesis[330] suggesting that it may have the potential to improve and standardize assessment of MS disease, provide pwMS and health care professionals a better understanding of disease progression, change the way MS is monitored in clinical trials and daily practice, and ultimately improve patient care. However, certain issues remain to be resolved before affirming such suggestion. First, there are no data on adherence and feasibility of using the Floodlight app in larger MS populations over a longer follow-up period. Second, the lack of longitudinal data analyses does not allow establishing which changes in the test's performance should be considered of clinical relevance. So, clinical interpretation of the app data captured over time remains to be defined. Finally, no study has attempted to explore the potential benefits of clinical remote monitoring versus standard clinical care on patient health outcomes in a real-world setting.

To this end, the next step is to conduct a randomized, double-blind study to confirm the feasibility, and assess the clinical interpretation and clinical utility of Floodlight MS compared to standard clinical monitoring. A total study period of 4 years is anticipated.

Although the project protocol is still in process, the objectives and outcomes are described below.

Objectives and outcomes

Primary objective

The main objective of the study is to assess the feasibility of RMT of MS using the Floodlight app in a large sample of patients, over a long follow-up period, in a real clinical setting. As well as identify which resources provided by the health centre / Hoffmann-La Roche Pharmaceutical (technical support / health

personnel, etc.) are necessary to guarantee the use of the Floodlight app in a sustained manner over time in routine clinical practice.

Primary endpoints:

- Percentage of adherent patients
- Number of monthly reminder calls made from the health centre (by the health data manager) to patients to encourage them to have a good adherence.
- Number of events reported per month by patients because of technical problems using the app (app failure) or problems with WIFI as the main reason for not sending the information captured by the app.

Secondary objective

To better understand and predict disease activity and progression using data captured from the Floodlight app leading to medical interventions

Secondary endpoint:

- Percent worsening (for example, 10%, 15%, 20%, or 25%) of Floodlight app test performance, confirmed in the consecutive 4 days, that leads in the majority of patients (>50%) to medical interventions such as requesting an additional MRI, starting or changing a treatment, referring to the rehabilitation unit, starting symptomatic treatment, etc.

Tertiary objectives

1- To assess the efficacy of Floodlight MS compared to standard clinical care on MRI activity.
2- To assess the efficacy of Floodlight MS compared to standard clinical care on: Disability progression / QoL / Brain atrophy.

Tertiary outcomes:

1a- Proportion of patients with new/enlarged T2 lesions from month (M)12 to M36.

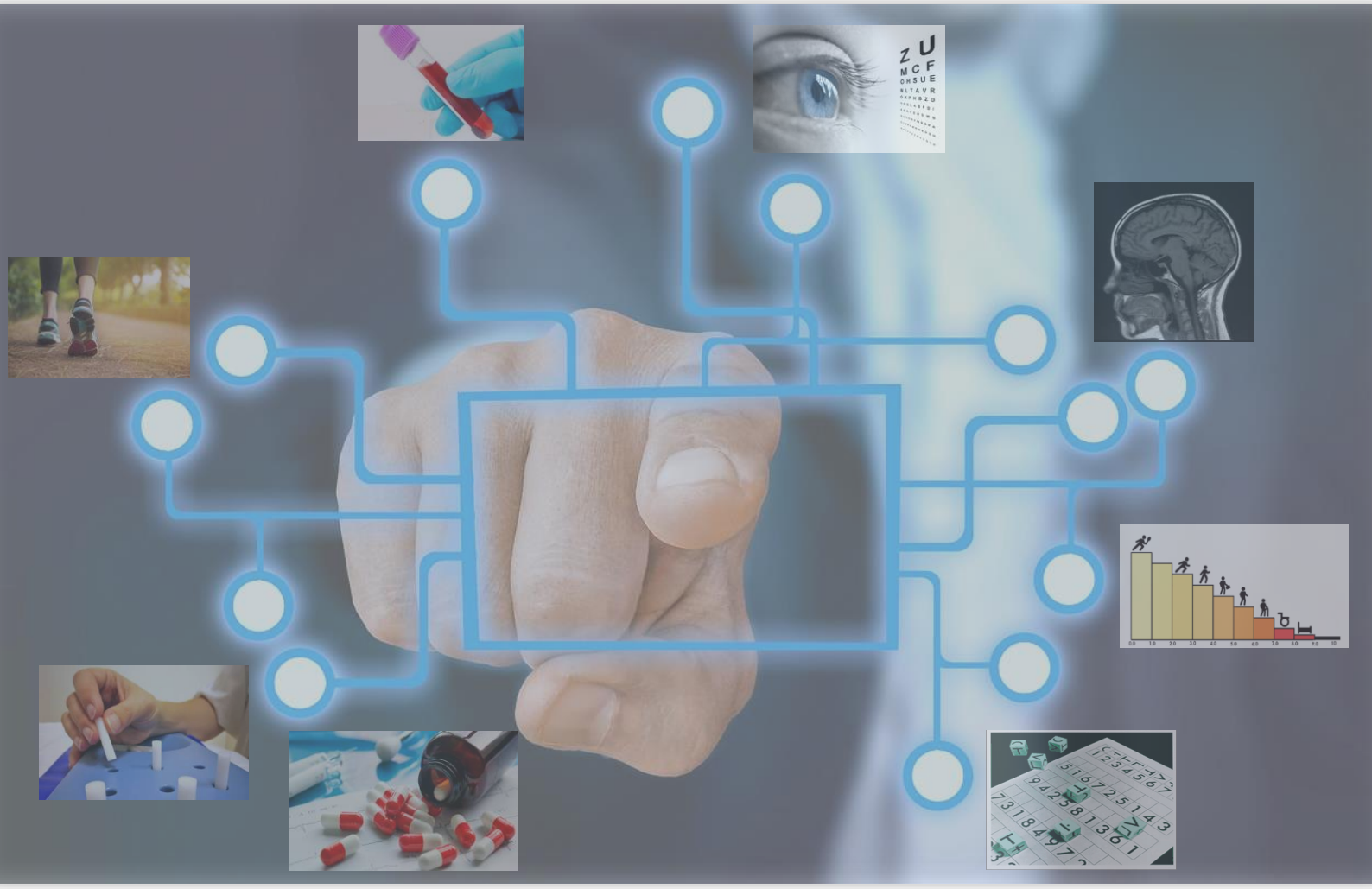
1b- Proportion of patients with new/enlarged T2 lesions from M24 to M36.

1c- Number of new/enlarged T2 lesions from M24 to M36.

2a- Time to 24-week confirmed EDSS worsening.

2b- Changes in Neuro-QoL questionnaire from M12 to M36.

2c- Global and grey matter brain volume loss from M12 to M36.



BIBLIOGRAPHY

9. BIBLIOGRAPHY

1- Reich DS, Lucchinetti CF, Calabresi PA. Multiple Slcerosis. N Engl J Med 2018;378:169-180.

2- Perez-Carmona N, Fernandez-Jover E and Sempere AP. Epidemiology of multiple sclerosis in Spain. Rev Neurol 2019;69:32–38.

3- Miller DH, Weinshenker BG, Filippi M, Banwell BL, Cohen JA, Freedman MS, Galetta SL, Hutchinson M, Johnson RT, Kappos L, Kira J, Lublin FD, McFarland HF, Montalban X, Panitch H, Richert JR, Reingold SC, Polman CH. Differential Diagnosis of Suspected Multiple Sclerosis: A Consensus Approach. Mult Scler 2008;14:1157-74.

4- Alonso A and Hernan MA. Temporal Trends in the Incidence of Multiple Sclerosis: A Systematic Review. Neurology 2008;71:129-35.

5- Alonso A, Hernan MA, Ascherio A. Allergy, Family History of Autoimmune Diseases, and the Risk of Multiple Sclerosis. Acta Neurol Scand 2008;117:15-20.

6- Koch-Henriksen N and Sorensen PS. "The Changing Demographic Pattern of Multiple Sclerosis Epidemiology. Lancet Neurol 2010;9:520-32.

7- Kotzamani D, Panou T, Mastorodemos V, Tzagournissakis M, Nikolakaki H, Spanaki C, Plaitakis A. Rising Incidence of Multiple Sclerosis in Females Associated with Urbanization. Neurology 2012;78:1728-35.

8- Dunn SE and Steinman L. The Gender Gap in Multiple Sclerosis: Intersection of Science and Society. JAMA Neurol 2013;70:634-5.

9- Ebers GC. Environmental Factors and Multiple Sclerosis. Lancet Neurol 2008;7:268-77.

10- Simpson Jr. S, Blizzard L, Otahal P, Van der Mei I, Taylor B. Latitude Is Significantly Associated with the Prevalence of Multiple Sclerosis: A Meta-Analysis. J Neurol Neurosurg Psychiatry 2011;82:1132-41.

11- Aguirre-Cruz L, Flores-Rivera J, de La Cruz-Aguilera DL, Rangel-Lopez E, Corona T. Multiple Sclerosis in Caucasians and Latino Americans. Autoimmunity 2011;44:571-5.

12- Giampaolo D, Bhigjee A, Retief C, Isaacs M, Britz M, Opperman D, Govender R, van Rensburg M. Guideline for the Diagnosis and Management of Multiple Sclerosis: A Southern African Perspective. S Afr Med J 2013;103:670-91.

13- Langer-Gould A, Brara SM, Beaber BE, Zhang JL. Incidence of Multiple Sclerosis in Multiple Racial and Ethnic Groups. Neurology 2013;80:1734-9.

14- Marrie RA. Environmental Risk Factors in Multiple Sclerosis Aetiology. Lancet Neurol 2004;3:709-18.

15- Hernan MA, Olek MJ, Ascherio A. Geographic Variation of MS Incidence in Two Prospective Studies of US Women. Neurology 1999;53:1711-8.

16- Hafler DA, Compston A, Sawcer S, Lander ES, Daly MJ, De Jager PL, de Bakker PI, Gabriel SB, Mirel DB, Ivinson AJ, Pericak-Vance MA, Gregory SG, Rioux JD, McCauley JL, Haines JL, Barcellos LF, Cree B, Oksenberg JR, Hauser SL. Risk Alleles for Multiple Sclerosis Identified by a Genomewide Study. N Engl J Med 2007;357:851-62.

17- De Jager PL, Jia X, Wang J, de Bakker PI, Ottoboni L, Aggarwal NT, Piccio L, et al. Meta-Analysis of Genome Scans and Replication Identify CD6, IRF8 and TNFRSF1A as New Multiple Sclerosis Susceptibility Loci. Nat Genet 2009;41:776-82.

18- Sawcer S, Hellenthal G, Pirinen M, Spencer CC, Patsopoulos NA, Moutsianas L, et al. "Genetic Risk and a Primary Role for Cell-Mediated Immune Mechanisms in Multiple Sclerosis." *Nature* 2011;476:214-9.

19- Baulina N, Kiselev I, Favorova O. Imprinted genes and multiple sclerosis: what do we know? Int J Mol Sci 2021;22:1346.

20- Ascherio A and Munger KL. Environmental Risk Factors for Multiple Sclerosis. Part Ii: Noninfectious Factors. Ann Neurol 2007;61:504-13.

21- Ascherio A, Munger KL, White R, Kochert K, Simon KC, Polman CH, et al. Vitamin D as an Early Predictor of Multiple Sclerosis Activity and Progression. *JAMA Neurol* 2014;71:306-14.

22- Munger KL, Zhang SM, O'Reilly E, Hernan MA, Olek MJ, Willett WC, Ascherio A. Vitamin D Intake and Incidence of Multiple Sclerosis. Neurology 2004;62:60-5.

23- Thacker EL, Mirzaei F, Ascherio A. Infectious Mononucleosis and Risk for Multiple Sclerosis: A Meta-Analysis. Ann Neurol 2006;59:499-503.

24- Levin LI, Munger KL, Rubertone MV, Peck CA, Lennette ET, Spiegelman D, Ascherio A. Temporal Relationship between Elevation of Epstein-Barr Virus Antibody Titers and Initial Onset of Neurological Symptoms in Multiple Sclerosis. Jama 2005;293:2496-500.

25- Ascherio A, Munger KL, Lennette ET, Spiegelman D, Hernan MA, Olek MJ, Hankinson SE, Hunter DJ. Epstein-Barr Virus Antibodies and Risk of Multiple Sclerosis: A Prospective Study. Jama 2001;286:3083-8.

26- Lunemann JD, Tintore M, Messmer B, Strowig T, Rovira A, Perkal H, Caballero E, Munz C, Montalban X, Comabella M. Elevated Epstein-Barr Virus-Encoded Nuclear Antigen-1 Immune Responses Predict Conversion to Multiple Sclerosis. Ann Neurol 2010;67:159-69.

27- Ascherio A and Munger KL. 99th Dahlem Conference on Infection, Inflammation and Chronic Inflammatory Disorders: Epstein-Barr Virus and Multiple Sclerosis: Epidemiological Evidence. Clin Exp Immunol 2010;160:120-4.

28- Bjornevik K, Cortese M, Healy BC, Kuhle J, Mina MJ, Leng Y, Elledge SJ, Niebuhr DW, Scher AI, Munger KL, Ascherio A. Longitudinal analysis reveals high prevalence of Epstein-Barr virus associated with multiple sclerosis. Science 2022; 375:296-301.

29- Bach JF. The Effect of Infections on Susceptibility to Autoimmune and Allergic Diseases. N Engl J Med 2002;347: 911-20.

30- Ponsonby AL, van der Mei I, Dwyer T, Blizzard L, Taylor B, Kemp A, Simmons R, Kilpatrick T. Exposure to Infant Siblings During Early Life and Risk of Multiple Sclerosis. Jama 2005;293: 463-9.

31- Hollenbach JA, Oksenberg JR. The immunogenetics of multiple sclerosis: a comprehensive review. J. Autoimmun 2015;64:13–25.

32- Turnbaugh PJ, Ridaura VK, Faith JJ, Rey FE, Knight R, Gordon JI. The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. Sci Transl Med 2009;1:6ra14.

33- Mandel I, Paperna T, Glass-Marmor L, Volkowich A, Badarny S, Schwartz I, Vardi P, Koren I, Miller A. Tight junction proteins expression and modulation in immune cells and multiple sclerosis. J Cell Mol Med 2012;16:765–775.

34- Chen J, Chia N, Kalari KR, Yao JZ, Novotna M, Soldan MMP, Luckey DH, Marietta EV, Jeraldo PR, Chen X. Multiple sclerosis patients have a distinct gut microbiota compared to healthy controls. Sci Rep 2016; 6:1–10.

35- Berer K, Gerdes LA, Cekanaviciute E, Jia X, Xiao L, Xia Z, Liu C, Klotz L, Stauffer U, Baranzini SE. Gut microbiota from multiple sclerosis patients enables spontaneous autoimmune encephalomyelitis in mice. Proc Natl Acad Sci 2017;114:10719–10724.

36- Buscarinu MC, Fornasiero A, Romano S, Ferraldeschi M, Mechelli R, Reniè R, Morena E, Romano C, Pellicciari G, Landi AC. The contribution of gut barrier changes to multiple sclerosis pathophysiology. Front. Immunol 2019;10:1916.

37- Schepici G, Silvestro S, Bramanti P, E. Mazzon E. The gut microbiota in multiple sclerosis: an overview of clinical trials. Cell Transplant 2019;28:1507–1527.

38- Franklin GM and Nelson L. Environmental Risk Factors in Multiple Sclerosis: Causes, Triggers, and Patient Autonomy. Neurology 2003;61: 1032-4.

39- Riise T, Nortvedt MW, Ascherio A. Smoking Is a Risk Factor for Multiple Sclerosis. Neurology 2003;61:1122-4.

40- Dobson R, Giovannoni G, Ramagopalan S. The Month of Birth Effect in Multiple Sclerosis: Systematic Review, Meta-Analysis and Effect of Latitude. J Neurol Neurosurg Psychiatry 2013;84: 427-32.

41- Fiddes B, Wason J, Kemppinen A, Ban M, Compston A and Sawcer S. Confounding Underlies the Apparent Month of Birth Effect in Multiple Sclerosis. Ann Neurol 2013;73:714-20.

42- Montalban X and Tintore M. Multiple Sclerosis in 2013: Novel Triggers, Treatment Targets and Brain Atrophy Measures. Nat Rev Neurol 2014;10:72-3.

43- Kleinewietfeld M, Manzel A, Titze J, Kvakan H, Yosef N, Linker RA, Muller DN, Hafler DA. Sodium Chloride Drives Autoimmune Disease by the Induction of Pathogenic Th17 Cells. Nature 2013;496: 518-22.

44- Wu C, Yosef N, Thalhamer T, Zhu C, Xiao S, Kishi Y, Regev A, Kuchroo VK. Induction of Pathogenic Th17 Cells by Inducible Salt-Sensing Kinase SGK1. Nature 2013;496:513-7.

45- Munger KL, Bentzen J, Laursen B, Stenager E, Koch-Henriksen N, Sorensen TI, Baker JL. Childhood Body Mass Index and Multiple Sclerosis Risk: A Long-Term Cohort Study. Mult Scler 2013;19: 1323-9.

46- Gianfrancesco MA, Acuna B, Shen L, Briggs FB, Quach H, Bellesis KH, Bernstein A, Hedstrom AK, Kockum I, Alfredsson L, Olsson T, Schaefer C, Barcellos LF. Obesity During Childhood and Adolescence Increases Susceptibility to Multiple Sclerosis after Accounting for Established Genetic and Environmental Risk Factors. Obes Res Clin Pract 2014;8: e435-47.

47- Huppke B, Ellenberger D, Hummel H, Stark W, Robl M, Gartner J, Huppke P. Association of obesity with multiple sclerosis risk and response to first line disease modifying drugs in children. Jama Neurol 2019;76:1157-1165.

48- Hedstrom AK, Lima Bomfim I, Barcellos L, Gianfrancesco M, Schaefer C, Kockum I, Olsson T, Alfredsson L. Interaction between Adolescent Obesity and HLA Risk Genes in the Etiology of Multiple Sclerosis. Neurology 2014;82:865-72.

49- Hedstrom AK, Lima Bomfim I, Hillert J, Olsson T, Alfredsson L. Obesity Interacts with Infectious Mononucleosis in Risk of Multiple Sclerosis. Eur J Neurol 2015;22:578-e38.

50- Compston A and Coles A. Multiple Sclerosis. Lancet 2008;372:1502-17.

51- Weiner HL. "Multiple Sclerosis Is an Inflammatory T-Cell-Mediated Autoimmune Disease." Arch Neurol 2004;61:1613-5.

52- Comabella M and Khoury SJ. Immunopathogenesis of Multiple Sclerosis. Clin Immunol 2012:142:2-8.

53- Frohman EM, Racke MK, Raine CS. Multiple Sclerosis--the Plaque and Its Pathogenesis. *N* Engl J Med 2006;354:942-55.

54- Lovett-Racke AE, Trotter JL, Lauber J, Perrin PJ, June CH, Racke MK. Decreased Dependence of Myelin Basic Protein-Reactive T Cells on CD28-Mediated Costimulation in Multiple Sclerosis Patients. A Marker of Activated/Memory T Cells. J Clin Invest 1998;101:725-30.

55- Munz C, Lunemann JD, Getts MT, Miller SD. Antiviral Immune Responses: Triggers of or Triggered by Autoimmunity? Nat Rev Immunol 2009;9:246-58.

56- Wucherpfennig KW and Strominger JL. Molecular Mimicry in T Cell-Mediated Autoimmunity: Viral Peptides Activate Human T Cell Clones Specific for Myelin Basic Protein. Cell 1995;80:695-705.

57- Zhang H, Podojil JR, Luo X, Miller SD. Intrinsic and Induced Regulation of the Age-Associated Onset of Spontaneous Experimental Autoimmune Encephalomyelitis. J Immunol 2008;181:4638-47.

58- Furtado GC, Marcondes MC, Latkowski JA, Tsai J, Wensky A, Lafaille JJ. Swift Entry of Myelin-Specific T Lymphocytes into the Central Nervous System in Spontaneous Autoimmune Encephalomyelitis. J Immunol 2008;181:4648-55.

59- Viglietta V, Baecher-Allan C, Weiner HL, Hafler DA. Loss of Functional Suppression by CD4+CD25+ Regulatory T Cells in Patients with Multiple Sclerosis. J Exp Med 2004;199:971-9.

60- Waubant E, Goodkin DE, Gee L, Bacchetti P, Sloan R, Stewart T, Andersson PB, Stabler G, Miller K. Serum MMP-9 and TIMP-1 levels are related to MRI activity in relapsing multiple sclerosis. Neurology 1999;53:1397-401

61- Holman DW, Klein RS, Ransohoff RM. The blood-brain-barrier, chemokines and Multiple Sclerosis. Biochim Biophys Acta 2011;1812:220-30.

62- Loleit V, Biberacher V, Hemmer B. Current and Future Therapies Targeting the Immune System in Multiple Sclerosis. Curr Pharm Biotechnol 2014;15:276-96.

63- McFarland HF and Martin R. Multiple Sclerosis: A Complicated Picture of Autoimmunity. Nat Immunol 2007;8:913-9.

64- Korn T, Bettelli E, Oukka M, Kuchroo VK. IL-17 and Th17 Cells. Annu Rev Immunol 2009;27: 485-517.

65- Langrish CL, Chen Y, Blumenschein WM, Mattson J, Basham B, Sedgwick JD, McClanahan T, Kastelein RA, Cua DJ. IL-23 Drives a Pathogenic T Cell Population That Induces Autoimmune Inflammation. J Exp Med 2005;201:233-40.

66- Iwakura Y and Ishigame H. The IL-23/IL-17 Axis in Inflammation. J Clin Invest 2006;116:1218-22.

67- Tzartos JS, Friese MA, Craner MJ, Palace J, Newcombe J, Esiri MM, and Fugger L. Interleukin-17 Production in Central Nervous System-Infiltrating T Cells and Glial Cells Is Associated with Active Disease in Multiple Sclerosis. Am J Pathol 2008;172:146-55.

68- Kebir H, Kreymborg K, Ifergan I, Dodelet-Devillers A, Cayrol R, Bernard M, Giuliani F, Arbour N, Becher B, Prat A. Human Th17 Lymphocytes Promote Blood-Brain Barrier Disruption and Central Nervous System Inflammation. Nat Med 2007;13:1173-5.

69- Hofstetter H, Gold R, Hartung HP. Th17 Cells in MS and Experimental Autoimmune Encephalomyelitis. Int MS J 2009;16: 12-8.

70- Frohman EM, Racke MK, Raine CS. Multiple Sclerosis--the Plaque and Its Pathogenesis. N Engl J Med 2006;354:942-55.

71- Raphael I, Nalawade S, Eagar TN, Forsthuber TG. T Cell Subsets and Their Signature Cytokines in Autoimmune and Inflammatory Diseases. Cytokine 2015;74:5-17.

72- Babbe H, Roers A, Waisman A, Lassmann H, Goebels N, Hohlfeld R, et al. Clonal Expansions of CD8(+) T Cells Dominate the T Cell Infiltrate in Active Multiple Sclerosis Lesions as Shown by Micromanipulation and Single Cell Polymerase Chain Reaction. J Exp Med 2000;192:393-404.

73- Johnson AJ, Suidan GL, McDole J, Pirko I. The CD8 T Cell in Multiple Sclerosis: Suppressor Cell or Mediator of Neuropathology? Int Rev Neurobiol 2007;79:73-97.

74- Huseby ES, Liggitt D, Brabb T, Schnabel B, Ohlen C, Goverman J. A Pathogenic Role for Myelin-Specific CD8(+) T Cells in a Model for Multiple Sclerosis. J Exp Med 2001;194:669-76.

75- Medana I, Martinic MA, Wekerle H, Neumann H. Transection of Major Histocompatibility Complex Class I-Induced Neurites by Cytotoxic T Lymphocytes. Am J Pathol 2001;159:809-15.

76- Gharibi T, Babaloo Z, Hosseini A, Marofi F, Ebrahimi-kalan A, Jahandideh S, Baradaran B. The rol pf B cells in the immunopathogenesis of multiple sclerosis. Immunology 2020;160:325-335.

77- McDonald WI, Compston A, Edan G, Goodkin D, Hartung HP, Lublin FD, et al. Recommended Diagnostic Criteria for Multiple Sclerosis: Guidelines from the International Panel on the Diagnosis of Multiple Sclerosis. Ann Neurol 2001;50:121-7.

78- Polman CH, Reingold SC, Edan G, Filippi M, Hartung HP, Kappos L, et al. Diagnostic Criteria for Multiple Sclerosis: 2005 Revisions to the "McDonald Criteria". Ann Neurol 2005;58:840-6.

79- Polman CH, Reingold SC, Banwell B, Clanet M, Cohen JA, Filippi M, et al. Diagnostic Criteria for Multiple Sclerosis: 2010 Revisions to the McDonald Criteria. Ann Neurol 2011;69:292-302.

80- Thompson AJ, Banwell B, Barkhof F, Carroll WM, Coetzee T, Comi G, et al. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. Lancet Neurol 2018;17:162-173

81- Magliozzi R, Howell O, Vora A, Serafini B, Nicholas R, Puopolo M, Reynolds R, Aloisi F. Meningeal B-Cell Follicles in Secondary Progressive Multiple Sclerosis Associate with Early Onset of Disease and Severe Cortical Pathology. Brain 2007;130:1089-104.

82- Hauser SL, Waubant E, Arnold DL, Vollmer T, Antel J, Fox RJ, Bar-Or A, Panzara M, Sarkar N, Agarwal S, Langer-Gould A, Smith CH; HERMES Trial Group. B-Cell Depletion with Rituximab in Relapsing-Remitting Multiple Sclerosis. N Engl J Med 2008;358:676-88.

83- Hemmer B, Archelos JJ, Hartung HP. New Concepts in the Immunopathogenesis of Multiple Sclerosis. Nat Rev Neurosci 2002;3:291-301.

84- Zajicek JP, Wing M, Scolding NJ, Compston DA. Interactions between Oligodendrocytes and Microglia. A Major Role for Complement and Tumour Necrosis Factor in Oligodendrocyte Adherence and Killing. Brain 1992;115:1611-31.

85- Williams A, Piaton G, Lubetzki C. Astrocytes--Friends or Foes in Multiple Sclerosis? Glia 2007;55:1300-12.

86- Cambron M, D'Haeseleer M, Laureys G, Clinckers R, Debruyne J, De Keyser J. White-Matter Astrocytes, Axonal Energy Metabolism, and Axonal Degeneration in Multiple Sclerosis. J Cereb Blood Flow Metab 2012;32: 413-24.

87- Miller DH, Weinshenker BG, Filippi M, Banwell BL, Cohen JA, Freedman MS, et al. Differential Diagnosis of Suspected Multiple Sclerosis: A Consensus Approach. Mult Scler 2008;14:1157-74.

88- Filippi M, Bar-Or A, Piehl F, Preziosa P, Solari A, Vukusic S, Rocca M. Multiple Sclerosis. Nature Rev 2018;4:43.

89- Lublin FD, Reingold SC. Defining the clinical course of multiple sclerosis: results of an international survey. Neurology 1996;46:907–911.

90- Lublin FD, Reingold SC, Cohen JA, Cutter GR, Sørensen PS, Thompson AJ, et al. Defining the clinical course of multiple sclerosis. Neurology 2014;83:278–286.

91- Lublin FD, Coetzee T, Cohen JA, Marrie RA, Thompson AJ; International Advisory Committee on Clinical Trials in MS. The 2013 clinical course descriptors for multiple sclerosis: A clarification. Neurology 2020;94:1088-1092.

92- Filippi M and Rocca MA. MR Imaging of Multiple Sclerosis. Radiology 2011;259:659-81.

93- De Stefano, N. et al. Radiologically isolated syndrome or subclinical multiple sclerosis: MAGNIMS consensus recommendations. Mult. Scler 2018;24:214–221.

94- Okuda DT, Mowry EM, Beheshtian A, Waubant E, Baranzini SE, Goodin DS, Hauser SL, Pelletier D. Incidental MRI Anomalies Suggestive of Multiple Sclerosis: The Radiologically Isolated Syndrome. Neurology 2009;72:800-5.

95- Okuda, DT, Siva A, Kantarci O, Inglese M, Katz I, Tutuncu M, et al. Radiologically isolated syndrome: 5-year risk for an initial clinical event. PLoS One 2014;9:e90509.

96- Yamout B, Khawajah. Radiologically Isolated Syn drome and multiple sclerosis. Mult Sler Relat Disord 2017;17:234-237.

97- Filippi M, Rocca MA, Ciccarelli O, De Stefano N, Evangelou N, Kappos L, et al, on behalf of the MAGNIMS Study Group. MRI criteria for the diagnosis of multiple sclerosis: MAGNIMS consensus guidelines. Lancet Neurol 2016;15:292-303.

98- Tintore M, Rovira A, Martinez MJ, Rio J, Diaz-Villoslada P, Brieva L, et al. Isolated Demyelinating Syndromes: Comparison of Different MR Imaging Criteria to Predict Conversion to Clinically Definite Multiple Sclerosis. AJNR Am J Neuroradiol 2000;21:702-6.

99- Swanton JK, Rovira A, Tintore M, Altmann DR, Barkhof F, Filippi M, et al. MRI Criteria for Multiple Sclerosis in Patients Presenting with Clinically Isolated Syndromes: A Multicentre Retrospective Study. Lancet Neurol 2007;6:677-86.

100- Rovira A, Swanton J, Tintore M, Huerga E, Barkhof F, Filippi M, et al. A Single, Early Magnetic Resonance Imaging Study in the Diagnosis of Multiple Sclerosis. Arch Neurol 2009;66:587-92.

101- Wattjes MP, Rovira À, Miller D, Yousry TA, Sormani MP, de Stefano MP, et al. Evidence-based guidelines: MAGNIMS consensus guidelines on the use of MRI in multiple sclerosis-establishing disease prognosis and monitoring patients. Nat Rev Neurol 2015;11:597–606.

102- Barkhof F, Filippi M, Miller DH, Scheltens P, Campi A, Polman CH, et al. Comparison of MRI Criteria at First Presentation to Predict Conversion to Clinically Definite Multiple Sclerosis. Brain 1997;120:2059-69.

103- Montalban X, Tintore M, Swanton J, Barkhof F, Fazekas F, Filippi M, et al. MRI Criteria for MS in Patients with Clinically Isolated Syndromes. Neurology 2010;74:427-34.

104- Brownlee WJ, Swanton JK, Miszkiel KA, Miller DH, Ciccarelli O. Should the symptomatic region be included in dissemination in space in MRI criteria for MS? Neurology 2016;87:680–83.

105- Tintore M, Otero-Romero S, Rio J, Arrambide G, Pujal B, Tur C, et al. Contribution of the symptomatic lesion in establishing MS diagnosis and prognosis. Neurology 2016;87:1368–74.

106- Miller DH, Barkhof F, Frank JA, Parker GJ, Thompson AJ. Measurement of atrophy in multiple sclerosis: pathological basis, methodological aspects and clinical relevance. Brain 2022;12:1676-1695.

107- Roca MA, Battaglini M, Benedict RH, De Stefano N, Geurts JJ, Henry RG, et al. Brain MRI atrophy quantification in MS: from methods to clinical application. Neurology 2017;88:403-413.

108- Pohl D, Rostasy K, Reiber H, Hanefeld F. CSF Characteristics in Early-Onset Multiple Sclerosis. Neurology 2004;63:1966-7.

109- McLean BN, Luxton RW, Thompson EJ. A Study of Immunoglobulin G in the Cerebrospinal Fluid of 1007 Patients with Suspected Neurological Disease Using Isoelectric Focusing and the Log IgG-Index. A Comparison and Diagnostic Applications. Brain 1990;113:1269-89.

110- Giesser BS. Diagnosis of Multiple Sclerosis. Neurol Clin 2011;29:381-8.

111- Dobson R, Ramagopalan S, Davis A, Giovannoni G. Cerebrospinal Fluid Oligoclonal Bands in Multiple Sclerosis and Clinically Isolated Syndromes: A Meta-Analysis of Prevalence, Prognosis and Effect of Latitude. J Neurol Neurosurg Psychiatry 2013;84:909-14.

112- Freedman MS, Thompson EJ, Deisenhammer F, Giovannoni G, Grimsley G, Keir G, et al. Recommended Standard of Cerebrospinal Fluid Analysis in the Diagnosis of Multiple Sclerosis: A Consensus Statement. Arch Neurol 2005;62:865-70.

113- Gronseth GS and Ashman EJ. Practice Parameter: The Usefulness of Evoked Potentials in Identifying Clinically Silent Lesions in Patients with Suspected Multiple Sclerosis (an Evidence-Based Review): Report of the Quality Standards Subcommittee of the American Academy of Neurology. Neurology 2000;54:1720-5.

114- Dobson R and Giovannoni G. Multiple sclerosis – A review. Eur J Neurol 2019; 26: 27–40.

115- Brownlee WJ, Hardy TA, Fazekas F, et al. Diagnosis of multiple sclerosis: Progress and challenges. Lancet 2017;389:1336–1346.

116- Geraldes R, Ciccarelli O, Barkhof F, et al. The current role of MRI in differentiating multiple sclerosis from its imaging mimics. Nat Rev Neurol 2018;14:199–213.

117- Solomon AJ, Naismith RT and Cross AH. Misdiagnosis of multiple sclerosis: Impact of the 2017 McDonald criteria on clinical practice. Neurology 2019; 92:26–33.

118-Ömerhoca S, Akkaş SY and İçen NK. Multiple sclerosis: Diagnosis and differential diagnosis. Noro Psikiyatr Ars 2018; 55(Suppl 1): S1-S9.

119-Thomson AJ, Banwell BL, Barkhof F, et al. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. Lancet Neurol 2018;17:162–173.

120-Gaitan MI and Correale J. Multiple sclerosis misdiagnosis: A persistent problem to solve. Front Neurol 2019;10:466.

121-Tintoré M, Vidal-Jordana A and Sastre-Garriga J. Treatment of multiple sclerosis success from bench to bedside. Nat Rev Neurol 2019;15:53–58.

122-Solomon AJ and Corboy JR. The tension between early diagnosis and misdiagnosis of multiple sclerosis. Nat Rev Neurol 2017;13:567–572.

123-Gobbin F, Zanoni M, Marangi A, et al. 2017 McDonald criteria for multiple sclerosis: Earlier diagnosis with reduced specificity. Mult Scler Relat Disord 2019;29:23–25.

124-Charil A, Yousry TA, Rovaris M, et al. MRI and the diagnosis of multiple sclerosis: Expanding the concept of "no better explanation." Lancet Neurol 2006;5:841–852.

125-Miller DH, Weinshenker BG, Filippi M, et al. Differential diagnosis of suspected multiple sclerosis: A consensus approach. Mult Scler 2008;14:1157–1174.

126-Mattson D. Update on the diagnosis of multiple sclerosis. Expert Rev Neurother 2002;2:319–328.

127-Solomon AJ, Bourdette DN, Cross AH, et al. The contemporary spectrum of multiple sclerosis misdiagnosis: A multicenter study. Neurology 2016;87:1393–1399.

128-Solomon AJ, Klein EP and Bourdette D. 'Undiagnosis' multiple sclerosis. Neurology 2012;78:1986–1991.

129-Midaglia L, Sastre-Garriga J, Pappolla A, Quibus L, Carvajal R, Vidal-Jordana A, Arrambide G, Rio J, Comabella M, Nos C, Castillo J, Galan I, Rodriguez-Acevedo B, Auger C, Tintore M, Montalban X, Rovira A. Mult Scler 2021;27:913-921.

130-Perez-Carmona N, Fernandez-Jover E and Sempere AP. Epidemiology of multiple sclerosis in Spain. Rev Neurol 2019;69:32–38.

131-Toledano M, Weinshenker BG and Solomon AJ. A clinical approach to the differential diagnosis of multiple sclerosis. Curr Neurol Neurosci Rep 2015;15:57.

132-Clarke MA, Samaraweera AP, Falah Y, et al. Single test to ARrive at multiple sclerosis (STAR-MS) diagnosis: A prospective pilot study assessing the accuracy of the central vein sign in predicting multiple sclerosis in cases of diagnostic uncertainty. Mult Scler 2020;26:433–441.

133-Absinta M, Maggi P, Sati P, et al. The "central vein sign" in patients with diagnostic "red flags" for multiple sclerosis: A prospective multicenter 3T study. Mult Scler 2020;26:421–432.

134-Solomon AJ and Weinshenker BG. Misdiagnosis of multiple sclerosis: Frequency, causes, effects, and prevention. Curr Neurol Neurosci Rep 2013;13:403.

135-Casanova B, Quintanilla-Bordás C, Gascón F. Escalation vs. early intense therapy in multiple sclerosis. J Pers Med 2022;12:119.

136- Montalban X, Hauser SL, Kappos L, Arnold DL, Bar-Or A, Comi G, et al. Ocrelizumab versus placebo in primary progressive multiple sclerosis. N Engl J Med 2017;376:209-20.

137- Kappos L, Bar-Or A, Cree BA, Fox RJ, Giovannoni G, Gold R, et al. EXPAND Clinical Investigators. Siponimod versus placebo in secondary progressive multiple sclerosis (EXPAND): a double-blind, randomised, phase 3 study. Lancet 2018;391:1263-1273.

138- Giovannoni G, Turner B, Gnanapavan S, Offiah C, Schmierer K, Marta M. Is it time to target no evident disease activity (NEDA) in multiple sclerosis? Mult Scler Relat Disord 2015;4:329-33.

139- Kaunzner UW, Al-Kawaz M, Gauthier SA. Defining disease activity and response to therapy in MS. Curr Treat Options Neurol 2017;19:20.

140- Naismith, R. T. Multiple sclerosis should be treated using a step-down strategy rather than a step-up strategy-NO. Mult Scler 2016;22:1400–1402.

141- Merkel, B., Butzkueven, H., Traboulsee, A. L., Havrdova, E. & Kalincik, T. Timing of high-efficacy therapy in relapsing-remitting multiple sclerosis: a systematic review. Autoimmun Rev 2017;16:658–665.

142- Tintoré M, Vidal-Jordana A, Sastre-Garriga J. Treatment of multiple sclerosis – success from bench to bedside. Nat Rev Neurol 2019;15:53-58.

143- Dillenseger A, Weidemann ML, Trentzsch K, Inojosa H, Hasse R, Schriefer D, Voigt I, Scholz M, Akgün K, Ziemssen T. Digital Biomarkers in Multiple Sclerosis. Brain Sci 2021;11:1519.

