Check for updates

Original research paper

MULTIPLE SCLEROSIS JOURNAL | MSJ

Validation of generative artificial intelligence techniques for synthetic data generation in multiple sclerosis research: A comparison with real-world evidence from the Italian multiple sclerosis registry

Multiple Sclerosis Journal
1–10
© The Author(s) 2026
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/13524585261492525
journals.sagepub.com/home/msj
Sage

Pietro Iaffaldano[1], Saverio D'Amico[2,3], Giuseppe Lucisano[1,4], Massimiliano Copetti[5], Tommaso Guerra[1], Maria A. Rocca[6,7,8], Francesco Patti[9,10], Giovanna De Luca[11], Diana Ferraro[12], Rocco Totaro[13], Vincenzo Brescia Morra[14], Giuseppe Salemi[15], Emilio Portaccio[16], Matteo Foschi[17,18], Matilde Inglese[19,20], Maria G. Coniglio[21], Clara G. Chisari[9,10], Francesca Caputo[1], Damiano Paolicelli[1], Mario A. Battaglia[22], Matteo G. Della Porta[2,23], Victor Savevski[2], Mattia Delleani[3], Filippo E. Colella[3,23], Elisabetta Sauta[2], Maria P. Amato[16,24], Massimo Filippi[6,7,8], and Maria Trojano[1]; on behalf of the Italian Multiple Sclerosis Register

Abstract

Background: The use of generative artificial intelligence to create synthetic data is an emerging strategy to enhance real-world evidence.

Objective: To assess the validity of artificial intelligence-generated synthetic data in replicating data collected in the Italian Multiple Sclerosis and Related Disorders Register for the evaluation of progression independent of relapse activity risk under early intensive treatment and escalation treatment strategies.

[1]Department of Translational Biomedicine and Neurosciences (DiBraiN), Multiple Sclerosis Center, Policlinico of Bari, University of Bari Aldo Moro, Italy
[2]IRCCS Humanitas Research Hospital, Milan, Italy
[3]Train S.r.l., Milan, Italy
[4]Center for Outcomes Research and Clinical Epidemiology–CORESEARCH SRL, Pescara, Italy
[5]Unit of Biostatistics, Fondazione IRCCS "Casa Sollievo della Sofferenza", San Giovanni Rotondo, Italy
[6]Neuroimaging Research Unit, Division of Neuroscience, IRCCS San Raffaele Scientific Institute, Milan, Italy
[7]Vita-Salute San Raffaele University, Milan, Italy
[8]Neurology Unit, IRCCS San Raffaele Scientific Institute, Milan, Italy
[9]Dipartimento di Scienze Mediche e Chirurgiche e Tecnologie Avanzate, GF Ingrassia, University of Catania, Italy
[10]UOS Sclerosi Multipla, AOU Policlinico "G Rodolico-San Marco", Università di Catania, Italy
[11]Centro Sclerosi Multipla, Clinica Neurologica, Policlinico SS Annunziata, Università "G. d'Annunzio", Chieti-Pescara, Italy
[12]Dipartimento di Neuroscienze, Ospedale Civile di Baggiovara, Azienda Ospedaliera-Università di Modena, Italy
[13]Centro Malattie Demielinizzanti, Ospedale San Salvatore, L'Aquila, Italy
[14]Department of Neurosciences, Reproductive and Odontostomatological Sciences, Multiple Sclerosis Clinical Care and Research Center, University of Naples Federico II, Italy
[15]Department of Biomedicine, Neurosciences and Advanced Diagnostics, University of Palermo, Italy
[16]Department of NEUROFARBA, University of Florence, Italy
[17]Neurology Unit, Department of Neuroscience, Multiple Sclerosis Center, Santa Maria delle Croci Hospital, Azienda USL Romagna, Ravenna, Italy
[18]Department of Biotechnological and Applied Clinical Sciences, University of L'Aquila, Italy
[19]Dipartimento Di Neuroscienze, Riabilitazione, Oftalmologia, Genetica E Scienze Materno–Infantili (DINOGMI), Università di Genova, Italy
[20]IRCCS, Ospedale Policlinico San Martino, Genova, Italy
[21]Center for Multiple Sclerosis, ASM P.O. "Madonna Delle Grazie", Matera, Italy
[22]Scientific Research Area, Italian Multiple Sclerosis Foundation (FISM), Genova, Italy
[23]Department of Biomedical Sciences, Humanitas University, Milan, Italy
[24]IRCCS Fondazione Don Carlo Gnocchi, Florence, Italy

Corresponding author:
Pietro Iaffaldano, Department of Translational Biomedicine and Neurosciences (DiBraiN), Multiple Sclerosis Center, Policlinico of Bari, University of Bari Aldo Moro, Bari 70121, Italy.
Email: pietro.iaffaldano@uniba.it

Data Availability Statement included at the end of the article.

Methods: Artificial intelligence generative models were trained on a sub-cohort of 1666 Italian Multiple Sclerosis and Related Disorders Register patients to generate a synthetic dataset of 4878 patients. Synthetic vAlidation FramEwork, Clinical Synthetic Fidelity, and Nearest Neighbor Distance Ratio were applied. Comparisons between treatment strategies were conducted using both real and synthetic data, focusing on progression independent of relapse activity risk.
Results: The synthetic data demonstrated fidelity and privacy preservation. Treatment-effect estimates for escalation treatment versus early intensive treatment were similar across the real and synthetic datasets. Cox proportional hazards models showed that synthetic data preserved the robustness of the estimates for the risk of a first progression independent of relapse activity event. Associations with radiological risk factors were broadly similar across the real and synthetic datasets.
Conclusion: Artificial intelligence-generated synthetic data reliably replicated treatment-effect outcomes of Italian Multiple Sclerosis and Related Disorders Register data, supporting their potential use in clinically defined research settings.

