# Repurposing EEG monitoring of general anaesthesia for building biomarkers of brain ageing: An exploratory study

David Sabbagh*[a,b], Jérôme Cartailler[a,c], Cyril Touchard[c], Jona Joachim[c], Alexandre Mebazaa[a,c], Fabrice Vallée[a,b,c], Étienne Gayat[a,c], Alexandre Gramfort[b], Denis A. Engemann*[b,d,e]

[a] Université de Paris, INSERM, U942 MASCOT, F-75006, Paris, France
[b] Université Paris-Saclay, Inria, CEA, Palaiseau, France
[c] Department of Anesthesia and Critical Care Medicine, AP-HP, Hôpital Lariboisière, F-75010, Paris, France
[d] Max Planck Institute for Human Cognitive and Brain Sciences, Department of Neurology, D-04103, Leipzig, Germany
[e] Roche Pharma Research and Early Development, Neuroscience and Rare Diseases, Roche Innovation Center Basel, F. Hoffmann–La Roche Ltd., Basel, Switzerland

Correspondence: *david.sabbagh@inria.fr, denis.engemann@roche.com

**Background:** EEG is a common tool for monitoring anaesthetic depth but is rarely reused at large for biomedical research. This study sets out to explore repurposing of EEG during anaesthesia to learn biomarkers of brain ageing.

**Methods:** We focused on brain age estimation as an example. Using machine learning, we reanalysed 4-electrodes EEG of 323 patients under propofol and sevoflurane. We included spatio-spectral features from stable anaesthesia for EEG-based age prediction applying recently published reference methods. Anaesthesia was considered stable when 95% of the total power was below a frequency between 8Hz and 13Hz.

**Results:** We considered moderate-risk patients (ASA <= 2) with propofol anaesthesia to explore predictive EEG signatures. Average alpha-band power (8-13Hz) was informative about age. Yet, state-of-the-art prediction performance was achieved by analysing the entire power spectrum from all electrodes (MAE = 8.2y, R2 = 0.65). Clinical exploration revealed that brain age was systematically linked with intra-operative burst suppression – commonly associated with age-related postoperative cognitive issues. Surprisingly, the brain age was negatively correlated with burst suppression in high-risk patients (ASA = 3), pointing at unknown confounding effects. Secondary analyses revealed that brain-age EEG signatures were specific to propofol anaesthesia, reflected by limited prediction performance under sevoflurane and poor cross-drug generalisation.

**Conclusions:** EEG from general anaesthesia may enable state-of-the-art brain age prediction. Yet, differences between anaesthetic drugs can impact the effectiveness of repurposing EEG from anaesthesia. To unleash the dormant potential of repurposing EEG-monitoring for clinical and health research, collecting larger datasets with precisely documented drug dosage will be key enabling factors.

**Keywords:** general anaesthesia, electroencephalogram (EEG), brain ageing, machine learning, burst suppression, propofol, sevoflurane

## Introduction

Electroencephalography (EEG) is often used during general anaesthesia (GA) for monitoring anaesthetic depth and adjusting drug dosage[1–5]. The increasing availability of EEG recordings from GA has stimulated cross-cutting research linking anesthesiology, neuroscience and cognition[6–8]. EEG dynamics during GA and the drug dosage required for achieving stable anaesthesia can depend on age and general health[9]. Cognitive decline has been associated with reduced drug requirement and EEG-power changes in the alpha frequency band (8-13Hz) [10,11]. These recent results point out the possibility that individual brain health can be reflected in EEG signals observed during GA[12].

Could EEG during GA, thus, be used beyond anaesthesia monitoring as a screening tool to assess the risk of developing neurodegenerative diseases? Several factors are favourable to this idea. First, EEG signals during stable anaesthesia can be of high signal quality as administration of curare minimises movement artefacts. Second, the availability of EEG recordings from GA keeps growing as GA is conducted in all age groups and everywhere in society, potentially avoiding selection bias of laboratory research[13,14]. Third, recent progress in advanced analytics and machine learning for EEG has led to development of novel biomarkers of cognition, brain function and health[15–17]. When combined, these factors could turn EEG monitoring data from GA into a valuable tool for clinical research.

In this study, we explore the possibility of repurposing existing EEG monitoring data from GA for building data-driven measures of brain aging[18]. Concretely, we propose to translate the previously developed brain-age framework[19] to the context of GA. By using machine learning, brain age can capture individual ageing. This framework provides quantitative comparisons of the chronological age of a person against their statistically expected age given their brain data ("this brain looks older/younger")[20,21]. Over the past years, numerous studies[19,20,22–24] have found that brain age estimates from the general population yield sensitive measures of neurodegenerative risk and disease severity in clinical populations[25].

The bulk of the brain age literature is based on MRI[19,24,26,27]. Nevertheless, a recent line of research has demonstrated that brain age can be estimated from EEG and may even contribute additional information over MRI[20,28,29]. It is currently unknown if brain age can be estimated from EEG during GA. Most EEG approaches for brain-age modelling relied on fine-grained spatial information provided by dozens to hundreds of electrodes and data collected in the laboratory setting[17,28]. On the other hand, brain age has recently been estimated from sleep EEG with fewer electrodes[18,30]. Another source of potential complications for estimating brain age from EEG during GA is related to the type of anaesthetic drug. For example, compared to widely used propofol anaesthesia, halogenated gases can shape EEG activity differently, increasing modelling complexity. Moreover, the amount of anaesthetic drug given to a patient to achieve a stable anaesthesia may by itself reflect the patient's cognitive health[10], potentially inducing confounding effects.

Our primary objective was to demonstrate the feasibility of modelling chronological age as a function of EEG power from intraoperative recordings in moderate-risk patients (ASA <= 2). We used state-of-the-art machine learning methods to estimate model prediction performance and hypothesised that the entire power spectrum was informative rather than specific frequency bands commonly used for monitoring anaesthetic depth. Our secondary objective was to investigate the association between chronological age, EEG-derived brain age, and intraoperative burst suppression in moderate-risk patients (ASA <= 2) and those suffering from pronounced health risk (ASA = 3). Finally, we investigated the specificity of the propofol-induced EEG signature of brain age through generalisation testing against EEG collected under sevoflurane anaesthesia.

# Methods

We reanalysed EEG and clinical data from the Probrain study registered under the ID NCT03876379 and approved by the SRLF (Société de réanimation de langue française) Ethics Advisory Committee (Chairperson Dr Jean Reignier, 48, avenue Claude Vellefaux, Paris, France) on the 5 January 2016, under the reference CE SRLF 11-356. The SRLF is the French national academic society for anaesthesia and critical care consulted by the department of anaesthesiology at the Lariboisière hospital (Paris, France). Patients were provided with an information letter. Verbal consent was recorded from every patient before anaesthesia.

## Patient selection

Between September 2017 and January 2020, patients eligible for an elective interventional surgery or interventional procedure under GA (for orthopaedic surgery or neuroradiology intervention for asymptomatic aneurysm) were selected to participate in the prospective, observational, monocentric Probrain study at the Lariboisiere hospital. The following exclusion criteria applied: pregnant women, age below 18 years, patients on sedation and mechanical ventilation before the procedure, history of bleeding aneurysm, neurodegenerative disease, neurological disorders, untreated depression.

## Anaesthetic protocol

GA was carried out based on standard practices. An opioid was administered (0.2ug/kg/h of sufentanil for orthopaedics and 3-5ng/ml of remifentanil for neuroradiology) followed by intravenous propofol and atracurium besylate non-depolarizing curare for GA induction. During maintenance either intravenous propofol or sevoflurane vapors were used as the main anaesthetic agents. Propofol was administered using total intravenous anaesthesia with a brain effect site concentration ranging from 3 to 3.5 ug/ml set by the physician and maintained automatically by syringe drivers under the Schnider model [13]. For sevoflurane, the mean alveolar concentration was set to obtain exhaled concentration at 1.5 to 2%. The choice of the anaesthetic drug for GA maintenance between sevoflurane and propofol, was left to the discretion of the anesthesiologist in charge. Propofol TCI (Target Controlled Infusion) and sevoflurane MAC (Minimum Alveolar Concentration) were adapted to keep a SEF95 between 8 and 15Hz during the maintenance.
The mean arterial pressure during the entire intervention was maintained at least 90% of its reference value and always above 65mm/Hg.

