

## ARTICLE OPEN



# Prediction of postoperative delirium after cardiac surgery by the interplay between preoperative plasma p-tau181 and IL-6 and heart-brain axis related factors: results from the prospective observational study FINDERI

Niels Hansen <sup>1</sup>✉, Clara Maria Knopp <sup>2</sup>, Hermann Esselmann <sup>1</sup>, Christopher M. Celano <sup>3,4</sup>, Carlotta Derad <sup>5</sup>, Thomas Asendorf <sup>5</sup>, Mohammed Chebbok <sup>2,6</sup>, Stephanie Heinemann <sup>2</sup>, Ihtaz Malik <sup>2</sup>, Barbara Morgado <sup>1</sup>, Matilda-Marie Becker <sup>1</sup>, Irina Günther <sup>1</sup>, Iryna Krasiuk <sup>1</sup>, Katharina Packroß <sup>1</sup>, Alina Isabel Rediske <sup>1</sup>, Nicholas Paul Süttmann <sup>1</sup>, Tobias Titsch <sup>1</sup>, Ingo Kutschka <sup>7,8</sup>, Hassina Baraki <sup>7,8</sup>, Jens Wiltfang <sup>1,9,10</sup>, Christine A. F. von Arnim <sup>2,8,12</sup>, Monika Sadlonova <sup>2,3,8,11,12</sup> and FINDERI investigators\*

© The Author(s) 2025, modified publication 2025

Postoperative delirium (POD) following cardiac surgery is a severe complication. There is evidence of a link between neuroinflammation and neurodegeneration in POD. We investigated the preoperative proinflammatory interleukin-6 (IL-6) and neuronal damage marker phosphorylated tau protein 181 (p-tau181) to POD while considering preoperative heart-brain axis related factors. The prospective FIND DELirium Risk factors (FINDERI) is an observational study in patients undergoing cardiac surgery. Biomarkers IL-6 and p-tau181 were measured in blood samples. For statistics, we utilized multiple logistic regression analyses and advanced machine learning techniques. In 491 patients, 106 (21.6%) developed POD. The age of patients with POD was significantly higher than that of patients without POD ( $p < 0.001$ ). Preoperative IL-6 and p-tau181 levels independently predicted POD [IL-6: area under the curve (AUC) = 0.605,  $p < 0.005$ ; p-tau181: AUC = 0.641,  $p < 0.0001$ ]. A multiple logistic regression analysis of preoperative log-transformed biomarkers levels (p-tau181, IL-6), female sex and cognitive performance increased the AUC (0.710,  $p < 0.0001$ ) in predicting POD. We created a decision tree prediction model including preoperative p-tau181, IL-6, and the severity of mitral valve disease (training data: AUC = 0.672,  $p < 0.0001$ ; validation data: AUC = 0.642,  $p < 0.05$ ). The LASSO regression showed an increased AUC in the training (0.751,  $p < 0.0001$ ) and validation dataset (0.652,  $p < 0.05$ ). Our results demonstrate that the combined assessment of preoperatively measured p-tau181 and IL-6, preoperative mitral valve disease, cognitive performance and female sex, significantly predicts POD. These findings provide evidence that neuroinflammation and neuronal cell damage are associated with POD.

*Molecular Psychiatry* (2026) 31:2509–2519; <https://doi.org/10.1038/s41380-025-03412-3>

## INTRODUCTION

Postoperative delirium (POD) is a frequent complication in cardiac surgery, with incidence rates ranging from 4 to 55% [1, 2]. The etiology involving aspects like neuroinflammation [3] and brain damage [4] remains incompletely understood. Emerging research highlights the importance of blood markers in elucidating POD mechanisms. Biomarkers are known to be closely associated with POD, such as preoperative [5, 6] and postoperative interleukin-6 (IL-6) levels [7], both of which implicate inflammation as one of the mechanisms contributing to POD development [8]. Given the link

between IL-6 and POD, we chose to investigate IL-6 as an inflammatory marker. Phosphorylated tau protein 181 (p-tau181) in plasma correlates closely with CSF p-tau181 and is a marker of structural brain damage [9, 10]. In addition, p-tau181 remains stable under the pre-analytical conditions of clinical practice [11] and is therefore suitable for blood sampling in the preoperative clinical setting. Elevated plasma tau (the term “tau” refers here to the middle region of tau as well as to truncated and complete forms of tau) was shown to correlate with delirium severity in POD patients [12] and accompanied by a simultaneous rise in IL-8 and

<sup>1</sup>Department of Psychiatry and Psychotherapy, University Medical Center Göttingen, Georg-August-University, Göttingen, Germany. <sup>2</sup>Department of Geriatrics, University Medical Center Göttingen, Georg-August-University, Göttingen, Germany. <sup>3</sup>Department of Psychiatry, Massachusetts General Hospital, Boston, MA, USA. <sup>4</sup>Department of Psychiatry, Harvard Medical School, Boston, MA, USA. <sup>5</sup>Department of Medical Statistics, University Medical Center Göttingen, Georg-August-University, Göttingen, Germany. <sup>6</sup>Department of Cardiology and Pneumology, University Medical Center Göttingen, Georg-August-University, Göttingen, Germany. <sup>7</sup>Department of Cardiovascular and Thoracic Surgery, University Medical Center Göttingen, Georg-August-University, Göttingen, Germany. <sup>8</sup>DZHK (German Center for Cardiovascular Research), partner site Göttingen, Göttingen, Germany. <sup>9</sup>German Center for Neurodegenerative Diseases (DZNE), Göttingen, Germany. <sup>10</sup>Neurosciences and Signaling Group, Institute of Biomedicine (iBIMED), Department of Medical Sciences, University of Aveiro, Aveiro, Portugal. <sup>11</sup>Department of Psychosomatic Medicine and Psychotherapy, University Medical Center Göttingen, Georg-August-University, Göttingen, Germany. <sup>12</sup>These authors contributed equally: Christine A. F. von Arnim, Monika Sadlonova. \*A list of authors and their affiliations appears at the end of the paper. ✉email: niels.hansen@med.uni-goettingen.de

Received: 26 March 2024 Revised: 19 November 2025 Accepted: 2 December 2025  
Published online: 22 December 2025

IL-10 [12] linking inflammation to neurodegeneration. While t-tau is a general biomarker for neurodegeneration, p-tau181 is specifically associated with the pathological hallmarks of Alzheimer disease (AD) [13] and POD severity [14]. A recent study showed that preoperative p-tau181 and p-tau217 plasma levels in 491 patients undergoing knee and hip replacement or laminectomy are predictive of POD [15], and of its severity. On the other hand, a recent study demonstrated that the distribution of measured p-tau181 levels in plasma does not differ significantly between patients with POD and those without it [16]. More research is required to clarify this controversy. Severe inflammation is likely to trigger neuronal cell damage and consequently raise the blood concentrations of IL-6 and p-tau181 in POD patients. However, conflicting evidence suggests that higher IL-6 levels correlate with reduced tau pathology in the cerebrospinal fluid [17], indicating a possible neuroprotective role of IL-6. Considering these complexities, our study examines the interplay between heart and brain disease, cognitive functioning, factors related to inflammation (IL-6 levels) such as immunotherapy, and factors related to neurodegeneration (p-tau181 levels) such as cognitive performance in relation to clinical outcomes and pathophysiology of POD following cardiac surgery. We hypothesize that the presence of both biomarkers – p-tau181 and IL-6 – along with preoperative factors related to heart and brain diseases and factors related to the immune system can predict POD in patients undergoing cardiac surgery. This approach postulates that inflammation and neurodegeneration may coexist in POD.

## METHODS

### Patient cohort recruitment and preoperative assessments

The FINd DELirium Risk factors (FINDERI, German Clinical Trials Registration number: DRKS00025095) study is an observational, prospective, and multidisciplinary investigation focused on POD in patients undergoing cardiac surgery [18–20]. Between 2021 and 2022, a total of 571 patients ( $\geq 50$  years) who consented to take part and underwent elective surgery in the Department of Cardiothoracic and Vascular Surgery at the University Medical Center Göttingen in Germany, were recruited. In addition to age and the requirement for elective surgery, fluency in the German language was defined as an inclusion criterion. Participants with a diagnosis of dementia -documented in the electronic medical records and impairing their ability to follow assessment instructions - were excluded. The Montreal Cognitive Assessment (MoCA) score did not serve as an exclusion criterion, since individuals may score low on the MoCA without meeting the diagnostic criteria for dementia according to the International Classification of Diseases, Tenth Revision (ICD-10). The current analysis centers on investigating the interplay between blood biomarkers such as IL-6 and p-tau181 and POD incidence. In addition, we take into account preexisting heart and brain comorbidities as well immune system factors to enhance the accuracy of POD prediction. We collected data related to heart disease, which includes a history of myocardial infarction, aortic, tricuspid and mitral valve disease and its severity, coronary artery disease, heart failure, the presence of pacemakers or defibrillators, aortic aneurysm, atrial fibrillation, and endocarditis. Factors for brain diseases and mental disorders collected from clinical records and not from those applied for modeling included carotid stenosis, history of cerebral ischemia, dementia, anxiety disorders, addiction, depression, and Parkinson's disease. Additionally, we examined other somatic factors such as peripheral artery disease, chronic obstructive pulmonary disease, renal insufficiency, recent acute dialysis, cancer within the last 5 years, thyroid disorders, and type 2 diabetes. All these factors were gathered from clinical data as part of the FINDERI study. Our study received ethical approval from the local ethics committee (approval: February 16, 2021, vote: November, 20, 2020). A study protocol was reported previously [18]. The study was in agreement with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement [21].

