Paths to Oblivion: Common Neural Mechanisms of Anaesthesia and Disorders of Consciousness

Andrea I. Luppi[a,b]*, Pedro A.M. Mediano[c], Fernando E. Rosas[d,e,f], Judith Allanson[b,g], John D. Pickard[b], Guy B. Williams[b,h], Michael M Craig[a,b], Paola Finoia[b], Alexander R.D. Peattie[a,b], Peter Coppola[a,b], Adrian Owen[i], Lorina Naci[j], David K. Menon[a,h], Daniel Bor[c], Emmanuel A. Stamatakis[a,b]

[a]Division of Anaesthesia, School of Clinical Medicine, University of Cambridge, United Kingdom

[b]Department of Clinical Neurosciences, University of Cambridge, Cambridge, United Kingdom

[c]Department of Psychology, University of Cambridge, Cambridge, United Kingdom

[d]Center for Psychedelic Research, Department of Brain Science, Imperial College London, London, United Kingdom

[e]Data Science Institute, Imperial College London, London, United Kingdom

[f]Centre for Complexity Science, Imperial College London, London, United Kingdom

[g]Department of Neurosciences, Cambridge University Hospitals NHS Foundation, Addenbrooke's Hospital, Cambridge, United Kingdom

[h]Wolfson Brain Imaging Centre, University of Cambridge, Cambridge, United Kingdom

[i]The Brain and Mind Institute, University of Western Ontario, London, Ontario, Canada

[j]Trinity College Institute of Neuroscience, Trinity College Dublin, Dublin, Ireland

[g]Division of Neurosurgery, School of Clinical Medicine, University of Cambridge, Cambridge, United Kingdom

*Corresponding author; email address: al857@cam.ac.uk

Abstract

The human brain generates a rich repertoire of spatiotemporal dynamics during normal wakefulness, supporting a wide variety of conscious experiences and cognitive functions. However, neural dynamics are reconfigured, in comparable ways, when consciousness is lost either due to anaesthesia or disorders of consciousness (DOC). Here, leveraging a neurobiologically realistic whole-brain computational model informed by functional MRI, diffusion MRI, and PET, we sought to identify the neurobiological mechanisms that explain the common reconfiguration of neural dynamics observed both for transient pharmacological intervention and chronic neuroanatomical injury. Our results show that, by incorporating local inhibitory action through a PET-based GABA receptor density map, our model can reproduce the brain dynamics of subjects undergoing propofol anaesthesia, and that this effect depends specifically on the spatial distribution of GABA receptors across cortical regions. Additionally, using a structural connectome obtained from DOC patients, we demonstrate how the dynamics that characterise loss of consciousness can emerge from changes in neuroanatomical connectivity. Crucially, we find that each of these two interventions generalises across datasets: a model with increased GABA-mediated inhibition can reproduce the dynamics of DOC patients' brains, and a model with a DOC connectome is also compatible with brain dynamics observed during propofol anaesthesia. These results demonstrate how increased inhibition and connectome randomisation represent different neurobiological paths towards the characteristic dynamics of the unconscious brain. Overall, the present findings begin to disentangle the neurobiological mechanisms by which highly dissimilar perturbations of the brain's neurodynamics can lead to unconsciousness.

Introduction

A central challenge of contemporary neuroscience is the quest to understand how the neurobiology and function of the human brain give rise to conscious experience [1,2]. One way to address this question is to identify changes in brain function that accompany changes in conscious state. However, the brain is a paradigmatic example of a complex system [3], and different perturbations of its precise functioning can serve as a path towards loss of consciousness. Examples of such perturbations range from transient pharmacological (general anaesthetic) interventions having widespread effects on neuromodulation [4–7], to chronic disorders of consciousness arising from traumatic or anoxic injuries of diverse location and extent, often including changes to the physical connectivity between brain regions [8–12].

In order to shed light on how the human brain supports consciousness, two fundamental challenges must be addressed. The first is to identify neurobiological signatures of consciousness and its loss that are generalisable across different paths to unconsciousness, rather than being specific to any of them [13]. The second is to provide a mechanistic understanding of the previously identified signatures of consciousness, in terms of the underlying neurobiological principles [14].

Substantial progress has recently been made towards answering the first question. The human brain generates a constantly changing repertoire of neural dynamics, supporting the rich variety of conscious experiences and cognitive functions [15,16,25–29,17–24]. These complex dynamics self-organise [30] according to the recursive local-global interplay of neuronal excitation and inhibition, and their intricate interactions across the network of the brain's anatomical connections (human connectome) [16,31–37].

Recently, converging evidence from multimodal neuroimaging has further shown that brain dynamics are substantially and consistently altered when consciousness is lost, which suggests that specific aspects of neural dynamics may play a central role in supporting human consciousness. In particular, striking similarities have been observed across brain dynamics induced by different ways of losing consciousness, such as acute pharmacological intervention with different anaesthetic drugs, but also chronic disorders of consciousness [13,38–44]. These findings suggest that spatio-temporal dynamics may provide a "common currency" between

brain and mind [45] capable of encoding the difference between consciousness and unconsciousness.

Despite recent achievements in finding common signatures of consciousness and its loss, the causal neurobiological mechanisms underlying these processes remain elusive. For example, anaesthesia is known to operate at the level of neurotransmission (propofol is a potent agonist of inhibitory GABA-A receptors [46,47]), without affecting the physical organisation of the human connectome. In contrast, disorders of consciousness (DOC) arise from severe brain injury, typically due to anoxia or head trauma, which cause brain lesions of variable severity and location often including alterations of the physical connectivity between patients' brain regions [10,48–56]. This similarity of outcomes and neural signatures [13,38–42] despite arising from radically different causes, begs the question: How can (transient) pharmacological and (chronic) structural perturbations converge to similar effects on brain dynamics, and the corresponding state of unconsciousness?

Here, we sought to obtain mechanistic insights into this fundamental question by employing whole-brain computational modelling, which is emerging as a powerful tool to investigate the neurobiological mechanisms underlying macroscale neural phenomena [14,35,57–60]. These models represent regional macroscale activity in terms of local dynamics influenced by inter-regional anatomical connectivity (obtained e.g. from diffusion MRI tractography), using approximations derived from oscillatory mechanisms (e.g. Kuramoto model), statistical mechanics (e.g. Ising model), or dynamical systems theory (e.g. Hopf model) [35,61].

Crucially, *in silico* models are uniquely suited to investigate how different neurobiological perturbations can induce similar alterations of brain dynamics [14]. Just like neuropsychological studies in human patients and experimental lesions in animal models have provided invaluable insights about brain organisation, function and dysfunction [62–65], whole-brain computational models can be systematically and reversibly lesioned to investigate the resulting alterations in macroscale brain dynamics. Being fully accessible to the researcher, the model's parameters can be systematically altered in ways that are still beyond the capabilities of experimental research, whether in humans or animals [14,66].

For these reasons, whole-brain computational modelling is becoming increasingly prominent as a tool to investigate the causal mechanisms that drive brain network organisation in healthy

and pathological conditions [35,57,75,67–74]. In particular, recent efforts have capitalised on the tractability of the Hopf and Ising models, beginning to shed light on the mechanisms of loss of consciousness during sleep, anaesthesia, and disorders of consciousness in terms of local-global synchronisation behaviour (Hopf model) [76–82] or statistical temperature (Ising model) [83–85].

Crucially, recent work has demonstrated that more detailed biophysical models that incorporate neurophysiologically realistic information about excitation, inhibition and neuromodulation – so-called *Dynamic Mean Field* (DMF) models – can provide insights about pharmacologically-induced changes in macroscale fMRI dynamics, in terms of the underlying neurobiology [86–88]. These models reduce the intricate dynamics of individual neurons to a set of coupled differential equations which approximate the detailed microscale neural properties of spiking neurons (incorporating realistic aspects of neurophysiology such as synaptic dynamics and membrane potential) [89] via a mean-field reduction [31,34,37,90,91]. Specifically, cortical regions are represented as macroscopic neural fields, whose local dynamics are coupled together by a network of anatomical connections [31,34,37]. An additional biophysical haemodynamic model can then be used to turn the DMF model's dynamics into a realistic simulator of BOLD signals [92]. Thus, neurobiologically realistic whole-brain computational modelling provides a principled way to bridge across scales, relating the macroscale neural dynamics of fMRI to the microscale neurophysiological mechanisms from which they emerge [14].

However, to date no studies have harnessed the power of DMF models to provide neurobiologically realistic accounts of pharmacological and chronic loss of consciousness. Such an effort is vital because as Sanz-Perl and colleagues remark, "*modeling efforts incorporating more complex dynamics could allow in silico rehearsal of interventions with ampler neurobiological interpretation*" [80]. Here, we leveraged a neurobiologically realistic DMF model informed by multimodal neuroimaging including empirical dynamics from functional MRI, anatomical connectivity obtained from diffusion MRI, and GABA-A receptor density estimated from positron emission tomography (PET) (Figure 1).

We used this modelling approach to simulate the empirical fMRI macroscale brain dynamics observed in the same N=16 subjects at baseline and during loss of consciousness induced by the intravenous anaesthetic, propofol. We also studied the fMRI dynamics of a cohort (N=21) of patients suffering from chronic disorders of consciousness (DOC) as a result of severe brain injury (traumatic or anoxic), comparing them with a group of N=20 healthy controls. By

subjecting the models to "virtual anaesthesia" (local modulation of inhibitory gain based on empirical GABA-A receptor distribution) and "virtual DOC" (alteration of the model's structural connectome), we sought to identify the neurobiological mechanisms underlying a fundamental question of modern neuroscience: how can transient perturbations of neurotransmission and chronic lesions to the structural connectome, both give rise to unconsciousness and its characteristic similar brain dynamics [13,38,39,41,93]?

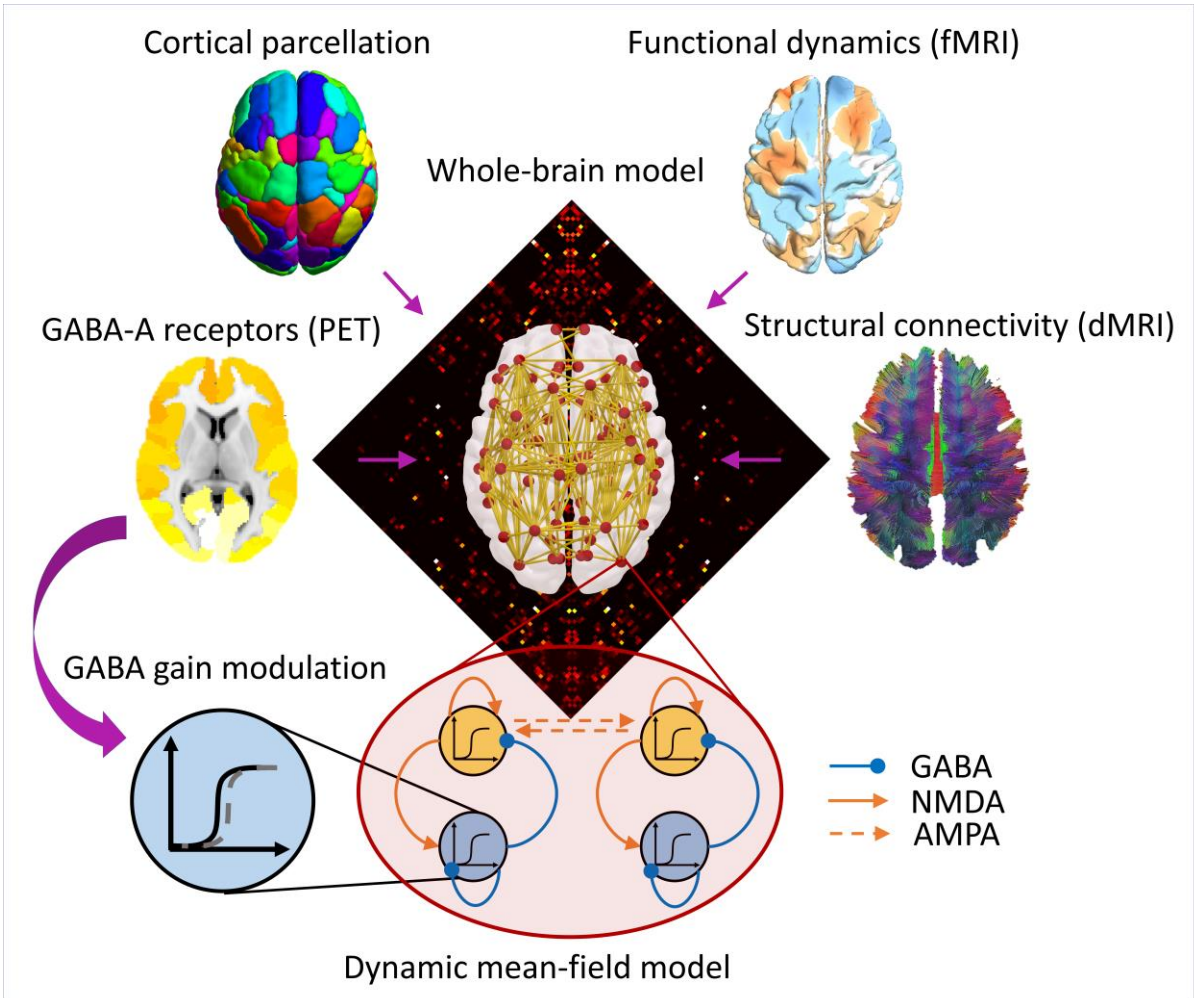

Figure 1. **Overview of whole-brain computational model incorporating multimodal neuroimaging data.** Based on a cortical parcellation with 68 regions of interest, each node (cortical region) is modelled through a neurophysiologically realistic biophysical model incorporating excitatory (NMDA) as well as inhibitory (GABA) synaptic dynamics. Nodes are connected by structural connectivity (from diffusion MRI) and the model's dynamics are fitted to simulate empirical brain dynamics (from functional MRI). Neurotransmitter information from PET can also be added in the model as modulating the local neuronal gain [86].

Results

The neurobiologically realistic dynamic mean-field (DMF) model only one free parameter: a global coupling parameter, denoted by *G*, which scales the excitatory-to-excitatory coupling between brain regions, as established by the empirical structural connectome. Thus, calibrating the model corresponds to finding the value of *G* that allows the model to best simulate observed fMRI dynamics of the human brain at rest (which is assessed in terms of the KS-distance between real and simulated data, see [86]). We followed this procedure for each of our two datasets (Figure 2): for the propofol dataset we optimised the model to fit the fMRI dynamics acquired during the baseline awake scan, and for the DOC dataset we optimised the model to fit the dynamics observed in the healthy controls. These two calibrated DMF models - with their corresponding global coupling values fitted to the dynamics of the conscious brain - constitute the starting point for our investigations.