## Data collection

Cerebral activity during GA was monitored using a Masimo™ device with a 4-frontal electrodes EEG montage (Fp1, Fp2, F7 and F8, referenced on Cz), sampled at 63 Hz by default. EEG sub-hairline electrodes were placed a few minutes before GA induction and removed shortly after recovery of consciousness. Intraoperative EEG data were then extracted from the device, anonymized and stored on a file server in EDF format. We collected demographic data (age, gender, weight, height, BMI), clinical information (type of anaesthetic drug, ASA score) from the anaesthesia-specific medical consultation. This resulted in n = 323 patients with both a proper concatenated EEG recording and metadata information. The resulting sample is described in Table 1.

**Table 1. Patient characteristics. Data are presented as median [25th–75th percentiles]**

| Variable | Number of patients or median |
|---|---|
| Female/male | 211/112 |
| Age (yrs) | 57 [39 - 70] |
| Body weight (kg) | 74 [61 - 85] |
| BMI (kg $m^{-2}$) | 25 [22 - 29] |
| Propofol/sevoflurane | 216/107 |
| ASA = 1 or 2 / ASA = 3 | 249/68 |

## EEG processing, model construction and clinical exploration

The EEG processing and feature extraction is outlined in Fig 1 and detailed in the supplementary materials. We applied brain age models and methods for model comparisons from previous reference publications[20,28]. Ridge regression was used as prediction algorithm with 100 (Monte Carlo) cross-validation (CV) splits. Model superiority was assessed via pair-wise ranking across CV splits[20,31]. For model construction, we focused on moderate-risk patients with low ASA scores under propofol anaesthesia (strictly below 3, n = 170, see ds1 in **Table 2**). For clinical exploration, we compared brain age to burst suppression (a common predictor of postoperative complications). We computed cross-validated brain age predictions[20] on the data used for model building (ASA <= 2, ds3 in **Table 2**) and then predicted brain age in a dataset from a high-risk population not used for model building (ASA = 3, ds4 in **Table 2**). The brain age variable was then obtained by concatenating the predictions across the two datasets (ds5 in **Table2**). For definition of burst suppression and details on statistical analyses and software, see supplementary materials.

**Table 2. Subsets of patients used for model construction and model exploration defined by ASA status and drug type (total including patients without burst suppression annotation)**

| Subset | Number of patients |
|---|---|
| ds1 [construction]: ASA <= 2, propofol | 170 |
| ds2 [exploration]: ASA <= 2, sevoflurane | 79 |
| ds3 [exploration]: ASA <= 2, propofol | 167 (170) |
| ds4 [exploration]: ASA = 3, propofol | 37 (40) |
| ds5 [exploration]: ASA <= 3, propofol | 204 (210) |

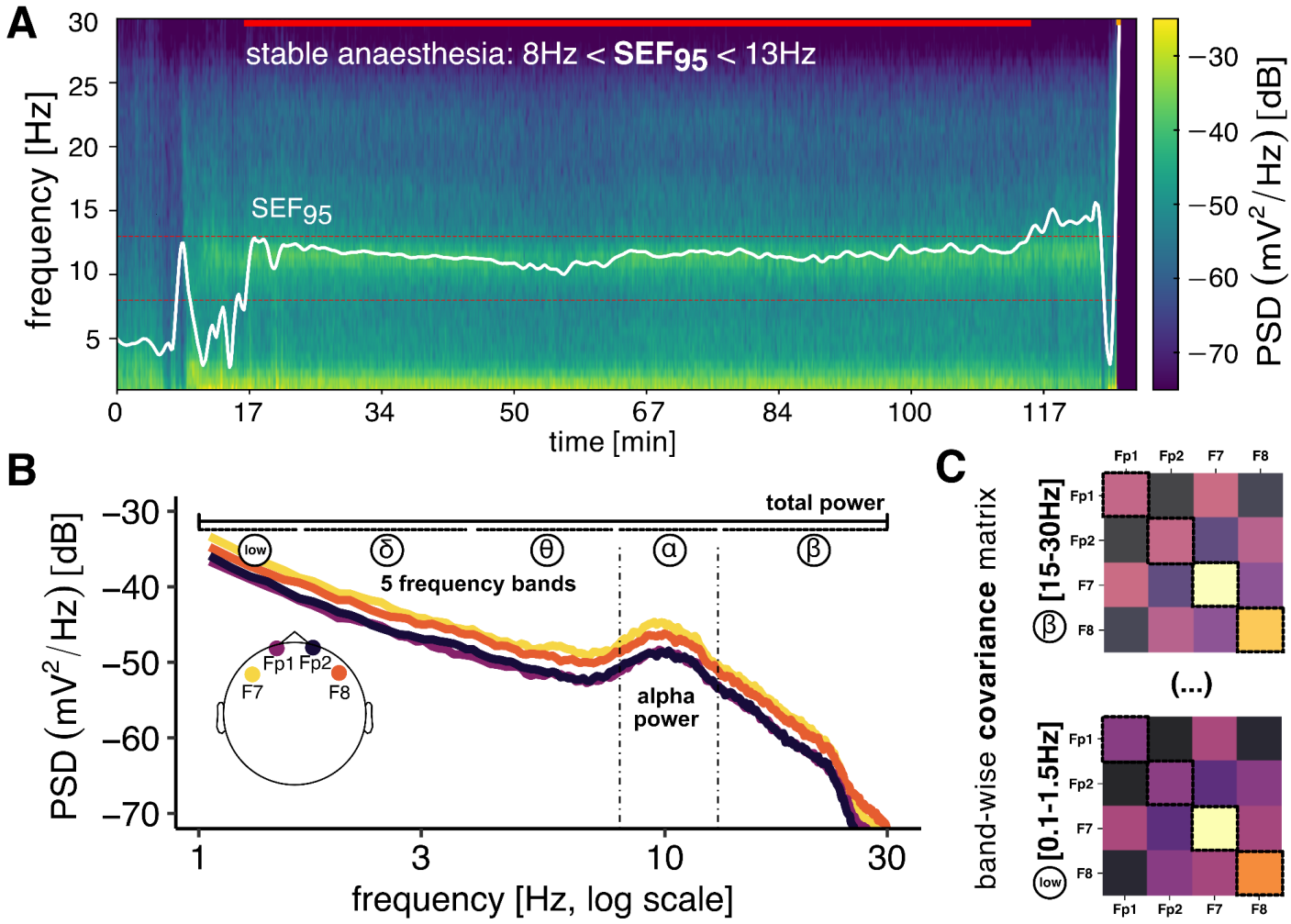


**Fig 1 Building biomarkers of brain aging from stable segments of GA. (A)** We selected the longest consecutive period of stable anaesthesia (propofol), defined by the spectral edge frequency (SEF) 95 falling in the 8-13 Hz range, the absence of high-voltage artifacts and of burst suppression. **(B)** We estimated the power-spectral density (PSD), which can reveal EEG signatures commonly used in anaesthetic monitoring such as total power or alpha power. Furthermore, we investigated the hypothesis that additional information on targets beyond clinical monitoring, *i.e.*, brain age is distributed across the entire power spectrum. For this purpose, we adaptively summarized the PSD over all frequencies and channels using supervised learning. Reference models for estimating brain age from EEG have exploited spatial patterns in different frequency bands presented by the covariance between EEG channels **(C)** whose diagonal summarises the PSD in the given frequency band. To assess the complementarity of these readouts, we used the stacking approach in which 1) 'alpha power', 2) 'total power', 3) the 'power spectrum' and 4) 'spatial patterns' from 5 frequency bands were gradually included. Model comparisons established the relative merit of these increasingly complex features.

