


# Long-MS-Diff: Towards Generating Anatomically Plausible Lesion Progression in MS Imaging using Diffusion Models

**Prateek Mathur**<sup>1,2</sup> 

PRATEEK.MATHUR@UCD.IE

<sup>1</sup> *School of Computer Science, University College Dublin, Dublin, Ireland*

<sup>2</sup> *Insight Research Ireland Centre for Data Analytics, UCD, Dublin, Ireland*

**Brendan S. Kelly**<sup>2,3</sup> 

BREANDAN.KELLY@INSIGHT-CENTRE.ORG

<sup>3</sup> *School of Medicine, University College Dublin, Dublin, Ireland*

**Ronan P. Killeen**<sup>3,4</sup> 

RP.KILLEEN@ST-VINCENTS.IE

<sup>4</sup> *St. Vincent's University Hospital, Dublin, Ireland*

**Aonghus Lawlor**<sup>1,2</sup> 

AONGHUS.LAWLOR@UCD.IE

**Editors:** Accepted for publication at MIDL 2025

## Abstract

Monitoring lesion progression in Multiple Sclerosis (MS) is vital for assessing disease activity and guiding treatment decisions. However, the limited availability of annotated longitudinal MRI data presents a challenge for developing robust deep learning models. We introduce Long-MS-Diff, a conditional diffusion-based framework for generating anatomically plausible follow-up brain scans in MS. The model is conditioned on baseline images, lesion change masks, and scalar lesion-level features. To address the extreme sparsity of new lesions, we incorporate an auxiliary segmentation task and propose an adaptive weighted loss to balance anatomical reconstruction with lesion-specific fidelity. A radiologist assessment of synthetic scans confirms high image quality, anatomical plausibility, and lesion adherence. In downstream segmentation experiments, moderate augmentation with Long-MS-Diff improves performance, outperforming models trained on real data alone. These results highlight the value of controlled synthetic data in modelling disease progression and demonstrate the utility of diffusion-based generation in data-scarce clinical settings.

**Keywords:** Diffusion Models, Synthetic Medical Imaging, Multiple Sclerosis

## 1. Introduction

Multiple Sclerosis (MS) is a chronic neurodegenerative disease marked by the development of white matter lesions, visible on brain MRI (McNamara et al., 2017b,a). Accurately modelling lesion progression is essential for tracking disease activity and informing treatment decisions. However, the scarcity of annotated longitudinal data limits the development of models that capture the spatial and temporal dynamics of MS lesion evolution.

Generative models offer a promising direction for addressing data scarcity by synthesising realistic follow-up scans for data augmentation. In particular, diffusion models have shown strong performance in generating high-fidelity medical images (Kazerouni et al., 2023). Yet, they often prioritise image realism over explicit guidance (Preechakul et al., 2021), making it challenging to model subtle lesion progression. Structural causal models provide controllability but rely on balanced datasets of healthy and diseased cases, which are rarely available (Ribeiro et al., 2023).

In this paper, we propose Long-MS-Diff, a conditional diffusion framework for generating anatomically realistic follow-up scans in MS. The model incorporates scalar lesion-level features, an auxiliary segmentation task, and an adaptive weighted loss to achieve both anatomical consistency and lesion-specific fidelity. Unlike prior diffusion approaches, Long-MS-Diff is specifically designed to model sparse lesion progression through lesion-aware conditioning and dropout. We validate the framework through expert evaluation and demonstrate improved segmentation performance with synthetic follow-up data.

## 2. Methods

We propose Long-MS-Diff, a conditional diffusion framework built on improved denoising diffusion models (Nichol and Dhariwal, 2021), for generating anatomically plausible follow-up MRI scans in MS. The model is conditioned on a baseline image, a binary lesion change mask, and scalar lesion-level features (lesion load, lesion count, and a binary change flag), which are spatially broadcast and concatenated before being passed to a UNet-based diffusion backbone (Figure 1).