Keywords
multiple sclerosis, synthetic data, disease registries, progression independent of relapse activity, treatment strategies

Received: 30 January 2026; revised: 5 September 2026; accepted: 11 September 2026

Introduction

Neuroinflammation and neurodegeneration phenomena define the course of multiple sclerosis (MS).[1,2] From the earliest stages of the disease and throughout its progression, irreversible neurological impairment accumulates as a result of both relapse-related worsening (RAW) and progression independent of relapse activity (PIRA) events.[3–5] PIRA is becoming more widely acknowledged as the primary cause of disability progression: understanding how disease-modifying therapies (DMTs) can impact PIRA events is crucial to unraveling the intricate pathways of silent progression and improving treatment approaches in MS.[6] Real-world evidence (RWE) from large MS registries provided significant insights into disease progression in recent years.[7] However, real-world datasets are frequently plagued by missing information and discrepancies, and data sharing between registries and other data sources is often restricted due to data privacy issues. Generative artificial intelligence (AI) offers a promising avenue to overcome such issues by producing synthetic data (SD), trained to learn the key features of a real source dataset, and offers a potential solution to circumvent some of these challenges.[8–10] The application of AI-driven SD in various research fields, such as hematology, oncology, and neurology, has shown encouraging results, demonstrating that SD can improve methodologies while preserving high fidelity to real-world distributions[11–13] and can accelerate research focused on personalized precision medicine. In this study, we aimed to (1) apply innovative synthetic data generation methods to the real data collected in the Italian Multiple Sclerosis and Related Disorders Register (RISM); (2) implement a validation framework to evaluate synthetic data fidelity and privacy preservability; and (3) evaluate the validity of AI-generated SD to replicate and to compare the risk of PIRA between early intensive treatment (EIT) versus escalation treatment strategy (ESC) in both real and synthetic MS cohorts.

Methods

Data source and study population

This was a study based on data extracted from the RISM database. RISM was approved by the ethical committee at the "Azienda Ospedaliero – Universitaria – Policlinico of Bari" (Study REGISTRO SM001—approved on 8 July 2016) and by local ethics committees in all participating centers. Patients signed an informed consent allowing the collection and use of their clinical data for research purposes. According to the Registry rules, the Scientific Committee of the RISM register approved to conduct this project and extract and use the registry data. Data extraction for this study was executed in March 2023. The selection criteria for patients included relapsing-onset MS patients with a minimum follow-up of 5 years, an initial DMT prescription within 3 years of disease onset, and at least three Expanded Disability Status Scale (EDSS) evaluations following treatment initiation. Patient data retrieved from the register included demographic, clinical, magnetic resonance imaging (MRI), and treatment-related characteristics. This study adhered to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) checklist for cohort studies.

Synthetic data generation and validation

Generative Adversarial Networks (GAN) and Tabular Generative Pretrained Transformer (T-GPT) were implemented as generative models and were trained

on a sub-cohort of patients with complete MRI tabular information to generate synthetic data of MS patients. MRI information refers to tabular variables extracted from clinical practice records and entered into the RISM. The T-GPT was chosen due to better performance and for the nature of clinical data. The generative model was trained on patients with complete data to generate new synthetic cohorts with complete information. Both GAN and T-GPT were trained on the real patient cohort to learn the joint multivariable distribution of the data, and synthetic patients were then generated by sampling from these learned cohort-level distributions rather than by using individual real patients as templates for one or more synthetic patients. Full architectures, preprocessing, hyperparameters, and training settings are provided in Supplementary Material S1 (SF3, SF4, SF5).

The synthetic cohort was evaluated using the Synthetic vAlidation FramEwork powered by Train (SAFE)[14] to assess statistical fidelity, clinical utility, and privacy preservability, ensuring that the generated data accurately reflected real-world distributions without compromising patient confidentiality. Statistical and clinical fidelity were measured using the Clinical Synthetic Fidelity (CSF) score, which averages multiple statistical tests; an optimal score is considered >90%. To assess privacy and the risk of a synthetic record matching an original patient, we first measured the identical match share (IMS) (exact matches between synthetic and original data). We also calculated the distance to the closest record (DCR), which is the absolute distance from a synthetic record to its nearest original counterpart. Furthermore, we used the Nearest Neighbor Distance Ratio (NNDR) as the ratio of the distances from a synthetic record to its nearest and second-nearest neighbors. The NNDR optimal range is 0.60–0.85, allowing for a fair comparison of inliers and outliers in the population.[15] SAFE library specifications and details are reported in Supplementary SF6. Variables supplied to the generative model included clinical, treatment, MRI, and outcome variables, including time to first PIRA (Supplementary File 5).

Clinical validation and statistical analysis

For clinical application of synthetic data, we replicated the risk of PIRA comparison analysis of EIT versus ESC in real and synthetic MS cohorts. Patients were categorized into two EIT and ESC groups based on the initial DMT prescribed. The EIT cohort included patients who initiated treatment with high efficacy (HE)-DMTs: natalizumab, alemtuzumab, ocrelizumab, cladribine, fingolimod, or mitoxantrone. Conversely, the ESC group included patients initially treated with moderate efficacy (ME)-DMTs, such as interferon-beta products, glatiramer acetate, teriflunomide, and dimethyl fumarate, and escalated to HE-DMTs after at least 1 year of ME-DMTs before escalation due to treatment failure. Patients in the EIT group were required to remain on treatment for at least 1 month before entering the analysis. Confirmed disability accrual (CDA) was defined as a confirmed 6-month disability increase from study baseline, measured by EDSS (increase $\geqslant$1.5 points with baseline EDSS = 0; $\geqslant$1.0 point with baseline EDSS $\geqslant$1.0 and $\leqslant$5.5; $\geqslant$0.5 point with baseline EDSS $\geqslant$6.0).

The date of CDA was assigned at the first EDSS when an increase was registered. PIRA, the primary outcome, was defined as a CDA event occurring more than 90 days after and more than 30 days before the onset of a relapse. The baseline was represented by the first prescription of DMT.

In descriptive analyses, categorical data were expressed as frequency and proportion. Continuous data were expressed as median and interquartile range (IQR).