## Results

We considered 323 patients for which we obtained actionable EEG (**Table 1**). A flow chart summarising the data collection process is shown in Supplementary Materials (**Fig S1**). Analysis-defining subsets of patients are listed in **Table 2**. We first explored the relationship between EEG activity and ageing across the frequency spectrum under propofol anaesthesia (**Fig 2**), which was more commonly administered in this study (**Table 1, Table2:** ds1). Binning the power spectra by age groups revealed age-related patterns (**Fig 2A**). Across frequencies, younger patients tended to show higher EEG power. We formally quantified this tendency using a linear mixed-

effects model regressing the log power (dB) on age, log frequency and their interaction (intercepts varying by patient). The analysis uncovered that regardless of frequency, EEG power declined on average by -0.10 dB, 95% CI [-0.13, -0.08] with every year of age. The full model is

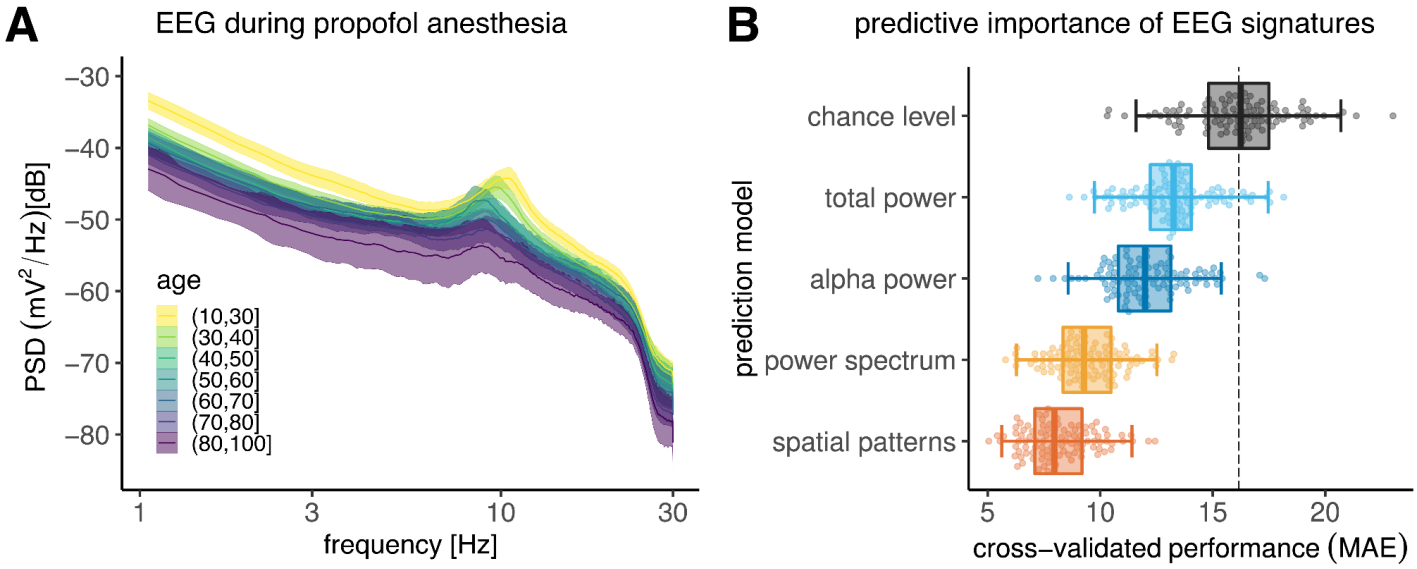


**Fig 2 Modelling brain ageing from spatio-spectral EEG signatures during propofol anaesthesia**. **(A)** depicts average power spectral density (PSD) by age group represented by different colours (n = 170, ASA <= 2). Shaded areas represent twice the standard error of the mean for every age group. Power and peak frequencies tend to decrease with age. **(B)** depicts cross-validated performance of age-prediction models based on progressive inclusion of different EEG signatures (Monte Carlo, 100 splits, MAE = mean absolute error). Dots represent scores at individual cross-validation splits, boxplots summarise the cross validation distribution, whiskers show the 2.5$^{th}$ and 97.5$^{th}$ percentiles. The dotted line indicates the expected performance for a chance-level model. Results are ordered by prediction error, descending from top to bottom. Models include EEG signatures from previous models (e.g. 'alpha power' refers to combining total power and alpha power). Informal analyses revealed that spatial patterns enabled virtually identical performance as the final model combining the diverse EEG signatures. For simplicity, we focused on the spatial patterns model in all subsequent analysis.

summarised in the supplementary materials (**Table S1**).

**Fig 2B** presents model comparisons of prediction models based on combinations of different EEG signatures. Chance level was estimated around 16 years of mean absolute error (MAE; $P_{25}$ = 14.8, $P_{75}$ = 17.5). The 'total power' model led to an improvement of the cross-validation (CV) score by about -2.9 years MAE on average ($P_{25}$ = -4.12, $P_{75}$ = -1.9), performing better than chance in 95/100 CV splits. As average alpha power was included, the CV scores improved by another -1.2 years MAE ($P_{25}$ = -1.9, $P_{75}$ = -0.34), beating the previous model on 86/100 CV splits. Adding fine-grained spectral information across all frequencies from 1-30Hz led to further improvements by -2.7 years MAE ($P_{25}$ = -3.7, $P_{75}$ = -1.86), out-performing the previous model on 95/100 CV splits. Finally, including information about the spatial patterns of covariance between the electrodes lowered the CV score by another -1.22 years MAE ($P_{25}$ = -1.93, $P_{75}$ = -0.47 superiority: 85/100 CV splits). In absolute terms, the final performance was about 8.2 years MAE ($P_{25}$ = 7.09, $P_{75}$ = 9.19) and corresponded to an $R^2$ score equivalent to 65% of explained variance ($P_{25}$ = 0.58, $P_{75}$ = 0.75).

For clinical exploration, we focused on the strong age-prediction signal under propofol anaesthesia in a subset of patients for whom burst suppression was annotated (n = 204, **Table 2:** ds3-ds5). We investigated the link between brain age, health status and undesirable intraoperative burst suppression. **Fig 3A** plots model-based age predictions against chronological age and burst-suppression proportion for the previously analysed cohort of patients with good general health (ASA score <= 2; n = 167, **Table 2:** ds3) as well as extrapolations to a distinct cohort of patients (ASA score = 3; n = 37, **Table 2:** ds4). Results suggested more complex relationships between EEG-predicted age and burst suppression across for patients of different health status and chronological age: One can observe that in younger patients with ASA <= 2, burst suppression was more frequent among those with higher brain age. This trend seems

inverted in older patients with ASA = 3 where bigger dots concentrate under the diagonal identity line. To formalise these relationships, we modelled the logit of the proportion of burst suppression as a weighted sum of scaled age, scaled brain age and health status and their respective interaction terms (**Fig 3B**). The analyses revealed a significant effect of brain age (beta = 0.85, SE = 0.20, t(196) = 4.4, p < 0.0001). Given a standard deviation of brain age of about 16y, this suggests that, across age and health status, burst suppression increased by 134% (exp(0.85) = 2.34) for every 16 years of brain age. Furthermore, a significant effect of brain age and health status emerged (beta = -0.99, SE = 0.44, t(196) = -3.7, p = 0.03), suggesting that the proportion of burst suppression was reduced by 63% (exp(-0.99) = 37%) for every 16 years of brain age in patients with ASA = 3 as compared to the other patients across all chronological ages. The full