### Fluid biomarkers

*Processing and storing biomaterial probes.* Blood samples were collected 1–2 days before and seven days after cardiac surgery and stored at the biomaterial bank of the Department of Psychiatry and Psychotherapy,

University Medical Center Göttingen in Germany. The serum samples underwent a 45-minute incubation period before centrifugation at 2000 x g for 10 minutes at 20°C. Ethylenediaminetetraacetic acid (EDTA) plasma centrifugation was carried out at 2000 x g for 10 minutes. Subsequently, the plasma samples were stored at -80°C in our biomaterial bank.

*IL-6 measurement.* The next generation automated sandwich immunoassay was used to detect IL-6 in human sera samples using Simple-Plex assay on Ella™ (Bio-technie, Minneapolis, USA). The assay was performed according to the manufacturer's instructions on a single-analyte Ella cartridge. In brief, sera samples were used after 2-fold dilution with sample diluent, and 50  $\mu$ l of this solution was added to each sample inlet of the Ella cartridge. Subsequently, 1 ml of wash buffer was added to the corresponding wells of the Ella cartridge. The sample results were produced using Simple Plex Runner v.3.7.2.0. The entire procedure took about 90 minutes. The lower limit of quantification (LLOQ) was defined according to the manufacturer. No sample was outside the range of the LLOQ and upper LOQ (ULOQ).

*Lumipulse based quantification of p-tau181.* To determine p-tau181 concentrations, we used commercially available p-tau181 immunoreaction cartridges for plasma on the fully automated Lumipulse G600II system. Quality control measurements were taken. To measure directly EDTA plasma, 400  $\mu$ l of centrifuged plasma with 2 mL microtubes (Sarstedt, Germany) were added to the device. Simple p-tau181 measurements were taken on the automated Lumipulse device with the same samples. The assays were done as single measurements according to the kit's instructions. The lower limit of quantification (LLOQ) was defined according to the manufacturer. All samples measured were above the LLOQ.

### Postoperative delirium screening and cognitive assessment

POD screening was conducted twice daily, consisting of morning and evening sessions. These POD assessments involved administration of the Confusion Assessment Method (CAM) for intensive care unit (ICU; CAM-ICU) [22, 23] and I-CAM [24, 25] in the intermediate care unit. The Richmond Agitation Sedation Scale (RASS) [26] was initially performed in the ICU to assess the levels of sedation and agitation. The screening of POD was done by study personnel on the usage of the CAM-ICU and I-CAM, as well as on the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DMS-5) and the ICD-10 diagnostic criteria of delirium. Furthermore, the study staff took part in supervision sessions (weekly or once every two weeks) with a senior consultant (MS) after completing the training [27, 28]. The CAM [22, 24, 25, 29] is a standardized, evidence-based instrument for POD screening. We performed a slightly altered form of the CAM algorithm. In particular, we utilized a short form of CAM [29, 30] by supplementing psychomotor alterations to the CAM algorithm (I-CAM) [24] with the following characteristics as worksheets for the evaluators: altered level of consciousness, inattention, acute onset, disorganized thinking [29], as well as psychomotor alterations. POD existed if the following features were documented [29]: acute change, altered level of consciousness, fluctuation, disorganized thinking or inattention. The psychomotor alterations served to underline the aforementioned POD existence, and to define the POD subtype (e.g., hypoactive or hyperactive POD) [25]. The CAM-ICU exhibits a high sensitivity (0.95–1.00) and a specificity (0.89–0.93) with an inter-rater reliability of 0.88 to 1.0 [22, 24, 25, 29]. The I-CAM has a sensitivity of 0.77 in a geriatric cohort with a high occurrence of dementia and a specificity between 0.96 and 1.00, together with an inter-rater reliability of 0.95 [22, 24, 25, 29]. Furthermore, we have used the MoCA to measure cognition prior to cardiac surgery [31].

### Statistical approach

Statistical analyses were performed using the programming environment R, version 4.3.1 [32]. The planned sample size ( $n = 500$ ) of the FINDERI study was estimated based on the area under the curve (AUC) of a delirium screening tool, considering a dropout rate of 20%. Categorical variables are presented as absolute or relative frequencies (%), while continuous variables are reported using the mean and standard deviation (SD). To compare patients with and without POD, Pearson's Chi-squared test was utilized for categorical variables, and the Welch Two Sample t-test was used for continuous variables. A two-sided p-value of less than 0.05 was regarded statistically significant. Additionally, univariate logistic regression analysis was conducted to examine relevant somatic factors, with a focus

on heart and brain diseases. Biomarkers IL-6 and p-tau181 were log-transformed (10 based log) prior to regression analysis. Odds ratios (ORs) are presented with 95% confidence intervals (CI). For plasma biomarkers p-tau181 and IL-6, individual evaluations as predictors of POD were conducted using Receiver Operating Characteristic (ROC) curves. The AUC is reported with log-transformed 95%-CI [33]. Optimal cut-off values for these biomarkers were determined by simultaneously maximizing specificity and sensitivity, and are reported alongside specificity, sensitivity, negative and positive predictive values [34, 35]. A linear mixed-effects model was applied to log-transformed IL-6 levels, incorporating POD variables, timing of sample collection, and their interaction, while including a random effect for individual patient variation. Furthermore, a multiple logistic regression was conducted with POD as the target variable. This regression included log-transformed preoperative plasma biomarkers and their interactions, along with age, female sex, MoCa results, the presence of tumors, and immunotherapy-related factors identified in the cohort (corticosteroids, colchicine, and cytostatic drugs) as explanatory POD variables. In the logistic regression models, only complete cases are considered. Supplementary Table 1 shows the number of observations for the analyses carried out. Supervised machine learning (ML) methods, specifically regularized logistic regression employing a least absolute shrinkage and selection operator (LASSO) and a classification decision tree, were also utilized to identify key parameters for POD [36, 37]. Input variables included age at baseline, preoperative IL-6 and p-tau181 levels, and somatic conditions. The classification decision tree had a max. depth of four as well as a min. split requirement of 20 observations, determined by the Gini index. LASSO was implemented with prior multiple imputation to address missing and unknown variables. Multiple imputation was conducted five times, with 50 iterations each, using predictive mean matching [38]. The machine learning approach utilized a tenfold cross-validation. A dataset containing 70% of the data was randomly selected as the training set, and the residual 30% of data were utilized for validation [39].

## RESULTS

### Core characteristics of the FINDERI cohort

**Demographics.** A total of 504 patients were included in our cohort analysis (Table 1); 491 of them could be thoroughly evaluated for the presence of POD (Supplementary Table 2). Thirteen patients were excluded from the detailed statistical analysis as a thorough POD assessment was not possible (for dropout analysis see supplementary Table 3). Among them, 106 patients (21.6%) were diagnosed with POD following cardiac surgery, based on the CAM-ICU and I-CAM. Detailed demographic data is presented in Table 1. Patients diagnosed with POD were significantly older than those without POD (mean age with POD:  $71.0 \pm 7.7$  years vs. mean age without POD:  $67.6 \pm 8.3$  years,  $p < 0.001$ , Table 1). Notable statistical differences were also found between the two groups in terms of MoCA findings, heart failure, implanted pacemaker or defibrillator, heart valve disease, the presence and severity of mitral valve disease, and type 2 diabetes, as detailed in Table 1. POD lasted an average of  $3.3 \pm 1.3$  days, diagnosed  $1.4 \pm 0.7$  days after cardiac surgery and ended  $3.8 \pm 1.3$  days after intervention. Most patients exhibited hypoactive delirium symptoms (45%). A form of mixed POD occurred slightly less frequently (37%); the fewest patients presented with hyperactive POD (18%). POD was diagnosed in 73% of patients via a positive CAM-ICU score and 47% applying the I-CAM), where some patients were positive in both CAM-ICU and I-CAM.

**Preoperative heart and brain disease related factors.** We conducted univariate logistic regression analysis to identify significant heart and brain disease related factors associated with the development of POD in those patients assessable for POD ( $n = 491$ , see supplementary Table 2). The analysis revealed several relevant factors: age (OR 1.05, 95%-CI: 1.02, 1.08,  $p < 0.001$ ), MoCA assessment results (OR 0.86, 95%-CI: 0.81, 0.92,  $p < 0.001$ ), the presence of heart failure (OR 2.29, 95%-CI: 1.26, 4.47,  $p = 0.01$ ), heart valve disease (OR 1.73, 95%-CI: 1.09, 2.81,  $p = 0.024$ ), mitral valve insufficiency (OR 1.75, 95%-CI: 1.12, 2.72,  $p = 0.013$ ), mitral valve stenosis (OR 7.22, 95%-CI: 1.17, 55.9,

$p = 0.033$ ), moderately severe mitral valve disease (OR 3.8, 95%-CI: 1.78, 8.00,  $p < 0.001$ ), extremely severe mitral valve disease (OR 2.03, 95%-CI: 1.08, 3.74,  $p = 0.025$ ), tricuspid valve insufficiency (OR 1.7, 95%-CI: 1.02, 2.82,  $p = 0.038$ ), moderate tricuspid valve disease (OR 3.13, 95%-CI: 1.09, 8.66,  $p = 0.028$ ), and type 2 diabetes (OR 2.12, 95%-CI: 1.35, 3.31,  $p < 0.001$ ). Additionally, the presence of a defibrillator was identified as a significant factor (OR 4.76, 95%-CI: 1.24, 19.5,  $p = 0.022$ ). These findings are presented in supplementary Table 2.