144- Wattjes MP, Ciccarelli O, Reich DS, Banwell B, de Stefano N, Enzinger C, et al., Magnetic Resonance Imaging in Multiple Sclerosis study group; Consortium of Multiple Sclerosis Centres; North American Imaging in Multiple Sclerosis Cooperative MRI guidelines working group. 2021 MAGNIMS-CMSC-NAIMS consensus recommendations on the use of MRI in patients with multiple sclerosis. Lancet Neurol 2021;20:653-670.

145- Edan G, Kappos L, Montalbán X, et al. Long-term impact of interferon beta-1b in patients with CIS: 8-year follow-up of BENEFIT. J Neurol Neurosurg Psychiatry 2014;85:1183–89.

146- Diebold M, Altersberger V, Décard BF, Kappos L, Derfuss T, Lorscheider J. A case of progressive multifocal leukoencephalopathy under dimethyl fumarate treatment without severe lymphopenia or immunosenescence. Mult Scler 2019;25:1682–85.

147- Nakahara J, Tomaske L, Kume K, Takata T, Kamada M, Deguchi K, et al. Three cases of non-carryover fingolimod-PML: Is the risk in Japan increased? Neurol Neuroimmunol Neuroinflamm 2019;6:e559.

148- Gerevini S, Capra R, Bertoli D, Sottini A, Imberti L. Immune profiling of a patient with alemtuzumab-associated progressive multifocal leukoencephalopathy. Mult Scler 2019;25:1196–201.

149- Dong-Si T, Richman S, Wattjes MP, et al. Outcome and survival of asymptomatic PML in natalizumab-treated MS patients. Ann Clin Transl Neurol 2014;1:755–64.

150- Grebenciucova E, Pruitt A. Infections in patients receiving multiple sclerosis disease-modifying therapies. Curr Neurol Neurosci Rep 2017;17:88.

151- Marrie RA. Comorbidity in multiple sclerosis: implications for patient care. Nat Rev Neurol 2017;13:375–82.

152- Thormann A, Magyari M, Koch-Henriksen N, Laursen B, Sørensen PS. Vascular comorbidities in multiple sclerosis: a nationwide study from Denmark. J Neurol 2016;263:2484–93.

153- Smith KA, Burkill S, Hiyoshi A, et al. Comorbid disease burden among MS patients 1968–2012: a Swedish register-based cohort study. Mult Scler 2021;27:268–80.

154- Filippi M, Preziosa P, Barkhof F, Chard DT, De Stefano N, Fox RJ, Gasperini C, Kappos L, Montalban X, Moraal B, Reich DS, Rovira À, Toosy AT, Traboulsee A, Weinshenker BG, Zeydan B, Banwell BL, Rocca MA. Diagnosis of Progressive Multiple Sclerosis From the Imaging Perspective: A Review. JAMA Neurol 2021;78:351-364.

155- Sastre-Garriga J, Pareto D, Battaglini M, Rocca MA, Ciccarelli O, Enzinger C, et al; MAGNIMS study group. MAGNIMS consensus recommendations on the use of brain and spinal cord atrophy measures in clinical practice. Nat Rev Neurol 2020;16:171-182.

156- Sormani, MP, Arnold, DL & De Stefano N. Treatment effect on brain atrophy correlates with treatment effect on disability in multiple sclerosis. Ann Neurol 2014;75:43–49.

157- Midaglia L, Sastre-Garriga J, Montalban X. Clinical monitoring of multiple sclerosis patients by means of digital technology, a field in the midst of a revolution. Rev Neurol 2021;73:210-218.

158- Kappos L, Wolinsky JS, Giovannoni G, Arnold DL, Wang Q, Bernasconi C, et al. Contribution of Relapse-Independent Progression vs Relapse-Associated Worsening to Overall Confirmed Disability Accumulation in Typical Relapsing Multiple Sclerosis in a Pooled Analysis of 2 Randomized Clinical Trials. JAMA Neurol 2020;77:1-9.

159- Confavreux C, Vukusic S, Moreau T, Adeleine P. Relapses and progression of disability in Multiple Sclerosis. N Engl J Med 2000;343:1430-8.

160- Confavreux C, Vukusic S, Adeleine P. Early clinical predictors and progression of irreversible disability in multiple sclerosis: an amnesic process. Brain 2003;126:770-82.

161- Kremenchutzky M, Rice GP, Baskerville J, Wingerchuk DM, Ebers GC. The natural history of multiple sclerosis: a geographically based study 9: observations on the progressive phase of the disease. Brain 2006;129:584–94.

162- Scalfari A, Neuhaus A, Degenhardt A, Rice GP, Muraro PA, Daumer M, Ebers GC. The natural history of multiple sclerosis: a geographically based study 10: relapses and long-term disability. Brain 2010;133:1914-29.

163- Giovannoni G, Popescu V, Wuerfel J, Hellwig K, Iacobeus E, Jensen MB, et al. Smouldering multiple sclerosis: the 'real MS'. Ther Adv Neurol Disord. 2022;15:17562864211066751.

164- Krajnc N, Bsteh G, Berger T. Clinical and Paraclinical Biomarkers and the Hitches to Assess Conversion to Secondary Progressive Multiple Sclerosis: A Systematic Review. Front Neurol 2021;12:666868.

165- Inojosa H, Proschmann U, Akgün K, Ziemssen T. A focus on secondary progressive multiple sclerosis (SPMS): challenges in diagnosis and defnition. J Neurol 2021;268:1210-1221.

166- Portaccio E, Bellinvia A, Fonderico M, Pastò L, Razzolini L, Totaro R, et al; Italian Multiple Sclerosis Register Centers Group. Progression is independent of relapse activity in early multiple sclerosis: a real-life cohort study. Brain. 2022 Mar 24;awac111.

167- IFNB Multiple Sclerosis Study Group. Interferon beta-1b is effective in relapsing-remitting multiple sclerosis. I. Clinical results of a multicenter, randomized, double-blind, placebo-controlled trial. The IFNB Multiple Sclerosis Study Group. Neurology 1993;43:655–61.

168- Jacobs LD, Cookfair DL, Rudick RA, Herndon RM, Richert JR, Salazar AM, et al. Intramuscular interferon beta-1a for disease progression in relapsing multiple sclerosis. The Multiple Sclerosis Collaborative Research Group (MSCRG). Ann Neurol 1996;39:285–94.

169- PRISMS Study Group. Randomised double-blind placebo-controlled study of interferon beta-1a in relapsing/remitting multiple sclerosis. PRISMS (Prevention of Relapses and Disability by Interferon beta-1a Subcutaneously in Multiple Sclerosis) Study Group. Lancet 1998;352:1498–504.

170- Rice GP, Filippi M, Comi G. Cladribine and progressive MS: clinical and MRI outcomes of a multicenter controlled trial. Cladribine MRI Study Group. Neurology 2000;54:1145–55.

171- DeLuca GC, Williams K, Evangelou N, Ebers GC, Esiri MM. The contribution of demyelination to axonal loss in multiple sclerosis. Brain 2006;129:1507–16.

172- Trapp BD, Nave KA. Multiple sclerosis: an immune or neurodegenerative disorder? Annu Rev Neurosci 2008;31:247–69.

173- Lublin FD, Häring DA, Ganjgahi H, Ocampo A, Hatami F, Čuklina J, Aarden P, Dahlke F, Arnold DL, Wiendl H, Chitnis T, Nichols TE, Kieseier BC, Bermel RA. How patients with multiple sclerosis acquire disability. Brain 2022 Feb 1:awac016.

174- Brownlee W, Hardy TA, Fazekas F, Miller DH. Diagnosis of multiple sclerosis: progress and challenges. Lancet 2017;389:1336–46.

175- Kurtzke JF. A new scale for evaluating disability in multiple sclerosis. Neurology (Minneap) 1955;5:580-3.

176- Kurtzke JF. Historical and clinical perspectives of the expanded disability status scale. Neuroepidemiology. 2008;31:1-9.

177- Kurtzke JF. Rating neurologic impairment in multiple sclerosis: An expanded disability status scale (EDSS). Neurology 1983;33:1444-52.

178- Cutter GR, Baier ML, Rudick RA, Cookfair DL, Fischer JS, Petkau J, et al. Development of a multiple sclerosis functional composite as a clinical trial outcome measure. Brain 1999;122:871–882.

179- Fischer JS, Rudick RA, Cutter GR, Reingold SC. The Multiple Sclerosis Functional Composite Measure (MSFC): an integrated approach to MS clinical outcome. National MS Society Clinical Outcomes Assessment Task Force. Mult Scler 1999;5:244–250.

180- Brochet B, Deloire MS, Bonnet M, Salort-Campana E, Ouallet JC, Petry KG, Dousset V. Should SDMT substitute for PASAT in MSFC? A 5-year longitudinal study. Mult Scler J 2008;14:1242–1249.

181- Meyer-Moock S, Feng Y-S, Maeurer M, Dippel F-W, Kohlmann T. Systematic literature review and validity evaluation of the expanded disability status scale (EDSS) and the multiple sclerosis functional composite (MSFC) in patients with multiple sclerosis. BMC Neurol 2014;14:58.

182- Brosseau L, Wolfson C. The inter-rater reliability and construct validity of the functional independence measure for multiple sclerosis subjects. Clin Rehabil 1994,8:107–115.

183- Sharrack B, Hughes R, Soudain S, Dunn G. The psychometric properties of clinical rating scales used in multiple sclerosis. Brain J Neurol 1999;122:141–159.

184- Hobart J, Freeman J, Thompson A. Kurtzke scales revisted: the application of psychometric methods to clinical intuition. Brain 2000;123:1027–1040.

185- Rothwell PM, McDowell Z, Wong CK, Dorman PJ. Doctors and patients don't agree: cross sectional study of patient's and doctors' perceptions and assessments of disability in multiple sclerosis. Br Med J 1997;314:1580–1583.

186- Amato MP, Ponziani G. Quantification of impairment in MS: discussion of the scales in use. Mult Scler 1999;5:216–219.

187- Ebers GC, Heigenhauser L, Daumer M, Lederer C, Noseworthy JH. Disability as an outcome in MS clinical trials. Neurology 2008;71:624–631.

188- Kalkers NF, Bergers L, de Groot V, Lazeron RH, van Walderveen MA, Uitdehaag BM, Polman CH, Barkhof F. Concurrent validity of the MS functional composite using MRI as a biological disease marker. Neurology 2001;56:215–219.

189- Hoogervorst ELJ, Kalkers NF, Cutter GR, Uitdehaag BM, Polman CH. The patient's perception of a (reliable) change in the Multiple Sclerosis Functional Composite. Mult Scler 2004;10:55–60.

190- Miller DM, Rudick RA, Cutter G, Baier M, Fischer JS. Clinical significance of the multiple sclerosis functional composite: relationship to patient-reported quality of life. Arch Neurol 2000;57:1319–1324.

191- Ozakbas S, Cagiran I, Ormeci B, Idiman E. Correlations between multiple sclerosis functional composite, expanded disability status scale and health-related quality of life during and after treatment of relapses in patients with multiple sclerosis. J Neurol Sci 2004;218:3–7.

192- Honarmand K, Akbar N, Kou N, Feinstein A. Predicting employment status in multiple sclerosis patients: the utility of the MS functional composite. J Neurol 2011;258:244–249.

193- Kalkers NF, de Groot V, Lazeron RH, Killestein J, Adèr HJ, Barkhof F, Lankhorst GJ, Polman CH. MS functional composite - relation to disease phenotype and disability strata. Neurology 2000;54:1233–1239.

194- Noseworthy JH, Vandervoort MK, Wong CJ, Ebers GC. Interrater variability with the Expanded Disability Status Scale (EDSS) and Functional Systems (FS) in a multiple sclerosis clinical. Neurology 1990;40:971–975.

195- Francis DA, Bain P, Swan AV, Hughes RA. An assessment of disability rating scales used in multiple sclerosis. Arch Neurol 1991;48:299–301.

196- Goodkin DE, Cookfair D, Wende K, Bourdette D, Pullicino P, Scherokman B, Whitham R. Inter- and intrarater scoring agreement using grades 1.0 to 3.5 of the Kurtzke Expanded Disability Status Scale (EDSS). Neurology 1992;42:859–863.

197- Hohol MJ, Orav EJ, Weiner HL. Disease steps in multiple-sclerosis - a simple approach to evaluate disease progression. Neurology 1995;45:251–255.

198- Cohen JA, Fischer JS, Bolibrush DM, Jak AJ, Kniker JE, Mertz LA, Skaramagas TT, Cutter GR. Intrarater and interrater reliability of the MS functional composite outcome measure. Neurology 2000;54:802–806.

199- Rosti-Otajarvi E, Hämäläinen P, Koivisto K, Hokkanen L. The reliability of the MSFC and its components. Acta Neurol Scand 2008;117:421–427.

200- Hutchinson J, Hutchinson M. The functional limitations profile may be a valid, reliable and sensitive measure of disability in multiple-sclerosis. J Neurol 1995;242:650–657.

201- Hohol MJ, Orav EJ, Weiner HL. Disease steps in multiple sclerosis: a longitudinal study comparing disease steps and EDSS to evaluate disease progression. Mult Scler 1999;5:349–354.

202- Schwid SR, Goodman AD, Apatoff BR, Coyle PK, Jacobs LD, Krupp LB, Miller AE, Wende KE, Brownscheidle CM, New York State Multiple Sclerosis. Are quantitative functional measures more sensitive to worsening MS than traditional measures? Neurology 2000;55:1901–1903.

203- Vaney C, Blaurock H, Gattlen B, Meisels C. Assessing mobility in multiple sclerosis using the rivermead mobility index and gait speed. Clin Rehabil 1996;10:216–226.

204- Ozakbas S, Ormeci B, Idiman E. Utilization of the multiple sclerosis functional composite in follow-up: relationship to disease phenotype, disability and treatment strategies. J Neurol Sci 2005;232:65–69.

205- Patzold T, Schwengelbeck M, Ossege LM, Malin JP, Sindern E. Changes of the MS functional composite and EDSS during and after treatment of relapses with methylprednisolone in patients with multiple sclerosis. Acta Neurol Scand 2002;105:164–168.

206- Cohen JA, Reingold SC, Polman CH, Wolinsky JS. Disability outcome measures in multiple sclerosis clinical trials: current status and future prospects. Lancet Neurol 2012;11:467–476.

207- Pascual AM, Boscá I, Coret F, Escutia M, Bernat A, Casanova B. Evaluation of response of multiple sclerosis (MS) relapse to oral high-dose methylprednisolone: usefulness of MS functional composite and Expanded Disability Status Scale. Eur J Neurol 2008;15:284–288.

208- Francis DA, Bain P, Swan AV, Hughes RA. An assessment of disability rating scales used in multiple sclerosis. Arch Neurol 1991;48:299–301.

209- European Medicines Agency (EMA): Guideline of clinical investigation of medical products for the treatment of Multiple Sclerosis. Draft. EMA/CHMP/771815/2011, Rev 2,20 September 2012. http://www.ema.europa.eu/ docs/en_GB/document_library/Scientific_guideline/2012/10/WC500133438.pdf.

210- Fischer JS, Rudick RA, Cutter GR, Reingold SC. The Multiple Sclerosis Functional Composite Measure (MSFC): an integrated approach to MS clinical outcome. National MS Society Clinical Outcomes Assessment Task Force. Mult Scler 1999;5:244–250.

211- Bosma LVAE, Kragt JJ, Brieva L, Khaleeli Z, Montalban X, Polman CH, Thompson AJ, Tintoré M, Uitdehaag BM. Progression on the multiple sclerosis functional composite in multiple sclerosis: what is the optimal cut-off for the three components? Mult Scler J 2010;16:862–867.

212- Lanska DJ. The history of reflex hammers. Neurology 1989;39:1542–1549.

213- Pearce JM. Early days of the tuning fork. J Neurol Neurosurg Psychiatry 1998;65:728733.

214- Glad SB, Nyland HI, Aarseth JH, Riise T, Myhr KM. Long-term follow-up of benign multiple sclerosis in Hordaland County, Western Norway. Mult Scler 2009;15:942-50.

215- Costelloe L, Thompson A, Walsh C, Tubridy N, Hutchinson M. Long-term clinical relevance of criteria for designating multiple sclerosis as benign after 10 years of disease. J Neurol Neurosurg Psychiatry 2008;79:1245-8.

216- Sayao AL, Devonshire V, Tremlett H. Longitudinal follow-up of "benign" multiple sclerosis at 20 years. Neurology 2007;68:496-500.

217- Pittock SJ, McClelland RL, Mayr WT, Jorgensen NW, Weinshenker BG, Noseworthy J, et al. Clinical implications of benign multiple sclerosis: a 20-year population-based follow-up study. Ann Neuro 2004;56:303-6.

218- D'Amico E, Haase R, Ziemssen T. Review: Patient-reported outcomes in multiple sclerosis care. Mult Scler Relat Disord 2019:33;61-66.

219- Nowinski CJ, Miller DM, Cella D. Evolution of Patient-Reported Outcomes and Their Role in Multiple Sclerosis Clinical Trials. Neurotherapeutics 2017;14:934-944.

220- Alberts JL. Hot Topic 2: e-MS and patients related outcomes. Utilizing mobile technology in the assessment and monitoring of MS patients. Mult Scler 2017;23(Suppl 3):S8-84.

221- Midaglia L, Mulero P, Montalban X, Graves J, Hauser SL, Julian L, et al. Adherence and Satisfaction of Smartphone- and Smartwatch-Based Remote Active Testing and Passive Monitoring in People With Multiple Sclerosis: Nonrandomized Interventional Feasibility Study. J Med Internet Res 2019;21:e14863.

222- Hobart J, Bowen A, Pepper G, et al. International consensus on quality standards for brain health-focused care in multiple sclerosis. Mult Scler 2019;25:1809–1818.

223- Rae-Grant A, Bennett A, Sanders AE, et al. Quality improvement in neurology: Multiple sclerosis quality measures: Executive summary. Neurology 2015;85:1904–1908.

224- Steelman AJ. Infection as an environmental trigger of multiple sclerosis disease exacerbation. Front Immunol 2015;6:520–520.

225- Mills EA, Mirza A and Mao-Draayer Y. Emerging approaches for validating and managing multiple sclerosis relapse. Front Neurol 2017;8:116.

226- Graves JS and Montalban X. Biosensors to monitor MS activity. Mult Scler 2020;26:605–608.

227- Gustavsen S, Olsson A, Søndergaard HB, Andresen SR, Sørensen PS, Sellebjerg F, Oturai A. The association of selected multiple sclerosis symptoms with disability and quality of life: a large Danish self-report survey. BMC Neurol 2021;21:317.

228- Klose, K., Kreimeier, S., Tangermann, U., et al. Patient- and person-reports on healthcare: preferences, outcomes, experiences, and satisfaction - an essay. Health Econ Rev 2016;6:18.

229- McKenna SP. Measuring patient-reported outcomes: moving beyond misplaced common sense to hard science. BMC Med 2011;9:86.

230- D'Amico E, Haase R, Ziemssen T. Review: Patient-reported outcomes in multiple sclerosis care. Mult Scler Relat Disord. 2019;33:61-66.

231- Brichetto G, Zaratin P. Measuring outcomes that matter most to people with multiple sclerosis: the role of patient-reported outcomes. Curr Opin Neurol 2020;33:295-299.

232- Golan D, Sagiv S, Glass-Marmor L, Miller A. Mobile-phone-based e-diary derived patient reported outcomes: Association with clinical disease. activity, psychological status and quality of life of patients with multiple sclerosis. Plos One 2021;16:e0250647.

233- Macaron G, Moss BP, Li H, Baldassari LE, Rao SM, Schindler D, et al. Technology-enabled assessments to enhance multiple sclerosis clinical care and research. Neurol Clin Pract 2020;10:222-231.

234- Matthews PM, Valerie, Block VJ, Leocani L. E-health and multiple sclerosis. Curr Opin Neurol 2020;33:271-276.

235- Cohen JA, Trojano M, Mowry EM, et al. Leveraging real-world data to investigate multiple sclerosis disease behavior, prognosis, and treatment. Mult Scler 2020;26:23–37.

236- Glaser A, Stahmann A, Meissner T, et al. Multiple sclerosis registries in Europe - an updated mapping survey. Mult Scler Relat Disord 2019;27:171–178.

237- Haase R, Schultheiss T, Kempcke R, Thomas K, Ziemssen T. Modern communication technology skills of patients with multiple sclerosis. Mult Scler 2013;19:1240–1241.

238- Boukhvalova AK, Fan O, Weideman AM, et al. Smartphone level test measures disability in several neurological domains for patients with multiple sclerosis. Front Neurol 2019;10: 358–358.

239- Chitnis T, Glanz BI, Gonzalez C, et al. Quantifying neurologic disease using biosensor measurements in-clinic and in free-living settings in multiple sclerosis. NPJ Digit Med 2019;2:123–123.

240- Naslund JA, Marsch LA, McHugo GJ, and Bartels SL. Emerging mHealth and eHealth interventions for serious mental illness: a review of the literature. J. Ment. Health 2015;24:321-32.

241- Gaggioli A and Riva G. From mobile mental health to mobile wellbeing: opportunities and challenges. Stud Health Technol Inform 2013;184:141-7.

242- Alexander S, Peryer G, Gray E, Barkhof F, Chataway J. Wearable technologies to measure clinical outcomes in multiple sclerosis: A scoping review. Mult Scler 2021;27:1643-1656.

243- Scholz M, Haase R, Schriefer D, Voigt I, Ziemssen T. Electronic health interventions in the case of multiple sclerosis: From theory to practice. Brain Sci 2021;11:180.

244- Lavorgna L, Russo A, De Stefano M, Lanzillo R, Esposito S, Moshtari F, Rullani F, Piscopo K, Buonanno D, Morra VB. Health-related coping and social interaction in people with multiple sclerosis supported by a social network: Pilot study with a new methodological approach. Interact. J Med Res 2017;6:e7402.

245- Allam A, Kostova Z, Nakamoto K, Schulz PJ. The effect of social support features and gamification on a Web-based intervention for rheumatoid arthritis patients: Randomized controlled trial. J Med Internet Res 2015;17:e3510.

246- Fernandez-Luque L, Elahi N, Grajales F, 3rd. An analysis of personal medical information disclosed in YouTube videos created by patients with multiple sclerosis. Stud Health Technol Inform 2009;150:292–296.

247- Biogen. Aby App—By above MS. Available online: https://www.abovems.com/en_us/home/ms-support-events/aby-app.html (accessed on 22 July 2021).

248- Settle JR, Maloni HW, Bedra M, Finkelstein J, Zhan M, Wallin MT. Monitoring medication adherence in multiple sclerosis using a novel web-based tool: A pilot study. J Telemed Telecare 2016;22:225–233.

249- De Angelis M, Lavorgna L, Carotenuto A, Petruzzo M, Lanzillo R, Brescia Morra V, Moccia M. Digital Technology in Clinical Trials for Multiple Sclerosis: Systematic Review. J Clin Med 2021;10:2328.

250- Motl RW, Pilutti L, Sandroff BM, Dlugonski D, Sosnoff JJ, Pula JH. Accelerometry as a measure of walking behavior in multiple sclerosis. Acta Neurol Scand 2013;127:384-90.

251- Shammas L, Zentek T, von Haaren B, Schlesinger S, Hey S, Rashid A. Home-based system for physical activity monitoring in patients with multiple sclerosis (Pilot study). Biomed Eng Online 2014;13:10.

252- Block V, Bove R, Zhao C, Garcha P, Graves J, Romeo AR, et al. Association of continuous assessment of step count by remote monitoring with disability progression among adults with multiple sclerosis. JAMA 2019;2:e190570.

253- Flachenecker F, Gabner H, Hannik J, Lee DH, Flachenecker 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.

254- Yousef A, Jonzzon S, Suleiman L, et al. Biosensing in multiple sclerosis. Expert Rev Med Devices 2017;14:901–912.

255- Craig JJ, Bruetsch AP, Lynch SG, et al. Instrumented balance and walking assessments in persons with multiple sclerosis show strong test-retest reliability. J Neuroeng Rehabil 2017;14:43.

256- Greene BR, Healy M, Rutledge S, et al. Quantitative assessment of multiple sclerosis using inertial sensors and the TUG test. Conf Proc IEEE Eng Med Biol Soc 2014;2014:2977–2980.

257- Bamberg SJ, Benbasat AY, Scarborough DM, et al. Gait analysis using a shoe-integrated wireless sensor system. IEEE T Inf Technol Biomed 2008;12:413–423.

258- Spain RI, St George RJ, Salarian A, et al. Bodyworn motion sensors detect balance and gait deficits in people with multiple sclerosis who have normal walking speed. Gait Posture 2012;35:573–578.

259- Reilmann R, Holtbernd F, Bachmann R, et al. Grasping multiple sclerosis: Do quantitative motor assessments provide a link between structure and function. J Neurol 2013;260:407–414.

260- Lopez-Martinez D and Picard R. Wearable technologies for multiple sclerosis: The future role of wearable stress measurement in improving quality of life. In: Proceedings of the second international conference on smart portable, wearable, implantable and disability-oriented devices and systems (SPWID'16), Valencia, 22–26 May 2016.

261- Yu F, Bilberg A and Stenager E. Wireless medical sensor measurements of fatigue in patients with multiple sclerosis. Conf Proc IEEE Eng Med Biol Soc 2010;2010:3763–3767.

262- Kumar P and Lee HJ. Security issues in healthcare applications using wireless medical sensor networks: A survey. Sensors 2012;12:55–91.

263- Smith A. The symbol-digit modalities test: A neuropsychologic test of learning and other cerebral disorders. Seattle: Helmuth J, editor. Learning disorders. Seattle: Special Child Publications;1968.pp.83–91.

264- Smith A. Symbol Digits Modalities Test: Manual. Los Angeles: Western Psychological Services;1982.

265- Rao SM, Losinski G, Mourany L, Schindler D, Mamone B, Reece C, et al. Processing speed test: validation of a self-administered, iPad-based tool for screening cognitive dysfunction in a clinic setting. Mult Scler 2017;23:1929–1937.

266- Berg KO, Wood-Dauphinee SL, Williams JI, Maki B. Measuring balance in the elderly: validation of an instrument. Can J Public Health 1992;83 Suppl 2:S7–S11.

267- Penner IK, Raselli C, Stöcklin M, Opwis K, Kappos L, Calabrese P. The Fatigue Scale for Motor and Cognitive Functions (FSMC): validation of a new instrument to assess multiple sclerosis-related fatigue. Mult Scler 2009;15:1509–1517.

268- Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001;16:606–613.

269- Learmonth YC, Motl RW, Sandroff BM, Pula JH, Cadavid D. Validation of patient determined disease steps (PDDS) scale scores in persons with multiple sclerosis. BMC Neurol 2013;13:37.

270- Hobart J, Cano S. Improving the evaluation of therapeutic interventions in multiple sclerosis: the role of new psychometric methods. Health Technol Assess 2009;13:1–177.

271- Hobart J, Lamping D, Fitzpatrick R, Riazi A, Thompson A. The Multiple Sclerosis Impact Scale (MSIS-29): a new patient-based outcome measure. Brain 2001;124(Pt 5):962–973.

272- Cheng W, Scotland A, Lipsmeier F, Kilchenmann T, Jin L, Schjodt-Eriksen J, et al. Human activity recognition from sensor-based large-scale continuous monitoring of Parkinson's disease patients. Proceedings of the 2nd IEEE/ACM International Conference on Connected Health; July 17-19, 2017; Philadelphia. 2017.

273- Jain S, Sima DM, Ribbens A, Cambron M, Maertens A, Van Hecke W, et al. Automatic segmentation and volumetry of multiple sclerosis brain lesions from MR images. Neuroimage Clin 2015;8:367–375.

274- Rakić M, Vercruyssen S, Van Eyndhoven S, de la Rosa E , Jain S, Van Huffel S, et al. icobrain ms 5.1: combining unsupervised and supervised approaches for improving the detection of multiple sclerosis lesions. Neuroimage Clin 2021;31:102707.

275- Liu Y, Lukas C, Steenwijk M, Daams M, Versteeg A, Duan Y, et al. Multicenter validation of mean upper cervical cord area measurements from head 3D T1-weighted MR imaging in patients with multiple sclerosis. AJNR Am J Neuroradiol 2016;37:749–75

276- Liu Z, Yaldizli Ö, Pardini M, Sethi V, Kearney H, Muhlert N, et al. Cervical cord area measurement using volumetric brain magnetic resonance imaging in multiple sclerosis. Mult Scler Relat Disord 2015;4:52–57.

277- Van Hecke W, Billiet T, Pedrosa de Barros N, et al. Evaluation of spinal cord atrophy in different MS subtypes using icospine. Poster presented at: *37th Congress of the European Committee for Treatment and Research in Multiple Sclerosis*; September 2019, Stockholm, Sweden.

278- Pedrosa de Barros N, Billiet T, Denissen S, et al. Assessing cervical spinal cord atrophy in T1-weighted brain scans using icobrain. Poster presented at: *American Society of Neuroimaging 43rd Annual Meeting*; March 2020, Atlanta, GA, USA.

279- De Leener B, Lévy S, Dupont SM, Fonov VS, Stikov N, Louis Collins D, et al. SCT: Spinal Cord Toolbox, an open-source software for processing spinal cord MRI data. Neuroimage Clin 2017;145:24–43.

280- MacKay DJC. Bayesian Interpolation. Neural Computation 1992;4:415-47.

281- Tipping ME. Sparse bayesian learning and the relevance vector machine. The Journal of Machine Learning Research 2001;1:211-44.

282- Bove R, White CC, Giovannoni G, Glanz B, Golubchikov V, Hujol J, et al. Evaluating more naturalistic outcome measures: a 1-year smartphone study in multiple sclerosis. Neurol Neuroimmunol Neuroinflamm 2015;2:e162.

283- Balto JM, Kinnett-Hopkins DL, Motl RW. Accuracy and precision of smartphone applications and commercially available motion sensors in multiple sclerosis. Mult Scler J Exp Transl Clin 2016;2:2055217316634754.

284- Kos D, Raeymaekers J, Van Remoortel A, D'hooghe MB, Nagels G, D'Haeseleer M, Peeters E, Dams T, Peeters T. Electronic visual analogue scales for pain, fatigue, anxiety and quality of life in people with multiple sclerosis using smartphone and tablet: a reliability and feasibility study. Clin Rehabil 2017;31:1215–1225.

285- D'hooghe M, Van Gassen G, Kos D, Bouquiaux O, Cambron M, Decoo D, et al.Improving fatigue in multiple sclerosis by smartphone-supported energy management: the MS TeleCoach feasibility study. Mult Scler Relat Disord 2018;22:90–96.

286- Goldman MD, LaRocca NG, Rudick RA, et al. Evaluation of multiple sclerosis disability outcome measures using pooled clinical trial data. Neurology 2019;93:e1921–e1931.

287- Pham L, Harris T, Varosanec M, et al. Smartphonebased symbol-digit modalities test reliably captures brain damage in multiple sclerosis. NPJ Digit Med 2021;4:36.

288- van Oirschot P, Heerings M, Wendrich K, et al. Symbol Digit Modalities Test variant in a smartphone app for persons with multiple sclerosis: Validation study. JMIR Mhealth Uhealth 2020;8:e18160.

289- Akhbardeh A, Arjona JK, Krysko KM, et al. Novel MS vital sign: Multi-sensor captures upper and lower limb dysfunction. Ann Clin Transl Neurol 2020;7:288–295.

290- Block VJ, Lizée A, Crabtree-Hartman E, et al. Continuous daily assessment of multiple sclerosis disability using remote step count monitoring. J Neurol 2017;264:316–326.

291- Maillart E, Labauge P, Cohen M, Maarouf A, Vukusic S, Donzé C, et al. MSCopilot, a new multiple sclerosis self-assessment digital solution: results of a comparative study versus standard tests. Eur J Neurol 2020;27:429-36.

292- Pratap A, Grant D, Vegesna A, Tummalacherla M, Cohan S, Deshpande C, et al. Evaluating the utility of smartphonebased sensor assessments in persons with multiple sclerosis in the real-world using an app (elevateMS): observational, prospective pilot digital health study. JMIR Mhealth Uhealth 2020;8:e22108.

293- Chard D and Trip SA. Resolving the clinicoradiological paradox in multiple sclerosis. F1000Res 2017;6:1828.

294- Miller DH, Grossman RI, Reingold SC, et al. The role of magnetic resonance techniques in understanding and managing multiple sclerosis. Brain 1998;121(Pt1):3–24.

295- Radue EW, Barkhof F, Kappos L, et al. Correlation between brain volume loss and clinical and MRI outcomes in multiple sclerosis. Neurology 2015;84:784–793.

296- Creagh AP, Simillion C, Scotland A, et al. Smartphone-based remote assessment of upper extremity function for multiple sclerosis using the Draw a Shape Test. Physiol Meas 2020;41:054002.

297- Creagh AP, Simillion C, Bourke A, et al. Smartphone- and smartwatch-based remote characterisation of ambulation in multiple sclerosis during the two-minute walk test. IEEE J Biomed Health Inform 2020;25:838–849.

298- Bourke AK, Scotland A, Lipsmeier F, et al. Gait Characteristics harvested during a smartphone-based self-administered 2-minute walk test in people with multiple sclerosis: Test–retest reliability and minimum detectable change. Sensors 2020;20:5906.

299- Broen MP, Marsman VA, Kuijf ML, et al. Unraveling the relationship between motor symptoms, affective states and contextual factors in Parkinson's disease: A feasibility study of the experience sampling method. PLoS One 2016;11:e0151195.

300- University of California San Francisco MS-EPIC Team, Cree BA, Hollenbach JA, et al. Silent progression in disease activity-free relapsing multiple sclerosis. Ann Neurol 2019;85:653–666.

301- Members of the MS in the 21st Century Steering Group, Rieckmann P, Centonze D, et al. Unmet needs, burden of treatment, and patient engagement in multiple sclerosis: A combined perspective from the MS in the 21st Century Steering Group. Mult Scler Relat Disord 2018;19:153-160.

302- Achiron A, Chapman J, Tal S, Bercovich E, Gil H. Superior temporal gyrus thickness correlates with cognitive performance in multiple sclerosis. Brain Struct Funct 2013;218:943–950.

303- Bisecco A, Stamenova S, Caiazzo G, et al. Attention and processing speed performance in multiple sclerosis is mostly related to thalamic volume. Brain Imaging Behav 2018;12:20–28.

304- Carmichael OT, Kuller LH, Lopez OL, et al. Cerebral ventricular changes associated with transitions between normal cognitive function, mild cognitive impairment, and dementia. Alzheimer Dis Assoc Disord 2007;21:14–24.

305- Engl C, Tiemann L, Grahl S, et al. Cognitive impairment in early MS: contribution of white matter lesions, deep grey matter atrophy, and cortical atrophy. J Neurol 2020;267:2307–2318.

306- Matías-Guiu JA, Cortés-Martínez A, Montero P, et al. Identification of Cortical and Subcortical Correlates of Cognitive Performance in Multiple Sclerosis Using Voxel-Based Morphometry. Front Neurol 2018;9:920.

307- Zhang X, Zhang F, Huang D, et al. Contribution of Gray and White Matter Abnormalities to Cognitive Impairment in Multiple Sclerosis. Int J Mol Sci 2016;181.

308- Batista S, Zivadinov R, Hoogs M, et al. Basal ganglia, thalamus and neocortical atrophy predicting slowed cognitive processing in multiple sclerosis. J Neurol 2012;259:139–146.

309- Riccitelli GC, Pagani E, Rodegher M, et al. Imaging patterns of gray and white matter abnormalities associated with PASAT and SDMT performance in relapsing-remitting multiple sclerosis. Mult Scler 2019;25:204–216.

310- Goossens N, Janssens L, Caeyenberghs K, Albouy G, Brumagne S. Differences in brain processing of proprioception related to postural control in patients with recurrent non-specific low back pain and healthy controls. Neuroimage Clin 2019;23:101881.

311- Maghzi AH, Revirajan N, Julian LJ, et al. Magnetic resonance imaging correlates of clinical outcomes in early multiple sclerosis. Mult Scler Relat Disord 2014;3:720–727.

312- Bergsland N, Benedict RHB, Dwyer MG, et al. Thalamic Nuclei Volumes and Their Relationships to Neuroperformance in Multiple Sclerosis: A Cross-Sectional Structural MRI Study. J Magn Reson Imaging 2021;53(3):731–739.

313- Daams M, Weiler F, Steenwijk MD, et al. Mean upper cervical cord area (MUCCA) measurement in long-standing multiple sclerosis: relation to brain findings and clinical disability. Mult Scler 2014;20:1860–1865.

314- Merchant H, Harrington DL, Meck WH. Neural basis of the perception and estimation of time. Ann Rev Neurosci 2013;36:313–336.

315- Feys P, Maes F, Nuttin B, et al. Relationship between multiple sclerosis intention tremor severity and lesion load in the brainstem. Neuroreport 2005;16:1379–1382.

316- Rocca MA, Horsfield MA, Sala S, et al. A multicenter assessment of cervical cord atrophy among MS clinical phenotypes. Neurology 2011;76:2096–2102.

317- Asemi A, Ramaseshan K, Burgess A, Diwadkar VA, Bressler SL. Dorsal anterior cingulate cortex modulates supplementary motor area in coordinated unimanual motor behavior. Front Hum Neurosci 2015;9:309.

318- Paus T, Petrides M, Evans AC, Meyer E. Role of the human anterior cingulate cortex in the control of oculomotor, manual, and speech responses: a positron emission tomography study. J Neurophysiol 1993;70:453–469.

319- Pantano P, Mainero C, Caramia F. Functional brain reorganization in multiple sclerosis: evidence from fMRI studies. J Neuroimaging 2006;16:104–114.

320- Rocca MA, Falini A, Colombo B, Scotti G, Comi G, Filippi M. Adaptive functional changes in the cerebral cortex of patients with nondisabling multiple sclerosis correlate with the extent of brain structural damage. Ann Neurol 2002;51:330–339.

321- Fleischer V, Gröger A, Koirala N, et al. Increased structural white and grey matter network connectivity compensates for functional decline in early multiple sclerosis. Mult Scler 2017;23:432–441.

322- Llufriu S, Martinez-Heras E, Solana E, et al. Structural networks involved in attention and executive functions in multiple sclerosis. Neuroimage Clin 2017;13:288–296.

