RESCIENCE C

Replication / Neuroscience

[Re] Spread of alpha-synuclein pathology through the brain connectome is modulated by selective vulnerability and predicted by network analysis

Thomas E. Paul[1,2, ID] and Mathieu Bourdenx[1,2, ID]

[1]Univ. de Bordeaux, Institut des Maladies Neurodégénératives, UMR 5293, Bordeaux, France – [2]CNRS, Institut des Maladies Neurodégénératives, UMR 5293, Bordeaux, France

1 Introduction

Parkinson's disease (PD) is a common motor neurodegenerative disorder, characterized in part by the progressive loss of pigmented dopaminergic neurons from the *Substantia Nigra pars compacta*. The main neuropathological hallmark of PD is the progressive accumulation of α-synuclein in fibrillar aggregates named Lewy Bodies (in the soma) or Lewy Neurites (in neurites) [1]. Interestingly, Lewy pathology follows a predictable pattern of progression in the brain suggesting that α-synuclein can propagate from neurons to neurons [2]. Such observation, common to many neurodegenerative diseases [3], prompted initiatives to model and predict pathology progression in the brain of patients with neurodegenerative disorders.

In 2012, Raj and colleagues developed a Network Diffusion Model (NDM) to predict atrophy progression in dementia [4]. For this purpose, they compared Magnetic Resonance Imaging acquisitions from healthy young subjects, healthy old subjects, and old patients suffering from either Frontotemporal Dementia or Alzheimer's Disease. Their study suggested that patterns of atrophy within the dementia spectrum could be explained by transmission of the disease along neuronal pathways. Using their model, they demonstrated a path for predicting future atrophy in individuals starting from baseline. The NDM model was later applied to other diseases using various kinds of data. In 2017, Mezias and Raj showed in a mouse model that the spread of amyloid-β pathology, involded in Alzheimer's disease, was driven by spatial proximity but not network connectivity [5]. At the opposite, progression of tau pathology, also involved in Alzheimer's disease and other tauopathies, was better recapitulated by connectivity [6]. In 2019, Henderson and colleagues reported that progression of α-synuclein pathology, similary to tau pathology, is largely predicted by anatomical connectivity between brain structures [7].

We here replicated the core of the NDM applied to α-synuclein spread in mice as implemented by Henderson and colleagues [7]. R scripts, connectivity matrix and experimental datasets were made fully accessible by the authors on GitHub allowing an easy process. We took inspiration from the original code in the way to build the connectivity matrix. It was then used as a tool to generate results and compare with the Python imple-

mentation. In order to ease comparison between the two codes, we kept some function and variable names similar. Overall, we were able to reproduce qualitatively and quantitatively the main results of the publication. The aim of the present replication is not to fully replicate the entire paper but to focus on the core model and immediate results (shown in Figures 4 and 5 in the original paper).

2 Background

2.1 Graph Theory and brain connectivity

Henderson et al. combined quantitative pathology mapping in the mouse brain with network modeling to understand the spatiotemporal pattern of spread of α-synuclein pathology [7]. Graph theory is a branch of mathematics applicable to Neuroscience. The brain is a complex system that can be modeled as a network or graph. When applied properly, Graph theory can offer important insights into different aspects of brain networks such as architecture, evolution, development, or clinical disorder [8].

Networks are defined as a collection of elements (or nodes) and their pairwise links (or edges) that can be summarized in the form of a connectivity matrix (named adjacency matrix). Depending on the type of graph, edges can have binary values (present: 1 - absent: 0) or actual weights reflecting connection strength. A graph is called undirected when an edge e reciprocally connects nodes V_a and V_b. Conversely, a graph is named directed when the edge e is projecting from V_a to V_b but not from V_b to V_a.

