Personalised Regional Modelling Predicts Tau Progression in the Human Brain

Pavanjit Chaggar[1], Jacob Vogel[2], Alexa Pichet Binette[2], Travis B. Thompson[3], Olof Strandberg[2], Niklas Mattsson-Carlgren[2, 7, 8], Linda Karlsson[2], Erik Stomrud[2], Saad Jbabdi[4], Stefano Magon[5], Gregory Klein[5], the Alzheimer's Disease Neuroimaging Initiative [*] , Oskar Hansson[2, 6], and Alain Goriely[1]

[1] *Mathematical Institute, University of Oxford*
[2] *Clinical Memory Research Unit, Department of Clinical Sciences Malmo, Lund University*
[3] *Department of Mathematics and Statistics, Texas Tech University, USA*
[4] *Wellcome Trust Centre for Integrative Neuroscience, University of Oxford*
[5] *F. Hoffmann-La Roche Ltd, Basel, Switzerland*
[6] *Memory Clinic, Skane University Hospital, Lund University*
[7] *Department of Neurology, Skåne University Hospital, Lund University*
[8] *Wallenberg Center for Molecular Medicine, Lund University*

June 4, 2025

Abstract

Aggregation of the hyperphosphorylated tau protein is a central driver of Alzheimer's disease, and its accumulation exhibits a rich spatiotemporal pattern that unfolds during the course of the disease, sequentially progressing through the brain across axonal connections. It is unclear how this spatiotemporal process is orchestrated, namely, to what extent the spread of pathologic tau is governed by transport between brain regions, local production, or both. To address this, we develop a mechanistic model from tau PET data to describe tau dynamics along the Alzheimer's disease timeline. Our analysis reveals longitudinal changes in production and transport dynamics in two independent cohorts, with subjects in the early stage of the disease exhibiting transport-dominated spread, consistent with an initial spread of pathological tau seeds, and subjects in the late stage disease characterised primarily by local tau production. Furthermore, we demonstrate that the

[*]Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: `http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf`

model can accurately predict subject-specific longitudinal tau accumulation at the regional level, potentially providing a new clinical tool to monitor and classify patient disease progression.

Teaser: A mechanistic model reveals tau protein dynamics in Alzheimer's, showing stage-specific shifts in transport and local production.

1 Introduction

Alzheimer's disease (AD) is a devastating neurological condition resulting in progressive brain atrophy and cognitive decline. The toxic forms of two proteins, amyloid-β (Aβ) and tau protein (tau) are believed to act in concert to drive AD progression [1, 2]. The pathological roles of these proteins in the human brain during AD have been investigated using positron emission tomography (PET), with radiotracers such as [^{18}F]florbetapir and [^{18}F]flortaucipir allowing for in-vivo quantification of Aβ and tau, respectively [3]. While Aβ tends to be more diffusely present throughout the cerebral cortex [4, 5, 6, 7], tau exhibits richer spatiotemporal dynamics, characterised by Braak staging [8]. Braak staging describes the trajectory of toxic tau, starting from the entorhinal cortex and sequentially progressing to the limbic regions, the basal temporal lobes, the broader association cortex, and finally the primary sensory cortex. This staging pattern has been validated using tau PET imaging [9, 10, 11] and has been shown to be highly correlated with atrophy and cognitive decline [12, 13], however, the mechanism for how tau staging is orchestrated remains unclear.

Growing evidence suggests that the progression of AD depends on two distinct factors: 1) the local production of toxic proteins; 2) the transport of toxic proteins throughout the brain. However, it has yet to be determined to what extent these factors contribute to the progression of AD and whether their contributions change over time. There is now substantial evidence that tau propagation follows a prion-like mechanism, progressively forming toxic oligomeric seeds and neurofibrillary tangles through an autocatalyic production process [14, 15]. The prion-like nature of tau has been demonstrated with transgenic animal models in which cortical injections of tau seeds induce the formation of tau aggregates that grow in concentration over time at the injection site and surrounding areas [16, 17]. In 2012, studies by Liu et al. and de Calignon et al. showed that transgenic mice overexpressing pathological human tau in the entorhinal cortex exhibit accumulation of tau aggregates and that tau invades axonally connected regions through transsynaptic transport to form seeds in otherwise healthy regions [18, 19]. Prion-like aggregation and axon-based tau transport have also been suggested in human postmortem studies [20] and in vivo studies using structural connectome-based models of tau PET capable of reproducing observed tau aggregation and spread [21, 22, 23, 24, 25, 26]. In a recent investigation, Meisl et al. analysed multimodal tau data from

Braak stage 3 onward and showed that tau production, not transport, is the main contributor of tau progression [27]. However, the study does not account for the spatial progression across individual brain regions or estimate dynamics across the full AD progression timeline. Unanswered questions remain about whether there are changes in tau production and transport rates over time and whether the balance of these two processes changes along the disease timeline. To address these outstanding questions, we develop a whole brain model capable of accurately describing longitudinal tau PET data and conduct a multicohort study to analyse tau dynamics across the full disease progression timeline.

To answer questions about temporal changes in AD tau dynamics in the human brain, an accurate and reliable model of longitudinal tau observations is needed. In the past decade, there have been numerous efforts to use mathematical models to better understand the spatiotemporal properties of AD pathology, ranging from linear diffusion models of tau [21, 22] to infinite-dimensional spatiotemporal models of toxic protein aggregation [28]. Each of these models makes different assumptions about the physical mechanisms of tau spread; however, there has not been a unifying effort to rigorously compare commonly used models to identify which are best able to accurately describe longitudinal tau PET observations. In addition, the models currently described in the literature do not account for regional variations in tau dynamics, which has been shown to influence tau progression [29, 30, 31] and are incapable of predicting longitudinal changes at the regional level. Here, we present a novel model that provides a qualitative account of regional vulnerability and its effect on tau progression. Using a previously developed Bayesian pipeline for longitudinal tau modelling [32, 26], we perform hypothesis-driven model selection on a family of common models from the AD modelling literature, including a new model that accounts for regional dynamics. We show that models that rely only on network diffusion or homogeneous tau production dynamics are not sufficient to model regional longitudinal data, whereas models accounting for regional variations in tau dynamics are able to accurately model longitudinal tau observations at a regional level.

We combine state-of-the-art modelling and inference methods with longitudinal tau PET data from two independent cohorts to address the outstanding question of how tau transport and production drive AD progression. To determine whether there are changes in tau transport and production rates during the progression of AD, we apply our model to three groups representing different stages of the AD timeline [33, 34]: (i) A^+T^+: amyloid positive, tau positive; (ii) A^+T^-: amyloid positive, tau negative;(iii) A^-T^-: amyloid negative tau negative, . We show that tau transport is faster in early stage disease (A^+T^-), and that there are primary and secondary increases in tau production dynamics throughout the disease timeline. We validate these results on an independent dataset using a different tau tracer, BioFINDER-2, on which the same results are obtained, further

showing that the model and results are robust and generalisable between datasets and the choice of tau tracer. Finally, we validate our model by showing that it can *forecast* regional rates of tau accumulation over time for individual patients. The combination of these methods provides a novel pipeline for analysing and understanding longitudinal tau data, allowing us to compare changes in disease dynamics throughout the AD timeline and predict subject-specific, region-specific changes in tau over time.

2 Results

2.1 Deriving a generative model of tau dynamics

We first extend previous work [35, 36, 28] to develop a mechanism-based model of tau dynamics in the human brain that can be calibrated using tau PET data. This model, called the *local FKPP model* (which is a network version of the well-known continuous Fisher-Kolmogorov–Petrovsky–Piskunov equation) and derived in full detail in Section 4.4, is given by a set of nonlinear ordinary differential equations on a structural connectome network of R nodes for the variables $s_i = s_i(t)$ for $i = 1, \ldots, R$, representing tau SUVR in different regions of interest:

$$\frac{\mathrm{d}s_i}{\mathrm{d}t} = \underbrace{-\rho \sum_{j=1}^{R} \mathcal{L}_{ij} \left(s_j - s_{0,j}\right)}_{transport} + \underbrace{\alpha \left(s_i - s_{0,i}\right) \left[\left(s_{\infty,i} - s_{0,i}\right) - \left(s_i - s_{0,i}\right)\right]}_{production}, \qquad i = 1, \ldots, R. \tag{1}$$

The first term represents the contribution of tau transport between brain regions through a graph Laplacian $\mathcal{L}$ with uniform rate ρ for the i-th region, consistent with previous work [21, 22, 24, 36]. The second term represents prion-like production at the i-th region with uniform rate α [26, 36]. Therefore, the local FKPP model describes future changes in tau SUVR as resulting from a combination of transport through axonal connections and prion-like production. In this study, regions of interest are given by the 68 cortical regions of the Desikan-Killiany-Tourville (DKT) atlas, in addition to bilateral hippocampus and amygdala, therefore, there are $R = 72$ nodes in the connectome model. We introduce two novel parameter vectors, regional baseline values, $s_{0,i}$, and carrying capacities $s_{\infty,i}$, that represent a healthy state and a late stage AD state, respectively, which add information about regional variations in production dynamics. To make the relationship to regional variability in production rates of tau more clear, consider a change of variables to $q_i = \left(s_i - s_{0,i}\right) / \left(s_{\infty,i} - s_{0,i}\right)$ for $i = 1 \ldots R$, then (1) becomes a standard (network) FKPP equation:

$$\frac{\mathrm{d}q_i}{\mathrm{d}t} = -\rho \sum_{j=1}^{R} \mathcal{L}_{ij} q_j + \hat{\alpha}_i q_i \left(1 - q_i\right) \qquad i = 1, \ldots, R, \tag{2}$$

with $\hat{\alpha}_i = \alpha\,(s_{\infty,i} - s_{0,i})$. In this form, regional tau evolves between $[0, 1]$ with a rate that depends on the difference between regional SUVR baseline values and carrying capacities, with larger differences equating to faster regional tau production, making the notion of regional vulnerability more explicit.

These parameters are derived from the Gaussian mixture modelling approach used in [24], in which a two-component Gaussian mixture model is used to describe healthy and pathological tau SUVR distributions. (see Section 4.5 for more details). An example of this is shown in Fig. 1b for the inferior temporal lobe, and the carrying capacities for the right hemisphere are shown in Fig. 1c. Since these parameters are estimated from tau PET, they also encode specific features of the tracer, such as regional differences in tracer uptake, specificity to 3R/4R tau pathology, on-target binding and off-target binding, and therefore allow us to model tau SUVR directly. An example of simulated regional trajectories is shown in Fig. 1a, encapsulating the effect of both transport and production. A consequence of the variation of carrying capacities is that the regional production rates also vary between regions, as seen in the middle panel of Fig. 1a, providing a qualitative account of regional vulnerability. Using the rescaled local FKPP model (2), we can also track the simulated concentration of tau in Braak regions, shown in the bottom panel of Fig. 1a, and we see that the model correctly predicts the invasion of tau into Braak regions. The ability to account for these regional variations extends previous models with homogeneous dynamics across regions [36, 23] providing a picture of tau progression that is more consistent with observed tau staging.

Figure 1: **Simulated transport and production dynamics in the local FKPP model.** **1a** Simulation from the local FKPP model using carrying capacities derived from Gaussian mixture models (shown in 1b). Simulations are initialised with a seed value of $(s_{0,i} + s_{\infty,i})/2$ in the bilateral entorhinal cortex, $i = \{27, 63\}$, with $\rho = 0.025$ and $\alpha = 0.25$. Each line in the middle panel represents the SUVR trajectory of one DKT atlas brain region. Values at time points $t = \{0, 15, 30, 45\}$ years are projected onto a cortical rendering in the top panel. Each line in the bottom panel represents concentration averaged over Braak regions, after rescaling simulated SUVR as $q_i = (s_i - s_{0,i}) \,/\, (s_{\infty,i} - s_{0,i})$. **1b** Two component Gaussian mixture model fit to a multi-cohort tau PET dataset [24] and data from ADNI for inferior temporal lobe. Baseline values for each region are taken as the mean of regional T^- distribution and the carrying capacity as the 99-th percentile of the regional T^+ distribution, and these are used to simulate the model in 1a and throughout this paper. **1c** Right hemisphere cortical rendering of the SUVR carrying capacities as determined through Gaussian mixture modelling.

2.2 Regional Heterogeneity Is Necessary for Longitudinal Prediction

To determine whether the local FKPP is capable of fitting observed AD trajectories, we compare it to tau PET data. For comparison, we also consider simpler models that can be obtained from the local FKPP model Eq. (1), namely *the global FKPP model*, Eq. (12), obtained by assuming

that none of the parameters vary locally, *the diffusion model* Eq. (11) obtained by taking $\alpha = 0$ in Eq. (1), and *the logistic model* Eq. (10) obtained by neglecting transport between regions, $\rho = 0$.

We use hierarchical Bayesian inference to calibrate each model to tau PET data, allowing us to quantify whether the model parameters can be identified from patient data and provide bounds on uncertainty for group- and individual-level parameters. We use tau PET data from ADNI, selecting Aβ^+ subjects who have at least three scans and are T^+ in the medial temporal or lateral temporal lobe (see Section 4.1 for details). We employ two metrics to compare models, the Bayesian information criteria to measure accuracy against in-sample data and the expected log predictive density (ELPD) to measure out-of-sample predictive accuracy, both provided in Table 1. Fig. 2a shows the in-sample longitudinal fit for each of the four models and Fig. 2b shows the regionally averaged out-of-sample longitudinal fit.

Figure 2: **Model Fit for In-sample and Out-of-Sample Data.**
2a Goodness of fit for four models, local FKPP, global FKPP, diffusion and logistic. For all panels, each point represents a region in the connectome model, averaged over subjects per scan. Top row shows estimated vs observed SUVR values. Bottom row shows estimated change vs observed change in SUVR. Only the local FKPP and logistic model are able to accurately capture longitudinal changes, while the global FKPP and diffusion models are each structurally incapable of describing heterogeneous production. **2b** Out-of-sample fits for four models of proteopathy. Top row: predicted vs observed out-of-sample SUVR. Bottom row: predicted change vs observed change from first in-sample scan to last out-of-sample scan. Each point represents a region in the connectome model, averaged over subjects.

The local FKPP performs best for both in-sample fit and out-of-sample predictive accuracy, followed by the logistic model for in-sample fit and the global FKPP model for out-of-sample fit. The results show that the diffusion model is not suitable for longitudinal modelling of tau PET data, clearly shown in Fig. 2a and Fig. 2b, since there is no mechanism for tau production or clearance and therefore the total concentration is conserved. The global FKPP model can capture changes in tau load, but it cannot describe the regional heterogeneity in tau production. The deficiencies in production dynamics of the diffusion and global FKPP models are addressed with the local FKPP model and the logistic model and the ability of these models to accurately describe longitudinal tau PET underlines the importance of including regionally specific production rates. The logistic model is capable of describing the trajectory of tau PET despite not being able to capture the transport of tau through the structural connectome, suggesting that A^+T^+ subjects may already have widespread invasion of tau seeds. However, the logistic model underfits changes in SUVR, particularly in the low SUVR range. The error in model fit relative to the final in-

	Local FKPP	**Global FKPP**	**Diffusion**	**Logistic**
BIC	**-33889.2**	-32823.1	-20539.1	-33138.7
ELPD	**638.22**	622.68	-175.579	607.14

Table 1: Assessment of model fit using the Bayesian information criteria, BIC, and the expected log predictive density, ELPD. Lower values are considered better for the BIC and higher values are better for the ELPD. The local FKPP model performs best in both metrics.

sample scan is shown in Fig. S1 for the local FKPP, global FKPP and logistic model. The residual analysis shows that the logistic model is prone to underestimate SUVR (mean error $= -0.02$, s.d. $= 0.02$), particularly for lower SUVR ranges, as shown in Fig. S1, while the error in the local FKPP (mean error $= -0.01$, s.d. $= 0.02$) and global FKPP (mean error $= 0.005$, s.d. $= 0.03$) is more balanced between regions. This indicates that initial tau deposition may be driven by transport between regions. Additionally, the global FKPP has a higher variance in predictions, often either overestimating or underestimating SUVR as a result of regionally homogeneous dynamics. Overall, the data support the use of the local FKPP model, as evidenced by its being the most capable of describing in-sample and out-of-sample data, while also capturing the role of both tau transport and local tau production.