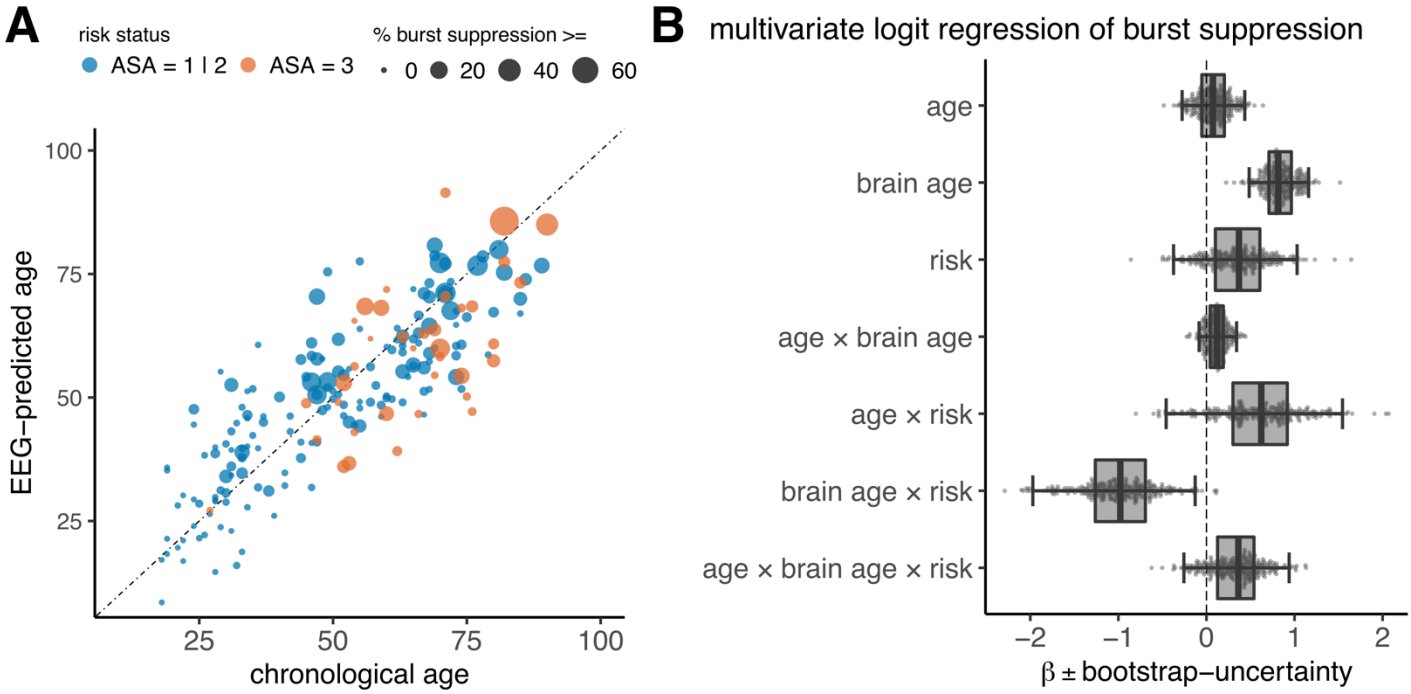


**Fig 3 Clinical exploration of Brain age derived from EEG during anaesthesia. (A)** plots chronological age (x axis) and brain age (y axis) against the proportion of the EEG showing burst suppression (marker size). Binary risk status (according to ASA scores) isi indicated by colour (n = 204). **(B)** presents a logit-regression model summarising these trends. Box plots represent the two-sided 95% CI (quantiles 2.5, 97.5), individual dots indicate possible coefficients implied by the statistical uncertainty of the model. Brain age and the interaction with brain age showed most consistent effects. Note that the exponential values of the coefficients can be interpreted as odds ratios of the proportion of burst suppression.

model is reported in the supplementary materials (**Table S2**).

Visual inspection of EEG activity under sevoflurane anaesthesia (**Fig 4A, Table2:** ds2) suggested higher EEG-power levels and weaker correlation between EEG power and age. We formally contrasted these trends with the previous results (**Fig 4A**) through a linear mixed-effects model regressing the log power (dB) on age, log frequency, drug type and their interactions; intercepts varied by patient. Independently of frequency or drug type, EEG power declined by -0.10 dB, 95% CI [-0.13, -0.08] for every year of age. Compared to propofol, sevoflurane led to 3.6 dB higher EEG power, CI 95% [1.21,6.10]. Notable interaction terms pointed at more complex EEG patterns: First, a two-way interaction suggested that differences between propofol and sevoflurane depended on frequency, implying that, under sevoflurane, the power declined, on average, by about -1.00 dB more than under propofol for every year of age, 95% CI [-1.27, -0.76]. Second, a three-way interaction suggested that under sevoflurane the effect of age may non-linearly increase by 0.01 dB per Hz, 95% CI [<0.01,0.01]. The full model is reported in the supplement materials (**Table S3**).

**Fig 4B** presents comparisons between age-prediction models that were either separately fitted and evaluated within each drug type, fitted within one drug type and evaluated on the other drug type, or fitted and evaluated pooling both drug types. When fitted separately,

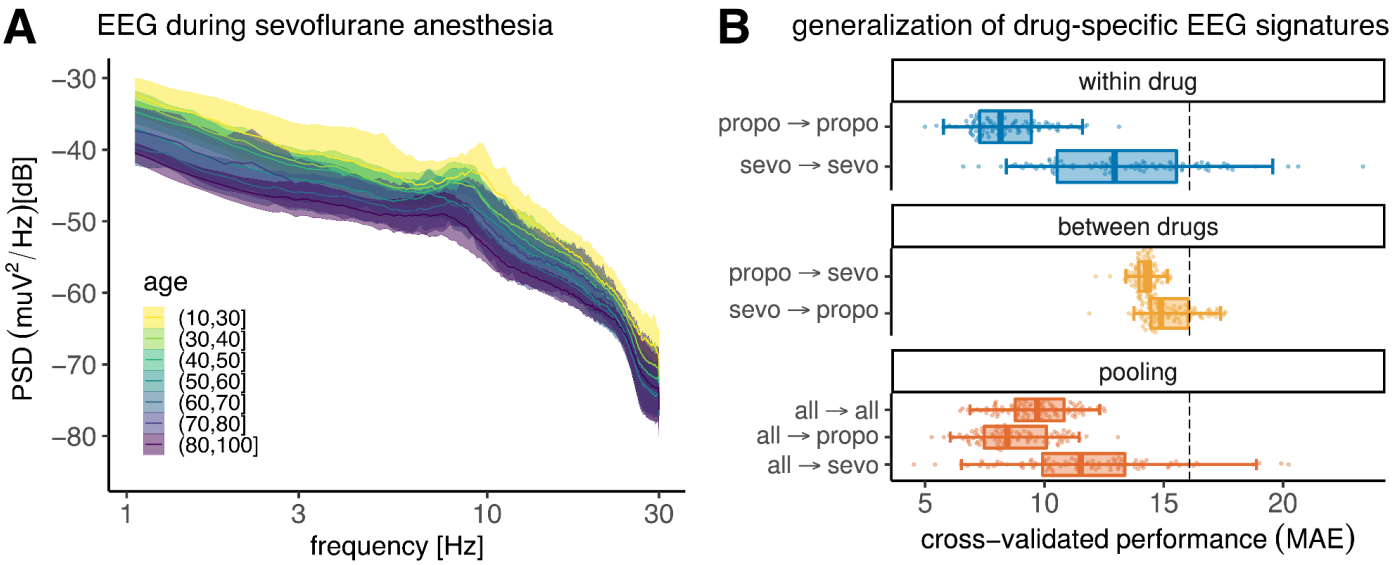


**Fig 4 Divergence and pooling of EEG-signatures by drug type. (A)** depicts the power spectral density of EEG under sevoflurane anaesthesia (n= 79, ASA <= 2). Visual conventions as in Fig 1A. Age appeared less consistently revealed by the PSD from sevoflurane as compared to propofol anaesthesia, raising the question if age prediction can be compared between these two drugs. **(B)** breaks down modelling performance of the 'spatial patterns' model by drug type (100 Monte Carlo splits, conventions as in Fig 1B). Within-drug models were fitted and evaluated separately by drug type (blue). Between-drug models were fitted on one drug type but evaluated on the other drug type (orange). Pooling was attempted (dark orange) by combining data for fitting and evaluation on both drug types (first row). Separate scoring was performed by drug type (rows 2-3). Results suggest that performance was best when using EEGs from propofol anaesthesia. While generalisation between drug types was poor, pooling EEGs from different drug types preserved the individual drug-specific performance.