**Plasma biomarker.** Plasma biomarkers p-tau181 and IL-6 were evaluated via univariate regression analysis to determine their significance in predicting POD in our cohort of 491 patients with assessed POD status (Supplementary Table 2). Our analysis indicated that preoperative IL-6 levels represented a significant risk factor (OR 1.38, 95%-CI 1.07, 1.77;  $p = 0.012$ ,  $n = 476$ ), while postoperative IL-6 levels (OR 1.07, 95%-CI 0.77, 1.47;  $p = 0.69$ ,  $n = 425$ ), and the difference between postoperative and preoperative IL-6 levels (OR 0.88, 95%-CI 0.68, 1.12,  $p = 0.29$ ,  $n = 422$ ) did not constitute significant risk factors in a univariate context (Supplementary Table 2). Similarly, preoperative levels of p-tau181 (OR 2.26, 95%-CI 1.46, 3.57;  $p < 0.001$ ,  $n = 483$ ) were found to be a significant predictor of POD in the FINDERI cohort.

### Plasma biomarker to predict postoperative delirium

The levels of preoperative IL-6 and preoperative p-tau181 were identified as significant predictors of POD. This is illustrated in Fig. 1; Supplementary Table 4. Supplementary Figure 1 also show the log-transformed biomarker levels of IL-6 and p-tau181. ROC analysis yielded an AUC of 0.605 (95%-CI: 0.544, 0.663,  $p = 0.0018$ ) for preoperative IL-6 levels and an AUC of 0.641 (95%-CI: 0.581, 0.698,  $p < 0.0001$ ) for preoperative p-tau181. The optimal cut-off value for preoperative IL-6 was 4.71, demonstrating a sensitivity of 60.4% and a specificity of 58.9%, as detailed in supplementary Table 5A (for postoperative IL-6 cut off value, see Table 4B). The optimal cut-off value determined for preoperative p-tau 181 was 1.57, with a sensitivity of 61.5% and a specificity 60.7%, as reported in supplementary Table 5C. However, it was observed that POD could not be predicted based on postoperative IL-6 levels or the difference between postoperative and preoperative IL-6 levels, as indicated in supplementary Table 4, supplementary Figures 1B, C and Fig. 1. Furthermore, a linear mixed-effects model for IL-6 demonstrated that while time ( $p < 0.0001$ ) and POD ( $p = 0.03$ ) were significant factors, the interaction between time and POD was not significant ( $p = 0.09$ ), as depicted in supplementary Figure 2. The mean group values in the difference of the log-transformed IL-6 levels for no POD were 1.809 (95%-CI: 1.71, 1.91,  $p < 0.0001$ ) and 1.624 for POD (95%-CI: 1.43, 1.82,  $p < 0.0001$ ).

### Predicting postoperative delirium by combining plasma biomarkers and heart and brain as well as immunotherapy related factors

To identify possible associations among biomarkers and various heart and brain as well as immune system related factors, we conducted a multiple logistic regression analysis in three different models (Table 2A). In the first model, IL-6, age, sex, tumor, corticosteroids, colchicine and cytostatic drugs were included as explanatory variables in the analysis; we found that age and log-transformed preoperative IL-6 facilitated the prediction of POD (Table 2B and Fig. 2, Model 1). We investigated the immunotherapy-associated variables in a model with IL-6, since the activity of the inflammatory marker IL-6 can be influenced by immunotherapy. Our first model was able to predict POD with reasonable accuracy, achieving an AUC of 0.658 (95%-CI: 0.596, 0.714,  $p < 0.0001$ ), as shown in Table 2B and Fig. 2, Model 1. Log p-tau181, age, gender, cognitive performance measured by MoCA and education were included as explanatory variables in our second model. It also included cognitive performance together

**Table 1.** Characteristics of the total FINDERI cohort and stratified by postoperative delirium: demographics, heart, brain and internal medicine disease factors and plasma biomarker.

Characteristic	N	N = 504 <sup>1</sup>	N	no POD, N = 385 <sup>1</sup>	N	POD, N = 106 <sup>1</sup>	p-value <sup>2</sup>
<b>Sex</b>	504		385		106		0.078
male		396 (78.57%)		309 (80.26%)		76 (71.70%)	
female		108 (21.43%)		76 (19.74%)		30 (28.30%)	
<b>Age</b>	504	68.3 ± 8.2	385	67.6 ± 8.3	106	71.0 ± 7.7	<0.001
<b>BMI</b>	496	28.3 ± 4.7	382	28.3 ± 4.6	102	28.2 ± 4.7	0.87
<b>MoCa</b>	499	23.8 ± 3.6	382	24.3 ± 3.4	104	22.3 ± 4.1	<0.001
<b>Coronary heart disease</b>	500	368 (73.60%)	382	279 (72%)	105	82 (77%)	0.57
<b>History of myocardial infarction</b>	499	126 (25.25%)	381	95 (24.93%)	105	29 (27.62%)	0.67
<b>Heart failure</b>	473	365 (77.17%)	363	268 (73.83%)	97	84 (86.60%)	<b>0.012</b>
<b>Implanted pacemaker or defibrillator</b>	503		384		106		<b>0.035</b>
No		482 (95.83%)		373 (97.14%)		98 (92.45%)	
pacemaker		12 (2.39%)		7 (1.82%)		3 (2.83%)	
defibrillator		9 (1.79%)		4 (1.04%)		5 (4.72%)	
<b>Heart valve disease</b>	502	318 (63.35%)	383	232 (60.57%)	106	77 (72.64%)	<b>0.030</b>
<b>Aortic valve disease</b>	501		382		106		0.61
No		294 (58.68%)		225 (58.90%)		61 (57.55%)	
insufficiency		69 (13.77%)		48 (12.57%)		18 (16.98%)	
stenosis		93 (18.56%)		75 (19.63%)		17 (16.04%)	
combined		45 (8.98%)		34 (8.90%)		10 (9.43%)	
<b>Severity of aortic valve disease</b>	488		372		104		0.88
none		294 (60.25%)		225 (60.48%)		61 (58.65%)	
mild		46 (9.43%)		33 (8.87%)		12 (11.54%)	
moderate		34 (6.97%)		27 (7.26%)		7 (6.73%)	
severe		114 (23.36%)		87 (23.39%)		24 (23.08%)	
<b>Mitral valve disease</b>	502		383		106		<b>0.012</b>
No		284 (56.57%)		231 (60.31%)		48 (45.28%)	
insufficiency		207 (41.24%)		146 (38.12%)		53 (50.00%)	
stenosis		5 (1.00%)		2 (0.52%)		3 (2.83%)	
combined		6 (1.20%)		4 (1.04%)		2 (1.89%)	
<b>Severity of mitral valve disease</b>	493		381		101		<0.001
none		284 (57.61%)		231 (60.63%)		48 (47.52%)	
mild		105 (21.30%)		86 (22.57%)		19 (18.81%)	
moderate		40 (8.11%)		19 (4.99%)		15 (14.85%)	
severe		64 (12.98%)		45 (11.81%)		19 (18.81%)	
<b>Tricuspid valve disease</b>	500		381		106		0.10
No		403 (80.60%)		314 (82.41%)		78 (73.58%)	
insufficiency		96 (19.20%)		66 (17.32%)		28 (26.42%)	
stenosis		1 (0.20%)		1 (0.26%)		0 (0.00%)	
combined		0 (0.00%)		0 (0.00%)		0 (0.00%)	
<b>Severity of tricuspid valve disease</b>	496		379		104		0.067
none		403 (81.25%)		314 (82.85%)		78 (75.00%)	
mild		66 (13.31%)		48 (12.66%)		18 (17.31%)	
moderate		16 (3.23%)		9 (2.37%)		7 (6.73%)	
severe		11 (2.22%)		8 (2.11%)		1 (0.96%)	
<b>Aortic aneurysm</b>	501	40 (7.98%)	383	29 (7.57%)	105	8 (7.62%)	>0.99
<b>Aortic dissection</b>	501	4 (0.80%)	383	1 (0.26%)	105	0 (0.00%)	>0.99
<b>Carotis artery stenosis</b>	491	69 (14.05%)	377	50 (13.26%)	102	17 (16.67%)	0.47
<b>Atrial fibrillation</b>	498	105 (21.08%)	381	73 (19.16%)	104	29 (27.88%)	0.072
<b>Endocarditis</b>	497	7 (1.41%)	380	4 (1.05%)	104	3 (2.88%)	0.36
<b>Peripheral arterial occlusion disease</b>	500	59 (11.80%)	381	41 (10.76%)	106	17 (16.04%)	0.19