323- Colato E, Stutters J, Tur C, et al. Predicting disability progression and cognitive worsening in multiple sclerosis using patterns of grey matter volumes. J Neurol Neurosurg Psychiatry 2021;92:995–1006.

324- Ellen M. Mowry, Robert A. Bermel, James R. Williams, et al. Harnessing Real-World Data to Inform Decision-Making: Multiple Sclerosis Partners Advancing Technology and Health Solutions (MS PATHS). Front Neurol 2020;11:632.

325- Rao SM, Losinski G, Mourany L, Schindler D, Mamone B, Reece C, Kemeny D, Narayanan S, Miller DM, Bethoux F, Bermel RA, Rudick R, Alberts. Processing speed test: Validation of a self-administered, iPad®-based tool for screening cognitive dysfunction in a clinic setting. Mult Scler 2017;23:1929-1937.

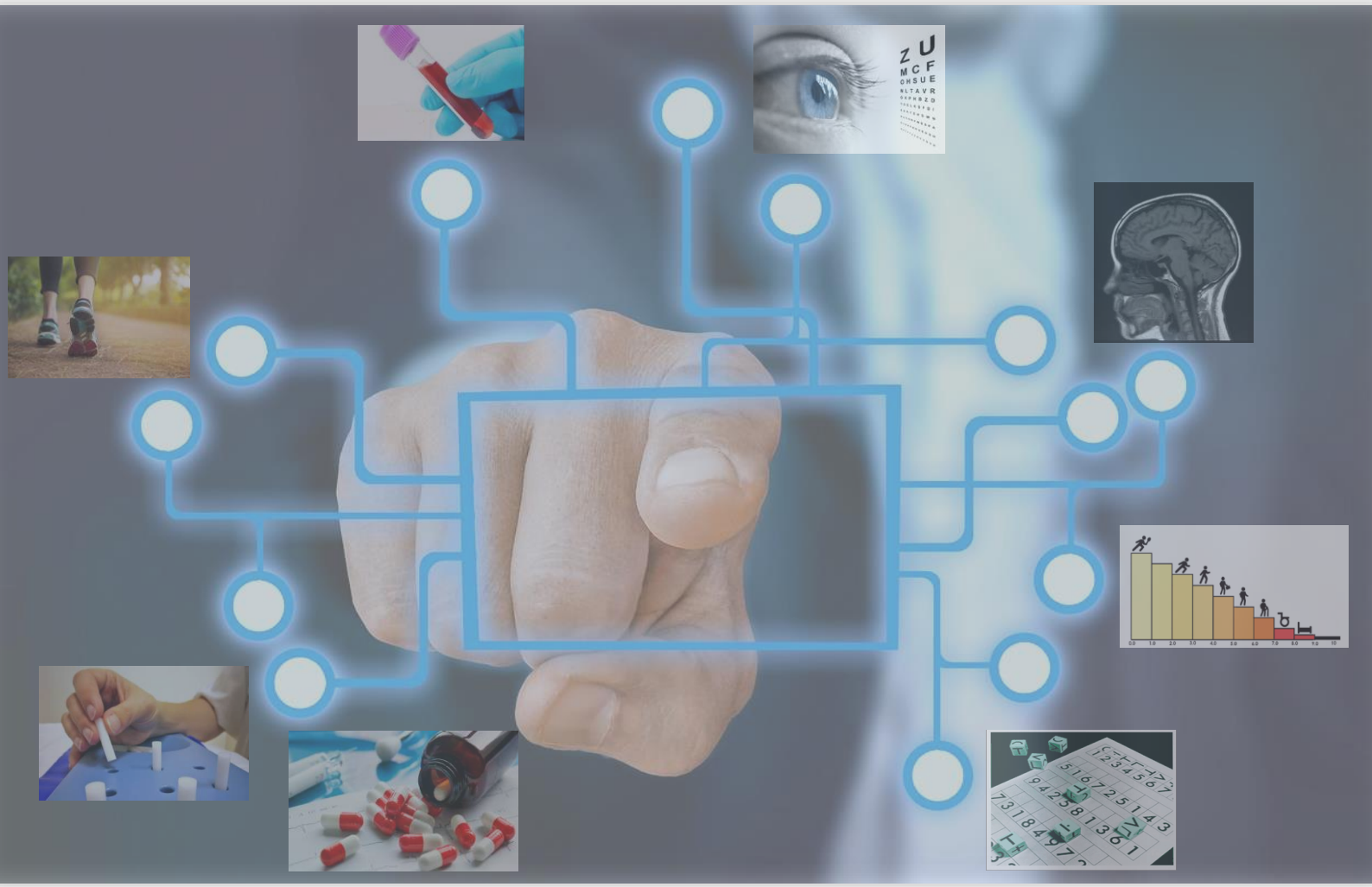
326- Rao SM, Galioto R, Sokolowski M, McGinley M, Freiburger J, Weber M, et al. Multiple Sclerosis Performance Test: validation of self-administered neuroperformance modules. Eur J Neurol 2020;27:878-886.

327- Midaglia L, Carbonel P, Carvajal R, Sastre-Garriga J, Rio J, Comabella M, et al. Defining controversies of benign MS using digital technology. Oral communication presented at: *38th Congress of the European Committee for Treatment and Research in Multiple Sclerosis*; September 2020, Washington, United States.

328- Ranjan Y, Rashid Z, Stewart C, Conde P, Begale M, Verbeeck D, Boettcher S; Hyve, Dobson R, Folarin A; RADAR-CNS Consortium. RADAR-Base: Open Source Mobile Health Platform for Collecting, Monitoring, and Analyzing Data Using Sensors, Wearables, and Mobile Devices. MIR Mhealth Uhealth 2019;7(8):e11734.

329- Sun S, Folarin AA, Ranjan Y, Rashid Z, Conde P, Stewart C, et al; RADAR-CNS Consortium. Using Smartphones and Wearable Devices to Monitor Behavioral Changes During COVID-19. J Med Internet Res 2020;22:e19992.

330- van der Walt A, Butzkueven H, Shin RK, Midaglia L, Capezzuto L, Lindemann M, et al. Developing a Digital Solution for Remote Assessment in Multiple Sclerosis: From Concept to Software as a Medical Device. Brain Sci 2021;11:1247.



APPENDIX

10. APPENDIX

10.1 Materials and Methods / Results

Trial design and participants

For women of childbearing potential, the eligibility criteria also included the agreement to use an acceptable birth control method during the study period. For HCs, the eligibility criteria were restricted to the ability to comply with the study protocol, age 18–55 years and weight 45–110 kg. Recruitment of HCs was planned to help validate the use of smartphone and sensor-based tests for use in MS versus normative data. Both childbearing potential and weight were restricted in this study as it involved physical tasks.

Data transfer

During active tests and passive monitoring, smartphone and smartwatch sensor data were recorded, including accelerometer, gyroscope, compass, Global Positioning System, ambient light (proximity), Wi-Fi access point signal strength and ID, and cell tower signal strength data. Touch data were recorded during the active tests that require pwMS and HCs to touch the screen. Participants' data from smartphone- and smartwatch-based assessments were asymmetrically encrypted and uploaded via the internet to a secure central server, maintained by the sponsor, each time the smartphone connected to Wi-Fi. Participants received instructions on how to connect their smartphone to the internet at home. If there was no Wi-Fi available, data were uploaded during the site visits.

Patient satisfaction

Analysis of the association between individual questions from the patient satisfaction questionnaire and pwMS population characteristics indicated that the following questions were significant: "How satisfied or dissatisfied are you with the guidance from the instructor?" correlated with the T25FW test (Spearman's correlation=–0.35, p = .003); "How satisfied or dissatisfied are you with how

easy it is to use the smartphone?" correlated with SDMT correct responses (Spearman's correlation=0.25, p = .03); and "How would you rate the frequency with which you were asked to do the active tests?" correlated with the T25FW test (Spearman's correlation=0.34, p = .004). Analysis of patient responses to the satisfaction questionnaire indicated that the 2MWT was the component of the Floodlight MS app that participants would most like to avoid. The analysis of this subgroup of pwMS who would prefer to avoid the 2MWT at study completion identified no significant association of this response with a specific pattern of MS disease or demographic features. The only significant associations were age (Spearman's correlation = –0.34, p = .006) and the T25FW test (Spearman's correlation = –0.26, p = .04), and only at the Week 12 assessment of the patient satisfaction questionnaire, where younger participants were more likely to prefer to avoid the 2MWT in the middle of the study.

10.2 Satisfaction questionnaire

Satisfaction questionnaire assessing patients' and healthy volunteers' experience regarding smartphone and smartwatch use and its impact on their daily activities.

Instructions: Please take some time to think about your level of satisfaction or dissatisfaction with the smartphone, the smartwatch and the apps of this clinical study. We are interested in your evaluation of the convenience of the smartphone, the smartwatch and the app over the duration of the study. For each question, please place a single check mark next to the response that most closely corresponds to your own experiences.

1. How satisfied or dissatisfied are you with how easy it is to use the smartphone?
   - Extremely Dissatisfied
   - Dissatisfied
   - Somewhat Dissatisfied
   - Neither Satisfied nor Dissatisfied
   - Somewhat Satisfied
   - Satisfied
   - Extremely Satisfied
2. How satisfied or dissatisfied are you with how easy it is to use the Floodlight MS app?
   - Extremely Dissatisfied
   - Dissatisfied
   - Somewhat Dissatisfied
   - Neither Satisfied nor Dissatisfied
   - Somewhat Satisfied
   - Satisfied
   - Extremely Satisfied
3. How satisfied or dissatisfied are you with how easy it is to put on and take off the smartwatch?
   - Extremely Dissatisfied
   - Dissatisfied

- Somewhat Dissatisfied
- Neither Satisfied nor Dissatisfied
- Somewhat Satisfied
- Satisfied
- Extremely Satisfied

4. How satisfied or dissatisfied are you with how easy it is to use the smartwatch?
   - Extremely Dissatisfied
   - Dissatisfied
   - Somewhat Dissatisfied
   - Neither Satisfied nor Dissatisfied
   - Somewhat Satisfied
   - Satisfied
   - Extremely Satisfied
5. How easy or hard is it to plan when you will do your active tests each time?
   - Very Hard
   - Hard
   - Easy
   - Very Easy
6. How would you rate the frequency with which you were asked to do the active tests?
   - Frequency of tests was unacceptable (too frequent)
   - Frequency of tests was acceptable
7. Taking all things into account, how satisfied or dissatisfied are you with the smartphone, smartwatch and apps?
   - Extremely Dissatisfied
   - Very Dissatisfied
   - Somewhat Dissatisfied
   - Neither Satisfied nor Dissatisfied
   - Somewhat Satisfied
   - Very Satisfied

- Extremely Satisfied

8. To what extent do you agree with the following statement? I would like to continue using the Floodlight MS app to understand my MS better and improve my disease management.
   - Completely agree
   - Somewhat agree
   - Somewhat disagree
   - Completely disagree
   - Not sure
9. To what extent do you agree with the following statement? In future I would like to be able to see the results of my tests straight after I've completed them and monitor over time.
   - Completely agree
   - Somewhat agree
   - Somewhat disagree
   - Completely disagree
   - Not sure
10. How satisfied or dissatisfied are you with the provided information/training materials?
    - Extremely Dissatisfied
    - Dissatisfied
    - Somewhat Dissatisfied
    - Neither Satisfied nor Dissatisfied
    - Somewhat Satisfied
    - Satisfied
    - Extremely Satisfied
11. How satisfied or dissatisfied are you with the guidance from the instructor?
    - Extremely Dissatisfied
    - Dissatisfied
    - Somewhat Dissatisfied
    - Neither Satisfied nor Dissatisfied
    - Somewhat Satisfied

- Satisfied
- Extremely Satisfied

12. Does the use of smartphone, smartwatch and the active tests have any impact on your daily activities?
    - None
    - Minimal
    - Acceptable
    - Substantial
    - Unacceptable
13. If you could avoid one of the components of Floodlight monitoring, which one would that be (check only one)?
    - None
    - Cognitive Test
    - Squeeze a Shape
    - Draw a Shape
    - Balance Test
    - U-Turn Test
    - Two-Minute Walk Test
    - Passive monitoring
14. Additional feedback/comments that you would like to share:

10.3 Participants discontinuing from study and excluded from adherence analysis.

Cohort	Participant ID	Reason for withdrawing from study	Excluded from adherence analysis
PwMS	1015	Monitoring frequency-related burden	
PwMS	1019	The participant did not have enough time to implement all study requirements	
HC	1025	Due to personal reasons, the participant was not able to continue with the study	
HC	1035	Lack of availability	X
HC	1036	Lack of availability because of new job	X
HC	1047	Privacy of data. The participant was not comfortable with the fact that the smartphone had GPS	
HC	1052	Participant's personal reasons	
PwMS	1061	Adverse event	X
HC	1062	Lack of availability to come to the visits	X
HC	1076	The participant started to work and did not have availability to come to CEMCAT for the final visit	X
PwMS	2001	Walking task was too exhausting. Concerned about accuracy of data (due to walking task)	
HC	2004	The participant did not want to complete tasks	
PwMS	2006	Travel time to get to study site	
PwMS	2009	Lack of time/interest	
HC	2014	Lives 4 hours away. Was unable to come for last visit	X

Legend: CEMCAT: Multiple Sclerosis Centre of Catalonia; GPS: Global Positioning System; HC: healthy control; PwMS: people with multiple sclerosis

10.4 PUBLICATIONS

10.4.1 Publication N1

Clinical monitoring of multiple sclerosis patients by means of digital technology, a field in the midst of a revolution. Midaglia L, Sastre-Garriga J, Montalban X. Rev Neurol. 2021 Sep 1;73(6):210-218. doi: 10.33588/rn.7306.2021136.

Monitorización clínica del paciente con esclerosis múltiple a través de la tecnología digital, un campo en plena revolución

Luciana Midaglia, Jaume Sastre-Garriga, Xavier Montalban

Departamento de Neurología e Inmunología. Centre d'Esclerosi Múltiple de Catalunya (Cemcat). Hospital Universitari Vall d'Hebron-Universitat Autònoma de Barcelona. Barcelona, España.

Correspondencia:
Dra. Luciana Midaglia. Departamento de Neurología e Inmunología. Centre d'Esclerosi Múltiple de Catalunya (Cemcat). Edifici Cemcat. P.º Vall d'Hebron, 119-129. E-08035 Barcelona.

E-mail:
lmidaglia@cem-cat.org

Conflicto de intereses:
L.M. declara no presentar conflicto de intereses. J.S.G. ha recibido compensación por participar en juntas asesoras, honorarios por conferencias y gastos de viaje para reuniones científicas, servicios de consultoría o apoyo a la investigación de Celgene, Novartis, Biogen, Teva, Merck, Almirall y Genzyme. X.M. ha recibido honorarios por conferencias y gastos de viaje para participar en reuniones científicas y ha sido miembro del comité directivo de ensayos clínicos o participado en juntas asesoras de ensayos clínicos en el pasado con Actelion, Almirall, Bayer, Biogen, Celgene, Genzyme, Hoffmann-La Roche, Novartis, Oryzon Genomics, Sanofi-Genzyme y Teva Pharmaceutical.

Aceptado tras revisión externa:
28.07.21.

Cómo citar este artículo:
Midaglia L, Sastre-Garriga J, Montalban X. Monitorización clínica del paciente con esclerosis múltiple a través de la tecnología digital, un campo en plena revolución. Rev Neurol 2021; 73: 210-8. doi: 10.33588/rn.7306.2021136.

Introducción. A pesar de los grandes avances acontecidos en el área del diagnóstico y el tratamiento de la esclerosis múltiple (EM), pocos cambios se han gestado respecto al seguimiento clínico. La escasez de tiempo y espacio en la práctica clínica dificulta la valoración de síntomas invisibles y ciertos síntomas motores, como la destreza manual y la capacidad de marcha, que presentan un claro impacto en la situación funcional del paciente.

Objetivo. Revisar el papel potencial de las herramientas tecnológicas en la monitorización clínica del paciente con EM.

Desarrollo. Se realizó una búsqueda bibliográfica a través de PubMed, seleccionando los estudios centrados en biosensores y herramientas digitales destinadas a la evaluación de la situación funcional general, y de aspectos concretos de la enfermedad o de determinados sistemas funcionales.

Resultados. Diferentes herramientas digitales en formato de biosensores, aplicaciones móviles o web, tanto de uso remoto como hospitalario, autocumplimentadas o administradas por el personal sanitario, parecen ofrecer una visión más 'completa y real' de la situación funcional de los pacientes. Algunos estudios han demostrado que la tecnología digital es capaz de detectar la progresión subclínica de la discapacidad, que las pruebas tradicionales, incluyendo la Escala expandida del estado de discapacidad, no consiguen reflejar, lo que favorece la adopción de medidas y acciones terapéuticas apropiadas de forma temprana y personalizada.

Conclusiones. Las herramientas digitales, capaces de brindar información clínica amplia y detallada, podrían ocupar un papel importante en la toma de decisiones y el seguimiento clínico del paciente afecto de EM.

Palabras clave. Dispositivos portátiles. e-Salud. Esclerosis múltiple. Herramientas digitales. Monitorización clínica. Tecnología digital.

Introducción

En los últimos años se han producido grandes avances en cuanto al proceso diagnóstico [1-8] y al abordaje terapéutico en el campo de la esclerosis múltiple (EM) [9-11].

Sin embargo, respecto al seguimiento clínico, desde el año 1983, cuando Kurtzke introdujo la Escala expandida del estado de discapacidad (EDSS) [12], pocos cambios en la práctica clínica se han gestado a este nivel. Si bien la EDSS consolidó su papel como herramienta de referencia a la hora de evaluar la situación funcional de los pacientes con EM, en una encuesta realizada recientemente, sólo un 25% de los neurólogos especialistas en el área respondió que la utilizaba de forma rutinaria, porcentaje que se redujo al 17% al preguntar específicamente si incluían la medición de la capacidad de marcha [13]. Entre las diferentes desventajas señaladas a lo largo del tiempo, cabe mencionar la baja fiabilidad interobservador, la exigencia de un consumo importante de tiempo y espacio, el hecho de reflejar un momento puntual de la situación funcional del paciente, que muchas veces no se correlaciona con su rendimiento físico a lo largo del día, y el evaluar fundamentalmente aspectos físicos y de capacidad de la marcha, sin incluir la valoración de áreas 'no visibles' de la enfermedad, como la cognición, el estado emocional, la fatiga, etc. [14,15]. Con la intención de mejorar la valoración funcional del paciente con EM, en 1994 se diseñó el *Multiple Sclerosis Functional Composite* (MSFC) como herramienta multidimensional, integrando test de velocidad de la marcha, destreza manual, velocidad de procesamiento de información y contraste visual. Si bien su correlación con la EDSS se ha estudiado ampliamente, su principal desventaja, además de precisar personal técnico debidamente entrenado y un tiempo medio de 20 minutos para su administración, es la utilización de *Z-scores* como medida de resultados, muy poco familiar para la mayoría de los clínicos. Su uso actualmente se limita, básicamente, al área de la investigación [16,17].

Basándonos en ello, podemos afirmar que, a pesar del paso de los años, continúa siendo difícil monitorizar con precisión la situación clínica de los pacientes. Concretamente, resulta complejo precisar el momento del inicio de una fase secundaria progresiva de la enfermedad, la presencia de progresión/empeoramiento en los fenotipos progresivos, así como determinar con exactitud la aparición de ciertos brotes o recaídas en las formas recurrentes que muchas veces se malinterpretan como fluctuaciones clínicas [18,19]. En este sentido, la tecnología digital podría ocupar un papel fundamental, brindando información detallada en el tiempo y el espacio sobre la experiencia real de la vida diaria de los pacientes, capturada de forma continua e independiente del evaluador, mejorando la atención personalizada [20] (Figura).

El presente trabajo tiene como finalidad abordar algunas de las potenciales herramientas tecnológicas que podrían tener su lugar en la práctica clínica, intentando solventar las limitaciones que actualmente persisten a la hora de monitorizar clínicamente al paciente con EM.

Figura. Clásicamente, el cuidado y el seguimiento del paciente tenían lugar en el centro sanitario de referencia. Sin embargo, la llegada de la tecnología digital podría desplazar el foco principal de la atención médica y la monitorización clínica al propio paciente, otorgándole un papel activo y protagónico en el cuidado de su propia salud. Figura modificada de [21].

Desarrollo

Para el desarrollo del contenido de este trabajo, se realizó una búsqueda bibliográfica en diciembre de 2020, mediante la plataforma científica PubMed, utilizando los siguientes términos MeSH: *'digital technology', 'digital tools', 'e-health', 'wearable devices'* en combinación con el término *'multiple sclerosis'*.

Se aplicaron los siguientes criterios de selección:

- *Diseño del estudio.* Se incluyeron trabajos de revisión, ensayos clínicos, estudios observacionales o series de casos.
- *Asociaciones evaluadas.* Se incluyeron estudios centrados en las siguientes asociaciones:
  - Biosensores utilizados en la monitorización de la EM.
  - Herramientas digitales de uso exclusivo en la monitorización de pacientes con EM destinadas a la evaluación de la situación funcional general, así como herramientas diseñadas para la evaluación de aspectos concretos de la enfermedad o de determinados sistemas funcionales.
- *Idioma.* Se consideraron artículos escritos en inglés y español.

Presentamos una descripción narrativa de los estudios seleccionados estructurados según los siguientes temas principales:

- Estudios que describen herramientas digitales (biosensores) utilizadas en la monitorización de la EM, aunque no específicas de la enfermedad.
- Estudios que describen herramientas digitales de uso exclusivo en pacientes con EM destinadas a la valoración de la discapacidad en general, así como a la monitorización de determinados aspectos o sistemas funcionales de la enfermedad, cuyo uso se limita al ámbito hospitalario. En estos casos, la información clínica puede ser recogida por el personal sanitario o bien proporcionada directamente por el propio paciente.
- Estudios que describen herramientas digitales para la monitorización remota o a distancia de determinados síntomas o aspectos de la enfermedad, con la información clínica directamente proporcionada por los pacientes.
- Estudios que describen herramientas digitales para la monitorización remota y, a su vez, continuada de la situación funcional general de los pacientes, con la información clínica directamente proporcionada por los pacientes.

Resultados

Un total de 51 artículos resultaron de la búsqueda *'digital technology and multiple sclerosis'*, 21 artículos de la combinación *'digital tools and multiple*

Tabla I. Clasificación de las herramientas digitales para la monitorización clínica del paciente con esclerosis múltiple.

Herramientas digitales para la monitorización clínica del paciente con esclerosis múltiple			
Biosensores	**De uso periódico en el ámbito hospitalario**	**De uso periódico a distancia o remoto**	**De uso remoto y continuado en el tiempo**
Acelerómetros Sensores manuales tipo guantes Giroscopios Actígrafos Acelerómetros triaxiales Medidores de la temperatura corporal Electrocardiógrafos	Valoración de la discapacidad en general	Valoración de la discapacidad en general	Valoración de la discapacidad en general
	• MSPT • MSCopilot	• MS monitor	• FLOODLIGHT • RADAR-CNS • ElevateMS
	Valoración de sistemas funcionales o aspectos concretos de la enfermedad	Valoración de sistemas funcionales o aspectos concretos de la enfermedad	
	• ASSESS MS • TabDMuS	• SymTrac • CSI	

CSI: Cognitive Stability Index; MSPT: Multiple Sclerosis Performance Test; RADAR-CNS: Remote Assessment of Disease and Relapse – Central Nervous System; TabDMuS: Tablet-based Data capture in Multiple Sclerosis.

sclerosis', 33 asociados a la búsqueda '*e-health and multiple sclerosis*' y, finalmente, 50 artículos en relación con los términos '*wearable devices and multiple sclerosis*'. Muchos artículos se repetían en las diferentes búsquedas. La mayoría de los artículos encontrados hacía referencia a temas no relacionados directamente con el objetivo principal de nuestra revisión. Por ejemplo, varios artículos correspondían a revisiones generales de 'la salud digital' aplicada en neurología o en esclerosis múltiple sin mencionar herramientas digitales en concreto, y muchos otros artículos hacían referencia al uso de herramientas digitales para la autogestión y el manejo de determinados síntomas de la enfermedad (como la espasticidad y fatiga), para la rehabilitación física y cognitiva, para la educación y formación de pacientes, así como para favorecer la adhesión a los tratamientos. Los artículos finalmente seleccionados aplicando los criterios de selección descritos en la sección de desarrollo se revisan detalladamente en los siguientes apartados (Tabla I).

Herramientas digitales (biosensores) utilizadas en la monitorización de la esclerosis múltiple, aunque no específicas de la enfermedad

Los acelerómetros son biosensores útiles para cuantificar los pasos, la posición del cuerpo y los patrones groseros de actividad (como estar de pie, sentado o en posición de decúbito) de las personas en general. En el área de la EM, ofrecen una medida confiable y válida para registrar la actividad/movilidad en la vida cotidiana de los pacientes [22,23]. La fiabilidad y la precisión del recuento de pasos con acelerómetros se han correlacionado con el recuento de pasos reales, particularmente en pacientes con discapacidad leve, así como también con la EDSS, aunque sólo en pacientes con valores de la EDSS entre 0 y 6,5 [24]. Sin embargo, se necesitan estudios a largo plazo para determinar si los acelerómetros son útiles para detectar el incremento de la discapacidad o la progresión de la enfermedad. En este sentido, un estudio piloto que incluyó a 11 pacientes con EM y otro más reciente con 95 pacientes mostraron durante el periodo de seguimiento de un año que un acelerómetro podría identificar las fluctuaciones y el deterioro de la situación funcional de los pacientes con mayor precisión y precocidad que el uso de las pruebas tradicionales, incluyendo la EDSS, aun cuando éstas se mantuvieran estables [25,26].

Otro estudio destinado a evaluar cambios precoces en la marcha a través de biosensores incluyó a 102 pacientes y a 22 controles sanos. Se les pidió que realizaran la prueba de la marcha de los 25 pies dos veces a una velocidad 'autoseleccionada', seguida de dos veces a una velocidad máxima. El análisis de subgrupos entre controles y pacientes mostró diferencias significativas en la longitud media del paso, la velocidad de la marcha, el tiempo de apoyo del pie y el tiempo de balanceo. Las diferencias entre los subgrupos según la EDSS fueron más pronunciadas en las pruebas con velocidad máxima de la marcha [27].

Los sensores en formato de guante son otro tipo de biosensores diseñados para evaluar la destreza manual. Un estudio que incluyó a 40 pacientes con EM y a 80 controles sanos demostró que esta herramienta era capaz de diferenciar entre los participantes de ambos grupos y, a su vez, mostró correlación con las medidas obtenidas mediante otras escalas de discapacidad como la EDSS y el MSFC [28].

También los acelerómetros y giroscopios, insertados en la muñeca, se han estudiado como herramientas de evaluación/medición de la funcionalidad del miembro superior en un grupo de sujetos sanos (n = 12) y personas con EM (n = 28). Los participantes debían realizar el programa de ejercicios *Action Reasearch Arm Test*, que incluye la ejecución de diferentes movimientos de agarre, pellizco y movimientos más groseros. Se observó que las tareas del *Action Reasearch Arm Test* ejecutadas por pacientes con EM fueron significativamente más lentas (aumento de la duración del 70%) y bruscas (aumento de sacudidas del 16%) con respecto a los controles [29].

Otros biosensores, como actígrafos, acelerómetros triaxiales, medidores de temperatura corporal y frecuencia cardíaca, aunque en fases más prematuras, también son objeto de estudio con el fin de obtener una información más completa y detallada de las diferentes funciones corporales y, por tanto, de la situación funcional del paciente con EM [22,23].

Herramientas digitales de uso exclusivo en la esclerosis múltiple para la monitorización de la discapacidad en general o de determinados sistemas funcionales, cuyo uso se limita al ámbito hospitalario

Para la valoración de la discapacidad en general

El *Multiple Sclerosis Performance Test* (MSPT) es una plataforma electrónica en formato de iPad® diseñada para ser implementada en el entorno clínico con una supervisión mínima, que integra los cuatro test tradicionales de función motora, cognitiva y visual, pertenecientes al MSFC en versión digital. Por tanto, captura datos de velocidad de la marcha, de contraste visual, de destreza manual y de velocidad de procesamiento de la información. Además, incluye el cuestionario de calidad de vida en enfermedades neurológicas, Neuro-Qol [30]. El MSPT se administra normalmente al inicio, cuando el paciente llega al centro, lo que permite al clínico disponer de los resultados el mismo día de la visita. Un estudio de validación que incluyó a 51 pacientes con EM y 49 controles sanos encontró que los datos obtenidos mediante la plataforma MSPT eran altamente reproducibles y se correlacionaron de manera significativa con los resultados de las pruebas tradicionales (MSFC) administradas por los técnicos. A su vez, dicha herramienta digital consiguió diferenciar la población de pacientes con EM de los controles sanos, y, dentro del grupo de pacientes, entre aquellos con discapacidad leve de aquellos con discapacidad alta [31]. Actualmente, como parte del proyecto internacional *Multiple Sclerosis Partners Advancing Technology Health Solutions*, existen 10 centros (Cleveland Clinic, Cleveland Clinic Nevada, New York University, Ohio Health Research Institute, Johns Hopkins University, University of Rochester y Washington University en St. Louis, en Estados Unidos; University Hospital of Giessen y Marburg, y University Hospital Carl Gustav Carus, en Alemania; y Centro de Esclerosis Múltiple de Catalunya, Hospital Universitario Vall d'Hebron, en España) que utilizan este dispositivo como una manera fácil de recopilar datos de forma estandarizada con el fin de aumentar la participación de los pacientes en el proceso de la atención médica y favorecer el uso de datos cuantitativos a la hora de tomar decisiones clínicas. Además, la recopilación y el análisis rigurosos de los datos del mundo real pueden mejorar nuestra comprensión de la enfermedad y acelerar el desarrollo de la medicina personalizada en el área de la EM [32].

MSCopilot es una aplicación móvil diseñada para la autoevaluación de pacientes con EM, que también combina las cuatro pruebas tradicionales: test de la marcha, destreza manual, cognición y vía visual, en versión electrónica. Un estudio francés, multicéntrico, abierto, aleatorizado y controlado, incluyó a 116 pacientes con EM y 69 controles sanos con el objetivo de validar la herramienta MSCopilot en comparación con el MSFC. Tanto los test de la aplicación móvil como los test tradicionales del MSFC se realizaron en el ámbito de la consulta hospitalaria. El estudio concluyó que MSCopilot no era inferior al MSFC (área bajo la curva, 0,92 y 0,89, respectivamente; p = 0,3) a la hora de evaluar la discapacidad en pacientes con EM remitente-recurrente o progresiva, confirmando la fiabilidad y la facilidad de uso en la práctica clínica de dicha herramienta digital para el seguimiento de pacientes con EM [33].

Para la valoración de determinados aspectos de la enfermedad o sistemas funcionales

ASSESS MS es un sistema de registro de imágenes en formato vídeo que permite la captura y el análisis del movimiento de forma consistente y detallada

por parte del personal sanitario, con el fin de automatizar la clasificación de la disfunción motora en pacientes con EM. Se trata de un proyecto puesto en marcha por un equipo multidisciplinar de investigadores del Reino Unido, Suiza y Holanda. En un estudio inicial, prospectivo, que contó con la colaboración de seis neurólogos y seis enfermeros especializados en EM, ASSESS MS se utilizó para registrar un conjunto predefinido de movimientos estándar en 51 pacientes voluntarios afectos de EM, generando datos de calidad suficiente para el análisis clínico. A su vez, la herramienta demostró mejorar la precisión en el momento de calcular el valor de la EDSS, ya que redujo el componente de subjetividad, por ejemplo, al evaluar el grado de dismetría o temblor durante la prueba dedo-nariz. La aceptación de ASSESS MS fue buena tanto por parte de los pacientes como por los profesionales de la salud, quienes indicaron que la herramienta era efectiva, eficiente y fácil de usar [34].

Tablet-based Data capture in Multiple Sclerosis (TaDiMuS) es una herramienta digital en formato de iPad© que captura información clínica proporcionada, a través de cuestionarios, directamente por los propios pacientes. Un estudio piloto, llevado a cabo en Australia, contó con la participación de 157 pacientes afectos de EM. Éstos debían completar de manera electrónica en la sala de espera, antes de ser atendidos, las subescalas del cuestionario de calidad de vida *MS Quality of Life Inventory* relacionadas con el control de los esfínteres urinario y fecal. Puntuaciones iguales o mayores a 2 generaban automáticamente un aviso a la enfermera clínica de la consulta para que ésta activase la atención médica necesaria. El estudio confirmó la validez de TaDiMuS como una herramienta de detección precoz de alteraciones esfinterianas, ofreciendo a los médicos un método eficiente, sensible y factible para su valoración y tratamiento tempranos [35].

Herramientas digitales para la monitorización a distancia de determinados síntomas o aspectos concretos de la enfermedad

MS Monitor es una web interactiva diseñada para que los pacientes lleven a cabo su propia monitorización clínica y multidisciplinaria de la enfermedad, mejorando el autocontrol y la autogestión de determinados síntomas, y favoreciendo la atención médica en general. Integra cuestionarios para evaluar el impacto de la EM y la calidad de vida, la fatiga, la ansiedad y la depresión, así como para vigilar la adhesión a los tratamientos. También dispone de diarios donde los pacientes pueden registrar la actividad física, los períodos de descanso, los síntomas urológicos, la micción y la ingesta de líquidos. Finalmente, otorga a los pacientes la posibilidad de realizar e-consultas a su médico. Un estudio piloto holandés, que incluyó a un total de 55 pacientes, evaluó los cambios/efectos a nivel de la fatiga utilizando dicha web. Se observó que la educación sanitaria proporcionada a través de esta herramienta electrónica era capaz de brindar a los pacientes con EM una mejor comprensión sobre la fatiga y los posibles factores relacionados con los niveles de percepción de ésta, como el bienestar subjetivo general, la medicación y la adhesión. Así, se consiguió de manera satisfactoria que los pacientes adoptasen estrategias beneficiosas para el autocontrol y la gestión de la fatiga [36].

SymTrac es una aplicación diseñada para que los pacientes puedan registrar sus síntomas y situación de bienestar a lo largo del tiempo. Esto les ayuda a identificar y, por tanto, a registrar detalladamente las posibles recaídas de la enfermedad. A su vez, la información recogida puede ser enviada al neurólogo por correo electrónico o ser utilizada en el momento de la visita médica y así evitar el olvido de algún detalle relevante de lo ocurrido. De esta forma, se evitaría el consecuente mal manejo terapéutico, es decir, que, en ocasiones, por olvidar mencionar determinados síntomas, el paciente sea mantenido bajo un tratamiento modificador de la enfermedad que pueda ser no del todo efectivo [37,38].

Cognitive Stability Index (CSI) es una prueba computarizada accesible mediante internet, de 30 minutos de duración, autocumplimentada por el paciente, que proporciona medidas objetivas y automatizadas de atención, memoria, velocidad de respuesta y velocidad de procesamiento de información para la evaluación inicial de la función cognitiva. Un estudio americano evaluó la utilidad del CSI a la hora de medir la afectación cognitiva de los pacientes. El estudio contó con la participación de 40 pacientes afectos de EM que presentaban quejas subjetivas de fallos de memoria y demostró que el CSI era tan específico y más sensible que otras herramientas tradicionalmente utilizadas para evaluar la función cognitiva en la EM, incluyendo una batería de pruebas neuropsicológicas, así como el *Paced Auditory Serial* [39].

Herramientas digitales para la monitorización remota y, a su vez, continuada de la situación funcional general de los pacientes

FLOODLIGHT es un estudio piloto prospectivo diseñado para evaluar la viabilidad de mediciones re-

Tabla II. Test activos y pasivos utilizados en el ensayo clínico FLOODLIGHT.

	Test activos							Monitorización pasiva
	Ánimo	Impacto de la EM	Cognición	Destreza manual	Marcha		Equilibrio	Patrón de la marcha
Aplicación FLOODLIGHT	Pregunta sobre el estado de ánimo	MSIS-29 (electrónica)	Versión digital del SDMT	Dibujar una forma y pellizcar una figura	Marcha durante dos minutos	Prueba de cinco vueltas en U	Prueba del equilibrio estático	Registro pasivo del comportamiento y patrón de la marcha
Test tradicionales en la visita médica		MSIS-29 (en papel)	Versión oral del SDMT	9-HPT	25FWT		*Berg Balance Test*	

9-HPT: *9 Hole Peg Test*; 25FWT: prueba de la marcha de los 25 pies; EM: esclerosis múltiple; MSIS-29: escala de impacto de la esclerosis múltiple-29; SDMT: *Symbol Digit Modalities Test*.

motas utilizando teléfonos y relojes inteligentes en pacientes con EM y controles sanos, y su correlación con las medidas tradicionales de monitorización clínica. Participaron dos centros, el Centro de Esclerosis Múltiple de Catalunya, Hospital Universitario Vall d'Hebron, Barcelona, y la Universidad de California, San Francisco. Se reclutó a 76 pacientes y 25 controles sanos. Los participantes debían utilizar los dispositivos electrónicos de forma diaria durante un período de 24 semanas. Dichos dispositivos contenían aplicaciones con cuestionarios y test que los participantes debían realizar de forma periódica (pregunta diaria sobre el estado de ánimo, escala de impacto de la esclerosis múltiple: 29, prueba del equilibrio estático, prueba cognitiva-versión digital del *Symbol Digit Modalities Test*, prueba de la marcha durante dos minutos, prueba cinco 5 vueltas en U, pruebas de destreza manual consistentes en dibujar una forma y pellizcar una figura), así como sensores que registraban pasivamente la información de movimiento (comportamiento y patrón de la marcha) durante la vida diaria. Por otra parte, todos los participantes debían acudir a las visitas programadas en el centro (basal, al mes 3 y al 6 del estudio) para realizar las pruebas tradicionales de cognición *(Symbol Digit Modalities Test)*, destreza manual *(9 Hole Peg Test)*, equilibrio *(Berg Balance Test)*, y marcha (prueba de la marcha de los 25 pies) administradas por el clínico, así como también, sólo en el caso de los pacientes, realizar la EDSS (Tabla II). El estudio demostró que la adhesión por parte de los pacientes y controles sanos fue satisfactoria y la mayoría de ellos respondió que el impacto del uso de los dispositivos electrónicos en sus actividades diarias fue entre nulo y aceptable. Los resultados de los test activos y la monitorización pasiva recogidos de manera digital mediante los dispositivos electrónicos se correlacionaron de manera adecuada con los resultados de las pruebas habituales de cognición, destreza manual, equilibrio y marcha llevadas a cabo en el propio centro durante las visitas programadas [40].