Thanks to the initiative from the Allen Institute for Brain Science's Mouse Connectivity Atlas (MCA), the mouse brain mesoscale connectome is now freely accessible for the neuroscience community [9]. In a matrix, connectivity is represented as "incoming" along rows and "outgoing" along columns. From the adjacency matrix, the in-degree and out-degree distributions can be computed as respectively the sum of all entries in the corresponding row and the sum of all entries in the corresponding column and are represented in diagonal matrices.

The Laplacian graph is a matrix used to explore the properties of a network. It is computed using the adjacency matrix and either the in-degree or out-degree graph depending on the applications.

2.2 Model description

The model used by Henderson et al. requires the mouse brain connectome (discussed in the previous section) and whole brain α-synuclein pathology quantification.

The experimental model of α-synuclein pathology propagation used in the study is the now classical α-synuclein pre-formed fibrils (PFFs - 5μg) unilateral injection in the dorsal striatum of non-transgenic mice (NTG) [10, 7]. Following inoculation, NTG mice were terminated at 3, 6, or 9 months post-injection (MPI). Brain pathology was assessed using traditional immunohistochemical methods to detect α-synuclein phosphorylated on Serine129 (p-α-syn). Quantification (as percent of region occupied by immunostaining) was performed on both brain hemispheres (ipsilaterally and contralaterally to the injection point) in 58 different regions. Thus, the p-α-syn dataset provided by the authors on their repository is an CSV sheet with values for 5 mice per group and timepoint for all 116 brain regions.

Using the dataset from [9] for synaptic connection, the authors generated a directed and weighed connectivity graph $G = V, E$ whose nodes V are N cortical and subcortical grey matter regions and whose edges $e_{ij} \in E$ represent an axonal projection from V_i to V_j. They then defined the weighed adjacency matrix of G as $A = [A_{ij}]$ generating a final parcellation of 116 regions.

The magnitude of the observed p-α-syn staining of all N nodes at a time t is the vector x_t. The predicted regional p-α-syn $\widehat{x_t}$ is a function of the adjacency matrix A and a seed region $seed$ ($\in E$) and is computed as shown in (1):

$$\widehat{x_t} = e^{-cLt}x_o \tag{1}$$

Where:

- c is a constant designed to achieve an optimal match with the data as the empirical diffusion constant for the pathology is *a priori* unknown.

- t stands for the timepoint of the prediction.

- L is the out-degree Laplacian matrix computed as shown in (2):

$$\begin{cases} -A_{ij} & \text{for } i \neq j \\ \sum_{j=1}^{N} A_{ij} & \text{for } i = j \end{cases} \tag{2}$$

- x_0 is the seed vector computed as shown in (3):

$$x_0 \begin{cases} 0 & \text{for } i \neq seed \\ 1 & \text{for } i = seed \end{cases} \tag{3}$$

3 Material and Methods

Henderson et al. coded their model using R and MATLAB. Codes and dataset were made fully accessible on GitHub by the authors. In order to only rely on opensource and free tools, we decided to reproduce the model in Python 3. We looked for operating systems intercompability and ran our code successfully on both Microsoft Windows 10 and MacOS 11.2.3. **Table 1** lists the required packages and their versions. Our code (and the associated datasets) is fully accessible in a GitHub repository https://github.com/MathieuBo/PathoSpreading.

Package	Version
NumPy	1.17.0
Pandas	1.1.5
Seaborn	0.11.1
Matplotlib	3.3.2
SciPy	1.5.2
Statsmodels	0.12.1
tqdm	4.7.2

Table 1. Description of required packages

4 Results

4.1 Network model of pathological α-synuclein spread

The model comprises a single free parameter c, the seed's diffusivity coefficient, that is used to scale the model to experimental results. The model is initiated by creating the seed vector x_0 for the inoculated region: here the caudoputamen (CPu). Then, the model iterates over possible the values of $c \in [0, 10]$. We selected the value of c that maximizes model fit, defined as the Pearson's correlation r between $\log_{10} x_t$ and $\log_{10} \widehat{x_t}$ averaged over all timepoints. **Table 2** summarizes the results including selected c value and correlation results. Correlation results for each timepoints are also plotted in **Figure 1A-C** (to be compared with Figure4a from the original article). Overall, we obtained a perfect replication, both quantitatively and qualitatively, of the original published results.