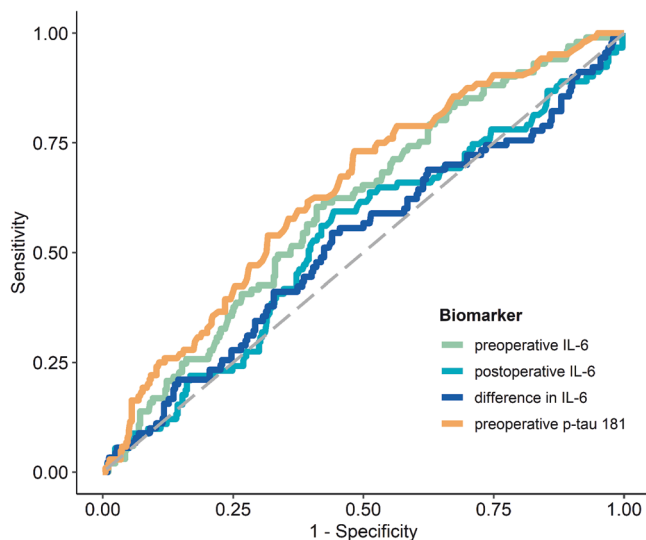
Table 1. continued

Characteristic	N	N = 504 <sup>1</sup>	N	no POD, N = 385 <sup>1</sup>	N	POD, N = 106 <sup>1</sup>	p-value <sup>2</sup>
History of stroke	499	54 (10.82%)	382	38 (9.95%)	104	15 (14.42%)	0.26
Renal insufficiency	502	64 (12.75%)	384	46 (11.98%)	105	17 (16.19%)	0.33
Acute dialysis requirement	503	3 (0.60%)	384	3 (0.78%)	106	0 (0.00%)	0.83
Tumor (during the past 5 years)	501	45 (8.98%)	383	32 (8.36%)	105	10 (9.52%)	0.86
Thyroid diseases	501	87 (17.37%)	382	64 (16.75%)	106	21 (19.81%)	0.56
Chronic obstructive pulmonary disease	500	40 (8.00%)	381	31 (8.14%)	106	8 (7.55%)	>0.99
Diabetes	503	150 (29.82%)	384	102 (26.56%)	106	46 (43.40%)	<b>0.001</b>
Depressive disorder	500	61 (12.20%)	382	41 (10.73%)	105	18 (17.14%)	0.11
Anxiety disorder	501	23 (4.59%)	383	19 (4.96%)	105	3 (2.86%)	0.51
Addiction (drugs/alcohol)	498	16 (3.21%)	381	13 (3.41%)	104	3 (2.88%)	>0.99
Dementia	503	1 (0.20%)	384	0 (0.00%)	106	1 (0.94%)	0.49
Parkinsons disease	503	2 (0.40%)	384	0 (0.00%)	106	2 (1.89%)	0.066
Log-transformed preoperative p-tau181	496	0.4 ± 0.5	379	0.4 ± 0.5	104	0.6 ± 0.5	<b>&lt;0.001</b>
Log-transformed preoperative IL-6	485	1.6 ± 0.8	375	1.5 ± 0.8	101	1.8 ± 0.8	<b>0.007</b>
Log-transformed postoperative IL-6	431	3.4 ± 0.7	334	3.3 ± 0.7	91	3.4 ± 0.7	0.71
Difference of log-transformed IL-6 (post-pre)	428	1.8 ± 1.0	332	1.8 ± 1.0	90	1.7 ± 0.9	0.26

<sup>1</sup>OR = Odds Ratio.

<sup>2</sup>Pearson's Chi-squared test or Welch Two Sample t-test.

n, number of patients, BMI body mass index, CI confidence interval, n number of patients, IL-6 interleukin 6, MoCa montreal cognitive assessment, POD postoperative delirium, p-tau 181 phosphorylated tau protein 181.



**Fig. 1** ROC curves of IL-6 and p-tau181 plasma biomarkers for predicting POD. ROC curves of preoperative IL-6 (orange), preoperative p-tau 181 (green), postoperative IL-6 (light blue) and the difference between post- and preoperative IL-6 (dark blue) blood levels for predicting POD following the CAM and CAM-ICU POD definition. Abbreviations: CAM = confusion assessment method, CAM-ICU = Confusion Assessment Method for the Intensive Care Unit, POD = postoperative delirium, IL-6 = interleukin 6, p-tau181 = phosphorylated tau protein 181.

with preoperative p-tau181 levels, as neurodegenerative processes are often associated with reduced cognition. This model revealed that log-transformed preoperative p-tau181, female sex, and cognitive performance were relevant factors. POD could be predicted with this model with moderate accuracy, namely an AUC of 0.694 (95%-CI: 0.637, 0.747,  $p < 0.0001$ , Table 2C, Fig. 2, Model 2). When we combined all explanatory variables (log-transformed IL-6, log-transformed p-tau181, age, sex, tumor, corticosteroids, colchicine, cytostatic drugs, MoCA findings,

education, preoperative log IL-6 and preoperative p-tau 181 and female sex) into one model (Model 3), female sex, education > 10 years and cognitive performance (MoCA) proved to be relevant factors. However, the interaction between the blood biomarkers p-tau181 and IL-6 in blood plasma was not a relevant POD predictor (OR 0.87, 95%-CI: 0.48, 1.56,  $p = 0.63$ , Table 2, Model 3). With this combined model 3, POD was predictable with moderate accuracy with an AUC of 0.709 (95% CI: 0.651, 0.763,  $p < 0.0001$ ; Table 2D, Fig. 2, Model 3).

### Predicting postoperative delirium taking a machine learning approach

**Decision tree.** To identify the most important guidelines for predicting POD, we created a decision showing four important rules. The rules were applied in the following sequence: (1) preoperative p-tau181 level greater than 1.4, (2) the presence of moderate or severe mitral valve disease, (3) a preoperative p-tau181 exceeding 1.8, and (4) preoperative IL-6 level above 5.8, as depicted in Fig. 3. The performance of this decision tree model was quantitatively evaluated, showing an AUC of 0.672 (95%-CI: 0.604, 0.735,  $p < 0.0001$ ) on the training set and 0.642 (95%-CI: 0.537, 0.738,  $p = 0.0108$ ) on the validation set. These results, along with detailed visual representations, are provided in supplementary Table 6, Fig. 3.

**LASSO.** During the application of the LASSO machine learning procedure, non-zero regression coefficient estimates were observed for two variables: age and preoperative p-tau181 levels. The performance of the LASSO model in predicting POD was assessed using ROC analysis, which demonstrated an AUC of 0.751 (95%-CI: 0.686, 0.805,  $p < 0.0001$ ) for the training and an AUC of 0.652 (95%-CI: 0.539, 0.747,  $p = 0.0086$ ) for the validation set, reflecting a moderate level of prediction accuracy. These findings, along with the corresponding graphical representation, are detailed in supplementary Table 6 and Fig. 4A, B.

### DISCUSSION

Our research in a substantially large cohort reveals that measuring preoperative p-tau181 and IL-6 levels in blood plasma in

**Table 2.** Multiple logistic regression models.

<b>A: Number of observations in the selected multivariate logistic regression models</b>			
<b>Model</b>	<b>Number of observations</b>		
Model 1	475		
Model 2	477		
Model 3	471		
<b>B: Model 1</b>			
<b>Characteristic<sup>1</sup></b>	<b>OR<sup>2</sup></b>	<b>95% CI<sup>2</sup></b>	<b>p-value</b>
<b>Log-transformed preoperative IL-6</b>	<b>1.35</b>	<b>1.03, 1.76</b>	<b>0.026</b>
<b>Age</b>	1.05	1.02, 1.08	<b>0.001</b>
<b>Sex: female</b>	1.62	0.95, 2.73	0.071
<b>Tumor</b>	1.10	0.48, 2.33	0.82
<b>Corticosteroids</b>	1.70	0.63, 4.29	0.27
<b>Colchicine</b>	0.00		0.99
<b>Cytostatics</b>	0.00		0.99
<b>C: Model 2</b>			
<b>Characteristic<sup>1</sup></b>	<b>OR<sup>2</sup></b>	<b>95% CI<sup>2</sup></b>	<b>p-value</b>
<b>Log-transformed preoperative p-tau181</b>	1.63	1.00, 2.65	<b>0.045</b>
<b>Age</b>	1.02	0.99, 1.05	0.18
<b>Sex: female</b>	1.86	1.08, 3.17	<b>0.024</b>
<b>MoCa</b>	0.87	0.81, 0.94	<b>&lt;0.001</b>
<b>Education<sup>3</sup></b>			
Finished secondary general school	0.92	0.25, 3.81	0.90
Finished secondary school	1.06	0.28, 4.55	0.94
Finished polytechnic secondary school	0.93	0.14, 5.76	0.94
Finished vocational school, technical college	1.64	0.30, 9.41	0.57
Finished academic secondary school	1.73	0.41, 8.17	0.47
<b>D: Model 3</b>			
<b>Characteristic<sup>1</sup></b>	<b>OR<sup>2</sup></b>	<b>95% CI<sup>2</sup></b>	<b>p-value</b>
<b>Log-transformed preoperative IL-6</b>	1.32	0.85, 2.00	0.20
<b>Log-transformed preoperative p-tau181</b>	1.93	0.59, 6.27	0.27
<b>Age</b>	1.02	0.99, 1.05	0.18
<b>Sex: female</b>	1.93	1.10, 3.34	<b>0.020</b>
<b>Tumor</b>	1.03	0.43, 2.26	0.95
<b>Corticosteroids</b>	1.96	0.69, 5.21	0.19
<b>Colchicine</b>	0.00		0.99
<b>Cytostatics</b>	0.00		0.99
<b>MoCa</b>	0.87	0.81, 0.94	<b>&lt;0.001</b>
<b>Education<sup>3</sup></b>			
Finished secondary general school	1.05	0.27, 4.60	0.94
Finished secondary school	1.17	0.29, 5.30	0.83
Finished polytechnic secondary school	1.25	0.18, 8.23	0.82
Finished vocational school, technical college	2.37	0.40, 14.7	0.34
Finished academic secondary school	2.19	0.49, 11.1	0.32
<b>Log-transformed preoperative IL-6* Log-transformed preoperative p-tau181</b>	0.87	0.48, 1.56	0.63

<sup>1</sup>AUC 0.6575, (95%-CI 0.5957, 0.7138;  $p < 0.0001$ ),

<sup>2</sup>OR odds ratio, CI confidence interval.