El *Remote Assessment of Disease and Relapse – Central Nervous System* es un proyecto europeo que tiene como finalidad explorar el potencial impacto de los datos clínicos recogidos por los propios pacientes a través de diferentes dispositivos electrónicos (Tabla III) a la hora de tomar decisiones médicas/terapéuticas en la práctica clínica habitual. El proyecto incluye la participación de pacientes afectos de EM, así como también de depresión y epilepsia. En el grupo de estudio de pacientes con EM están participando tres centros, el Centro de Esclerosis Múltiple de Catalunya, Hospital Universitario Vall d'Hebron, Barcelona, el Hospital San Rafael de Milán, y la Universidad de Copenhague y Rigshospitalet. El estudio pretende determinar la utilidad de la tecnología digital para detectar cambios en el estado de ánimo y en la marcha de pacientes con EM, así como identificar predictores de recaídas o progresión. Para ello se diseñaron dos subestudios: 'de depresión', enfocado en la aparición de brotes, la evaluación de la actividad y participación social, el patrón del sueño, la cognición y el patrón de la voz, para el cual se prevé incluir a un total de 240 pacientes con diagnóstico de EM reciente; y 'de discapacidad', enfocado en la aparición de brotes, la progresión de la discapacidad (EDSS), la evaluación de la marcha y el equilibrio, la fatiga, la cognición y el patrón de la voz, para el cual se

Tabla III. Herramientas digitales utilizadas en el estudio RADAR-CNS junto con sus contenidos y funciones.

	Funciones/recogida de datos	Participación del paciente
RADAR-CNS app	Contiene test y cuestionarios para evaluar la fatiga, la cognición y otras áreas clínicas de la EM, así como para evaluar el habla/patrón de voz	Activa
Fitbit	Evalúa el movimiento, las actividades diarias y las características del sueño	Pasiva
Sensores del móvil inteligente	Recoge información sobre el ruido y la luz ambiental, la duración de las llamadas telefónicas y la cantidad de mensajes de texto enviados. Dispone de un GPS y un giroscopio para el control de las distancias y el patrón de marcha	
eFaros180	Recoge información de equilibrio al realizar la maniobra de Romberg y la marcha en tándem, así como de la variabilidad de la marcha, el recuento de pasos y las caídas. A su vez, permite el registro electrocardiográfico	

RADAR-CNS: *Remote Assessment of Disease and Relapse – Central Nervous System.*

prevé incluir a un total de 400 pacientes con EDSS 2-6. De momento no disponemos de resultados sobre su utilidad ni estudios de validación [41].

Elevate MS es un estudio piloto, prospectivo, de 12 semanas de duración, que incluyó a una cohorte de pacientes con EM (n = 495) y controles sanos (n = 134) de Estados Unidos. Está dirigido a evaluar la viabilidad y la utilidad de capturar datos de salud relacionados con la EM del mundo real y de forma remota utilizando la aplicación 'elevateMS' para investigar las asociaciones entre la gravedad de la EM notificada por los propios pacientes y las mediciones de pruebas funcionales activas obtenidas mediante sensores, así como para evaluar el impacto de las condiciones meteorológicas locales en la carga de la morbilidad. Dicha aplicación incluía un cuestionario de calidad de vida (Neuro-QoL), la escala de discapacidad física autoinformada *(Patient Determined Disease Steps)* y registros de salud diaria (síntomas, potenciales desencadenantes, movilidad, dolor). Las pruebas funcionales activas consistían en el golpeteo de dedos, caminar, la prueba del equilibrio, el test cognitivo *The voice-controlled Digit Symbol Substitution Test* y la prueba dedo-nariz. Los datos meteorológicos locales se recopilaron cada vez que los participantes completaron una tarea activa. Los síntomas más comúnmente descritos por los pacientes fueron fatiga (62,6%), debilidad (44,8%), problemas de memoria/atención (42,2%) y dificultad para caminar (41,4%), mientras que los desencadenantes encontrados con mayor frecuencia fueron la temperatura ambiental alta (52,3%), el estrés (50,5%) y el dormir tarde por las noches (44,6%) %). La *Patient Determined Disease Steps* notificada por los pacientes y las puntuaciones del Neuro-QoL se asociaron significativamente con los resultados de las pruebas funcionales. Finalmente, la temperatura local se asoció significativamente con el rendimiento de las pruebas activas de los participantes [42].

Conclusiones

La tecnología digital ha llegado al área de la salud y está claro que cambiará de manera franca nuestra forma de brindar atención médica. Su papel a la hora de monitorizar clínicamente al paciente parece ofrecer numerosas ventajas. Concretamente, en el seguimiento clínico del paciente con EM es capaz de proporcionar una visión más 'completa y real' de su situación funcional, incluyendo la valoración de síntomas 'no visibles' de la enfermedad, como la cognición, el estado afectivo y la fatiga, así como también evaluaciones de funciones físicas, como la destreza manual y la capacidad de la marcha, habitualmente no evaluadas en la práctica clínica por falta de tiempo y espacio. De esta manera, nos encaminamos a detectar de forma precoz los pequeños y progresivos cambios en la discapacidad, muchas veces referidos por los pacientes, pero que las pruebas clínicas tradicionales, incluyendo la EDSS, no consiguen reflejar. En este sentido, el uso de la tecnología digital nos permitirá adoptar medidas y acciones terapéuticas apropiadas de forma oportuna y personalizada. En el futuro, las herramientas digitales, capaces de brindar información clínica amplia y detallada, podrían ocupar un papel importante en la toma de decisiones y el seguimiento clínico del paciente afecto de EM.

Bibliografía

1. Klocke S, Hahn N. Multiple sclerosis. Ment Health Clin 2019; 9: 349-58.
2. Oh J, Vidal-Jordana A, Montalban X. Multiple sclerosis: clinical aspects. Curr Opin Neurol 2018; 31: 752-9.
3. Rovira A. Advances in the diagnosis, characterization, and monitoring of multiple sclerosis. Neuroimaging Clin N Am 2017; 27: xvii-xviii.
4. Hauser SL, Cree BA. Treatment of multiple sclerosis: a review. Am J Med 2020; 133: 1380-90.e2.

5. Young IR. Significant events in the developments of MRI. Magn Reson Imaging 2004; 19: 523-6.
6. Miller DH, Filippi M, Fazekas F, Frederiksen JL, Matthews PM, Montalban X, et al. Role of magnetic resonance imaging within diagnostic criteria for multiple sclerosis. Ann Neurol 2006; 56: 273-8.
7. Thompson AJ, Banwell BL, Barkhof F, Carroll WM, Coetzee T, Comi G, et al. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. Lancet Neurol 2018; 17: 162-73.
8. Chen Y, Haacke EM, Bernitsas E. Imaging of the spinal cord in multiple sclerosis: past, present, future. Brain Sci 2020; 10: 857.
9. Derfuss T, Mehling M, Papadopoulou A, Bar-Or A, Cohen JA, Kappos L. Advances in oral immunomodulating therapies in relapsing multiple sclerosis. Lancet Neurol 2020; 19: 336-47.
10. Montalban X, Hauser SL, Kappos L, Arnold DL, Bar-Or A, Comi G, et al; ORATORIO clinical investigators. ocrelizumab versus placebo in primary progressive multiple sclerosis. N Engl J Med 2017; 376: 209-20.
11. Giovannoni G. Disease-modifying treatments for early and advanced multiple sclerosis: a new treatment paradigm. Curr Opin Neurol 2018; 31: 233-43.
12. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology 1983; 33: 1444-52.
13. Alberts JL. Hot Topic 2: e-MS and patients related outcomes. Utilizing mobile technology in the assessment and monitoring of MS patients. Mult Scler 2017; 23 (Suppl 3): 58-84.
14. Inojosa H, Schriefer D, Ziemssen T. Clinical outcome measures in multiple sclerosis: a review. Autoimmun Rev 2020; 19: 102512.
15. Amato MP, Portaccio E. Clinical outcome measures in multiple sclerosis. J Neurol Sci 2007; 259: 118-22.
16. Fischer JS, Rudick RA, Cutter GR, Reingold SC. The Multiple Sclerosis Functional Composite Measure (MSFC): an integrated approach to MS clinical outcome assessment. National MS Society Clinical Outcomes Assessment Task Force. Mult Scler 1999; 5: 244-50.
17. Rudick RA, Cutter G, Reingold S. The Multiple Sclerosis Functional Composite: a new clinical outcome measure for multiple sclerosis trials. Mult Scler 2002; 8: 359-65.
18. Filippi M, Preziosa P, Barkhof F, Chard DT, De Stefano N, Fox RJ, et al. Diagnosis of progressive multiple sclerosis from the imaging perspective: a review. JAMA Neurol 2021; 78: 351-64.
19. Lublin FD, Reingold SC, Cohen JA, Cutter GR, Sørensen PS, Thompson AJ, et al. Defining the clinical course of multiple sclerosis: the 2013 revisions. Neurology 2014; 83: 278-86.
20. Lavorgna L, Brigo F, Moccia M, Leocani L, Lanzillo R, Clerico M, et al. e-Health and multiple sclerosis: An update. Mult Scler 2018; 24: 1657-64.
21. Brichetto G, Pedullà L, Podda J, Tacchino A. Beyond center-based testing: understanding and improving functioning with wearable technology in MS. Mult Scler 2019; 25: 1402-11.
22. Yousef A, Jonzzon S, Suleiman L, Arjuna J, Graves JS. Biosensing in multiple sclerosis. Expert Rev Med Devices. 2017; 14: 901-12.
23. Bradshaw MJ, Farrow S, Motl RW, Chitnis T. Wearable biosensors to monitor disability in multiple sclerosis. Neurol Clin Pract 2017; 7: 354-62.
24. Motl RW, Pilutti L, Sandroff BM, Dlugonski D, Sosnoff JJ, Pula JH. Accelerometry as a measure of walking behavior in multiple sclerosis. Acta Neurol Scand. 2013; 127: 384-90.
25. Shammas L, Zentek T, von Haaren B, Schleisinger S, Hey S, Rashid A. Home-based system for physical activity monitoring in patients with multiple sclerosis (Pilot study). Biomed Eng Online 2014; 13: 10.
26. Block V, Bove R, Zhao C, Garcha P , Graves J, Romeo AR, et al. Association of continuous assessment of step count by remote monitoring with disability progression among adults with multiple sclerosis. JAMA 2019; 2: e190570.
27. Flachenecker F, Gaberel H, Hannik J, Lee DH, Flachenecker 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.
28. Bonzano L, Sormani MP, Tacchino A, Abate L, Lapucci C, Mancardi GL, et al. Quantitative assessment of finger motor impairment in multiple sclerosis. PLoS One 2013; 8: e65225.
29. Carpinella I, Cattaneo D, Ferrarin M. Quantitative assessment of upper limb motor function in multiple sclerosis using an instrumented Action Research Arm Test. J Neuroeng Rehabil 2014; 11: 67.
30. Rudick RA, Miller D, Bethoux F, Rao SM, Lee JC, Stough D, et al. The Multiple Sclerosis Performance Test (MSPT): an iPad-based disability assessment tool. J Vis Exp 2014; 88: e51318.
31. Rao SM, Galioto R, Sokolowski M, McGinley M, Fretburger J, Weber M, et al. Multiple Sclerosis Performance Test: validation of self-administered neuroperformance modules. Eur J Neurol 2020; 27: 878-86.
32. Baldassari LE, Nakamura K, Moss BP, Macaron G, Li H, Weber M, et al. Technology-enabled comprehensive characterization of multiple sclerosis in clinical practice. Mult Scler Relat Disord 2020; 38: 101525.
33. Maillart E, Labauge P, Cohen M, Maarouf A, Vukusic S, Donzé C, et al. MSCopilot, a new multiple sclerosis self-assessment digital solution: results of a comparative study versus standard tests. Eur J Neurol 2020; 27: 429-36.
34. Morrison C, D'Souza M, Huckvale K, Dorn JF, Buggraaff J, Kamm CP, et al. Usability and acceptability of ASSESS MS: assessment of motor dysfunction in multiple sclerosis using depth-sensing computer vision. JMIR Hum Factors 2015; 1: e1.
35. Bradnall HN, Kuppanda KE, O'Connell A, Hardy TA, Reddel SW, Barnett MH. Tablet-based screening improves continence management in multiple sclerosis. Ann Clin Transl Neurol 2015; 2: 679-87.
36. Jongen PJ, Sinnige LG, van Geel BM, Verheul F, Verhagen WI, van der Kruijk RA, et al. The interactive web-based program MSmonitor for self-management and multidisciplinary care in multiple sclerosis: concept, content, and pilot results. Patient Prefer Adherence 2015; 9: 1741-50.
37. Hopkins C. SymTrac. Nurs Stand 2015; 29: 33.
38. Novartis. SymTrac URL: https: //www.novartis.co.uk/symtrac.shtml. Fecha última consulta: 15.12.2020.
39. Younes M, Hill J, Quinlen J, Kilduff M, Pang B, Cook SD, et al. Internet-based cognitive testing in multiple sclerosis. Mult Scler 2007; 13: 1011-9.
40. Midaglia L, Mulero P, Montalban X, Graves J, Hauser SL, Julian L, et al. Adherence and satisfaction of smartphone- and smartwatch-based remote active testing and passive monitoring in people with multiple sclerosis: nonrandomized interventional feasibility study. J Med Internet Res 2019; 21: e14863.
41. Ranjan Y, Rashid Z, Stewart C, Conde P, Begale M, Verbeeck D, et al, RADAR-CNS Consortium. RADAR-Base: open source mobile health platform for collecting, monitoring, and analyzing data using sensors, wearables, and mobile devices. JMIR Mhealth Uhealth 2019; 7: e11734.
42. Pratap A, Grant D, Vegesna A, Tummalacherla M, Cohan S, Deshpande C, et al. Evaluating the utility of smartphone-based sensor assessments in persons with multiple sclerosis in the real-world using an app (elevateMS): observational, prospective pilot digital health study. JMIR Mhealth Uhealth 2020; 8: e22108.

Clinical monitoring of multiple sclerosis patients by means of digital technology, a field in the midst of a revolution

Introduction. Despite the great advances that have occurred in the diagnosis and treatment of multiple sclerosis (MS), few changes have taken place in terms of clinical monitoring. The lack of time and space in clinical practice limits the assessment of invisible symptoms and certain motor symptoms as manual dexterity and walking ability, which have a clear impact on the patient functional situation.

Objective. To review the potential role of technological tools in the clinical monitoring of MS patients.

Development. A bibliographic search was carried out through PubMed, selecting those studies focused on biosensors and digital tools aimed at evaluating the general functional situation, and specific aspects of the disease or certain functional systems.

Results. Different digital tools such as biosensors, mobile or web applications, both for remote and hospital use, self-completed or administered by healthcare personnel, seem to offer a more 'complete and real' picture of the functional situation of patients. Some studies have shown that digital technology could detect subclinical disability progression, which traditional tests, including the EDSS, fail to reflect, favouring the adoption of appropriate therapeutic measures and actions in an early and personalized manner.

Conclusions. Digital tools, capable of collecting detailed and extensive clinical information, could play an important role in decision-making and clinical monitoring of patients with MS.

Key words. Clinical monitoring. Digital technology. Digital tools. e-Health. Multiple sclerosis. Wearable devices.

10.4.2 Publication N2

Adherence and Satisfaction of Smartphone- and Smartwatch-Based Remote Active Testing and Passive Monitoring in People With Multiple Sclerosis: Nonrandomized Interventional Feasibility Study. Midaglia L, Mulero P, Montalban X, Graves J, Hauser SL, Julian L, Baker M, Schadrack J, Gossens C, Scotland A, Lipsmeier F, van Beek J, Bernasconi C, Belachew S, Lindemann M.J Med Internet Res. 2019 Aug 30;21(8):e14863. doi: 10.2196/14863.

Original Paper

Adherence and Satisfaction of Smartphone- and Smartwatch-Based Remote Active Testing and Passive Monitoring in People With Multiple Sclerosis: Nonrandomized Interventional Feasibility Study

Luciana Midaglia[1,2], MD; Patricia Mulero[1], MD; Xavier Montalban[1,3], MD, PhD; Jennifer Graves[4], MAS, MD, PhD; Stephen L Hauser[5], MD; Laura Julian[6], PhD; Michael Baker[7], MSc; Jan Schadrack[7], MD; Christian Gossens[7], MBA, PhD; Alf Scotland[7], MSc; Florian Lipsmeier[7], PhD; Johan van Beek[7], PhD; Corrado Bernasconi[7], MD, PhD; Shibeshih Belachew[7], MD, PhD; Michael Lindemann[7,8], MBA, PhD

[1]Department of Neurology-Neuroimmunology, Multiple Sclerosis Centre of Catalonia, Vall d'Hebron University Hospital, Barcelona, Spain
[2]Department of Medicine, Autonomous University of Barcelona, Barcelona, Spain
[3]Division of Neurology, University of Toronto, Toronto, ON, Canada
[4]Department of Neurology, University of California, San Diego, San Diego, CA, United States
[5]Department of Neurology, University of California, San Francisco, San Francisco, CA, United States
[6]Genentech Inc, South San Francisco, CA, United States
[7]F Hoffmann–La Roche Ltd, Basel, Switzerland
[8]Department of Economics, Baden-Wuerttemberg Cooperative State University, Loerrach, Germany

Corresponding Author:
Christian Gossens, MBA, PhD
F Hoffmann–La Roche Ltd
124 Grenzacherstrasse
Basel,
Switzerland
Phone: 41 61 687 5113
Fax: 41 61 691 9391
Email: christian.gossens@roche.com

Related Article:
This is a corrected version. See correction statement: https://www.jmir.org/2019/10/e16287

Abstract

Background: Current clinical assessments of people with multiple sclerosis are episodic and may miss critical features of functional fluctuations between visits.

Objective: The goal of the research was to assess the feasibility of remote active testing and passive monitoring using smartphones and smartwatch technology in people with multiple sclerosis with respect to adherence and satisfaction with the FLOODLIGHT test battery.

Methods: People with multiple sclerosis (aged 20 to 57 years; Expanded Disability Status Scale 0-5.5; n=76) and healthy controls (n=25) performed the FLOODLIGHT test battery, comprising active tests (daily, weekly, every two weeks, or on demand) and passive monitoring (sensor-based gait and mobility) for 24 weeks using a smartphone and smartwatch. The primary analysis assessed adherence (proportion of weeks with at least 3 days of completed testing and 4 hours per day passive monitoring) and questionnaire-based satisfaction. In-clinic assessments (clinical and magnetic resonance imaging) were performed.

Results: People with multiple sclerosis showed 70% (16.68/24 weeks) adherence to active tests and 79% (18.89/24 weeks) to passive monitoring; satisfaction score was on average 73.7 out of 100. Neither adherence nor satisfaction was associated with specific population characteristics. Test-battery assessments had an at least acceptable impact on daily activities in over 80% (61/72) of people with multiple sclerosis.

Conclusions: People with multiple sclerosis were engaged and satisfied with the FLOODLIGHT test battery. FLOODLIGHT sensor-based measures may enable continuous assessment of multiple sclerosis disease in clinical trials and real-world settings.

Trial Registration: ClinicalTrials.gov: NCT02952911; https://clinicaltrials.gov/ct2/show/NCT02952911

(J Med Internet Res 2019;21(8):e14863) doi: 10.2196/14863

KEYWORDS

multiple sclerosis; patient adherence; patient satisfaction; smartphone; wearable electronic devices; mobile phone

Introduction

Disease progression throughout the clinical course of multiple sclerosis (MS) is measured using clinician-reported outcomes, most commonly the Expanded Disability Status Scale (EDSS) [1], magnetic resonance imaging (MRI), and patient-reported outcomes (PROs). Conventional assessment of the clinical course of MS relies on relapse-associated and periodic in-clinic visits. However, the current intermittently conducted clinic-based outcome measures in MS have limitations in studying the insidiously subtle progression in MS and may fail to comprehensively capture transient symptomatic and performance fluctuations that affect people with multiple sclerosis.

The ability of consumer wearable technology to measure functional impairment associated with various neurological disease symptoms through smartphone-based assessments is an important area of research [2-5]. Smartphones as vehicles for sensor-based technologies offer the potential for enhanced active and passive real-time data capture that may fundamentally shift traditional paradigms of clinical monitoring [6-9]. A recent large-scale study demonstrated that more than 95% of people with multiple sclerosis have access to a mobile device and most use it routinely [10]. Recently published studies have described the use of technologies in developing tools to assess people with multiple sclerosis [7,9,11-13]. In a recent study, it was reported that both healthy participants and people with multiple sclerosis were capable of completing daily tasks on a smartphone for 1 year [7]. Collection of data on a variety of cognitive and motor tests via the smartphone may represent a feasible way to gather highly granular data to accurately describe the MS disease course outside of the clinic [7].

The FLOODLIGHT study [NCT02952911] was a prospective pilot study to assess the feasibility of remote measurements using smartphones and smartwatches in people with multiple sclerosis and healthy controls (HCs). The smartwatch and smartphone contained apps that prompted the user to perform various assessments and protocols, referred to as active tests. The app also passively recorded sensor data during daily life, referred to as passive monitoring. The novel smartphone- and smartwatch-based FLOODLIGHT active tests were developed to be self-administered by people with multiple sclerosis to capture MS symptoms, including hand motor function, gait and posture, mood, and cognitive impairment.

The primary objectives of the FLOODLIGHT study, which began in November 2016, were to evaluate participant adherence to smartphone- and smartwatch-based assessments and collect feedback from people with multiple sclerosis and HCs on the smartphone and smartwatch schedule of assessments and its impact on their daily activities using a patient satisfaction questionnaire. Other objectives of the study, which will be addressed in subsequent publications, are to determine the association between exploratory sensor-based outcomes derived from the respective components of the FLOODLIGHT test battery and conventional MS clinical outcomes and explore whether the FLOODLIGHT test battery can differentiate between participants with and without MS.

Methods

Trial Design and Participants

After providing written informed consent, people with multiple sclerosis and HCs, preferentially partners or cohabitants, were evaluated for eligibility for enrollment in the FLOODLIGHT study. Eligibility criteria for people with multiple sclerosis included the ability to comply with the study protocol, age 18 to 55 years, diagnosis of MS (2010 revised McDonald criteria, treated or untreated) [14], EDSS score 0 to 5.5 (inclusive), and weight from 99 to 243 lbs (45 to 110 kg). An EDSS score of 5.5 as a maximum limit was meant to ensure any patient with relapsing or progressive MS would not have any significant difficulty in participating in the proposed testing as per study protocol. Further details on eligibility criteria are provided in Multimedia Appendix 1.

This study was conducted at two sites in two countries with a total of 76 people with multiple sclerosis and 25 HCs; 60 people with multiple sclerosis and 20 HCs were recruited from the Multiple Sclerosis Centre of Catalonia, Vall d'Hebron University Hospital, Barcelona, Spain, and 16 people with multiple sclerosis and 5 HCs were recruited from the University of California, San Francisco, California.

The protocol, informed consent forms, any information given to the participants, and relevant supporting information were reviewed and approved by the institutional review board/ethics committee before the study was initiated. Confidentiality was maintained by assigning each participant enrolled in the study a unique identification number.

Study Design

The FLOODLIGHT study combines continuous sensor data capture with smartphones and smartwatches and standard clinical outcome measures. Eligible people with multiple sclerosis and HCs were enrolled in the study and assessed clinically at the enrollment visit, week 12, and termination visit (week 24). In addition, participants were asked to perform a set of daily active tests and contribute sensor data via passive

monitoring with smartphone and smartwatches over a period of 24 weeks (Figure 1).

Figure 1. FLOODLIGHT study design. PRO: patient-reported outcome. "a" indicates that active tests were administered weekly or every two weeks (see next figure for schedule). "b" indicates that brain magnetic resonance imaging was performed in people with multiple sclerosis.

In-Clinic Assessments

At each scheduled in-clinic visit (enrollment, week 12, and week 24), the following reference clinical tests were performed for all participants: 9-Hole Peg Test (9HPT), oral version of Symbol Digit Modalities Test (SDMT) [15-17], Timed 25-Foot Walk (T25FW) test, Berg Balance Scale (BBS) [18], Fatigue Scale for Motor and Cognitive Functions (FSMC) [19], and Patient Health Questionnaire–9 (PHQ-9) [20]. For people with multiple sclerosis only, disability was measured by EDSS [1], Patient Determined Disease Steps (PDDS) [21], and Multiple Sclerosis Impact Scale–29 (MSIS-29; version 2) [22-24]. While performing some of the in-clinic tests, people with multiple sclerosis and HCs were asked to carry or wear the smartphone and smartwatch to collect sensor data alongside the in-clinic measures. On the scheduled in-clinic visits, the smartphone and smartwatch FLOODLIGHT active tests were performed under investigator supervision. The satisfaction questionnaire (Multimedia Appendix 2) assessed people with multiple sclerosis' and HCs' experience regarding smartphone and smartwatch use and its impact on their daily activities at the week 12 visit and at the study termination/early discontinuation visit. Brain MRI was performed in people with multiple sclerosis at the enrollment visit and at week 24.

Smartphone and Smartwatch Testing

At the enrollment visit, people with multiple sclerosis and HCs were provided with the FLOODLIGHT solution that included a smartphone and smartwatch preconfigured so participants could only run the FLOODLIGHT software. A belt bag was also provided for participants to carry their smartphone in an anterior medial position. The smartphone and smartwatch pair contained preinstalled apps that prompted the user to perform various assessments, referred to as active tests. The apps also passively recorded sensor data, referred to as passive monitoring. At the enrollment visit, participants received training on the use of the smartphone and smartwatch and were provided with supporting content to help them complete the tests successfully. Participants were instructed to complete the active tests at approximately the same time each day and carry the smartphone and smartwatch throughout the day, recharging the devices overnight. Data transfer from the smartphone and smartwatch is described in Multimedia Appendix 1.

FLOODLIGHT Active Tests

People with multiple sclerosis and HCs were asked to perform various active tests (daily, weekly, every two weeks, or on demand) via the smartphone (Figure 2 and Table 1). These novel active tests were developed to be self-administered by people with multiple sclerosis to capture MS symptoms. A range of clinical and sensor-based assessments were chosen to capture the most prominent symptoms of MS from a broad spectrum of symptoms. People with multiple sclerosis and HCs were required to wear the smartwatch throughout the active tests.

Figure 2. FLOODLIGHT active tests and their schedule frequency. DMQ: Daily Mood Question; MSIS-29: Multiple Sclerosis Impact Scale–29; SBT: Static Balance Test; SDMT: Symbol Digit Modalities Test; ST: Symptom Tracker; 2MWT: Two-Minute Walk Test; 5UTT: 5 U-Turn Test.

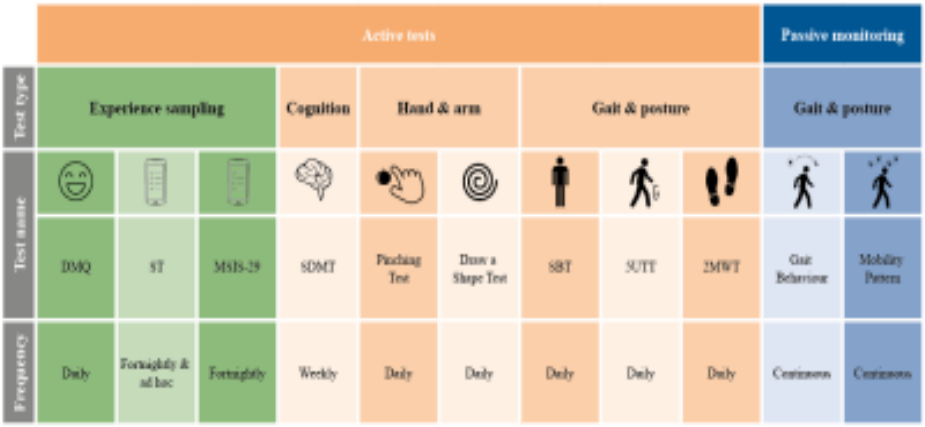


Table 1. FLOODLIGHT active tests.

Domain and test	Short description
Daily hand motor function tests[a]	
Draw a Shape (DaS) Test	The aim of the DaS Test is to assess fine finger/manual dexterity while the participants are instructed to hold the mobile device in the untested hand and draw on the smartphone touchscreen six prewritten alternating shapes of increasing complexity (linear, rectangular, circular, sinusoidal, and spiral) with the second finger of the tested hand as fast and as accurately as possible within a maximum time (30 seconds for each of the two attempts per shape).
Pinching Test	The aim of the Pinching Test is to assess fine pinching/grasping dexterity while the participants are instructed to hold the mobile device in the untested hand and touch the screen with two fingers from the tested hand (thumb + second or thumb + third finger preferred) to squeeze/pinch as many round shapes (ie, tomatoes) as they can during 30 seconds.
Daily gait tests[b]	
Two-Minute Walk Test (2MWT)	Participants are instructed to walk as fast and as long as they can for 2 minutes but walk safely. The 2MWT is a simple test that is required to be performed on an even ground in a place where participants have identified they could walk straight for as far as ≥200 meters without U-turns. Participants are allowed to wear regular footwear and an assistive device and/or orthotic as needed.
5 U-Turn Test (5UTT)	The aim of this test is to assess difficulties or unusual patterns in performing U-turns while walking on a short distance at comfortable pace. The 5UTT can be performed indoors or outdoors, on an even ground where participants are instructed to walk safely and perform five successive U-turns going back and forward between two points a few meters apart for 1 minute. Participants are allowed to wear regular footwear and an assistive device and/or orthotic as needed.
Static Balance Test (SBT)	Participants are asked to stand still unsupported for 30 seconds with relaxed arms straight alongside the body if possible.
Weekly cognitive test	
Electronic version of the Symbol Digit Modalities Test (SDMT) [15-17]	The aim of SDMT testing is to detect impairment of key neurocognitive functions that underlie many substitution tasks.
Patient-reported outcomes (PROs)	
Daily Mood Question (DMQ)	This test represents an assessment of participants' perceived overall state by responding daily to the question "How do you feel now?" on a 5-item Likert scale, ranging from excellent to horrible.
Electronic version of the Multiple Sclerosis Impact Scale–29, version 2 (MSIS-29) [22-24]; people with multiple sclerosis only	This questionnaire measures the physical and psychological impact of multiple sclerosis.
Multiple Sclerosis Symptom Tracker (MSST); people with multiple sclerosis only	Patients are asked if they experienced new or significantly worsening symptoms during the last 2 weeks. If yes, onset of the symptoms and the patients' perception to whether they think they experienced a relapse (yes, no, or unsure) are recorded.

[a]Tests alternatingly performed with right and left hand; users are instructed on daily alternation.

[b]Recommended position of smartphone in an anterior medial position in the belt bag.

FLOODLIGHT Passive Monitoring

Passive monitoring collected metrics on gait and mobility throughout the day in a continuous and unobtrusive manner. Participants were instructed to carry their smartphone preferably in an anterior medial position in a belt bag or, alternatively, in their pocket and wear the smartwatch all day as they went about their daily routine until the devices ran out of charge.

Statistical Analyses

The analyses of the primary objectives of this study were descriptive. Statistical tests were exploratory and conducted at the two-sided 5% significance level without adjustment for multiple comparisons. The analyses were based on all enrolled patients (full analysis set [FAS]). Patients who prematurely withdrew from the study for any reason were still included in the FAS. Supportive analyses of selected variables were carried out in the per-protocol population, which included people with multiple sclerosis who completed at least 1 week in the study and did not discontinue due to an adverse event or a reason unrelated to the use of the FLOODLIGHT solution (Multimedia Appendix 3).

Adherence was evaluated for the following tests and test groups: all FLOODLIGHT active tests, Two-Minute Walk Test (2MWT), all active tests except 2MWT, smartphone use, smartwatch use, and for the per-protocol population.

Adherence to active tests was measured as the proportion of study weeks with at least 3 days of completed testing (study co-primary endpoint). Adherence to sensor-based passive

monitoring was measured as the proportion of study weeks with at least 3 days of passive monitoring for at least 4 hours per day while the devices were worn by the participant [25] (study co-primary endpoint). Descriptive statistics of calculated adherence were reported for all active tests, all active tests except 2MWT, 2MWT, smartphone use, and smartwatch use. Categorical and numeric variables with fewer than five values were tested for association with adherence using the Kruskal-Wallis test. The association between continuous variables was assessed using the Spearman rank correlation.

Participant complete abandoning of active testing and passive monitoring was also investigated in a time-to-event survival analysis based on the Kaplan-Meier method along the FLOODLIGHT study. The abandoning event was defined as the last week in which the participant was adherent according to the definitions above for active tests and passive monitoring. Active tests performed on days of in-clinic visits were not considered in the adherence calculation to focus the abandoning analysis on the remote use. Participants leaving the study before the terminal visit were considered as censored. The impact of different characteristics on adherence was assessed using Cox regression.

A satisfaction score was developed from the satisfaction questionnaire (Multimedia Appendix 2) that sums the individual answers to questions 1-7 and 10-12 rescaled to 0-100 from their original Likert scale. An interquestion correlation analysis was performed to ensure questions are equally correlated and can be combined. Descriptive statistics of satisfaction score and items are reported, along with covariate analyses of demographics and disease state. Analysis of the change in satisfaction score between week 12 and week 24 are reported using the Wilcoxon signed-rank test.

Patient baseline characteristics incorporated as covariates in the analysis of correlation with FLOODLIGHT adherence and satisfaction outcomes were age, gender, body mass index, time since first MS symptom onset, EDSS, T25FW time, 9HPT time, and the oral SDMT correct responses. A descriptive analysis of safety variables, including adverse events and serious adverse events, was carried out in the FAS.

Results

Baseline Demographics and Characteristics

Participants' baseline demographics for the FAS are described in Table 2. There was an expected greater proportion of females among the people with multiple sclerosis compared with HCs. The majority (69/76, 91%) of people with multiple sclerosis had relapsing-remitting MS (RRMS), with a mild EDSS score (mean 2.4) and presumably "normal" hand/arm function based on an upper limit of normal range defined as the average 9HPT time for HCs plus two standard deviations [26] (Table 2).

Overall, 92% (70/76) of people with multiple sclerosis and 64% (16/25) of HCs who enrolled in the FLOODLIGHT study reached the week 24 visit. Reasons for discontinuation from the study are described in Multimedia Appendix 3.

Table 2. Demographics and characteristics of people with multiple sclerosis and healthy controls (HCs) at baseline.

Parameter	People with multiple sclerosis (n=76)	HCs (n=25)
Age (years), mean (SD)	39.5 (7.9)	34.9 (9.3)
Female, n (%)	53 (70)	7 (28)
Multiple sclerosis (MS) diagnosis, n (%)		
Primary progressive multiple sclerosis	3 (4)	—[a]
Secondary progressive multiple sclerosis	4 (5)	—
Relapsing-remitting multiple sclerosis	69 (91)	—
Time since MS symptom onset (years), mean (SD)	11.3 (7.0)[b]	—
Proportion of people with multiple sclerosis with ≥1 relapse in the past year, n (%)	18 (24)	—
Expanded Disability Status Scale, mean (SD)	2.4 (1.4)	—
Proportion of people with multiple sclerosis with ≥1 T1 Gd[c]-enhancing lesion, n (%)	2 (3)[d]	—
Total FLAIR[e] lesion volume (mL), mean (SD)	6.3 (7.5)[f]	—
9-Hole Peg Test (seconds), mean (SD)		
Dominant hand	22.1 (4.6)[g]	18.9 (2.1)
Nondominant hand	22.8 (4.9)[h]	19.5 (2.0)
Timed 25-Foot Walk (seconds), mean (SD)	6.0 (2.1)[h]	5.0 (1.0)
Symbol Digit Modalities Test (correct responses), mean (SD)	53.8 (11.8)[h]	63.8 (10.0)
Berg Balance Scale, mean (SD)	52.5 (5.7)[i]	56.0 (0)[j]
Patient Determined Disease Steps, mean (SD)	1.5 (1.6)	—
Fatigue Scale for Motor and Cognitive Functions (total score), mean (SD)	59.1 (22.7)[g]	25.5 (6.0)
Patient Health Questionnaire–9, mean (SD)	8.3 (6.1)[k]	2.4 (2.9)[l]
Participants with any previous medications, n (%)	46 (61)	6 (24)
Previous disease-modifying treatment[m], n (%)		
Daclizumab (Zinbryta)	0 (0)	—
Glatiramer acetate (Copaxone)	12 (16)	—
Glatiramer acetate (Glatopa)	1 (1)	—
IFN[n] β-1a IM[o] (Avonex)	4 (5)	—
IFN β-1a SC[p] (Rebif)	5 (7)	—
IFN β-1b SC (Betaseron/Betaferon)	6 (8)	—
IFN β-1b SC (Extavia)	1 (1)	—
Pegylated IFN β-1a (Plegridy)	2 (3)	—
Dimethyl fumarate (Tecfidera)	9 (12)	—
Fingolimod (Gilenya)	9 (12)	—
Teriflunomide (Aubagio)	3 (4)	—
Alemtuzumab (Lemtrada)	2 (3)	—
Mitoxantrone (Novantrone)	1 (1)	—
Natalizumab (Tysabri)	19 (25)	—
Other[q]	5 (7)	—

[a]Not applicable.
[b]n=75.

[c]Gd: gadolinium.
[d]n=68.
[e]FLAIR: fluid-attenuated inversion recovery.
[f]n=70.
[g]n=73.
[h]n=74.
[i]n=71.
[j]n=22.
[k]n=60.
[l]n=20.
[m]Total baseline disease-modifying treatment history.
[n]IFN: interferon.
[o]IM: intramuscular.
[p]SC: subcutaneous.
[q]Hidroferol; Radiance study (RPC1063 versus IFN β-1a); Rituximab (Rituxan).