2.3 Local model forecasts regional tau progression

A major possible benefit of the mathematical modelling of AD lies in its application to clinical and pharmacological research, particularly by predicting the progression of the patient's disease. Here, we show that the local FKPP model can be used to accurately *forecast* the trajectory of tau PET. We divide the ADNI A^+T^+ cohort into train and test cohorts. We define a fixed training set comprising $n = 41$ subjects who have three longitudinal scans. For the remaining $N = 16$ subjects with more than three longitudinal scans, we perform inference three times, each time adding an additional scan to the training set, starting with a single scan. The resulting posterior predictive trajectories for the left inferior temporal lobe are shown in Fig. 3. In Fig. S2 we provide the posterior predictive trajectories without observation noise, highlighting the uncertainty in model parameters. The results show that a single scan is often insufficient to provide meaningful forecast accuracy, despite benefiting from information pooled across subjects in the hierarchical Bayesian model. This greatly improves with the addition of a second data point; however, in some cases this produces inaccurate forecasts of future data if there is a decrease in tau SUVR, perhaps due to atrophy. With the addition of a third data point, the results generally converge with low uncertainty and accurately forecast future observations. Similar results are shown in the supplementary information for the entorhinal cortex Fig. S3 and Fig. S4. The results suggest that a more constrained approach which

Figure 3: **Out-of-sample fit and posterior predictive plots for the left inferior temporal region.**
The local FKPP model was iteratively calibrated to a A^+T^+ ADNI cohort with 41 in-sample subjects and 16 test subjects. Three iteration were run where for each iteration an additional scan from the test subjects were included, starting with a single scan. Posterior predictive trajectories for left inferior temporal lobe are shown for each iteration (neglecting observation noise). In the above figure, each panel represents one of the 16 test subjects. Each point represents a data point added for a training iteration; trajectories are color matched to correspond to the number of longitudinal data points included for training.

more heavily weights the effect of population priors on forecasts may prove fruitful in circumstances in which there is insufficient data to capture an individual's trajectory. Nonetheless, these results demonstrate the power of a simple model based on key biological priors (namely the connectome, regional baseline values, and regional carrying capacities) and two free parameters to forecast the regional progression of tau PET that may provide benefit to clinical and pharmaceutical researchers.

2.4 Early AD progression is driven by tau transport

Next, we sought to determine whether there are any changes in tau production and transport dynamics across the AD progression timeline. To do so, we use two cohorts of tau PET data, ADNI and BioFINDER-2 (BF2), each divided into three groups, A^+T^+, A^+T^-, A^-T^-, representing different stages of AD. Since BF2 uses a different tau PET radiotracer, we rerun the Gaussian mixture modelling analysis to recover the tracer-specific baseline and carrying capacities. We then apply the NUTS sampling algorithm to the hierarchical Bayesian model to obtain samples from the posterior distribution, Eq. (17), providing distributions for model parameters at the population level and individual level.

The distributions of the population parameters for the A^+T^+, A^+T^-, A^-T^- groups are shown in Fig. 4a for ADNI and Fig. 4b for BF2 and are summarised in Table 2. The posterior distributions between cohorts are qualitatively the same, with changes likely reflecting differences in cohort and tracers. The inferred parameters show an increase in the transport rate for the A^+T^- group compared to the A^+T^+ and A^-T^- groups, suggesting that tau spreads more easily between regions during the early stages of the disease and is minimal in the later stages of AD. The inferred posterior distributions for the production parameter show a progressive increase in the production rate along the disease timeline, with a primary increase from the A^-T^- to the A^+T^- group and a secondary increase from the A^+T^- group to the A^+T^+ group. The negative production rate for the A^-T^- group indicates that the signal, on average, decreases. This could reflect changes in noise due to

off-target, nonspecific binding, or atrophy from non-AD related neurodegeneration. These results suggest that in early AD tau begins in a transport-dominated phase ($\rho > \alpha$) and then switches to a production-dominated phase ($\alpha > \rho$).

Figure 4: **Inferred population level parameters using ADNI and BF2 tau data**.
4a & 4b Population production and transport parameters across A^+T^+, A^+T^- and A^-T^- groups for ADNI (4a) and BF2 (4b) tau PET data. **4c & 4d** Inferred population production and transport parameters from spatially shuffled data (shown in grey) compare to inferred distributions from true data for A^+T^+ (4c), A^+T^- (4d) ADNI groups. Asterisks between groups denotes distributions are significantly different ($p < 0.01$), tested using the Mann-Whitney U test.

To confirm that the parameter distributions reflect meaningful dynamics present in the data and are not as a result of statistical patterns, we rerun the analysis on the A^+T^- and A^+T^+ groups using spatially shuffled data. A comparison of the posterior distributions obtained shuffled data and those obtained from true data are shown in Figs. 4c and 4d. Note that we only spatially shuffle the data and therefore expect minimal changes to the estimated production parameters. We note a marked difference in the estimated transport dynamics in both the A^+T^- and A^+T^+ groups between the true and shuffled data, confirming that the dynamics present in the data are not a consequence of statistical patterns in the data, but represent tau dynamics measured through PET.

To examine whether the negative production rates observed in Fig. 4 are a result of regional atrophy, we rerun the analysis in the ADNI cohort with partial volume correction applied [37, 38]. The population-level posterior distributions resulting from this analysis are shown in Fig. S5a. The results are qualitatively the same as those obtained using non-PVC data Fig. 4, with increases in tau production along the AD progression timeline and a faster transport rate in A^+T^- subjects compared to A^+T^+ and A^-T^- subjects. However, using PVC data results in an increased production rate for the A^-T^- and A^+T^- groups, largely eliminating the negative production rate observed with non-PVC data. This suggests that the negative production rate observed in Fig. 4 primarily results from brain atrophy outpacing tau production, resulting in a net reduction of SUVR.

Additionally, we rerun the analysis using an eroded white matter reference region, which has been suggested as a more effective reference region for longitudinal analysis [39]. The posterior distributions for this dataset are shown in Fig. S5b. The results show a similar transport effect to Fig. 4. However, there are significant differences in the inferred production parameters. The A^+T^- and A^+T^+ groups both display higher production than the A^-T^- group, however, there is no significant difference between the A^+T^- and A^+T^+ groups. After further analysis, we believe this effect is due to positive correlations in the reference region and the target SUVR that reduce group

differences and limit interpretation of the longitudinal analysis (analysis provided in supplementary information Section S1.3).

The large parcels present in the DKT atlas may affect the detail in which production and transport dynamics are present in the data, since data is averaged over large parcels with heterogeneous volumes. To address this, we rerun the analysis using the Schaefer-200 atlas [40] provided in the ENIGMA toolbox [41]. The population-level model parameters for the A^+T^+, A^+T^-, and A^-T^- groups are shown in Fig. S7. The transport parameters between groups are qualitatively similar to those shown in Fig. 4, further supporting an initial phase of accelerated tau transport. The production parameter was higher for the A^+T^+ group compared to the A^+T^- and A^-T^- groups, consistent with the DKT atlas. However, there is no longer a substantial increase from A^-T^- to A^+T^-, with the two groups showing similar tau production rates, possibly as a result of greater sensitivity to tau-related atrophy effects due to smaller parcel sizes.

The transport dominated phase of the early AD subjects supports evidence showing tau seeds are present throughout the cortex before symptom onset [20, 42] and, together with the small role of transport in the A^+T^+ groups, helps explain the strong performance of the logistic model in Section 2.2. Overall the results reveal temporal changes in the dynamics of tau progression, with an initial transport dominated phase, perhaps in which seeds are deposited around the cortex, followed by a production dominated phase indicative of secondary tauopathy, likely due to spatial colocalisation with Aβ catalysing tau production.

	Group	ρ_μ	ρ_σ	α_μ	α_σ
ADNI	A^+T^+	0.01	0.01	0.13	0.1
	A^+T^-	0.02	0.04	0.02	0.23
	A^-T^-	0.02	0.11	−0.04	0.19
BF2	A^+T^+	0.004	0.01	0.11	0.09
	A^+T^-	0.02	0.02	0.02	0.22
	A^-T^-	0.02	0.08	−0.03	0.20

Table 2: Summary of inferred posterior distributions providing the means of inferred population parameters for ADNI and BF2, where ρ_μ is the average of population transport parameter, ρ_σ is the standard deviation of the population transport parameter, α_μ is the average population production parameter, α_σ is the standard deviation of the population production parameter. Parameters are shown for each of the A^+T^+, A^+T^- and A^-T^- groups.

3 Discussion

We have derived a physics-based generative model to describe tau PET data in terms of underlying tau dynamics and applied it data from ADNI and BF2 to understand how it compares to other mod-

els present in the literature and how it can inform us about tau transport and production dynamics in the human brain. We have shown that this model can forecast accurately longitudinal regional tau PET progression in AD subjects. Furthermore, by performing inference across different patient groups across the AD disease timeline, we uncover temporal changes in transport and production dynamics, showing an initial transport dominated phase associated with primary tauopathy and seeding, followed by an accelerated production dominated phase indicative of secondary tauopathy.

Several studies have proposed different models of proteopathy in AD, a key difference between them being descriptions of the tau production process, which vary widely in complexity [22, 28, 24, 43, 27, 25, 26]. Here we present a parsimonious model of tau transport that relies on regionally specific carrying capacities and show, through model selection Section 2.2, that it is capable of outperforming other models proposed in the literature. A possible cause of regional heterogeneity in carrying capacities is heterogeneity in regional risk factors that promote tau proliferation, the most likely of which is Aβ. Aβ has its own spatial topography within AD patients, being particularly present throughout the fronto-partietal-temporal default mode network and stimulating neuronal hyper-activation [11, 44, 45]. The presence of Aβ will have a two-fold effect on tau dynamics, first through a catalysing effect on tau production [46, 47, 48] and second through the promotion of activity-dependent spread and production through functional networks [49, 50]. In Thompson et al. 2020, we formulated a model describing the dynamic interaction between Aβ and tau, that predicts an increase in carrying capacities based on Aβ concentration [51], however, further work toward simplifying the model will be necessary before it can be used for inference with patient data.

Another key set of factors that contribute to regional vulnerability are genetic markers. It has already been shown in mice models of AD that gene expression patterns can inform tau spread [31] and human models of Parkinson's disease have shown how gene expression patterns can inform regional vulnerability to create a model of toxic protein spread in Parkinson's disease [52]. There are several candidate genes for modelling regional vulnerability in AD, most notably the microtubule association protein tau (MAPT), as a proxy for relative baseline tau vulnerability [53] and apolipoprotein-E (APOE) for those patients with the APOEϵ4 mutation [29, 54, 55]. Although there are many other candidate genes that can influence regional vulnerability, care should be taken to avoid creating overdetermined models. In summary, while the work here provides compelling evidence for the need for regional vulnerability, further work should seek to explain the mechanisms through which regional carrying capacities emerge from a culmination of regional risk factors, such as Aβ deposition and gene expression patterns.

There is extensive evidence of tau transport and production throughout the brain, however, it has not yet been determined whether one of these processes dominates the other and whether their

relative contributions to disease progression change over time. To this end, we sought to determine whether inferred parameters of our model change in groups across the disease timeline. We find that during the early stages of the disease, when there is a low concentration of tau in the medial temporal lobe, tau dynamics are transport-dominated but become production-dominated later in the disease. This supports previous work by Meisl et al. [27] who show through an analysis of multiple datasets and methods of tau quantification that tau dynamics are production dominated from Braak stage 3 onwards. This is consistent with our work, considering individuals who are positive on tau PET in early Braak stage regions may already show fairly advanced Braak stages at autopsy [56] analogous to individuals at middle Braak stages used by Meisl et al. 2021. These results suggest that, in early AD, tau seeds invade connected regions from the medial temporal lobe, but the overall concentration does not grow substantially. Only in later stages of AD is the spread production-dominated and driven by fast increases in concentration gradients, leading to progressive Braak-like staging. These results also support the utility of the logistic production model in being able to describe longitudinal A^+T^+ data (Fig. 2a & Fig. 2b), since seeds would already be densely present around the cortex and progression is driven by tau production. The results are also consistent with experimental evidence showing that tau seeds are present before tau pathology [20, 42].

Together the results indicate an intrinsically spatiotemporal process, with variations in both tau transport and production along the AD timeline. Our results suggest that there is an initial transport-dominated tauopathy that results in tau seeds spreading from the medial temporal lobe to axonally connected regions. This initial phase is followed by a secondary production-dominated tauopathy with accelerated regional accumulation and slower transport, capable of reproducing sequential Braak like staging such as in Fig. 1a. This contrasts to the largely temporal process of Aβ, described by [57], where initially Aβ is present throughout the brain, but increases concentration at different rates due to regional vulnerabilities. Our results suggest that the early period of AD during which tau is more easily transported between brain regions may be a critical time for intervention. Many immunotherapies currently being developed act on extracellular tau [58] and should therefore interrupt tau transmission through the extracellular space of the synaptic junction. If AD is a consequence of first tau spread and then tau production, it will be crucial that these immunotherapies are administered early in the AD process to halt the widespread transmission of tau before accelerated local production can occur. In contrast, therapies that act to reduce intracellular tau concentration should be effective in slowing AD progression throughout the AD continuum, regardless of whether widespread tau transmission has occurred [59].

This work presents a step forward in whole brain tau modelling, however, there are still many obstacles that are not addressed here. There are limitations that pertain to the sparsity of longitudinal data. In this work, we fix a number of parameters to ensure the practical identifiability of the models given the available data. In particular, we fix baseline values and carrying carrying in the dynamical system, and subject initial conditions in the probabilistic model. By fixing baseline values and carrying capacities, we are unable to determine whether these also undergo dynamical changes. This also limits the direct application of the model to other tauopathies that exhibit different tau PET profiles. A related limitation is the particular choice of parameterisation for regional vulnerability. Here, we have chosen to parameterise regional vulnerability as changes in regional carrying capacities scaled by a single free parameter. Future work should investigate whether more expressive models, such as regional free parameters or dynamically changing parameters influenced by other risk factors, such as amyloid, are more informative about longitudinal biomarker data. Second, there are limitations related to the scale at which we are modelling. While the model we present here is derived from a physics-based model, the model reduction comes at the cost of a loss in mechanistic insight into precise transport and production mechanisms. This will remain a hard limitation while we work with macroscale brain data. In addition, tau PET data is intrinsically limited by resolution, inability to detect early changes, and non-specific and off-target binding sources, collectively providing a source of uncertainty that affects parameter identifiability of intricate processes such as transport. Therefore, while the modelling results suggest changes to transport and production across the AD continuum, our conclusions are limited by the nature of PET measurements and require experimental validation. A potential avenue to address this limitation will be the development of multi-scale models that rely on in-vitro or animal studies for calibration and permit macro-scale reduced order models. Third, as increasing longitudinal scans are added to inference, the overall uncertainty decreases, as shown in Fig. 3 and Figs. S2 to S4, and does not always accurately capture out-of-sample regional data points. This likely results from the assumption of independent and identically distributed noise across subjects, scans and regions. As a result of the sequential staging of tau, there are likely to be more regions that have low signal and low longitudinal change and therefore the grouped noise parameter will be driven by these regions and underestimate noise in regions of high tau production. Future work should incorporate more specific noise models that account for changes in observation noise with the amount of signal, for example a multiplicative noise model that assumes that noise scales proportionally to signal. Finally, and perhaps most importantly, the model as presented here has limited application to real world clinical data, since most patients are not monitored with longitudinal tau PET and would therefore lack the necessary data with which to calibrate the model. Therefore, while the

model can provide useful mechanistic insights, it may not provide immediate clinical utility. We envisage two main applications of the current work. First, to provide motivation for experimental work into understanding changes in tau spreading in-vivo across the AD timeline. Second, direct applications to pharmaceutical trials, in which longitudinal tau PET is more readily available, to inform optimal intervention periods and testing patient trajectories against model predictions. The limitation on data availability may be eased by future work to incorporate fluid biomarker data into the modelling pipeline, as a less invasive and more readily available measure of tau burden. Fluid biomarkers can provide valuable information on tau production levels in-vivo [60, 61] that could be used to calibrate simplified spreading models such as the logistic model presented here, or similar asymptotic approximation to the local FKPP model that require only a production parameter [62].