		Henderson et al.	Replication
Best c		1.625	1.625
MPI 1			
	Pearson's r	0.559	0.559
	p-value	$8.085e^{-9}$	$8.085e^{-9}$
MPI 3			
	Pearson's r	0.696	0.696
	p-value	$3.451e^{-17}$	$3.451e^{-17}$
MPI6			
	Pearson's r	0.648	0.648
	p-value	$2.638e^{-14}$	$2.638e^{-14}$

Table 2. Data correspondence between the reproduced results and the original article. The results have been rounded up to the third decimal. P-values for Pearson's correlation tests were corrected for multiple comparisons using Bonferroni's method. MPI: month(s) post injection.

Then, we performed some control experiments on the model to assess its robustness. First, as in the original article, we randomly seeded the diffusion model in all brain regions and calculated the value c. The 'true' seeded region (CPu) was among the best fits (97th percentile), at each timepoint (**Figure 1D**). Then, we implemented 2 additional controls: we randomly shuffled (over 150 iterations) the connectivity matrix (**Figure 1E**) and the p-α-syn matrix (**Figure 1F**), as in [11]. For each control experiment, we plotted the corresponding value of r. One can appreciate the profound performance loss in the different control experiments. Our results confirm that the spatiotemporal spread of pathological α-synuclein is driven by connectivity.

4.2 Differential regional vulnerability is correlated with α-synuclein expression

Intrinsic regional vulnerability is hypothesized to be a critical factor in the development of brain pathology related to α-synuclein or other pathogenic proteins [12]. Henderson et al. thought to investigate that aspect using the connectivity-based model. To do so, they used the difference between prediction and experiment (i.e. the residues of the linear regression) as a measure of relative vulnerability. Regions with higher p-α-syn staining than predicted were defined as vulnerable while the ones with lower values than predicted were considered resilient (**Figure 2A**). We here present an example at 6 months post-injection (**Figure 2B**). We used a 'lollipop' plot to help appreciate the distance to linear regression. Residues presented a bell shape-like distribution centered on 0 (**Figure 2C**). We plotted the vulnerability averaged over the 3 timepoints (**Figure 2D**). Similarly to the original article, we observed that anterior cortical regions (motor
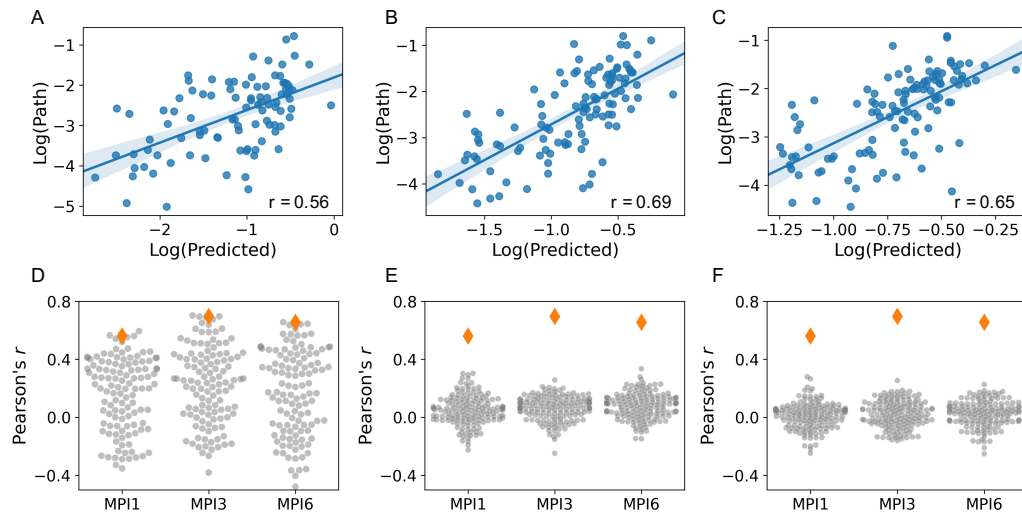
Figure 1. Network diffusion model based on anatomical connectivity explains pathological α-synuclein spread. (**A-C**) Scatterplots of $\log_{10}$(predicted) p-α-syn surface versus experimentally-measured $\log_{10}$(p-α-syn surface) values at 1 month (**A**), 3 months (**B**), and 6 months (**C**). The blue lines are the linear correlations between experimental and predicted values (shaded areas: 95% confidence interval). Values at the bottom right of each panel are the Pearson's correlation coefficient r between experiments and model's predictions. (**D-F**) Experimental controls: random seeding in other brain regions (**D**); shuffle of connectivity matrix (**E**); and shuffle of p-α-syn matrix (**F**). For each control experiment, the orange marker represents the actual model fit. MPI: month(s) post injection. Path: experimentally-measured p-α-syn surface.