Clinical characteristics were compared and tested between the EIT and ESC groups by the chi-square and *t*-test, for categorical and continuous variables, respectively. For continuous variables, the normal distribution assumption was checked by the Shapiro–Wilk test and visual inspection of the Q-Q plot.

For correlation analysis, the relationships between the columns of the dataset were evaluated by building correlation matrices. Since the clinical and demographic features had different data types (i.e. continuous and categorical), we used different methods to calculate the correlation metric. We calculated the correlation of continuous-continuous features using Pearson's correlation. For continuous-categorical features, we used the Correlation Ratio, and for categorical-categorical features, we performed the Uncertainty Coefficient proposed by Theil's U. We created two correlation matrices for real and synthetic datasets, respectively, and then we calculated the difference between those matrices. Metric was calculated as the average of all values in the difference matrix. Cox proportional hazards (PH) models were used to assess the risk of PIRA in EIT and ESC groups on a cohort of patients with non-missing MRI information ($n = 1666$).

The following variables have been included as covariates in the models: gender (male vs female), age and EDSS at first DMT start, time to first DMT start from onset, number of relapses in the 2 years before DMT start (0, 1, or $\geqslant$2), onset type (multifocal vs monofocal), T2 lesion number (0, 1–2, 3–8, or $\geqslant$9), gadolinium (GD)-enhancing T1 lesions (presence vs absence) at the MRI performed before DMT start, and treatment strategy (ESC as the reference).

In addition, the Kaplan–Meier estimates of time to PIRA were obtained for the real and synthetic cohorts, stratified by treatment strategy. The PH assumption was checked by testing the logarithm of time interaction with covariates and by Schoenfeld residuals. Results of Cox regression models were expressed as hazard ratio (HR) and 95% confidence interval (95%CI) of reaching the

outcomes. The homogeneity of HR across datasets was assessed to validate SD.

p-values < 0.05 result in statistical significance. The software used to perform the analyses is R 4.2.3.

Results

Clinical data of 79,001 patients were available in the RISM at the time of data extraction. After applying the inclusion criteria, we retrieved a cohort of 4878 relapsing MS patients. Among them, 914 patients were treated with an EIT approach, while 3964 patients followed an ESC strategy. Baseline characteristics, stratified on the basis of the treatment strategy, are shown in Table 1. The two groups differed significantly in several baseline characteristics. Patients in the EIT group were slightly older at the initiation of the first DMT compared to those in the ESC group (30.85 vs 29.70 years, $p = 0.0018$). The proportion of patients over 40 years old was higher in the EIT group (25.82% vs 20.41%), whereas the proportion of younger patients (up to 20 years old) was lower (11.27% vs 13.07%). No significant difference was found in gender distribution between the two groups ($p = 0.9549$). Patients in the EIT group presented higher median EDSS scores compared to the ESC group at the first DMT start (2.5 vs 1.5, respectively). At baseline, the lesion burden was also significantly higher for patients in the EIT group, with a larger proportion presenting with at least nine T2 lesions compared to the ESC group (21.44% vs 16.15%, respectively).

A sub-cohort of 1666 MS patients with non-missing MRI information (tabularized) was selected and was used to train a T-GPT to generate a synthetic dataset of 4878 patients with complete data. The flowchart reporting patients' selection procedure and cohorts' definition is shown in Figure 1. SD demonstrated a strong resemblance to real distributions, with high alignment in variables such as age at DMT initiation, sex distribution, baseline EDSS, lesion load at baseline MRI, and time to first PIRA event (Figure 2). SD closely mirrored real distributions, with a CSF score of 97% confirming high fidelity. Privacy was preserved, with an NNDR value of 0.61 suggesting a low risk of identifiable patient data leakage. The robustness of the synthetic data cohort was also highlighted by the Spearman/Pearson/Kendall correlation matrices shown in Figure 3. Both the strength and direction of correlation were very similar in SD and real data, and the correlation differences between SD and real data were close to zero, as shown in Figure 3.

Cox PH models were used to assess the impact of treatment strategy on PIRA risk in the real cohort and SD group. The analyses included all covariates reported in the "Methods" section. The age at DMT start emerged as a risk factor of PIRA (HR, 95% CI = 1.02, 1.01–1.03, $p < 0.001$ for the real data; 1.02, 1.01–1.02, $p < 0.001$ for SD). The Cox models showed broadly similar HR estimates and preserved the overall pattern of associations between the real and synthetic cohorts. The Kaplan–Meier estimates of time to PIRA showed similar temporal patterns in the real and synthetic cohorts (Figure 4). The EIT approach in reducing the risk of PIRA compared to ESC was not significant in the real cohort (HR = 0.88; 95% CI = 0.71–1.08, $p = 0.22$) and significant in the SD (0.86, 0.76–0.98, $p = 0.02$). A longer time to the initiation of the first DMT was also associated with a higher risk of PIRA in the SD (HR = 1.10, 95% CI = 1.04–1.17, $p < 0.001$). The effect of MRI variables on PIRA risk retrieved significant findings. Patients with ⩾9 T2 lesions at baseline had a significantly higher risk of PIRA (HR = 1.88, 95% CI = 1.20–2.96, $p = 0.01$), and conversely, the presence of GD-enhancing T1 lesion (HR = 0.82, 95% CI = 0.73–0.91, $p < 0.001$) was a protective factor of PIRA events. Conversely, MRI parameters were not significant in the real cohort group. Forest plots reporting HR in real and synthetic datasets are shown in Figure 5. Exact HR with CIs is provided for both real and synthetic datasets in Supplementary Material (SF7).