The primary contribution of this work has been to provide a parsimonious account of regional tau dynamics in AD. Future work should seek to build on this, adding more information and data to probe the unexplained dynamics in AD. Most pressingly, these include dynamical interactions between Aβ and tau in a sufficiently simple way to accommodate the ability to perform inference with patient data. Furthermore, the study sheds light on potential avenues of clinical investigation of anti-tau therapies by showing how targeting different tau processes (transport or local production) at different times during the AD continuum may be essential for effective intervention.

4 Methods

4.1 Data Processing

We use PET from the Alzheimer's Disease Neuroimaging Initiative (`adni.loni.usc.edu`). ADNI is a public-private partnership with the aim of using serial biomarkers to measure the progression of AD. For up-to-date information, see `www.adni-info.org`. We download fully processed tau PET and magnetic resonance image (MRI) data, summarised as standardised uptake value ratios (SUVR) and volumes for each of the regions in the Desikan-Killiany-Tourville (DKT) atlas. We renormalise individual SUVR using an inferior cerebellum SUVR reference region. Amyloid status for ADNI subjects was also downloaded from ADNI and used to classify subjects, with Aβ positivity requiring cortical summary florbetapir SUVR > 0.78,using the composite reference region provided by ADNI, comprising the whole cerebellum, brain stem and eroded white matter.

We also use data from the Swedish BioFINDER-2 study (NCT03174938), which uses the RO948 tau PET radiotracer. All participants were recruited at Skåne University Hospital and the Hospital of Ängelholm, Sweden and the cohort covers the full spectrum of AD, ranging from cognitively normal individuals, patients with mild cognitive impairment (MCI) and with dementia. All details

about the cohort have been described previously [63]. Amyloid status was determined by amyloid-PET (flutemetamol) with the pons used as the reference region, based on a previously established cutoff from Gaussian mixture modelling as detailed in [60], with positivity requiring flutemetamol SUVR > 1.03. The tau data are analysed using the analysis pipeline detailed in [63]. Briefly, SUVR images were generated using the inferior cerebellum as a reference region, and average SUVR was extracted for regions in the DKT atlas. Four subjects in the BF2 A^+T^- group are removed due to high off-target binding in the skull/meninges or MRI registration problems. In both cohorts we select only subjects who have at least three tau PET scans to allow for inference on the time-series model. ADNI and BF2 tau PET data are summarised in Table 3.

	Group	**N Sub.**	**Age**	**Female**	**Edu.**	**CN**	**MCI**	**AD**	**Aβ Mean**	**Aβ s.d.**
ADNI	A^+T^+	57	72.61	0.54	16.37	0.28	0.53	0.19	82.26	34.01
	A^+T^-	37	76.23	0.49	16.49	0.70	0.32	0.03	49.15	38.42
	A^-T^-	53	71.76	0.48	16.48	0.57	0.42	0.03	0.14	13.43
BF2	A^+T^+	54	73.9	0.65	11.6	0.2	0.43	0.37	70.68	30.59
	A^+T^-	18	72.1	0.5	13.2	0.61	0.31	0.08	24.56	31.89
	A^-T^-	53	66.9	0.5	12.8	0.81	0.19	0.0	-7.31	5.98

Table 3: Demographics for ADNI and BF2 cohorts. Aβ mean and standard deviations are provided in centiloids.

We perform inference over three groups: A^-T^-, A^+T^-, A^+T^+ We distinguish between T^- and T^+ using a tau PET SUVR cut-off for two composite regions, as detailed in [34]. There are two cut-offs, one for determining tau positivity in the medial temporal lobe (MTL, defined as the mean of the bilateral entorhinal and amygdala), and another for neocortical positivity (defined as the middle temporal and inferior temporal gyri). The thresholds for the composite regions are based on regional Gaussian mixture models, as previously described [24]. For each composite, we average the SUVR values from the constituent regions and fit a two component Gaussian mixture model. The threshold for the region is then set to the SUVR at which there is a 50% chance of being T^+. For ADNI, the thresholds are 1.375 and 1.395 and for BF2 they are 1.248 and 1.451 for the MTL and cortical composites, respectively. We define a subject as being T^+ if their last scan is suprathreshold in either the MTL or cortical tau PET SUVR and T^- if the SUVR value is below both SUVR thresholds.

4.2 Mathematical Models of proteopathy

4.3 Structural Connectome Modelling

We use the structural connectome to model the transport of tau between brain regions. To generate structural connectomes, we use diffusion weighted MRI images of 150 healthy individuals from the Human Connectome Project (HCP) [65, 66]. From these data, connectomes are derived using the probabilistic tractography algorithm probtrackx [67], available in FSL, using 10000 samples per voxel, randomly sampled from a sphere around the voxel centre. The number of streamlines between each of R regions in the DKT atlas are summarised as an adjacency matrix, $\mathbf{A}$, that defines our connectome graph, G. To model transport of tau between regions, we use the graph Laplacian of G, given by:

$$\mathbf{L} = \mathbf{D} - \mathbf{A}, \tag{3}$$

where $\mathbf{D}$ is the degree matrix, $\mathbf{D} = \text{diag}(\mathbf{A} \cdot \mathbf{1})$. To ensure a transport process respects mass conservation across regions of varying volumes, we weight the graph Laplacian by regional volumes,

$$\mathcal{L} = \mathbf{V}^{-1}\mathbf{L} \tag{4}$$

where $\mathbf{V} = \text{diag}\,(\mathbf{v}/v_r)$, and $\mathbf{v} = (v_1, v_2, \dots v_R)$ is a vector of regional volumes and v_r is a reference region. In Section 2, we model three groups of subjects, A^-T^-, A^+T^- and A^+T^+, to reflect changes in volume across the disease timeline and variation in individual brain volumes, we define $\mathbf{v}$ and v_r on a group and individual level, respectively. For a given group with N subjects, we define the normalised volume of a region v_i as

$$v_i = \frac{1}{N}\sum_{n}^{N} \frac{v_i^n}{v_r^n}, \tag{5}$$

taking v_i^n as the initial volume of the ith region and nth subject and v_r^n is as the maximum initial regional volume for the nth subject. Then $\mathbf{v}$ is the average normalised volume per subject in a cohort.

The graph Laplacian is used in the next section to derive models of tau propagation on the brain network.

4.4 Local model of Tau Proliferation

We start with a coupled model of healthy and toxic protein, the heterodimer model, from which we aim to derive a simplified model of toxic protein dynamics that includes regional information. The

heterodimer model on a network is:

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} = -\rho \sum_{j=1}^{R} \mathcal{L}_{ij} p_j + k_0 - k_1 p_i - k_{12} p_i \tilde{p}_i, \qquad i = 1, \ldots, R, \tag{6a}$$

$$\frac{\mathrm{d}\tilde{p}_i}{\mathrm{d}t} = -\rho \sum_{j=1}^{R} \mathcal{L}_{ij} \tilde{p}_j \qquad - \tilde{k}_1 \tilde{p}_i + k_{12} p_i \tilde{p}_i, \qquad i = 1, \ldots, R, \tag{6b}$$

where p_i , $\tilde{p}$ are, respectively, the healthy and toxic protein concentration at node i, k_0 is the natural production rate of healthy protein, k_1 and $\tilde{k}_1$ are the clearance rates of healthy and toxic proteins, respectively, and k_{12} is the rate of conversion from healthy proteins into toxic proteins [36]. To simplify the heterodimer model, we can follow a similar procedure to that presented in [36], by linearising around a healthy state. Assuming an homogenous state with $\tilde{p}_i \ll p_i$, implies $\frac{\mathrm{d}p_i}{\mathrm{d}t} = 0$ and $-\sum_j^R \mathcal{L}_{ij} p_j = 0$ for $i = 1 \ldots R$. Then, linearising around $\tilde{p} = 0$, we have

$$p_i(\tilde{p}_i) \approx \frac{k_0}{k_1} \left(1 - \frac{k_{12}}{k_1} \tilde{p}_i\right).$$

Substituting this expression for p_i into equation (6b) we obtain,

$$\frac{\mathrm{d}\tilde{p}_i}{\mathrm{d}t} = -\rho \sum_{j=1}^{R} \mathcal{L}_{ij} \tilde{p}_j + \alpha \tilde{p}_i - \beta \tilde{p}_i{}^2, \qquad i = 1, \ldots, R, \tag{7}$$

where

$$\beta = \frac{k_0}{k_1} k_{12} - \tilde{k}_1 \qquad \text{and} \qquad \alpha = \frac{k_0 k_{12}^2}{k_1^2}. \tag{8}$$

From here we derive a model for tau PET SUVR, $s_i = s_i(t)$ for $i = 1, \ldots, R$, with regional carrying capacities and baseline values to model with the requirement that at node i, the healthy state corresponds to a baseline value of $s_i = s_{0,i}$ at $t = 0$ and the fully toxic state has asymptotic value $s_i = s_{\infty,i}$ as $t \to \pm\infty$. To accommodate regionally varying carrying capacities and a regionally uniform production rate, we asssume reginoally varying clearance, $\tilde{k}_1 \to \tilde{k}_{1,i}$ for $i = 1, \ldots, R$, making $\beta \to \beta_i$ regionally depedent. With this assumption, we can introduce the *local FKPP* model

$$\frac{\mathrm{d}s_i}{\mathrm{d}t} = -\rho \sum_{j=1}^{R} \mathcal{L}_{ij} \left(s_j - s_{0,j}\right) + \alpha \left(s_i - s_{0,i}\right) \left[\left(s_{\infty,i} - s_{0,i}\right) - \left(s_i - s_{0,i}\right)\right], \qquad i = 1, \ldots, R, \tag{9}$$

where $(s_i - s_{0,i})$ represents a shift to regionally dependent nonzero baseline SUVR values and $s_{\infty,i} = \beta_i \alpha^{-1} + s_{0,i}$ is the region carrying capacity at node i.

Then, the *logistic model* is then simply obtained by taking $\rho = 0$:

$$\frac{\mathrm{d}s_i}{\mathrm{d}t} = \alpha \left(s_i - s_{0,i}\right) \left[\left(s_{\infty,i} - s_{0,i}\right) - \left(s_i - s_{0,i}\right)\right], \qquad i = 1, \ldots, R, \tag{10}$$

where at each node the variable s_i connects asymptotically, for $t \to \pm\infty$, the healthy state $s_{0,i}$ to the toxic state $s_{\infty,i}$ (which implies there is no mechanism for propagation from node to node in this model). The diffusion model assumes $\alpha = 0$ in Eq. (9):

$$\frac{\mathrm{d}s_i}{\mathrm{d}t} = -\rho \sum_{j=1}^{R} \mathcal{L}_{ij} \left(s_j - s_{0,j}\right), \qquad i = 1, \ldots, R, \tag{11}$$

where there is no mechanism for production. The *global FKPP* model is taken by assuming regionally homogeneous baseline values and carrying capacities across all nodes:

$$\frac{\mathrm{d}s_i}{\mathrm{d}t} = -\rho \sum_{j=1}^{R} \mathcal{L}_{ij}(s_i - s_0) + \alpha(s_i - s_0)\left[(s_\infty - s_0) - (s_i - s_0)\right], \qquad i = 1, \ldots, R. \tag{12}$$

A method for determining the values of $\mathbf{s}_\infty$, $\mathbf{s_0}$, s_∞, and s_0 are provided in the next section.

4.5 Estimating fixed model parameters

To estimate the fixed parameters for $\mathbf{s}_0$ and $\mathbf{s}_\infty$, we fit a two component Gaussian mixture model to population level data of regional SUVR. For regions in which a reliable measure of tau SUVR can be obtained, we expect to see two separate distributions, a T^- distribution capturing the expected tau load in a given region, and a T^+ distribution describing the pathological tau load [24]. Using the fitted Gaussian mixture models, we approximate $s_{0,i}$ as the mean of the T^- distributions for the i-th region and $s_{\infty,i}$ as the 99-th percentile of the T^+ distributions for the i-th. These parameters are used to simulate Eqs. (9) to (11). To simulate from Eq. (12) we take s_∞ as $\max(\mathbf{s}_\infty)$ and s_0 as $\min(\mathbf{s_0})$. For subcortical regions it is not possible to obtain reliable tau PET signal due to off-target binding [68, 69] and we therefore exclude these regions from our model, leaving a total of 72 regions. For ADNI, we use the multi-tau cohort of AV1451 PET data, detailed in [24]. For BF2 data, we use all available RO948 PET scans. For all regions used, including the bilateral hippocampus and amygdala, a Gaussian mixture with two components provided a better fit to the data than a single component data, compared using the AIC score.

4.6 Probabilistic model

For each of the three groups, A^+T^+, A^+T^-, A^-T^-, we use a hierarchical model, factoring over patients and scans. In each group there are N subjects, each of whom have T_n scans, for $n = 1 \ldots N$ subjects, summarised over R regions, ($R = 72$). The observations times, i.e. scan dates, are denoted by $\mathbf{t} = t_j^n$ for $j = 1 \ldots T_n$, $n = 1 \ldots N$. We denote the full data set for a group as $\mathbf{Y}$ and individual

subject data as Y_{ij}^n, corresponding to the nth subject, at scan j and region i. For a single subject, we have the following data generating function:

$$\mathbf{Y}^n = \mathbf{f}(\mathbf{y}_0^n, \theta^n, \mathbf{t}^n) + \epsilon. \tag{13}$$

where $\mathbf{Y}^n$ is the individual data for R regions and T_n time points, with initial conditions $\mathbf{y}_0$, model parameters θ, and observations times $\mathbf{t}$. The data are generated by a dynamical systems, $\mathbf{f}$, with observation error ϵ. To derive a likelihood function from Eq. (13), we assume the observations errors are independently and identically distributed and sampled from a Gaussian distribution with standard deviation σ. The data generating distribution for a single observation from a subject is then:

$$\epsilon \sim \mathcal{N}(0, \sigma^2) \tag{14}$$

$$\mathbf{Y}^n \sim \mathcal{N}(\mathbf{f}(\mathbf{y}_0^n, \theta, \mathbf{t}^n), \sigma^2 \mathbf{I}) \tag{15}$$

To extend this to a hierarchical population model, we define random variables, $\Theta = \{(\rho_i, \alpha_i)\}_{i=1}^N$, encoding subject specific model parameters and hierarchical population parameters, $\Omega = \{\rho_\mu, \rho_\sigma, \alpha_\mu, \alpha_\sigma\}$, upon which each Θ_i depends. In Sections 2.2 and 2.4, we assume fixed initial conditions, $\mathbf{y}_0^n$, and observations times, $\mathbf{t}^n$, taken as the first tau PET scan and scan dates respectively. Then the likelihood function for a single subject under the hierarchical model is:

$$p(\mathbf{Y}^n, \Theta_n \mid \Omega, \sigma, \mathbf{y}_0^n, \mathbf{t}^n) = \prod_j^{T_n} \prod_i^R p(Y_{ij}^n \mid \Theta_n, \sigma, \mathbf{y}_0^n, t_j^n)\, p(\Theta_n \mid \Omega) \tag{16}$$

where the first term inside the product on the right hand side is the contribution of the subject level model and the second term is the hierarchical model. Then the posterior for all subjects, hierarchical parameters, subject specific parameters and observation noise is:

$$p(\Theta, \Omega, \sigma \mid \mathbf{Y}, \mathbf{t}, \mathbf{y}_0) \propto \prod_n^N p(\mathbf{Y}^n, \Theta_n \mid \Omega, \sigma, \mathbf{y}_0^n, \mathbf{t}^n)\, p(\Omega, \sigma). \tag{17}$$

4.7 Inference Algorithm

We run inference for each patient group separately using a Hamiltonian Monte Carlo No-U-Turn Sampler (NUTS) to sample from the group posterior distribution. We use the same priors across patient groups, provided in table 4. We use weakly informative priors based on scales at which we expect to observe parameter values and ensure the transport parameter is positive. The NUTS sampler is initialised with a unit diagonal Euclidean metric and a target acceptance ratio of 0.8.