IL-6 interleukin 6

<sup>1</sup>AUC 0.6939, (95%-CI 0.6365, 0.7463;  $p < 0.0001$ ), <sup>2</sup>OR = Odds Ratio, CI = Confidence Interval,

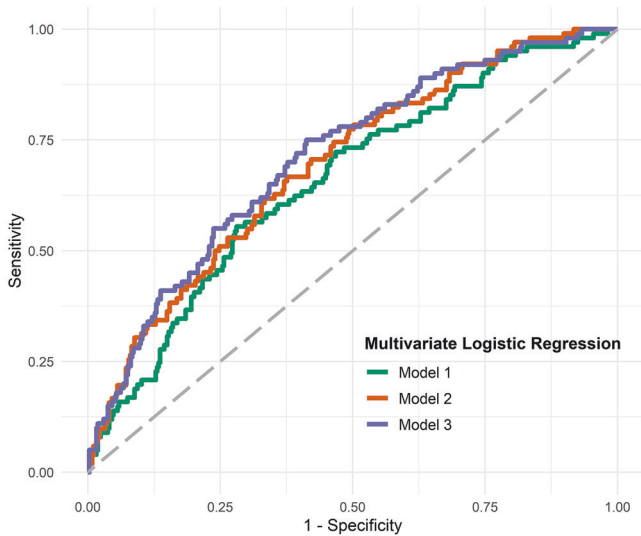
<sup>3</sup>The German secondary education system with secondary educational institutions in Germany is depicted in supplementary table 7.

MoCa montreal cognitive assessment, p-tau181 = phosphorylated tau protein 181

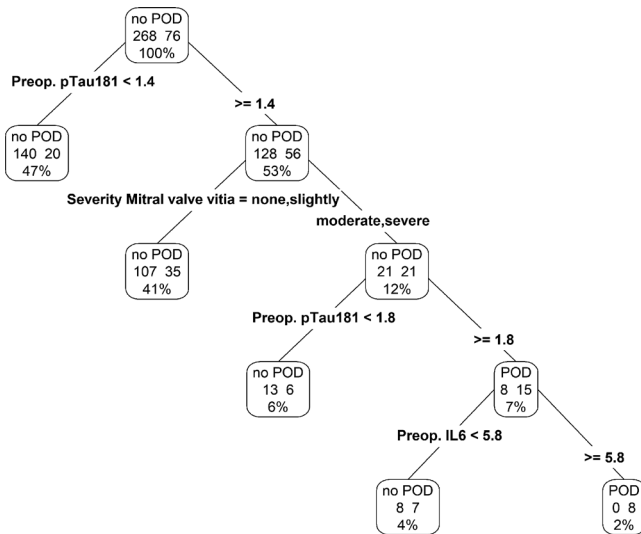
<sup>1</sup>AUC 0.7085, (95%-CI 0.6492, 0.7619;  $p < 0.0001$ ), <sup>2</sup>OR odds ratio, CI confidence interval,

<sup>3</sup>The German secondary education system with secondary educational institutions in Germany is depicted in supplementary table 7.

IL-6 = interleukin 6, MoCa = Montreal Cognitive Assessment, p-tau181 = phosphorylated tau protein 181



**Fig. 2 Multiple logistic regression models for predicting POD.** Model 1: Plasma log transformed preoperative IL-6 in conjunction with age predict POD with an AUC of 0.658 (95%-CI, 0.596, 0.714,  $p < 0.0001$ ). Model 2: log transformed preoperative IL-6 in conjunction with female sex and cognitive performance (MoCa assessment findings) predict POD with an AUC of 0.694 (95%-CI, 0.637, 0.747,  $p < 0.0001$ ). Model 3: preoperative log-transformed levels of p-tau181 and IL-6, sex and cognitive performance determine POD prediction with an AUC of 0.710 (95%-CI, 0.651, 0.763,  $p < 0.0001$ ). Abbreviation: AUC = area under the curve, CI = confidence interval, MoCa = Montreal Cognitive Assessment, ROC = receiver operating characteristics.



**Fig. 3 Decision tree for POD prediction.** Decision tree with the four most predictive variables for POD, ie, preoperative p-tau181 value, preoperative IL-6 and severity of mitral valve disease. Abbreviation: POD = postoperative delirium, p-tau 181 = phosphorylated tau protein 181, IL-6 = interleukin 6.

conjunction with mitral valve disease, cognitive performance, female sex, and age are important to enable the prediction of POD in patients undergoing cardiac surgery. This insight carries significant potential clinical relevance even though some factors such as mitral valve disease are known to be independent predictors of poor outcome [40]. A data-driven ML approach utilizing decision trees was instrumental in identifying specific

assessment criteria. These criteria, which include preoperative p-tau181 and IL-6 levels, as well as severe mitral valve disease, proved effective in predicting POD. Moreover, incorporating these blood biomarkers and human disease factors related to heart and brain diseases into a LASSO procedure yielded the highest accuracy in our analysis with an AUC of 0.751 in the training set, although the calculation in the validation set could not uphold the same AUC level with an AUC of 0.652, indicating a certain degree of overfitting. To date, no other study has explored the combined analysis of p-tau181 and IL-6 levels along with a broad range of heart and brain related factors as predictors of POD. Consequently, our predictive model entailing specific p-tau181 and IL-6 cut-off levels within the decision tree stands as a unique contribution to this field, offering a novel perspective in the assessment and management of POD risk.

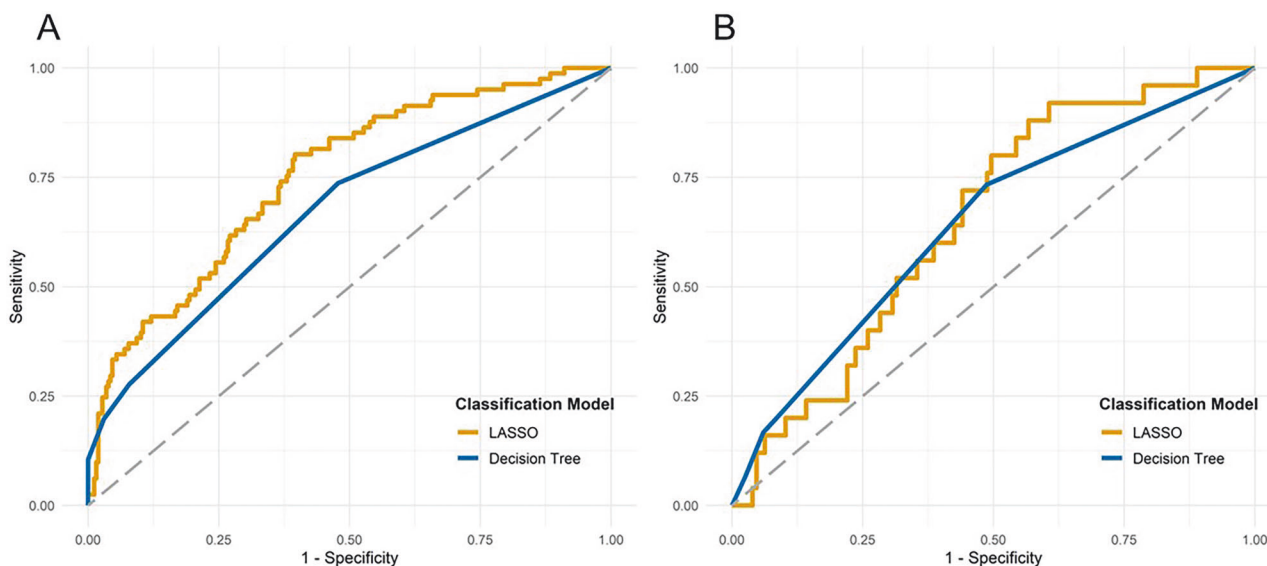
**Biomarker p-tau181 for predicting postoperative delirium**

In our study, we explored the potential of p-tau181 levels as a blood biomarker for predicting POD. We observed a significant predictive capability with an AUC of 0.641 in the ROC analysis. However, studies investigating p-tau181’s role in predicting non-cardiac POD showed a considerably higher AUC of 0.885 [15] and 0.869 [41]. This notable difference in AUC may be attributed to technical variations in the assays used: Liang et al. utilized established nanoneedle technology, whereas our study employed an immunoassay approach with p-tau181 measurement using Fujirebio’s Lumipulse device.

