

RESEARCH ARTICLE

WILEY

Impact of sample size and regression of tissue-specific signals on effective connectivity within the core default mode network

Alexander N. Silchenko¹  | Felix Hoffstaedter^{1,2}  | Simon B. Eickhoff^{1,2} 

¹Institute of Neuroscience and Medicine, Brain & Behaviour (INM-7), Research Center Jülich, Jülich, Germany

²Institute of Systems Neuroscience, Medical Faculty, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

Correspondence

Alexander N. Silchenko, Institute of Neuroscience and Medicine, Brain & Behaviour (INM-7), Research Center Jülich, Jülich, Germany.
Email: a.silchenko@fz-juelich.de

Funding information

Horizon 2020 Framework Programme, Grant/Award Numbers: 945539, 826421; IRTG 2150; US National Institute of Health, Grant/Award Number: R01MH074457; SFB, Grant/Award Number: 1451; DFG

Abstract

Interactions within brain networks are inherently directional, which are inaccessible to classical functional connectivity estimates from resting-state functional magnetic resonance imaging (fMRI) but can be detected using spectral dynamic causal modeling (DCM). The sample size and unavoidable presence of nuisance signals during fMRI measurement are the two important factors influencing the stability of group estimates of connectivity parameters. However, most recent studies exploring effective connectivity (EC) have been conducted with small sample sizes and minimally pre-processed datasets. We explore the impact of these two factors by analyzing clean resting-state fMRI data from 330 unrelated subjects from the Human Connectome Project database. We demonstrate that both the stability of the model selection procedures and the inference of connectivity parameters are highly dependent on the sample size. The minimum sample size required for stable DCM is approximately 50, which may explain the variability of the DCM results reported so far. We reveal a stable pattern of EC within the core default mode network computed for large sample sizes and demonstrate that the use of subject-specific thresholded whole-brain masks for tissue-specific signals regression enhances the detection of weak connections.

1 | INTRODUCTION

Resting-state functional magnetic resonance imaging (fMRI) is the standard tool for the investigation of brain network connectivity. Most studies characterizing the intrinsic organization of resting-state networks are based on functional connectivity (FC), which is defined as statistical dependencies among observed neurophysiological responses (Biswal et al., 1995). FC has been shown to encode intellectual performance (Song et al., 2008; Van Den Heuvel et al., 2009), monitoring of external environment (Gusnard & Raichle, 2001), and emergence of stimulus-independent thoughts (Mason et al., 2007).

Despite the great promise for advancing our understanding of the cognitive abilities that underlie intellectual feats, such brain-wide association studies require consortium-level ($N > 1000$) sample sizes to avoid under-powered correlations and statistical errors (Marek et al., 2022). The expense and effort necessary for obtaining and analyzing large fMRI datasets have resulted in numerous small-sample studies with exceedingly low replicability (Poldrack et al., 2017; Szucs & Ioannidis, 2020). Remarkably, sampling variability alone is sufficient to generate nominally significant ($p < 0.05$) but inflated correlations (Marek et al., 2022), whereas the common removal of global nuisance alters the variance of the residual signal, thereby modifying

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2023 The Authors. *Human Brain Mapping* published by Wiley Periodicals LLC.

the resulting correlation structure (Fox et al., 2009; Friston, 2011). Thus, the stability and reliability of FC analyses are strongly limited by the requirement of large sample sizes and control of noise in fMRI signals. Another disadvantage of resting-state FC is its inability to capture the directionality of the observed interactions between brain regions. A dynamic causal modeling (DCM) approach can be applied to infer causal (directed) influences between regions (Friston et al., 2003). The question remains as to whether there are similar limitations for the stable detection of causal interactions between different brain regions as in FC?

The basic idea of DCM is to use Bayesian model inversion to estimate the effective connectivity (EC) among neuronal populations from observed BOLD signals (Friston et al., 2003; Stephan et al., 2010). Stochastic DCM, initially developed to model task-driven changes in the EC between brain regions (Li et al., 2012) has been successfully modified for the analysis of resting-state fMRI data (Friston et al., 2014). The deterministic spectral DCM version uses a power-law function in the spectral domain to model the neuronal fluctuations. Such fitting of second-order data features makes this scheme estimation computationally and statistically much more efficient than stochastic DCM (Razi et al., 2015). Spectral DCM has recently been validated using simulated and empirical data for both small and large brain networks (Razi et al., 2017). Additionally, the within-subject reliability of spectral DCM with respect to subjects' conditions and various processing parameters (region of interest [ROI] size and global signal regression [GSR]) was demonstrated in a longitudinal study (Almgren et al., 2018). Despite the growing popularity of DCM, small sample sizes (20–30) are still common in recent studies, possibly leading to divergent estimates of connectivity parameters at the group level. Indeed, even in the case of relatively small networks such as the core default mode network (DMN), the patterns of EC found in different studies are not fully consistent (Di & Biswal, 2014; Friston et al., 2014; Li et al., 2012; Razi et al., 2015; Sharaev et al., 2016). Thus, the problem of a minimal sample size, required for a robust DCM analysis, is raised for studies of EC. The main problem with under-sampled DCM studies, however, is that having few subjects leads to a relatively small joint model space, which impedes or even prohibits true parameters estimates due to high inter-subject variability. In other words, the correct parameter estimates can be left to chance or always overestimated because of strong inter-individual variability (Button et al., 2013; Lindquist et al., 2013).