cortex for example), the piriform cortex or the amygdalar region are highly vulnerable.

To determine whether model performance could be improved by the addition of other components, Henderson et al. seek out for factors that could contribute to regional vulnerability. An immediate assumption is that local levels of messenger-RNA (mRNA) encoding for α-synuclein could influence the future accumulation of α-synuclein in aggregates. To directly test that hypothesis, the model was modified to account for regional levels of α-synuclein mRNA (**Figure 2E**) by direclty modifying the adjacency matrix $A = S \times A$, where S is the diagonal matrix of the vector R containing the regional levels of α-synuclein mRNA for each brain region of A such as:

$$ S = \begin{cases} R_i & \text{for } i = j \\ 0 & \text{otherwise} \end{cases} $$

(4)

Model fitting and performance estimation remained similar. Incorporation of α-synuclein trancript's levels in the network diffusion model improved model performance by up 12% of added explained variance (**Figure 2F-H**). Here also, we perfectly replicated the original results both qualitatively (shown in **Figure 2F-H**) and quantitatively (shown in **Table 3**). Overall, these data suggest that *SNCA* gene expression is an important factor impacting the vulnerability of regions and subsequently pathology spread.

	Henderson et al.	Replication
Best c	0.414	0.414
MPI 1		
Pearson's r	0.580	0.580
p-value	$1.401e^{-9}$	$1.401e^{-9}$
MPI 3		
Pearson's r	0.727	0.727
p-value	$2.421e^{-19}$	$2.421e^{-19}$
MPI6		
Peasron's r	0.739	0.739
p-value	$2.965e^{-20}$	$2.965e^{-20}$

Table 3. Data correspondence between the original article and the reproduced results using the model incorporating α-synuclein mRNA levels. The results have been rounded up to the third decimal. P-values for Pearson's correlation tests were corrected for multiple comparisons using Bonferroni's method. MPI: month(s) post injection.

4.3 In silico seeding of alternative regions

To examine the generalizability of the model (*i.e.* the ability to apply to other conditions), Henderson et al. proceeded to *in silico* seeding of alternative regions such as the Piriform Cortex (Pir) and the *Substantia Nigra* (SN). **Figure 3** displays the p-α-syn spread at 1, 3 and 6 month(s) post-injection after seeding the model with either the Pir or the SN. We successfully obtained similar predictions after seeding in either the Pir or the SN. *In silico* injection in the Pir remarkably corresponds to previous semi-quantitative p-α-syn grading study [14] (**Figure 3A**). On the other hand, *in silico* injection in the SN resulted in slow propagation through the nigrostrial tract to first the caudoputamen and then to cortical regions (**Figure 3B**). The observed pattern following SN injection mimicks previous experimental results and the staging of human cases [2, 15, 16]. Taken together, these findings support the ability of the Diffusion Network model to generalize data predictions.