Adherence

Over a period of 18 months (November 2016-April 2018), more than 6 terabytes of raw data were collected from 76 people with multiple sclerosis and 25 HCs. Participants performed 67,544 active tests, of which 9787 were the 2MWT, and recorded 200,171 hours of passive monitoring, of which 113,165 hours were captured with the smartwatch. Over 24 weeks, most participants performed 5 to 7 active tests per week, including the 2MWT (Figure 3). Adherence of people with multiple sclerosis to completing active tests and passive monitoring was good and remained stable over time after week 6 (Figures 4 and 5). Even in the last week of the 24-week study, participants completed all active tests on average 4 out of 7 days per week (Figure 4), and recorded at least 4 hours of data via passive monitoring on average 4 out of 7 days per week (Figure 5). The lowest average adherence over 24 weeks was observed for active tests including the 2MWT and the 2MWT only, with participants showing highest average adherence for passive monitoring (Figure 6). A total of 70% (16.68/24 weeks) of participants were adherent to all active tests, 75% (17.95/24 weeks) to all active tests except 2MWT, 71% (17.13/24 weeks) to 2MWT, 79% (18.89/24 weeks) to smartphone- or smartwatch-based passive monitoring, 66% (15.74/24 weeks) to smartphone-based passive monitoring, and 74% (17.69/24 weeks) to smartwatch-based passive monitoring.

Figure 3. Adherence of people with multiple sclerosis to active tests for individual participants: number of performed active tests per week [level of activity (light green: high; dark green/grey: low] over individual study weeks [columns]).

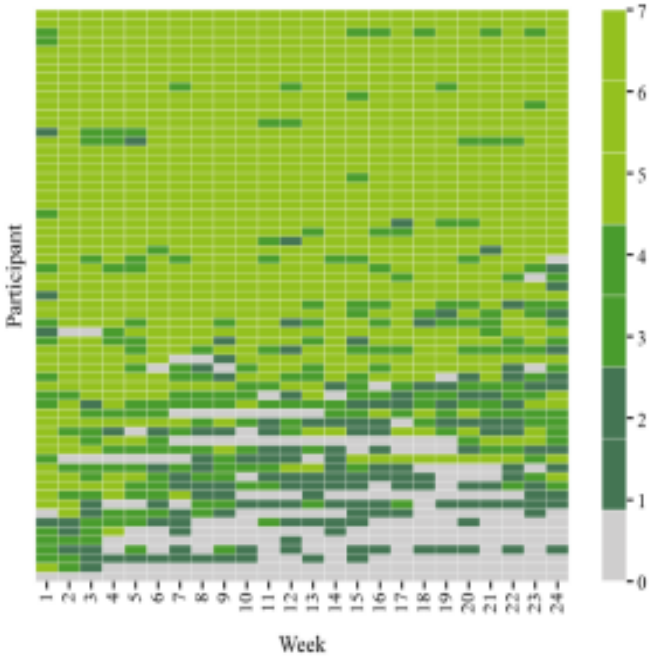


Figure 4. Adherence of people with multiple sclerosis to active tests. 2MWT: Two-Minute Walk Test.

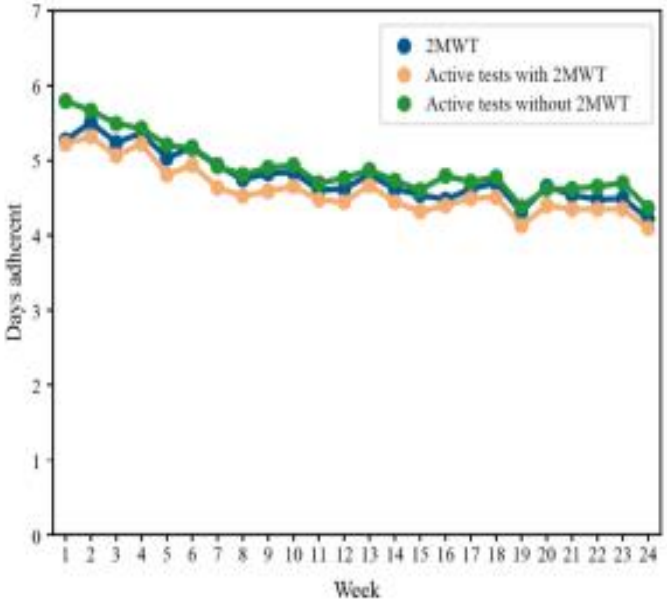


Figure 5. Adherence of people with multiple sclerosis to smartphone and smartwatch passive monitoring. Days with more than 4 hours of passive monitoring on a device are considered as adherent.

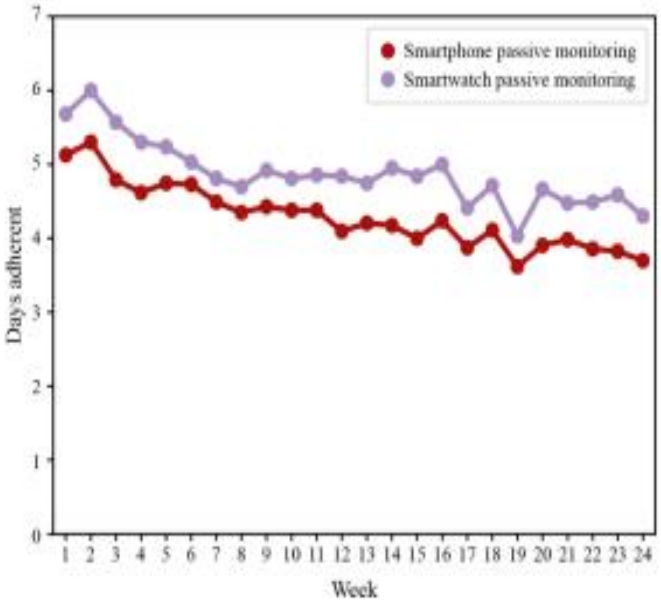
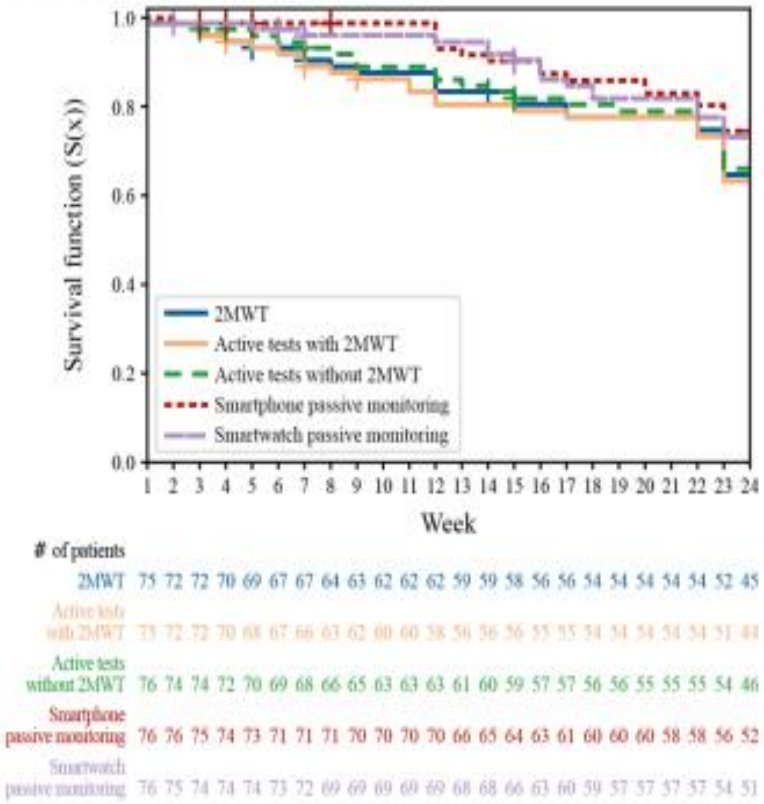


Figure 6. Adherence of people with multiple sclerosis to active tests and passive monitoring. The results of the time-to-event survival analysis based on the Kaplan–Meier method along the FLOODLIGHT study. The abandoning event was defined as the last week in which the participant was adherent according to the definitions for active tests and passive monitoring. Active tests performed on days of in-clinic visits were not considered in the adherence calculation, to focus the abandoning analysis on the remote use. Participants leaving the study before the terminal visit were considered as censored. 2MWT: Two-Minute Walk Test.



Correlation was explored between adherence measures and people with multiple sclerosis population characteristics. Only disease duration showed significant small negative correlation with measures of adherence (Spearman rank correlations; 2MWT adherence: –0.42, $P<.001$; smartphone passive monitoring adherence: –0.29, $P=.02$; all active tests except 2MWT adherence: –0.37, $P=.003$; and smartwatch passive monitoring adherence: –0.27, $P=.04$), indicating that disease severity and demographics did not appear to play a significant role in adherence.

Patient Satisfaction

The average overall satisfaction score among people with multiple sclerosis who completed the study at week 12 (n=64) was 74.1 out of a possible 100 and remained stable at week 24 (study termination/early discontinuation visit [n=68]) with 73.7 out of 100 (Wilcoxon signed-rank test $P=.71$). There was one significant association between overall satisfaction score and gender ($P=.04$). Individual questions from the satisfaction questionnaire (Multimedia Appendix 2) were analyzed for their association with people with multiple sclerosis population characteristics, described in Multimedia Appendix 1.

Implications for the use of the FLOODLIGHT test battery in people with multiple sclerosis were assessed from individual questions from the patient satisfaction questionnaire. When asked to rate the impact of the smartphone, smartwatch, and active tests on daily living, more than 80% (61/72) of people with multiple sclerosis perceived the FLOODLIGHT test battery to have at least acceptable impact on daily activities (Figure 7). Nearly 50% (32/71) of participants had no issue with any of the active tests, and only one-third would prefer to avoid the 2MWT, most likely due to increased burden from execution—for example, having to find a place to perform the test or not wanting to go outside in bad weather (Figure 8). Without providing any data feedback to the people with multiple sclerosis throughout the study, more than 60% (46/72) of participants would have liked to continue using FLOODLIGHT "to understand my MS better and improve my disease management" (Figure 9). Approximately 90% (65/72) of people with multiple sclerosis indicated their interest to see the results of the tests, which will be addressed in future Roche-sponsored studies using FLOODLIGHT (CONSONANCE [NCT03523858] and FLOODLIGHT Open [floodlightopen.com]; Figure 10). Analysis of patient responses to the satisfaction questionnaire is described in Multimedia Appendix 1.

Figure 7. Implications of FLOODLIGHT in people with multiple sclerosis for "impact on daily activities" from the patient satisfaction questionnaire.

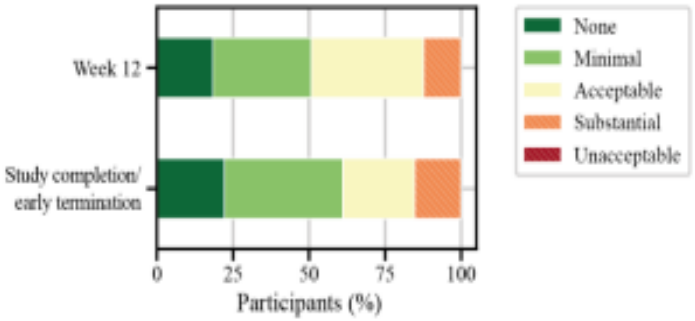


Figure 8. Implications of FLOODLIGHT in people with multiple sclerosis for "avoiding one component of FLOODLIGHT" from the patient satisfaction questionnaire. 2MWT: Two-Minute Walk Test.

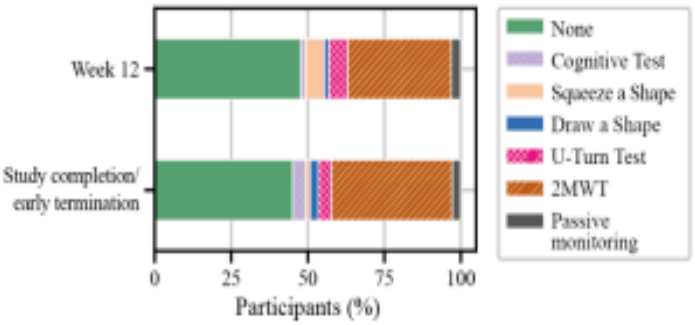


Figure 9. Implications of FLOODLIGHT in people with multiple sclerosis for "desire to continue using the FLOODLIGHT app" from the patient satisfaction questionnaire.

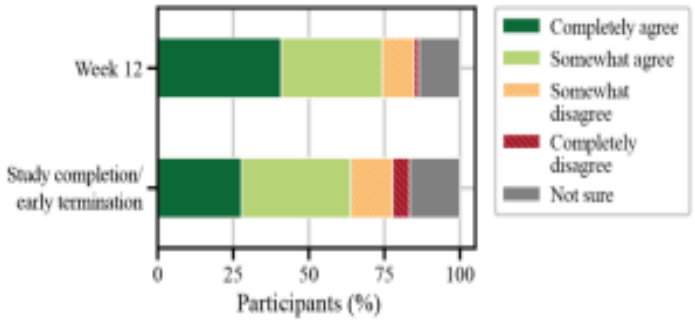
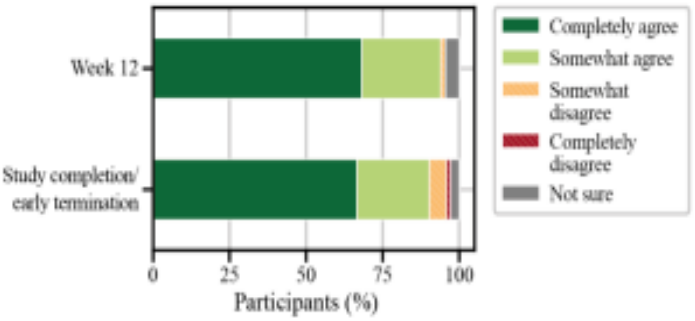


Figure 10. Implications of FLOODLIGHT in people with multiple sclerosis for "prefer to see results immediately to monitor" from the patient satisfaction questionnaire.



Discussion

Principal Findings

This study demonstrates that the use of smartphones and smartwatches for remote daily active testing and continuous passive monitoring is feasible over 6 months and provides further support to earlier studies, which have shown that healthy participants and people with multiple sclerosis were capable of completing daily tasks on a smartphone [7]. This study provides further evidence for the use of digital technology, including smartphones, for data collection. Other studies in MS have used smartphone apps to (1) assess steps when walking on a treadmill [9]; (2) assess pain, fatigue, anxiety, and quality of life [13]; and (3) assess the feasibility of gathering passive and active performance data [7]. Together with the current analyses, these studies document the focus toward developing digital measures to continuously monitor and assess the MS disease.

In this protocol, the FLOODLIGHT solution collected metrics on cognition, mood, upper extremity function, and gait and posture by instructing participants to perform a set of daily active tests, which should take approximately 5 minutes in total to complete and capture activity data via passive monitoring over a period of 24 weeks. A previous study has shown that 51% of participants (22/38 of people with multiple sclerosis and 17/38 of healthy participants) completed 12 months of daily data collection, where participants were prompted to complete one assigned test [7]. In the context of the FLOODLIGHT study, we observed that overall adherence to active tests was 70% (16.68/24 weeks), which appears to be higher than the adherence of participants to 12 months of daily data collection (39/76, 51%) from Bove et al [7]. However, comparisons between the studies are limited, as the study design and burden of testing are different—for example, the app from Bove et al contained 19 different tests, of which participants were prompted to complete one each day. As the FLOODLIGHT app was integrated into standalone devices in this study, deployment of the app on participants' own mobile devices may increase adherence because it removes the need to carry a separate, dedicated device and decreases burden on the individual.

Limitations

As this study remains a pilot investigation designed to collect first experiences from continuous sensor data capture, the main limitation is the small sample size and short duration of follow-up. Future ongoing FLOODLIGHT studies (CONSONANCE and FLOODLIGHT Open) will collect longer term data on smartphone-based sensor data capture in a larger number of participants from a broader disability spectrum. Additionally, whether physical and cognitive limitations in people with secondary progressive MS (SPMS) and people with primary progressive MS (PPMS) differentially impacts adherence compared with people with relapsing MS (RMS) cannot be gleaned from the current data set due to the low numbers of advanced patients enrolled in the study; however, this important question warrants future research exploring remote monitoring in patients with more advanced MS. The importance of continuous monitoring in RMS should also not be overlooked, as the sensitivity of this novel approach aiming at detecting progression in a real-world setting may provide an earlier window into disease progression outside of the clinic.

Comparison With Prior Work

A recent study assessing the feasibility of the MS TeleCoach, a novel intervention offering telemonitoring of fatigue and telecoaching of physical activity in people with multiple sclerosis, showed that participants were highly engaged, with 76% (57/75) of participants completing the study, and 91% (21/23) of a subset of completers showing a median of quite satisfied in the patient satisfaction questionnaire [27]. During the 12-week study period, use of the MS TeleCoach improved fatigue levels in people with multiple sclerosis with moderate to severe fatigue, suggesting that implementation of digital technologies can enhance patient performance. Together with the data presented here, these results indicate that the use of consumer devices by people with multiple sclerosis for sensor data capture fulfills the prerequisites of people with multiple sclerosis satisfaction and acceptable adherence to daily active tests and passive monitoring for potential integration in long-term clinical trials and treatment monitoring. Regarding future studies, for example FLOODLIGHT Open, attempts will be made to improve participant adherence throughout the study by introducing controlled app variations, such as reminders, types of achievements, and fun metrics.

Conclusions

In summary, these analyses showed that people with multiple sclerosis are highly engaged with performing active tests and capturing continuous data via passive monitoring and are satisfied with the FLOODLIGHT test battery. Neither satisfaction nor adherence showed strong correlation with study population characteristics. More than 60% (46/72) of people with multiple sclerosis indicated their interest to continue to use FLOODLIGHT, and approximately 90% (65/72) wanted to see the results of their tests in real time as biofeedback, which was implemented in future studies using the FLOODLIGHT solution. These findings indicate that smartphone-based FLOODLIGHT outcomes may represent a promising avenue to enable a more accurate and continuous assessment of MS disease in clinical trials and real-world practice settings and may eventually also contribute to informing and guiding clinical research and clinical practice in the future.

Acknowledgments

We would like to thank all patients, their families, and the investigators who participated in this trial. This research was funded by F Hoffmann–La Roche Ltd, Basel, Switzerland. We would like to thank the following employees from F Hoffmann–La Roche Ltd who supported and contributed to the study: Atieh Bamdadian, Alessandro Barbato, Jan Beckmann, Sandro Fritz, Nicholas Pierce Heinemeier, Timothy Kilchenmann, Lito Kriara, Bernd Laub, Grégoire Pointeau, Caroline Polakowska, Marcin Puhacz, Cedric Simillion, Jens Schjodt-Eriksen, Jörg Sprengel, Ralf Stubner, and Krzysztof Trybus. Writing and editorial assistance for

this manuscript was provided by Heather Latimer from Articulate Science, United Kingdom, and funded by F Hoffmann–La Roche Ltd.

Qualified researchers may request access to individual patient-level data through the clinical study data request platform (https://clinicalstudydatarequest.com). Further details on Roche's criteria for eligible studies are available at https://clinicalstudydatarequest.com/Study-Sponsors/Study-Sponsors-Roche.aspx. For further details on Roche's Global Policy on the Sharing of Clinical Information and how to request access to related clinical study documents, see https://www.roche.com/research_and_development/who_we_are_how_we_work/clinical_trials/our_commitment_to_data_sharing.htm.

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: F Hoffmann–La Roche Ltd, Basel, Switzerland, provided financial support for the study and publication of this manuscript.

Conflicts of Interest

XM has received speaker honoraria and travel expense reimbursement for participation in scientific meetings, been a steering committee member of clinical trials, or served on advisory boards of clinical trials for Actelion, Biogen, Celgene, Merck, Novartis, Oryzon, Roche, Sanofi Genzyme, and Teva Pharmaceutical. JG has received grants or research support from Biogen, Genentech Inc, and S3 Group and has received compensation for a nonbranded resident and fellow education seminar supported by Biogen. SLH serves on the scientific advisory boards for Annexon, Symbiotix, Bionure, and Molecular Stethoscope, is on the board of trustees for Neurona Therapeutics, and has received travel reimbursement and writing assistance from F Hoffmann–La Roche Ltd for CD20-related meetings and presentations. LJ is an employee of Genentech Inc and a shareholder of F Hoffmann–La Roche Ltd. MB, JS, and CG are employees and shareholders of F Hoffmann–La Roche Ltd. AS, FL, and JvB are employees of F Hoffmann–La Roche Ltd. CB and ML are contractors for F Hoffmann–La Roche Ltd. SB was an employee of F Hoffmann–La Roche Ltd during the completion of the work related to this manuscript. SB is now an employee of Biogen (Cambridge, MA), which was not in any way associated with this study. LM and PM have nothing to disclose.

Multimedia Appendix 1

Additional trial information.
[PDF File (Adobe PDF File), 278KB-Multimedia Appendix 1]

Multimedia Appendix 2

Satisfaction questionnaire.
[PDF File (Adobe PDF File), 238KB-Multimedia Appendix 2]

Multimedia Appendix 3

Participant flow table.
[PDF File (Adobe PDF File), 138KB-Multimedia Appendix 3]

References

1. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology 1983 Nov;33(11):1444-1452. [Medline: 6685237]
2. Bradshaw MJ, Farrow S, Motl RW, Chitnis T. Wearable biosensors to monitor disability in multiple sclerosis. Neurol Clin Pract 2017 Aug;7(4):354-362. [doi: 10.1212/CPJ.0000000000000382] [Medline: 29185551]
3. Lavorgna L, Brigo F, Moccia M, Leocani L, Lanzillo R, Clerico M, et al. e-Health and multiple sclerosis: An update. Mult Scler 2018 Nov;24(13):1657-1664. [doi: 10.1177/1352458518799629] [Medline: 30231004]
4. Busis N. Mobile phones to improve the practice of neurology. Neurol Clin 2010 May;28(2):395-410. [doi: 10.1016/j.ncl.2009.11.001] [Medline: 20202500]
5. Lipsmeier F, Taylor KI, Kilchenmann T, Wolf D, Scotland A, Schjodt-Eriksen J, et al. Evaluation of smartphone-based testing to generate exploratory outcome measures in a phase 1 Parkinson's disease clinical trial. Mov Disord 2018 Dec;33(8):1287-1297 [FREE Full text] [doi: 10.1002/mds.27376] [Medline: 29701258]
6. Prasad S, Ramachandran R, Jennings C. Development of smartphone technology to monitor disease progression in multiple sclerosis. Neurology 2012 Apr 22;78(Meeting Abstracts 1):P01.144. [doi: 10.1212/WNL.78.1_MeetingAbstracts.P01.144]
7. Bove R, White CC, Giovannoni G, Glanz B, Golubchikov V, Hujol J, et al. Evaluating more naturalistic outcome measures: a 1-year smartphone study in multiple sclerosis. Neurol Neuroimmunol Neuroinflamm 2015 Dec;2(6):e162 [FREE Full text] [doi: 10.1212/NXI.0000000000000162] [Medline: 26516627]
8. Boukhvalova AK, Kowalczyk E, Harris T, Kosa P, Wichman A, Sandford MA, et al. Identifying and quantifying neurological disability via smartphone. Front Neurol 2018;9:740 [FREE Full text] [doi: 10.3389/fneur.2018.00740] [Medline: 30233487]

9. Balto JM, Kinnett-Hopkins DL, Motl RW. Accuracy and precision of smartphone applications and commercially available motion sensors in multiple sclerosis. Mult Scler J Exp Transl Clin 2016;2:2055217316634754 [FREE Full text] [doi: 10.1177/2055217316634754] [Medline: 28607720]
10. Haase R, Schultheiss T, Kempcke R, Thomas K, Ziemssen T. Modern communication technology skills of patients with multiple sclerosis. Mult Scler 2013 Aug;19(9):1240-1241. [doi: 10.1177/1352458512471882] [Medline: 23388164]
11. Rudick RA, Miller D, Bethoux F, Rao SM, Lee J, Stough D, et al. The Multiple Sclerosis Performance Test (MSPT): an iPad-based disability assessment tool. J Vis Exp 2014 Jun 30(88):e51318 [FREE Full text] [doi: 10.3791/51318] [Medline: 25046650]
12. Gholami F, Trojan DA, Kovecses J, Haddad WM, Gholami B. A Microsoft Kinect-based point-of-care gait assessment framework for multiple sclerosis patients. IEEE J Biomed Health Inform 2017 Dec;21(5):1376-1385. [doi: 10.1109/JBHI.2016.2593692] [Medline: 27455529]
13. Kos D, Raeymaekers J, Van Remoortel A, D'hooghe MB, Nagels G, D'Haeseleer M, et al. Electronic visual analogue scales for pain, fatigue, anxiety and quality of life in people with multiple sclerosis using smartphone and tablet: a reliability and feasibility study. Clin Rehabil 2017 Sep;31(9):1215-1225. [doi: 10.1177/0269215517692641] [Medline: 28786335]
14. Polman CH, Reingold SC, Banwell B, Clanet M, Cohen JA, Filippi M, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. Ann Neurol 2011 Feb;69(2):292-302 [FREE Full text] [doi: 10.1002/ana.22366] [Medline: 21387374]
15. Smith A. The symbol-digit modalities test: A neuropsychologic test of learning and other cerebral disorders. Seattle: Helmuth J, editor. Learning disorders. Seattle: Special Child Publications; 1968:83-91.
16. Smith A. Symbol Digits Modalities Test: Manual. Los Angeles: Western Psychological Services; 1982.
17. Rao SM, Losinski G, Mourany L, Schindler D, Mamone B, Reece C, et al. Processing speed test: validation of a self-administered, iPad-based tool for screening cognitive dysfunction in a clinic setting. Mult Scler 2017 Dec;23(14):1929-1937. [doi: 10.1177/1352458516688955] [Medline: 28080262]
18. Berg KO, Wood-Dauphinee SL, Williams JI, Maki B. Measuring balance in the elderly: validation of an instrument. Can J Public Health 1992;83 Suppl 2:S7-S11. [Medline: 1468055]
19. Penner IK, Raselli C, Stöcklin M, Opwis K, Kappos L, Calabrese P. The Fatigue Scale for Motor and Cognitive Functions (FSMC): validation of a new instrument to assess multiple sclerosis-related fatigue. Mult Scler 2009 Dec;15(12):1509-1517. [doi: 10.1177/1352458509348519] [Medline: 19995840]
20. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [Medline: 11556941]
21. Learmonth YC, Motl RW, Sandroff BM, Pula JH, Cadavid D. Validation of patient determined disease steps (PDDS) scale scores in persons with multiple sclerosis. BMC Neurol 2013 Apr 25;13:37 [FREE Full text] [doi: 10.1186/1471-2377-13-37] [Medline: 23617555]
22. Hobart J, Cano S. Improving the evaluation of therapeutic interventions in multiple sclerosis: the role of new psychometric methods. Health Technol Assess 2009 Feb;13(12):1-177 [FREE Full text] [doi: 10.3310/hta13120] [Medline: 19216837]
23. Hobart J, Lamping D, Fitzpatrick R, Riazi A, Thompson A. The Multiple Sclerosis Impact Scale (MSIS-29): a new patient-based outcome measure. Brain 2001 May;124(Pt 5):962-973. [doi: 10.1093/brain/124.5.962] [Medline: 11335698]
24. Riazi A, Hobart JC, Lamping DL, Fitzpatrick R, Thompson AJ. Multiple Sclerosis Impact Scale (MSIS-29): reliability and validity in hospital based samples. J Neurol Neurosurg Psychiatry 2002 Dec;73(6):701-704 [FREE Full text] [doi: 10.1136/jnnp.73.6.701] [Medline: 12438473]
25. Cheng W, Scotland A, Lipsmeier F, Kilchenmann T, Jin L, Schjodt-Eriksen J, et al. Human activity recognition from sensor-based large-scale continuous monitoring of Parkinson's disease patients. 2017 Presented at: Proceedings of the 2nd IEEE/ACM International Conference on Connected Health; July 17-19, 2017; Philadelphia. [doi: 10.1109/CHASE.2017.87]
26. Wang Y, Bohannon RW, Kapellusch J, Garg A, Gershon RC. Dexterity as measured with the 9-Hole Peg Test (9-HPT) across the age span. J Hand Ther 2015;28(1):53-59. [doi: 10.1016/j.jht.2014.09.002] [Medline: 25449717]
27. D'hooghe M, Van Gassen G, Kos D, Bouquiaux O, Cambron M, Decoo D, et al. Improving fatigue in multiple sclerosis by smartphone-supported energy management: the MS TeleCoach feasibility study. Mult Scler Relat Disord 2018 May;22:90-96 [FREE Full text] [doi: 10.1016/j.msard.2018.03.020] [Medline: 29649789]

Abbreviations

2MWT: Two-Minute Walk Test
5UTT: 5 U-Turn Test
9HPT: 9-Hole Peg Test
BBS: Berg Balance Scale
DMQ: Daily Mood Question
EDSS: Expanded Disability Status Scale
FAS: full analysis set
FLAIR: fluid-attenuated inversion recovery

FSMC: Fatigue Scale for Motor and Cognitive Functions
HCs: healthy controls
MRI: magnetic resonance imaging
MS: multiple sclerosis
MSIS-29: Multiple Sclerosis Impact Scale–29
MSST: Multiple Sclerosis Symptom Tracker
PDDS: Patient Determined Disease Steps
PHQ-9: Patient Health Questionnaire–9
PPMS: primary progressive multiple sclerosis
PRO: patient-reported outcome
RMS: relapsing multiple sclerosis
RRMS: relapsing-remitting multiple sclerosis
SBT: Static Balance Test
SDMT: Symbol Digit Modalities Test
SPMS: secondary progressive multiple sclerosis
T25FW: Timed 25-Foot Walk

Edited by G Eysenbach; submitted 11.06.19; peer-reviewed by A Nguyen, L Visser, L Lavorgna, T Clavier; comments to author 29.06.19; revised version received 11.07.19; accepted 19.07.19; published 30.08.19

Please cite as:
Midaglia L, Mulero P, Montalban X, Graves J, Hauser SL, Julian L, Baker M, Schadrack J, Gossens C, Scotland A, Lipsmeier F, van Beek J, Bernasconi C, Belachew S, Lindemann M
Adherence and Satisfaction of Smartphone- and Smartwatch-Based Remote Active Testing and Passive Monitoring in People With Multiple Sclerosis: Nonrandomized Interventional Feasibility Study
J Med Internet Res 2019;21(8):e14863
URL: http://www.jmir.org/2019/8/e14863/
doi: 10.2196/14863
PMID: 31471961

10.4.3 Publication N3

A smartphone sensor-based digital outcome assessment of multiple sclerosis. Montalban X, Graves J, Midaglia L, Mulero P, Julian L, Baker M, Schadrack J, Gossens C, Ganzetti M, Scotland A, Lipsmeier F, van Beek J, Bernasconi C, Belachew S, Lindemann M, Hauser SL.Mult Scler. 2022 Apr;28(4):654-664. doi: 10.1177/13524585211028561. Epub 2021 Jul 14.

MULTIPLE SCLEROSIS JOURNAL | MSJ

Original Research Paper

A smartphone sensor-based digital outcome assessment of multiple sclerosis

Xavier Montalban, Jennifer Graves, Luciana Midaglia, Patricia Mulero, Laura Julian, Michael Baker, Jan Schadrack, Christian Gossens, Marco Ganzetti, Alf Scotland, Florian Lipsmeier, Johan van Beek, Corrado Bernasconi, Shibeshih Belachew, Michael Lindemann* and Stephen L Hauser*

Multiple Sclerosis Journal
1–11
DOI: 10.1177/13524585211028561

Abstract

Background: Sensor-based monitoring tools fill a critical gap in multiple sclerosis (MS) research and clinical care.

Objective: The aim of this study is to assess performance characteristics of the Floodlight Proof-of-Concept (PoC) app.

Methods: In a 24-week study (clinicaltrials.gov: NCT02952911), smartphone-based active tests and passive monitoring assessed cognition (electronic Symbol Digit Modalities Test), upper extremity function (Pinching Test, Draw a Shape Test), and gait and balance (Static Balance Test, U-Turn Test, Walk Test, Passive Monitoring). Intraclass correlation coefficients (ICCs) and age- or sex-adjusted Spearman's rank correlation determined test–retest reliability and correlations with clinical and magnetic resonance imaging (MRI) outcome measures, respectively.

Results: Seventy-six people with MS (PwMS) and 25 healthy controls were enrolled. In PwMS, ICCs were moderate-to-good (ICC(2,1) = 0.61–0.85) across tests. Correlations with domain-specific standard clinical disability measures were significant for all tests in the cognitive ($r = 0.82$, $p < 0.001$), upper extremity function ($|r| = 0.40$–0.64, all $p < 0.001$), and gait and balance domains ($r = -0.25$ to -0.52, all $p < 0.05$; except for Static Balance Test: $r = -0.20$, $p > 0.05$). Most tests also correlated with Expanded Disability Status Scale, 29-item Multiple Sclerosis Impact Scale items or subscales, and/or normalized brain volume.

Conclusion: The Floodlight PoC app captures reliable and clinically relevant measures of functional impairment in MS, supporting its potential use in clinical research and practice.

Keywords: Multiple sclerosis, smartphone, sensors, digital health technology, wearable electronic devices, mobile phone

Date received: 12 February 2021; revised: 1 June 2021; accepted: 3 June 2021

Correspondence to:
C Gossens
F. Hoffmann-La Roche Ltd, 124 Grenzacherstrasse, 4070 Basel, Switzerland.
christian.gossens@roche.com

Xavier Montalban
Patricia Mulero
Department of Neurology-Neuroimmunology, Centre d'Esclerosi Múltiple de Catalunya (Cemcat), Hospital Universitari Vall d'Hebron, Barcelona, Spain

Jennifer Graves
Department of Neurosciences, University of California San Diego, San Diego, CA, USA

Luciana Midaglia
Department of Neurology-Neuroimmunology, Centre d'Esclerosi Múltiple de Catalunya (Cemcat), Hospital Universitari Vall d'Hebron, Barcelona, Spain and Department of Medicine, Autonomous University of Barcelona, Barcelona, Spain

Laura Julian
Genentech, Inc., South San Francisco, CA, USA

Michael Baker
Jan Schadrack
Christian Gossens
Marco Ganzetti
Alf Scotland
Florian Lipsmeier
Johan van Beek
Corrado Bernasconi
Shibeshih Belachew
Michael Lindemann
F. Hoffmann-La Roche Ltd, Basel, Switzerland

Stephen L Hauser
UCSF Weill Institute for Neurosciences and Department of Neurology, University of California San Francisco, San Francisco, CA, USA

* Shared senior authorship.

Introduction

The course of multiple sclerosis (MS) was traditionally categorized as relapsing or progressive, with relapses associated with focal inflammation and progression with neurodegeneration. Recent advances in therapeutics have increased control of MS relapses in most patients but have also unmasked an underlying insidious progression, termed progression independent of relapses, in many patients previously thought to have relapsing–remitting disease.[1] Thus, progression can now be viewed as a universal feature of MS that is present throughout the disease course. Consequently, optimal control of MS would require that progression is minimized or eliminated in all patients. To achieve this goal, sensitive tools for monitoring disability in all people with MS (PwMS) are required, even for patients whose disease activity appears to be superficially under control in terms of relapses.[2] Furthermore, detection of progression onset or worsening is critical to optimally adapt the therapeutic strategy. Additional challenges in MS disability assessment include detection of pseudo-relapses and symptoms that often fluctuate with illness, fatigue, or changes in body temperature.[3–5] These fluctuations limit the utility of once- or twice-yearly in-clinic

monitoring; a more frequent estimate of function during daily life routine is likely to have greater value in tracking MS impairment.

Smartphone sensor-based, remotely administered digital tests represent a promising new avenue to capture disability burden with quantitative accuracy.[6] Such tests typically allow the remote capture of ecologically valid measures in a frequent, time- and cost-efficient fashion with minimal patient burden. In addition, they quantify multiple aspects of nervous system function during a given task, in opposition to clinician-administered disability tests that only assess the capacity to complete a task with a performance summary score.[7,8] Consequently, sensor-based tests can potentially disentangle intra- and inter-individual differences in underlying patterns of overall identical levels of MS-related functional impairment and thereby augment the resolution of disability measures compared with traditional scores.[7] They are also rater-independent, offering more objective measures of functional ability.[8–10]

In recent years, sensor-based tests have been increasingly used to assess functional ability in MS,[7–14] as well as in Parkinson's disease[15] and Huntington's disease.[16,17] The Floodlight Proof-of-Concept (PoC) app was designed to remotely measure, at home in an unsupervised setting, functional ability in cognitive, upper extremity, and gait and balance domains using a smartphone device without additional hardware.[9] Consequently, assessment with the Floodlight PoC app can be performed more frequently than clinician-administered tests, which potentially allows the capture of subtle changes in function that occur in free-living environments between clinic visits, including changes that would not necessarily trigger clinical interventions. The study "Monitoring of Multiple Sclerosis Participants With the Use of Digital Technology (Smartphones and Smartwatches)—A Feasibility Study" was the first clinical trial to implement the Floodlight PoC app.[9] Using data from this study, we previously demonstrated that PwMS were engaged and highly satisfied with this app.[9] We expand on this prior work by presenting here the critical evaluation of the app's test–retest reliability and its correlations with standard clinical and magnetic resonance imaging (MRI) measures in PwMS.

Methods

Study design and participants

The present study is a 24-week, prospective study (clinicaltrials.gov: NCT02952911) aimed to assess the feasibility of remotely monitoring PwMS with the Floodlight PoC app, which was developed for use in this study on a provisioned smartphone and smartwatch. The full study design and inclusion/exclusion criteria have been previously described.[9] Both PwMS and healthy controls (HC) were 18–55 years of age. PwMS (untreated or treated) were diagnosed according to the 2010 revised McDonald criteria[18] and had an Expanded Disability Status Scale (EDSS) score at baseline between 0.0 and 5.5, inclusive. All patients provided written informed consent. At each scheduled clinic visit (baseline, Week 12, Week 24), all participants underwent a clinical evaluation. In addition, PwMS were assessed by MRI at baseline and Week 24.

Participants were provided with a preconfigured smartphone (Samsung Galaxy S7) and smartwatch (Motorola 360 Sport) with the Floodlight PoC app installed;[9] for this report, only the smartphone data are reported. The smartphone prompted all participants to perform daily or weekly remote, sensor-based tests, referred to as active tests (Table 1). In addition, for passive monitoring, sensor data were passively recorded to examine ambulation and overall mobility in daily life (Table 1). Participants were encouraged to continuously carry the device to gather data for passive monitoring.