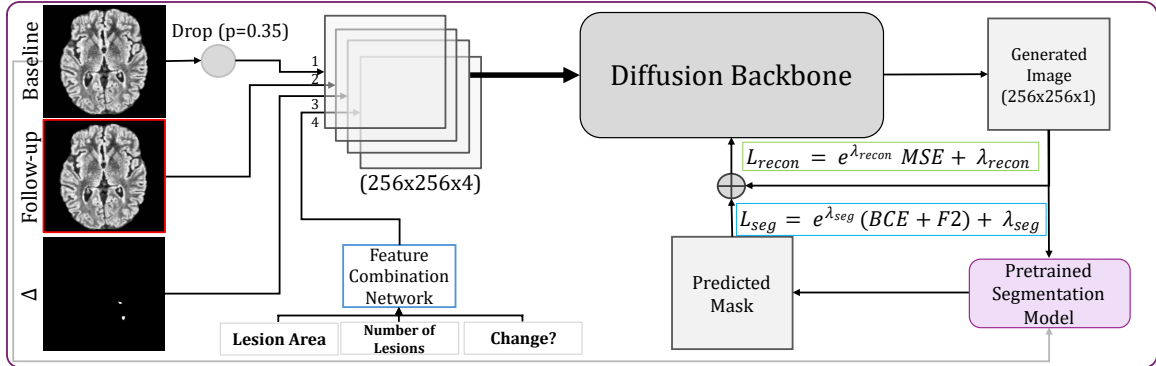


Figure 1: Long-MS-Diff model for generating anatomically plausible MS follow-up scans.

Recent advances in generative modelling have focused on conditional generation, where auxiliary guidance enforces semantic structure in the output (Konz et al., 2024; Saragih et al., 2024). In MS, a key challenge is the extremely low prevalence of new lesions, typically under 0.0013% of the image, which leads models to favour image realism over lesion fidelity. To address this, we introduce an auxiliary segmentation task, where a pretrained model segments new lesions in the generated scan, and the predicted mask is compared against ground truth using a segmentation loss ( $L_{seg}$ )

Diffusion models optimise for pixel-wise fidelity via a reconstruction loss  $L_{recon}$ , which can overpower lesion-specific supervision. We mitigate this with two strategies. First, we introduce an *adaptive weighted loss* that dynamically balances  $L_{recon}$  and  $L_{seg}$ , ensuring both anatomical and lesion-level fidelity. Second, we implement a *conditioning dropout mechanism* by randomly replacing the baseline image with noise during training. This encourages the model to rely more on the lesion change mask and scalar features, promoting stronger lesion-aware generation.

### 3. Experiments and Results

We evaluate Long-MS-Diff on an internal longitudinal MS dataset comprising 170 patients, split into 110/30/30 for train/validation/test. Each patient has at least one baseline and one follow-up FLAIR MRI scan ( $256 \times 256$ ), along with expert-annotated lesion change masks. The diffusion model is trained for approximately 18 hours<sup>1</sup>. A bi-temporal UNet serves as the pretrained segmentation model used for auxiliary supervision.

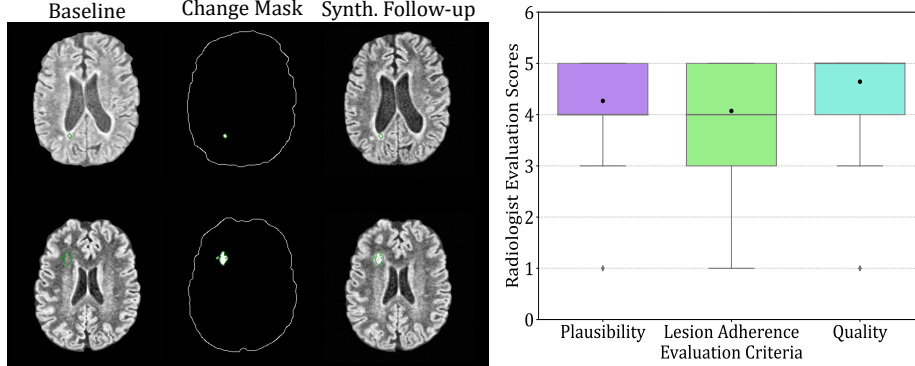


Figure 2: Examples of Generated Synthetic lesion progression (left). Radiologist evaluation of image quality, anatomical plausibility and lesion guidance adherence (right).