While the sample sizes in both studies were relatively comparable, the type of surgery (cardiac vs. hip vs. laparoscopic) differed, potentially contributing to the observed discrepancies in accurately diagnosing POD. Another recent study by Leung et al. [16] identified significant difference in p-tau181 levels between POD and non-POD patients. However, their study did not find a significant difference in distribution of p-tau181 levels between these groups in a Kolomorogov-Smirnov analysis. This contrasts with our findings and those of Liang et al. [15], possibly due to the fact that Leung et al. [16] examined a distinct population not undergoing cardiac surgery. Our investigation into the role of p-tau181 levels as a marker for POD in the context of cardiac surgery is novel, underscoring the biomarker’s potential importance in predicting susceptibility to neurodegeneration associated with POD. In addition, cognitive performance and female sex should be considered as factors for predicting POD when p-tau181 blood plasma levels serve as a blood biomarker-based strategy for predicting POD.

**Blood IL-6 as a postoperative delirium predictor**

Our study’s findings regarding the prediction of POD using preoperative versus postoperative IL-6 levels are in alignment with those from numerous other studies [5–7, 42–46]. Specifically, we validated the use of pre- and postoperative IL-6 levels as predictors of POD, obtaining similar to slightly lower AUC values in the ROC analysis, with preoperative IL-6 at 0.73 and postoperative IL-6 at 0.72 in patients undergoing cardiac surgery [5]. Our analysis revealed that preoperative IL-6 levels play a particularly significant role in predicting POD, with an AUC of 0.605 in the ROC curve. Recent research suggests that the soluble ectodomain of trigger receptor 2 (sTREM2) expressed on myeloid cells, which plays an important role in neuroglial inflammation, is modulated by IL-6 levels, which are already elevated in POD patients on the first postoperative day [47], thus revealing a key indication of a pathophysiological link between possible glial neuroinflammation and inflammation detectable in the blood via a rise in the proinflammatory marker IL-6. Moreover, IL-6 levels have been observed to rise steadily during initial hours following a coronary artery bypass graft surgery. Particularly, an IL-6 level exceeding a value of 583 pg/ ml after the 18th postoperative hour was associated with an increased incidence of POD [43]. Our



**Fig. 4** ROC curves of the classification models regularized LASSO and decision trees. ROC analysis revealed that the classification model decision tree yielded significantly predictive accuracy of POD for the training (A) and validation (B) datasets [A, B:  $p < 0.05$ ; AUC for the training dataset of 0.672 (95%-CI: 0.604, 0.735,  $p < 0.0001$ ); AUC for the validation dataset of 0.642 (95%-CI: 0.537, 0.738,  $p = 0.0108$ )]. LASSO also predicted POD accurately with an AUC of 0.751 (95%-CI: 0.686, 0.805,  $p < 0.0001$ ) for the training dataset (A) and 0.652 (95%-CI: 0.538, 0.747,  $p = 0.0086$ ) for the validation dataset (B). Abbreviations: CI = confidence interval, ROC = receiver operating characteristics, POD = postoperative delirium.

results, therefore, suggest the need for a more nuanced combination of blood biomarkers to improve the accuracy of POD prediction, as indicated by the outcomes of our ROC analysis. It is also important to consider age as a factor in predicting POD when relying on the blood biomarker IL-6 to predict delirium.

#### Combining plasma biomarkers with other factors to predict POD

The observed elevated postoperative IL-6 levels in patients with POD underscores the critical role of ongoing inflammation in the development of POD. This connection was recently confirmed in a study by Vasunilashorn et al. [48], who identified inflammation through different proteomics as a common pathway of progression and risk for POD. Complementing this perspective, a study by Liu et al. [42] reported that patients experiencing delirium showed increases in plasma neurofilament light chains and glial fibrillary acid protein, markers indicative of axonal injury and glial inflammation, a pattern not seen in non-delirium patients. The novelty of our study is not that we investigated the general neuroinflammatory marker IL-6, but that we combined it with a cell destruction marker and the link between increased neuroinflammation associated with POD and subsequent axonal brain damage. To our knowledge, no prior study has integrated such biomarkers. Although our ROC analysis showed only a modest increase in the AUC (0.671) compared to studies focusing solely on IL-6 and p-tau181 biomarkers, this outcome may be partly attributed to the significant influence of age, as revealed by the logistic regression analysis. Future research should therefore include subgroup analyses of potential patients undergoing heart surgery, particularly focusing on the incidence of POD. Such studies could identify patient subpopulations that might benefit most from POD prediction using these combined biomarkers. Taken together, we demonstrate that patients who developed POD presented higher preoperative plasma concentrations of IL-6 or p-tau181 than those who did not develop POD. The preoperative plasma concentration of IL-6 or p-tau181 predicted the presence and severity of POD, with p-tau181 being most closely associated with that outcome. Our multiple logistic regression model further revealed female sex as a relevant POD-

predicting factor in this context. Therefore, female gender should be included in predicting delirium in conjunction with the plasma biomarkers IL-6 and p-tau181 in future studies.

#### Age, cognitive performance, and blood biomarkers predicting POD

Previous research suggests that both age and preoperative cognitive status are risk factors for the development of POD [49–51], but few studies have examined these relationships while accounting for the effects of inflammation and neuronal cell injury. In our analyses, age was associated significantly with POD in Model 1 (when IL-6 was included in the model), but not in Model 2 (when adjusted for p-tau181) or Model 3 (when adjusted for IL-6, ptau181, and MoCA). These results suggest that age is related to the POD risk even when accounting for preoperative inflammation (as determined by blood IL-6 levels; Model 1). Cognitive performance (assessed by MoCA score), in contrast, emerged as a particularly relevant risk factor when considering the presence of preoperative neuronal injury (p-tau181, Model 2) and both inflammatory and neurodegenerative markers simultaneously (Model 3). While both age and cognitive performance are associated with POD risk, our analyses also indicate that the contribution of these factors is less pronounced than those of inflammation (IL-6, Model 1: OR 1.35) and neurodegeneration (p-tau181, Model 2: OR 1.63). These findings are in line with previous studies [50, 51] confirming that both age and cognitive status are important risk factors for POD, but they also highlight that their effects should be interpreted in the context of biological markers of inflammation and neurodegeneration. Notably, another study confirmed - consistent with Model 2 - that preoperative plasma p-tau181 is an age-independent predictor of POD [41]. The odds ratios suggest that pre-existing neuronal cell damage (p-tau181, OR = 1.63, Model 2) may be a more decisive factor for POD than preoperative inflammation (IL-6, OR = 1.35, Model 1).

#### Limitations

Despite the fact that all our study participants were undergoing cardiac surgery, our cohort was quite heterogeneous. This diversity stemmed from a variety of heart surgeries performed

and the broad range of ages among patients. This heterogeneity required us to account for a multitude of somatic factors that were evidently influential in the patients' need for cardiac surgery. Furthermore, the study was conducted at a single center, which may have caused institutional bias. These limitations underscore the need for further research, possibly enhancing the predictive accuracy for POD. Another limitation concerns age-related cognitive decline, dementia, and neurodegeneration as potential confounders of POD. To address this, we incorporated age into our multiple regression analyses, lowered the minimum inclusion age to 50 years, and excluded individuals with a preoperative diagnosis of dementia or neurodegeneration. Overfitting is another problem that must be considered when selecting predictors [52]. The LASSO procedure enforces sparsity of the model and reduces the number of covariables as a result. By applying a LASSO procedure in combination with cross-validation, overfitting can be reduced. This can ultimately increase the generalizability of our conclusions.

## CONCLUSIONS

In this novel analysis, which included preoperative cardiac factors and markers of inflammation and neurodegeneration, we found that a combination of plasma biomarkers (p-tau181 and IL-6) along with specific heart related factors such as mitral valve disease, cognitive performance, age, and female sex can help predict the occurrence of POD in patients undergoing cardiac surgery. Our data suggest that pre-existing neuronal injury may be a slightly stronger predictor of POD than preoperative inflammation. Furthermore, our findings indicate that the interaction between preoperative neuronal injury and inflammation does not appear to contribute meaningfully to POD risk. ML approaches, such as decision trees and LASSO, offer additional tools for predicting and understanding the risk factors associated with POD. Our findings emphasize the importance of taking a comprehensive approach to assess and manage the POD risk in this patient population. The associations and predictors of POD we identified provide a foundation for further research in this field. Confirming these associations and delving into the underlying mechanisms will contribute to a deeper understanding of how POD manifests in the context of cardiac surgery. Such research can pave the way for developing more effective interventions for preventing or managing POD in this specific patient population. This might have potential clinical relevance in identifying patients carrying a higher risk for POD before surgery, aiming to prevent life-threatening complications.

## DATA AVAILABILITY

The FINDERI dataset generated and/or analyzed for the present article is available upon request to the corresponding author. The investigators follow the FAIR Principles for scientific data management and stewardship.