the prediction model from sevoflurane EEG performed about 4.9 years of MAE ($P_{25}$ = 2.7, $P_{75}$ = 7.0) worse than propofol, inferior on 95/100 CV splits. Similarly, cross-drug generalisation of prediction models led to low prediction performance between 14 and 15 years of MAE. By comparison, generalisation from propofol to sevoflurane EEG was by -1 year of MAE more successful ($P_{25}$ = -1.8, $P_{75}$ = -0.25, superiority: 86/100 CV splits). Prediction models trained on the pooled EEG data from both drug types led to worse performance than the model trained under propofol EEG with about 1.3 year higher MAE ($P_{25}$ = -0.31, $P_{75}$ = 2.8, inferiority: 68/100 CV splits). On the other hand, results improved over the sevoflurane-based model by -3.7 years of MAE ($P_{25}$ = -6.4, $P_{75}$ = -1.3, superiority: 85/100 CV splits). Considering the subscores of the pooled model for the observations under sevoflurane EEG, suggested that combining both drugs

## Discussion

This study explored the feasibility of repurposing EEG-monitoring data from GA for building measures of brain ageing by importing machine learning approaches for brain age prediction originally developed in the laboratory setting. Under propofol anaesthesia, we reached prediction performance comparable to reference studies using research-grade EEG[28,32]. Model comparisons revealed that age-related information was present in spatial activity patterns distributed across the entire power spectrum. Clinical exploration of the well-performing propofol model highlighted important associations between brain-predicted age and the probability of burst suppression, which, however, depended on health status. Additional results suggested that the brain age signature was specific to propofol and may not generalise to sevoflurane.

Clinical studies have identified a link between EEG power spectra in the alpha band (8-13Hz), pre-operative cognitive decline[10], intraoperative burst suppression[33,34] and postoperative cognitive disorders[35]. This body of work pointed at the possibility that the EEG response to anaesthetic drugs may reveal the presence of neurodegenerative risk. Here we extended this prior art by directly applying brain-age prediction models for EEG[28] on a larger dataset of EEG during anaesthesia. Under propofol, this approach has led to performance matching recent work

with high-density MEG and EEG[20,32]. Previous studies during anaesthesia have instead focussed on EEG-signatures closely related to anaesthesia monitoring, with particular emphasis on total power and alpha power[6,36]. Our results have shown that the entire power spectrum and fine-grained correlations between signals collected at different electrodes may contain information relevant for learning biomarkers beyond anaesthesia monitoring.

On the other hand, prediction models constructed from EEG under sevoflurane anaesthesia were far less convincing and did not combine well with EEG collected under propofol anaesthesia. This may be intrinsically related to differences between both drugs regarding the mechanism of action: Propofol selectively activates $GABA_A$ receptors[37] whereas sevoflurane acts on several synaptic pathways, potentially increasing the complexity of the signal across age [38]. An alternative explanation might be that the dosage of sevoflurane was more constant across patients as it relied on standard MAC target values. Instead, propofol dosage was determined by a personalised TCI target value, chosen to stabilise GA i.e. to keep SEF95 in 8-13 Hz. This could, in principle, open the door to confounding effects, if propofol requirement for stable anaesthesia depends on age and health: Higher propofol dosage and brain age both dampen EEG power while patients at risk of developing burst suppression might receive lower doses of propofol. In turn, this should lead to reduced propofol-induced dampening, cancelling out the effect of brain-age on EEG power. That would be in line with observed changes of direction of association between brain age and burst suppression when comparing moderate-risk (ASA <= 2) versus high-risk (ASA = 3) patients, for whom, higher brain age went along with less burst suppression. Unfortunately, this hypothesis cannot be readily disambiguated using the present study, as neither the propofol dosage nor TCI parameters at a given moment in the EEG recording were available for deconfounding analysis.

The present study has successfully applied concepts and methods from laboratory research in cognitive neuroscience[19,23] on EEG monitoring data collected during GA. Therein, our systematic model comparisons of EEG signatures, drug types and patient populations have extended the scope of brain-age research to clinical real-world EEG collected in the absence of consciousness. Our findings motivate future research in clinical and cognitive neuroscience in anesthesiology beyond monitoring of anaesthesia depth. These strengths of our work have to be put in perspective with the important limitation that inconsistent findings regarding the variability of brain-age effects across drug types could not be resolved. Despite successful age-prediction results, the current work, therefore, does not yet present a ready-to-use brain age measure as, for example, the MRI-based brain age delta[19]. To push this exploratory effort to the next level, future studies with, ideally, larger samples must focus on precise control and measurement of drug dosage at any moment during anaesthesia. Preferably, propofol should be administered at constant dosage to rule out confounding by clinical factors. Validation against pre-surgical brain age estimates derived from gold-standard anatomical MRI[24] or research-grade high-density EEG[28] will be essential.

In conclusion, the present study points out the general feasibility of repurposing EEG from anaesthesia for learning biomarkers of brain ageing and health beyond the imminent perimeter of patient monitoring. To unleash this dormant potential of EEG-monitoring data for clinical and public health research, collecting larger datasets with precisely documented drug dosage will be key enabling factors.

## Author contributions

authors in alphabetical order

**Conceptualization**: A.M., A.G, D.E., E.G., F.V.

**Data curation**: C.T., D.S., J.J., F.V.

**Software**: A.G., D.E., D.S., J.C.

**Formal analysis**: A.G., D.E., D.S., J.C.

**Supervision**: A.G., D.E., E. G.

**Funding acquisition**: D.E., E.T.

**Validation**: A.M., C.T., D.E., D.S., J.C.

**Investigation**: D.E, D.S

**Visualisation**: D.E.

**Methodology**: D.E., D.S.

**Project administration**: D.E.

**Writing—original draft**: D.E., D.S.

**Writing—review and editing**: A.G., A.M., C.T, D.E., D.S., E.T, F.V., J.C, J.J.

## Acknowledgments

This work was supported by a 2018 “médecine numérique” (for digital medicine) thesis grant issued by Inserm (French national institute of health and medical research) and Inria (French national research institute for the digital sciences). It was also supported by the Inserm-Inria AI chair EDS-PeriOP and the ANR BrAIN AI chair (ANR-20-CHIA-0016).

## Declaration of conflicts of interest

D.E. is a full-time employee of F. Hoffmann-La Roche Ltd.

## References

1. Gibbs FA, Gibbs EL, Lennox WG: EFFECT ON THE ELECTRO-ENCEPHALOGRAM OF CERTAIN DRUGS WHICH INFLUENCE NERVOUS ACTIVITY. Arch Intern Med 1937; 60:154–66
2. Schneider G, Gelb AW, Schmeller B, Tschakert R, Kochs E: Detection of awareness in surgical patients with EEG-based indices—bispectral index and patient state index. Br J Anaesth 2003; 91:329–35
3. Purdon PL, Sampson A, Pavone KJ, Brown EN: Clinical electroencephalography for anesthesiologists: part I: background and basic signatures. Anesthesiology 2015; 123:937–60
4. Myers RR, Stockard JJ, Fleming NI, France CJ, Bickford RG: THE USE OF ON-LINE TELEPHONIC COMPUTER ANALYSIS OF THE E.E.G. IN ANAESTHESIA. Br J Anaesth 1973; 45:664–70
5. Radtke FM, Franck M, Lendner J, Krüger S, Wernecke KD, Spies CD: Monitoring depth of anaesthesia in a randomized trial decreases the rate of postoperative delirium but not postoperative cognitive dysfunction. Br J Anaesth 2013; 110 Suppl 1:i98-105
6. Purdon PL, Pavone KJ, Akeju O, Smith AC, Sampson AL, Lee J, Zhou DW, Solt K, Brown EN: The Ageing Brain: Age-dependent changes in the electroencephalogram during propofol and sevoflurane general anaesthesia. Br J Anaesth 2015; 115 Suppl 1:i46–57
7. Mashour GA, Palanca BJA, Basner M, Li D, Wang W, Blain-Moraes S, Lin N, Maier K, Muench M, Tarnal V, Vanini G, Ochroch EA, Hogg R, Schwartz M, Maybrier H, Hardie R, Janke E, Golmirzaie G, Picton P, McKinstry-Wu AR, Avidan