Discussion

Our findings support the validity of AI-generated SD to replicate clinical and MRI real data as those collected in the RISM with high-fidelity (CSF = 97%) and privacy preservation (NNDR = 0.61) values. Recent literature highlights the increasing role of SD in various fields of biomedical research, particularly in enhancing study reproducibility and overcoming data-sharing restrictions.[8,12,13,16–18] Research on neurological diseases has also been greatly impacted by recent developments in generative AI modeling, which have opened up novel possibilities for data augmentation, privacy protection, and personalized medicine. Variational autoencoders and GANs have proven to be effective in producing high-fidelity synthetic neuroimaging data, which has been successfully applied in neurodegenerative diseases.[11,19,20] Synthetic data can also facilitate external validation of predictive models, providing robust insights into treatment effects.[21] For research purposes, synthetic datasets that replicate a portion of the Public Health England Cancer Registry and the Netherlands Cancer Registry are now accessible in Europe[22,23] and the United States.[24]

The treatment-effect estimates were concordant in magnitude between the real and synthetic cohorts, with HRs of 0.88 and 0.86, respectively. Although statistical significance was reached only in the synthetic cohort, the similar effect estimates indicate preservation of the direction and magnitude of the association. In the synthetic dataset, the Cox model showed a significantly reduced risk of PIRA with the EIT strategy. Several previous observational RWE studies from large cohorts reported significant reductions

Table 1. Baseline clinical and demographic characteristics of EIT and ESC cohorts.

Variable	ESC ($n = 3964$)	EIT ($n = 914$)	p-value
Age at first DMT start	29.70 (23.60–38.05)	30.85 (24.20–40.50)	0.0018
0.0–20.0	518 (13.07)	103 (11.27)	0.0012
20.1–40.0	2637 (66.52)	575 (62.91)	
>40.0	809 (20.41)	236 (25.82)	
Female sex	2633 (66.42)	608 (66.52)	0.9549
Time from the start of the first ME DMT to escalation (months)	51.71 (29.05–80.46)	NA	NA
Time from first visit to first DMT start (months)	9.08 (4.41–17.35)	8.06 (3.95–17.50)	0.1579
Time spent on DMT (years)	7.80 (5.40–10.60)	5.60 (2.30–7.90)	<0.0001
Follow–up time (years)	9.45 (7.28–11.84)	7.55 (6.14–9.63)	<0.0001
Number of EDSS evaluations (visits density)	16.00 (10.00–24.00)	15.00 (9.00–22.00)	0.0007
Patients with relapse before first DMT start (%)	1896 (47.83)	506 (55.36)	<0.0001
Multifocal onset			
No	3312 (83.55)	733 (80.20)	0.0479
Yes	492 (12.41)	134 (14.66)	
NA	160 (4.04)	47 (5.14)	
EDSS score at the first DMT start	1.50 (1.00–2.50)	2.50 (1.50–3.50)	<0.0001
Patients with T1 GD + lesion (at least 1) at baseline MRI, *n* (%)			
No	2121 (53.51)	513 (56.13)	0.3291
Yes	724 (18.26)	153 (16.74)	
NA	1119 (28.23)	248 (27.13)	
T2 lesion number at baseline MRI, *n* of patients (%)			
0	57 (1.44)	23 (2.52)	<0.0001
1–2	84 (2.12)	17 (1.86)	
3–8	612 (15.44)	83 (9.08)	
⩾9	640 (16.15)	196 (21.44)	
NA	2571 (64.86)	595 (65.10)	
PIRA	1747 (44.07)	354 (38.73)	0.0033
Time to PIRA (years)	2.70 (1.34–4.70)	3.09 (1.63–4.85)	0.06

DMT: disease-modifying treatment; EDSS: Expanded Disability Status Scale; ESC: escalation strategy; EIT: early intensive treatment; GD: gadolinium; ME DMT: moderate efficacy DMT; MS: multiple sclerosis; MRI: magnetic resonance imaging; PIRA: progression independent of relapse activity.

in long-term disability with EIT compared to ESC paradigms.[25,26] Another study from the RISM research group assessed long-term disability trajectories and demonstrated that the mean yearly delta-EDSS values were significantly higher following ESC approaches, emphasizing the therapeutic effects of early initiation of HE-DMTs.[27] More recently, a study from the RISM network confirmed the long-term superiority of EIT strategies, showing significantly lower risks of CDA, PIRA, and RAW compared with escalation strategies.[28] The findings of the present study are consistent with these previous observations and extend them with real and synthetic cohorts showing a similar direction and magnitude of the treatment association.

Most importantly, this study highlights the potential of SD to complement real-world data. Indeed, AI-generated SD allowed the evaluation of MRI-derived risk factors for disability progression within the synthetic cohort. Previous work using register data has often not allowed the use of MRI data, because it is frequently missed due to the lack of a systematic acquisition of these data.[29] The synthetic cohort was derived exclusively from the 1666 MRI-complete patients and therefore reflects this subgroup rather than patients with missing MRI data. MRI-derived associations showed broadly similar patterns across the real and synthetic analyses, although some differences in effect estimates and statistical significance were observed. In the synthetic cohort, high lesion load was associated with increased PIRA risk,[30,31] while the presence of GD-enhancing T1 lesions was associated with a lower risk of PIRA. Therefore, our findings suggest that SD may constitute a valuable tool for investigating