5 Conclusion

In this article, we successfully replicated the computational model created by Henderson et al. [7] in Python.

The Network Diffusion Model initially developed by Raj et al. [4] and recently adapted by Henderson et al. [7] is a powerful tool to understand the spread of p-α-syn in the brain. In its initial form, the model is solely based on anatomical connectivity between brain regions. Despite its simplicity, the model is able to predict 31% to 47% of the total variance. In a second version, the endogenous α-synuclein expression (at the mRNA level) is added to the model. The addition improves model performance (up to 12% of added explained variance), although a quantitative analysis should be performed in the future.

Similarly to what is observed in preclinical models of synucleinopathies or in the brain of Parkinson's disease's patients, α-synuclein pathology is composed of 'islands' of vulnerable neurons susceptible to developing pathogenic protein inclusions. By comparing the observed and predicted p-α-syn staining based on anatomical connectivity, the model allows to explore regional vulnerability. Finally, seeding in different regions such as the Pir and the SN supports the network diffusion model generalizability. Overall, this model is attractive as it is simple, easily handled, and replicable.

Nonetheless, the model is not without limitations. Other experimental datasets (investigating other seeding regions or using alternative markers for pathological α-synuclein) are still required to quantitatively assess the ability of the model to generalize the prediction of α-synuclein spread. Another barrier comes from the available connectivity data [9]. The connectivity matrix used by Henderson et al. and ourselves is at the mesoscale. One implication, for example, is that brain structures are considered homogenous, hence not taking into account the laminar cortical organization for example. The observation of a distinct distribution of the α-synuclein pathology in cortical layer suggests that lower scale connectivity data could help to capture overlooked events.

Modeling disease propagation in the brain is challenging and often reductionist *per se*. Indeed, during the course of disease, the brain changes according to at least two parallel processes: the "Dynamics **On** Network" and the "Dynamics **Of** Network" [17]. The first one is related to processes that occur atop a static structure as, for instance, α-synuclein spread. The second defines the brain network changes over time (e.g. neurodegeneration or extracellular remodelling [18]). A major challenge remains in modeling both these "on" and "of" dynamics. If we were only to understand one dynamic and leave the other dynamic unexplored, we could imagine to still be a step away from the reality of the pathology.