MS, Kelz MB: Recovery of consciousness and cognition after general anesthesia in humans. Elife 2021; 10
8. Supp GG, Siegel M, Hipp JF, Engel AK: Cortical hypersynchrony predicts breakdown of sensory processing during loss of consciousness. Curr Biol 2011; 21:1988–93
9. Cartailler J, Parutto P, Touchard C, Vallée F, Holcman D: Alpha rhythm collapse predicts iso-electric suppressions during anesthesia. Communications biology 2019; 2:1–10
10. Touchard C, Cartailler J, Levé C, Serrano J, Sabbagh D, Manquat E, Joachim J, Mateo J, Gayat E, Engemann D, Vallée F: Propofol Requirement and EEG Alpha Band Power During General Anesthesia Provide Complementary Views on Preoperative Cognitive Decline. Front Aging Neurosci 2020; 12:593320
11. Tanabe S, Mohanty R, Lindroth H, Casey C, Ballweg T, Farahbakhsh Z, Krause B, Prabhakaran V, Banks MI, Sanders RD: Cohort study into the neural correlates of postoperative delirium: the role of connectivity and slow-wave activity. Br J Anaesth 2020; 125:55–66
12. Palanca BJA, Wildes TS, Ju YS, Ching S, Avidan MS: Electroencephalography and delirium in the postoperative period. Br J Anaesth 2017; 119:294–307
13. Fry A, Littlejohns TJ, Sudlow C, Doherty N, Adamska L, Sprosen T, Collins R, Allen NE: Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. Am J Epidemiol 2017; 186:1026–34
14. Henrich J, Heine SJ, Norenzayan A: The weirdest people in the world? Behav Brain Sci 2010; 33:61–83; discussion 83-135
15. Engemann DA, Raimondo F, King J-R, Rohaut B, Louppe G, Faugeras F, Annen J, Cassol H, Gosseries O, Fernandez-Slezak D, Laureys S, Naccache L, Dehaene S, Sitt JD: Robust EEG-based cross-site and cross-protocol classification of states of consciousness. Brain 2018; 141:3179–92
16. Gaubert S, Raimondo F, Houot M, Corsi M-C, Naccache L, Diego Sitt J, Hermann B, Oudiette D, Gagliardi G, Habert M-O, Dubois B, De Vico Fallani F, Bakardjian H, Epelbaum S, Alzheimer's Disease Neuroimaging Initiative: EEG evidence of compensatory mechanisms in preclinical Alzheimer's disease. Brain 2019; 142:2096–112
17. Gemein LAW, Schirrmeister RT, Chrabąszcz P, Wilson D, Boedecker J, Schulze-Bonhage A, Hutter F, Ball T: Machine-learning-based diagnostics of EEG pathology. Neuroimage 2020; 220:117021
18. Ye E, Sun H, Leone MJ, Paixao L, Thomas RJ, Lam AD, Westover MB: Association of Sleep Electroencephalography-Based Brain Age Index With Dementia. JAMA Netw Open 2020; 3:e2017357
19. Cole JH, Franke K: Predicting Age Using Neuroimaging: Innovative Brain Ageing Biomarkers. Trends Neurosci 2017; 40:681–90
20. Engemann DA, Kozynets O, Sabbagh D, Lemaître G, Varoquaux G, Liem F, Gramfort A: Combining magnetoencephalography with magnetic resonance imaging enhances learning of surrogate-biomarkers. Elife 2020; 9:e54055
21. Spiegelhalter D: How old are you, really? Communicating chronic risk through 'effective age'of your body and organs. BMC Med Inform Decis Mak 2016; 16:1–6
22. Dosenbach NUF, Nardos B, Cohen AL, Fair DA, Power JD, Church JA, Nelson SM, Wig GS, Vogel AC, Lessov-Schlaggar CN, Others: Prediction of individual brain maturity using fMRI. Science 2010; 329:1358–61
23. Liem F, Varoquaux G, Kynast J, Beyer F, Masouleh SK, Huntenburg JM, Lampe L, Rahim M, Abraham A, Craddock RC, Riedel-Heller S, Luck T, Loeffler M, Schroeter ML, Witte AV, Villringer A, Margulies DS: Predicting brain-age from

multimodal imaging data captures cognitive impairment. Neuroimage 2017; 148:179–88
24. Cole JH, Ritchie SJ, Bastin ME, Hernández MCV, Maniega SM, Royle N, Corley J, Pattie A, Harris SE, Zhang Q, Others: Brain age predicts mortality. Mol Psychiatry 2018; 23:1385
25. Cole JH, Marioni RE, Harris SE, Deary IJ: Brain age and other bodily 'ages': implications for neuropsychiatry. Mol Psychiatry 2019; 24:266–81
26. Franke K, Luders E, May A, Wilke M, Gaser C: Brain maturation: predicting individual BrainAGE in children and adolescents using structural MRI. Neuroimage 2012; 63:1305–12
27. Jónsson BA, Bjornsdottir G, Thorgeirsson TE, Ellingsen LM, Walters GB, Gudbjartsson DF, Stefansson H, Stefansson K, Ulfarsson MO: Brain age prediction using deep learning uncovers associated sequence variants. Nat Commun 2019; 10:1–10
28. Sabbagh D, Ablin P, Varoquaux G, Gramfort A, Engemann DA: Predictive regression modeling with MEG/EEG: from source power to signals and cognitive states. Neuroimage 2020; 222:116893
29. Xifra-Porxas A, Ghosh A, Mitsis GD, Boudrias M-H: Estimating brain age from structural MRI and MEG data: Insights from dimensionality reduction techniques. Neuroimage 2021; 231:117822
30. Sun H, Paixao L, Oliva JT, Goparaju B, Carvalho DZ, Leeuwen KG van, Akeju O, Thomas RJ, Cash SS, Bianchi MT, Westover MB: Brain age from the electroencephalogram of sleep. Neurobiol Aging 2019; 74:112–20
31. Sabbagh D, Ablin P, Varoquaux G, Gramfort A, Engemann DA: Manifold-regression to predict from MEG/EEG brain signals without source modeling 2019 at <http://arxiv.org/abs/1906.02687>
32. Engemann DA, Mellot A, Höchenberger R, Banville H, Sabbagh D, Gemein L, Ball T, Gramfort A: A reusable benchmark of brain-age prediction from M/EEG resting-state signals 2021:p 2021.12.14.472691 doi:10.1101/2021.12.14.472691
33. Touchard C, Cartailler J, Levé C, Parutto P, Buxin C, Garnot L, Matéo J, Kubis N, Mebazaa A, Gayat E, Vallée F: EEG power spectral density under Propofol and its association with burst suppression, a marker of cerebral fragility. Clin Neurophysiol 2019; 130:1311–9
34. J. Cartailler CT, P. Parutto CP E Gayat, Vallee F: Brain fragility among middle-aged and elderly patients from electroencephalogram during induction of anaesthesia 2021; 38:1–3
35. Cartailler J, Loyer C, Vanderlynden E, Nizard R, Rabuel C, Coblentz Baumann L, Hourregue C, Dumurgier J, Paquet C: Determinants of Post-Operative Cognitive Decline in Elderly People. J Prev Alzheimers Dis 2021; 8:322–228
36. Bojak I, Liley DTJ: Modeling the effects of anesthesia on the electroencephalogram. Phys Rev E Stat Nonlin Soft Matter Phys 2005; 71:041902
37. Yip GMS, Chen Z-W, Edge CJ, Smith EH, Dickinson R, Hohenester E, Townsend RR, Fuchs K, Sieghart W, Evers AS, Franks NP: A propofol binding site on mammalian GABAA receptors identified by photolabeling. Nat Chem Biol 2013; 9:715–20
38. Kreuzer M, Stern MA, Hight D, Berger S, Schneider G, Sleigh JW, García PS: Spectral and Entropic Features Are Altered by Age in the Electroencephalogram in Patients under Sevoflurane Anesthesia. Anesthesiology 2020; 132:1003–16

# SUPPLEMENTARY MATERIAL

## Supplementary Methods

### EEG processing and feature engineering

We anonymized the monitoring EEG and converted into the BIDS format[1] using the MNE-BIDS package[2]. We epoched the signal of each patient using 60-second sliding windows (10 seconds shift) using the MNE-Python software[3] and the MNE-BIDS pipeline[1] For every epoch we computed two different types of power-spectral features using the coffeine package[2]. We estimated the power-spectral density (PSD) using Welch's method and Hamming windows of 8 seconds (4 seconds shift), which can reveal EEG signatures commonly used in anaesthetic monitoring and clinical research. PSDs were estimated in 244 frequency bins between 0 and 30 Hz, and Hamming windows w ere averaged using a trimmed mean (cut = 25%) to increase robustness to artefacts. To explore the importance of spatio-spectral patterns, we computed the covariance between all 4 electrodes in 5 frequency bands adapted from reference publications on brain age-prediction[4, 5].
Subsequently, epochs with peak-to-peak amplitude below 0.1 µV on any of the four electrodes were discarded, avoiding inclusion of artefacts or burst suppression. We then selected epochs of stable anaesthesia, defined by the spectral edge frequency (SEF) 95 falling in the [8-13]Hz range[6]. We retained the longest period of consecutive epochs of stable anaesthesia using the average SEF95 across the 4 channels. The power spectral density, the covariance matrices were then averaged across the selected epochs.