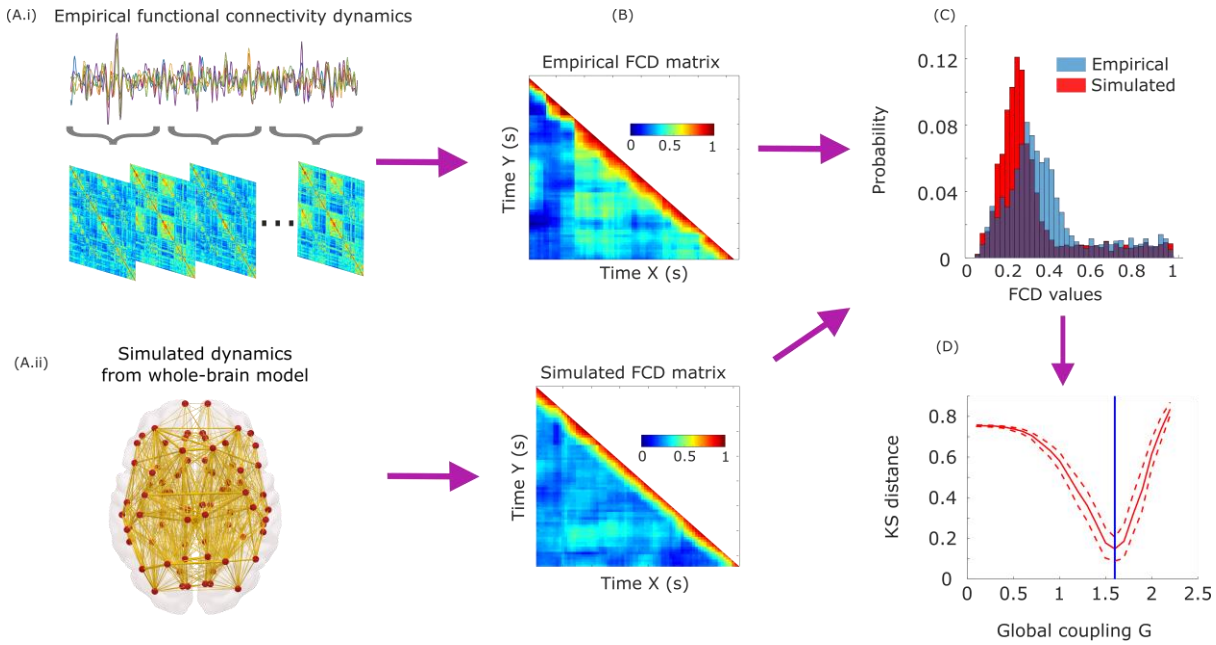

Figure 2. Overview of model fitting procedure. (A.i) Time-resolved matrices of functional connectivity are obtained from empirical functional MRI via the sliding-window approach: regional BOLD time-series are partitioned into windows of 30 TRs, sliding by 3 TRs at a time, following the same approach as previous work using the DMF model [86]; functional connectivity between each pair of regions is computed within each window by means of Pearson correlation, generating a stack of FC matrices representing the evolution of FC over time. (A.ii) The same procedure is repeated for the simulated BOLD timeseries produced by the model with various levels of the global coupling parameter, *G*. (B) For both the empirical and simulated functional connectivity dynamics (FCD), a time-versus-time FCD matrix is computed by correlating the time-dependent FC matrices centred at each timepoint. (C) Histograms of the distribution of FCD values in each matrix are obtained over all participants (blue) and for each simulation (red), and their similarity is evaluated by means of the Kolmogorov-

Smirnov distance. (D) Across values of the global coupling parameter *G*, we compute the KS-distance between the empirical FCD and the FCD of each simulation (red solid line; dashed lines indicate standard deviation). The optimum value of *G* for the model (blue vertical line) is chosen as the one that minimises the average KS-distance across 100 simulations (here shown for the awake condition of the propofol dataset, where *G*=1.6). This procedure determines the value that allows the model to best simulate the empirical dynamics of functional connectivity in the healthy human brain.

Inhibitory modulation from GABA-A receptor distribution reveals a shared mechanism for loss of consciousness

The effects of propofol anaesthesia on the brain were modelled by capitalising on the recently built whole-brain map of GABA-A regional receptor density, generated on the basis of benzodiazepine receptor (BZR) density measured from [^{11}C]flumazenil Positron Emission Tomography (PET) [94]. Incorporating this information in the DMF model allowed us to evaluate the extent to which the dynamics of the anaesthetised brain can be explained in terms of propofol-induced alterations in the detailed balance of local excitation and inhibition.

In previous work Deco and colleagues [86] modelled the effects of the serotonergic drug LSD by locally modulating the neuronal gain of each excitatory population in the model according to the empirical distribution of 5HT-2A receptors across brain regions [86]. Inspired by their approach, here we show for the first time that the influence of regional GABA-A receptor density on functional dynamics can be modelled using a DMF model informed by regional GABA-A receptor density.

The strategy followed in [86] was to first calibrate the model on baseline data to obtain a global coupling value, and then fit a secondary excitatory parameter separately on baseline and post-dose data. Our approach follows Deco's but differs in one key respect: given the inhibitory nature of GABA, we modulated the inhibitory (rather than excitatory) local gain.[1] To do so, we introduced an inhibitory gain scaling parameter in the model, denoted by s_I. This parameter allowed us to scale the inhibitory gain at each region according to the empirical local density of GABA-A receptors, as quantified based on PET-derived maps of receptor density [94].

[1] Note that the excitatory and inhibitory populations within each region in the biophysical model are mutually and recursively coupled, and hence both excitation and inhibition are eventually affected by this procedure.

This procedure allowed us to ask whether adjusting the value of inhibitory gain s_I according to local GABA-A receptor density would allow the model to simulate the characteristic dynamics of acute propofol-induced unconsciousness. A positive answer to this question would implicate regional GABA-ergic inhibition as a neurobiological mechanism behind the action of propofol (a known GABA-ergic agonist) in inducing the characteristic macroscale dynamics observed during loss of consciousness due to propofol anaesthesia [13,38,40,41].

To address this issue, we studied whether some appropriate value of s_I (which scales the gain related to the local GABA-A receptor density) would improve the model's ability to simulate the dynamics of deep propofol anaesthesia. For this purpose, we used the previously calibrated DMF model to generate simulations for each value of s_I between 0 (corresponding to the model without local GABA inhibitory modulation) and 1, in increments of 0.02. Then, for each value of s_I, we computed the KS distance between the model's simulated macroscale dynamics and the empirical dynamics observed in the anaesthetised subjects. The optimal value of s_I, was then identified as the value that resulted in the minimum mean KS distance between empirical and simulated dynamics (across N=10 simulations for each value of s_I), thereby establishing a *propofol model*.

Note that a model incorporating additional neurobiological information (in this case, information about the regional distribution of GABA-A receptors) might produce generally better-fitting simulations just in virtue of its increased complexity. To control for this possibility, we compared the propofol model with an analogous model, which also incorporated regional GABA-A receptor density, but whose inhibitory gain scaling parameter was chosen as the one that best fitted the empirical dynamics observed in awake subjects. We refer to this as the *baseline model*.

Having completed the fitting procedure for our models, we then proceeded to analyse the models' performance. To this end, we generated 100 simulations from each of the propofol and baseline models. For both models, we then computed the KS distance between each simulation, and the empirical dynamics observed during wakefulness, and during anaesthesia. This provided us with a way to quantify the ability of each model (in terms of goodness of fit, i.e. low KS distance) to simulate the empirical brain dynamics observed during wakefulness, and the empirical brain dynamics observed during propofol-induced loss of consciousness.

Results show that incorporating the regional distribution of GABA-A receptors as local modulators of the global inhibitory gain, substantially improved the DMF model's fit to empirical brain dynamics observed during propofol anaesthesia (Figure 3B and Table 1). This improvement could not be explained as a generic effect due to additional model complexity: if this had been the case, improvements should have been observed for the model's ability to fit both awake and anaesthetised dynamics. Instead, the improvement was specific to anaesthetised dynamics, and the propofol model actually had poorer ability to fit the awake dynamics. In other words, the addition of this inhibition parameter in accordance with its empirical distribution across brain regions, makes the model capable of switching between simulating awake or anaesthetised brain dynamics - and when it is simulating anaesthetised brain dynamics, the model's fit to awake brain dynamics deteriorates. Since propofol is a well-known GABA-ergic agonist, these results confirm that taking into account GABA agonism (local modulation of inhibitory gain by regional GABA-A receptor density) is sufficient to recapitulate the known effects of the GABA-ergic agent propofol on empirical brain dynamics, leading to dynamics that are known to characterise the state of unconsciousness.

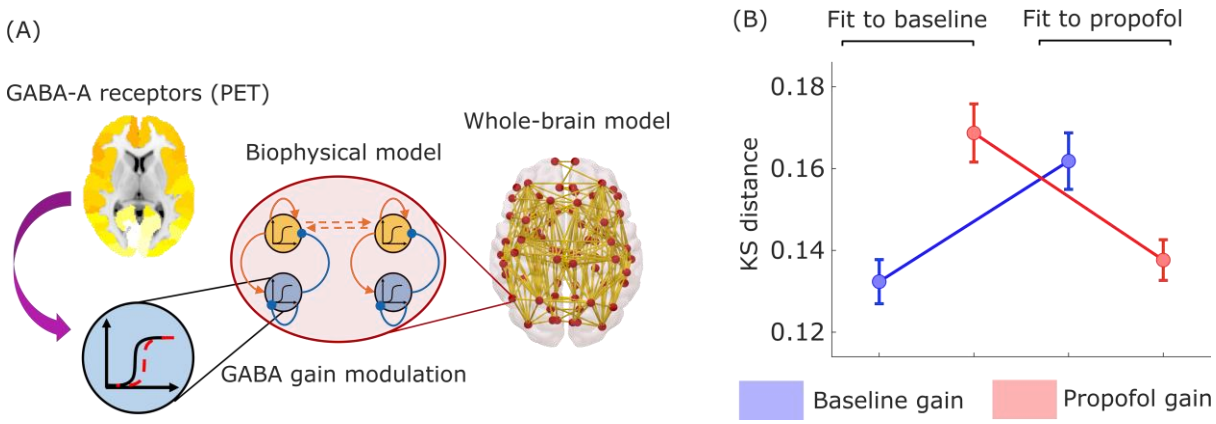

Figure 3. **Modulation of inhibitory gain by empirical GABA-A receptor density improves model fit to propofol dynamics.** (A) The inhibitory gain of each node in the balanced DMF model is modulated by the regional density of GABA-A receptors, estimated from PET. (B) Mean model fit for 100 simulations for the propofol dataset, quantified as KS-distance to the baseline (awake) and propofol (anaesthetised) conditions, using a value of gain for inhibitory scaling s_I derived from calibrating the model with either the baseline or propofol conditions. Error bars represent the standard error of the mean.

Table 1. Statistical testing for the effects of local inhibitory GABA modulation.

	Original GABA map			Scrambled GABA map			Uniform GABA map		
	Estimate	Standard error	p-value	Estimate	Standard error	p-value	Estimate	Standard error	p-value
Model	0.04	0.009	<.0001	-0.01	0.006	.081	$2x10^{-17}$	0.007	1.0
Target condition	0.03	0.009	.0008	0.003	0.006	.610	0.02	0.007	.001
Interaction	- 0.06	0.01	<.0001	--	--	--	--	--	--

Crucially, we also confirmed that the improved fit to anaesthetised dynamics is not merely the result of increasing overall inhibition in the model: rather, regional information about the distribution of GABA receptor density plays a key role in the model's improved fit. To demonstrate this point, we show that the results are not replicated if the PET-derived regional distribution of GABA-A receptor density is reshuffled across regions (Figure 4A), or if uniform values are used for each region (i.e. by setting all regions to have a value equal to the mean of the distribution; Figure 4B). In both cases, the model's ability to fit anaesthetised dynamics is not improved with respect to the baseline model (Figure 4A,B and Table 1). Therefore, our results show that the specific regional distribution of GABA-A receptors across the cortex plays a key role in generating the brain dynamics characteristic of unconsciousness induced by propofol administration.

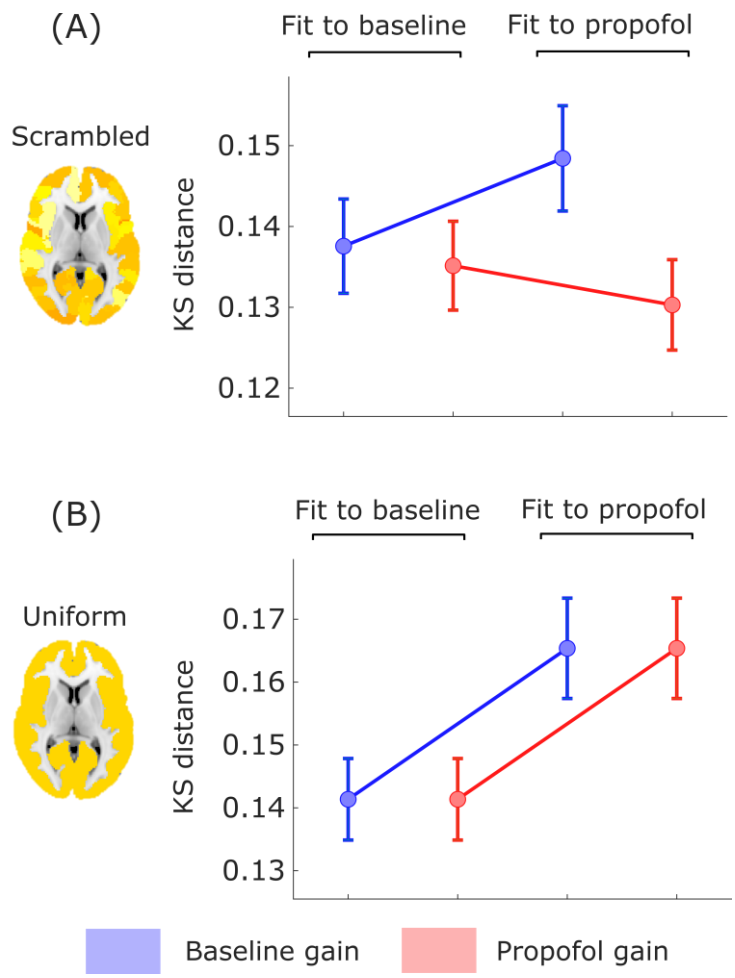

Figure 4. **Modulation of inhibitory gain by reshuffled or uniform GABA-A receptor density**. (A) The inhibitory gain of each node in the balanced DMF model is modulated by the regional density of GABA-A receptors, reshuffled across cortical regions. Plot shows the mean model fit for 100 simulations, quantified as KS-distance between simulated and empirical dynamics, for each combination of condition and gain. (B) The inhibitory gain of each node in the balanced DMF model is modulated by the regional density of GABA-A receptors, set to a uniform value (mean of the empirical distribution) across cortical regions. Plot shows the mean model fit for 100 simulations, quantified as KS-distance between and empirical dynamics, for each combination of condition and gain. Error bars represent the standard error of the mean.