6 Acknowledgements

References

1. B. Dehay et al. "Targeting α-synuclein for treatment of Parkinson's disease: mechanistic and therapeutic considerations." In: **Lancet Neurol** 14.8 (Aug. 2015), pp. 855–866. DOI: 10.1016/S1474-4422(15)00006-X.
2. H. Braak, K. Del Tredici, U. Rüb, R. A. de Vos, E. N. Jansen Steur, and E. Braak. "Staging of brain pathology related to sporadic Parkinson's disease." In: **Neurobiol Aging** 24.2 (2003), pp. 197–211. DOI: 10.1016/s0197-4580(02)00065-9.
3. M. Jucker and L. C. Walker. "Propagation and spread of pathogenic protein assemblies in neurodegenerative diseases." In: **Nat Neurosci** 21.10 (Oct. 2018), pp. 1341–1349. DOI: 10.1038/s41593-018-0238-6.
4. A. Raj, A. Kuceyeski, and M. Weiner. "A network diffusion model of disease progression in dementia." In: **Neuron** 73.6 (Mar. 2012), pp. 1204–1215. DOI: 10.1016/j.neuron.2011.12.040.
5. C. Mezias and A. Raj. "Analysis of Amyloid-β Pathology Spread in Mouse Models Suggests Spread Is Driven by Spatial Proximity, Not Connectivity." In: **Front Neurol** 8 (2017), p. 653. DOI: 10.3389/fneur.2017.00653.
6. C. Mezias, E. LoCastro, C. Xia, and A. Raj. "Connectivity, not region-intrinsic properties, predicts regional vulnerability to progressive tau pathology in mouse models of disease." In: **Acta Neuropathol Commun** 5.1 (Aug. 2017), p. 61. DOI: 10.1186/s40478-017-0459-z.
7. M. X. Henderson, E. J. Cornblath, A. Darwich, B. Zhang, H. Brown, R. J. Gathagan, R. M. Sandler, D. S. Bassett, J. Q. Trojanowski, and V. M. Y. Lee. "Spread of α-synuclein pathology through the brain connectome is modulated by selective vulnerability and predicted by network analysis." In: **Nat Neurosci** 22.8 (Aug. 2019), pp. 1248–1257. DOI: 10.1038/s41593-019-0457-5.
8. O. Sporns. "Graph theory methods: applications in brain networks." In: **Dialogues Clin Neurosci** 20.2 (June 2018), pp. 111–121. DOI: 10.31887/DCNS.2018.20.2/osporns.
9. S. W. Oh et al. "A mesoscale connectome of the mouse brain." In: **Nature** 508.7495 (Apr. 2014), pp. 207–214. DOI: 10.1038/nature13186.
10. K. C. Luk, V. Kehm, J. Carroll, B. Zhang, P. O'Brien, J. Q. Trojanowski, and V. M. Lee. "Pathological α-synuclein transmission initiates Parkinson-like neurodegeneration in nontransgenic mice." In: **Science** 338.6109 (Nov. 2012), pp. 949–953. DOI: 10.1126/science.1227157.
11. P. Pandya, C. Mezias, and A. Raj. "Predictive Model of Spread of Progressive Supranuclear Palsy Using Directional Network Diffusion." In: **Front Neurol** 8 (Dec. 2017), p. 692. DOI: 10.3389/fneur.2017.00692.

12. H. Fu, J. Hardy, and K. E. Duff. "Selective vulnerability in neurodegenerative diseases." In: **Nat Neurosci** 21.10 (Oct. 2018), pp. 1350–1358.

13. F. Claudi, A. L. Tyson, L. Petrucco, T. W. Margrie, R. Portugues, and T. Branco. "Visualizing anatomically registered data with Brainrender." In: **eLife** 10 (Mar. 2021). DOI: 10.7554/elife.65751.

14. N. L. Rey, J. A. Steiner, N. Maroof, K. C. Luk, Z. Madaj, J. Q. Trojanowski, V. M.-Y. Lee, and P. Brundin. "Widespread transneuronal propagation of alpha-synucleinopathy triggered in olfactory bulb mimics prodromal Parkinson's disease." In: **Journal of Experimental Medicine** 213.9 (Aug. 2016), pp. 1759–1778. DOI: 10.1084/jem.20160368.

15. A. Recasens et al. "Lewy body extracts from Parkinson disease brains trigger α-synuclein pathology and neurodegeneration in mice and monkeys." In: **Ann Neurol** 75.3 (Mar. 2014), pp. 351–362.

16. M. Bourdenx et al. "Identification of distinct pathological signatures induced by patient-derived α-synuclein structures in nonhuman primates." In: **Sci Adv** 6.20 (May 2020), eaaz9165.

17. A. Raj and F. Powell. "Models of Network Spread and Network Degeneration in Brain Disorders." In: **Biological Psychiatry: Cognitive Neuroscience and Neuroimaging** 3.9 (Sept. 2018), pp. 788–797. DOI: 10.1016/j.bpsc.2018.07.012.