### Construction of prediction models, model comparisons and statistical inference

For constructing brain age models, we focused on moderate-risk patients with low ASA scores under propofol anaesthesia (strictly below 3, n = 170, see ds1 in **Table 2**), considering four alternative models highlighting distinct summaries of EEG power spectra. The 'total power' model summed EEG power (1-30Hz), averaged across all 4 electrodes (1 feature), representing the hypothesis of spectrally uniform information. The 'alpha power' model added EEG power averaged between (8-13Hz) across all 4 electrodes (1 additional feature), representing the hypothesis of a local spectral effect in the frequency range used for monitoring anaesthetic depth. The 'power spectrum' model presented uniformly spaced frequencies between (1-30Hz) averaged across all 4 electrodes (16 features), representing the hypothesis of distributed spectral effects. Finally, the 'spatial patterns' model implemented the brain age model from a reference publication[4] which analysed the covariance in the above 5 frequency bands (low [0.1-1.5]Hz, delta [1.5-4]Hz, theta [4-8]Hz, alpha [8-15]Hz and beta [15-30]Hz), representing the hypothesis of distributed spatio-spectral effects (35 features). The model concatenated the band-specific EEG covariances after vectorization with Riemannian embeddings that accounted for non-linearities caused by EEG-volume conduction[7].
As in previous reference models[4, 8], ridge regression[9] was used as the prediction algorithm. ts regularisation parameter was tuned by generalised cross validation[10] on a logarithmic grid of 100 values between [$10^{-10}$, $10^{100}$] on each training set. For fair comparisons between models based on different numbers of features, the stacking method[11] was used as in previous brain-age publications[4, 12]. As the sample size was limited, instead of the non-linear random forest algorithm, we also used ridge regression in the stacking layer.
To estimate the expected generalisation performance, Monte Carlo cross validation (CV) with 100 splits and 20% testing data was applied with fixed random seeds supporting pairwise model comparisons. To date, no generally accepted procedure exists for defining statistically justified null-hypothesis significance tests of observed performance rankings between pairs of prediction models. However, the cross-validation distribution itself provides valuable uncertainty estimates. To gauge practical significance, we followed the strategy from previous work[4, 8] and plotted or reported percentiles of the CV distribution ($P_{2.5}$, $P_{25}$, $P_{75}$, $P_{97.5}$) alongside rank statistics counting

[1] https://mne.tools/mne-bids-pipeline/
[2] https://github.com/coffeine-labs/coffeine

on how many splits one model was better than the other. Chance level was assessed by a dummy regressor predicting the mean outcome of the training data for each sample.

## Detection of burst suppression

To detect iso-electrical burst suppression (first part of the burst suppression pattern) from intraoperative EEG we adapted the method of a recent publication[13]. For each EEG, a trained clinician identified intra-operative periods based on the alpha-band. From this intraoperative EEG signal, we discarded flat artefacts searching for segments below 0.1 µV and high-amplitude epochs above 80 µV lasting at least 1 second. After the smoothing using a 30-seconds rolling average, we collected regions below 2.5 $\mu V$ amplitude, then we sequentially applied a 0.2 second erosion, a 1 second dilation and an 0.8 second erosion. From this output we estimated the fraction of time spent in burst suppression during induction (the first 25 min) and maintenance periods (after 25 min). We focused on the maintenance period for the better statistical properties of the signal (more samples due to a longer period, less artefacts) and potentially lower false positive rate (drug-induced burst suppression can arise at induction more often). Burst-suppression events were automatically excluded from analysis of EEG via artefact rejection. Annotations were not provided for 6 recordings (ds3 - ds5 in **Table 2**).

## Statistical analysis

To investigate correlations between age and the power spectrum, we computed linear mixed effect models using the lm4 package in R[14]. Statistical inference was obtained using confidence intervals and p-values approximated from the t-statistic as implemented by the sjPlot package (https://cran.r-project.org/web/packages/sjPlot/index.html). To investigate the associations between age, brain age and burst suppression across clinical groups, ordinary least squares regression was used on logit-transformed burst-suppression proportions as outcome. This approach allowed us to gauge complementarity of brain age and age in one modelling step[8, 15]. Next to obtain parameter estimates and p-values, we reported confidence intervals based on the parametric bootstrap implemented in the arm package as in a previous publication[15].