Simulated brain injury induces unconscious-like dynamics

Whole-brain computational models provide a unique tool to understand the effects of connectome alterations on macroscale brain dynamics [70,71,83]. Connectome replacement allows us to determine which of two conditions is more compatible with a given perturbation of the connectome, in terms of the connectome's capacity to support the corresponding brain dynamics. Specifically, a smaller impact of connectome perturbation on the model's fit to condition X than condition Y, indicates that the perturbed connectome is better suited to supporting the dynamics of condition X than Y. We term this procedure "*Connectome Replacement Analysis*", which is illustrated in Figure 5A.

Leveraging this capability, we subjected the DMF model to the virtual equivalent of severe brain injury: namely, we replaced the underlying connectivity matrix governing the long-range interactions between brain regions with a connectome obtained from diffusion-weighted imaging of N=21 patients with chronic DOC due to severe brain injury. This procedure imparts on the model with effects akin to what severe brain injury does on anatomical connectivity. This "virtual DOC" provides a way to isolate the effects over brain dynamics of connectivity disruptions that result in loss of human consciousness.

Note that such substantial perturbations are expected to deteriorate the model's ability to replicate the dynamics of awake healthy brains (i.e. increasing the KS distance, corresponding to decreased goodness-of-fit): our models were optimised with biophysical parameters pertaining to healthy brains, and using a healthy connectome. However, our hypothesis was that the dynamics generated by the model with DOC connectome should be more similar (lower KS distance, indicating a better fit) to the empirical dynamics of DOC patients' brains, than to the dynamics of conscious, healthy brains.

Our results supported these predictions. As expected, connectome replacement led to a reduction in the model's ability to fit control brain dynamics (Figure 5B). Also as expected, the model with the healthy connectome was better able to simulate conscious than unconscious dynamics (Figure 5B, blue plots). Crucially, however, this pattern reversed following replacement of the healthy connectome with the DOC connectome (Figure 5B, red plots): dynamics generated using the DOC connectome were more similar to the empirical brain dynamics of DOC patients, than to conscious controls (Table 2). This observation supports our hypothesis, demonstrating that unconscious brain dynamics are more compatible with the DOC connectome than conscious dynamics; below, we also demonstrate that this result is not specific to the chronic unconsciousness that characterises disorders of consciousness, but rather it generalises to the transient unconsciousness caused by propofol anaesthesia, too.

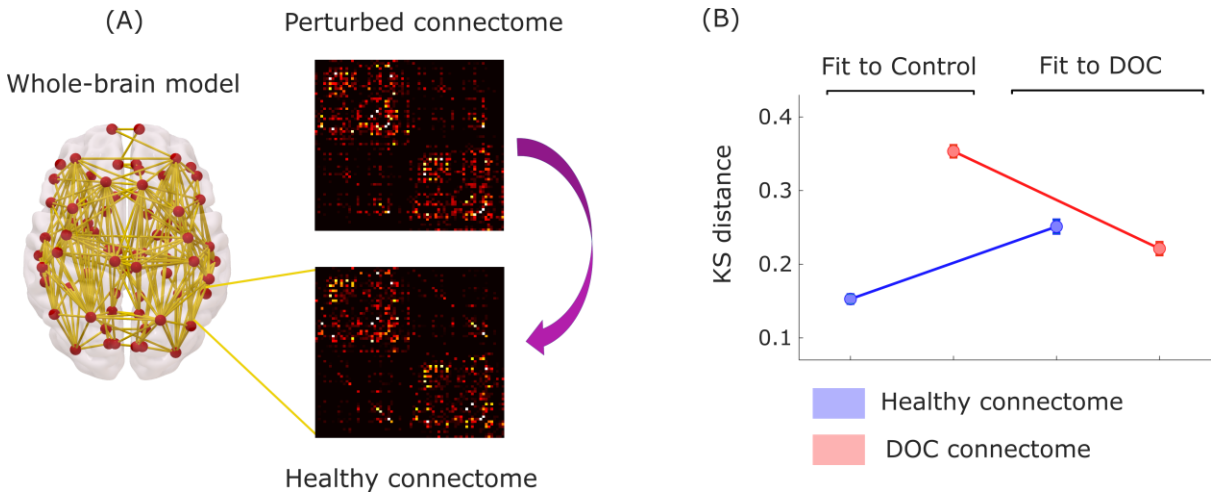

Figure 5. **Connectome replacement analysis with DOC connectome.** (A) The original healthy connectome of the model is replaced with the group-average connectome obtained from diffusion MRI of N=21 DOC patients,

and the resulting model is used to generate 100 simulations. (B) Plot shows the mean model fit for 100 simulations, for the original and perturbed models for the DOC dataset, quantified as the KS-distance to the control (conscious) and DOC (unconscious) conditions. Error bars represent the standard error of the mean.

Table 2. Statistical testing for the effects of connectome replacement.

	DOC connectome			Random connectome			Lattice connectome		
	Estimate	Standard error	p-value	Estimate	Standard error	p-value	Estimate	Standard error	p-value
Model	0.20	0.012	<.0001	0.29	0.012	<.0001	0.06	0.012	<.0001
Target condition	0.10	0.012	<.0001	0.10	0.012	<.0001	0.10	0.012	<.0001
Interaction	- 0.23	0.016	<.0001	- 0.21	0.017	<.0001	0.05	0.016	.007

Remarkably, these results could be replicated by replacing the original healthy structural connectome with a randomised version having the same average connectivity [95]. After perturbation, the model's fit to DOC brain dynamics became better than the model's fit to conscious dynamics (Figure 6A and Table 2), suggesting that unconscious dynamics are more compatible than conscious dynamics with a randomised connectome. In contrast, the opposite effect was observed when the original connectome was rewired into a regular (lattice) network, which had a larger impact on the model's fit to DOC than control brain dynamics (Figure 6B and Table 2). Crucially, the fact that lattice networks further degrade the fit of DOC dynamics demonstrates that not every deviation from the original healthy connectome leads to a higher compatibility with unconscious over conscious dynamics: rather, compatibility depends on the specific topology of the perturbed connectome. Together, these results suggest that DOC dynamics are more compatible with an unstructured connectome.

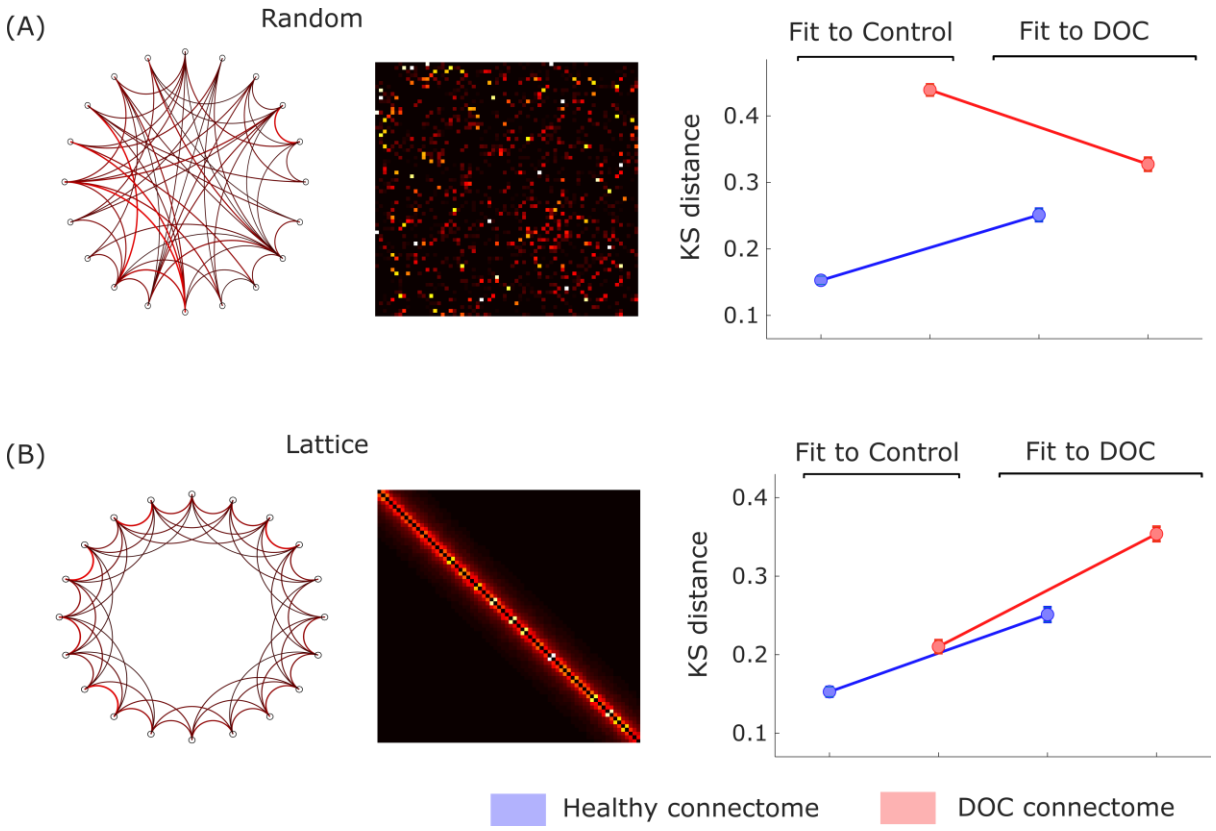

Figure 6. **Connectome replacement analysis with random and lattice networks.** (A) Plots show the mean model fit for 100 simulations, quantified as the KS distance to each condition for the original and perturbed models using a randomised connectome. (B) Plots show the mean model fit for 100 simulations, quantified as the KS distance to each condition for the original and perturbed models using a lattice-like connectome. Error bars represent the standard error of the mean.

Generalisation across datasets

Having identified the role of GABA-mediated inhibition for propofol anaesthesia, we next sought to determine to what extent inhibition can also explain the dynamics of unconsciousness arising from severe brain injury. Our rationale was that, even though these patients have not been exposed to GABA-ergic agents but rather owe their condition to severe brain injury, recent evidence suggests similarities of brain dynamics during anaesthesia and disorders of consciousness [13,38,39,42,96]. A positive answer to this question would further implicate a change in the excitation-inhibition balance, not just in the generation of dynamics pertaining to propofol anaesthesia, but more broadly as a general mechanism responsible for the characteristic dynamics of unconscious states - whether due to anaesthesia or brain injury.

Therefore, we followed the same "virtual anaesthesia" procedure with empirical data from healthy controls and DOC patients. Intriguingly, we observed analogous results: local modulation of inhibitory gain based on GABA-A receptor density allowed the model to substantially improve its fit to DOC patients' brain dynamics (Figure 7A and Table 3). However, in contrast with propofol anaesthesia, the improvements were also observed when the regional receptor map was scrambled, or replaced by a uniform map (Figure 7B,C and Table 3). Thus, whereas propofol anaesthesia depends on the specific distribution of GABA-A receptors across the cortex, indicating that these receptors are mediating the effects of propofol, the characteristic dynamics of DOCs are less selective, and appear to correspond to a non-specific increase in global inhibition.

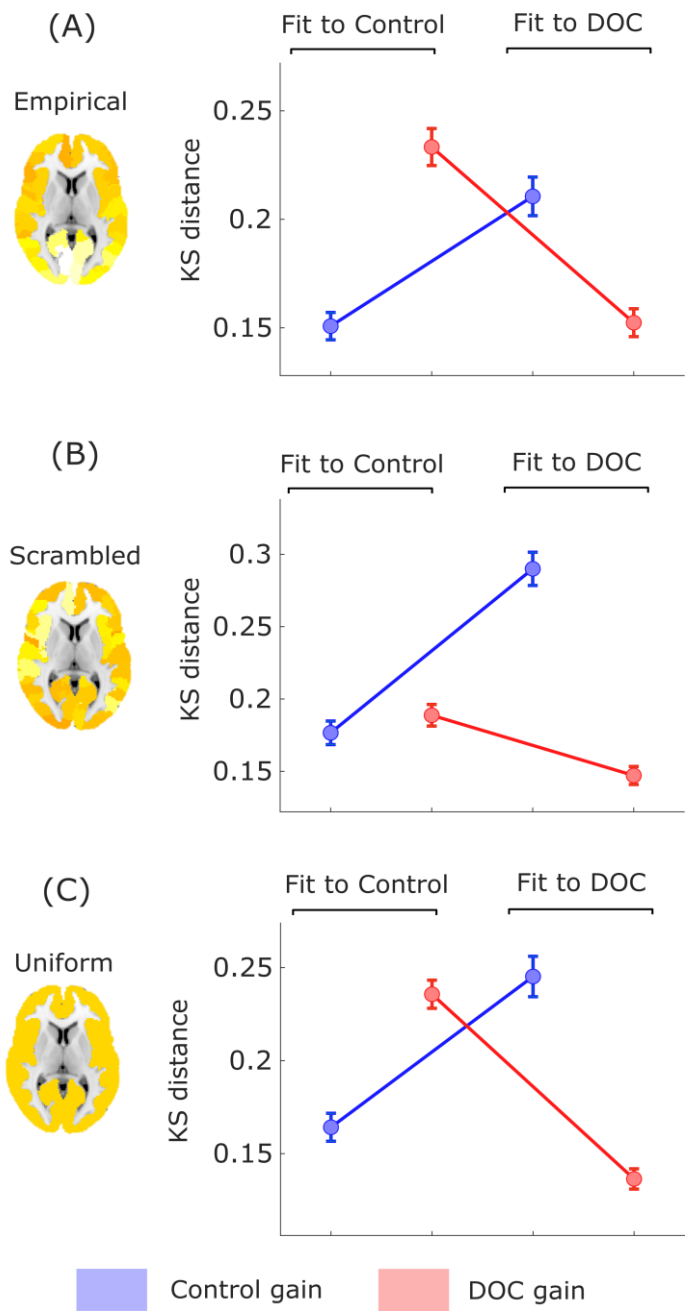

Figure 7. **Modulation of inhibitory gain by empirical GABA-A receptor density improves model fit to DOC brain dynamics.** (A) Mean model fit for 100 simulations for the DOC dataset, evaluating KS distance (lower means better fit) to the control (conscious) and DOC (unconscious) conditions, using a value of inhibitory gain scaling s_I derived from either the control subjects or the DOC patients. (B) The inhibitory gain of each node in the balanced DMF model is modulated by the regional density of GABA-A receptors, reshuffled across cortical regions. Plot shows the mean model fit (KS-distance) between simulated and empirical dynamics, for each combination of condition and gain. (C) The inhibitory gain of each node in the balanced DMF model is modulated by the regional density of GABA-A receptors, set to a uniform value (mean of the empirical distribution) across cortical regions. Plot shows the mean model fit (KS-distance) between simulated and empirical dynamics, for each combination of condition and gain. Error bars show the standard error of the mean.

Table 3. Generalisation of local GABA modulation results to DOC patients.