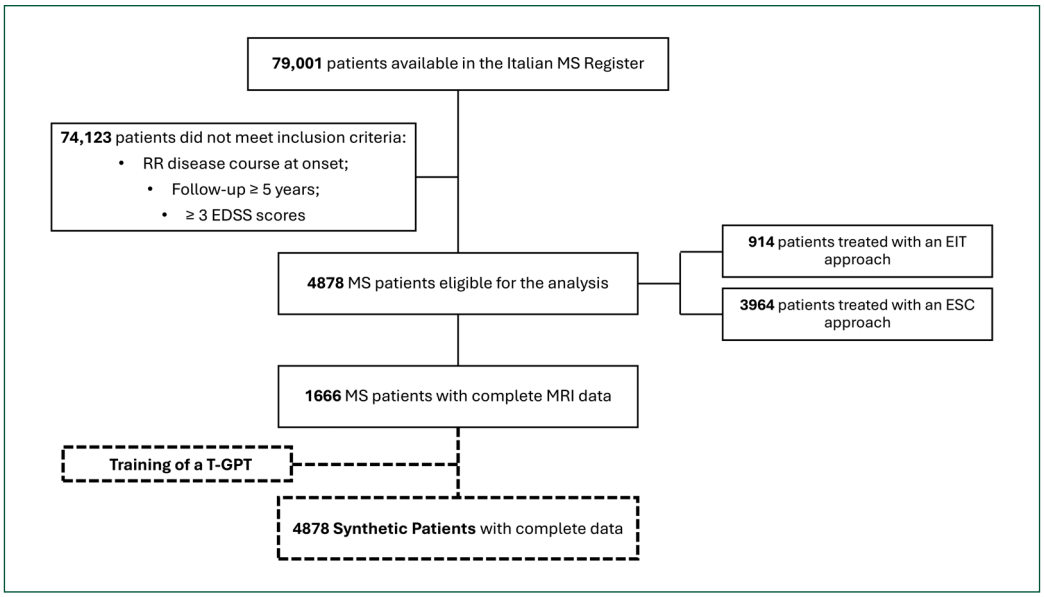

Figure 1. Flowchart of patient selection procedure.
EIT: early intensive treatment; ESC: escalation treatment strategy; EDSS: Expanded Disability Status Scale; MRI: magnetic resonance imaging; RRMS: relapsing-remitting multiple sclerosis; T-GPT: Tabular Generative Pretrained Transformer.

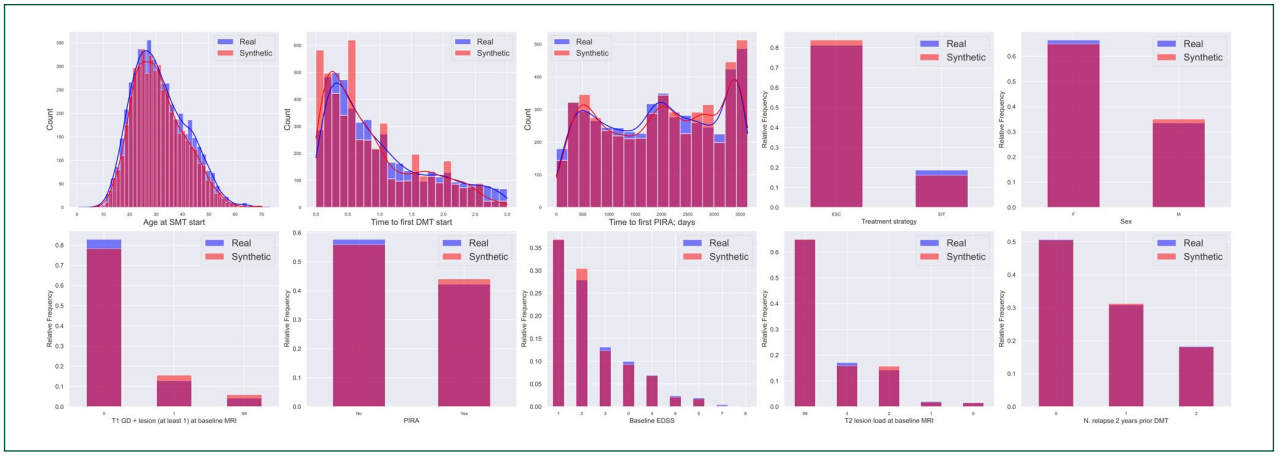

Figure 2. Comparison of distributions for key clinical variables between real and synthetic datasets.
Distributions of key demographic and clinical variables are shown for both real and synthetic data. Variables include age at start of disease-modifying therapy (DMT), sex, baseline Expanded Disability Status Scale (EDSS), treatment strategy (ESC vs EIT), presence of T1 gadolinium-enhancing lesions at baseline MRI, T2 lesion load at baseline, number of relapses in the 2 years prior to DMT initiation, time to first DMT initiation, and time to first progression independent of relapse activity (PIRA) event. Synthetic data closely mirror the real data across all distributions, supporting the fidelity of the generative model.

MRI-based prognostic associations, offering a novel framework to enhance the robustness of RWE in MS. The ability of SD to mimic the effects of real-world therapy on PIRA has become particularly relevant as PIRA is increasingly acknowledged as a significant contributor to disability accrual in MS.[1–6,32,33] In addition, as shown in the Supplementary, the MRI-complete cohort differed from excluded patients in several clinical and follow-up characteristics, which should be considered when interpreting the representativeness of the cohort.

One of the most promising applications of generative modeling in neurological disease research suggested by this study is its potential to contribute to precision medicine by creating digital twins of patients. In MS research, digital twin approaches based on SD could enable real-time decision-making for clinicians, helping to tailor treatment strategies to individual patient profiles while maintaining data privacy.[11]

While the advantages of generative AI in medical research are substantial, challenges remain, and limitations

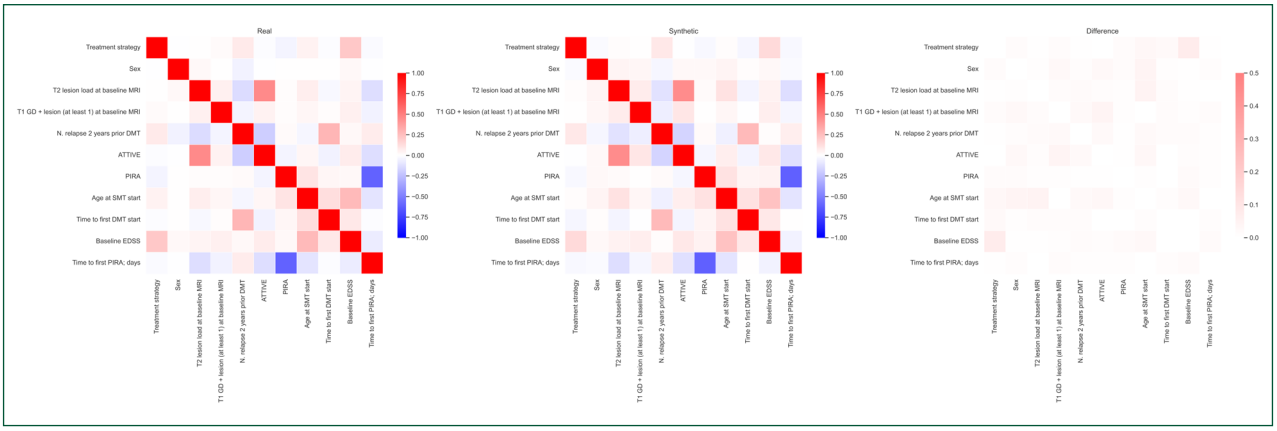

Figure 3. Correlation comparison between the real and synthetic data.