For each patient group, we collected four chains each with 2000 samples. All chains showed good convergence (measured by $0.99 < \hat{r} < 1.01$) with no post warm-up numerical errors associated with the NUTS sampler.

Parameter	**Prior**	**Support**
ρ_μ	$Lognormal(0, 1)$	$[0, \infty]$
ρ_σ	$Lognormal(0, 1)$	$[0, \infty]$
α_μ	$\mathcal{N}(0, 1)$	$[\infty, \infty]$
α_σ	$Lognormal(0, 1)$	$[0, \infty]$
ρ_i	$\mathcal{N}(\rho_\mu, \rho_\sigma)$	$[0, \infty]$
α_i	$\mathcal{N}(\alpha_\mu, \alpha_\sigma)$	$[-\infty, \infty]$
σ	$Lognormal(0, 1)$	$[0, \infty]$

Table 4: Prior distributions for hierarchical model parameters.

4.8 Model Assessment

In Section 2.2 we use two metrics to compare a family of models, the Bayesian information criteria (BIC) and the expected log predictive density (ELPD). The BIC is used to compare the in-sample accuracy of each of the model's fit to the data using a NUTS sampler, and is given by:

$$BIC = k \log(n) - 2 \log L, \tag{18}$$

where

$$L = p(\mathbf{Y}, \Theta^\star \mid \Omega^\star, \sigma^\star, \mathbf{y}_0, \mathbf{t}) \tag{19}$$

is the total log-likelihood of all data, $\mathbf{Y}$, over subjects, scans and regions, calculated using the parameters $\Theta^\star$, $\Omega^\star$ and $\sigma^\star$. Parameters are chosen from the posterior samples collected during inference as those that maximise the log-likelihood. k is the number combined number of parameters between $\Theta^\star$, $\Omega^\star$ and $\sigma^\star$ ($k = 119$ for the local and global FKPP models; $k = 60$ for the diffusion and logistic models); $n = 13536$ is the total number of observations. .

The ELPD is used for estimating the out-of-sample predictive accuracy and is adapted from [70]. To do this, we use A^+T^+ subjects who have more than three tau PET scans (N = 10), using only the first three scans for training and remaining scans to measure predictive accuracy. For our model, the ELPD is then calculated as:

$$\text{ELPD} = \sum_i^N \left[\log \left(\frac{1}{S} \sum_s^S p\left(\mathbf{Y}^n \mid \Theta_s^n, \mathbf{y}_0^n, \mathbf{t}^n\right) \right) \right] \tag{20}$$

where $\mathbf{Y}^n$ are the unobserved data, $\Theta_s^n = \{(\rho, \alpha)\}_{s=1}^S$ are posterior samples of model parameters, $\mathbf{y}_0^n$ are subjects initial condition and $\mathbf{t}^n$ are scan dates, each for $n = 1 \dots N$ subjects.

The final method we use for model assessment is the comparison to shuffled data to examine whether the posterior distributions generated inferred from the true data are due to meaningful tau signal or statistical properties of the data. We perform this test on the A^+T^+ and A^+T^- groups by random spatial shuffling of the data. The same random permutation are applied to the regional baseline values and carrying capacities. The inference algorithm in Section 4.7 was then applied to the shuffled dataset and 1000 posterior samples were collected. This process was repeated 10 times for each group. In the A^+T^+ positive group, one chain failed to converge and was discarded.

Acknowledgement

The work of P. Chaggar was supported by funding from the Engineering and Physical Sciences Research Council grant EP/L016044/1 and F. Hoffmann-La Roche. Author J. Vogel was supported by the SciLifeLab and Wallenberg Data Driven Life Science Program (grant: KAW 2020.0239). A Pichet-Binette is supported by a postdoctoral fellowship from the Fonds de recherche en Santé Québec (298314). S. Jbabdi is supported by Wellcome SRF (221933/Z/20/Z) and Wellcome Collaborative Award (215573/Z/19/Z). G. Klein and S. Magon are supported by F. Hoffmann-La Roche. The BioFINDER study was supported by the Alzheimer's Association (SG-23-1061717), Swedish Research Council (2022-00775), ERA PerMed (ERAPERMED2021-184), the Knut and Alice Wallenberg foundation (2017-0383), the Strategic Research Area MultiPark (Multidisciplinary Research in Parkinson's disease) at Lund University, the Swedish Alzheimer Foundation (AF-980907), the Swedish Brain Foundation (FO2021-0293), The Parkinson foundation of Sweden (1412/22), the Cure Alzheimer's fund, the Konung Gustaf V:s och Drottning Victorias Frimurarestiftelse, the Skåne University Hospital Foundation (2020-O000028), Regionalt Forskningsstöd (2022-1259) and the Swedish federal government under the ALF agreement (2022-Projekt0080). The precursor of 18F-flutemetamol was sponsored by GE Healthcare. The precursor of 18F-RO948 was provided by Roche. The work of A. Goriely was supported by the Engineering and Physical Sciences Research Council grant EP/R020205/1. The funding sources had no role in the design and conduct of the study; in the collection, analysis, interpretation of the data; or in the preparation, review, or approval of the manuscript.

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute

on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research and Development, LLC.; Johnson and Johnson Pharmaceutical Research and Development LLC.; Lumosity; Lundbeck; Merck and Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Data were provided (in part) by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

For the purpose of Open Access, the authors will apply a CC BY public copyright license to any Author Accepted Manuscript (AAM) version arising from this submission.

Disclosures

OH has acquired research support (for the institution) from ADx, AVID Radiopharmaceuticals, Biogen, Eli Lilly, Eisai, Fujirebio, GE Healthcare, Pfizer, and Roche. In the past 2 years, he has received consultancy/speaker fees from AC Immune, Amylyx, Alzpath, BioArctic, Biogen, Cerveau, Eisai, Eli Lilly, Fujirebio, Merck, Novartis, Novo Nordisk, Roche, Sanofi and Siemens.

Data Availability

Pseudonymized data from the BioFINDER study can be shared by request from a qualified academic as long as data transfer is in agreement with EU legislation on the general data protection regulation and decisions by the Swedish Ethical Review Authority and Region Skåne, which should be regulated in a material transfer agreement.

Code Availability

Code for the analysis and visualisations produced in this manuscript are available at `https://github.com/PavanChaggar/local-fkpp`. Analysis was conducted using the Julia programming language [71] and we make particular use of DifferentialEquations.jl [72], Turing.jl [73] and Makie.jl [74].

Author Contribution

PC, JV, OH, TBT, AG conceived the research and the theory in discussion with all authors. JV, APB, OS, ES, OH provided the BIOFINDER the data. PC carried all simulations. PC,JV, APB, OS, ES, OH, AG analysed the data. All authors were involved in the discussion of the results and in drafting the manuscript.

References

[1] Lary C Walker and Mathias Jucker. Neurodegenerative diseases: expanding the prion concept. *Annual review of neuroscience*, 38:87–103, 2015.

[2] Michel Goedert, Masami Masuda-Suzukake, and Benjamin Falcon. Like prions: the propagation of aggregated tau and α-synuclein in neurodegeneration. *Brain*, 140(2):266–278, 2017.

[3] Oskar Hansson. Biomarkers for neurodegenerative diseases. *Nature Medicine*, 27:954 – 963, 2021.

[4] Dietmar R Thal, Udo Rüb, Mario Orantes, and Heiko Braak. Phases of aβ-deposition in the human brain and its relevance for the development of ad. *Neurology*, 58(12):1791–1800, 2002.

[5] Agneta Nordberg. Pet imaging of amyloid in Alzheimer's disease. *The lancet neurology*, 3(9):519–527, 2004.

[6] Azadeh Firouzian, Alex Whittington, Graham E Searle, Ivan Koychev, Giovanna Zamboni, Simon Lovestone, and Roger N Gunn. Imaging aβ and tau in early stage Alzheimer's disease with [18f] av45 and [18f] av1451. *EJNMMI research*, 8(1):1–9, 2018.

[7] Niklas Mattsson, Sebastian Palmqvist, Erik Stomrud, Jacob Vogel, and Oskar Hansson. Staging β-amyloid pathology with amyloid positron emission tomography. *JAMA neurology*, 76(11):1319–1329, 2019.

[8] Heiko Braak and Eva Braak. Neuropathological stageing of Alzheimer-related changes. *Acta neuropathologica*, 82(4):239–259, 1991.

[9] Michael Schöll, Samuel N. Lockhart, DanielvR. Schonhaut, James P. O'Neil, Mustafa Janabi, Rik Ossenkoppele, Suzanne L. Baker, Jacob W. Vogel, Jamie Faria, Henry D. Schwimmer, Gil D. Rabinovici, and William J. Jagust. Pet imaging of tau deposition in the aging human brain. *Neuron*, 89(5):971 – 982, 2016.

[10] Rik Ossenkoppele, Daniel R. Schonhaut, Michael Schöll, Samuel N. Lockhart, Nagehan Ayakta, Suzanne L. Baker, James P. O'Neil, Mustafa Janabi, Andreas Lazaris, Averill Cantwell, Jacob Vogel, Miguel Santos, Zachary A. Miller, Brianne M. Bettcher, Keith A. Vossel, Joel H. Kramer, Maria L. Gorno-Tempini, Bruce L. Miller, William J. Jagust, and Gil D. Rabinovici. Tau PET patterns mirror clinical and neuroanatomical variability in Alzheimer's disease. *Brain*, 139(5):1551–1567, 03 2016.

[11] Hanna Cho, Jae Yong Choi, Mi Song Hwang, You Jin Kim, Hye Mi Lee, Hye Sun Lee, Jae Hoon Lee, Young Hoon Ryu, Myung Sik Lee, and Chul Hyoung Lyoo. In vivo cortical spreading pattern of tau and amyloid in the Alzheimer disease spectrum. *Annals of neurology*, 80(2):247–258, 2016.

[12] Renaud La Joie, Adrienne V Visani, Suzanne L Baker, Jesse A Brown, Viktoriya Bourakova, Jungho Cha, Kiran Chaudhary, Lauren Edwards, Leonardo Iaccarino, Mustafa Janabi, et al. Prospective longitudinal atrophy in Alzheimer's disease correlates with the intensity and topography of baseline tau-pet. *Science translational medicine*, 12(524):eaau5732, 2020.

[13] Chenjie Xia, Sara J Makaretz, Christina Caso, Scott McGinnis, Stephen N Gomperts, Jorge Sepulcre, Teresa Gomez-Isla, Bradley T Hyman, Aaron Schultz, Neil Vasdev, et al. Association of in vivo [18f] av-1451 tau pet imaging results with cortical atrophy and symptoms in typical and atypical Alzheimer disease. *JAMA neurology*, 74(4):427–436, 2017.

[14] Mathias Jucker and Lary C Walker. Self-propagation of pathogenic protein aggregates in neurodegenerative diseases. *Nature*, 501(7465):45–51, 2013.

[15] Mathias Jucker and Lary C Walker. Propagation and spread of pathogenic protein assemblies in neurodegenerative diseases. *Nature neuroscience*, 21(10):1341, 2018.

[16] Florence Clavaguera, Tristan Bolmont, R Anthony Crowther, Dorothee Abramowski, Stephan Frank, Alphonse Probst, Graham Fraser, Anna K Stalder, Martin Beibel, Matthias Staufenbiel,

et al. Transmission and spreading of tauopathy in transgenic mouse brain. *Nature cell biology*, 11(7):909–913, 2009.

[17] Florence Clavaguera, Hiroyasu Akatsu, Graham Fraser, R Anthony Crowther, Stephan Frank, Jürgen Hench, Alphonse Probst, David T Winkler, Julia Reichwald, Matthias Staufenbiel, et al. Brain homogenates from human tauopathies induce tau inclusions in mouse brain. *Proceedings of the National Academy of Sciences*, 110(23):9535–9540, 2013.

[18] Li Liu, Valerie Drouet, Jessica W Wu, Menno P Witter, Scott A Small, Catherine Clelland, and Karen Duff. Trans-synaptic spread of tau pathology in vivo. *PloS one*, 7(2):e31302, 2012.

[19] Alix De Calignon, Manuela Polydoro, Marc Suárez-Calvet, Christopher William, David H Adamowicz, Kathy J Kopeikina, Rose Pitstick, Naruhiko Sahara, Karen H Ashe, George A Carlson, et al. Propagation of tau pathology in a model of early Alzheimer's disease. *Neuron*, 73(4):685–697, 2012.

[20] Sarah L DeVos, Bianca T Corjuc, Derek H Oakley, Chloe K Nobuhara, Riley N Bannon, Alison Chase, Caitlin Commins, Jose A Gonzalez, Patrick M Dooley, Matthew P Frosch, et al. Synaptic tau seeding precedes tau pathology in human Alzheimer's disease brain. *Frontiers in neuroscience*, 12:267, 2018.

[21] Ashish Raj, Amy Kuceyeski, and Michael Weiner. A network diffusion model of disease progression in dementia. *Neuron*, 73(6):1204–1215, 2012.

[22] Ashish Raj, Eve LoCastro, Amy Kuceyeski, Duygu Tosun, Norman Relkin, Michael Weiner, Alzheimer's Disease Neuroimaging Initiative (ADNI, et al. Network diffusion model of progression predicts longitudinal patterns of atrophy and metabolism in Alzheimer's disease. *Cell reports*, 10(3):359–369, 2015.

[23] Yasser Iturria-Medina, Roberto C Sotero, Paule J Toussaint, Alan C Evans, Alzheimer's Disease Neuroimaging Initiative, et al. Epidemic spreading model to characterize misfolded proteins propagation in aging and associated neurodegenerative disorders. *PLoS Comput Biol*, 10(11):e1003956, 2014.

[24] Jacob W Vogel, Yasser Iturria-Medina, Olof T Strandberg, Ruben Smith, Elizabeth Levitis, Alan C Evans, and Oskar Hansson. Spread of pathological tau proteins through communicating neurons in human Alzheimer's disease. *Nature Communications*, 11(1):1–15, 2020.

[25] Amelie Schäfer, Mathias Peirlinck, Kevin Linka, Ellen Kuhl, Alzheimer's Disease Neuroimaging Initiative (ADNI, et al. Bayesian physics-based modeling of tau propagation in Alzheimer's disease. *Frontiers in Physiology*, page 1081, 2021.

[26] Amelie Schäfer, Pavanjit Chaggar, Alain Goriely, and Ellen Kuhl. Correlating tau pathology to brain atrophy using a physics-based bayesian model. *Engineering with Computers*, pages 1–11, 2022.

[27] Georg Meisl, Eric Hidari, Kieren Allinson, Timothy Rittman, Sarah L DeVos, Justin S Sanchez, Catherine K Xu, Karen E Duff, Keith A Johnson, James B Rowe, et al. In vivo rate-determining steps of tau seed accumulation in Alzheimer's disease. *Science advances*, 7(44):eabh1448, 2021.

[28] S. Fornari, A. Schäfer, E. Kuhl, and A. Goriely. Spatially-extended nucleation-aggregation-fragmentation models for the dynamics of prion-like neurodegenerative protein-spreading in the brain and its connectome. *Journal of Theoretical Biology*, 486:110102, 2020.

[29] Dunja Mrdjen, Edward J Fox, Syed A Bukhari, Kathleen S Montine, Sean C Bendall, and Thomas J Montine. The basis of cellular and regional vulnerability in Alzheimer's disease. *Acta Neuropathologica*, 138:729–749, 2019.

[30] Hongjun Fu, John Hardy, and Karen E Duff. Selective vulnerability in neurodegenerative diseases. *Nature neuroscience*, 21(10):1350–1358, 2018.