	Original GABA map			Scrambled GABA map			Uniform GABA map		
	Estimate	Standard error	p-value	Estimate	Standard error	p-value	Estimate	Standard error	p-value
Model	0.08	0.011	<.0001	0.01	0.012	<.0001	0.07	0.011	<.0001
Target condition	0.06	0.011	<.0001	0.11	0.012	.317	0.08	0.011	<.0001
Interaction	- 0.14	0.015	<.0001	- 0.15	0.017	<.0001	- 0.18	0.016	<.0001

Additionally, if propofol and severe injury are different ways by which the human brain can be pushed towards unconsciousness, then inducing a virtual DOC via connectome replacement should also lead to a model that is better able to simulate the dynamics of an anaesthetised brain, than an awake brain - thereby recapitulating what we previously observed with DOC brain dynamics. Remarkably, results show that - as previously observed with DOC patients - simulated dynamics generated from a model using the DOC connectome are more compatible (lower KS distance) with the dynamics of propofol anaesthesia than with the dynamics of awake subjects' brains (Figure 8A and Table 4).

Furthermore, the importance of topology of the perturbed connectome was also observed for propofol dynamics, with randomisation of the connectome having a smaller effect on the model's ability to fit propofol dynamics than awake dynamics (Figure 8B and Table 4). Conversely, the decrease in the model's ability to fit empirical data was exacerbated for the propofol condition, when the original connectome was replaced with a lattice network (Figure 8C and Table 4).

These findings generalise our DOC results to propofol anaesthesia, indicating that the DOC connectome is not only more compatible with DOC patients' brain dynamics than with healthy subjects' dynamics. Rather, the generalisation to propofol anaesthesia suggests that the DOC connectome may be more compatible with unconscious dynamics in general: whether arising from brain injury or pharmacological intervention.

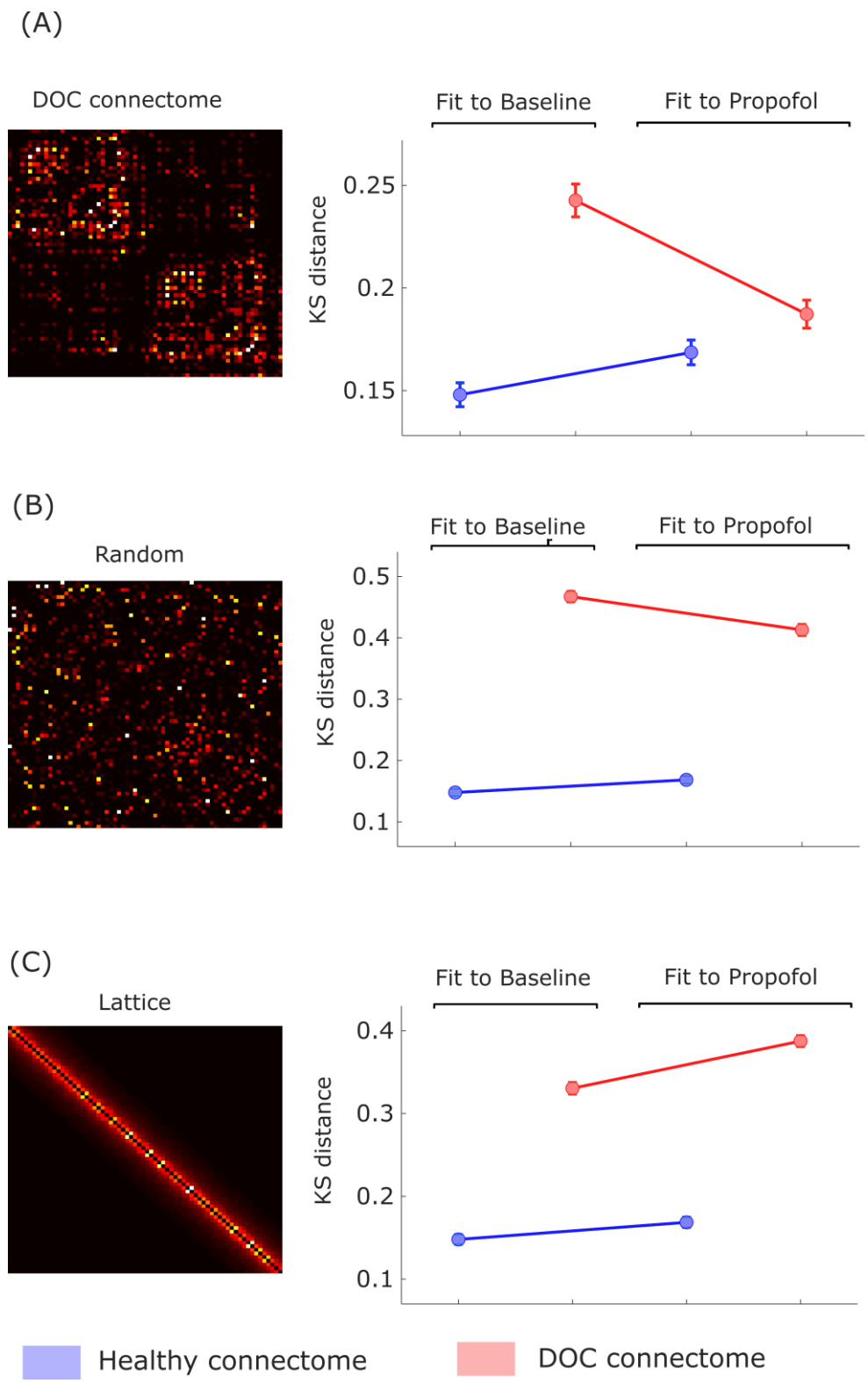

Figure 8. **Connectome replacement analysis with DOC connectome generalises to propofol anaesthesia.** (A) Plot shows the mean model fit for 100 simulations from the original and perturbed models for the DOC dataset, evaluated as KS-distance to the baseline (conscious) and propofol (unconscious) conditions. (B) Plot shows the mean model fit (KS distance) for 100 simulations, showing the fit to each condition for the original and perturbed models using a randomised connectome. (C) Plot shows the mean model fit (KS distance) for 100 simulations, showing the fit to each condition for the original and perturbed models using a lattice-like connectome. Error bars represent the standard error of the mean.

Table 4. Generalisation of connectome replacement results to anaesthetised volunteers.

	DOC connectome			Random connectome			Lattice connectome		
	Estimate	Standard error	p-value	Estimate	Standard error	p-value	Estimate	Standard error	p-value
Model	0.09	0.010	<.0001	0.32	0.010	<.0001	0.20	0.006	<.0001
Target condition	0.02	0.010	.031	0.02	0.010	.0495	0.04	0.006	<.0001
Interaction	- 0.08	0.013	<.0001	- 0.07	0.015	<.0001	-	-	-

Discussion

The human brain generates a constantly varying set of neural dynamics, supporting a rich repertoire of conscious experiences and cognitive functions [15–18,20], which unfold out of an interplay between the structural connectome and the brain's neuromodulatory mechanisms [16,31–37]. This paper sought to identify neurobiological mechanisms that are capable of explaining how highly dissimilar causes - such as transient perturbations of neurotransmission *versus* chronic lesions to brain anatomy and connectivity - can give rise to loss of consciousness and its characteristic brain dynamics [14].

To this end, we employed a whole-brain *Dynamic Mean Field* model that simulates macroscale functional dynamics of the human brain by means of neurobiologically realistic biophysical modelling, which integrates empirical neuronal dynamics from functional MRI, anatomical connectivity obtained from diffusion MRI, and neurotransmitter receptor density estimated from positron emission tomography. In addition to providing insights about the mechanisms that are responsible for observed macroscale neurobiology, DMF models offer a powerful tool to understand the causal mechanisms underlying pharmacologically-induced macroscale changes in neural dynamics [86–88]. The effect of inhibition was assessed by enriching the DMF model, modulating the neuronal gain of each inhibitory population according to the empirical density of GABA-A receptors across cortical regions, quantified using in-vivo Positron Emission Tomography autoradiography [94]. Additionally, the role of lesions to brain anatomy was studied by considering the connectome observed in a population of DOC patients.

Our results demonstrate that GABA-mediated inhibition plays a mechanistic role in the emergence of the characteristic macroscale dynamics observed during propofol-induced unconsciousness. These results align with neurophysiological evidence indicating that propofol is primarily a GABA-A receptor agonist [46,47]. Indeed, our results further indicate that propofol anaesthesia is crucially dependent on the specific regional distribution of GABA-A receptors across the cortex, since neither reshuffling this distribution across regions nor setting all regions to equal density values could reproduce the same effect. Future research may build on this finding of regional specificity by seeking to identify whether GABA-A receptor density at specific regions plays an especially prominent role in propofol-induced anaesthesia, or whether concerted action across the entire cortex is required. The prominent involvement of the brain's default mode network in anaesthesia induced with the GABA-ergic agents propofol and

sevoflurane [13,39,40,93,97,98] and more specifically the precuneus/posterior cingulate cortices, suggests that GABA receptor density at these regions may be an especially promising candidate for predicting anaesthetic effects.

Remarkably, our PET-informed results showed that considering GABA-mediated scaling of regional inhibitory gain also improved the model's ability to simulate the characteristic dynamics of DOC patients' brains, even though these patients owe their chronic condition to severe brain injury rather than pharmacological intervention. This observation suggests a change of excitatory-inhibitory balance in favour of inhibition, not just in the generation of dynamics pertaining to propofol anaesthesia, but more broadly as a general neurobiological mechanism for the dynamics that characterise unconsciousness - whether due to anaesthesia or brain injury. Indeed, there is evidence that physiologically awake but unconscious DOC patients show cortical OFF-periods analogous to those observed in healthy individuals during sleep [99], possibly arising from reduced cortico-cortical connectivity and a resulting shift in excitatory-inhibitory balance towards excessive inhibition [100], as is observed at a local level after stroke [101]. And indeed, both disorders of consciousness and general anaesthesia are known to correspond to reduced cerebral metabolism, as measured with PET [102,103].

Crucially, the present results are consistent with the recent findings of [41], where the framework of connectome harmonic decomposition revealed how propofol anaesthesia and disorders of consciousness can both be explained in terms of coarse-grained harmonic modes of the connectome providing increased contributions to brain dynamics. An increase in the contribution of coarse-grained harmonic modes had been previously generated *in silico* by increasing inhibition in a Wilson-Cowan computational model [16]. Hence, taken together these studies suggest that a change in the global balance between excitation and inhibition in favour of inhibition seems to play a role in the characteristic dynamics of unconsciousness - which the present work formally demonstrates by combining empirical fMRI data on propofol anaesthesia and DOC, with whole-brain computational modelling informed by GABA-A receptor density.

Nevertheless, a key difference emerged between anaesthesia and DOC: whereas anaesthesia critically depends on propofol's specific pattern of local inhibition across the cortex, incorporating regional specificity of GABA receptor density distribution did not further improve the model's ability to simulate DOC patients' brain dynamics, beyond the

improvement provided by using a uniform or scrambled GABA-A receptor map. Therefore, whereas propofol anaesthesia is causally mediated by GABA-A receptors and their specific distribution across the cortex, it appears that a global increase in inhibition is sufficient to generate the characteristic dynamics of disorders of consciousness.

Our results from connectome replacement in the whole-brain model informed by empirical anatomical connectivity (dMRI), point to injury-induced randomisation of the connectome as one such candidate mechanism in DOC patients. Specifically, our findings show that (a) unconscious fMRI dynamics (whether due to propofol anaesthesia or brain injury) are more compatible with the empirical DOC connectome, than conscious dynamics; and (b) unconscious dynamics are also more compatible with a random connectome than conscious dynamics (whereas the opposite holds for a lattice-like connectome). It makes sense that the effect of rewiring the network into a lattice should have the opposite effect of randomisation: lattice networks are characterised by high clustering but also high average shortest path between pairs of nodes - the opposite of random networks. Indeed, the well-known "small-world" organisation that characterises many real-world networks, including the human brain [104,105] (but see [106]) is defined as being the topological middle ground between these two extremes of possible network configurations [95,107].

It is also remarkable that the same results from connectome replacement - greater compatibility of unconscious dynamics with the DOC connectome perturbation - could be generalised to the propofol dataset. For DOC patients, such a result may perhaps be expected, since the initial connectome used in the model was obtained from healthy controls, whereas the perturbed DOC connectome was obtained by combining the individual connectomes of the same DOC patients. However, observing the same result in the propofol dataset is a powerful validation of our approach, demonstrating that the results are specific to the presence vs absence of consciousness, rather than being influenced by the specific dataset used.

Thus, thanks to connectome replacement we can infer that the increased neuronal inhibition that characterises both disorders of consciousness and anaesthesia, is functionally equivalent to randomisation of the connectome. However, propofol's anaesthetic effects are mediated by GABA-A receptors according to their specific regional distribution, whereas disorders of consciousness can be explained in terms of a more generic increase in global inhibition - possibly arising from randomisation of the connectome due to anatomical lesions, whose extent

and location are do not follow uniform patterns. Anaesthesia may be expected to operate similarly across individuals, in terms of which regions are more or less affected by propofol. In contrast, each DOC patient is unique in the cause, extent and location of their brain injury. As a result, whereas anaesthesia may depend on specific localised patterns, it stands to reason that the characteristic macroscale dynamics of DOC patients' brains should arise from global-scale neurobiological mechanisms, which may originate from a variety of causes without necessarily depending on specific locations for injury. On the other hand, future research may also explicitly investigate whether lesions leading to disorders of consciousness show an association with regional distribution of GABA-A receptors.

Our results combining fMRI (neural dynamics), dMRI (anatomical connectivity) and PET (neurotransmitter system) demonstrate that human consciousness arises from the delicate balance of local excitation and inhibition, interacting across an intricate network of anatomical connections. Many paths can lead to unconsciousness by disturbing this balance, whether by influencing the nodes' activity (through inhibitory modulation) or the connectivity between them (through connectome randomisation). As befits such a complex dynamical system as the human brain, it is likely that other paths to unconsciousness will also exist, explaining phenomena such as regular sleep-wake alternation, epileptic seizures, and the effects of non-GABAergic anaesthetics such as ketamine – some of which have already started to be explored using whole-brain computational modelling [68,76,78,79,81]. Extending the present framework to account for additional ways of losing consciousness will be a crucial endeavour, informing both our understanding of brain network function and of human consciousness. Likewise, it is vital to combine multimodal neuroimaging and whole-brain modelling to identify paths from unconsciousness back to consciousness, using our understanding of post-anaesthetic recovery to restore consciousness in DOC patients, whether by means of custom-designed drugs or deep brain stimulation [57,58,86].

Limitations

A well-known adage asserts that "All models are wrong", and the present work is no exception. Models of neurobiological function can vary in complexity related to the level of physiological detail and scale, with both aspects incurring costs in terms of computational resources and time. Thus, trade-offs between realism and complexity are unavoidable [35]. At the end of the spectrum

favouring physiological accuracy, microscale spiking models of individual neurons can incorporate thousands of parameters to account for numerous aspects of their neurobiology, but at present such models are not yet able to demonstrate computational scalability and theoretical tractability that could make them useful for the purposes of the present investigation of macroscale brain dynamics [35]. At the other end of the spectrum, important insights about the spatial and temporal organisation of macroscale brain networks can be attained from abstract statistical modelling based on Hidden Markov Models, PCA/ICA, and clustering approaches [17,18,20,23,28,29,108]. Rather than seeking to provide plausible neurobiological mechanisms of empirical brain dynamics, these approaches aim to obtain low-dimensional descriptions of the data (in terms of "metastates", components or clusters), which can lead to conceptually and computationally more tractable accounts that have been proven useful to study questions about the macroscale organisation of the brain.