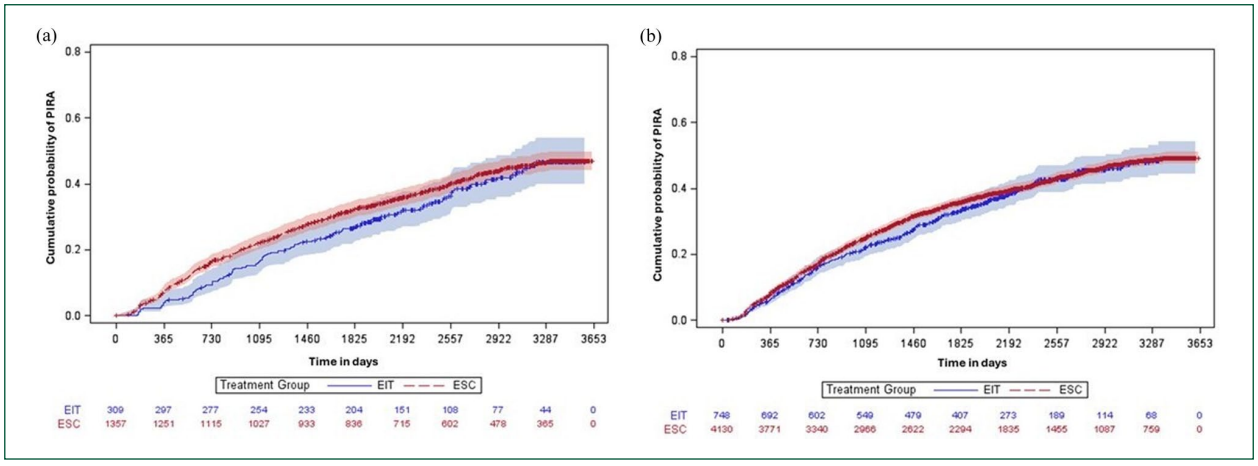

Figure 4. Kaplan–Meier estimates of time to PIRA in the real-world (a) and synthetic (b) cohorts, stratified by treatment strategy.

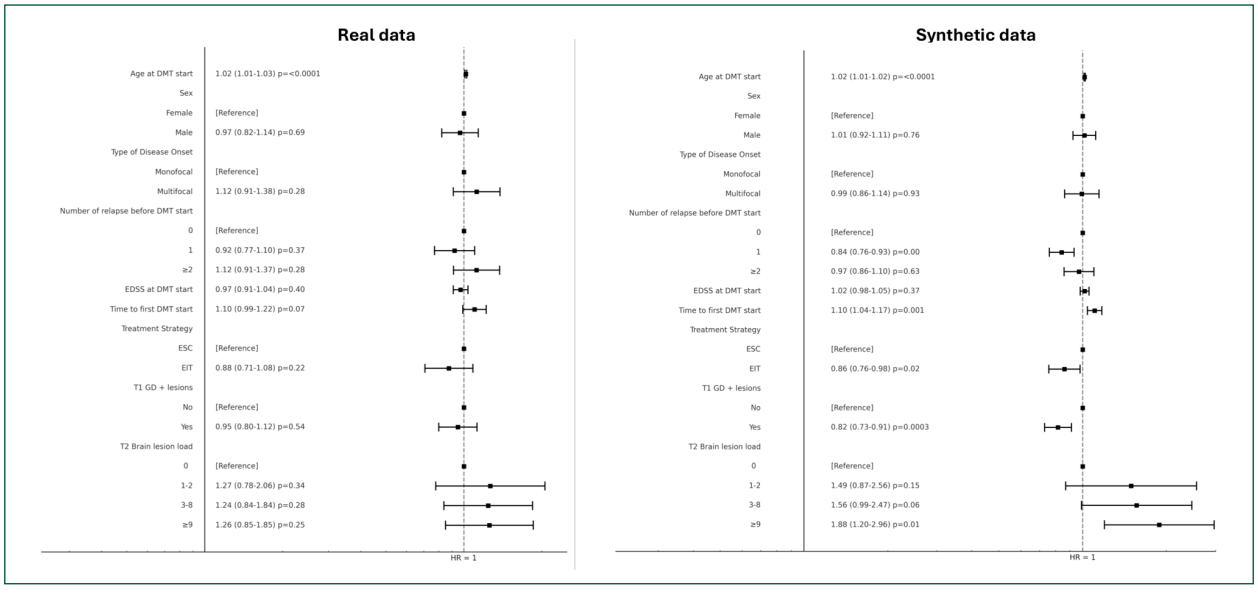

Figure 5. Forest plots for Cox proportional hazards model applied to the real and synthetic datasets to investigate risk factors of PIRA.

must be acknowledged. Subtle biases in the original dataset may be perpetuated, and the quality of the real-world training data determines how well SD is generated. The selection of the study cohort reflects the real-world nature of the RISM, where heterogeneous data availability across participating centers[34] requires data selection and curation before developing and validating generative AI models. Accordingly, the present findings should be interpreted within the boundaries of this predefined analytical use case, as the synthetic cohort was designed to preserve relationships relevant to the same clinical question of this study rather than to serve as a general-purpose representation of the registry. Model evaluation, data fidelity, and privacy concerns must be rigorously addressed to ensure clinical relevance and adherence to high ethical standards. A potential limitation of synthetic data is that the greater statistical precision observed in larger synthetic cohorts may partly reflect the increased effective sample size. Furthermore, a limitation of the present study is that privacy was assessed using distance-based metrics only; future work will incorporate adversarial privacy evaluations, including membership inference attacks, to provide a more comprehensive assessment. Differences in statistical significance between the real and synthetic cohorts should be interpreted considering all the limitations of source sample and synthesis process. In addition, further studies are warranted to assess the applicability of AI-generated synthetic data across different MS populations, treatment paradigms, and statistical methodologies.