[31] Eli J Cornblath, Howard L Li, Lakshmi Changolkar, Bin Zhang, Hannah J Brown, Ronald J Gathagan, Modupe F Olufemi, John Q Trojanowski, Danielle S Bassett, Virginia MY Lee, et al. Computational modeling of tau pathology spread reveals patterns of regional vulnerability and the impact of a genetic risk factor. *Science Advances*, 7(24):eabg6677, 2021.

[32] Amelie Schäfer, Pavanjit Chaggar, Travis B Thompson, Alain Goriely, Ellen Kuhl, Alzheimer's Disease Neuroimaging Initiative, et al. Predicting brain atrophy from tau pathology: A summary of clinical findings and their translation into personalized models. *Brain Multiphysics*, 2:100039, 2021.

[33] Clifford R Jack, David A Bennett, Kaj Blennow, Maria C Carrillo, Billy Dunn, Samantha Budd Haeberlein, David M Holtzman, William Jagust, Frank Jessen, Jason Karlawish, et al. Niaaa research framework: Toward a biological definition of Alzheimer's disease. *Alzheimer's & Dementia*, 14(4):535–562, 2018.

[34] Rik Ossenkoppele, Alexa Pichet Binette, Colin Groot, Ruben Smith, Olof Strandberg, Sebastian Palmqvist, Erik Stomrud, Pontus Tideman, Tomas Ohlsson, Jonas Jögi, et al. Amyloid and tau pet-positive cognitively unimpaired individuals are at high risk for future cognitive decline. *Nature Medicine*, pages 1–7, 2022.

[35] Johannes Weickenmeier, Ellen Kuhl, and Alain Goriely. The multiphysics of prion-like diseases: progression and atrophy. *Phys Rev Lett*, 121(158101), 2018.

[36] Sveva Fornari, Amelie Schäfer, Mathias Jucker, Alain Goriely, and Ellen Kuhl. Prion-like spreading of Alzheimer's disease within the brain's connectome. *Journal of the Royal Society Interface*, 16(159):20190356, 2019.

[37] Douglas N. Greve, David H. Salat, Spencer L. Bowen, David Izquierdo-Garcia, Aaron P. Schultz, Ciprian Catana, John A. Becker, Claus Svarer, Gitte Moos Knudsen, Reisa A. Sperling, Reisa A. Sperling, Keith A. Johnson, and Keith A. Johnson. Different partial volume correction methods lead to different conclusions: An 18f-fdg-pet study of aging. *NeuroImage*, 132:334–343, 2016.

[38] Suzanne L. Baker, Anne Maass, and William J. Jagust. Considerations and code for partial volume correcting [18f]-av-1451 tau pet data. *Data in Brief*, 15:648 – 657, 2017.

[39] Christina B. Young, Susan M. Landau, Theresa M. Harrison, Kathleen L. Poston, and Elizabeth C. Mormino. Influence of common reference regions on regional tau patterns in cross-sectional and longitudinal [18f]-av-1451 pet data. *NeuroImage*, 243:118553 – 118553, 2021.

[40] Alexander Schaefer, Ru Kong, Evan M. Gordon, Timothy O. Laumann, Xi-Nian Zuo, Avram J. Holmes, Simon B. Eickhoff, and B. T. Thomas Yeo. Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity mri. *bioRxiv*, 2017.

[41] Sara Larivière, Casey Paquola, Bo yong Park, Jessica Royer, Yezhou Wang, Oualid M. Benkarim, Reinder Vos de Wael, Sofie Louise Valk, Sophia I. Thomopoulos, Matthias Kirschner, Lindsay B. Lewis, Alan C. Evans, Sanjay M. Sisodiya, Carrie R. McDonald, Paul M. Thompson, and Boris C. Bernhardt. The enigma toolbox: multiscale neural contextualization of multisite neuroimaging datasets. *Nature Methods*, 18:698 – 700, 2021.

[42] Martí Colom-Cadena, Caitlin Davies, Sònia Sirisi, Ji-Eun Lee, Elizabeth M Simzer, Makis Tzioras, Marta Querol-Vilaseca, Érika Sánchez-Aced, Ya Yin Chang, Kristjan Holt, et al. Synaptic oligomeric tau in alzheimer's disease—a potential culprit in the spread of tau pathology through the brain. *Neuron*, 2023.

[43] Amelie Schäfer, Elizabeth C Mormino, and Ellen Kuhl. Network diffusion modeling explains longitudinal tau pet data. *Frontiers in Neuroscience*, 14:1370, 2020.

[44] Sebastian Palmqvist, Michael Schöll, Olof Strandberg, Niklas Mattsson, Erik Stomrud, Henrik Zetterberg, Kaj Blennow, Susan Landau, William Jagust, and Oskar Hansson. Earliest accumulation of β-amyloid occurs within the default-mode network and concurrently affects brain connectivity. *Nature communications*, 8(1):1214, 2017.

[45] Randy L Buckner, Abraham Z Snyder, Benjamin J Shannon, Gina LaRossa, Rimmon Sachs, Anthony F Fotenos, Yvette I Sheline, William E Klunk, Chester A Mathis, John C Morris, et al. Molecular, structural, and functional characterization of Alzheimer's disease: evidence for a relationship between default activity, amyloid, and memory. *Journal of neuroscience*, 25(34):7709–7717, 2005.

[46] Amy M. Pooler and Bradley T. et. al. Hyman. Amyloid accelerates tau propagation and toxicity in a model of early Alzheimer's disease. *Acta Neuropathologica Communications*, 3(1):14, 2015.

[47] Rachel E Bennett, Sarah L DeVos, Simon Dujardin, Bianca Corjuc, Rucha Gor, Jose Gonzalez, Allyson D Roe, Matthew P Frosch, Rose Pitstick, George A Carlson, et al. Enhanced tau aggregation in the presence of amyloid β. *The American journal of pathology*, 187(7):1601–1612, 2017.

[48] Zhuohao He, Jing L Guo, Jennifer D McBride, Sneha Narasimhan, Hyesung Kim, Lakshmi Changolkar, Bin Zhang, Ronald J Gathagan, Cuiyong Yue, Christopher Dengler, et al. Amyloid-β plaques enhance Alzheimer's brain tau-seeded pathologies by facilitating neuritic plaque tau aggregation. *Nature medicine*, 24(1):29, 2018.

[49] Amy M Pooler, Emma C Phillips, Dawn HW Lau, Wendy Noble, and Diane P Hanger. Physiological release of endogenous tau is stimulated by neuronal activity. *EMBO reports*, 14(4):389–394, 2013.

[50] Jessica W Wu, S Abid Hussaini, Isle M Bastille, Gustavo A Rodriguez, Ana Mrejeru, Kelly Rilett, David W Sanders, Casey Cook, Hongjun Fu, Rick ACM Boonen, et al. Neuronal activity enhances tau propagation and tau pathology in vivo. *Nature neuroscience*, 19(8):1085–1092, 2016.

[51] Travis B. Thompson, Pavanjit Chaggar, Ellen Kuhl, Alain Goriely, and for the Alzheimer's Disease Neuroimaging Initiative. Protein-protein interactions in neurodegenerative diseases: A conspiracy theory. *PLOS Computational Biology*, 16(10):1–41, 10 2020.

[52] Ying-Qiu Zheng, Yu Zhang, Yvonne HC Yau, Yashar Zeighami, Kevin Larcher, Bratislav Misic, and Alain Dagher. Connectome architecture, gene expression and functional co-activation shape the propagation of misfolded proteins in neurodegenerative disease. *bioRxiv*, page 449199, 2018.

[53] MGSM Goedert, Maria Grazia Spillantini, R Jakes, D Rutherford, and RA Crowther. Multiple isoforms of human microtubule-associated protein tau: sequences and localization in neurofibrillary tangles of Alzheimer's disease. *Neuron*, 3(4):519–526, 1989.

[54] Aylin Dincer, Charles D Chen, Nicole S McKay, Lauren N Koenig, Austin McCullough, Shaney Flores, Sarah J Keefe, Stephanie A Schultz, Rebecca L Feldman, Nelly Joseph-Mathurin, et al. Apoe $\varepsilon 4$ genotype, amyloid-β, and sex interact to predict tau in regions of high apoe mrna expression. *Science translational medicine*, 14(671):eabl7646, 2022.

[55] Jorge Sepulcre, Michel J Grothe, Federico d'Oleire Uquillas, Laura Ortiz-Terán, Ibai Diez, Hyun-Sik Yang, Heidi IL Jacobs, Bernard J Hanseeuw, Quanzheng Li, Georges El-Fakhri, et al. Neurogenetic contributions to amyloid beta and tau spreading in the human cortex. *Nature medicine*, 24(12):1910–1918, 2018.

[56] Vikas Kotari, Sudeepti Southekal, Michael Navitsky, Ian A Kennedy, Ming Lu, Amanda Morris, Jennifer Ann Zimmer, Adam S Fleisher, Mark A Mintun, Michael D Devous Sr, et al. Early tau detection in flortaucipir images: validation in autopsy-confirmed data and implications for disease progression. *Alzheimer's Research & Therapy*, 15(1):41, 2023.

[57] Alex Whittington, David J Sharp, and Roger N Gunn. Spatiotemporal distribution of β-amyloid in Alzheimer disease is the result of heterogeneous regional carrying capacities. *Journal of Nuclear Medicine*, 59(5):822–827, 2018.

[58] LA Sandusky-Beltran and EM Sigurdsson. Tau immunotherapies: Lessons learned, current status and future considerations. *Neuropharmacology*, 175:108104, 2020.

[59] Catherine J Mummery, Anne Börjesson-Hanson, Daniel J Blackburn, Everard GB Vijverberg, Peter Paul De Deyn, Simon Ducharme, Michael Jonsson, Anja Schneider, Juha O Rinne, Albert C Ludolph, et al. Tau-targeting antisense oligonucleotide maptrx in mild alzheimer's disease: a phase 1b, randomized, placebo-controlled trial. *Nature Medicine*, pages 1–11, 2023.

[60] Alexa Pichet Binette, Nicolai Franzmeier, Nicola Spotorno, Michael Ewers, Matthias Brendel, Davina Biel, Olof Strandberg, Shorena Janelidze, Sebastian Palmqvist, et al. Amyloid-associated increases in soluble tau relate to tau aggregation rates and cognitive decline in early alzheimer's disease. *Nature Communications*, 13(1):6635, 2022.

[61] Linda Karlsson, Jacob Vogel, Ida Arvidsson, Kalle Åström, Olof Strandberg, Jakob Seidlitz, Richard AI Bethlehem, Erik Stomrud, Rik Ossenkoppele, Nicholas J Ashton, et al. A machine learning-based prediction of tau load and distribution in alzheimer's disease using plasma, mri and clinical variables. *medRxiv*, pages 2024–05, 2024.

[62] Prama Putra, Hadrien Oliveri, Travis Thompson, and Alain Goriely. Front propagation and arrival times in networks with application to neurodegenerative diseases. *SIAM Journal on Applied Mathematics*, 83(1):194–224, 2023.

[63] Antoine Leuzy, Ruben Smith, Rik Ossenkoppele, Alexander Santillo, Edilio Borroni, Gregory Klein, Tomas Ohlsson, Jonas Jögi, Sebastian Palmqvist, Niklas Mattsson-Carlgren, et al. Diagnostic performance of ro948 f 18 tau positron emission tomography in the differentiation of Alzheimer disease from other neurodegenerative disorders. *JAMA neurology*, 77(8):955–965, 2020.

[64] Sebastian Palmqvist, Shorena Janelidze, Yakeel T Quiroz, Henrik Zetterberg, Francisco Lopera, Erik Stomrud, YI Su, Yinghua Chen, Geidy E Serrano, Antoine Leuzy, et al. Discriminative accuracy of plasma phospho-tau217 for alzheimer disease vs other neurodegenerative disorders. *Jama*, 324(8):772–781, 2020.

[65] David C. Van Essen, Kâmil Uğurbil, Edward J. Auerbach, Deanna M. Barch, Timothy Edward John Behrens, Richard D. Bucholz, Acer Y-C Chang, Liyong Chen, Maurizio Corbetta, Sandra W. Curtiss, Stefania Della Penna, David A. Feinberg, Matthew F. Glasser, Noam Y. Harel, A. C. Heath, Linda J. Larson-Prior, Daniel S. Marcus, G. Michalareas, Steen Moeller, Robert Oostenveld, Steven E. Petersen, Fred W. Prior, Bradley L. Schlaggar, Stephen M. Smith, Abraham Z. Snyder, Junqian Xu, and Essa Yacoub. The human connectome project: A data acquisition perspective. *NeuroImage*, 62:2222–2231, 2012.

[66] Stamatios N. Sotiropoulos, Saâd Jbabdi, Junqian Xu, Jesper L. R. Andersson, Steen Moeller, Edward J. Auerbach, Matthew F. Glasser, Moisés Hernández, Guillermo Sapiro, Mark Jenkinson, David A. Feinberg, Essa Yacoub, Christophe Lenglet, David C. Van Essen, Kâmil Uğurbil, and Timothy Edward John Behrens. Advances in diffusion mri acquisition and processing in the human connectome project. *NeuroImage*, 80:125–143, 2013.

[67] Timothy EJ Behrens, H Johansen Berg, Saad Jbabdi, Matthew FS Rushworth, and Mark W Woolrich. Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *Neuroimage*, 34(1):144–155, 2007.

[68] Jae Yong Choi, Hanna Cho, Sung Jun Ahn, Jae Hoon Lee, Young Hoon Ryu, Myung Sik Lee, and Chul Hyoung Lyoo. Off-target 18f-av-1451 binding in the basal ganglia correlates with age-related iron accumulation. *Journal of Nuclear Medicine*, 59(1):117–120, 2018.

[69] Ruben Smith, Michael Schöll, Antoine Leuzy, Jonas Jögi, Tomas Ohlsson, Olof Strandberg, and Oskar Hansson. Head-to-head comparison of tau positron emission tomography tracers [18f] flortaucipir and [18f] ro948. *European journal of nuclear medicine and molecular imaging*, 47(2):342–354, 2020.

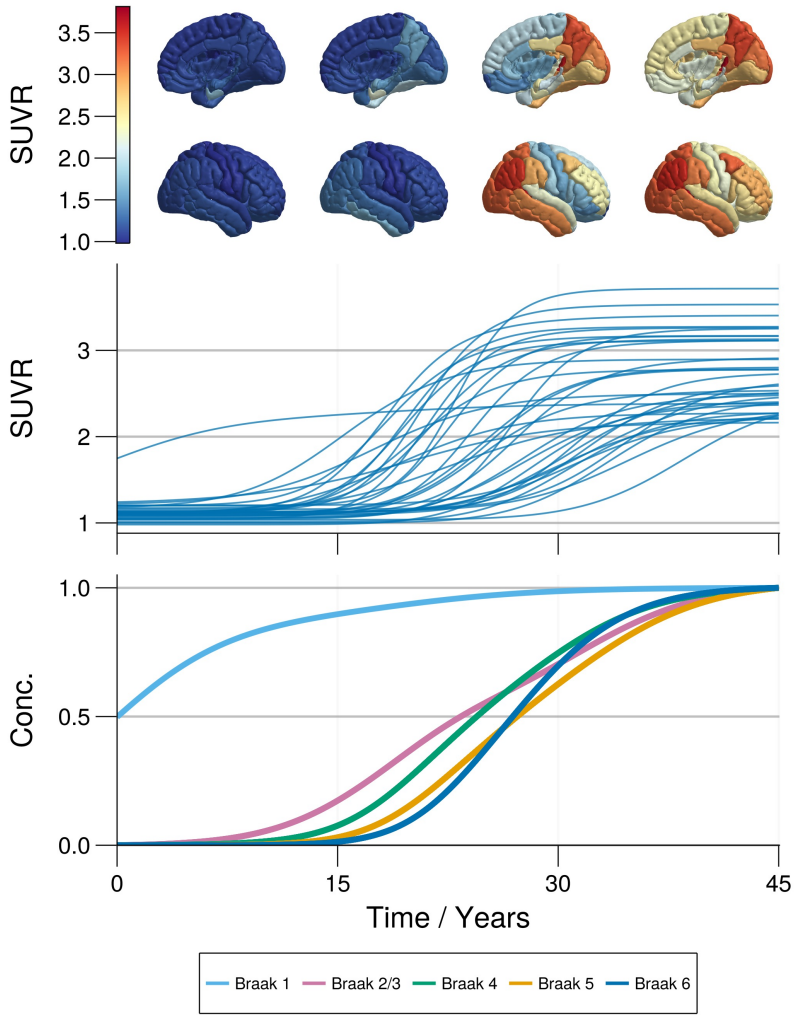
[70] Andrew Gelman, Jessica Hwang, and Aki Vehtari. Understanding predictive information criteria for bayesian models. *Statistics and computing*, 24(6):997–1016, 2014.