Note that the same adage also admits that "some models are useful": following insights from statistical mechanics, which has demonstrated how macroscopic phenomena can sometimes be understood when atoms are modelled as aggregates rather than individuals. The present work uses a dynamic mean-field model to reduce the complexity of spiking neuron models to a more tractable set of differential equations that represent the mean activity of macroscopic neuronal ensembles [37,90]. This approach allowed us to simulate realistic whole-brain dynamics in terms of a small number of neurobiologically-informed parameters, thus combining tractability with causal insight.

Nevertheless, we acknowledge that a wide variety of other modelling approaches are possible; indeed, even among DMF models, alternatives have been developed that incorporate additional information about regional neurobiology [109,110] or use different fitting procedures [87]. Likewise, alternative models (e.g. Hopf, Ising) have recently been used to investigate loss of consciousness during sleep [76–81] anaesthesia [78–80,82,84,85] and also disorders of consciousness [80,82,83]. Though less neurobiologically detailed, such models have been able to provide insights about different aspects of brain function, such as criticality and the predicted effects of applying external perturbations to individual regions. Thus, it is clear that further complementary insights may be obtained by considering additional neurobiological mechanisms and multiple levels of explanation - each of which may require a different modelling approach [14,35,58].

We also assumed that the characteristic dynamics of propofol anaesthesia and disorders of consciousness would correspond to loss of consciousness. Although this is a robust finding supported by converging empirical evidence, the assumption may not hold for every state of unconsciousness, and active investigation in the area remains ongoing. Finally, it is important to note that the diversity of disorders of consciousness in terms of aetiology and severity can greatly benefit from an individual-subject approach. Since we considered the cohort of DOC patients as a whole, it is possible that the similar effects of perturbation using a random connectome or the DOC connectome may in fact arise because we obtained a single "DOC connectome" from the combination of several patients, whose individual lesions may be specific but distinct. Likewise, each patient may only exhibit increased inhibition in a specific region, but if such regions differ across patients, then considering them together may result in apparently uniform inhibition. Indeed, analysing such a diverse cohort as a group is always challenging; having demonstrated the efficacy of our modelling approach at the group level, we expect that a fruitful avenue for future research will be to refine our results by considering the specificity of each unique patient condition. In this regard, it is intriguing that some DOC patients can be paradoxically awakened by administration of the drug zolpidem, which is a GABA-ergic agonist [111], which suggests that - at least some patients - the causative neurobiological mechanisms may be substantially different from those identified here based on a group-average DOC connectome. Thus, in future efforts we will apply the frameworks developed here to individual patients, to explore their specific deficits and potential avenues to promote recovery at a finer-grained level.

It is also not just the DOC patients who could benefit from an individualised approach: the GABA-A receptor map used here was obtained from an independent sample of volunteers [94] and therefore the present work could not take into account individual differences in regional GABA-A receptor density distribution. Such differences may however play an important role in explaining individual susceptibility to anaesthesia with GABA-ergic agents - and potentially predict individual risk of experiencing post-anaesthetic complications, such as emergence delirium [112,113]. Investigating this possibility will be an important avenue of future research.

Conclusion

Taken together, our results from a neurobiologically plausible whole-brain computational model demonstrate fundamental similarities, not just between the macroscale brain dynamics that characterise anaesthesia and disorders of consciousness [13,38–42] but also between the neurobiological mechanisms from which they can arise - despite the fact that anaesthesia is a transient pharmacological intervention and DOCs are the result of permanent neuroanatomical injury. Both disorders of consciousness and propofol anaesthesia were shown to arise from neurobiological mechanisms that are functionally equivalent to connectome randomisation, and both involve increased perturbed excitation-inhibition balance, as indicated by incorporating into the model information about regional GABA-A receptor density estimated from PET. However, differences also emerged: a global increase in inhibition suffices to explain the macroscale dynamics that characterise disorders of consciousness, whereas the anaesthetic effects of propofol are mediated by GABA-A receptors according to their specific distribution across cortical regions.

Overall, the present findings begin to unravel the neurobiological mechanisms by which different perturbations of the brain's structure and function - transient pharmacological intervention and chronic neuroanatomical injury - can lead to unconsciousness. Having demonstrated the power of whole-brain computational modelling to address this challenge, the same framework may also prove fruitful to address the reverse problem: namely, how the recovery of consciousness after anaesthesia can inform our ability to restore consciousness in DOC patients.

Materials and Methods

Anaesthesia Data

The propofol data employed in this study have been published before [3,13,114]. For clarity and consistency of reporting, where applicable we use the same wording as our previous study [13].

Recruitment

As previously reported [13], "The propofol data were collected at the Robarts Research Institute in London, Ontario (Canada) between May and November 2014. A total of 19 (18–40 years; 13 males) healthy, right- handed, native English speakers, with no history of neurological disorders were recruited. Each volunteer provided written informed consent, following relevant ethical guidelines, and received monetary compensation for their time. The Health Sciences Research Ethics Board and Psychology Research Ethics Board of Western University (Ontario, Canada) ethically approved this study. Due to equipment malfunction or physiological impediments to anaesthesia in the scanner, data from three participants (1 male) were excluded from analyses, leaving 16" [13].

Procedure

Resting-state fMRI data were acquired at no sedation (Awake), and Deep sedation (anaesthetised: Ramsay score of 5). As previously reported [13]: "Ramsay level was independently assessed by two anaesthesiologists and one anaesthesia nurse in the scanning room before fMRI acquisition began, in each condition. Additionally, participants performed two tests: a computerised auditory target-detection task and a memory test of verbal recall, to evaluate their level of wakefulness independently of the assessors. For the Awake condition, participants did not receive a Ramsey score, as this scale is designed for patients in critical care. Instead, they had to be fully awake, alert and communicating appropriately. An infrared camera located inside the scanner was used to monitor wakefulness. For the Deep sedation condition, propofol was administered intravenously using an AS50 auto syringe infusion pump (Baxter Healthcare, Singapore); step-wise sedation increments sedation were achieved using

an effect-site/plasma steering algorithm combined with the computer-controlled infusion pump. Further manual adjustments were performed as required to reach target concentrations of propofol, as predicted by the TIVA Trainer (European Society for Intravenous Aneaesthesia, eurosiva.eu) pharmacokinetic simulation program. This software also specified the blood concentrations of propofol, following the Marsh 3-compartment model, which were used as targets for the pharmacokinetic model providing target-controlled infusion. The initial propofol target effect-site concentration was 0.6 μg mL^{-1}, with oxygen titrated to maintain SpO2 above 96%. Concentration was then increased by increments of 0.3 μg mL^{-1}, and Ramsay score was assessed: if lower than 5, a further increment occurred. Participants were deemed to have reached Ramsay level 5 once they stopped responding to verbal commands, were unable to engage in conversation, and were rousable only to physical stimulation. Data acquisition began once loss of behavioural responsiveness occurred for both tasks, and the three assessors agreed that Ramsay sedation level 5 had been reached. The mean estimated effect-site and plasma propofol concentrations were kept stable by the pharmacokinetic model delivered via the TIVA Trainer infusion pump; the mean estimated effect-site propofol concentration was 2.48 (1.82-3.14) μg mL^{-1}, and the mean estimated plasma propofol concentration was 2.68 (1.92- 3.44) μg mL^{-1}. Mean total mass of propofol administered was 486.58 (373.30- 599.86) mg. These values of variability are typical for the pharmacokinetics and pharmacodynamics of propofol. At Ramsay 5 sedation level, participants remained capable of spontaneous cardiovascular function and ventilation. However, since the sedation procedure did not take place in a hospital setting, airway security could not be ensured by intubation during scanning, although two anaesthesiologists closely monitored each participant. Consequently, scanner time was minimised to ensure return to normal breathing following deep sedation. No state changes or movement were noted during the deep sedation scanning for any of the participants included in the study" [13].

Design

As previously reported [13]: "In the scanner, subjects were instructed to relax with closed eyes, without falling asleep; 8 minutes of fMRI scan without any task ("resting-state") were acquired for each participant. Additionally, a separate 5-minute long scan was also acquired while a plot-driven story was presented through headphones to participants, who were instructed to listen while keeping their eyes closed" [13]. The present analysis focuses on the resting-state data

only; the story scan data have been published separately [84] and will not be discussed further here.

Data Acquisition

As previously reported [13]: "MRI scanning was performed using a 3-Tesla Siemens Tim Trio scanner (32-channel coil), and 256 functional volumes (echo-planar images, EPI) were collected from each participant, with the following parameters: slices = 33, with 25% inter-slice gap; resolution = 3mm isotropic; TR = 2000ms; TE = 30ms; flip angle = 75 degrees; matrix size = 64x64. The order of acquisition was interleaved, bottom-up. Anatomical scanning was also performed, acquiring a high-resolution T1- weighted volume (32-channel coil, 1mm isotropic voxel size) with a 3D MPRAGE sequence, using the following parameters: TA = 5min, TE = 4.25ms, 240x256 matrix size, 9 degrees FA" [13].

Disorders of Consciousness Patient Data

The DOC patient functional data employed in this study have been published before [13,41,115]. For clarity and consistency of reporting, where applicable we use the same wording as our previous study [13].

Recruitment

As previously reported [13]: "A sample of 71 DOC patients was included in this study. Patients were recruited from specialised long-term care centres. To be invited to the study, patients must have had a DOC diagnosis, written informed consent to participation from their legal representative, and were capable of being transported to Addenbrooke's Hospital. The exclusion criteria included any medical condition that made it unsafe for the patient to participate (decision made by clinical personnel blinded to the specific aims of the study) or any reason they are unsuitable to enter the MRI scanner environment (e.g. non-MRI-safe implants), significant pre-existing mental health problems, or insufficient English pre injury. After admission, each patient underwent clinical and neuroimaging testing. Patients spent a total of five days (including arrival and departure days) at Addenbrooke's Hospital. Coma Recovery Scale-Revised (CRS-R) assessments were recorded at least daily for the five days of admission. If behaviours were indicative of awareness at any time, patients were classified as

MCS; otherwise UWS. We assigned MCS- or MCS+ sub-classification if behaviours were consistent throughout the week. The most frequent signs of consciousness in MCS- patients are visual fixation and pursuit, automatic motor reactions (e.g. scratching, pulling the bed sheet) and localisation to noxious stimulation whereas MCS+ patients may, in addition, follow simple commands, intelligibly verbalise or intentionally but inaccurately communicate [53,54]. Scanning occurred at the Wolfson Brain Imaging Centre, Addenbrooke's Hospital, between January 2010 and December 2015; medication prescribed to each patient was maintained during scanning. Ethical approval for testing patients was provided by the National Research Ethics Service (National Health Service, UK; LREC reference 99/391). All clinical investigations were conducted in accordance with the Declaration of Helsinki. As a focus of this study was on graph-theoretical properties of the brain, patients were systematically excluded from the final cohort analysed in this study based on the following criteria: 1) large focal brain damage (i.e. more than 1/3 of one hemisphere) as stated by an expert in neuroanatomy blinded to the patients' diagnoses; 2) excessive head motion during resting state scanning (i.e. greater than 3mm in translation and/or 3 degrees in rotation); 3) suboptimal segmentation and normalization of images; 4) incomplete brain acquisition. A total of 21 adults (13 males; 17-70 years; mean time post injury: 13 months) meeting diagnostic criteria for Unresponsive Wakefulness Syndrome/Vegetative State or Minimally Conscious State due to brain injury were included in this study." [13] (Table 5).

Table 5: Demographic information for patients with Disorders of Consciousness.

Sex	Age	Months post injury	Aetiology	Diagnosis	CRS-R Score	Scan
M	46	23	TBI	UWS	6	12 dir
M	57	14	TBI	MCS-	12	12 dir
M	35	34	Anoxic	UWS	8	12 dir
M	17	17	Anoxic	UWS	8	12 dir
F	31	9	Anoxic	MCS-	10	12 dir
F	38	13	TBI	MCS	11	12 dir
M	29	68	TBI	MCS	10	63 dir
M	23	4	TBI	MCS	7	63 dir
F	70	11	Cerebral bleed	MCS	9	63 dir

F	30	6	Anoxic	MCS-	9	63 dir
F	36	6	Anoxic	UWS	8	63 dir
M	22	5	Anoxic	UWS	7	63 dir
M	40	14	Anoxic	UWS	7	63 dir
F	62	7	Anoxic	UWS	7	63 dir
M	46	10	Anoxic	UWS	5	63 dir
M	21	7	TBI	MCS	11	63 dir
M	67	14	TBI	MCS-	11	63 dir
F	55	6	Hypoxia	UWS	12	63 dir
M	28	14	TBI	MCS	8	63 dir
M	22	12	TBI	MCS	10	63 dir
F	28	8	ADEM	UWS	6	63 dir

CRS-R, Coma Recovery Scale-Revised; UWS, Unresponsive Wakefulness Syndrome; MCS, Minimally Conscious State; TBI, Traumatic Brain Injury.

FMRI Data Acquisition

As previously reported [13]: "Resting-state fMRI was acquired for 10 minutes (300 volumes, TR=2000ms) using a Siemens Trio 3T scanner (Erlangen, Germany). Functional images (32 slices) were acquired using an echo planar sequence, with the following parameters: 3 x 3 x 3.75mm resolution, TR = 2000ms, TE = 30ms, 78 degrees FA. Anatomical scanning was also performed, acquiring high-resolution T1-weighted images with an MPRAGE sequence, using the following parameters: TR = 2300ms, TE = 2.47ms, 150 slices, resolution 1 x 1 x 1mm".

Acquisition of Diffusion-Weighted Data

As the data were acquired over the course of several years, two different diffusion-weighted image acquisition schemes were used for the DOC patients. The first acquisition scheme (used for the N=7 patients whose data were acquired earliest in time) used an echo planar sequence (TR = 8300 ms, TE = 98 ms, matrix size = 96 x 96, 63 slices, slice thickness = 2 mm, no gap, flip angle = 90°). This included diffusion sensitising gradients applied along 12 non-collinear directions with 5 b-values that ranged from 340 to 1590 s/mm2 and 5 b = 0 images. One of

these patients only had partial brain coverage for the diffusion MRI acquisition, and their DWI data were excluded from further analysis. The more recent acquisition scheme (used for the more recently scanned patients, and for all healthy controls) instead involved the use of 63 directions with a b-value of 1000 s/mm2. Both DWI acquisition types have been used before in the context of structural connectivity analysis in DOC patients [51,53].