18. F. N. Soria et al. "Synucleinopathy alters nanoscale organization and diffusion in the brain extracellular space through hyaluronan remodeling." In: **Nat Commun** 11.1 (July 2020), p. 3440.
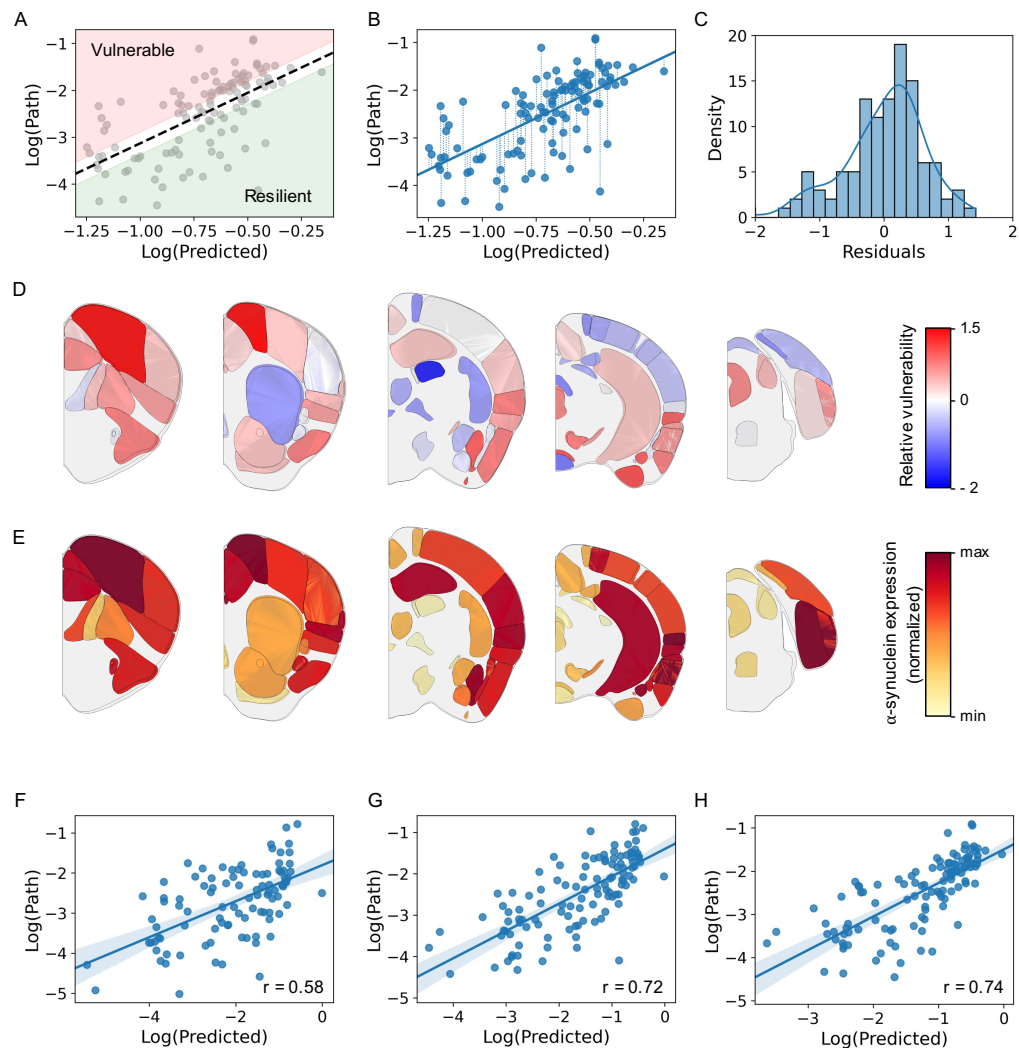
Figure 2. Differential regional vulnerability is correlated with α-synuclein expression. (**A**) Conceptual framework for the assessment of regional vulnerability of brain regions. We here show the scatterplot of $\log_{10}$(Predicted) and $\log_{10}$(Measured) at 6 months post-injection. The black dashed line is the linear regression between prediction and experimental observation. The red area indicates brain regions with higher measured p-α-syn staining than predicted by the model, hence 'vulnerable'. Conversely, the green area indicates brain regions with lower observed p-α-syn staining than predicted by connectivity, hence 'resilient' brain regions. (**B**) Example of a lollipop plot for the 6 months timepoint highlighting the distance to the mean linear regression (i.e. residues - dashed blue lines). (**C**) Distribution of linear regression residues. The solid blue line is the kernel density estimation of the distribution. (**D**) Heatmap of the residuals between the predicted and measured p-α-syn staining plotted on an anatomical mouse brain as a measure of the relative vulnerability of regions averaged over the 3 experimental timepoints, using the *Brainrender* package [13]. **Note**: the differences with the original article arise mostly from the use of different rendering method. As shown by the numerical comparison