## REFERENCES

1. Staveski SL, Pickler RH, Khoury PR, Ollberding NJ, Donnellan AL, Mauney JA, et al. Prevalence of ICU delirium in postoperative pediatric cardiac surgery patients. *Pediatr Crit Care Med*. 2021;22:68–78.
2. Chen H, Mo L, Hu H, Ou Y, Luo J. Risk factors of postoperative delirium after cardiac surgery: a meta-analysis. *J Cardiothorac Surg*. 2021;16:113.
3. O'Neal JB, Shaw AD. Predicting, preventing, and identifying delirium after cardiac surgery. *Perioper Med (Lond)*. 2016;5:7.
4. van Munster BC, Korse CM, de Rooij SE, Bonfrer JM, Zwinderman AH, Korevaar JC. Markers of cerebral damage during delirium in elderly patients with hip fracture. *BMC Neurol*. 2009;9:21.
5. Lv XC, Lin Y, Wu QS, Wang L, Hou YT, Dong Y, et al. Plasma interleukin-6 is a potential predictive biomarker for postoperative delirium among acute type a

- aortic dissection patients treated with open surgical repair. *J Cardiothorac Surg*. 2021;16:146.
6. Ayob F, Lam E, Ho G, Chung F, El-Beheiry H, Wong J. Preoperative biomarkers and imaging tests as predictors of post-operative delirium in non-cardiac surgical patients: a systematic review. *BMC Anesthesiol*. 2019;19:25.
7. Zhang S, Ji MH, Ding S, Wu Y, Feng XW, Tao XJ, et al. Inclusion of interleukin-6 improved performance of postoperative delirium prediction for patients undergoing coronary artery bypass graft (POD-CABG): A derivation and validation study. *J Cardiol*. 2022;79:634–41.
8. Xiao MZ, Liu CX, Zhou LG, Yang Y, Wang Y. Postoperative delirium, neuroinflammation, and influencing factors of postoperative delirium: A review. *Medicine (Baltimore)*. 2023;102:e32991.
9. Nabizadeh F, Balabandian M, Rostami MR, Ward RT, Ahmadi N, Pourhamzeh M, Alzheimer's disease Neuroimaging Initiative (ADNI). Plasma p-tau181 associated with structural changes in mild cognitive impairment. *Aging Clin Exp Res*. 2022;34:2139–47.
10. Fong TG, Vasunilashorn SM, Kivisäkk P, Metzger ED, Schmitt EM, Marcantonio ER, et al. Biomarkers of neurodegeneration and neural injury as potential predictors for delirium. *Int J Geriatr Psychiatry*. 2024;39:e6044.
11. Sunde AL, Alsnes IV, Aarsland D, Ashton NJ, Tovar-Rios DA, De Santis G, et al. Preanalytical stability of plasma biomarkers for Alzheimer's disease pathology. *Alzheimers Dement (Amst)*. 2023;15:e12439.
12. Ballweg T, White M, Parker M, Casey C, Bo A, Farahbakhsh Z, et al. Association between plasma tau and postoperative delirium incidence and severity: a prospective observational study. *Br J Anaesth*. 2021;126:458–66.
13. Jack CR Jr, Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeberlein SB, et al. NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimers Dement*. 2018;14:535–62. Contributors
14. Parker M, White M, Casey C, Kunkel D, Bo A, Blennow K, et al. Cohort Analysis of the Association of Delirium Severity With Cerebrospinal Fluid Amyloid-Tau-Neurodegeneration Pathologies. *J Gerontol A Biol Sci Med Sci*. 2022;77:494–501. <https://doi.org/10.1093/gerona/glab203>
15. Liang F, Baldyga K, Quan Q, Khatri A, Choi S, Wiener-Kronish J, et al. Preoperative Plasma Tau-PT217 and Tau-PT181 Are Associated With Postoperative Delirium. *Ann Surg*. 2023;277:e1232–e1238.
16. Leung JM, Rojas JC, Tang C, Chan B, Lario-Lago A, Boxer AL, et al. Presence of preoperative neurodegeneration biofluid markers in patients with postoperative delirium. *Anesthesiology*. 2023;139:432–43. Perioperative Medicine Research Group
17. Capogna E, Watne LO, Sørensen Ø, Guichelaar CJ, Idland AV, Halaas NB, et al. Associations of neuroinflammatory IL-6 and IL-8 with brain atrophy, memory decline, and core AD biomarkers - in cognitively unimpaired older adults. *Brain Behav Immun*. 2023;113:56–65.
18. Sadlonova, Vogelgsang M, Lange J, Günther C, Wiesent I, Eberhard A, et al. Identification of risk factors for delirium, cognitive decline, and dementia after cardiac surgery (FINDERI-find delirium risk factors): a study protocol of a prospective observational study. *BMC Cardiovasc Disord*. 2022;22:299.
19. Sadlonova M, Hansen N, Esselmann H, Celano CM, Derad C, Asendorf T, et al. Preoperative delirium risk screening in patients undergoing a cardiac surgery: results from the prospective observational FINDERI study. *Am J Geriatr Psychiatry*. 2024;32:835–51. FINDERI investigators
20. Itting PT, Sadlonova M, Santander MJ, Knierim M, Derad C, Asendorf T, et al. and FINDERI Investigators. Intra- and early postoperative predictors of delirium risk in cardiac surgery: results from the prospective observational FINDERI study. *Int J Surg*. 2025;111:2872–85.
21. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Int J Surg*. 2014;12:1495–9.
22. Ely EW, Inouye SK, Bernard GR, Gordon S, Francis J, May L, et al. Delirium in mechanically ventilated patients: validity and reliability of the confusion assessment method for the intensive care unit (CAM-ICU). *JAMA*. 2001;286:2703–10.
23. Ely EW, Margolin R, Francis J, May L, Truman B, Dittus R, et al. Evaluation of delirium in critically ill patients: validation of the confusion assessment method for the intensive care unit (CAM-ICU). In: *Critical care medicine*. 2001;29:1370–9.
24. Inouye SK, van Dyck CH, Alessi CA, Balkin S, Siegel A, et al. Clarifying confusion: the confusion assessment method. A new method for detection of delirium. *Annals of internal medicine*. 1990;113:941–8.
25. Thomas C, Kreisel SH, Oster P, Driessen M, Arolt V, Inouye SK. Diagnosing delirium in older hospitalized adults with dementia: adapting the confusion assessment method to international classification of diseases, tenth revision, diagnostic criteria. *J Am Geriatr Soc*. 2012;60:1471–7.

26. Ely EW, Truman B, Shintani A, Thomason JW, Wheeler AP, Gordon S, et al. Monitoring sedation status over time in ICU patients: reliability and validity of the Richmond Agitation-Sedation Scale (RASS). *JAMA*. 2003;289:2983–91.
27. Green JR, Smith J, Teale E, Collinson M, Avidan MS, Schmitt EM, et al. Use of the confusion assessment method in multicentre delirium trials: training and standardisation. *BMC Geriatr*. 2019;19:107.
28. Inouye SK, Foreman MD, Mion LC, Katz KH, Cooney LM Jr. Nurses' recognition of delirium and its symptoms: comparison of nurse and researcher ratings. *Arch Intern Med*. 2001;161:2467–73.
29. Inouye SK, Sharon K, Kosar CM, Tommet D, Schmitt EM, Puelle MR, et al. The CAM-S: development and validation of a new scoring system for delirium severity in 2 cohorts. *Annals of internal medicine*. 2014;160:526–33.
30. Hestermann U, Backenstrass M, Gekle I, Hack M, Mundt C, Oster P, et al. Validation of a German version of the confusion assessment method for delirium detection in a sample of acute geriatric patients with a high prevalence of dementia. *Psychopathology*. 2009;42:270–6.
31. Nasreddine Z, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The montreal cognitive assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*. 2005;53:695–9.
32. R Core Team. R: A Language and Environment for Statistical Computing. Version 4.3.1, 2023. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
33. Konietzschke F, Placzek M, Schaarschmidt F, Hothorn LA. nparcomp An R software package for nonparametric multiple comparisons and simultaneous confidence intervals. *J. Stat. Soft.* 2015;64:1–17.
34. López-Ratón M, Rodríguez-Álvarez MX, Suárez CC, Sampedro FG. Optimal cut-points An R package for selecting optimal cutpoints in diagnostic tests. *J. Stat. Soft.* 2014;61:1–36.
35. Stevenson M, Sergeant E, Nunes T, Heuer C, Marshall J, Sanchez J, et al. epiR: Tools for the analysis of epidemiological data. R package version 2.0.63; 2023.
36. Therneau, T & Atkinson, B rpart: Recursive Partitioning and Regression Trees: R package version 4.1.19; 2022.
37. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J. Stat. Soft.* 2010. <https://doi.org/10.18637/jss.v033.i01>
38. van Buuren, S & Groothuis-Oudshoorn, K mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*. 2011. <https://doi.org/10.18637/jss.v045.i03>.
39. Kuhn M Building predictive models in R using the caret package. *J. Stat. Soft.* 2008. <https://doi.org/10.18637/jss.v028.i05>.
40. Oliveira FR, Oliveira VH, Oliveira IM, et al. Hypertension, mitral valve disease, atrial fibrillation and low education level predict delirium and worst outcome after cardiac surgery in older adults. *BMC Anesthesiol*. 2018;18:15.
41. Wu LX, Su JB, Li GL, Zhuang M, Sun HJ, Zhang JR Preoperative Plasma p-tau231, p-tau181 and p-tau217 are associated with Postoperative Delirium: a prospective study. *J Gerontol A Biol Sci Med Sci*. 2025;glaf058.
42. Lin JL, Zheng GZ, Chen LW, Luo ZR. A nomogram model for assessing predictors and prognosis of postoperative delirium in patients receiving acute type A aortic dissection surgery. *BMC Cardiovasc Disord*. 2023;23:72.
43. Chen Y, Lu S, Wu Y, Shen Y, Zhao H, Ding S, et al. Change in serum level of interleukin 6 and delirium after coronary artery bypass graft. *Am J Crit Care*. 2019;28:462–70.
44. Liu P, Li YW, Wang XS, Zou X, Zhang DZ, Wang DX, et al. High serum interleukin-6 level is associated with increased risk of delirium in elderly patients after noncardiac surgery: a prospective cohort study. *Chin Med J (Engl)*. 2013;126:3621–7.
45. Adamis D, van Gool WA, Eikelenboom P. Consistent patterns in the inconsistent associations of Insulin-like growth factor 1 (IGF-1), C-Reactive Protein (C-RP) and Interleukin 6 (IL-6) levels with delirium in surgical populations. A systematic review and meta-analysis. *Arch Gerontol Geriatr*. 2021;97:104518.
46. Imai T, Morita S, Hasegawa K, Goto T, Katori Y, Asada Y. Postoperative serum interleukin-6 level as a risk factor for development of hyperactive delirium with agitation after head and neck surgery with free tissue transfer reconstruction. *Auris Nasus Larynx*. 2023;50:777–82.
47. Ruhnau J, Müller J, Nowak S, Strack S, Sperlich D, Pohl A, et al. Serum biomarkers of a Pro-neuroinflammatory state may define the pre-operative risk for post-operative delirium in spine surgery. *Int J Mol Sci*. 2023;24:10335.
48. Vasunilashorn SM, Dillon ST, Marcantonio ER, Libermann TA. Application of multiple omics to understand postoperative delirium pathophysiology in humans. *Gerontology*. 2023;69:1369–84. <https://doi.org/10.1159/000533789>
49. Li J, Zou P, Zhou Y, Wang J, Zhang Y, Tan X, et al. Advanced age is significantly associated with poor outcomes of thoracic endovascular aortic repair: a systematic review and meta-analysis. *BMC Surg*. 2025;25:289 <https://doi.org/10.1186/s12893-025-02990-x>
50. Choi Y, Kim EM, Lee H. Predicting postoperative delirium in orthopedic surgery using explainable artificial intelligence. *Stud Health Technol Inform*. 2025;329:1940–1. <https://doi.org/10.3233/SHTI251289>
51. Wu SJ, Sharma N, Bauch A, Yang HC, Hect JL, Thomas C, et al. Predicting post-operative delirium in older patients before elective surgery: Multicenter retrospective cohort study. *JMIR Aging*. 2025;8:e67958 <https://doi.org/10.2196/67958>. PAWEL Study Group
52. McNeish DM. Using lasso for predictor selection and to assuage overfitting: a method long overlooked in behavioral sciences. *Multivariate Behav Res*. 2015;50:471–84. <https://doi.org/10.1080/00273171.2015.1036965>