[71] Jeff Bezanson, Stefan Karpinski, Viral B Shah, and Alan Edelman. Julia: A fast dynamic language for technical computing. *arXiv preprint arXiv:1209.5145*, 2012.

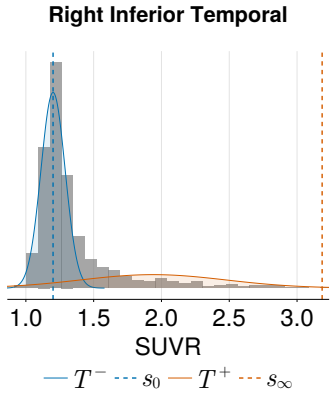
[72] Christopher Rackauckas and Qing Nie. Differentialequations. jl–a performant and feature-rich ecosystem for solving differential equations in julia. *Journal of Open Research Software*, 5(1), 2017.

[73] Mohamed Tarek, Kai Xu, Martin Trapp, Hong Ge, and Zoubin Ghahramani. Dynamicppl: Stan-like speed for dynamic probabilistic models. *arXiv preprint arXiv:2002.02702*, 2020.

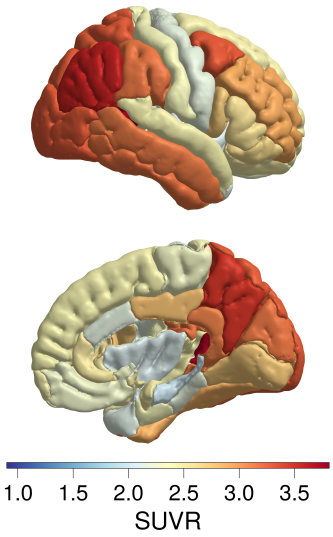
[74] Simon Danisch and Julius Krumbiegel. Makie.jl: Flexible high-performance data visualization for julia. *Journal of Open Source Software*, 6(65):3349, 2021.



(a)

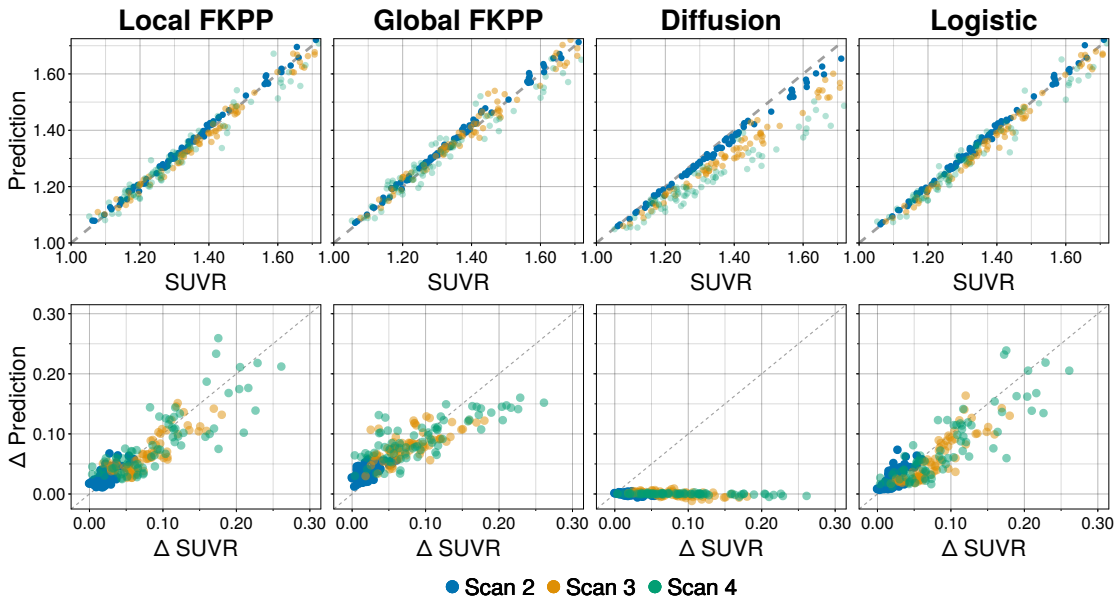


(b)

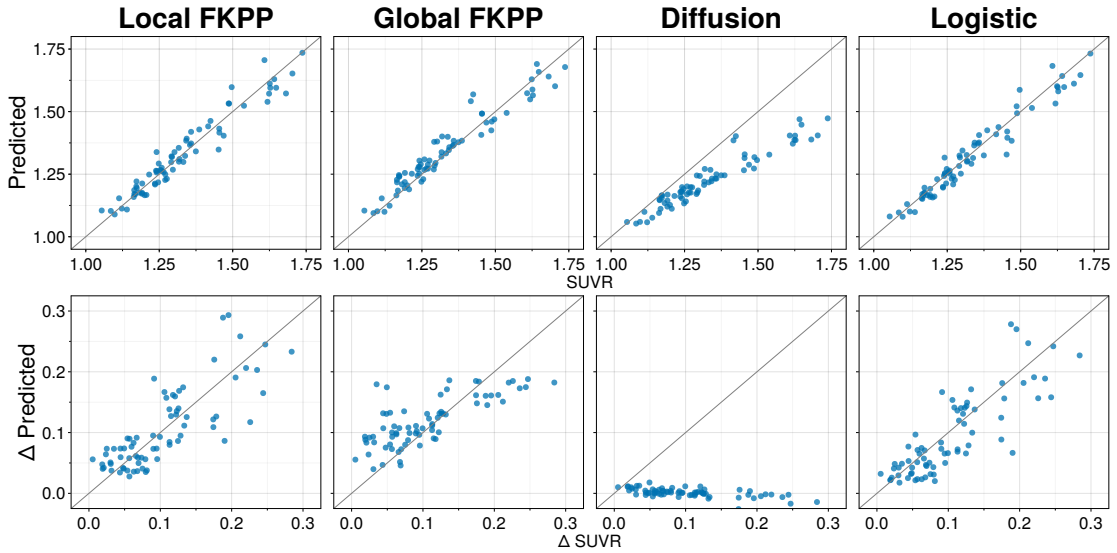


(c)

Figure 1: **Simulated transport and production dynamics in the local FKPP model. 1a** Simulation from the local FKPP model using carrying capacities derived from Gaussian mixture models (shown in 1b). Simulations are initialised with a seed value of $(s_{0,i} + s_{\infty,i})/2$ in the bilateral entorhinal cortex, $i = \{27, 63\}$, with $\rho = 0.025/yr$ and $\alpha = 0.25/yr$. Each line in the middle panel represents the SUVR trajectory of one DKT atlas brain region. Values at time points $t = \{0, 15, 30, 45\}$ years are projected onto a cortical rendering in the top panel. Each line in the bottom panel represents concentration averaged over Braak regions, after rescaling simulated SUVR as $q_i = (s_i - s_{0,i}) \,/\, (s_{\infty,i} - s_{0,i})$. **1b** Two component Gaussian mixture model fit to a multi-cohort tau PET dataset [24] and data from ADNI for inferior temporal lobe. Baseline values for each region are taken as the mean of regional T^- distribution and the carrying capacity as the 99-th percentile of the regional T^+ distribution, and these are used to simulate the model in 1a and throughout this paper. **1c** Right hemisphere cortical rendering of the SUVR carrying capacities as determined through Gaussian mixture modelling.



(a) In-sample model fits.



(b) Out-of-sample model fits.

Figure 2: **Model Fit for In-sample and Out-of-Sample Data. 2a** Goodness of fit for four models, local FKPP, global FKPP, diffusion and logistic. For all panels, each point represents a region in the connectome model, averaged over subjects per scan. Top row shows estimated vs observed SUVR values. Bottom row shows estimated change vs observed change in SUVR. **2b** Out-of-sample fits for four models of proteopathy. Top row: predicted vs observed out-of-sample SUVR. Bottom row: predicted change vs observed change from first in-sample scan to last out-of-sample scan. Each point represents a region in the connectome model, averaged over subjects. Only the local FKPP and logistic model are able to accurately capture longitudinal changes, while the global FKPP and diffusion models provide poor forecasts of heterogeneous tau production.

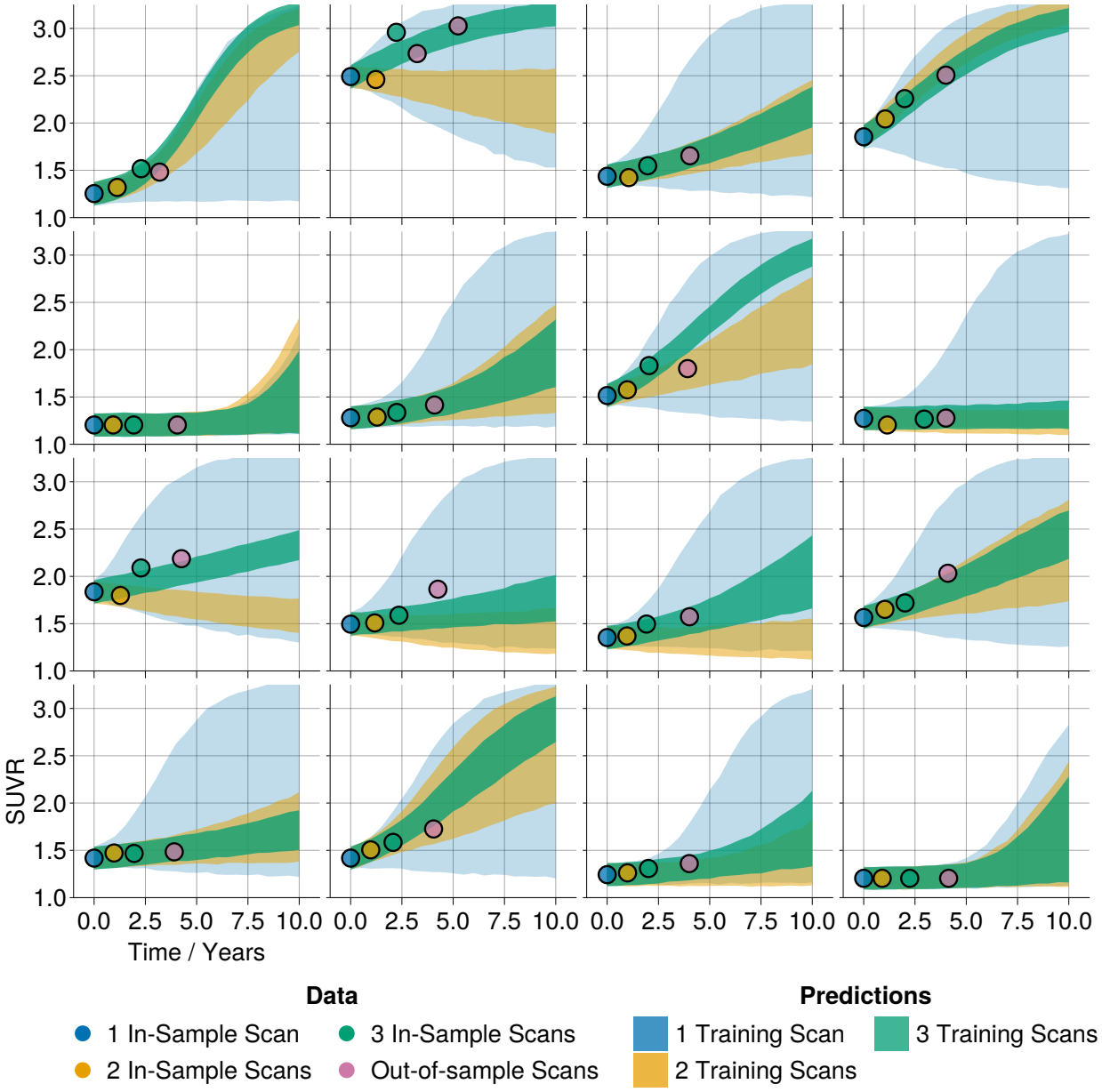


Figure 3: **Out-of-sample fit and posterior predictive plots for the left inferior temporal cortex.** The local FKPP model was iteratively calibrated to a A^+T^+ ADNI cohort with 41 in-sample subjects and 16 test subjects. Three iteration were run where for each iteration an additional scan from the test subjects were included, starting with a single scan. Posterior predictive trajectories for left inferior temporal lobe are shown for each iteration. In the above figure, each panel represents each of the 16 test subjects. Each point represents a data point added for training iteration; trajectories are color matched to correspond to the number of longitudinal data points included for training.

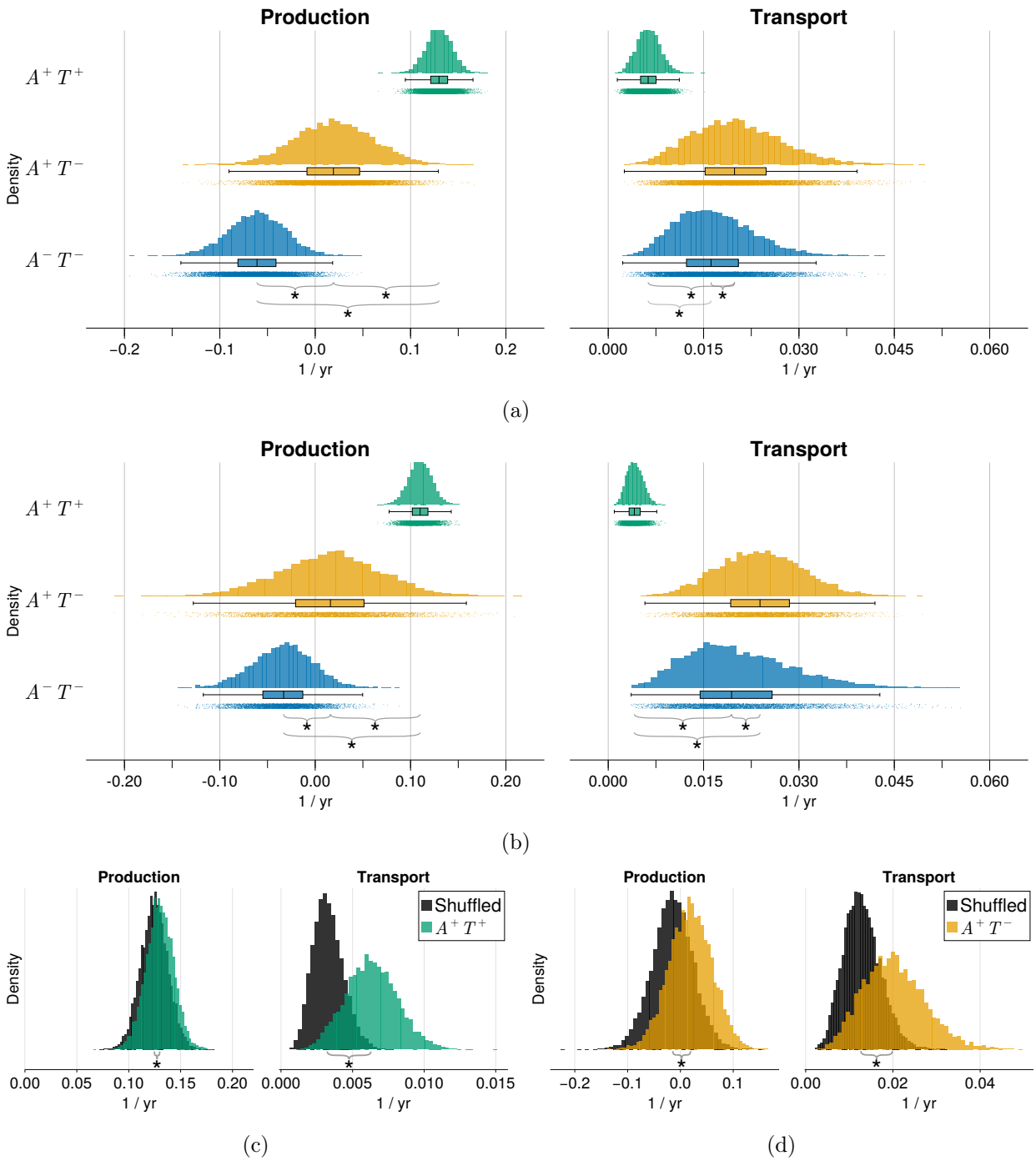


Figure 4: **Inferred population level parameters using ADNI and BF2 tau data**. **4a & 4b** Population production and transport parameters across A^+T^+, A^+T^- and A^-T^- groups for ADNI (4a) and BF2 (4b) tau PET data. **4c & 4d** Inferred population production and transport parameters from spatially shuffled data (shown in grey) compare to inferred distributions from true data for A^+T^+ (4c), A^+T^- (4d) ADNI groups. Asterisks between groups denotes distributions are significantly different ($p < 0.01$), tested using the Mann-Whitney U test.