in **Table1**, the ouput values were absolutely similar. (**E**) Heat map of the α-synuclein mean expression energy values obtained from the Allen Brain Atlas *in situ* hybridization study for each of the designated region. Expression was normalized using min/max to ease visualization. **Note**: Henderson et al. used a divergent colormap to represent a continous value, hence making it difficult to interpret regional variations. We here chose a sequential continuous colormap to improve the understanding. (**F-H**) Scatterplots of $\log_{10}$(predicted) p-α-syn versus experimentally-measured $\log_{10}$(p-α-syn) values at 1 month (**F**), 3 months (**G**), and 6 months (**H**). The blue lines are the linear correlations between experimental and predicted values (shaded areas: 95% confidence interval). Values at the bottom right of each panel are the Pearson's correlation coefficient r between experimental and predicted pathologies. Path: experimentally-measured p-α-syn surface.
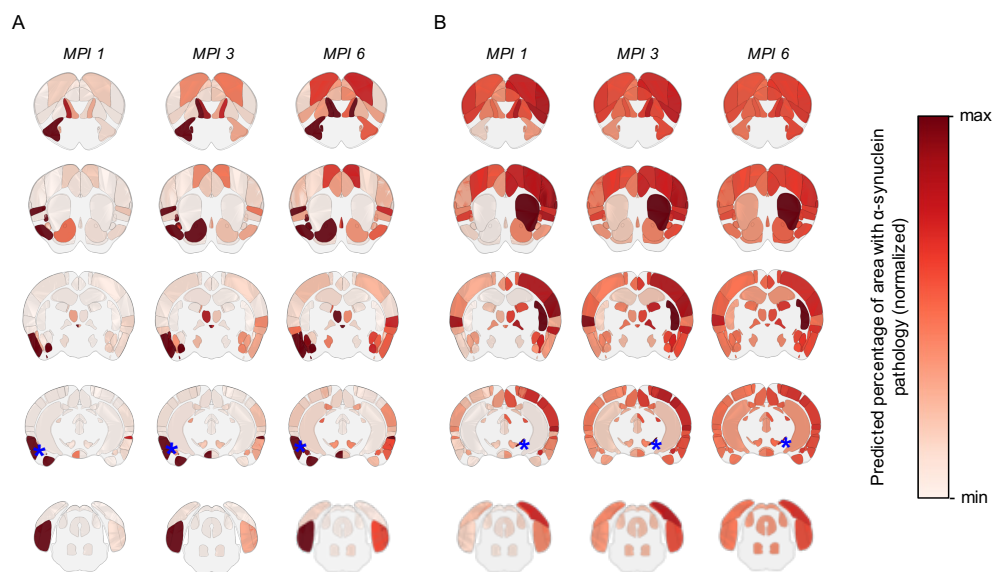
Figure 3. *In silico* seeding of alternative regions in the mouse brain Predicted p-α-syn staining after *in silico* seeding in the Piriform Cortex (**A**) or the Substantia Nigra (**B**) The injection site is indicated on each slice by a blue asterisk. Both Figure A and B are displayed using the *Brainrender* package [13]. **Note**: Henderson et al. used a divergent colormap to represent a continous value, hence making it difficult to interpret regional variations. We here chose a sequential continuous colormap to improve the understanding.