## Supplementary Figures

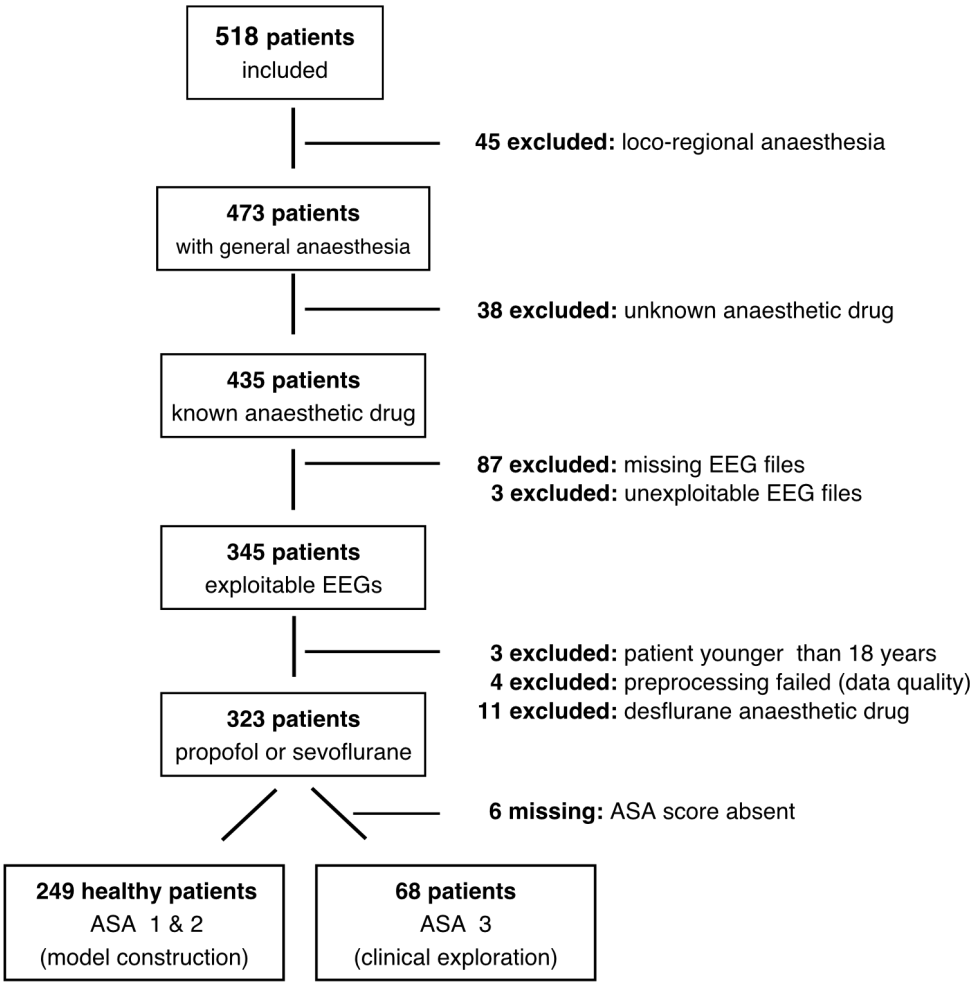


**Fig S1. Flowchart illustrating data selection.**

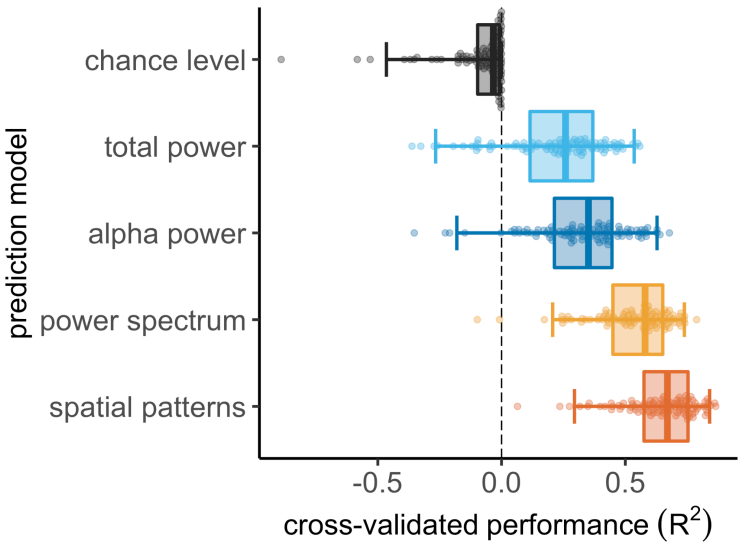


**Fig. S2: Model comparisons using variance explained ($R^2$) as error metric.** Same model and visual conventions as in **Fig 2B**.

**Table S1. Linear Mixed Effect model of EEG log power as a function of frequency and age.**

| | | psd | |
|---|---|---|---|
| *Predictors* | *Estimates* | *CI* | *p* |
| (Intercept) | -30.51 | -31.80 – -29.22 | **<0.001** |
| freqs [log] | -8.62 | -8.77 – -8.48 | **<0.001** |
| age | -0.10 | -0.13 – -0.08 | **<0.001** |
| freqs [log] * age | 0.00 | -0.00 – 0.00 | 0.286 |
| **Random Effects** | | | |
| $\sigma^2$ | 25.73 | | |
| $\tau_{00\ participant_id}$ | 8.03 | | |
| $N_{participant_id}$ | 166 | | |
| Observations | 42496 | | |

**Table S2. Generalised linear model of logit proportion of burst suppression as a function of age and brain age and risk status captured by the ASA score (sc = scaled).**

| | | **qlogis(y)** | |
|---|---|---|---|
| *Predictors* | *Estimates* | *CI* | *p* |
| (Intercept) | -4.31 | -4.59 – -4.04 | **<0.001** |
| age sc | 0.07 | -0.30 – 0.44 | 0.718 |
| brain age sc | 0.86 | 0.47 – 1.25 | **<0.001** |
| risk [high risk] | 0.33 | -0.35 – 1.01 | 0.340 |
| age sc * brain age sc | 0.13 | -0.10 – 0.35 | 0.265 |
| age sc * risk [high risk] | 0.58 | -0.31 – 1.47 | 0.200 |
| brain age sc * risk [high risk] | -0.99 | -1.85 – -0.13 | **0.025** |
| (age sc * brain age sc) * risk [high risk] | 0.34 | -0.25 – 0.93 | 0.254 |
| Observations | 204 | | |
| $R^2$ / $R^2$ adjusted | 0.349 / 0.326 | | |

**Table S3. Linear Mixed Effect model of EEG log power as a function of frequency, age and drug type.**

| | | psd | |
|---|---|---|---|
| *Predictors* | *Estimates* | *CI* | *p* |
| (Intercept) | -30.51 | -31.87 – -29.16 | **<0.001** |
| freqs [log] | -8.62 | -8.77 – -8.48 | **<0.001** |
| age | -0.10 | -0.13 – -0.08 | **<0.001** |
| drug type [sevoflurane] | 3.65 | 1.21 – 6.10 | **0.003** |
| freqs [log] * age | 0.00 | -0.00 – 0.00 | 0.286 |
| freqs [log] * drug type [sevoflurane] | -1.02 | -1.27 – -0.76 | **<0.001** |
| age * drug type [sevoflurane] | -0.01 | -0.06 – 0.03 | 0.591 |
| (freqs [log] * age) * drug type [sevoflurane] | 0.01 | 0.00 – 0.01 | **0.002** |
| **Random Effects** | | | |
| $\sigma^2$ | 25.69 | | |
| $\tau_{00\ participant_id}$ | 8.96 | | |
| $N_{participant_id}$ | 244 | | |
| Observations | 62464 | | |

## References


1. Pernet CR, Appelhoff S, Gorgolewski KJ, et al. EEG-BIDS, an extension to the brain imaging data structure for electroencephalography. *Scientific data* Nature Publishing Group; 2019; **6**: 1–5

2. Appelhoff S, Sanderson M, Brooks TL, et al. MNE-BIDS: Organizing electrophysiological data into the BIDS format and facilitating their analysis. *The Journal of Open Source Software* [Internet] 2019; **4** Available from: https://mne.tools/mne-bids/

3. Gramfort A, Luessi M, Larson E, et al. MNE software for processing MEG and EEG data. *Neuroimage* 2014; **86**: 446–60

4. Sabbagh D, Ablin P, Varoquaux G, Gramfort A, Engemann DA. Predictive regression modeling with MEG/EEG: from source power to signals and cognitive states. *Neuroimage* 2020; **222**: 116893

5. Engemann DA, Mellot A, Höchenberger R, et al. A reusable benchmark of brain-age prediction from M/EEG resting-state signals [Internet]. bioRxiv. 2021 [cited 2022 Jan 25]. p. 2021.12.14.472691 Available from: https://www.biorxiv.org/content/biorxiv/early/2021/12/16/2021.12.14.472691

6. Bruhn J, Bouillon TW, Radulescu L, Hoeft A, Bertaccini E, Shafer SL. Correlation of approximate entropy, bispectral index, and spectral edge frequency 95 (SEF95) with clinical signs of “anesthetic depth” during coadministration of propofol and remifentanil. *The Journal of the American Society of Anesthesiologists* The American Society of Anesthesiologists; 2003; **98**: 621–7

7. Congedo M, Barachant A, Bhatia R. Riemannian geometry for EEG-based brain-computer interfaces; a primer and a review. *Brain-Computer Interfaces* Taylor & Francis; 2017; **4**: 155–74

8. Engemann DA, Kozynets O, Sabbagh D, et al. Combining magnetoencephalography with magnetic resonance imaging enhances learning of surrogate-biomarkers. *Elife* eLife Sciences Publications Limited; 2020; **9**: e54055

9. Hoerl AE, Kennard RW. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics* Taylor & Francis Group; 1970; **12**: 55–67

10. Golub GH, Heath M, Wahba G. Generalized Cross-Validation as a Method for Choosing a Good Ridge Parameter. *Technometrics* [Taylor & Francis, Ltd., American Statistical Association, American Society for Quality]; 1979; **21**: 215–23

11. Wolpert DH. Stacked generalization. *Neural Netw* Elsevier; 1992; **5**: 241–59

12. Liem F, Varoquaux G, Kynast J, et al. Predicting brain-age from multimodal imaging data captures cognitive impairment. *Neuroimage* 2017; **148**: 179–88

13. Cartailler J, Parutto P, Touchard C, Vallée F, Holcman D. Alpha rhythm collapse predicts iso-electric suppressions during anesthesia. *Communications biology* Nature Publishing Group; 2019; **2**: 1–10

14. Bates D, Sarkar D, Bates MD, Matrix L. The lme4 package. *R package version* 2007; **2**: 74

15. Dadi K, Varoquaux G, Houenou J, Bzdok D, Thirion B, Engemann D. Population modeling with machine learning can enhance measures of mental health. *Gigascience* Oxford Univeristy Press; 2021; **10**: 1–16