AI-generated SD in this study reliably reproduced treatment-effect estimates for PIRA obtained from a large real-world MS registry and preserved several MRI-derived prognostic associations from the source cohort. Synthetic cohorts may complement traditional RWE, enhance reproducibility, and expand opportunities for privacy-preserving research in MS. Future research should explore the broader application of SD in MS research, including its potential use in multicenter studies and regulatory decision-making and refining SD methodologies to maximize their impact.

Acknowledgements

The authors thank the research assistants of the RISM register.

ORCID iDs

Pietro Iaffaldano https://orcid.org/0000-0003-2308-1731
Massimiliano Copetti https://orcid.org/0000-0002-7960-5947
Tommaso Guerra https://orcid.org/0000-0002-5248-5981
Maria A. Rocca https://orcid.org/0000-0003-2358-4320
Diana Ferraro https://orcid.org/0000-0003-4818-3806
Emilio Portaccio https://orcid.org/0000-0002-9662-1762
Damiano Paolicelli https://orcid.org/0000-0002-8645-1763
Maria P. Amato https://orcid.org/0000-0003-3325-3760
Massimo Filippi https://orcid.org/0000-0002-5485-0479
Maria Trojano https://orcid.org/0000-0002-6329-8946

Ethical considerations

RISM was approved by the ethical committee at the "Azienda Ospedaliero – Universitaria – Policlinico of Bari" (Study REGISTRO SM001—approved on 8 July 2016) and by local ethics committees in all participating centers.

Consent to participate

Patients signed an informed consent to collect and use their clinical data for research purposes.

Consent for publication

Not applicable (no identifiable individual data are presented).

Author contributions

P.I., S.D.A., G.L., M.C., M.T., and T.G. contributed to the conception and design of the study. P.I., S.D.A., G.L., M.C., M.T., T.G., V.S., M.D., F.E.C., and E.S. contributed to the acquisition and analysis of data. P.I., S.D.A., G.L., M.C., M.T., T.G., V.S., M.D., F.E.C., and E.S. contributed to drafting the text or preparing the figures.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Declaration of conflicting interests

The authors declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: They note, however, that patients included in the analysis received various disease-modifying therapies, and several authors have received honoraria, travel grants, consulting fees, research support, or participated in advisory boards sponsored by pharmaceutical companies, including Actelion, Allergan, Almirall, Alexion, Amgen, Bayer Schering, Biogen, Bristol Myers Squibb, Celgene, Excemed, Genzyme, Forward Pharma, Ipsen, MedDay, Merck, Mylan, Novartis, Sanofi, Roche, Teva, and their affiliates.

Data availability statement

Anonymized data that support the findings of this study will be shared on reasonable request from a qualified investigator.

Supplemental material

Supplemental material for this article is available online.

References

1. Filippi M, Amato MP, Avolio C, et al. Towards a biological view of multiple sclerosis from early subtle to clinical progression: an expert opinion. *J Neurol* 2025; 272(2): 179. https://doi.org/10.1007/s00415-025-12917-4
2. Jakimovski D, Bittner S, Zivadinov R, et al. Multiple sclerosis. *Lancet* 2024; 403(10422): 183–202. https://doi.org/10.1016/S0140-6736(23)01473-3
3. Iaffaldano P, Portaccio E, Lucisano G, et al. Multiple sclerosis progression and relapse activity in children