S1 Supplementary Information

S1.1 Residual Analysis Of Local Fkpp And Logistic Models

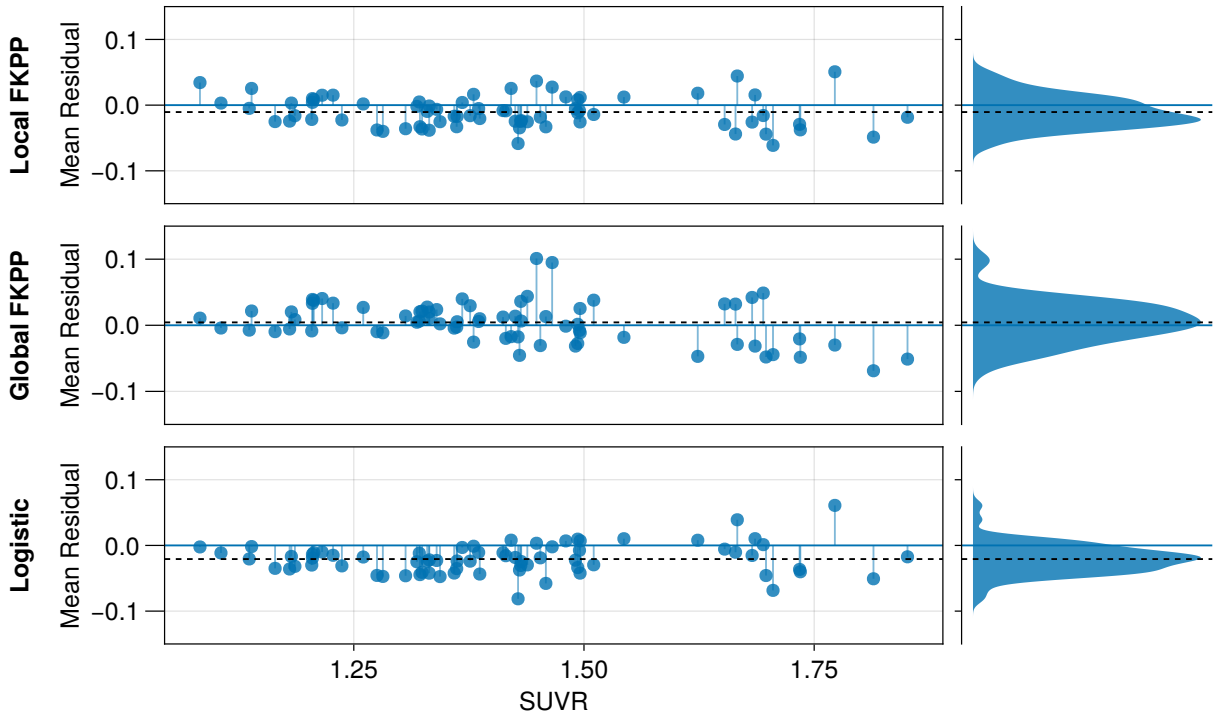


Figure S1: **Residual analysis for in-sample model fit.** Regional residual averaged over subjects in the A^+T^+ cohort, showing the regionally averaged difference between the SUVR of final scans and the corresponding prediction using the local FKPP, global FKPP and logistic models. Black dashed line represents the mean error averaged over subjects and regions. Blue solid line highlights zero error.

S1.2 Regional Forecasting of Tau SUVR

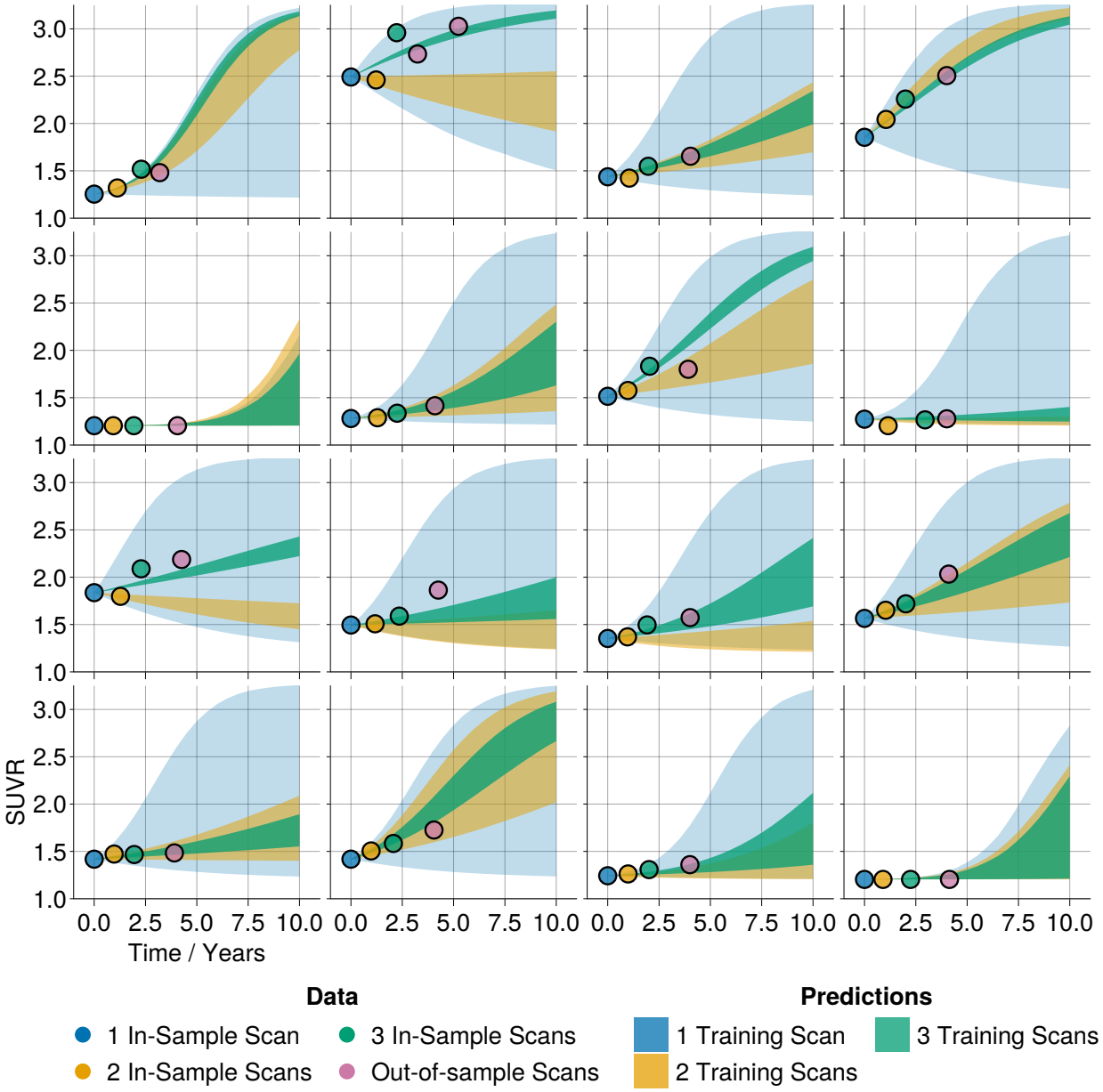


Figure S2: **Out-of-sample fit and posterior predictive plots for the left inferior temporal region.** The local FKPP model was iteratively calibrated to a A^+T^+ ADNI cohort with 41 in-sample subjects and 16 test subjects. Three iteration were run where for each iteration an additional scan from the test subjects were included, starting with a single scan. Posterior predictive trajectories for left inferior temporal lobe are shown for each iteration (neglecting observation noise). In the above figure, each panel represents one of the 16 test subjects. Each point represents a data point added for a training iteration; trajectories are color matched to correspond to the number of longitudinal data points included for training.

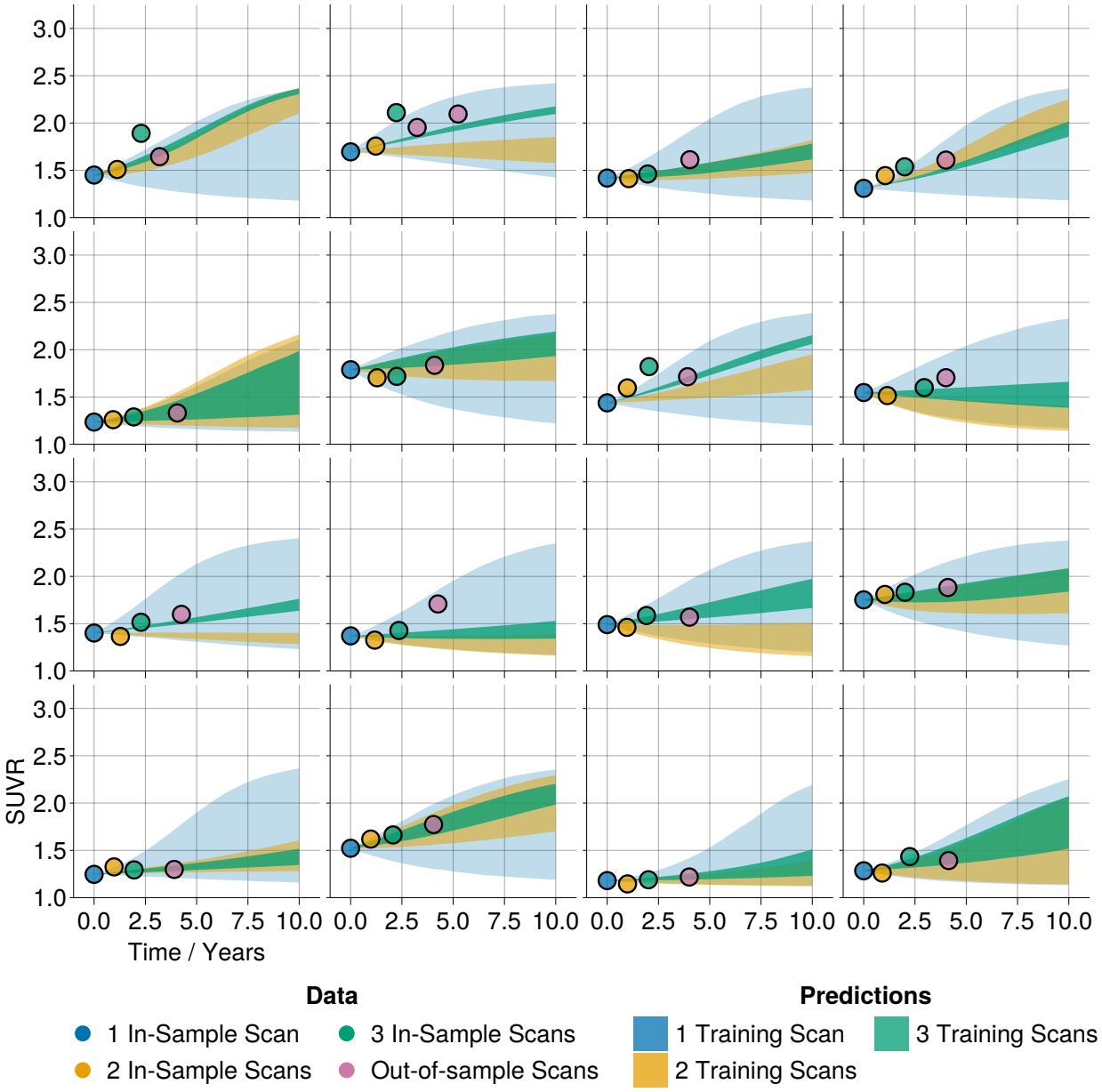


Figure S3: **Out-of-sample fit and posterior predictive plots for the left entorhinal cortex.** The local FKPP model was iteratively calibrated to a A^+T^+ ADNI cohort with 41 in-sample subjects and 16 test subjects. Three iteration were run where for each iteration an additional scan from the test subjects were included, starting with a single scan. Posterior predictive trajectories for left entorhinal cortex are shown for each iteration. In the above figure, each panel represents each of the 16 test subjects. Each point represents a data point added for training iteration; trajectories are color matched to correspond to the number of longitudinal data points included for training.