Healthy controls

We also acquired diffusion MRI data from N=20 healthy volunteers (13 males; 19-57 years), with no history of psychiatric or neurological disorders. The mean age was not significantly different between healthy controls (M = 35.75; SD = 11.42) and DOC patients (M = 38.24; SD = 15.96) ($t(39) = -0.57$, $p = 0.571$, Hedges's $g = -0.18$; permutation-based t-test).

FMRI Data Acquisition

Resting-state fMRI was acquired for 5:20 minutes (160 volumes, TR=2000ms) using a Siemens Trio 3T scanner (Erlangen, Germany). The acquisition parameters were the same as those for the DOC patients: Functional images (32 slices) were acquired using an echo planar sequence, with the following parameters: 3 x 3 x 3.75mm resolution, TR = 2000ms, TE = 30ms, 78 degrees FA. High-resolution T1-weighted anatomical images were also acquired, using an MPRAGE sequence with the following parameters: TR = 2300ms, TE = 2.47ms, 150 slices, resolution 1 x 1 x 1mm. Two subjects had incomplete fMRI acquisition, leaving N=18 subjects for the final analysis.

Acquisition of diffusion-weighted imaging data

The diffusion-weighted acquisition scheme was the same 63-directions scheme used for the DOC patients: TR = 8300 ms, TE = 98 ms, matrix size = 96 x 96, 63 slices, slice thickness = 2 mm, no gap, flip angle = 90°, 63 directions with a b-value of 1000 s/mm2.

Functional MRI preprocessing and denoising

Following our previous work [13], we preprocessed the functional imaging data using a standard pipeline, implemented within the SPM12-based (http://www.fil.ion.ucl.ac.uk/spm) toolbox CONN (http://www.nitrc.org/projects/conn), version 17f [116]. As described, "The pipeline comprised the following steps: removal of the first five scans, to allow magnetisation to reach steady state; functional realignment and motion correction; slice-timing correction to account for differences in time of acquisition between slices; identification of outlier scans for subsequent regression by means of the quality assurance/artifact rejection software *Artifact Detection Toolbox (art;* (http://www.nitrc.org/projects/artifact_detect); spatial normalisation to Montreal Neurological Institute (MNI-152) standard space with 2mm isotropic resampling resolution, using the segmented grey matter image from each volunteer's high-resolution T1-weighted image, together with an *a priori* grey matter template" [13].

For the DOC patients, due to the presence of deformations caused by brain injury, rather than relying on automated pipelines, each patient's brain was individually preprocessed using SPM12, with visual inspections after each step. To further reduce potential movement artifacts, data underwent despiking with a hyperbolic tangent squashing function. Since the controls had a shorter scan duration than DOC patients, to ensure comparability between the two cohorts the DOC functional scans were truncated to be of the same length as the control subjects' functional scans (after removal of the initial scans).

To reduce noise due to cardiac and motion artifacts, we applied the anatomical CompCor method of denoising the functional data [117]. The anatomical CompCor method (also implemented within the CONN toolbox) involves regressing out of the functional data the following confounding effects: the first five principal components attributable to each individual's white matter signal, and the first five components attributable to individual cerebrospinal fluid (CSF) signal; six subject-specific realignment parameters (three translations and three rotations) as well as their first- order temporal derivatives; the artifacts identified by *art;* and main effect of scanning condition [117]. Linear detrending was also applied, and the subject-specific denoised BOLD signal timeseries were band-pass filtered to eliminate both low-frequency drift effects and high-frequency noise, thus retaining frequencies between 0.008 and 0.09 Hz.

DWI Preprocessing and tractography

The diffusion data were preprocessed with MRtrix3 tools [118]. After manually removing diffusion-weighted volumes with substantial distortion [53], the pipeline involved the following steps: (i) DWI data denoising by exploiting data redundancy in the PCA domain [119] (*dwidenoise* command); (ii) Correction for distortions induced by eddy currents and subject motion by registering all DWIs to b0, using FSL's *eddy* tool (through MRtrix3 *dwipreproc* command); (iii) rotation of the diffusion gradient vectors to account for subject motion estimated by *eddy* [120]*;* (iv) b1 field inhomogeneity correction for DWI volumes (*dwibiascorrect* command); (v) generation of a brain mask through a combination of MRtrix3 *dwi2mask* and FSL *BET* commands.

DTI data were reconstructed from the preprocessed DWIs using DSI Studio (www.dsi-studio.labsolver.org), which implements q-space diffeomorphic reconstruction (QSDR [121]), an established methodology to investigate structural networks in DOC patients [54]. As explained in the original publication [121], "QSDR is a model-free method that calculates the orientational distribution of the density of diffusing water in a standard space, to conserve the diffusible spins and preserve the continuity of fiber geometry for fiber tracking. QSDR first reconstructs diffusion-weighted images in native space and computes the quantitative anisotropy (QA) in each voxel [121]. These QA values are used to warp the brain to a template QA volume in Montreal Neurological Institute (MNI) space using the statistical parametric mapping (SPM) nonlinear registration algorithm. Once in MNI space, spin density functions (SDFs) were again reconstructed with a mean diffusion distance of 1.25 mm using three fiber orientations per voxel" [121].

Following previous work [122], "After reconstruction with QSDR, we used deterministic fiber tracking with a high-performing "FACT" algorithm with 1,000,000 streamlines, to identify the connections between brain regions, following previously established parameters [122,123]: angular cutoff = 55◦, step size = 1.0 mm, tract length between 10mm (minimum) and 400mm (maximum), no spin density function smoothing, and QA threshold determined by DWI signal in the cerebro-spinal fluid. DSI Studio automatically applies a default anisotropy threshold of 0.6 Otsu's threshold to the anisotropy values of the spin density function, in order to generate a white matter mask which is then used for automatic screening of each streamline, to exclude streamlines improper termination locations [122,123]".

Brain parcellation

For both BOLD and DWI data, brains were parcellated into 68 cortical regions of interest (ROIs), according to the Desikan-Killiany anatomical atlas [124], in line with previous whole-brain modelling work [109].

Functional Connectivity Dynamics

Following [86], functional connectivity dynamics (FCD) were quantified in terms of Pearson correlation between regional BOLD timeseries, computed within a sliding window of 30 TRs with increments of 3 TRs. Subsequently, the resulting matrices of functional connectivity at times t_x and t_y were themselves correlated, for each pair of timepoints t_x and t_y, thereby obtaining an FCD matrix of time-versus-time correlations. Thus, each entry in the FCD matrix represents the similarity between functional connectivity patterns at different points in time.

Group Structural Connectivity

The structural connectivity (SC) for the DMF model was obtained by following the procedure described in [109] to derive a group-consensus structural connectivity matrix. Separately for the healthy controls and DOC patients, a consensus matrix C was obtained as follows. For each pair of regions i and j, if more than half of subjects had non-zero connection i and j, C_{ij} was set to the average across all subjects with non-zero connections between i and j. Otherwise, C_{ij} was set to zero.

Whole-brain computational modelling

Whole-brain spontaneous brain activity (as quantified using blood oxygen level dependent (BOLD) signal data from functional MRI) was simulated using a neurobiologically realistic Dynamic Mean Field (DMF) model. The DMF model [31,34,37] uses an empirically validated

mathematical mean-field approach to represent the collective behaviour of integrate-and-fire neurons by means of coupled differential equations, providing a neurobiologically plausible account of regional neuronal firing rate.

Specifically, the model simulates local biophysical dynamics of excitatory (NMDA) and inhibitory (GABA) neuronal populations, interacting over long-range neuroanatomical connections (white matter tracts obtained from diffusion MRI). The model further incorporates multimodal neuroimaging information about empirical brain dynamics (measured using functional MRI) and neurotransmitter receptor density, estimated from positron emission tomography (PET) [86].

Each cortical area *n* (defined by a parcellation scheme) is represented in terms of two reciprocally coupled neuronal masses, one excitatory and the other inhibitory, with the synaptic connections between excitatory neuronal populations in different regions given by the weight of structural connectivity, to account for the number and density of interregional axon fibers. Additional factors that can influence the long-range excitatory-to-excitatory coupling between brain regions, such as neurotransmission but also synaptic plasticity mechanisms, are accounted for by a global coupling parameter, *G*. Since conductivity of the white matter fibers is assumed to be constant across the brain, *G* constitutes the only free parameter in the model.

The following differential equations therefore govern the model's behaviour:

$$I_n^{(E)} = W_E I_0 + w_+ J_{NMDA} S_n^{(E)} + G J_{NMDA} \sum_p^N C_{np} S_p^{(E)} - J_n^{FIC} S_n^{(I)}$$

$$I_n^{(I)} = W_I I_0 + J_{NMDA} S_n^{(E)} - S_n^{(I)}$$

$$r_n^{(E)} = F\left(I_n^{(E)}\right) = \frac{g_E\left(I_n^{(E)} - I_{thr}^{(E)}\right)}{1 - exp\left(-d_E g_E\left(I_n^{(E)} - I_{thr}^{(E)}\right)\right)}$$

$$r_n^{(I)} = F\left(I_n^{(I)}\right) = \frac{g_{NM} g_I\left(I_n^{(I)} - I_{thr}^{(I)}\right)}{1 - exp\left(-d_I g_{NM} g_I\left(I_n^{(I)} - I_{thr}^{(I)}\right)\right)}$$

$$\frac{dS_n^{(E)}(t)}{dt} = \frac{S_n^{(E)}}{\tau_{NMDA}} + \left(1 + S_n^{(E)}\right)\gamma r_n^{(E)} + \sigma \nu_n(t)$$

$$\frac{dS_n^{(I)}(t)}{dt} = \frac{S_n^{(I)}}{\tau_{GABA_A}} + r_n^{(I)} + \sigma\nu_n(t)$$

$$g_n^{NM} = 1 + s_I d_n^{GABA}$$

Following previous work [86,88], "for each excitatory *(E)* and inhibitory *(I)* neural mass, the quantities $I_n^{(E.I)}$, $r_n^{(E,I)}$, and $S_n^{(E,I)}$ represent its total input current (nA), firing rate (Hz) and synaptic gating variable, respectively. The function F(·) is the transfer function (or *F–I curve*), representing the non-linear relationship between the input current and the output firing rate of a neural population. Finally, J_n^{FIC} is the local feedback inhibitory control of region *n*, which is optimized to keep its average firing rate at approximately 3Hz [37,88], and ν_n is uncorrelated Gaussian noise injected to region *n*". The model's fixed parameters are reported in Table 6 [37,86,88]. Additionally, g_n^{NM} is the *neuromodulatory scaling factor* modulating the transfer function for each cortical region in the model as a function of d_n^{GABA}, the regional density of GABA-A receptors (see below for details) and an inhibitory gain scaling parameter s_I. The *baseline* model (corresponding to a DMF model without GABA-A regional inhibitory neuromodulation) is obtained by setting s_I to zero, in which case *G* remains the sole free parameter in the model. Details for optimisation of the s_I parameter for the GABA-A modulated model are provided below.

Table 6. Dynamic Mean Field model parameters from [37,86,88].

Parameter	Symbol	Value
External current	I_0	0.382 nA
Excitatory scaling factor for I_0	W_E	1
Inhibitory scaling factor for I_0	W_I	0.7
Local excitatory recurrence	w_+	1.4
Excitatory synaptic coupling	J_{NMDA}	0.15 nA
Threshold for F ($I_n^{(E)}$)	$I_{thr}^{(E)}$	0.403 nA

Threshold for F ($I_n^{(I)}$)	$I_{thr}^{(I)}$	0.288 nA
Gain factor of F ($I_n^{(E)}$)	g_E	310 nC 1
Gain factor of F ($I_n^{(I)}$)	g_I	615 nC 1
Shape of F ($I_n^{(E)}$) around $I_{thr}^{(E)}$	dE	0.16 s
Shape of F ($I_n^{(I)}$) around $I_{thr}^{(I)}$	dI	0.087 s
Excitatory kinetic parameter	γ	0.641
Amplitude of uncorrelated Gaussian noise v_n	σ	0.01 nA
Time constant of NMDA	τ_{NMDA}	100 ms
Time constant of GABA	τ_{GABA_A}	10 ms

A Balloon-Windkessel (BW) hemodynamic model [92] was then used to turn simulated regional neuronal activity into simulated regional BOLD signal. The Balloon-Windkessel model considers the BOLD signal as a nonlinear function of the normalized total deoxyhemoglobin voxel content, normalized venous volume, resting net oxygen extraction fraction by the capillary bed, and resting blood volume fraction. The BOLD-signal estimation for each brain area is computed from the level of neuronal activity in that particular area. Finally, simulated regional BOLD signal was bandpass filtered in the same range as the empirical data (0.008-0.09 Hz).

Implementation

The code used to run all the simulations in this study was written in optimised C++ using the high-performance library Eigen. The C++ core of the code, together with Python and Octave/Matlab interfaces is publicly available at http://www.gitlab.com/concog/fastdmf.

To simulate BOLD data, FastDMF splits the problem in two steps: integrating the coupled differential equations underlying the DMF model, to obtain excitatory firing rates in each brain region; and using these firing rates to integrate the (uncoupled) differential equations of the BW hemodynamic model and obtain BOLD timeseries.

Integration of the DMF equations is performed with the Euler-Maruyama method, and it is highly parallelizable and bounded by the O(N^2) complexity of the matrix-vector multiplication corresponding to the excitatory-to-excitatory coupling between brain regions.

Simulated excitatory firing rates are stored in a cylindrical array with a fixed buffer size to limit memory requirements.

In addition, a further set of threads is spawned to solve the BW model using the simulated excitatory firing rates. Since the BW solver reads from the same cylindrical array, it interfaces with the DMF solver with a controlled multi-threaded architecture. Every TR-equivalent in simulation time the value of all BOLD signals is copied to a pre-allocated array, to be returned at the end of the requested simulation time.

In a standard laptop, FastDMF attains a speed-up of between 5x and 10x over publicly available Matlab implementations, due to the speed of Eigen and the parallelisation of DMF and BW solvers. In addition, due to the cylindrical buffer, this implementation is able to simulate arbitrarily long BOLD time series with a fixed memory overhead, thereby allocating orders of magnitude less memory than a naive Matlab implementation.

Finally, the library includes interface functions for Matlab (via its C Matrix API) and Python (via the Boost.Python library). In both languages the function returns a standard array (numpy.ndarray in the case of Python) that can be easily processed for further analysis.

Fitting of the G parameter

In order to identify appropriate parameters for the simulations, early whole-brain modelling efforts used the grand average FC as target for fitting the model to empirical data. However, it has since become apparent that the macroscale neural signals measured by functional MRI are not static, even on the timescale of a few tens of seconds: rather, they exhibit a wide range of dynamics. Therefore, in order to properly take into account the time-dependencies of FC, it is advantageous to fit the model to empirical functional connectivity dynamics (FCD). Doing so ensures that the simulated BOLD data will exhibit realistic patterns of time-evolving functional connectivity [86,91].