## ACKNOWLEDGEMENTS

The authors wish to thank Anke Jahn-Brodmann and Hans Klafki for training and supervision according to measurement of biomarkers. The authors gratefully acknowledge the contributions of the patients, family members, nurses, physicians, staff members, and the FINDERI study group (Claudia Lange, Adriana Wiesent, Jessika Schmitz, Frederike E. Bauer, Julia Ehrentraut, Annika Graß, Rebecca Arentz, Maria Dornieden, Charlotte Eberhard, Mirjam F. Schröder, Anne-Catherine Thiel, Melanie Hoteit, Sophie Dörfler, Maria Alexy, Theresa Kruck, Tendai Chiwakata, Marianne Richter, Julia Kühnle, Paul Thomas Itting, Maria Knierim, Manuel Johannes Santander Martinez). We thank Carole Cürten for editing as well as proofreading the English language in this manuscript.

## AUTHOR CONTRIBUTIONS

Contribution: Conceptualization, MS, NH, HE, TA, MC, IK, JW, HB, CAFvA; investigation, MS, JS, AW, FEB, JE; methodology, project administration, data curation, MS, CAFvA, HB, JW, NH, HE, CD, TA; statistical analyzes, CD, TA, MS, CMC; writing—original draft, MS, CMC, CD, TA; writing—review & editing, NH, SH, JW, HB, CAFvA, funding acquisition, MC, MS; supervision MS, HB, CAFvA, JW, NH, HE. All authors have read and approved the final manuscript.

## FUNDING

This study is funded by a grant from the German Heart Foundation/German Foundation for Heart Research (#F/65/21). CMC is supported by the National Heart, Lung, and Blood Institute through grant R01HL155301. JW is supported by an Ilídio Pinho professorship, iBiMED (UIDB/04501/2020) at the University of Aveiro, Portugal. Open Access funding enabled and organized by Projekt DEAL.

## COMPETING INTERESTS

CMC has received funding from the National Heart, Lung, and Blood Institute (grant R01HL155301) and stipends from Elsevier for editorial work for General Hospital Psychiatry. JW has served at scientific advisory boards for Abbot, Biogen, Boehringer-Ingelheim, Eli Lilly, F. Hoffmann-La Roche, Immunogenetics, MSD SHARP & DOHME, he has also received honorarium for lectures sponsored by Eli Lilly, Pfizer, Janssen, MSD SHARP & DOHME, Amgen, Roche Pharma, Actelion Pharmaceuticals, Guangzhou Glorylen Medical Technology Co. (China), Beijing Yibai Science and Technology Ltd. CAFvA has received honoraria from serving on the scientific advisory board of Biogen, Roche, Lilly and Dr. Willmar Schwabe GmbH & Co. CAFvA has also received funding for travel and speaker honoraria from Biogen, Roche diagnostics AG, Novo Nordisk, MedTriX GmbH, Landesvereinigung für Gesundheit und Akademie für Sozialmedizin Niedersachsen e. V., Medical Tribune Verlagsgesellschaft mbH, Novartis, FomF GmbH, Paracelsus-Kliniken Deutschland, Schön Klinik Bad Aibling and Dr. Willmar Schwabe GmbH & Co. NH has received funding support from the DFG (project number 5302297989) and receives travel support from Eli Lilly and Company. MS, HE, CD, TA, MC, SH, AW, JS, FEB, JE, IG, HB report no competing interests.

## ETHICS APPROVAL AND CONSENTS TO PARTICIPATE

The study protocol was approved by the Ethics Committee of the University of Göttingen Medical Center on February 16, 2021 (#20/11/20). Written informed consent is obtained from each participant before any trial-related procedures are performed and patients receive a copy of the signed and dated written consent form. All methods were performed in accordance with the relevant guidelines and regulations.

**ADDITIONAL INFORMATION**

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41380-025-03412-3>.

**Correspondence** and requests for materials should be addressed to Niels Hansen.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025, modified publication 2025

**FINDERI INVESTIGATORS**

Adriana Wiesent<sup>2</sup>, Alina Isabel Rediske<sup>1</sup>, Anke Jahn-Brodmann<sup>1</sup>, Anne-Catherine Thiel<sup>2</sup>, Annika Gaß<sup>2</sup>, Barbara Morgado<sup>1</sup>, Björn Hendrik Schott<sup>1</sup>, Britta Albert<sup>2</sup>, Carlotta Derad<sup>5</sup>, Charlotte Eberhard<sup>7</sup>, Christine A. F. von Arnim<sup>2,8</sup>, Christopher M. Celano<sup>3,4</sup>, Clara Maria Knopp<sup>2</sup>, Frederike E. Bauer<sup>2</sup>, Hassina Baraki<sup>7,8</sup>, Hermann Esselmann<sup>1</sup>, Ihtaz Malik<sup>2</sup>, Ingo Kutschka<sup>7,8</sup>, Irina Günther<sup>1</sup>, Iryna Krasiuk<sup>1</sup>, Jens Wiltfang<sup>1,9,10</sup>, Jessica Schmitz<sup>2</sup>, Julia Ehrentraut<sup>2</sup>, Katharina Packroß<sup>1</sup>, Maike Hohberg<sup>5</sup>, Manuel J. Santander<sup>7</sup>, Maria Alexy<sup>2</sup>, Maria Dornieden<sup>2</sup>, Maria Knierim<sup>7</sup>, Marianne Richter<sup>5</sup>, Matilda-Marie Becker<sup>1</sup>, Melania Hotheit<sup>7</sup>, Michael Moser<sup>2</sup>, Miriam F. Schröder<sup>7</sup>, Mohammed Chebbok<sup>2,6</sup>, Monika Sadlonova<sup>2,3,8,11,12</sup>, Nicholas Paul Süttmann<sup>1</sup>, Niels Hansen<sup>1</sup>, Paul T. Itting<sup>7</sup>, Rebecca Arentz<sup>2</sup>, Sandra Kastenbauer<sup>2</sup>, Santander Martinez<sup>7</sup>, Manuel Johannes<sup>7</sup>, Sophie Doerfler<sup>2</sup>, Stephanie Heinemann<sup>1,2</sup>, Tendai Chiwakata<sup>2</sup>, Theresa Maria Kruck<sup>2</sup>, Thomas Asendorf<sup>5</sup> and Tobias Titsch<sup>1</sup>