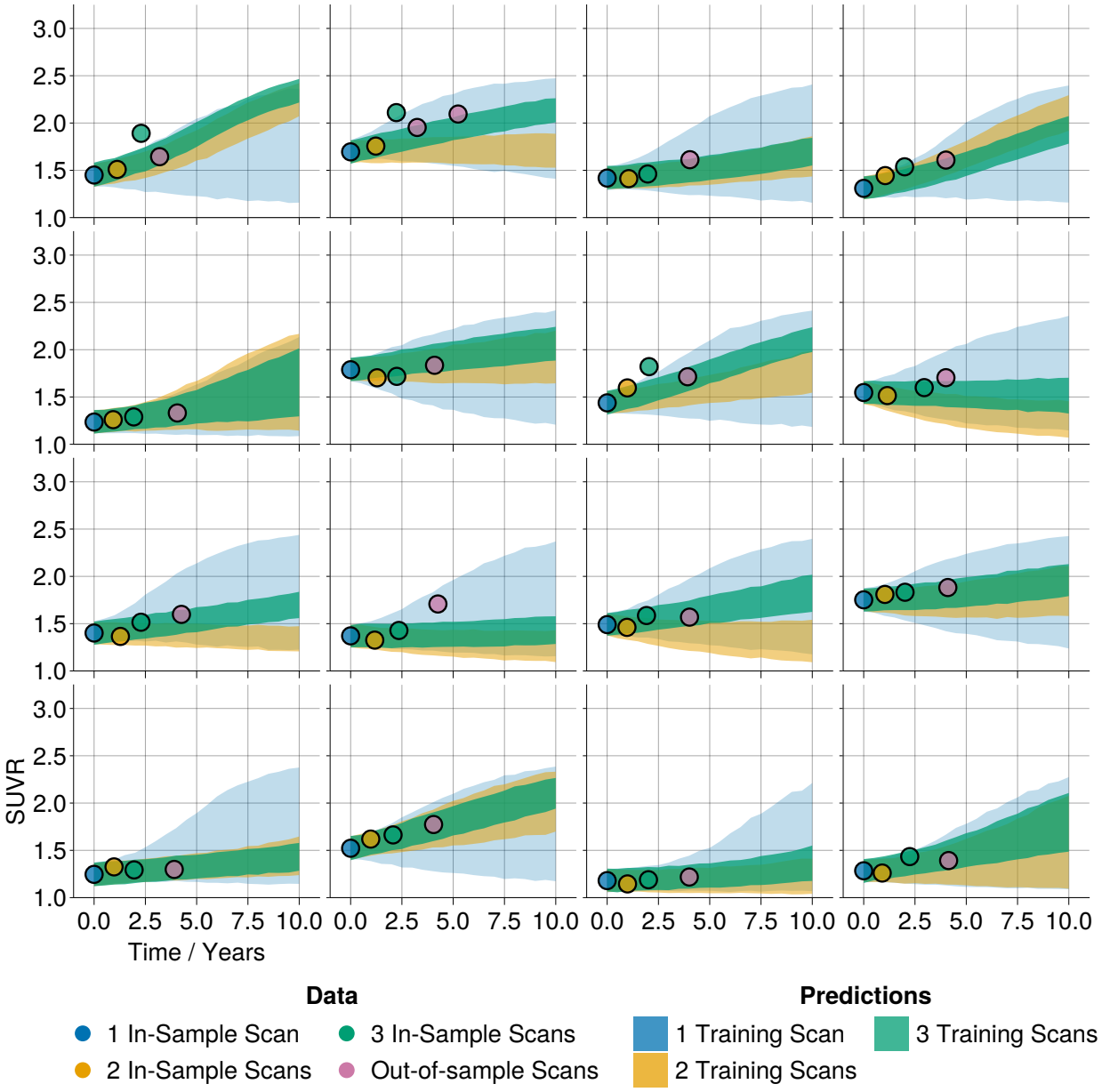


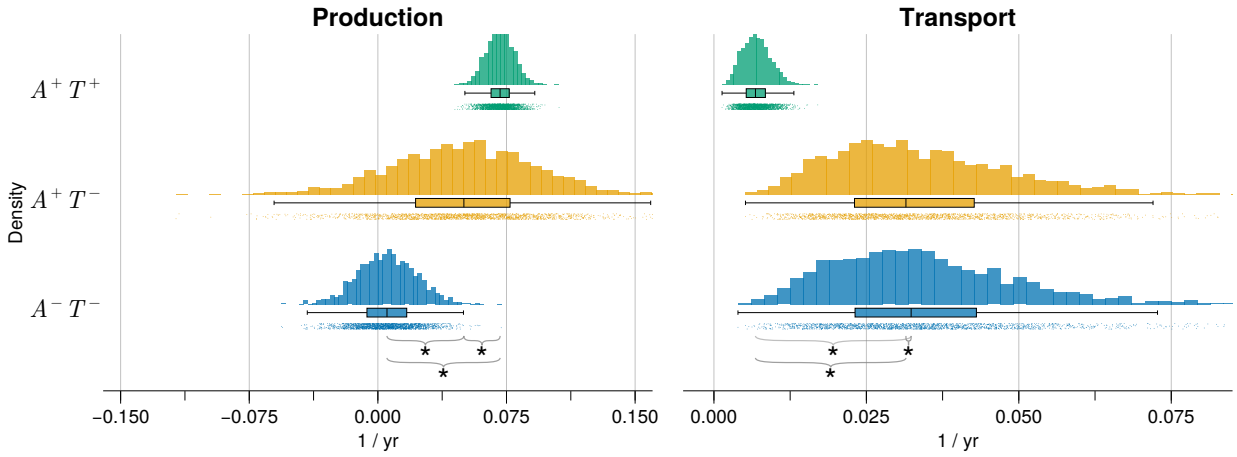
Figure S4: **Out-of-sample fit and posterior predictive plots for the left entorhinal cortex with the addition of observation noise.** The local FKPP model was iteratively calibrated to a A^+T^+ ADNI cohort with 41 in-sample subjects and 16 test subjects. Three iteration were run where for each iteration an additional scan from the test subjects were included, starting with a single scan. Posterior predictive trajectories for left entorhinal cortex are shown for each iteration. In the above figure, each panel represents each of the 16 test subjects. Each point represents a data point added for training iteration; trajectories are color matched to correspond to the number of longitudinal data points included for training.

S1.3 Effects Of Data Processing Choices On Longitudinal Modelling

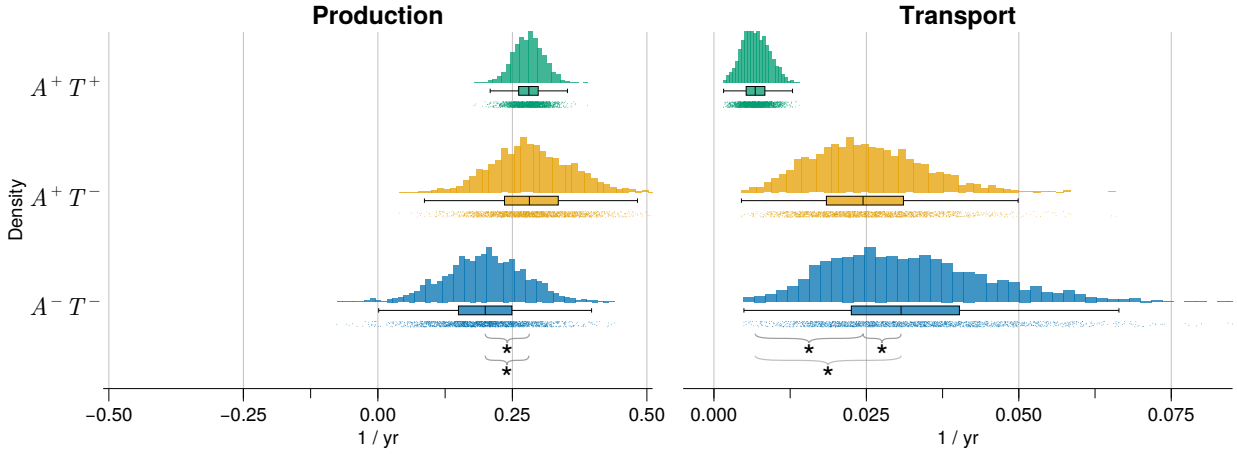
We reapplied the local FKPP to ADNI data (using methods in Section 4.7) with two different processing options: (1) partial volume correction and (2) an eroded white matter reference region. The posterior distributions of this analysis are shown in Fig. S5.

In an effort to explain difference between the changes in the production rate across AD groups with the eroded white matter, we conducted a further analysis of the longitudinal behviour of both the eroded white matter and inferior cerebellar reference region. To perform this analysis, we use preprocessed data from ADNI, for which SUVR values are provided using the inferior cerebellum reference for both target regions and alternative reference regions, including the eroded white matter. In Fig. S6a, we show the longitudinal change in SUVR for regions in the DKT atlas, averaged over subjects, using both the inferior cerebellum and the eroded white matter reference regions. The use of the eroded white matter reference region eliminates the negative longitudinal changes observed with the inferior cerebellar reference region for the A^+T^- and A^-T^- groups, consistent with the results of model inference in Fig. S5. We hypothesized that this is due to atrophy-related effects in the reference region that result in decreasing reference region SUVR. To test this, we conduct ed further analysis into the longitudinal behaviour of the different reference regions. First, we examine longitudinal volume changes in the reference regions for each subject from baseline scans to final scans, shown in Fig. S6b. The results show that the volume of the inferior cerebellar reference region is more stable longitudinally than the eroded white matter, which shows greater variability and is more prone to atrophy. Next, we examined longitudinal changes in reference region SUVR from baseline to final scan. Since ADNI data are already processed using the inferior cerebellum reference region, we use the cerebellar gray as a proxy for inferior cerebellar signal. In Fig. S6c we compare the subject-wise longitudinal change in SUVR of the eroded white matter and cerebellar gray to longitudinal change in medial temporal SUVR (comprising the bilateral entorhinal cortex and amygdala), where the SUVR is for all regions used is calculated using the inferior cerebellum, as provided by ADNI. We see that there is a strong positive correlation between longitudinal eroded white matter SUVR change and MTL SUVR and that decreasing MTL SUVR is associated with decreases in reference region SUVR. This effect is not observed with the the cerebellar gray region, which shows negligible longitudinal change. Due to this positive correlation, additional normalisation with the eroded white matter reference region would diminish any longitudinal changes in target regions, such as we see in Fig. S6a. In sum, these analyses show that the eroded white matter has less stable longitudinal SUVR and volume compared to the inferior cerebellum and its longitudinal SUVR is positively correlated with MTL SUVR, likely resulting in the diminished production

effect observed in the inferred posterior distributions in Fig. S5b.



(a)



(b)

Figure S5: **Inferred posterior distributions for PVC and EWM ADNI data. S5a**. Posterior distributions from A^+T^+, A^+T^- and A^-T^- groups for ADNI data processed using an inferior cerebellar reference region and with partial volume correction. **S5b**. Posterior distributions from A^+T^+, A^+T^- and A^-T^- groups using ADNI data processed using the eroded white matter reference region. Posterior distributions were obtained using Eq. (17) and an identical inference set up to those presented in Section 4.7. We collected 2000 posterior samples from each gropu; there were no divergences and samples shows good effective sample size and $0.99 \leq \hat{r} \leq 1.01$, indicating good convergence. Significance between distributions is given by the Mann-Whitney U test with $P < 0.01$.

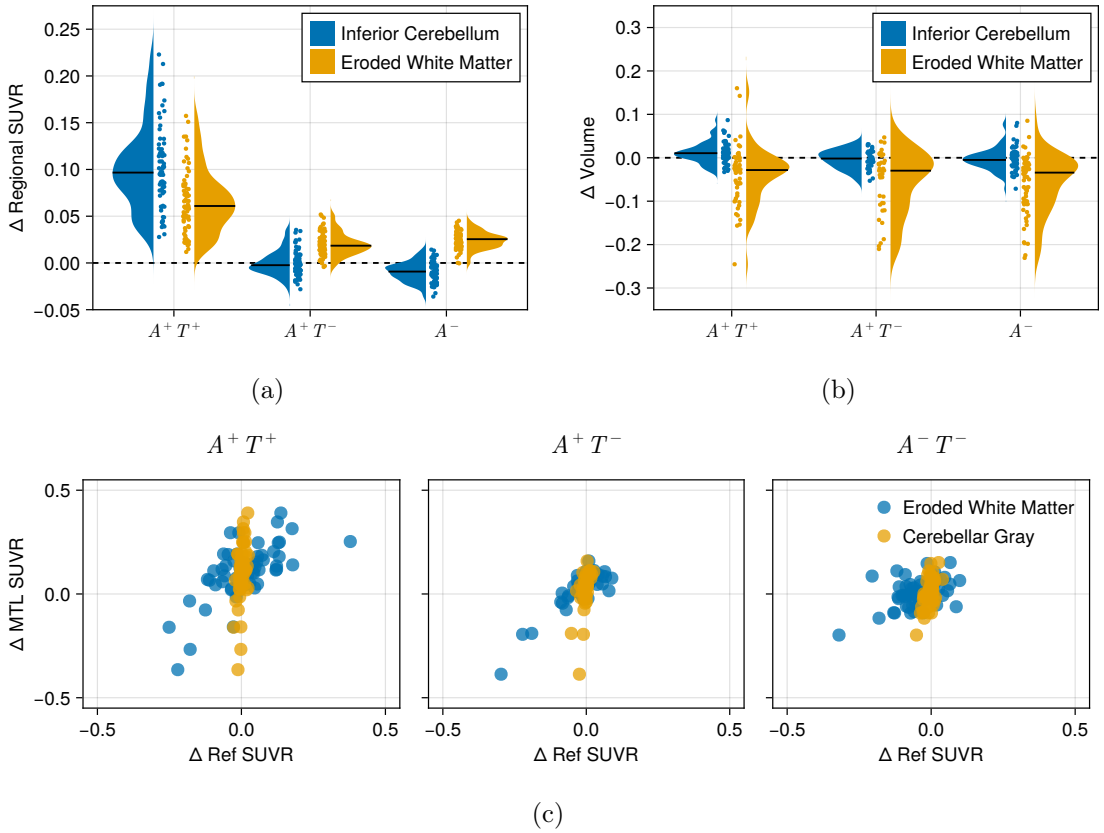


Figure S6: **Effects of reference region on longitudinal tau PET. S6a**. Longitudinal change in target regions SUVR from first to last scan in the A^+T^+, A^+T^- and A^-T^- ADNI cohorts with inferior cerebellar and eroded white matter reference regions. Each point represents the longitudinal change in SUVR for a region in the DKT atlas, averaged over subjects in each cohort. The density of these changes is shown for each group and reference region, with the median change highlighted with a black line through the density. **S6b**. Longitudinal change in reference region volume relative to baseline scan and normalised by total intracranial volume. Each point represents a subject in the ADNI cohort. **S6c**. Longitudinal change in reference region SUVR vs change in MTL (bilateral amygdala and entorhinal cortex) SUVR per subject in the each of the ADNI cohorts. Each point represents a subject in each ADNI cohort. All tau PET SUVR values by ADNI are pre-normalised by the inferior cerebellar reference region, and therefore the all SUVR values for the inferior cerebellum are equal to one. Therefore, we use the cerebellar gray region to approximate the inferior cerebellar SUVR. For the cerebellar gray region, eroded white matter and MTL regions, we use the SUVR values based on the inferior cerebellar reference region to highlight the effect of further normalisation using a different reference region, namely the eroded white matter.

S1.4 Re-Analysis on Schaefer-200 Atlas

We investigate the role of parcellation and connectome on the results presented in Section 2.4. We rerun our analysis using the Schaefer-200 atlas provided by the ENIGMA project [41]. The

larger parcellations means individual parcels have more uniform volumes. We expect this to more accurately detect transport dynamics that may be lost averaging signal over large regions, such as those in the DKT atlas. We rerun this analysis on the BF-2 dataset, using the same subjects cohorts used in Section 2.4. First, we estimate regional parameter vectors using Gaussian mixture modelling, as described in Section 4.5. Excluding regions that are not best captured by a two-component Gaussian mixture model there are $R = 198$ regions in total. We re-apply the model Eq. (17) to the A^+T^+, A^+T^- and A^-T^- BF-2 groups, collecting 2000 samples using a No-U-Turn sampler. The population level parameters are shown in Fig. S7.

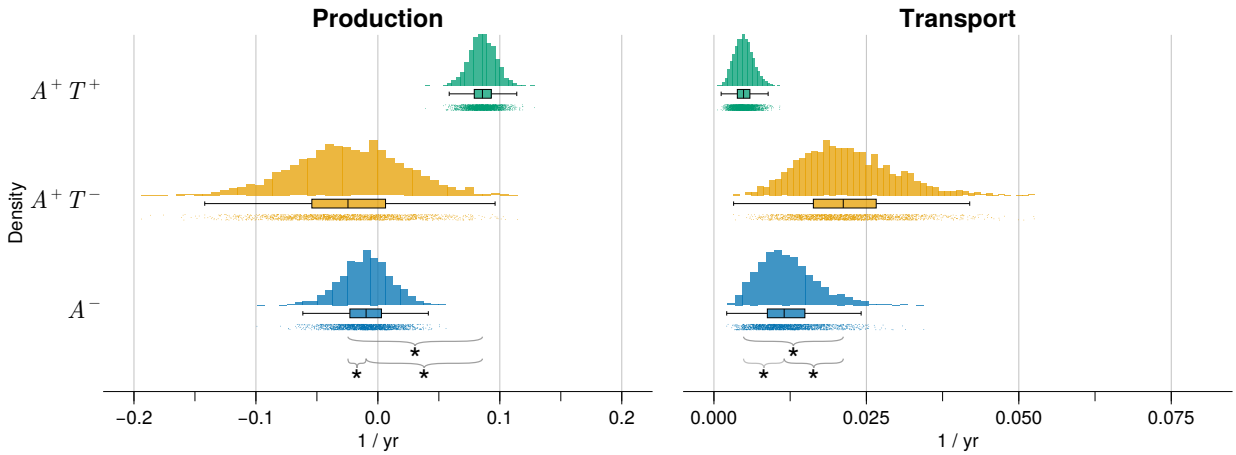


Figure S7: **Posterior distributions for BF2 and Schaefer atlas**. Population-level distributions for A^+T^+, A^+T^- and A^-T^- BF-2 cohorts, using the local FKPP model and the Schaefer-200 atlas. Differences between distributions are tested using a two sample Mann-Whitney U test and significance $p < 0.01$ is denoted by $*$.