Floodlight PoC app

The Floodlight PoC app was designed to assess functional abilities across three key domains affected by MS: cognition, upper extremity function, and gait and balance.[9] Cognition was assessed using the electronic Symbol Digit Modalities Test (e-SDMT). As with the oral Symbol Digit Modalities Test (SDMT), it assessed impairment of key neurologic functions underlying cognitive information processing speed. The aim was to correctly match a maximum number of symbols to their paired digits within 90 seconds.

The Pinching Test and Draw a Shape Test assessed upper extremity function. The Pinching Test evaluated fine pinching or grasping dexterity and instructed participants to successfully pinch as many circular tomato shapes appearing on the smartphone screen at different positions as possible within 30 seconds. The Draw a Shape Test, which required participants to draw six prewritten shapes of increasing complexity (two diagonal lines, a square, circle, a figure-of-8, and a spiral), assessed fine finger or manual dexterity

Finally, the Static Balance Test (SBT), U-Turn Test (UTT), Walk Test, and Passive Monitoring examined

Table 1. Test features from the Floodlight PoC app included in the analysis.

Functional domain	Test/feature	Definitions[a]	Disability concept assessed by the feature	Sensor	Testing frequency	Functional meaning of higher scores
Cognition	**e-SDMT**					
	Number of correct responses (*n*)	Number of correct responses given within 90 seconds	Cognitive speed to process information	Touchscreen	Weekly	Better
Upper extremity function	**Pinching Test**					
	Double touch asynchrony (s)	Gap duration between the first and the second finger touching the screen	Ability to synchronously perform fine distal upper extremity movement	Touchscreen	Daily	Worse
	Number of successful pinches (*n*)	Number of tomatoes successfully pinched, or squeezed, within 30 seconds	Ability to perform fast and accurate finger opposition movement	Touchscreen	Daily	Better
	Draw a Shape Test					
	Overall mean trace accuracy	Proportion of sample points that overlap with the reference shape	Accuracy of fine distal upper extremity movement	Touchscreen	Daily	Better
	Overall mean trace celerity (s^{-1})	Ratio of overall mean trace accuracy and the time needed to draw the shape	Speed and accuracy of fine distal upper extremity movement	Touchscreen	Daily	Better
Gait and balance	**SBT**					
	Sway path (m/s^2)	Sum of the accelerometer signals in the *x*-, *y*-, and *z*-axis.	Ability to maintain stable orthostatic posture	Accelerometer	Daily	Worse
	UTT					
	Turn speed (rad/s)	Angular velocity while performing U-turns	Ability to turn while walking	Accelerometer and gyroscope	Daily	Better
	Walk Test					
	Step power	Integral of mean-centered acceleration magnitude signal over time divided by number of steps	Energy invested per step during regular walking	Accelerometer	Daily	Better
	Passive Monitoring					
	Turn speed (rad/s)	Angular velocity while performing U-turns	Ability to turn while walking	Accelerometer and gyroscope	Continuously	Better
	Step power	Integral of mean-centered acceleration magnitude signal over time divided by number of steps	Energy invested per step during regular walking	Accelerometer	Continuously	Better

PoC: Proof of Concept; e-SDMT: electronic Symbol Digit Modalities Test; SBT: Static Balance Test; UTT: U-Turn Test.
[a]See Table S2 in the supplementary appendix for the mathematical definitions.

gait and balance. The SBT was designed to study balance by asking the participants to stand, unsupported, as still as possible for 30 seconds. The UTT instructed participants to perform five successive U-turns separated by at least 4 meters within 1 minute at a comfortable pace to evaluate difficulties or unusual patterns in turning while walking and dynamic balance. The Walk Test aimed to assess gait while walking as fast as possible but also safely for 2 minutes on an even ground without performing U-turns. Finally, Passive Monitoring examined both aspects of gait—turning while walking and regular, straight walking—throughout the day. While performing the gait assessments, participants carried the smartphone in a running belt or in their trouser pocket and were permitted to use an assistive device and/or orthotic as needed.

For each active test and passive monitoring, one to two test features that are illustrative of the test and probe different neurologic concepts were extracted from the raw sensor data (Table 1).

Signal processing

As the smartphone-based tests were performed without supervision by a physician or study coordinator, quality control steps were applied to identify and exclude individual assessments that were performed incorrectly. This ensures the measurements are both reliable and accurate. To exclude such incorrectly performed assessments, quality control flags were defined for each test (Table S1). In addition, only sufficiently adherent participants, that is, those who contributed at least six individual assessments in the course of the study, were included in the analysis of that particular test. Applying these two quality control steps resulted in the final dataset consisting of valid assessments.

Next, all valid assessments contributed by a participant were aggregated to study test–retest reliability and Spearman's rank correlations in a cross-sectional analysis. As the tests were performed once daily at most, the test–retest analysis was based on the median test performance on the active tests and passive monitoring during 12 two-week windows. Two-week windows were chosen to reduce variability that is independent of general disease status and might be attributed to good or bad days or to differences between weekdays and weekends. For the cross-sectional correlation analysis, the median test performance across the entire study duration as well as the mean of the three in-clinic assessments (mean of two assessments for MRI) were computed.

Statistical analysis

Test–retest reliability was evaluated with intraclass correlation coefficients (ICC[2,1]) separately in PwMS and HC, which considered all consecutive 2-week windows. Generally, at least three valid individual assessments were required for each 2-week period. An exception was made for the e-SDMT to accommodate its weekly testing schedule; only one valid e-SDMT assessment for each 2-week window was considered sufficient. Test–retest reliability was considered as poor (ICC < 0.5), moderate (ICC $= 0.5$–0.75), good (ICC $= 0.75$–0.9), or excellent (ICC > 0.9).[19]

To examine the agreement of the test features from the Floodlight PoC app with standard clinical and MRI measures in PwMS, the test features were correlated against domain-specific clinical measures (oral SDMT, Nine-Hole Peg Test [9HPT], Berg Balance Scale [BBS], Timed 25-Foot Walk [T25FW]), EDSS, 29-item Multiple Sclerosis Impact Scale (MSIS-29) items or subscales, T2 FLAIR (fluid-attenuated inversion recovery lesion volume, and normalized brain volume using Spearman's rank correlation. In addition, test features from UTT and Walk Test were correlated against corresponding features obtained from Passive Monitoring. The strength of correlation was considered as good-to-excellent ($|r| > 0.75$), moderate-to-good ($|r| = 0.5$–0.75), fair ($|r| = 0.2$–0.49), or not correlated ($|r| < 0.25$), where $|r|$represents the absolute value.[20]

Partial correlation analyses were performed for the upper extremity function tests and gait assessments to investigate the contribution of each test and test feature in predicting 9HPT times and T25FW times, respectively. All test features from the Pinching Test and the Draw a Shape Test as well as 9HPT time were included in the upper extremity function model. Each of these features was correlated against each other while controlling for the remaining features in the model. The partial correlation for the gait assessments was run in a similar fashion but included all test features from the UTT, Walk Test, and Passive Monitoring in addition to T25FW time instead.

All correlations were adjusted for age and sex with a robust linear model. Statistical significance was set at $p < 0.05$ without correction for multiple comparisons.

Results

Full baseline demographics and disease characteristics have been previously described[9] and are summarized in Table 2. In total, 76 PwMS were enrolled between

Table 2. Baseline demographics and disease characteristics.

Variable	$N = 76$
Age (years), mean (SD)	39.5 (7.9)
Female, *n* (%)	53 (69.7)
Diagnosis, *n* (%)	
Primary progressive multiple sclerosis	3 (3.9)
Secondary progressive multiple sclerosis	4 (5.3)
Relapsing–remitting multiple sclerosis	69 (90.8)
Expanded Disability Status Scale	
Mean (SD)	2.4 (1.4)
Median	2.0
IQR	1.5–3.5
Range	0.0–5.5
Oral Symbol Digit Modalities Test score	
Mean (SD)	53.8 (11.8)
Median	55.0
IQR	46.0–63.0
Range	26.0–77.0
Nine-Hole Peg Test time (s)	
Mean (SD)	22.2 (4.0)
Median	21.4
IQR	19.7–23.7
Range	16.4–39.6
Berg Balance Scale total score	
Mean (SD)	52.5 (5.7)
Median	56.0
IQR	51.0–56.0
Range	31.0–56.0
Timed 25-Foot Walk time (s)	
Mean (SD)	6.0 (2.1)
Median	5.3
IQR	4.6–6.8
Range	3.5–12.1
T2 FLAIR lesion volume (mL)	
Mean (SD)	6.3 (7.5)
Median	3.0
IQR	1.0–9.3
Range	0.1–31.2
Normalized brain volume (mL)	
Mean (SD)	1474.5 (75.6)
Median	1477.2
IQR	1439.6–1525.6
Range	1137.3–1628.1

SD: standard deviation; IQR: interquartile range; FLAIR: fluid-attenuated inversion recovery.

November 2016 and May 2018. Most PwMS had mild disease (mean baseline EDSS: 2.4; EDSS range: 0.0–5.5) and were diagnosed with relapsing–remitting MS. Valid data were available in 68–73 PwMS. PwMS and in 18–24 HC across the 10 tests considered (Figure S1), using test-specific quality control criteria. Similarly, 81.5–99.9% of assessments in PwMS and 76.0–99.9% of assessments in HC were considered valid (Figure S2). During the 24-week study, only eight participants exhibited a change in EDSS ≥ 1, which is defined as clinically meaningful,[21] an insufficient number to enable further analysis of changes in EDSS (Figure S3).

Test–retest reliability

Test–retest reliability was assessed in PwMS and HC with valid assessments in all consecutive 2-week windows (range: 32–46 PwMS and 8–11 HC). In PwMS, ICCs(2,1) were moderate or good (Table 3), suggesting that reliable data can be captured with the Floodlight PoC app. In HC, where the group sizes were lower, ICCs(2,1) were mostly poor to good.

Correlation with clinical and MRI measures

The age- and sex-adjusted Spearman's rank correlation analysis in PwMS is summarized in Table 3 and Figure 1. All statistically significant correlations were in the expected direction. Thus, increasing levels of MS-related disability were associated with worse performance on the Floodlight PoC app.

Overall, strongest correlations of test features were observed with the respective domain-specific standard clinical disability measures. These correlations were good-to-excellent in the cognitive domain ($r = 0.82$) and fair or moderate-to-good in the upper extremity function domain ($|r| = 0.40$–0.64) and gait and balance domain ($r = -0.25$ to -0.52, all $p < 0.05$). Only the SBT did not correlate with its domain-specific standard clinical measure, the BBS ($r = -0.20$, $p > 0.05$). Most test features also correlated with EDSS (all $p < 0.05$ except for Draw a Shape Test overall mean trace celerity and Passive Monitoring step power) and their respective MSIS-29 subscale or items (all $p < 0.05$ except for Draw a Shape Test overall mean trace celerity, Passive Monitoring turn speed, and Passive Monitoring step power). Normalized brain volume correlated significantly with test features across all domains with the strongest association found with e-SDMT ($r = 0.54$, $p < 0.001$). Similar results were obtained with unadjusted measures (Table S3).

Table 3. Test–retest reliability in PwMS and HC and age- and sex-adjusted Spearman's rank correlation analysis of the Floodlight PoC app in PwMS.

Domain	Test	Feature	Test–retest reliability ICC (2,1) (95% CI)		Spearman's rank correlations for PwMS				
			HC	PwMS	Domain-specific standard clinical measure	EDSS	MSIS-29 subscale/ items[a]	T2 FLAIR lesion volume (mL)	Normalized brain volume (mL)
Cognition	e-SDMT	Number of correct responses (*n*)	0.55 (0.34–0.80)[b]	0.85 (0.76–0.91)[c]	Oral SDMT	0.82***	−0.43***	−0.52***	−0.42***	0.54***
Upper extremity function	Pinching Test	Double touch asynchrony (s)	0.72 (0.53–0.90)[d]	0.71 (0.62–0.80)[c]	9HPT	0.64***	0.30*	0.35**	0.17	−0.26*
	Pinching Test	Number of successful pinches (*n*)	0.81 (0.65–0.94)[d]	0.72 (0.61–0.82)[c]	9HPT	−0.52***	−0.26*	−0.33**	−0.12	0.32**
	Draw a Shape Test	Overall mean trace accuracy	0.53 (0.32–0.79)[b]	0.85 (0.79–0.90)[c]	9HPT	−0.48***	−0.40***	−0.40***	−0.26*	0.33**
	Draw a Shape Test	Overall mean trace celerity (s^{-1})	0.45 (0.25–0.73)[b]	0.81 (0.73–0.87)[c]	9HPT	−0.40***	−0.08	0.03	−0.26*	0.24*
Gait and balance	SBT	Sway path (m/s^2)	0.40 (0.20–0.73)[e]	0.71 (0.61–0.80)[g]	BBS	−0.20	0.24*	0.31**	0.21	−0.05
	UTT	Turn speed (rad/s)	0.45 (0.24–0.75)[d]	0.83 (0.76–0.89)[h]	T25FW	−0.52***	−0.45***	−0.39***	−0.13	0.27*
	Walk Test	Step power	0.85 (0.70–0.95)[e]	0.78 (0.70–0.86)[g]	T25FW	−0.31**	−0.28*	−0.24*	0.04	−0.02
	Passive Monitoring	Turn speed (rad/s)	0.66 (0.42–0.89)[e]	0.72 (0.61–0.82)[h]	T25FW	−0.25*	−0.27*	−0.12	−0.11	0.14
	Passive Monitoring	Step power	0.63 (0.39–0.88)[e]	0.61 (0.48–0.74)[h]	T25FW	−0.33**	−0.19	−0.22	0.09	0.00

PwMS: people with multiple sclerosis; HC: healthy controls; PoC: Proof of Concept; ICC: intraclass correlation coefficient; CI: confidence interval; EDSS: Expanded Disability Status Scale; MSIS-29: 29-item Multiple Sclerosis Impact Scale; FLAIR: fluid-attenuated inversion recovery; e-SDMT: electronic Symbol Digit Modalities Test; 9HPT: Nine-Hole Peg Test; SBT: Static Balance Test; BBS: Berg Balance Scale; UTT, U-Turn Test; T25FW Timed 25-Foot Walk.
Colored background indicates significant correlation in expected direction.
[a]The e-SDMT was correlated against the psychological subscale, the Pinching and Draw a Shape Tests against the arm-related items (items 2, 6 and 15), and all other tests against the physical subscale.
[b]$n = 11$.
[c]$n = 46$.
[d]$n = 10$.
[e]$n = 44$.
[f]$n = 9$.
[g]$n = 42$.
[h]$n = 41$.
[i]$n = 39$.
[j]$n = 8$.
[k]$n = 32$.
$^*p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

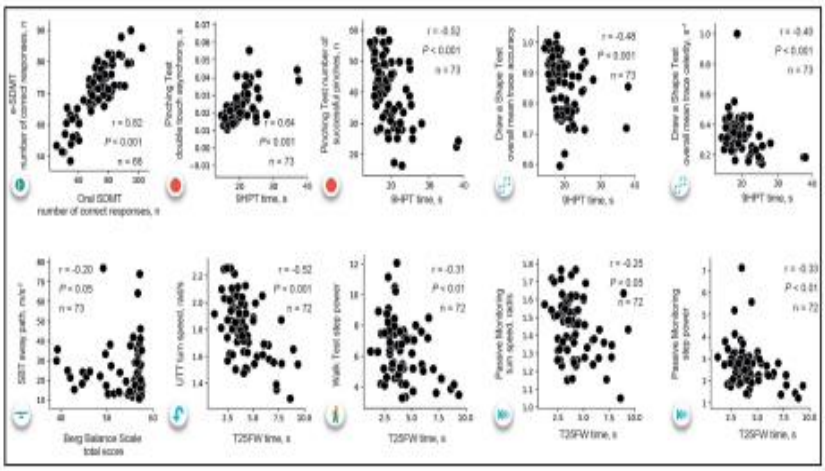


Figure 1. Age- and sex-adjusted Spearman's rank correlations between active tests and passive monitoring (vertical axis) and their respective domain-specific standard clinical measures (horizontal axis).
e-SDMT: electronic Symbol Digit Modalities Test; SDMT: Symbol Digit Modalities Test; 9HPT: Nine-Hole Peg Test; SBT: Static Balance Test; UTT: U-Turn Test; T25FW: Timed 25-Foot Walk.

Figure 2. Age- and sex-adjusted Spearman's rank correlations between passive monitoring and active gait features.
UTT: U-Turn Test; NS: not significant.
***$p < 0.001$.

Next, we assessed correlations between active gait tests and passive monitoring. UTT turn speed showed moderate-to-good correlation with Passive Monitoring turn speed ($r = 0.43$, $p < 0.001$). Stronger, good-to-excellent positive correlations were observed between Walk Test step power and Passive Monitoring step power ($r = 0.76$, $p < 0.001$; Figure 2).

The partial correlation analysis revealed that both the Pinching Test (double touch asynchrony: partial $r = 0.37$, $p < 0.001$) and Draw a Shape Test (overall mean trace accuracy: partial $r = -0.40$, $p < 0.001$; overall mean trace celerity: partial $r = -0.30$, $p < 0.01$) contain independent information in predicting 9HPT time (Figure S4). Similarly, the UTT carries unique information in predicting T25FW time when correcting for the other gait features (partial $r = -0.31$, $p < 0.01$; Figure S5).

Comparing the performance of the e-SDMT and oral SDMT in PwMS and HC showed that both HC and PwMS had fewer correct responses on the e-SDMT compared with the oral SDMT (Figure S6).

Discussion

Here, we provide the first evidence that the Floodlight PoC app can reliably capture clinically relevant data measures of functional impairment in PwMS. By leveraging smartphone-based consumer technology for clinical research, the Floodlight PoC app assesses key neurological domains affected by MS and provides a more objective and detailed picture of the disease than is possible with standard "point of care" clinical assessments. Results from this study indicate that test features derived from the Floodlight PoC app hold potential for use in clinical research and practice.

Test–retest reliability was consistent with ICCs reported for standard clinical measures in PwMS.[22] As

anticipated, statistically significant correlations were observed between test features from the Floodlight PoC app and related standard clinical and MRI measures. Correlations of similar strength between comparable smartphone sensor-based remote monitoring tests and clinician-administered tests have been reported in the cognitive,[11,12] upper extremity function,[13] and gait domains,[10] despite differences in test design and features. The active tests and passive monitoring also showed mostly fair correlations with EDSS and respective MSIS-29 subscales and items, indicating that measurements obtained with the Floodlight PoC app agree both with the overall level of MS-related disability and participant's perception of the impact of their disease. T2 FLAIR lesion volume did not show any strong correlation with either active tests or passive monitoring. This is not surprising given the clinico-radiological paradox,[23] which describes the mismatch between the white matter lesion volume and the clinical outcomes in MS, and the subsequent poor cross-sectional correlation between T2-weighted imaging and MS disability measures in a relatively mild MS population.[24] Normalized brain volume correlated with test features from all assessed domains. This is in line with previously reported correlations between normalized brain volume and measures of MS-related disability.[25]

Not surprisingly, the e-SDMT most closely resembles its domain-specific standard clinical measure. The Spearman rho of 0.82 is comparable to the previously reported correlation between other smartphone-based versions of the SDMT and the pen-and-paper version of the SDMT ($r = 0.71–0.85$).[11,12] In our study, we noted that the e-SDMT scores tended to be lower than the oral SDMT scores in PwMS and HC. This is likely due to the longer time required to select the correct response on a smartphone display compared with saying the correct response out loud. Another possible reason is that the e-SDMT displays only one symbol at a time. The oral SDMT, on the contrary, provides participants the entire symbol sequence printed on a sheet of paper,[26] thus allowing them to work ahead and use their working memory to a greater extent. This difference also makes the oral SDMT more dependent on eye tracking than the e-SDMT.

Given the different concepts assessed by the Floodlight PoC app versus clinical measures, 1:1 correlations were not necessarily expected. For example, mean trace celerity did not correlate with EDSS or the arm-related MSIS-29 items. This is likely because these clinical measures do not capture the time component as overall mean trace accuracy correlated significantly with both clinical measures.

The partial correlation analysis presented here revealed that test features from both the Pinching Test and the Draw a Shape Test independently correlate with 9HPT time. This supports the concept that specific sensor-based test features can capture performance outcome information currently not recorded with commonly used in-clinic assessments. This exemplifies the potential of sensor data to characterize functional impairment beyond a single summary score that is typically recorded for in-clinic performance outcome measures. Future work should explore the use of this technology in broader clinical applications and focus on establishing the clinical relevance for the additional information it can provide. It is possible that richer information can be extracted by incorporating additional test features. Initial results on a more comprehensive multidimensional feature space have been previously reported for the Draw a Shape Test[27] and Walk Test.[28,29]

In addition to the active tests, the Floodlight PoC app also assesses gait in a free-living situation, or in daily life, through passive monitoring. It has been suggested that signs of gait alteration may be more pronounced during daily life than in conventional in-clinic metrics,[30] thereby highlighting the importance of capturing out-of-clinic performance through passive monitoring. As such, passive monitoring may improve the translation of clinical findings to meaningful care as it informs on the patients' true abilities during daily life activities.[9] A recent study demonstrated the feasibility of passively monitoring gait in PwMS using three biosensors attached to the wrist, ankle, and sternum.[8] Moreover, our study revealed significant correlations between gait features from Passive Monitoring and T25FW, highlighting the feasibility of capturing more ecological measures of everyday functional ability through Passive Monitoring using a single smartphone device.

Several limitations to this study exist. Most enrolled PwMS had mild disease with limited MS-related disability; the mean EDSS at baseline was 2.4. The current analysis assessed the performance characteristics of the Floodlight PoC app in PwMS with EDSS scores in the range of 0.0–5.5. However, previously it has been shown that wearable sensors might not accurately capture step detection at slow walking speeds, particularly at EDSS score 6.5. This feature should be considered when assessing any wearable monitoring technology.[31] In addition, the analyses presented here were cross-sectional. Due to the relative short duration of the study (24 weeks), a longitudinal analysis on change in functional ability, disease progression, and relapses was not possible. Future studies will lend greater clarity into the use in a broader patient population, including people with more advanced disease, and the test performance over time.

Furthermore, test–retest reliability analysis was conducted in 2-week windows, in which no disease progression was assumed, as each assessment was done at most once per day. Same-day test–retest reliability analysis will be addressed in future work using data from subsequent studies. Further work will also be needed on the development of domain-specific and overall MS outcome measures based on digital health technology.[32]

Conclusion

Using a consumer smartphone device, self-administered at-home active tests and passive monitoring assessed the functional ability across three key domains affected by MS: cognition, upper extremity function, and gait and balance. This study demonstrated that the Floodlight PoC app provides reliable measures that align with standard clinical and MRI measures used to quantify MS functional impairment and overall disability. Test–retest reliability was moderate-to-good, and significant correlations in the expected direction were observed between the test features from the Floodlight PoC app and standard clinical and MRI measures. While active tests were conducted daily or weekly, passive monitoring permitted the continuous assessment of gait during daily life activities. The higher temporal resolution and multidimensional feature space of functional data collected by this platform hold the potential to capture subtle, potentially disease-related information which are not readily discriminated by clinician-administered assessments. It also has the potential to improve and standardize assessment of MS disease over time, provide PwMS and health care professionals in both specialty and primary care environments a better understanding of disease progression, change the way MS is monitored in clinical trials and daily practice, and ultimately improve patient care. The current iteration of the app, Floodlight™ MS, is available for public use in selected countries, and a rolling release schedule is now in process to provide access in the near future to the wider MS community across the world.

Acknowledgements

We would like to thank all patients, their families, and the investigators who participated in this trial. We would also like to thank the following colleagues at F. Hoffmann-La Roche Ltd for their contributions and support to the study: Jan Beckmann, Frank Dondelinger, Sandro Fritz, Nicholas Pierce Heinemeier, Sven Holm, Timothy Kilchenmann, Lito Kriara, Bernd Laub, Grégoire Pointeau, Caroline Polakowska, Cedric Simillion, Jens Schjodt-Eriksen, Jörg Sprengel, Ralf Stubner, Krzysztof Trybus, and Mattia Zanon. Writing and editorial assistance for this manuscript was provided by Samuel Bestall, PhD, and Marina Dragovic from Articulate Science, United Kingdom, and funded by F. Hoffmann-La Roche Ltd.

Declaration of Conflicting Interests

X.M. has received speaking honoraria and travel expenses for participation in scientific meetings, has been a steering committee member of clinical trials, or participated in advisory boards of clinical trials in the past years with Actelion, Alexion, Bayer, Biogen, Celgene, EMD Serono, Genzyme, Immunic, Medday, Merck, Mylan, Nervgen, Novartis, Roche, Sanofi Genzyme, Teva Pharmaceutical, TG Therapeutics, Excemed, MSIF, and NMSS. L.M. and P.M. have nothing to disclose. J.G. over the past year has received grant/contract research support from the National MS Society, Biogen, and Octave Biosciences. She serves on a steering committee for a trial supported by Novartis. She has received honoraria for a non-promotional, educational activity for Sanofi Genzyme. She has received speaker fees from Alexion and BMS and served on an advisory board for Genentech. L.J. is an employee of Genentech, Inc., and a shareholder of F. Hoffmann-La Roche Ltd. M.B., J.S., and C.G. are employees and shareholders of F. Hoffmann-La Roche Ltd. M.G. and C.B. are contractors for F. Hoffmann-La Roche Ltd. A.S. was a consultant to F. Hoffmann-La Roche Ltd via Inovigate during the completion of the work related to this manuscript. A.S. is now an employee of Inovigate (Basel, Switzerland). F.L. is an employee of F. Hoffmann-La Roche Ltd. J.v.B. and S.B. were employees of F. Hoffmann-La Roche Ltd during the completion of the work related to this manuscript. Both are now employees of Biogen (Cambridge, MA), which was not in any way associated with this study. M.L. is a consultant to F. Hoffmann-La Roche Ltd via Inovigate. S.L.H. serves on the scientific advisory boards for Alector, Annexon, Bionure, and Molecular Stethoscope; is on the Board of Directors for Neurona Therapeutics; and has received travel reimbursement and writing assistance from F. Hoffmann-La Roche Ltd and Novartis for CD20-related meetings and presentations.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: F. Hoffmann-La Roche Ltd, Basel, Switzerland, provided financial support for the study and publication of this manuscript.

Data availability

Qualified researchers may request access to individual patient-level data through the clinical study data

request platform (https://vivli.org). Further details on Roche's criteria for eligible studies are available at https://vivli.org/members/ourmembers. For further details on Roche's Global Policy on the Sharing of Clinical Information and how to request access to related clinical study documents, see https://www.roche.com/research_and_development/who_we_are_how_we_work/clinical_trials/our_commitment_to_data_sharing.htm.

ORCID iD

Shibeshih Belachew https://orcid.org/0000-0003-3976-1950

Supplemental material

Supplemental material for this article is available online.

References

1. Kappos L, Wolinsky JS, Giovannoni G, et al. Contribution of relapse-independent progression vs relapse-associated worsening to overall confirmed disability accumulation in typical relapsing multiple sclerosis in a pooled analysis of 2 randomized clinical trials. *JAMA Neurol* 2020; 77: 1132–1140.
2. Hobart J, Bowen A, Pepper G, et al. International consensus on quality standards for brain health-focused care in multiple sclerosis. *Mult Scler* 2019; 25: 1809–1818.
3. Rae-Grant A, Bennett A, Sanders AE, et al. Quality improvement in neurology: Multiple sclerosis quality measures: Executive summary. *Neurology* 2015; 85: 1904–1908.
4. Steelman AJ. Infection as an environmental trigger of multiple sclerosis disease exacerbation. *Front Immunol* 2015; 6: 520–520.
5. Mills EA, Mirza A and Mao-Draayer Y. Emerging approaches for validating and managing multiple sclerosis relapse. *Front Neurol* 2017; 8: 116–116.
6. Graves JS and Montalban X. Biosensors to monitor MS activity. *Mult Scler* 2020; 26: 605–608.
7. Boukhvalova AK, Fan O, Weideman AM, et al. Smartphone level test measures disability in several neurological domains for patients with multiple sclerosis. *Front Neurol* 2019; 10: 358–358.
8. Chitnis T, Glanz BI, Gonzalez C, et al. Quantifying neurologic disease using biosensor measurements in-clinic and in free-living settings in multiple sclerosis. *NPJ Digit Med* 2019; 2: 123–123.
9. Midaglia L, Mulero P, Montalban X, et al. Adherence and satisfaction of smartphone- and smartwatch-based remote active testing and passive monitoring in people with multiple sclerosis: Nonrandomized interventional feasibility study. *J Med Internet Res* 2019; 21: e14863.
10. Block VJ, Lizée A, Crabtree-Hartman E, et al. Continuous daily assessment of multiple sclerosis disability using remote step count monitoring. *J Neurol* 2017; 264: 316–326.
11. Pham L, Harris T, Varosanec M, et al. Smartphone-based symbol-digit modalities test reliably captures brain damage in multiple sclerosis. *NPJ Digit Med* 2021; 4: 36.
12. van Oirschot P, Heerings M, Wendrich K, et al. Symbol Digit Modalities Test variant in a smartphone app for persons with multiple sclerosis: Validation study. *JMIR Mhealth Uhealth* 2020; 8: e18160.
13. Akhbardeh A, Arjona JK, Krysko KM, et al. Novel MS vital sign: Multi-sensor captures upper and lower limb dysfunction. *Ann Clin Transl Neurol* 2020; 7: 288–295.
14. Angelini L, Hodgkinson W, Smith C, et al. Wearable sensors can reliably quantify gait alterations associated with disability in people with progressive multiple sclerosis in a clinical setting. *J Neurol* 2020; 267: 2897–2909.
15. Lipsmeier F, Taylor KI, Kilchenmann T, et al. Evaluation of smartphone-based testing to generate exploratory outcome measures in a phase 1 Parkinson's disease clinical trial. *Mov Disord* 2018; 33: 1287–1297.
16. Tortelli R, Simillion C, Lipsmeier F, et al. The digital-HD study: Smartphone-based remote testing to assess cognitive and motor symptoms in Huntington's Disease (1816). *Neurology* 2020; 94(Suppl. 15): 1816.
17. McLaren B, Andrews SC, Glikmann-Johnston Y, et al. Feasibility and initial validation of "HD-Mobile," a smartphone application for remote self-administration of performance-based cognitive measures in Huntington's disease. *J Neurol* 2021; 268: 590–601.
18. Polman CH, Reingold SC, Banwell B, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Ann Neurol* 2011; 69: 292–302.
19. Koo TK and Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 2016; 15: 155–163.
20. Portney LG and Watkins MP. *Foundations of clinical practice research: Applications to practice*. London: Pearson, 2009.
21. Kragt JJ, Nielsen JM, van der Linden FA, et al. How similar are commonly combined criteria for EDSS progression in multiple sclerosis. *Mult Scler* 2006; 12: 782–786.
22. Goldman MD, LaRocca NG, Rudick RA, et al. Evaluation of multiple sclerosis disability outcome

measures using pooled clinical trial data. *Neurology* 2019; 93: e1921–e1931.

23. Chard D and Trip SA. Resolving the clinico-radiological paradox in multiple sclerosis. *F1000Res* 2017; 6: 1828.
24. Miller DH, Grossman RI, Reingold SC, et al. The role of magnetic resonance techniques in understanding and managing multiple sclerosis. *Brain* 1998; 121(Pt1): 3–24.
25. Radue EW, Barkhof F, Kappos L, et al. Correlation between brain volume loss and clinical and MRI outcomes in multiple sclerosis. *Neurology* 2015; 84: 784–793.
26. Smith A. *Symbol Digit Modalities Test: Manual*. Los Angeles, CA: Western Psychological Services, 1982.
27. Creagh AP, Simillion C, Scotland A, et al. Smartphone-based remote assessment of upper extremity function for multiple sclerosis using the Draw a Shape Test. *Physiol Meas* 2020; 41: 054002.
28. Creagh AP, Simillion C, Bourke A, et al. Smartphone- and smartwatch-based remote characterisation of ambulation in multiple sclerosis during the two-minute walk test. *IEEE J Biomed Health Inform* 2020; 25: 838–849.
29. Bourke AK, Scotland A, Lipsmeier F, et al. Gait Characteristics harvested during a smartphone-based self-administered 2-minute walk test in people with multiple sclerosis: Test–retest reliability and minimum detectable change. *Sensors* 2020; 20: 5906.
30. Broen MP, Marsman VA, Kuijf ML, et al. Unraveling the relationship between motor symptoms, affective states and contextual factors in Parkinson's disease: A feasibility study of the experience sampling method. *PLoS ONE* 2016; 11: e0151195.
31. Storm FA, Nair KPS, Clarke AJ, et al. Free-living and laboratory gait characteristics in patients with multiple sclerosis. *PLoS ONE* 2018; 13: e0196463.
32. Taylor KI, Staunton H, Lipsmeier F, et al. Outcome measures based on digital health technology sensor data: Data- and patient-centric approaches. *NPJ Digit Med* 2020; 3: 97.

10.4.4 Publication N4

Developing a Digital Solution for Remote Assessment in Multiple Sclerosis: From Concept to Software as a Medical Device. van der Walt A, Butzkueven H, Shin RK, Midaglia L, Capezzuto L, Lindemann M, Davies G, Butler LM, Costantino C, Montalban X.Brain Sci. 2021 Sep 21;11(9):1247. doi: 10.3390/brainsci11091247.

brain sciences

MDPI

Review

Developing a Digital Solution for Remote Assessment in Multiple Sclerosis: From Concept to Software as a Medical Device

Anneke van der Walt [1,2,*], Helmut Butzkueven [1], Robert K. Shin [3], Luciana Midaglia [4], Luca Capezzuto [5], Michael Lindemann [5], Geraint Davies [5], Lesley M. Butler [5], Cristina Costantino [5] and Xavier Montalban [6]

1 Department of Neuroscience, Central Clinical School, Monash University, Melbourne, VIC 3004, Australia; helmut.butzkueven@monash.edu
2 The Alfred, Melbourne, VIC 3004, Australia
3 MedStar Georgetown University Hospital, Washington, DC 20007, USA; Robert.K.Shin@gunet.georgetown.edu
4 Servei de Neurologia-Neuroimmunologia, Centre d'Esclerosi Múltiple de Catalunya (Cemcat), Institut de Recerca Vall d'Hebron (VHIR), Hospital Universitari Vall d'Hebron, Universitat Autònoma de Barcelona, 08035 Barcelona, Spain; lmidaglia@cem-cat.org
5 F. Hoffmann-La Roche Ltd., 4070 Basel, Switzerland; luca.capezzuto@roche.com (L.C.); michael.lindemann@roche.com (M.L.); geraint.davies.gd1@roche.com (G.D.); lesley.butler@roche.com (L.M.B.); costantino.cristina@gene.com (C.C.)
6 Multiple Sclerosis Centre of Catalonia (Cemcat), Department of Neurology/Neuroimmunology, Hospital Universitari Vall d'Hebron, Universitat Autònoma de Barcelona, 08035 Barcelona, Spain; xavier.montalban@cem-cat.org
* Correspondence: anneke.vanderwalt@monash.edu; Tel.: +61-3-99030555

check for updates

Citation: van der Walt, A.; Butzkueven, H.; Shin, R.K.; Midaglia, L.; Capezzuto, L.; Lindemann, M.; Davies, G.; Butler, L.M.; Costantino, C.; Montalban, X. Developing a Digital Solution for Remote Assessment in Multiple Sclerosis: From Concept to Software as a Medical Device. *Brain Sci.* **2021**, *11*, 1247. https://doi.org/10.3390/brainsci11091247

Academic Editors: Michelle Ploughman and Stephen D. Mierney

Received: 13 August 2021
Accepted: 16 September 2021
Published: 21 September 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Abstract: There is increasing interest in the development and deployment of digital solutions to improve patient care and facilitate monitoring in medical practice, e.g., by remote observation of disease symptoms in the patients' home environment. Digital health solutions today range from non-regulated wellness applications and research-grade exploratory instruments to regulated software as a medical device (SaMD). This paper discusses the considerations and complexities in developing innovative, effective, and validated SaMD for multiple sclerosis (MS). The development of SaMD requires a formalised approach (design control), inclusive of technical verification and analytical validation to ensure reliability. SaMD must be clinically evaluated, characterised for benefit and risk, and must conform to regulatory requirements associated with device classification. Cybersecurity and data privacy are also critical. Careful consideration of patient and provider needs throughout the design and testing process help developers overcome challenges of adoption in medical practice. Here, we explore the development pathway for SaMD in MS, leveraging experiences from the development of Floodlight™ MS, a continually evolving bundled solution of SaMD for remote functional assessment of MS. The development process will be charted while reflecting on common challenges in the digital space, with a view to providing insights for future developers.

Keywords: multiple sclerosis; software as a medical device; digital health; participatory health; monitoring; smartphone-based assessments; clinical validation; technical validation; MS apps; digital health solution development

1. Introduction

Multiple sclerosis (MS) is an inflammatory demyelinating and degenerative disease [1] characterised by a wide clinical variability in disease trajectory between individuals [2]. Clinical monitoring is intermittently, and often inconsistently [3,4], applied via in-clinic measures, such as the Expanded Disability Status Scale (EDSS) [5] and magnetic resonance imaging; detecting early disease progression is thus challenging [3,4]. Progressive worsening in specific domains (e.g., cognition [6]) can be subtle or subclinical, especially in

the early stages of the disease, but tends to increase in frequency and severity over time. The worsening of disability is a multidimensional process and difficult to detect [7,8]. At present, the diagnosis of progression in MS is typically retrospective with a heavy reliance on clinical history, requiring progressive worsening for more than 6 months based on EDSS score, without evidence of relapses [3,4]. Monitoring in MS relies on infrequent outpatient assessments (typically occurring once or twice annually) with a lack of objective assessments of progression available to healthcare professionals (HCPs). New clinical and research tools are therefore needed to address the unmet need of early detection and ongoing assessment of progressive worsening, rendering this an inviting area for innovation in the digital health space.