To qualitatively assess the realism of the generated scans, we synthesised 15,000 follow-up images using randomly sampled change masks. A random subset was reviewed by an expert radiologist, who rated each image on three criteria—overall image quality, anatomical plausibility, and lesion guidance adherence using a 5-point Likert scale. The images received high average scores—4.6 for quality, 4.2 for plausibility, and 4.0 for lesion adherence—indicating that the generated data is realistic and clinically meaningful.

Table 1: Performance of ViT (ResNet-50) trained with varying real-to-synthetic data.

Real:Synthetic	Dice	Precision	Recall	Correct	False Pos.	False Neg.
1:0	0.245	0.241	0.360	47	305	51
1:1	0.370	<b>0.383</b>	0.495	52	<b>273</b>	46
1:2	<b>0.382</b>	0.367	<b>0.593</b>	<b>66</b>	460	<b>32</b>
1:3	0.311	0.277	0.520	57	471	41

Additionally, we conduct a new lesion detection task using a Vision Transformer trained with different real-to-synthetic data ratios. Evaluation is carried out on real test sets following an established protocol (Kelly et al., 2024). Significant performance gains are observed (Table 1) with synthetic data augmentation peaking at a 1:2 ratio compared to no augmentation (1:0). A decline at 1:3 suggests overfitting to synthetic data and reduced generalisability. Despite modest performance, these results highlight the utility of Long-MS-Diff in generating clinically meaningful data for robust MS progression modelling.

### Acknowledgments

This research was supported by Science Foundation Ireland (SFI) under Grant Number SFI/12/RC/2289\_P2 and the Irish Centre for High-End Computing (ICHEC).

1. Model was trained on 4×40GB NVIDIA A100 GPUs

## References

- Amirhossein Kazerooni, Ehsan Khodapanah Aghdam, Moein Heidari, Reza Azad, Mohsen Fayyaz, Ilker Hacıhaliloğlu, and Dorit Merhof. Diffusion models in medical imaging: A comprehensive survey. *Medical Image Analysis*, 88:102846, 2023. ISSN 1361-8415. doi: 10.1016/j.media.2023.102846.
- Brendan S Kelly, Prateek Mathur, Ronan P Killeen, and Aonghus Lawlor. A siamese u-transformer for change detection on MRI brain for multiple sclerosis, a model development and external validation study. *medRxiv*, page 2024.04.05.24305386, 2024. doi: 10.1101/2024.04.05.24305386.
- Nicholas Konz, Yuwen Chen, Haoyu Dong, and Maciej A Mazurowski. Anatomically-controllable medical image generation with segmentation-guided diffusion models. *arXiv*, 2024. doi: 10.48550/arxiv.2402.05210.
- C. McNamara, G. Sugrue, B. Murray, and P.J. MacMahon. Current and emerging therapies in multiple sclerosis: Implications for the radiologist, part 2—surveillance for treatment complications and disease progression. *American Journal of Neuroradiology*, 38(9):1672–1680, 2017a. ISSN 0195-6108. doi: 10.3174/ajnr.a5148.
- C. McNamara, G. Sugrue, B. Murray, and P.J. MacMahon. Current and emerging therapies in multiple sclerosis: Implications for the radiologist, part 1—mechanisms, efficacy, and safety. *American Journal of Neuroradiology*, 38(9):1664–1671, 2017b. ISSN 0195-6108. doi: 10.3174/ajnr.a5147.
- Alex Nichol and Prafulla Dhariwal. Improved denoising diffusion probabilistic models. *arXiv*, 2021.
- Konpat Preechakul, Nattanat Chatthee, Suttisak Wizadwongsa, and Supasorn Suwajanakorn. Diffusion autoencoders: Toward a meaningful and decodable representation. *arXiv*, 2021. doi: 10.48550/arxiv.2111.15640.
- Fabio De Sousa Ribeiro, Tian Xia, Miguel Monteiro, Nick Pawlowski, and Ben Glocker. High fidelity image counterfactuals with probabilistic causal models. *arXiv*, 2023. doi: 10.48550/arxiv.2306.15764.
- Daniel G. Saragih, Atsuhiko Hibi, and Pascal N. Tyrrell. Using diffusion models to generate synthetic labeled data for medical image segmentation. *International Journal of Computer Assisted Radiology and Surgery*, 19(8):1615–1625, 2024. doi: 10.1007/s11548-024-03213-z.