Unlike matrices of inter-regional connectivity, where each brain region is the same across different scans, FCDs are represented as matrices encoding the relationship between brain dynamics at different timepoints. Since timepoints are not the same across individuals or scans, FCD matrices cannot be compared by means of simple correlation. Therefore, to evaluate

model performance in terms of producing meaningful temporal dynamics, here we follow the approach of [86], using the Kolmogorov-Smirnov distance to compare the histograms of empirical and simulated FCD values (obtained from the upper triangular FCD matrix), to find the G parameter that results in the best match between empirical and simulated functional connectivity dynamics.

To find the value of *G* that generates simulations whose FCD best match empirical FCD, we generated 100 simulations for each value of *G* between 0.1 and 2.5, using increments of 0.1. For each simulation at each value of G, we computed the KS distance between empirical (group-wise) and simulated FCD. Finally, we set the model's G parameter to the value that minimised the mean KS distance - corresponding to the model that is best capable of simulating the temporal dynamics of functional connectivity observed in the healthy human brain at rest.

This procedure was performed separately for the propofol dataset (with 250 TRs) and the DOC dataset, which was truncated to the number of TRs available for the healthy controls (155 TRs).

Local inhibitory gain modulation from GABA-A maps

We modulated local inhibitory gain based on the recent high-resolution quantitative atlas of human brain GABA-A receptors, generated on the basis of benzodiazepine receptor (BZR) density measured from [^{11}C]flumazenil Positron Emission Tomography (PET) autoradiography [94]. Based on this atlas, we obtained a quantitative measure of GABA-A receptor density for each region of the Desikan-Killiany cortical parcellation. Following [86] regional density values were normalised between 0 and 1 by dividing each by the maximum value.

We then used the previously calibrated DMF model to generate simulations for each value of s_I between 0 (corresponding to the model without local GABA-A inhibitory modulation, i.e. the original DMF model) and 1, in increments of 0.02. Then, for each value of s_I, we computed the KS distance between the model's simulated macroscale dynamics and the empirical dynamics observed in each condition (baseline or propofol, control or DOC). For each condition, the optimal value of s_I, was then identified as the value that resulted in the minimum mean KS distance between empirical and simulated dynamics (across N=10 simulations for each value of s_I).

As validation analysis, we also repeated the same procedure, optimising the inhibitory gain scaling s_I, but with two different kinds of receptor density maps: a "scrambled" map, whereby the values of GABA-A receptor density obtained from PET were randomised across regions; and a "uniform" map, whereby each region was set to the same value, corresponding to the mean of the distribution of PET-derived receptor densities.

Connectome Replacement

Connectome replacement was performed using the initial balanced DMF model (i.e. with optimised *G* parameter, but without additional inhibitory gain modulation), based on the consensus connectome from diffusion imaging of healthy controls (referred to as the "healthy connectome").

Three perturbed connectomes were used. Firstly, the consensus connectome obtained from diffusion imaging of N=21 DOC patients, referred to as the "DOC connectome". Secondly, the original healthy connectome was randomised according to the weight-preserving procedure of [95] to generate a "random connectome" that differs from the original in terms of topology, but preserves the weight distribution. Thirdly, we used the procedure described in [95] to turn the healthy connectome into a lattice network with the same weight distribution - providing a different and opposite perturbation of the network's topology.

For each perturbed connectome, the DMF model was used to generate 100 simulations, using the optimal global coupling *G*, but with inter-regional connectivity given by the perturbed connectome rather than the original connectome. This was repeated for each dataset (propofol and DOC) and the resulting simulations were compared with each condition (baseline/propofol and control/DOC) in terms of KS-distance.

Statistical analysis

To test the effect of the various procedures over the DMF model, we performed linear regression modelling using the KS distance as target variable, while considering model and target condition as predictor variables. For each scenario, models with and without interactions between model and target condition were constructed; then, selection between these models

was conducted according to the Akaike information criterion (AIC). All analyses were conducted using RStudio (Version 1.3.1093; http://www.rstudio.com/).

Acknowledgements

The authors would like to thank all the participants for their contribution to this study. This work was supported by grants from the UK Medical Research Council [U.1055.01.002.00001.01 to AMO and JDP]; The James S. McDonnell Foundation [to AMO and JDP]; and the Canada Excellence Research Chairs program (215063 to AMO); the National Institute for Health Research (NIHR, UK), Cambridge Biomedical Research Centre and NIHR Senior Investigator Awards [to DKM], the Stephen Erskine Fellowship (Queens' College, Cambridge, to EAS), the L'Oreal-Unesco for Women in Science Excellence Research Fellowship to LN; the British Oxygen Professorship of the Royal College of Anaesthetists [to DKM] the Gates Cambridge Trust (to AIL) and the Vice-Chancellor Award (to PC). PAM and DB are funded by the Wellcome Trust (grant no. 210920/Z/18/Z). FR is funded by the Ad Astra Chandaria foundation. The research was also supported by the NIHR Brain Injury Healthcare Technology Co-operative based at Cambridge University Hospitals NHS Foundation Trust and University of Cambridge. AMO and DKM are Fellows of the CIFAR Brain, Mind, and Consciousness Programme.

Author Contributions

AIL: conceived the study; analysed data; wrote first draft of the manuscript. PAM: conceived the study; contributed to data analysis and interpretation of results; reviewed and edited the manuscript. FR: conceived the study; contributed to data analysis and interpretation of results; reviewed and edited the manuscript. M.M.C., P.C., A.R.D.P: contributed to data analysis. DKM: reviewed the manuscript. DB: reviewed and edited the manuscript. EAS: conceived the study; reviewed and edited the manuscript. P.F., G.B.W., J.A., J.D.P., A.M.O., L.N., D.K.M. and E.A.S. were involved in designing the original studies for which the present data were collected. P.F., M.M.C., G.B.W., J.A., L.N. and E.A.S. all participated in data collection.

Competing Interests

The authors declare no competing interests.

Data and Code Availability

The CONN toolbox is freely available online (http://www.nitrc.org/projects/conn).

The C++ core of the DMF code, together with Python and Octave/Matlab interfaces is publicly available at http://www.gitlab.com/concog/fastdmf.

The propofol and DOC patient data that support the findings of this study are available from Dr. Emmanuel Stamatakis, University of Cambridge (email: eas46@cam.ac.uk) upon reasonable request.

References

1. Dehaene, S., Lau, H. & Kouider, S. What is consciousness, and could machines have it? *Science (80-.).* **358**, 486–492 (2017).
2. Koch, C., Massimini, M., Boly, M. & Tononi, G. Neural correlates of consciousness: progress and problems. *Nat. Rev. Neurosci.* **17**, 307–321 (2016).
3. Varley, T. F. *et al.* Consciousness & Brain Functional Complexity in Propofol Anaesthesia. *Sci. Rep.* **10**, (2020).
4. Kelz, M. B. & Mashour, G. A. The Biology of General Anesthesia from Paramecium to Primate. *Current Biology* **29**, R1199–R1210 (2019).
5. Mashour, G. A. & Hudetz, A. G. Bottom-Up and Top-Down Mechanisms of General Anesthetics Modulate Different Dimensions of Consciousness. *Front. Neural Circuits* **11**, 44 (2017).
6. Hudetz, A. G. & Mashour, G. A. Disconnecting Consciousness: Is There a Common Anesthetic End Point? *Anesth. Analg.* **123**, 1228–1240 (2016).
7. Hemmings, H. C. *et al.* Towards a Comprehensive Understanding of Anesthetic Mechanisms of Action: A Decade of Discovery. *Trends in Pharmacological Sciences* **40**, 464–481 (2019).
8. Hannawi, Y., Lindquist, M. A., Caffo, B. S., Sair, H. I. & Stevens, R. D. Resting brain activity in disorders of consciousness: a systematic review and meta-analysis. *Neurology* **84**, 1272–1280 (2015).
9. Giacino, J. T., Fins, J. J., Laureys, S. & Schiff, N. D. Disorders of consciousness after acquired brain injury: The state of the science. *Nature Reviews Neurology* **10**, 99–114 (2014).
10. Cavaliere, C. *et al.* Diffusion tensor imaging and white matter abnormalities in patients with disorders of consciousness. *Front. Hum. Neurosci.* **8**, 6–12 (2015).
11. Billeri, L. *et al.* Toward improving diagnostic strategies in chronic disorders of consciousness: An overview on the (re-)emergent role of neurophysiology. *Brain Sci.* **10**, (2020).
12. Song, M., Zhang, Y., Cui, Y., Yang, Y. & Jiang, T. Brain Network Studies in Chronic Disorders of Consciousness: Advances and Perspectives. *Neuroscience Bulletin* **34**, 592–604 (2018).

13. Luppi, A. I. *et al.* Consciousness-specific dynamic interactions of brain integration and functional diversity. *Nat. Commun.* **10**, (2019).
14. Cofré, R. *et al.* Whole-brain models to explore altered states of consciousness from the bottom up. *Brain Sci.* **10**, 1–29 (2020).
15. Atasoy, S., Donnelly, I. & Pearson, J. Human brain networks function in connectome-specific harmonic waves. *Nat. Commun.* **7**, (2016).
16. Atasoy, S., Deco, G., Kringelbach, M. & Pearson, J. Harmonic Brain Modes: A Unifying Framework for Linking Space and Time in Brain Dynamic. *Neurosci.* (2018). doi:10.1177/1073858417728032
17. Shine, J. M. *et al.* The Dynamics of Functional Brain Networks: Integrated Network States during Cognitive Task Performance. *Neuron* **92**, 544–554 (2016).
18. Shine, J. M. *et al.* Human cognition involves the dynamic integration of neural activity and neuromodulatory systems. *Nat. Neurosci.* **22**, 289–296 (2019).
19. Fukushima, M. *et al.* Structure–function relationships during segregated and integrated network states of human brain functional connectivity. *Brain Struct. Funct.* **223**, 1091–1106 (2018).
20. Allen, E. A. *et al.* Tracking whole-brain connectivity dynamics in the resting state. *Cereb. Cortex* **24**, 663–676 (2014).
21. Preti, M. G., Bolton, T. A. & Van De Ville, D. The dynamic functional connectome: State-of-the-art and perspectives. *Neuroimage* **160**, 41–54 (2017).
22. Hutchison, R. M., Hutchison, M., Manning, K. Y., Menon, R. S. & Everling, S. Isoflurane induces dose-dependent alterations in the cortical connectivity profiles and dynamic properties of the brain's functional architecture. *Hum. Brain Mapp.* **35**, 5754–5775 (2014).
23. Hutchison, R. M. *et al.* Dynamic functional connectivity: Promise, issues, and interpretations. *Neuroimage* **80**, 5–79 (2013).
24. Barttfeld, P. *et al.* Signature of consciousness in the dynamics of resting-state brain activity. *Proc. Natl. Acad. Sci.* **112**, 887–892 (2015).
25. Atasoy, S., Vohryzek, J., Deco, G., Carhart-harris, R. L. & Kringelbach, M. L. Common neural signatures of psychedelics: Frequency-specific energy changes and repertoire expansion revealed using connectome-harmonic decomposition. *Prog. Brain Res.* **242**, (2018).
26. Atasoy, S. *et al.* Connectome-harmonic decomposition of human brain activity reveals dynamical repertoire re-organization under LSD. *Sci. Rep.* **7**, 1–18 (2017).
27. Lord, L. D. *et al.* Dynamical exploration of the repertoire of brain networks at rest is modulated by psilocybin. *Neuroimage* **199**, 127–142 (2019).
28. Vohryzek, J., Deco, G., Cessac, B., Kringelbach, M. L. & Cabral, J. Ghost Attractors in Spontaneous Brain Activity: Recurrent Excursions Into Functionally-Relevant BOLD Phase-Locking States. *Front. Syst. Neurosci.* **14**, 20 (2020).
29. Vidaurre, D., Smith, S. M. & Woolrich, M. W. Brain network dynamics are hierarchically organized in time. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 12827–12832 (2017).
30. Rosas, F., Mediano, P. A. M., Ugarte, M. & Jensen, H. J. An Information-Theoretic Approach to Self-Organisation: Emergence of Complex Interdependencies in Coupled Dynamical Systems. (2018). doi:10.3390/e20100793
31. Deco, G. & Jirsa, V. K. Ongoing cortical activity at rest: Criticality, multistability, and ghost attractors. *J. Neurosci.* **32**, 3366–3375 (2012).
32. Deco, G., Kringelbach, M. L., Jirsa, V. K. & Ritter, P. The dynamics of resting fluctuations in the brain: Metastability and its dynamical cortical core. *Sci. Rep.* **7**, (2017).
33. Deco, G., Jirsa, V. K. & McIntosh, A. R. Emerging concepts for the dynamical organization of resting-state activity in the brain. *Nature Reviews Neuroscience* **12**, 43–56 (2011).
34. Deco, G. *et al.* Resting-state functional connectivity emerges from structurally and dynamically shaped slow linear fluctuations. *J. Neurosci.* **33**, 11239–11252 (2013).
35. Cabral, J., Kringelbach, M. L. & Deco, G. Functional connectivity dynamically evolves on multiple time-scales over a static structural connectome: Models and mechanisms. *Neuroimage* **160**, 84–96 (2017).
36. Honey, C. J., Kötter, R., Breakspear, M. & Sporns, O. Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proc. Natl. Acad. Sci. U. S. A.* **104**,