Remote digital solutions such as smartphone-based apps, wearables, and decision support algorithms are increasingly utilised in research and clinical trial settings [9] and are beginning to emerge in routine medical care. This paper will focus on smartphone technology, which is ubiquitous and broadly accessible [10,11], making it a viable approach for facilitating remote assessment [12–14]. Smartphones can be used in a patients' home environment as frequently as required and their use is increasingly familiar and unobtrusive. Further, most off-the-shelf smartphones contain sensors with the capacity to gather objective data unaffected by inter- and intra-rater variability. Measurements and patient-reported information captured with smartphone technology have the potential to enable more frequent, decentralised, and home-based care to supplement the infrequent in-clinic assessments typically offered to patients. Mutually sharing this information with patients can help focus the clinical conversation or empower shared decision making. Smartphone-based digital solutions are thus ideally placed to contribute to improving clinical care management for people living with MS (PLwMS) [15,16] and providing personalised healthcare [17].

Today, smartphone applications, performing a variety of functions, are available to support PLwMS. Many of these tools are non-regulated wellness applications designed to support day-to-day disease management, for example through symptom or medication intake tracking, visit-scheduling, provision of disease education, and connectivity to supportive care facilities or patient social media networks [18–23]. Other smartphone-based solutions enable assessment of functional parameters affected by the disease, such as mobility and cognition, or therapeutic benefit, such as for fatigue or depression [24]. Data and digital biomarkers collected by patient-facing apps may provide clinical value by generating new insights into the MS disease course, ultimately improving the understanding of individual disease trajectories and response to intervention. Despite their promise, however, smartphone-based solutions have not yet been fully integrated into routine medical practice.

The development journey of a smartphone-based solution for remote assessment of PLwMS will be presented here as a case study to illustrate the design and development process, validation, regulatory and clinical requirements, as well as deployment in the emergent digital health landscape. The process will be discussed from the perspective of industry developers and academic collaborators, from ideation through to technical solution development and version iteration, certification, and deployment. The Floodlight programme is a Roche-led initiative that aims to create digital solutions to facilitate functional assessment in MS. The first Floodlight app was an assessment suite for clinical research that required provisioned smartphones. More recent versions have been developed under design control to ensure that they meet the regulatory standards of reliability and meaningfulness associated with software as a medical device (SaMD)—standalone software that can perform medical functions without being part of a specific medical device hardware [25]—and to enable access for use on personal smartphones in a variety of integrated MS care settings.

2. Concept, Proof of Concept, and Assessment of Unmet Needs

The MS digital health space is still largely uncharted. Close partnerships between developers, researchers, HCPs, and PLwMS, from inception and throughout the design process, is essential to ensure that technical solutions, such as smartphone apps, are grounded in science and adequately address unmet patient and/or healthcare needs. Technical development typically begins with the identification and prioritisation of user needs, then the ideation of possible solutions, followed by a "design, test, and iterate" build cycle to ensure those needs are fulfilled. For SaMD development, this creative cycle must also be balanced with clinical, technical, and regulatory processes to ensure the required rigour is achieved. In parallel, it must be established that the solution provides output that is meaningful to both PLwMS and HCPs and that can be readily embedded in the relevant healthcare system.

Technical development for SaMD may begin after proof of concept (PoC) has already been established in a research setting. In the Floodlight PoC study, sensor-based measurement was shown to effectively capture reliable and clinically relevant measures of functional impairment in three domains: cognition, gait, and balance, and hand motor function [26,27]. The Floodlight PoC study constituted sufficient evidence to allow for the use of these assessments in a research setting, but further development under design control was required for deployment in a clinical setting as SaMD. Implementing a secondary, more rigorous technical design step also provided evidence to support face validity and inform features that would facilitate the user experience.

The Double Diamond (DD) model (Figure 1), an iterative approach commonly used in software development, was utilised to guide the process of gaining user insights for the design of the new Floodlight solution. DD is a non-linear model based on divergent–convergent thinking, where a topic is first explored more widely or deeply (divergent) before a focused approach is taken with a singular design solution (convergent) [28]. The iterative aspect of development is then retained through the ongoing acquisition and utilisation of new real-world user insights, experiences, and behaviours to inform subsequent refinement of the solution.

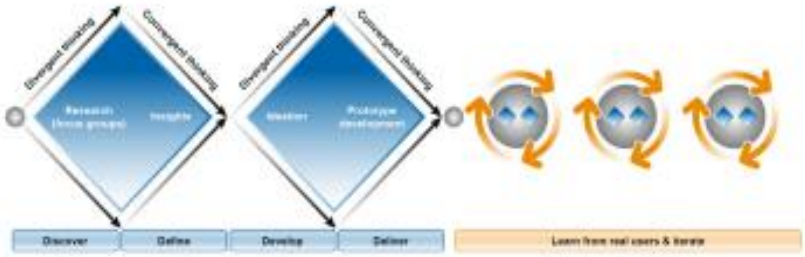


Figure 1. Double Diamond model, an iterative approach with rapid prototyping.

In the initial divergent phase, focus groups with MS experts and PLwMS were conducted to identify the signs and symptoms that might best represent the emergence of progression and to define the current in-clinic standards used to assess functional loss. This process then fed into the convergent phase, which prioritised the assessment of hand motor function, gait, and cognition, all domains frequently affected in PLwMS with worsening disease [29–37]. These findings served to substantiate the selection of domain assessments tested in the Floodlight PoC study. During the second divergent phase, exploration of how to technically design the solution and implement the assessments took place, followed by the prioritisation and consolidation of a singular, defined approach for technical development. In order to inform iterative updates and advancements to the solution in the future, mechanisms, such as an analytics platform, were then incorporated to collect insights from real-world users.

This design effort yielded a preliminary structure for the new Floodlight solution, named Floodlight™ MS (currently v1.2). Floodlight MS would provide five assessments for measurement of function across three domains, as well as a Patient Journal (Figure 2). A "bundling approach", wherein the five assessments would be verified, validated, and independently registered as SaMD, was taken to enable flexibility to change, update, and add new features or assessments without compromising the solution as a whole.

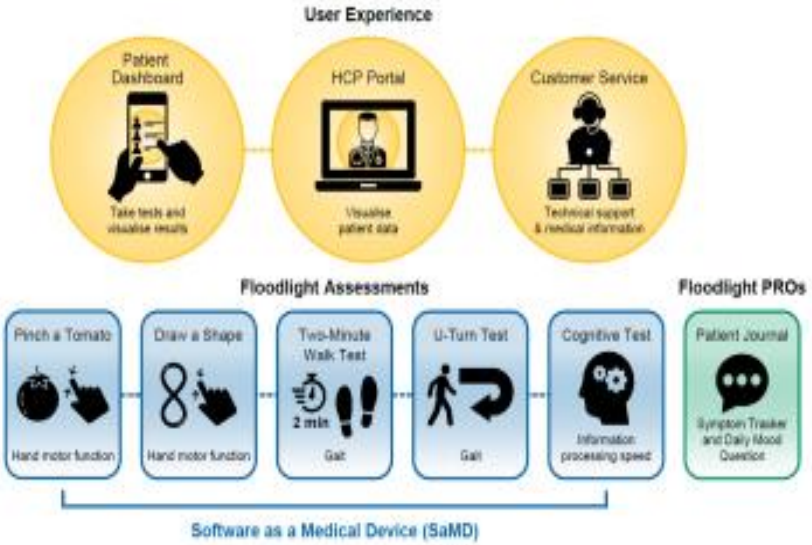


Figure 2. Illustration of current version of Floodlight™ MS v1.2 app and assessments. HCP, healthcare professional; PRO, patient-reported outcome.

3. Desirability: Challenges in Developing a Digital Solution That PLwMS and HCPs Need and Use

Identifying and balancing the needs and desires of different users when creating a digital solution can be challenging. Prior to initiating design control, the Jobs-to-be-Done (JTBD) framework [38] was used to define concrete user need statements for Floodlight MS. JTBD is an outcome-driven innovation strategy used to provide an in-depth understanding of user goals in a structured manner. Core functional desired outcomes (e.g., "Minimize the time it takes to determine how the patient's past symptoms have changed since their last consultation"), as well as emotional and related jobs that might impact the ability to achieve an outcome (e.g., "Avoid feeling guilty for not spending enough time with a patient"), were collected for each user type. For Floodlight MS solution design, users were defined as (1) individuals with MS who are trying to live their lives while managing their MS, and (2) neurologists who are maintaining MS patients' quality of life.

JTBD outcome statements reframed the needs related to management of MS into user needs that can be addressed through a technical solution and that can be used to establish parameters for device quality system requirements. To determine which of the needs was most underserved, the desired outcome statements were quantitatively ranked by 202 PLwMS and 211 HCPs in terms of importance of the outcome and current satisfaction in performing the job. For PLwMS, the most underserved needs concerned gaining a better understanding of their health status and treatment management. For HCPs the most underserved needs included monitoring changes in the health status of PLwMS, assessing the impact of MS on daily life, and driving patient compliance. For Floodlight MS, facilitation of improved conversations between PLwMS and their neurologists emerged as a defining priority.

JTBD analysis also clarified factors that might limit the ability of PLwMS to interact with an app, such as comorbidities and disability status. The findings indicated that assessments within the solution must be convenient, with a reasonable duration and frequency. Different levels of user ability in terms of digital skills, as well as aspects such as dexterity and cognitive and visual impairments, would be likely to impact engagement. For many commercial applications, engagement is a key performance indicator and revenue driver. In digital health, however, solutions should only strive for sufficient engagement to support successful outcomes, in order to strike the balance between benefit and burden to the users.

To ensure safety and effectiveness during use, human factors that may affect an individual user's performance need to be identified and addressed. Errors are frequently caused by the design of the user interface with which users interact. Formative testing with an additional cohort of users is a step in the design control process that serves to identify potential hazardous situations, assess overall usability, and ensure that the interface can be clearly understood and operated per intended use. In formative testing for Floodlight MS, PLwMS expressed satisfaction with an MS-specific solution, which they identified would be a key part of the conversation with their neurologist. PLwMS also indicated that they were more likely to utilise the solution if it were prescribed by an HCP. Formative testing with neurologists was used to assess HCP willingness to adopt the solution and potential barriers to adoption. Neurologists reported that they saw the solution as complementary to their current processes, as long as the data were readily interpretable and easily accessible.

Formative testing informed a significant decision in the developmental journey for the Floodlight programme: the adoption of a prescription-based model for Floodlight MS. This model prioritised partnering with HCPs in a coordinated care setting to identify appropriate patient users, support onboarding and oversee generation and interpretation of patient data. These findings were substantiated by Floodlight Open, a global open-access study, entirely operated via digital interfaces, that was designed to assess adherence to using the app and the feasibility of a "bring your own device" research version of the Floodlight assessment suite provided directly to PLwMS. In line with the adherence issues reported in similar fully digital studies conducted in real-world settings [39], overall adherence in Floodlight Open was low. This contrasts with the controlled environment of the Floodlight PoC study, where good adherence and patient satisfaction were observed [26]. Moreover, in Floodlight Open, adherence rates were positively impacted by concomitant studies that provided clinical coordination. Together, these findings suggested that a supportive clinical care environment would be required to maintain long-term use of the Floodlight MS solution.

Insufficient adherence to remote digital health solutions often presents a challenge to long-term engagement [39]. This is a significant obstacle for developers of apps intended for users with MS, where engagement may be required throughout the user's lifetime. Adherence to the use of a digital health solution over time may be regarded as a behaviour, determined by factors such as the user's motivation, ability, and other aspects such as forgetfulness. Behavioural design is based on insights from behavioural science, which can be implemented to aid in evoking desired user behaviour, and is recognised as a key element for development of digital health solutions to increase the likelihood of achieving the desired outcomes [40–42]. For example, the concept of the "neurological loop" has been used to explain how habits are formed via a three-step loop composed of cue, routine, and reward; solutions can thus be designed to provide users with strategic rewards to elicit repeated behaviour, based on specific cues [43]. A behavioural design approach was adopted to identify features that might enable users of Floodlight MS to achieve the outcome of improving clinical conversations. Fogg's behaviour model (Behaviour = Motivation × Ability × Prompt [44]) was used as a framework to audit the design to identify facilitators and barriers to engagement in terms of motivation, ability, and prompts [44,45]. Feedback architecture was then designed to ensure appropriate communication with users and rewards (motivational prompts) for short-, medium-, and

long-term outcomes. For example, the PLwMS interface home screen was designed to incorporate prompts for action, a progress indicator, and an appointment calendar to orient use of Floodlight MS around the care conversation. Further, notification and content architecture were also devised to sustain motivation across different use cases.

As there is great variability in symptomology and disease course between individuals living with MS, solutions designed for these users need to accommodate diverse characteristics and varied needs, preferences, and behaviours when utilising smartphones. The complexity of addressing individual preferences and needs in a "one-size-fits-all" approach is typified in the end-user reaction to app gamification. The utility of gamification (the use of game design elements in other contexts) is widely discussed in relation to digital health solution development, as it may aid in increasing motivation and sustaining usage (i.e., increasing adherence [46]); however, any elements need to be applied cautiously in the context of healthcare and must support the desired outcome, which, for Floodlight MS, is the use of data for a care conversation. The topic of gamification—where, in the context of the Floodlight programme experience, some users considered it an inappropriate approach to disease assessment—may represent an example of possible divergent perspectives from different users and user types, which further illustrates the importance of behavioural strategies in studying use patterns. Even after final solution design, regular testing of the applied concepts should be conducted to ensure the usability of the features for all users, aligning with the specific needs of PLwMS. Moreover, careful consideration must be given to how the solution is implemented to ensure effective use. The application of behavioural science, for example through built-in analytics, whilst continuing to develop, test, and iterate on a periodic release cycle, will be important to enable iterative development throughout the solution lifecycle.

4. Regulatory Standards: Data Security, Verification and Validation

Challenges and compromises are involved in creating a digital solution that is not only meaningful to end users, but also technically and scientifically robust and aligned with regulatory standards. Digital solutions producing measures adoptable in medical care must meet the standards of device regulatory agencies on design control, cybersecurity and data privacy, risk analysis, and clinical evaluation—all elements considered in the certification of SaMD [25].

To satisfy regulatory requirements, each of the assessments provided by Floodlight MS were subject to technical verification, as well as clinical and analytical validation (Figure 3). Individual assessments and data features were selected for SaMD certification based upon evidence obtained from the Floodlight PoC study and insights from DD.

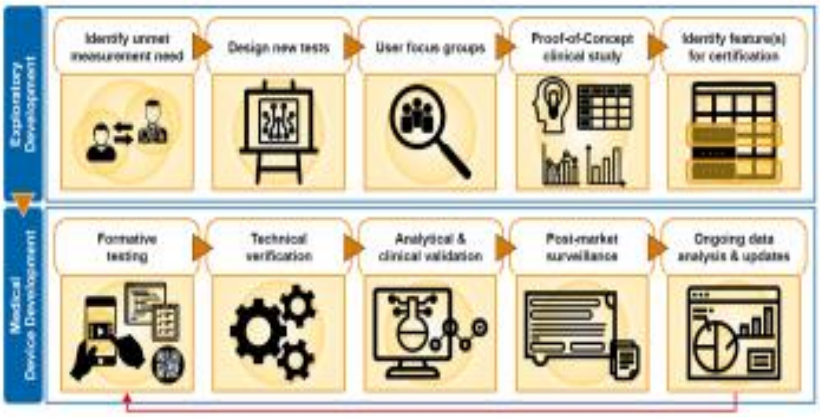


Figure 3. Development of Floodlight™ MS software as a medical device.

Technical verification requires assurance that the software is built to the specified requirements. These include elements such as: software unit testing to ensure the functionality of each component, software system testing of new features or components to expose defects in interfaces and interactions between integrated components, and software verification testing to check that the software meets the specified requirements. In modern software development, an agile process is typically implemented to embed quality testing into development, including efforts such as in-sprint level manual testing (in which incremental development and testing occur in tandem) and automation tests to support regression efforts (a software testing practice that ensures an application still functions as expected after any change).

Analytical validation is needed to ensure SaMD output is reliable, accurate, and precise: it demonstrates how well SaMD fulfil their intended use by accurately measuring the desired parameters and generating the correct outputs [47,48]. Analytical validation includes testing the user experience of the solution. This does not require patient or disease-specific assessment, so testing can be conducted in healthy individuals and/or via simulations. For each of the Floodlight MS assessments, robot testing was conducted across 26 mobile devices, representing over 70% of the global smartphone market. Acceptance criteria consisted of three components: observed variability when the same test is repeated on the same smartphone multiple times (within device error), observed variability when the same test is repeated on different smartphones (between device error), and distance between mean measurement of all smartphones to a theoretical ground truth (systemic bias). All smartphones tested passed acceptance criteria with the exception of the Alcatel 7 phone, which failed due to a device chipset issue where screen sizes are not properly reported by the device, rendering the tomatoes in the Floodlight MS "Pinch A Tomato Test" larger than the acceptable range defined in SaMD specifications. All operating systems were also validated. Testing demonstrated that the operating system of the mobile device did not influence data captured. Finally, a series of security tests were conducted, ranging from threat modelling to penetration testing, to ensure data security in the final product.

Whereas analytical validation establishes reliability, a process of clinical evaluation establishes clinical association and validation and is used to determine the sufficiency of evidence and the requirement for further clinical investigation. Clinical evaluation is a systematic and methodologically sound process used to continuously generate, collect, analyse, and assess the clinical data pertaining to a device in order to verify and validate the safety and performance of the device, including any clinical benefits, in the target user population and when used as intended [49]. Each assessment in the Floodlight MS solution was subject to evaluation, supported by multiple evidence sources, including the Floodlight PoC study and an observational study with PLwMS that provided an assessment by clinical content area experts that the process of, and results from, the Floodlight MS assessments achieved their intended purpose.

Post-marketing surveillance of SaMD is also required to provide ongoing monitoring of any defects and/or safety concerns, in order to ensure that solutions are safe and effective during real-world use. The post-market clinical follow-up plan, which is part of the clinical evaluation, specifies methods and procedures for collection and evaluation of clinical data from on-market use to confirm safety and performance. In addition to implementing subsequent clinical trials and real-world evidence generation initiatives, customer support should be present to capture reportable SaMD events for investigation, such as technical defects and safety issues, in order to fully comply with regulatory requirements. For Floodlight MS, customer support was tailored to respond to users' needs (HCPs and PLwMS) in a specific geography, for example by offering local language support, in addition to addressing technical and medical questions.

Significant effort from the developer is required to achieve robust, regulatory-grade clinical validation. Given the rapidly evolving nature of digital technology, it is also critical for developers to find effective approaches to continuously advance solution design. Likewise, continual technological advancement presents a challenge to regulators, who must

concurrently advance policy in order to foster growth of digital innovation for better disease management [17]. To this end, regulatory agencies are actively facilitating collaborative initiatives within the digital community to advance digital health innovation [50].

Two key aspects of this dialogue are data safety and cybersecurity, as it is crucial to not only establish and maintain robust data privacy and security, but also to adapt them to comply with local requirements across geographies (e.g., General Data Protection Regulation in the EU, the Health Insurance Portability and Accountability Act in the USA, etc.). Data security can be divided into technical (obtaining and storage of data), methodological (the software application and infrastructure used to deliver it), and procedural aspects (data usage, data access, and security breaches), and each of these must be carefully considered at each stage of design and development [51]. Demonstrating a robust approach to personal data security is also a means to build users' trust: 45% of users worry about the unwanted use of their data when using mobile devices for health-related activities [52,53], and there are legitimate concerns over user identification, data sharing with third parties, or accidental data leakage [52]. As data privacy and security provisions must be placed at the fore from the start of a user's interaction with a digital solution, Floodlight MS users are presented with a data privacy notice during the sign-up process. This contains detailed information on the treatment of their personal information, which is in line with the applicable regulatory frameworks. Different types of security measures are in place for Floodlight MS, including password-protected access with automatic logout after a period of inactivity, and data encryption in transit and at rest. Moreover, the legal manufacturer of the device ensures that appropriate training and processes for data management operators are set up and that action plans are ready in case of any incidents.

After regulatory clearance has been achieved, the hurdles of adoption into medical practice, making the solution accessible, and providing ongoing monitoring and managing the system need to be overcome. All these aspects will involve the collaborative efforts of multiple healthcare stakeholders such as PLwMS, HCPs, and payers, as these hurdles cannot be overcome without participation from all relevant parties [17].

5. Taking an Adaptive Approach

Agility is key when developing digital solutions, requiring a fluid approach to facilitate an iterative developmental process that aligns with the design control and requisite regulatory requirements. The rapidly changing technological environment contrasts markedly with classical drug development with its careful, largely linear and standardised processes [54,55]. Once a new solution is developed and deployed, post-market data can serve to further validate clinical effectiveness, evolve technical capabilities, and refine the user experience, and may even support subsequent regulatory engagement and reassessment.

Real-world evidence generation and non-interventional studies are often more efficient than, and can be complementary to, interventional trials. For the Floodlight programme, non-interventional studies and real-world evidence generated with Floodlight MS serves to complement assessment of Floodlight test technology in more formal clinical trial settings. Research is also being conducted to improve the clinical utility of the test suite, support clinical analysis, develop quality control features, and advance the understanding of sensor data [56–59].

The development path outlined here culminated in the release of Floodlight MS, which contains the five assessments registered as SaMD. Additional features, such as the Patient Journal, designed to help users reach the outcome of improved care conversations, are grouped separately and are thus able to be more frequently and flexibly iterated and improved upon based on user feedback, without necessitating resubmission to regulatory authorities. This continual iteration is enabled by recurring development cycles, which allow improvements to be implemented frequently, generating updated versions several times throughout the year. Floodlight MS is set to continue evolving, and thus certain topics covered in this paper may be revisited as knowledge advances.

The use of different technical deployment models, tailoring the means by which the solution is provided, may also provide the flexibility needed to meet local regulatory requirements and enable interoperability and integration in a complex and fragmented electronic health records landscape [60,61]. Three deployment models were designed for Floodlight MS: (i) a standalone solution with an HCP-facing web-based portal for data access, (ii) a standalone solution that can be integrated with electronic health records, and (iii) a software development kit (SDK). The SDK enables rapid, tailored integration of the Floodlight MS assessments into a third-party solution, for example into the DreaMS digital research tool advanced by the Research Center for Clinical Neuroimmunology and Neuroscience, Basel (RC2NB) [62]. The SDK approach also allows for integration into solutions developed in-house. Tailoring the deployment model for a digital solution may enable greater interoperability and the capability to address diverse and local needs on a greater scale.

6. Future Horizons

The provision and adoption of technological solutions and the sharing of information globally has the potential to drive knowledge acquisition and positively affect healthcare worldwide [63]. Digital solutions offer great promise in delivering increasingly individualised, easily accessible, and effective healthcare, with the capacity to evolve with time and adapt to the changing needs of PLwMS and HCPs. The impact of the COVID-19 pandemic has given additional proof of such versatility and usefulness, highlighting how barriers can be overcome through the adoption of digital tools [64], where capturing digital data remotely may mean that symptom tracking can be maintained even when clinic visits are not possible. Ultimately, digital solutions must contribute to the long-term resolution of broader health system challenges, such as lack of access to care, lack of frequent monitoring, costly and ineffective treatment, and delayed diagnosis of MS disease progression.

Fundamentally, digital solutions such as Floodlight MS aim to improve outcomes for PLwMS. To reach this objective, continued collaboration and partnership with the entire MS community is needed, not only to continually refine individual solutions but also to create robust standards for implementation, interpretation, and interoperability. Ongoing investment into the clinical development of a digital solution will also enable continuous improvement, enhancing the clinical utility and sustained actionability of a given solution. This is especially relevant for solutions such as Floodlight MS, which generate data that may be used to improve our understanding of the disease or create digital biomarkers. The use of such data could serve to bridge the gap between clinical trials and medical care, for example, by enabling the creation of a baseline dataset in routine care prior to clinical trial enrolment; or by enabling a more immediate comparison of population outcomes to individual performance using the same reliable, objective outcome measures.

The sharing and secondary use of the data collected using digital solutions will be important for shaping the future of research, regulation, and policymaking in the digital healthcare sphere. Platforms such as the European Health Data Space are being developed to facilitate data sharing across sectors [65]. Key areas centre around health data exchange, access to health data for research and policymaking, and a single market for digital health products and services. Structuring projects to generate data for secondary use in collaboration with the scientific community will help to shape the digital space by furthering research, as well as from a regulatory perspective in terms of enabling better fit-for-purpose and evidence-based policymaking.

The Floodlight programme has undertaken a path of SaMD design and development to support safe and effective use of Floodlight MS, a bundled digital assessment solution for PLwMS. The incorporation of design strategies commonly used in software development informed features to support user adherence, clinical utility, and readiness for integration into today's fragmented global healthcare landscape. This effort is underpinned by an iterative and collaborative clinical validation, technical refinement, and deployment approach intended to drive continual evolution of the technology. Ultimately, the emergence

of robust digital solutions may help to change the way that disease progression is measured in MS, enabling optimisation of care, and helping to bridge clinical trial and medical practice data.

Author Contributions: Writing—Review and Editing, A.v.d.W., H.B., R.K.S., L.M., L.C., M.L., G.D., L.M.B., C.C. and X.M. All authors have read and agreed to the published version of the manuscript.

Funding: The author(s) disclosed receipt of the following financial support for the research, authorship and/or publication of this article: F. Hoffmann-La Roche Ltd., Basel, Switzerland provided financial support for the publication of this manuscript.

Acknowledgments: We would like to thank all the Floodlight partnering institutions and scientific collaborators. We would also like to thank the F. Hoffmann-La Roche Ltd. Floodlight team members, particularly Mike Baker, Camille André, JP Guyon, Seya Colloud, Jorge Cancela, Licinio Craveiro, Alexandre Breton, Susanne Clinch, Jens Schjodt-Eriksen, Catherine Wu, Shibir Desai, Teri Kanerva, Christian Gossens, Timothy Kilchenmann, and Christine Eighteen. Writing and editorial assistance for this manuscript was provided by Terri Desmonds of Articulate Science, UK, and funded by F. Hoffmann-La Roche Ltd.

Conflicts of Interest: The authors declared the following potential conflicts of interest with respect to the research, authorship and/or publication of this article: A.v.d.W. received travel support, speaker honoraria and served on advisory boards for Biogen, Merck, Genzyme, Novartis and Teva. H.B. has received institutional (Monash University) funding from Biogen, F. Hoffmann-La Roche Ltd., Merck and Novartis; has carried out contracted research for Novartis, Merck, F. Hoffmann-La Roche Ltd. and Biogen; has taken part in speakers' bureaus for Biogen, Genzyme, F. Hoffmann-La Roche Ltd. and Merck; has received personal grants from Oxford PharmaGenesis and Biogen (prior to 30 June 2018). R.K.S. has received consulting fees and speaker honoraria from Biogen, Bristol Myers Squibb, EMD Serono, Genentech, Novartis and Sanofi Genzyme. L.M. has nothing to disclose. L.C. is an employee of F. Hoffmann-La Roche Ltd. M.L. is a consultant for F. Hoffmann-La Roche Ltd. via Inovigate. G.D. is an employee of F. Hoffmann-La Roche Ltd. L.M.B. is an employee of F. Hoffmann-La Roche Ltd. C.C. is an employee and shareholder of F. Hoffmann-La Roche Ltd. X.M. has received speaking honoraria and/or travel expenses for participation in scientific meetings, and/or has been a steering committee member of clinical trials, and/or participated in advisory boards of clinical trials in the past years with Actelion, Alexion, Bayer, Biogen, Bristol Myers Squibb/Celgene, EMD Serono, Genzyme, F. Hoffmann-La Roche Ltd., Immunic Therapeutics, Janssen Pharmaceuticals, MedDay, Merck, Mylan, NervGen Pharma Corp., Novartis, Sanofi Genzyme, Teva Pharmaceuticals, TG Therapeutics, EXCEMED, Multiple Sclerosis International Federation and National Multiple Sclerosis Society.

References

1. Filippi, M.; Bar-Or, A.; Piehl, F.; Preziosa, P.; Solari, A.; Vukusic, S.; Rocca, M.A. Multiple sclerosis. *Nat. Rev. Dis. Primers* **2018**, *4*, 43. [CrossRef]
2. Lorscheider, J.; Buzzard, K.; Jokubaitis, V.; Spelman, T.; Havrdova, E.; Horakova, D.; Trojano, M.; Izquierdo, G.; Girard, M.; Duquette, P. Defining secondary progressive multiple sclerosis. *Brain* **2016**, *139*, 2395–2405. [CrossRef]
3. Allen-Philbey, K.; Middleton, R.; Tuite-Dalton, K.; Baker, E.; Sternett, A.; Albor, C.; Schmierer, K. Can we improve the monitoring of people with multiple sclerosis using simple tools, data sharing, and patient engagement? *Front. Neurol.* **2020**, *11*, 464. [CrossRef] [PubMed]
4. Lublin, F.D.; Reingold, S.C.; Cohen, J.A.; Cutter, G.R.; Sørensen, P.S.; Thompson, A.J.; Wolinsky, J.S.; Balcer, L.J.; Banwell, B.; Barkhof, F. Defining the clinical course of multiple sclerosis: The 2013 revisions. *Neurology* **2014**, *83*, 278–286. [CrossRef]
5. Kurtzke, J.F. Rating neurologic impairment in multiple sclerosis: An expanded disability status scale (EDSS). *Neurology* **1983**, *33*, 1444. [CrossRef]
6. Amato, M.P.; Portaccio, E.; Goretti, B.; Zipoli, V.; Hakiki, B.; Giannini, M.; Pastò, L.; Razzolini, L. Cognitive impairment in early stages of multiple sclerosis. *Neurol. Sci.* **2010**, *31*, 211–214. [CrossRef] [PubMed]
7. Inojosa, H.; Proschmann, U.; Akgün, K.; Ziemssen, T. Should we use clinical tools to identify disease progression? *Front. Neurol.* **2021**, *11*, 1890. [CrossRef] [PubMed]
8. Manouchehrinia, A.; Zhu, F.; Piani-Meier, D.; Lange, M.; Silva, D.G.; Carruthers, R.; Glaser, A.; Kingwell, E.; Tremlett, H.; Hillert, J. Predicting risk of secondary progression in multiple sclerosis: A nomogram. *Mult. Scler.* **2019**, *25*, 1102–1112. [CrossRef]
9. Inan, O.; Tenaerts, P.; Prindiville, S.; Reynolds, H.; Dizon, D.; Cooper-Arnold, K.; Turakhia, M.; Pletcher, M.; Preston, K.; Krumholz, H. Digitizing clinical trials. *Jpn. Digit. Med.* **2020**, *3*, 1–7. [CrossRef]

10. Haase, R.; Schultheiss, T.; Kempcke, R.; Thomas, K.; Ziemssen, T. Modern communication technology skills of patients with multiple sclerosis. *Mult. Scler.* **2013**, *19*, 1240. [CrossRef]
11. Center, P.R. Cell Phone and Smartphone Ownership Demographics. Available online: www.pewinternet.org/data-trend/mobile/cell-phone-and-smartphone-ownership-demographics/ (accessed on 15 June 2021).
12. Prasad, S.; Ramachandran, R.; Jennings, C. Development of Smartphone Technology to Monitor Disease Progression in Multiple Sclerosis (P01. 144). *Neurology* **2012**, *78*. [CrossRef]
13. Bove, R.; White, C.C.; Giovannoni, G.; Glanz, B.; Golubchikov, V.; Hujol, J.; Jennings, C.; Langdon, D.; Lee, M.; Legedza, A. Evaluating more naturalistic outcome measures: A 1-year smartphone study in multiple sclerosis. *Neurol. Neuroimmun. Neuroinflamm.* **2015**, *2*, e162. [CrossRef] [PubMed]
14. Boukhvalova, A.K.; Kowalczyk, E.; Harris, T.; Kosa, P.; Wichman, A.; Sandford, M.A.; Memon, A.; Bielekova, B. Identifying and quantifying neurological disability via smartphone. *Front. Neurol.* **2018**, *9*, 740. [CrossRef] [PubMed]
15. Feys, P.; Giovannoni, G.; Dijsselbloem, N.; Centonze, D.; Eelen, P.; Lykke Andersen, S. The importance of a multi-disciplinary perspective and patient activation programmes in MS management. *Mult. Scler.* **2016**, *22*, 34–46. [CrossRef] [PubMed]
16. Marziniak, M.; Brichetto, G.; Feys, P.; Meyding-Lamadé, U.; Vernon, K.; Meuth, S.G. The use of digital and remote communication technologies as a tool for multiple sclerosis management: Narrative review. *JMIR Rehabil. Assist. Technol.* **2018**, *5*, e7805. [CrossRef] [PubMed]
17. Cancela, J.; Charlafti, I.; Colloud, S.; Wu, C. Digital health in the era of personalized healthcare: Opportunities and challenges for bringing research and patient care to a new level. *Digit. Health* **2021**, 7–31. [CrossRef]
18. Scholz, M.; Haase, R.; Schriefer, D.; Voigt, I.; Ziemssen, T. Electronic health interventions in the case of multiple sclerosis: From theory to practice. *Brain Sci.* **2021**, *11*, 180. [CrossRef]
19. Lavorgna, L.; Russo, A.; De Stefano, M.; Lanzillo, R.; Esposito, S.; Moshtari, F.; Rullani, F.; Piscopo, K.; Buonanno, D.; Morra, V.B. Health-related coping and social interaction in people with multiple sclerosis supported by a social network: Pilot study with a new methodological approach. *Interact. J. Med. Res.* **2017**, *6*, e7402. [CrossRef]
20. Allam, A.; Kostova, Z.; Nakamoto, K.; Schulz, P.J. The effect of social support features and gamification on a Web-based intervention for rheumatoid arthritis patients: Randomized controlled trial. *J. Med. Internet Res.* **2015**, *17*, e3510. [CrossRef]
21. Fernandez-Luque, L.; Elahi, N.; Grajales, F., 3rd. An analysis of personal medical information disclosed in YouTube videos created by patients with multiple sclerosis. *Stud. Health Technol. Inform.* **2009**, *150*, 292–296.
22. Biogen. Aby App—By above MS. Available online: https://www.abovems.com/en_us/home/ms-support-events/aby-app.html (accessed on 22 July 2021).
23. Settle, J.R.; Maloni, H.W.; Bedra, M.; Finkelstein, J.; Zhan, M.; Wallin, M.T. Monitoring medication adherence in multiple sclerosis using a novel web-based tool: A pilot study. *J. Telemed. Telecare* **2016**, *22*, 225–233. [CrossRef]
24. De Angelis, M.; Lavorgna, L.; Carotenuto, A.; Petruzzo, M.; Lanzillo, R.; Brescia Morra, V.; Moccia, M. Digital Technology in Clinical Trials for Multiple Sclerosis: Systematic Review. *J. Clin. Med.* **2021**, *10*, 2328. [CrossRef] [PubMed]
25. International Medical Device Regulators Forum. Available online: http://www.imdrf.org/docs/imdrf/final/technical/imdrf-tech-140918-samd-framework-risk-categorization-141013.pdf (accessed on 12 July 2021).
26. Midaglia, L.; Mulero, P.; Montalban, X.; Graves, J.; Hauser, S.L.; Julian, L.; Baker, M.; Schadrack, J.; Gossens, C.; Scotland, A. Adherence and satisfaction of smartphone-and smartwatch-based remote active testing and passive monitoring in people with multiple sclerosis: Nonrandomized interventional feasibility study. *J. Med. Internet Res.* **2019**, *21*, e14863. [CrossRef]
27. Montalban, X.; Graves, J.; Midaglia, L.; Mulero, P.; Julian, L.; Baker, M.; Schadrack, J.; Gossens, C.; Ganzetti, M.; Scotland, A. A smartphone sensor-based digital outcome assessment of multiple sclerosis. *Mult. Scler.* **2021**, 1352458521028561. [CrossRef] [PubMed]
28. British Design Council. What Is the Framework for Innovation? Design Council's Evolved Double Diamond. Available online: https://www.designcouncil.org.uk/news-opinion/what-framework-innovation-design-councils-evolved-double-diamond (accessed on 12 July 2021).
29. Heesen, C.; Böhm, J.; Reich, C.; Kasper, J.; Goebel, M.; Gold, S. Patient perception of bodily functions in multiple sclerosis: Gait and visual function are the most valuable. *Mult. Scler.* **2008**, *14*, 988–991. [CrossRef] [PubMed]
30. Bethoux, F.; Bennett, S. Evaluating walking in patients with multiple sclerosis: Which assessment tools are useful in clinical practice? *Int. J. MS Care* **2011**, *13*, 4–14. [CrossRef] [PubMed]
31. Dubuisson, N.; Baker, D.; Thomson, A.; Marta, M.; Gnanapavan, S.; Turner, B.; Giovannoni, G.; Schmierer, K. Disease modification in advanced MS: Focus on upper limb function. *Mult. Scler.* **2017**, *23*, 1956–1957. [CrossRef]
32. Ortiz-Rubio, A.; Cabrera-Martos, I.; Rodriguez-Torres, J.; Fajardo-Contreras, W.; Díaz-Pelegrina, A.; Valenza, M.C. Effects of a home-based upper limb training program in patients with multiple sclerosis: A randomized controlled trial. *Arch. Phys. Med. Rehabilit.* **2016**, *97*, 2027–2033. [CrossRef]
33. Kister, I.; Bacon, T.E.; Chamot, E.; Salter, A.R.; Cutter, G.R.; Kalina, J.T.; Herbert, J. Natural history of multiple sclerosis symptoms. *Int. J. MS Care* **2013**, *15*, 146–156. [CrossRef] [PubMed]
34. Costa, S.L.; Genova, H.M.; DeLuca, J.; Chiaravalloti, N.D. Information processing speed in multiple sclerosis: Past, present, and future. *Mult. Scler.* **2017**, *23*, 772–789. [CrossRef]
35. Ovejа-Guevara, C.; Ayuso Blanco, T.; Brieva Ruiz, L.; Hernández Pérez, M.Á.; Meca-Lallana, V.; Ramió-Torrentà, L. Cognitive dysfunctions and assessments in multiple sclerosis. *Front. Neurol.* **2019**, *10*, 581. [CrossRef]