[published correction appears in JAMA Neurol. 2024 Jan 1;81(1):88. DOI: 10.1001/jamaneurol.2023.5461]. *JAMA Neurol* 2024; 81(1): 50–58. https://doi.org/10.1001/jamaneurol.2023.4455
4. Portaccio E, Betti M, De Meo E, et al. Progression independent of relapse activity in relapsing multiple sclerosis: impact and relationship with secondary progression [published correction appears in J Neurol. 2024 Oct;271(10):7066-7068. DOI: 10.1007/s00415-024-12627-3]. *J Neurol* 2024; 271(8): 5074–5082. https://doi.org/10.1007/s00415-024-12448-4
5. Calabrese M, Preziosa P, Scalfari A, et al. Determinants and biomarkers of progression independent of relapses in multiple sclerosis. *Ann Neurol* 2024; 96(1): 1–20. https://doi.org/10.1002/ana.26913
6. Guerra T and Iaffaldano P. A window into new insights on progression independent of relapse activity in multiple sclerosis: role of therapies and current perspective. *Int J Mol Sci* 2025; 26(3): 884. https://doi.org/10.3390/ijms26030884
7. Trojano M, Kalincik T, Iaffaldano P, et al. Interrogating large multiple sclerosis registries and databases: what information can be gained? *Curr Opin Neurol* 2022; 35(3): 271–277. https://doi.org/10.1097/WCO.0000000000001057
8. D'Amico S, Dall'Olio D, Sala C, et al. Synthetic data generation by artificial intelligence to accelerate research and precision medicine in hematology. *JCO Clin Cancer Inform* 2023; 7: e2300021. https://doi.org/10.1200/CCI.23.00021
9. Chen RJ, Lu MY, Chen TY, et al. Synthetic data in machine learning for medicine and healthcare. *Nat Biomed Eng* 2021; 5(6): 493–497. https://doi.org/10.1038/s41551-021-00751-8
10. Rajotte JF, Bergen R, Buckeridge DL, et al. Synthetic data as an enabler for machine learning applications in medicine. *iScience* 2022; 25(11): 105331. https://doi.org/10.1016/j.isci.2022.105331
11. Seiler M and Ritter K. Pioneering new paths: the role of generative modelling in neurological disease research. *Pflugers Arch* 2025; 477(4): 571–589. https://doi.org/10.1007/s00424-024-03016-w
12. Jacobs F, D'Amico S, Benvenuti C, et al. Opportunities and challenges of synthetic data generation in oncology. *JCO Clin Cancer Inform* 2023; 7: e2300045. https://doi.org/10.1200/CCI.23.00045
13. D'Amico S, Dall'Olio L, Rollo C, et al. MOSAIC: an artificial intelligence-based framework for multimodal analysis, classification, and personalized prognostic assessment in rare cancers. *JCO Clin Cancer Inform* 2024; 8: e2400008. https://doi.org/10.1200/CCI.24.00008
14. Zazzetti E, D'Amico S, Jacobs F, et al. Longitudinal synthetic data generation by artificial intelligence to accelerate clinical and translational research in breast cancer. *JCO Clin Cancer Inform* 2025; 9: e2500033. https://doi.org/10.1200/CCI-25-00033
15. Zhao Z, Kunar A, Birke R, et al. Chen proceedings of the 13th Asian conference on machine learning. *PMLR* 2021; 157: 97–112. https://doi.org/10.48550/arXiv.2102.08369
16. Krauland MG, Frankeny RJ, Lewis J, et al. Development of a synthetic population model for assessing excess risk for cardiovascular disease death. *JAMA Netw Open* 2020; 3(9): e2015047. https://doi.org/10.1001/jamanetworkopen.2020.15047
17. Suresh K, Cohen MS, Hartnick CJ, et al. Making use of artificial intelligence-generated synthetic tympanic membrane images. *JAMA Otolaryngol Head Neck Surg* 2023; 149(6): 555–556. https://doi.org/10.1001/jamaoto.2023.0218
18. El Emam K. Status of synthetic data generation for structured health data. *JCO Clin Cancer Inform* 2023; 7: e2300071. https://doi.org/10.1200/CCI.23.00071
19. Wang R, Bashyam V, Yang Z, et al. Applications of generative adversarial networks in neuroimaging and clinical neuroscience. *Neuroimage* 2023; 269: 119898. https://doi.org/10.1016/j.neuroimage.2023.119898
20. Ghebrehiwet I, Zaki N, Damseh R, et al. Revolutionizing personalized medicine with generative AI: a systematic review. *Artif Intell Rev* 2024; 57: 128. https://doi.org/10.1007/s10462-024-10768-5
21. Sauta E, Robin M, Bersanelli M, et al. Real-world validation of molecular international prognostic scoring system for myelodysplastic syndromes. *J Clin Oncol* 2023; 41(15): 2827–2842. https://doi.org/10.1200/JCO.22.01784
22. Netherlands Comprehensive Cancer Organisation (IKNL): synthetic dataset Netherlands Cancer Registry (NCR), https://iknl.nl/en/ncr
23. Horvat P, Gray CM, Lambova A, et al. Comparing findings from a Friends of Cancer Research exploratory analysis of real-world end points with the Cancer Analysis System in England. *JCO Clin Cancer Inform* 2021; 5: 1155–1168.
24. Chen J, Chun D, Patel M, et al. The validity of synthetic clinical data: a validation study of a leading synthetic data generator (Synthea) using clinical quality measures. *BMC Med Inform Decis Mak* 2019; 19(1): 44. https://doi.org/10.1186/s12911-019-0793-0
25. He A, Merkel B, Brown JWL, et al. Timing of high-efficacy therapy for multiple sclerosis: a retrospective observational cohort study. *Lancet Neurol* 2020; 19(4): 307–316.
26. Spelman T, Magyari M, Piehl F, et al. Treatment escalation vs immediate initiation of highly effective treatment for patients with relapsing-remitting multiple sclerosis: data from 2 different national strategies. *JAMA Neurol* 2021; 78(10): 1197–1204.
27. Iaffaldano P, Lucisano G, Caputo F, et al. Long-term disability trajectories in relapsing multiple sclerosis patients treated with early intensive or escalation treatment strategies. *Ther Adv Neurol Disord* 2021; 14: 17562864211019574.
28. Iaffaldano P, Lucisano G, Guerra T, et al. Early intensive versus escalation approach: ten-year impact on disability in relapsing multiple sclerosis. *Ann Clin Transl Neurol* 2025; 12(10): 2012–2019. https://doi.org/10.1002/acn3.70131
29. Iaffaldano P, Lucisano G, Guerra T, et al. A comparison of natalizumab and ocrelizumab on disease progression in multiple sclerosis. *Ann Clin Transl Neurol* 2024; 11(8): 2008–2015. https://doi.org/10.1002/acn3.52118
30. Cagol A, Benkert P, Melie-Garcia L, et al. Association of spinal cord atrophy and brain paramagnetic rim lesions with progression independent of relapse activity in people with MS. *Neurology* 2024; 102(1): e207768.
31. Müller J, Cagol A, Lorscheider J, et al. Harmonizing definitions for progression independent of relapse activity in multiple sclerosis: a systematic review. *JAMA Neurol* 2023; 80(11): 1232–1245.

32. Lublin FD, Häring DA, Ganjgahi H, et al. How patients with multiple sclerosis acquire disability. *Brain* 2022; 145(9): 3147–3161. https://doi.org/10.1093/brain/awac016
33. Simone M, Lucisano G, Guerra T, et al. Disability trajectories by progression independent of relapse activity status differ in pediatric, adult and late-onset multiple sclerosis. *J Neurol* 2024; 271(10): 6782–6790. https://doi.org/10.1007/s00415-024-12638-0
34. Mosconi P, Guerra T, Paletta P, et al. Data monitoring roadmap. The experience of the Italian multiple sclerosis and related disorders register. *Neurol Sci* 2023; 44(11): 4001–4011. https://doi.org/10.1007/s10072-023-06876-9