10240–10245 (2007).
37. Deco, G. *et al.* How local excitation-inhibition ratio impacts the whole brain dynamics. *J. Neurosci.* **34**, 7886–7898 (2014).
38. Demertzi, A. *et al.* Human consciousness is supported by dynamic complex patterns of brain signal coordination. *Sci. Adv.* **5**, 1–12 (2019).
39. Huang, Z., Zhang, J., Wu, J., Mashour, G. A. & Hudetz, A. G. Temporal circuit of macroscale dynamic brain activity supports human consciousness. *Sci. Adv.* **6**, 87–98 (2020).
40. Golkowski, D. *et al.* Changes in Whole Brain Dynamics and Connectivity Patterns during Sevoflurane- and Propofol-induced Unconsciousness Identified by Functional Magnetic Resonance Imaging. *Anesthesiology* **130**, 898–911 (2019).
41. Luppi, A. I. *et al.* Connectome Harmonic Decomposition of Human Brain Dynamics Reveals a Landscape of Consciousness. *bioRxiv* (2020). doi:10.1101/2020.08.10.244459
42. Campbell, J. M. *et al.* Pharmacologically informed machine learning approach for identifying pathological states of unconsciousness via resting-state fMRI. *Neuroimage* **206**, 116316 (2020).
43. Noirhomme, Q. *et al.* Brain Connectivity in Pathological and Pharmacological Coma. *Front. Syst. Neurosci.* **4**, (2010).
44. Huang, Z. *et al.* Decoupled temporal variability and signal synchronization of spontaneous brain activity in loss of consciousness: An fMRI study in anesthesia. *Neuroimage* **124**, 693–703 (2016).
45. Northoff, G., Wainio-Theberge, S. & Evers, K. Is temporo-spatial dynamics the “common currency” of brain and mind? In Quest of “Spatiotemporal Neuroscience”. *Physics of Life Reviews* **33**, 34–54 (2020).
46. Yip, G. M. S. *et al.* A propofol binding site on mammalian GABA A receptors identified by photolabeling. *Nat. Chem. Biol.* **9**, 715–720 (2013).
47. Jurd, R. *et al.* General anesthetic actions *in vivo* strongly attenuated by a point mutation in the $GABA_A$ receptor β3 subunit. *FASEB J.* **17**, 250–252 (2003).
48. Lant, N. D., Gonzalez-Lara, L. E., Owen, A. M. & Fernández-Espejo, D. Relationship between the anterior forebrain mesocircuit and the default mode network in the structural bases of disorders of consciousness. *NeuroImage Clin.* **10**, 27–35 (2016).
49. Newcombe, V. F. J. *et al.* Aetiological differences in neuroanatomy of the vegetative state: Insights from diffusion tensor imaging and functional implications. *J. Neurol. Neurosurg. Psychiatry* **81**, 552–561 (2010).
50. Weng, L. *et al.* Abnormal structural connectivity between the basal ganglia, thalamus, and frontal cortex in patients with disorders of consciousness. *Cortex* **90**, 71–87 (2017).
51. Wang, L. *et al.* White matter integrity correlates with residual consciousness in patients with severe brain injury. *Brain Imaging Behav.* **12**, 1669–1677 (2018).
52. Wu, X. *et al.* White Matter Deficits Underlying the Impaired Consciousness Level in Patients with Disorders of Consciousness. *Neurosci. Bull.* **34**, 668–678 (2018).
53. Zheng, Z. S., Reggente, N., Lutkenhoff, E., Owen, A. M. & Monti, M. M. Disentangling disorders of consciousness: Insights from diffusion tensor imaging and machine learning. *Hum. Brain Mapp.* **38**, 431–443 (2017).
54. Tan, X. *et al.* Structural connectome alterations in patients with disorders of consciousness revealed by 7-tesla magnetic resonance imaging. *NeuroImage Clin.* **22**, (2019).
55. Bodart, O. *et al.* Global structural integrity and effective connectivity in patients with disorders of consciousness. *Brain Stimul.* **11**, 358–365 (2018).
56. Ferraro, S. *et al.* Interhemispherical Anatomical Disconnection in Disorders of Consciousness Patients. *J. Neurotrauma* **36**, 1535–1543 (2019).
57. Deco, G. & Kringelbach, M. L. Great expectations: Using whole-brain computational connectomics for understanding neuropsychiatric disorders. *Neuron* **84**, 892–905 (2014).
58. Kringelbach, M. L. & Deco, G. Brain States and Transitions: Insights from Computational Neuroscience. *Cell Reports* **32**, 108128 (2020).
59. Sanz-Leon, P., Knock, S. A., Spiegler, A. & Jirsa, V. K. Mathematical framework for large-scale brain network modeling in The Virtual Brain. *Neuroimage* **111**, 385–430 (2015).
60. Bensaid, S., Modolo, J., Merlet, I., Wendling, F. & Benquet, P. COALIA: a computational

model of human EEG for consciousness research. *Front. Syst. Neurosci.* **13**, 575043 (2019).
61. Marinazzo, D. *et al.* Information transfer of an Ising model on a brain network. *BMC Neurosci.* **14**, P376 (2013).
62. Keifer, J. & Summers, C. H. Putting the "biology" back into "neurobiology": The strength of diversity in animal model systems for neuroscience research. *Front. Syst. Neurosci.* **10**, 69 (2016).
63. Bisiach, E. & Luzzatti, C. Unilateral Neglect of Representational Space. *Cortex* **14**, 129–133 (1978).
64. Scoville, W. & Milner, B. Loss of recent memory after bilateral hippocampal lesions. *J. Neurol. Neurosurg. Psychiatry* **20**, 11–21 (1957).
65. Weiskrantz, L. Blindsight: A Case Study and Implications. (1986).
66. Deco, G., Tononi, G., Boly, M. & Kringelbach, M. L. Rethinking segregation and integration: Contributions of whole-brain modelling. *Nature Reviews Neuroscience* **16**, 430–439 (2015).
67. Adhikari, M. H. *et al.* Decreased integration and information capacity in stroke measured by whole brain models of resting state activity. *Brain* **140**, 1068–1085 (2017).
68. Hashemi, M. *et al.* The Bayesian Virtual Epileptic Patient: A probabilistic framework designed to infer the spatial map of epileptogenicity in a personalized large-scale brain model of epilepsy spread. *Neuroimage* **217**, 116839 (2020).
69. Aerts, H. *et al.* Modeling brain dynamics after tumor resection using The Virtual Brain. *Neuroimage* **213**, 116738 (2020).
70. Váša, F. *et al.* Effects of lesions on synchrony and metastability in cortical networks. *Neuroimage* **118**, 456–467 (2015).
71. Hellyer, P. J., Scott, G., Shanahan, M., Sharp, D. J. & Leech, R. Cognitive flexibility through metastable neural dynamics is disrupted by damage to the structural connectome. *J. Neurosci.* **35**, 9050–9063 (2015).
72. Honey, C. J. & Sporns, O. Dynamical consequences of lesions in cortical networks. *Hum. Brain Mapp.* **29**, 802–809 (2008).
73. Alstott, J., Breakspear, M., Hagmann, P., Cammoun, L. & Sporns, O. Modeling the Impact of Lesions in the Human Brain. *PLoS Comput. Biol.* **5**, e1000408 (2009).
74. Cabral, J., Kringelbach, M. L. & Deco, G. Exploring the network dynamics underlying brain activity during rest. *Prog. Neurobiol.* **114**, 102–131 (2014).
75. Hellyer, P. J. *et al.* The control of global brain dynamics: Opposing actions of frontoparietal control and default mode networks on attention. *J. Neurosci.* **34**, 451–461 (2014).
76. Deco, G. *et al.* Awakening: Predicting external stimulation to force transitions between different brain states. *Proc. Natl. Acad. Sci.* 201905534 (2019). doi:10.1073/pnas.1905534116
77. Jobst, B. M. *et al.* Increased Stability and Breakdown of Brain Effective Connectivity During Slow-Wave Sleep: Mechanistic Insights from Whole-Brain Computational Modelling. *Sci. Rep.* **7**, 4634 (2017).
78. Hahn, G. *et al.* Signature of Consciousness in Brain-Wide Synchronization Patterns of Monkey and Human fMRI Signals. *Neuroimage* **226**, 117470 (2020).
79. Ipiña, I. P. *et al.* Modeling regional changes in dynamic stability during sleep and wakefulness. *Neuroimage* **215**, (2020).
80. Sanz Perl, Y. *et al.* Perturbations in dynamical models of whole-brain activity dissociate between the level and stability of consciousness. *bioRxiv* (2020). doi:10.1101/2020.07.02.185157
81. Deco, G., Tagliazucchi, E., Laufs, H., Sanjuán, A. & Kringelbach, M. L. Novel Intrinsic Ignition Method Measuring Local-Global Integration Characterizes Wakefulness and Deep Sleep. *Eneuro* **4**, ENEURO.0106-17.2017 (2017).
82. López-González, A. *et al.* Loss of consciousness reduces the stability of brain hubs and the heterogeneity of brain dynamics. *bioRxiv* 2020.11.20.391482 (2020). doi:10.1101/2020.11.20.391482
83. Abeyasinghe, P. M. *et al.* Consciousness and the Dimensionality of DOC Patients via the Generalized Ising Model. *J. Clin. Med.* **9**, 1342 (2020).
84. Kandeepan, S. *et al.* Modeling an auditory stimulated brain under altered states of consciousness using the generalized ising model. *Neuroimage* **223**, 117367 (2020).

85. Stramaglia, S. *et al.* Ising model with conserved magnetization on the human connectome: Implications on the relation structure-function in wakefulness and anesthesia. *Chaos* **27**, (2017).
86. Deco, G. *et al.* Whole-Brain Multimodal Neuroimaging Model Using Serotonin Receptor Maps Explains Non-linear Functional Effects of LSD. *Curr. Biol.* **28**, 3065–3074 (2018).
87. Kringelbach, M. L. *et al.* Dynamic coupling of whole-brain neuronal and neurotransmitter systems. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 9566–9576 (2020).
88. Herzog, R. *et al.* A mechanistic model of the neural entropy increase elicited by psychedelic drugs. *Sci. Rep.* **10**, 17725 (2020).
89. Izhikevich, E. M. & Edelman, G. M. Large-scale model of mammalian thalamocortical systems. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 3593–3598 (2008).
90. Wong, K. F. & Wang, X. J. A recurrent network mechanism of time integration in perceptual decisions. *J. Neurosci.* **26**, 1314–1328 (2006).
91. Hansen, E. C. A., Battaglia, D., Spiegler, A., Deco, G. & Jirsa, V. K. Functional connectivity dynamics: Modeling the switching behavior of the resting state. *Neuroimage* **105**, 525–535 (2015).
92. Stephan, K. E., Weiskopf, N., Drysdale, P. M., Robinson, P. A. & Friston, K. J. Comparing hemodynamic models with DCM. *Neuroimage* **38**, 387–401 (2007).
93. Luppi, A. I. *et al.* A Synergistic Workspace for Human Consciousness Revealed by Integrated Information Decomposition. *bioRxiv* 2020.11.25.398081 (2020). doi:10.1101/2020.11.25.398081
94. Nørgaard, M. *et al.* A High-Resolution In Vivo Atlas of the Human Brain's Benzodiazepine Binding Site of GABA A Receptors. *bioarXiv* (2020). doi:10.1101/2020.04.10.035352
95. Muldoon, S. F., Bridgeford, E. W. & Bassett, D. S. Small-world propensity and weighted brain networks. *Sci. Rep.* **6**, (2016).
96. Lu, C.-W., Lin, T.-Y. & Wang, S.-J. Facilitation of Glutamate Release From Rat Cerebral Cortex Nerve Terminal by Subanesthetic Concentration Propofo.
97. Ranft, A. *et al.* Neural Correlates of Sevoflurane-induced Unconsciousness Identified by Simultaneous Functional Magnetic Resonance Imaging and Electroencephalography. *Anesthesiology* **125**, 861–872 (2016).
98. Boveroux, P., Vanhaudenhuyse, A. & Phillips, C. Breakdown of within- and between-network Resting State during Propofol-induced Loss of Consciousness. *Anesthesiology* **113**, 1038–1053 (2010).
99. Rosanova, M. *et al.* Sleep-like cortical OFF-periods disrupt causality and complexity in the brain of unresponsive wakefulness syndrome patients. *Nat. Commun.* **9**, (2018).
100. Funk, C. M. *et al.* Role of somatostatin-positive cortical interneurons in the generation of sleep slow waves. *J. Neurosci.* **37**, 9132–9148 (2017).
101. Fanciullacci, C. *et al.* Delta Power Is Higher and More Symmetrical in Ischemic Stroke Patients with Cortical Involvement. *Front. Hum. Neurosci.* **11**, 385 (2017).
102. Boveroux, P. *et al.* Brain function in physiologically, pharmacologically, and pathologically altered states of consciousness. *Int. Anesthesiol. Clin.* **46**, 131–146 (2008).
103. Laureys, S., Owen, A. M. & Schiff, N. D. Brain function in coma, vegetative state, and related disorders. *Lancet Neurology* **3**, 537–546 (2004).
104. Bullmore, E. & Sporns, O. The economy of brain network organization. *Nat. Rev. Neurosci.* **13**, 336–349 (2012).
105. Bassett, D. S. & Bullmore, E. T. Small-World Brain Networks Revisited. *Neuroscientist* **23**, 499–516 (2017).
106. Papo, D., Zanin, M., Martínez, J. H. & Buldú, J. M. Beware of the Small-World Neuroscientist! *Front. Hum. Neurosci.* **10**, 96 (2016).
107. Humphries, M. D. & Gurney, K. Network 'small-world-ness': a quantitative method for determining canonical network equivalence. *PLoS One* **3**, e0002051 (2008).
108. Smith, S. M. *et al.* Temporally-independent functional modes of spontaneous brain activity. *Proc. Natl. Acad. Sci.* **109**, 3131–3136 (2012).
109. Wang, P. *et al.* Inversion of a large-scale circuit model reveals a cortical hierarchy in the dynamic resting human brain. *Sci. Adv.* **5**, 1–12 (2019).

110. Demirtaş, M. *et al.* Hierarchical Heterogeneity across Human Cortex Shapes Large-Scale Neural Dynamics. *Neuron* **101**, 1181-1194.e13 (2019).
111. Sutton, J. A. & Clauss, R. P. A review of the evidence of zolpidem efficacy in neurological disability after brain damage due to stroke, trauma and hypoxia: A justification of further clinical trials. *Brain Injury* **31**, 1019–1027 (2017).
112. Radtke, F. M. *et al.* Risk factors for inadequate emergence after anesthesia: Emergence delirium and hypoactive emergence. *Minerva Anestesiol.* **76**, 394–404 (2010).
113. Xará, D., Silva, A., Mendonça, J. & Abelha, F. Inadequate emergence after anesthesia: emergence delirium and hypoactive emergence in the Postanesthesia Care Unit. *J. Clin. Anesth.* **25**, 439–446 (2013).
114. Naci, L. *et al.* Functional diversity of brain networks supports consciousness and verbal intelligence. *Sci. Rep.* **8**, (2018).
115. Varley, T. F. *et al.* Fractal dimension of cortical functional connectivity networks & severity of disorders of consciousness. *PLoS One* **15**, (2020).
116. Whitfield-Gabrieli, S. & Nieto-Castanon, A. Conn: A Functional Connectivity Toolbox for Correlated and Anticorrelated Brain Networks. *Brain Connect.* **2**, 125–141 (2012).
117. Behzadi Y, Restom K, Liau J & Liu TT. A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *Neuroimage* **37**, 90–101 (2007).
118. Tournier, J. D. *et al.* MRtrix3: A fast, flexible and open software framework for medical image processing and visualisation. *NeuroImage* **202**, (2019).
119. Veraart, J. *et al.* Denoising of diffusion MRI using random matrix theory. *Neuroimage* **142**, 394–406 (2016).
120. Leemans, A. & Jones, D. K. The B-matrix must be rotated when correcting for subject motion in DTI data. *Magn. Reson. Med.* **61**, 1336–1349 (2009).
121. Yeh, F.-C., Wedeen, V. J. & Tseng, W.-Y. I. Estimation of fiber orientation and spin density distribution by diffusion deconvolution. *Neuroimage* **55**, 1054–1062 (2011).
122. Luppi, A. I. & Stamatakis, E. A. Combining network topology and information theory to construct representative brain networks. *Netw. Neurosci.* 1–46 (2020). doi:10.1162/netn_a_00170
123. Medaglia, J. D. *et al.* Cognitive Control in the Controllable Connectome. *arXiv* (2016).
124. Desikan, R. S. *et al.* An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* **31**, 968–980 (2006).