36. Patti, F.; Amato, M.; Trojano, M.; Bastianello, S.; Tola, M.; Goretti, B.; Caniatti, L.; Di Monte, E.; Ferrazza, P.; Brescia Morra, V. Cognitive impairment and its relation with disease measures in mildly disabled patients with relapsing—Remitting multiple sclerosis: Baseline results from the Cognitive Impairment in Multiple Sclerosis (COGIMUS) study. *Mult. Scler.* **2009**, *15*, 779–788. [CrossRef] [PubMed]
37. Krishnan, V.; Jaric, S. Hand function in multiple sclerosis: Force coordination in manipulation tasks. *Clin. Neurophysiol.* **2008**, *119*, 2274–2281. [CrossRef] [PubMed]
38. Ulwick, T. What Is Jobs-to-be-Done? Available online: https://jobs-to-be-done.com/what-is-jobs-to-be-done-fea59c8e39eb (accessed on 12 July 2021).
39. Pratap, A.; Neto, E.C.; Snyder, P.; Stepnowsky, C.; Elhadad, N.; Grant, D.; Mohebbi, M.H.; Mooney, S.; Suver, C.; Wilbanks, J. Indicators of retention in remote digital health studies: A cross-study evaluation of 100,000 participants. *Jpn. Digit. Med.* **2020**, *3*, 1–10. [CrossRef] [PubMed]
40. Burrus, O.; Gupta, C.; Ortiz, A.; Zulkiewicz, B.; Furberg, R.; Uhrig, J.; Harshbarger, C.; Lewis, M.A. Principles for developing innovative HIV digital health interventions: The case of Positive Health Check. *Med. Care* **2018**, *56*, 756–760. [CrossRef]
41. Pagoto, S.; Bennett, G.G. How behavioral science can advance digital health. *Transl. Behav. Med.* **2013**, *3*, 271–276. [CrossRef]
42. Klonoff, D.C. Behavioral theory: The missing ingredient for digital health tools to change behavior and increase adherence. *J. Diabetes Sci. Technol.* **2019**, *13*, 276–281. [CrossRef] [PubMed]
43. Duhigg, C. *The Power of Habit: Why We Do What We Do and How to Change*; Random House: New York, NY, USA, 2013; pp. 1–371.
44. Fogg, B.J. *Tiny Habits: The Small Changes That Change Everything*; Eamon Dolan Books: Boston, MA, USA, 2019. [CrossRef]
45. Fogg, B.J. A behavior model for persuasive design. In Proceedings of the 4th International Conference on Persuasive Technology, Claremont, CA, USA, 26–29 April 2009; pp. 1–7. [CrossRef]
46. Cugelman, B. Gamification: What it is and why it matters to digital health behavior change developers. *JMIR Serious Games* **2013**, *1*, e3139. [CrossRef]
47. Mathews, S.C.; McShea, M.J.; Hanley, C.L.; Ravitz, A.; Labrique, A.B.; Cohen, A.B. Digital health: A path to validation. *Jpn. Digit. Med.* **2019**, *2*, 1–9. [CrossRef]
48. Food and Drug Administration. Software as a Medical Device (SAMD). Clinical Evaluation-Guidance for Industry and Food and Drug Administration Staff. Retrieved June; 2017; Volume 29, p. 2020. Available online: https://www.fda.gov/files/medical%20devices/published/Software-as-a-Medical-Device-%28SAMD%29--Clinical-Evaluation---Guidance-for-Industry-and-Food-and-Drug-Administration-Staff.pdf (accessed on 15 September 2021).
49. Article 2 of Regulation (EU) 2017/745-MDR. *Off. J. Eur. Union* **2017**, L 117/1–L 117/175. Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ%3AL%3A2017%3A117%3ATOC (accessed on 15 September 2021).
50. FDA. FDA Digital Devices: Digital Health Center of Excellence Services. Available online: https://www.fda.gov/medical-devices/digital-health-center-excellence/digital-health-center-excellence-services (accessed on 12 July 2021).
51. Bennett, K.; Bennett, A.J.; Griffiths, K.M. Security considerations for e-mental health interventions. *J. Med. Internet Res.* **2010**, *12*, e61. [CrossRef] [PubMed]
52. Commission, E. *Green Paper on Mobile Health ("mHealth")*. 2014, pp. 1–20. Available online: https://digital-strategy.ec.europa.eu/en/library/green-paper-mobile-health-mhealth (accessed on 12 July 2021).
53. Recruitment, B.C.P. *Leveraging Mobile Health Technology for Patient Recruitment*. 2012, pp. 1–16. Available online: https://docplayer.net/10235751-Leveraging-mobile-health-technology-for-patient-recruitment-an-emerging-opportunity.html (accessed on 12 July 2021).
54. Leclerc, O.; Smith, J. How New Biomolecular Platforms and Digital Technologies Are Changing R&D. Available online: https://www.mckinsey.com/industries/pharmaceuticals-and-medical-products/our-insights/how-new-biomolecular-platforms-and-digital-technologies-are-changing-r-and-d (accessed on 12 July 2021).
55. Izmailova, E.S.; Wagner, J.A.; Ammour, N.; Amondikar, N.; Bell-Vlasov, A.; Berman, S.; Bloomfield, D.; Brady, L.S.; Cai, X.; Calle, R.A. Remote digital monitoring for medical product development. *Clin. Transl. Sci.* **2021**, *14*, 94–101. [CrossRef]
56. Cheng, W.-Y.; Bourke, A.K.; Lipsmeier, F.; Bernasconi, C.; Belachew, S.; Gossens, C.; Graves, J.S.; Montalban, X.; Lindemann, M. U-turn speed is a valid and reliable smartphone-based measure of multiple sclerosis-related gait and balance impairment. *Gait Posture* **2021**, *84*, 120–126. [CrossRef] [PubMed]
57. Creagh, A.; Simillion, C.; Scotland, A.; Lipsmeier, F.; Bernasconi, C.; Belachew, S.; van Beek, J.; Baker, M.; Gossens, C.; Lindemann, M. Smartphone-based remote assessment of upper extremity function for multiple sclerosis using the Draw a Shape Test. *Physiol. Meas.* **2020**, *41*, 054002. [CrossRef]
58. Creagh, A.P.; Simillion, C.; Bourke, A.K.; Scotland, A.; Lipsmeier, F.; Bernasconi, C.; van Beek, J.; Baker, M.; Gossens, C.; Lindemann, M. Smartphone-and smartwatch-based remote characterisation of ambulation in multiple sclerosis during the two-minute walk test. *IEEE J. Biomed. Health Inform.* **2020**, *25*, 838–849. [CrossRef]
59. Bourke, A.K.; Scotland, A.; Lipsmeier, F.; Gossens, C.; Lindemann, M. Gait characteristics harvested during a smartphone-based self-administered 2-minute walk test in people with multiple sclerosis: Test-retest reliability and minimum detectable change. *Sensors* **2020**, *20*, 5906. [CrossRef]
60. Bernat, J.L. Ethical and quality pitfalls in electronic health records. *Neurology* **2013**, *80*, 1057–1061. [CrossRef]
61. Romero, M.R.; Staub, A. Specialty Task Force: A Strategic Component to Electronic Health Record (EHR) Optimization. *Stud. Health Technol. Inform.* **2016**, *225*, 1051–1052.

62. D'Souza, M.; Papadopoulou, A.; Girardey, C.; Kappos, L. Standardization and digitization of clinical data in multiple sclerosis. *Nat. Rev. Neurol.* **2021**, *17*, 119–125. [CrossRef]
63. Organisation, W.H. *Digital Implementation Investment Guide (DIIG): Integrating Digital Interventions into Health Programmes*. 2020, pp. 1–182. Available online: https://www.who.int/publications/i/item/9789240010567 (accessed on 12 July 2021).
64. Breuer, R.; Zurkiya, D.N.; Samorezov, J.; Patangay, A.; Zerbi, C.; Company, F.M. Omnichannel Engagement in Medtech: The Time Is Now. Available online: https://www.mckinsey.com/industries/pharmaceuticals-and-medical-products/our-insights/omnichannel-engagement-in-medtech-the-time-is-now (accessed on 12 July 2021).
65. Iacob, N.; Simonelli, F. Towards a European Health Data Ecosystem. *Eur. J. Risk Regul.* **2020**, *11*, 884–893. [CrossRef]

10.5 OTHER CEMCAT EFFORTS IN THE DEVELOPMENT AND DEPLOYMENT OF DIGITAL TECHNOLOGY FOR MS MONITORING IN CLINICAL PRACTICE

Cemcat's commitment to innovation and the implementation of digital technology in the clinical setting is strong and growing. This section describes other projects, beyond Floodlight study, that support it.

10.5.1 Multiple Sclerosis Partners Advancing Technology Health Solutions (MS PATHS)

MS PATHS[324] is a collaborative network including 10 MS centres in the United States and Europe that combines technology and patient participation, providing access to clinical data from a large cohort of people with MS (Table 10.5.1A).

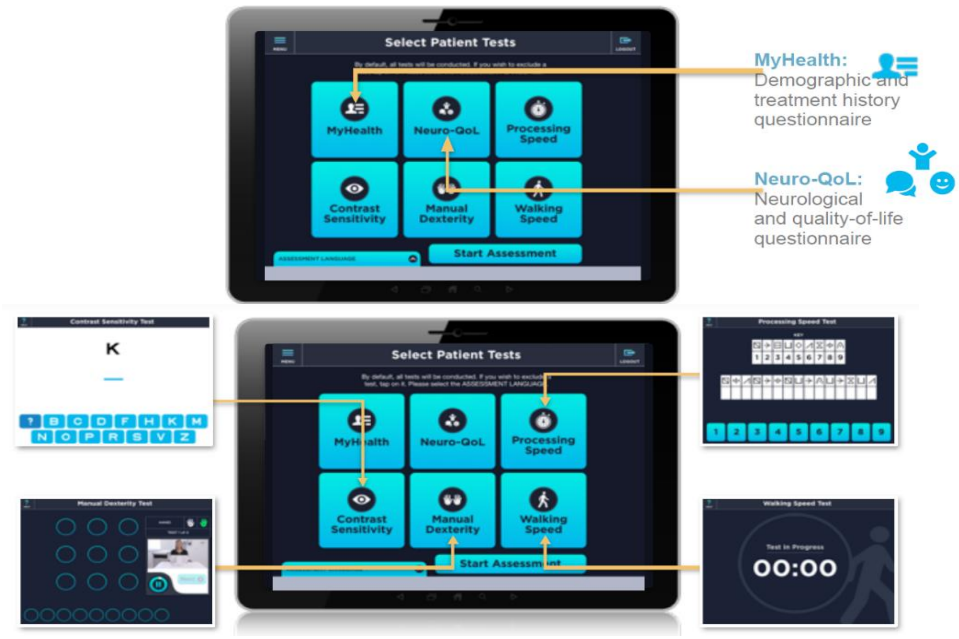
Clinical data are collected by using the Multiple Sclerosis Performance Test (MSPT)[325], an iPad-based assessment designed to be implemented in the clinical setting with minimal supervision. The MSPT (Fig. 10.5.1A) quantifies major MS-associated motor, visual and cognitive symptoms and quality of life outcomes. As part of a structured patient history, requested variables are gender, ethnicity, level of education, employment status, living situation, smoking status, age at MS onset and diagnosis, PDDS, patient-based MS phenotype, MS relapses in the last 12 months and use of DMT for MS as referred by patients. The MSPT also incorporates neuroperformance tests as an electronic adaptation of the MSFC integrating (1) Processing Speed Test (PST), which evaluates cognitive function including elements of attention, psychomotor speed, visual processing, and working memory, and is adapted from the SDMT; (2) Contrast Sensitivity Test (CST), which measures visual acuity with 100% and 2.5% levels of contrast, and is adapted from the Sloan Low Contrast Visual Acuity Test; (3) Manual Dexterity Test (MDT), which assesses upper extremity motor function and is adapted from the 9HPT; and (4) Walking Speed Test (WST), which assesses lower extremity motor function, and is adapted from the T25FW. Finally, patients are asked to complete the Neuro-QoL questionnaire (12 subscales), including the domains of

physical symptoms, fatigue, emotional and cognitive health and social aspects. In addition to clinical data reported by patients, MS PATHS includes the exchange of MRI-standardized data and blood samples for MS biomarkers and genetics research.

Table 10.5.1A. Centres participating in the MS PATHS network.

Europe	United States
University Hospital of Giessen and Marburg Marburg, Germany University Hospital Carl Gustav Carus Dresden, Germany Multiple Sclerosis Centre of Catalonia, Vall d'Hebron Hospital Barcelona, Spain	Cleveland Clinic Cleveland, Ohio New York University New York, New York OhioHealth Columbus, Ohio Cleveland Clinic Lou Ruvo Center for Brain Health Las Vegas, Nevada Johns Hopkins University Baltimore, Maryland University of Rochester Rochester, New York Washington University in St. Louis St. Louis, Missouri

Figure 10.5.1A. All assessments included in the Multiple Sclerosis Performance Test



Rao et al[326], published the validation of the self-administered neuroperformance tests included in the MSPT. They assessed test-retest reliability, practice effects and convergent validity of the CST, MDT and WST in 30 MS patients and 30 healthy controls, examined sensitivity to MS disability in over 600 MS patients as part of their routine clinic assessment and compared performance on the PST in research volunteers and clinical samples. The CST, MDT and WST were shown to be reliable, valid and sensitive to MS outcomes. Performance was comparable to technician-administered testing. PST performance was poorer in the clinical sample compared with the research volunteer sample.

Since 2017, MSPT has been available in Cemcat as part of the clinical practice to portray a more complete picture of the functional situation of the patients. In this way, it is intended that digital technology be a complementary tool that allows, in a simple way, to collect standardized and relevant data, increase the participation of patients in the medical care process and favor the use of quantitative data for clinical decision-making. Additionally, the rigorous collection

and analysis of real-world data may improve our understanding of the disease and accelerate the development of personalized medicine in MS.

MSPT tests are self-managed; raw data is transferred automatically to the cloud, where raw scores and normative adjusted scores are calculated; and the results, graphically and easily interpreted, can be automatically transferred to the electronic medical record in real time for clinical use (Fig. 10.5.1B).

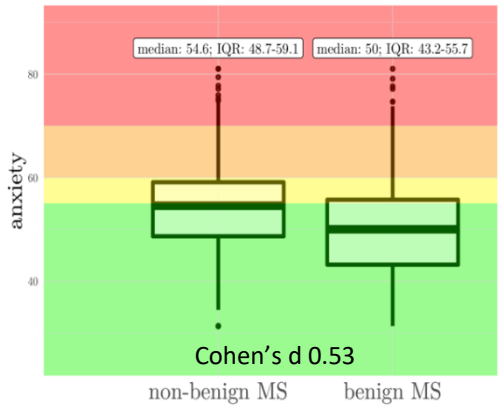
Figure 10.5.1B. Circuit of MS patients during their stay at Cemcat. 1- Arrival of the patient at the Centre. 2- Completion of the self-performed tests through the MSPT with minimal supervision of health personnel. 3- Medical visit with the neurologist who has the results of the MSPT integrated into the electronic history of the patient.

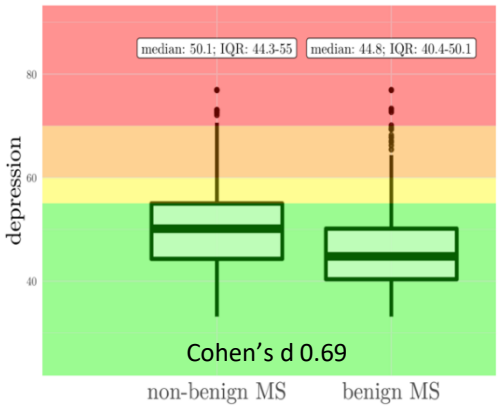
As part of the clinical research promoted by the MS PATHS network, Cemcat have carried out various projects. One of them, "*Defining controversies of benign MS using digital technology*", was presented as an oral communication in the ECTRIMS held in 2020[327]. The aim of the study was to identify benign MS (BMS) patients using PDDS, a PRO of physical disability, as a proxy for EDSS; support the "benign status" describing its physical and non-physical characteristics considering the neuroperformance tests and Neuro-QoL included in the MSPT; and explore the features, among all variables considered, that best discriminate BMS. The results of the study are shown below (Table 10.5.1B and Table 10.5.1C, Fig. 10.5.1C and Fig. 10.5.1D)

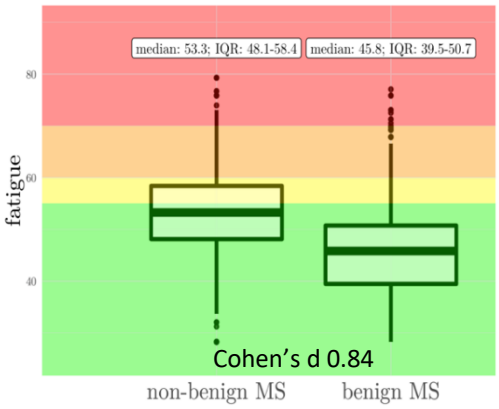
Table 10.5.1B. Results of the neuroperformance tests.

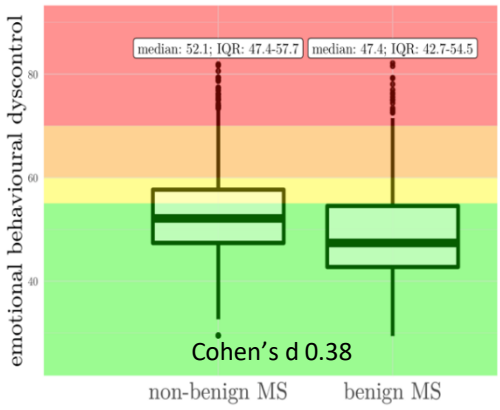
	BMS (*n* = 3852)	**Non-BMS (*n* = 4497)**	**Cohen's d**
Median (IQR) PST	49 (42–57) *n* = 3573 (92.7%)	40 (31–47) *n* = 3880 (86.3%)	0.82
Median (IQR) CST 2.5%	38 (30–44) *n* = 2289 (39.4%)	30 (20–39) *n* = 1970 (43.8%)	0.56
Median (IQR) MDT	24.4 (21.7–28.1) *n* = 3365 (87.3%)	30.6 (26.1–36.1) *n* = 3172 (70.5%)	0.97
Median (IQR) WST	5.6 (4.8–6.8) *n* = 3507 (91%)	8.3 (6.6–11.3) *n* = 3222 (71.6%)	0.81

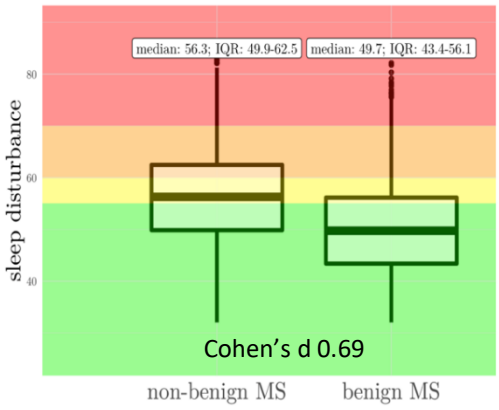
BMS: benign MS, CST: Contrast Sensitivity Test, IQR: *interquartile range*, MDT: Manual Dexterity Test, PST: Processing Speed Test, WST: Walking Speed Test.

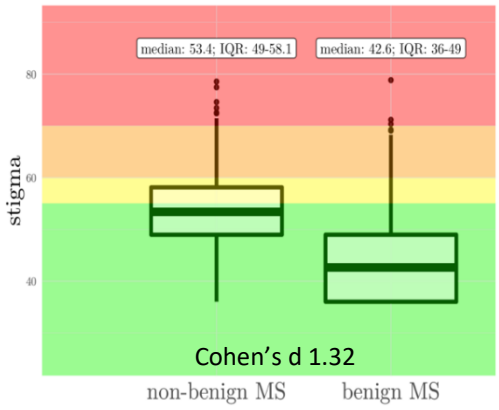

anxiety
median: 54.6; IQR: 48.7-59.1
median: 50; IQR: 43.2-55.7
Cohen's d 0.53
non-benign MS
benign MS

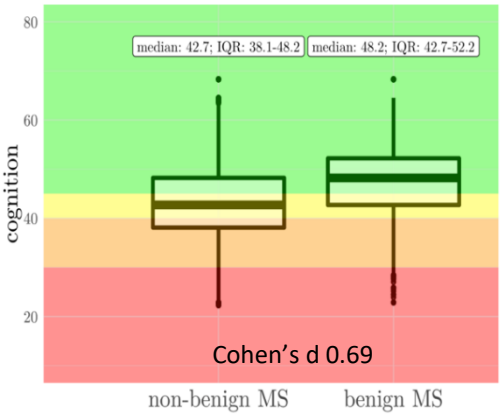

depression
median: 50.1; IQR: 44.3-55
median: 44.8; IQR: 40.4-50.1
Cohen's d 0.69
non-benign MS
benign MS

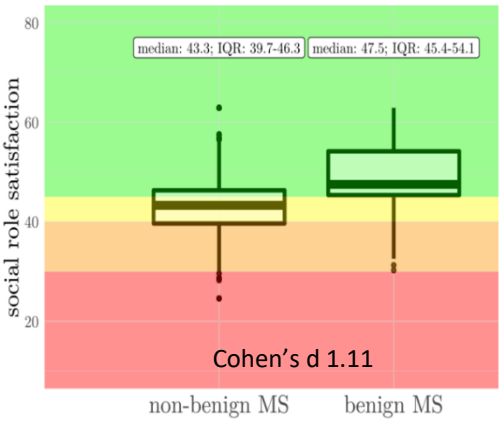

fatigue
median: 53.3; IQR: 48.1-58.4
median: 45.8; IQR: 39.5-50.7
Cohen's d 0.84
non-benign MS
benign MS

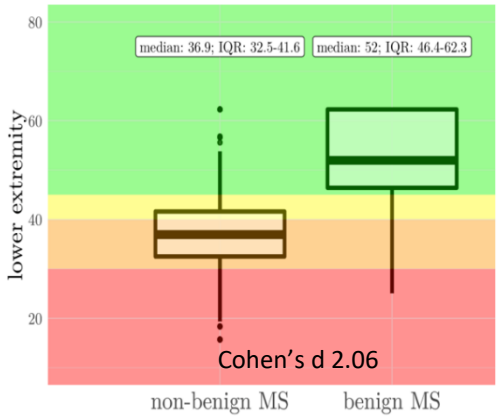

emotional behavioural dyscontrol
median: 52.1; IQR: 47.4-57.7
median: 47.4; IQR: 42.7-54.5
Cohen's d 0.38
non-benign MS
benign MS

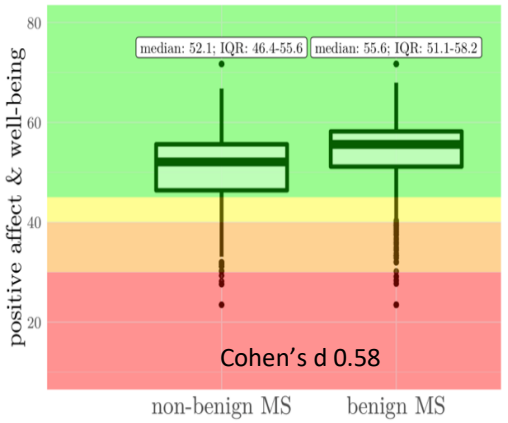

sleep disturbance
median: 56.3; IQR: 49.9-62.5
median: 49.7; IQR: 43.4-56.1
Cohen's d 0.69
non-benign MS
benign MS

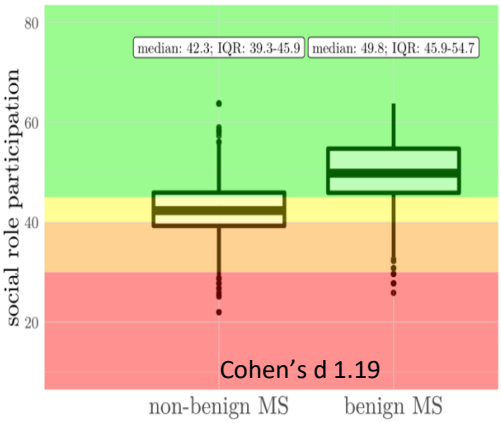

stigma
median: 53.4; IQR: 49-58.1
median: 42.6; IQR: 36-49
Cohen's d 1.32
non-benign MS
benign MS


cognition
median: 42.7; IQR: 38.1-48.2
median: 48.2; IQR: 42.7-52.2
Cohen's d 0.69
non-benign MS
benign MS


social role satisfaction
median: 43.3; IQR: 39.7-46.3
median: 47.5; IQR: 45.4-54.1
Cohen's d 1.11
non-benign MS
benign MS


lower extremity
median: 36.9; IQR: 32.5-41.6
median: 52; IQR: 46.4-62.3
Cohen's d 2.06
non-benign MS
benign MS


positive affect & well-being
median: 52.1; IQR: 46.4-55.6
median: 55.6; IQR: 51.1-58.2
Cohen's d 0.58
non-benign MS
benign MS


social role participation
median: 42.3; IQR: 39.3-45.9
median: 49.8; IQR: 45.9-54.7
Cohen's d 1.19
non-benign MS
benign MS

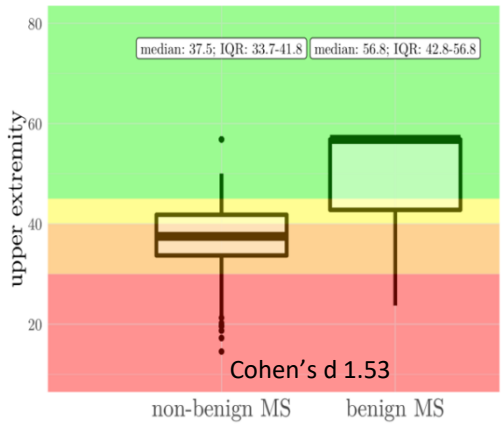

upper extremity
median: 37.5; IQR: 33.7-41.8
median: 56.8; IQR: 42.8-56.8
Cohen's d 1.53
non-benign MS
benign MS

Figure 10.5.1C. Neuro-QoL T-scores: results of the 12 subscales (including physical and non-physical domains) for non-BMS and BMS patients. *Reference population. Scores 0.5–1.0 SD worse than the mean = mild symptoms/impairment (yellow). Scores 1.0–2.0 SD worse than the mean = moderate symptoms/impairment (orange). Scores ≥2.0 SD worse than the mean = severe symptoms/impairment (red).

Table 10.5.1C. Percentages of BMS and non-BMS patients with at least mild symptoms/impairment on the Neuro-QoL subscales.

	BMS (*n* = 3187/3852)	**Non-BMS (*n* = 3377/4497)**
Physical domain (%)		
Lower extremity motor function	21.7	88.7
Upper extremity motor function	31.6	87.3
Non-physical domain (%)		
Anxiety	27.9	46.7
Depression	8.2	24.8
Fatigue	17.1	43.3
Emotional behavioural dyscontrol	24.4	37.7
Sleep disturbance	28.6	55.7
Stigma	6.9	40.0
Cognition	34.2	62.2
Social role satisfaction	23.7	69.6
Positive affect and well-being	7.6	20.0
Social role participation	21.6	70.3

BMS: benign MS

Of note, although BMS patients behaved better than those considered non-BMS for all the analyzed tests and questionnaire from MSPT, also many patients self-defined as BMS reported at least mild impairment or symptoms in different domains evaluated by Neuro-QoL. This suggests that a definition of benignity mainly based on physical aspects, predominantly gait, does not reflect the real functional status of the MS patient. In addition, when we looked for the variables that best discriminated a population with BMS, we found that those related to the patient's QoL, such as motor function of the extremities, as well as social aspects and stigma, had a greater impact. This suggests that PROs (in this case the Neuro-QoL) outperform objective tests of MS (such as WST or MDT) when exploring the most representative characteristics of BMS.

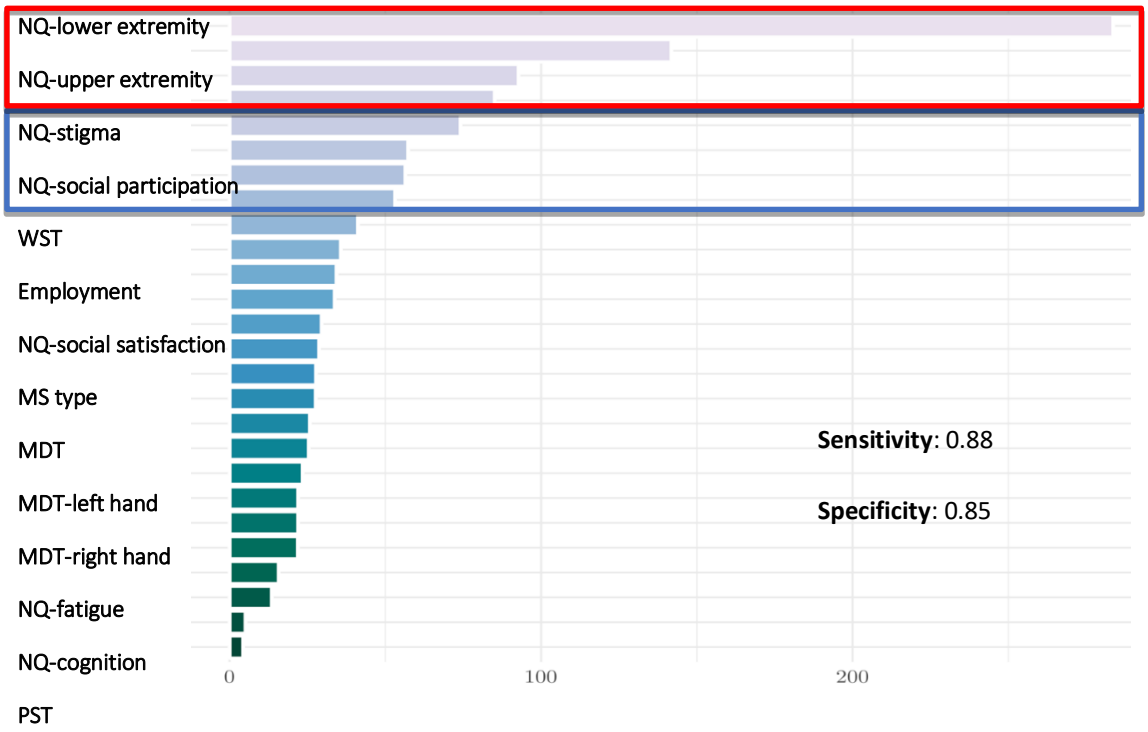


Figure 10.5.1D. Patient characteristics that better discriminate between BMS and non-BMS by the random forest model. NQ: Neuro-QoL. WST: Walking speed test. MS: multiple sclerosis. MDT: Manual dexterity test. PST: Processing speed test.

10.5.2 RADAR-CNS Remote Assessment of Disease and Relapse - Central Nervous System

Other technological initiative launched in 2018 is the Remote Assessment of Disease and Relapse-Central Nervous System (RADAR-CNS)[328], a major international academic-industry research program aimed at developing novel methods and infrastructure for monitoring major depressive disorder, epilepsy, and MS using wearable devices and smartphone technology. Cemcat's collaboration is focused on the MS area.

During the study, participants are required to regularly use the RADAR-tool (Fig. 10.5.2A), which is a cross-platform Cordova app for active monitoring (active remote monitoring technology) through conscious action (eg, questionnaires, audio questions, timed tests) and a native Android app for passive monitoring via phone and wearable sensors (passive remote monitoring technology). RADAR-

tool also includes capabilities for data aggregation, management of studies, and real-time visualizations.

The RADAR-tool platform can be deployed both in local settings, such as a hospital, for local data collection or for ambulatory studies through remote deployment for centralized data collection.

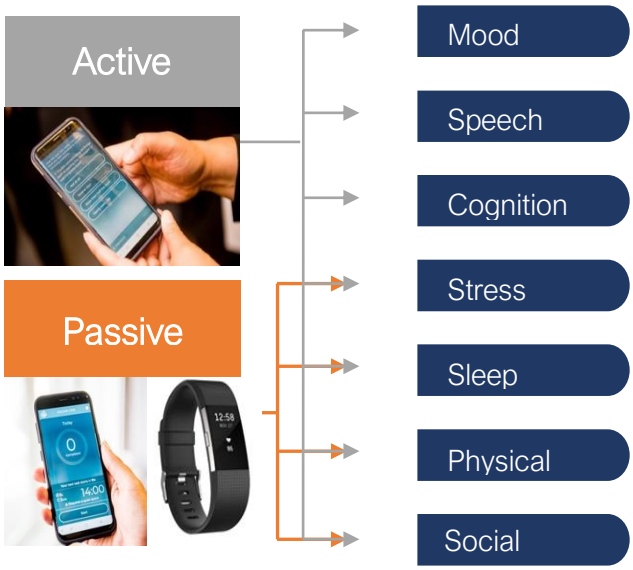


Figure 10.5.2A. Domains explored in the RADAR-CNS project.

The main objectives of the RADAR-CNS project[328] are to determine the usability, feasibility and acceptability of, and adherence to, a RMT to provide real-time objective multidimensional indications of clinical state in pwMS and to explore the potential impact of clinical data collected through different electronic devices when making medical / therapeutic decisions in routine clinical practice.

In the MS area, the RADAR-CNS project presents 2 sub-studies:

1- Depression substudy, focused on the appearance of relapses, the evaluation of social activity and participation, the pattern of sleep, cognition and the pattern of the voice.

2- Disability substudy, focused on flare-ups, disability progression (EDSS), evaluation of gait and balance, fatigue, cognition, and voice pattern.

One of the first published studies related to this project, showed the capacity of the RADAR-CNS tool for monitoring the behavior of MS patients during the SARS-CoV-2 virus pandemic[329]. For that purpose, data were extracted from smartphone and wearable devices, and managed by the RADAR-base from 1,062 participants recruited in Italy, Spain, Denmark, the United Kingdom, and the Netherlands. Nine features collected on a daily basis including time spent at home, maximum distance travelled from home, the maximum number of Bluetooth-enabled nearby devices (as a proxy for physical distancing), step count, average heart rate, sleep duration, bedtime, phone unlock duration, and social app use duration were analyzed. Kruskal-Wallis tests followed by post hoc Dunn tests were performed to assess differences in these features among baseline, prelockdown, and during lockdown periods. The analyses showed (see Fig. 10.5.2B for results of participants from Spain) reduced sociality as measured through mobility features and increased virtual sociality through phone use. People were more active on their phones (p<0.001 for Italy, Spain, and the United Kingdom), spending more time using social media apps (p<0.001 for Italy, Spain, the United Kingdom, and the Netherlands), particularly around major news events. Furthermore, participants had a lower heart rate (p<0.001 for Italy and Spain; p=0.02 for Denmark), went to bed later (p<0.001 for Italy, Spain, the United Kingdom, and the Netherlands), and slept more (p<0.001 for Italy, Spain, and the United Kingdom). The study also found that young people had longer homestay than older people during the lockdown and fewer daily steps. So, the study suggests that RADAR-base can be used to rapidly quantify and provide a holistic view of people behavioral changes, in this study, of MS patients.

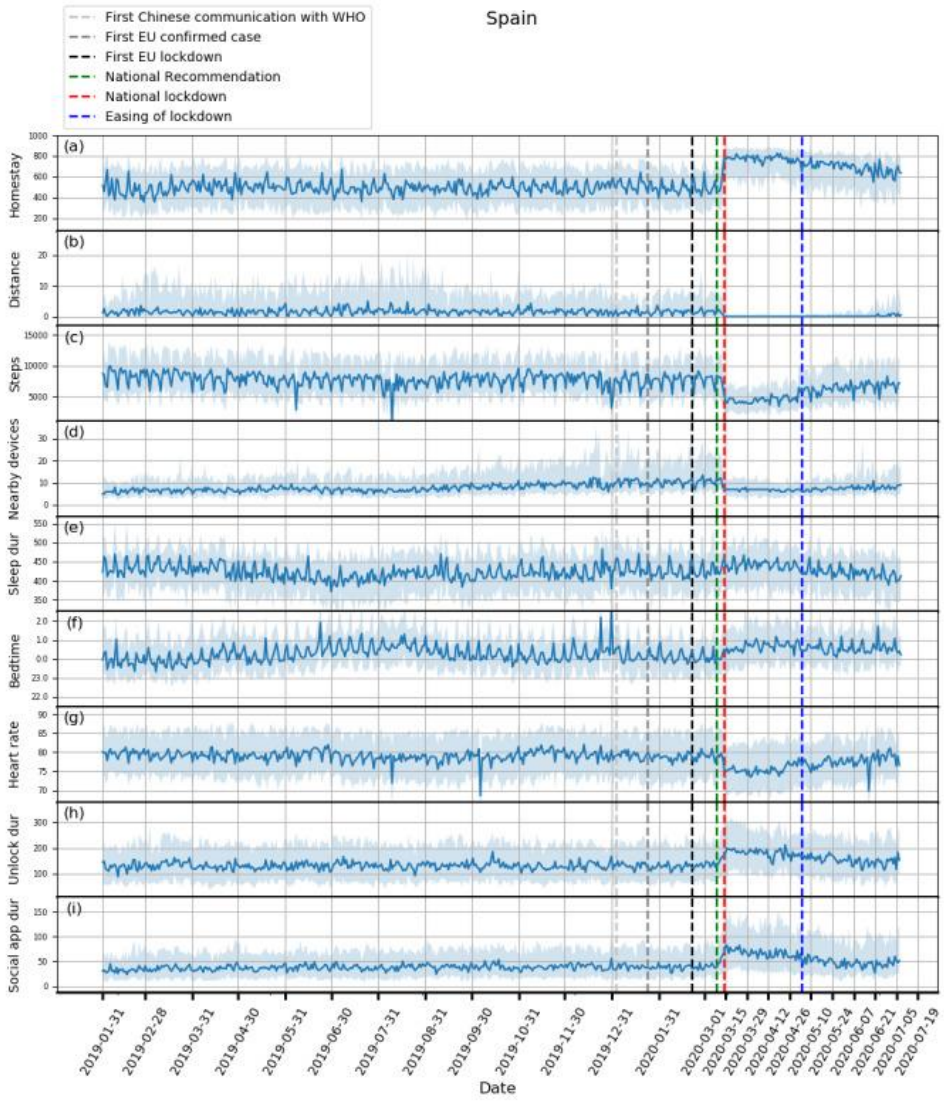


Figure 10.5.2B. Behavioral changes for Spain (329 participants). (a) homestay duration, (b) maximum distance from home, (c) Fitbit step count, (d) maximum number of nearby devices, (e) total sleep duration, (f) bedtime, (g) heart rate, (h) unlock duration, (i) social app duration. Solid line: median; shade: 25th percentile to 75th percentile. dur: duration; WHO: World Health Organization.