In the current study, we evaluate the impact of sample size on the stability of causal interactions between brain regions in Human Connectome Project (HCP) data to investigate whether this is also a severe limitation as reported for classical resting-state FC. We explore the stability of the group EC estimates within the core DMN with increasing sample size and different combinations of GSR in the comparatively large HCP sample. Here, we used public resting-state fMRI data from 330 unrelated subjects from the HCP database of young adults. Time series are made available and cleaned from movement-related and physiological artifacts using the well-tested ICA-FIX method (Salimi-Khorshidi et al., 2014). This effective noise reduction should permit a representative definition of the lower boundary of

sample sizes for resting-state DCM, for which common inter-subject variability allows for stable connectivity estimates.

The core DMN is a well-established intrinsic resting-state network consisting of the ventral medial prefrontal cortex (mPFC), precuneus/posterior cingulate cortex (PCC) and is also widely used in recent EC studies (Di & Biswal, 2014; Friston et al., 2014; Li et al., 2012; Razi et al., 2015; Sharaev et al., 2016). In his influential review, Raichle suggested that the core DMN mediates internal modes of cognitive activity (Raichle, 2015). In a review and meta-analysis on variability in the healthy DMN, Mak et al. (2017) found that connectivity strength seems to follow an inverse U-shape with the strongest coupling in adulthood and weaker FC in children and the elderly. With regard to cognitive performance, there is some evidence for an association between cognitive performance and FC, as well as with task-induced deactivation of DMN regions, while these effects are pronounced with aging and clinical conditions in particular. Our study of EC during the resting state aims to reveal a stable connectivity pattern and shed light on the roles played by different DMN regions in the formation of sustained activity patterns. The HCP datasets for young adults cleaned from artifacts can be the ideal choice for such study due to the expected strong interaction within the core DMN, allowing to reveal the patterns of causal interactions.

The presence of noise in fMRI data is a manifestation of many different confounding sources (Behzadi et al., 2007; Liu, 2016; Liu et al., 2017). A method widely used to denoise fMRI signals is to correct resting-state fMRI time-series for fluctuations in the global signal, which is the average signal across all voxels of the entire magnetic resonance imaging (MRI) volume. GSR has been shown to enhance the efficiency of detecting significant FC (Fox et al., 2009; Liu et al., 2017). Moreover, the effects of GSR on EC within and between resting-state networks have recently been studied in Almgren et al. (2020). It was found that the effect of GSR on between-network EC, averaged over all connections, was negligible to small, whereas the effect of GSR on individual connections was moderate but non-negligible (Almgren et al., 2020).

Here, we concentrated on tissue-specific nuisance signals from the white matter (WM) and cerebrospinal fluid (CSF), which are commonly used for filtering fMRI data, because signals of neuronal origin are not present in these compartments. These nuisance signals are computed via extraction of fMRI time series from the CSF and WM masks, and subsequent computation of representative signals using principal component analysis (Liu, 2016). The cleanup of physiological noise is known to improve FC results and is an essential step in analysis pipelines (Murphy et al., 2013). In particular, the use of individual (subject-specific) masks with a high probability threshold during the extraction of WM and CSF signals enhances filtering performance by increasing the temporal signal-to-noise ratio (Bartoň et al., 2019). Is there any impact of such individual (subject-specific) brain masking on the stability of the EC estimates? To reveal the impact of tissue-specific signals regression (TSSR) on the results of spectral DCM and its possible influence on the optimal sample size estimates, all measures were computed for the different whole-brain masks and combinations of WM, CSF and grey matter (GM) global signals regression.

2 | METHODS

2.1 | Datasets and preprocessing

The datasets used in our study included 330 unrelated subjects (163 females, mean age 28.48 and SD 3.43) and were obtained from the 1200 subjects ICA-FIX denoised release of the HCP's database (Van Essen et al., 2012). All HCP imaging data were acquired on a customized Siemens 3T Skyra at Washington University (St Louis) using a multiband sequence. Whole-brain resting-state fMRI images were acquired with a spatial resolution $2 \times 2 \times 2$ mm and a temporal resolution 0.72 s. Two sessions of resting-state fMRI data were collected on consecutive days for each participant, and each session consisted of two runs. The length of each resting-state fMRI scan was 14.4 min (1200 frames). Details of data collection can be found elsewhere (Smith et al., 2013; Van Essen et al., 2013). To minimize the possible influence of different artifacts on the observed patterns of EC, we used the ICA-FIX denoised HCP dataset comprising four sessions for each subject (Salimi-Khorshidi et al., 2014). Subsequently, a quality check of the ICA-FIX denoised HCP datasets was performed for every particular 4D dataset using the DVARS toolbox. The intensity of the HCP data was normalized and spatially transformed to MNI152 space using FMRIB Software Library (Glasser et al., 2013). We further increased the signal-to-noise ratio of the fMRI data in spatial parametric mapping package (SPM12b revision 7219) by applying spatial smoothing using a 5 mm Gaussian kernel (Hillebrandt et al., 2014). The SPM12b package (revision 7219) was used to fit the first level GLMs and compute BOLD signals as the principal eigenvariates of ROIs masked by spheres (8 mm radius) for each of the nodes comprising the core DMN. The ROI center coordinates were ($x=3$, $y=54$, $z=-2$) for mPFC, ($x=0$, $y=-52$, $z=26$) for PCC, ($x=-50$, $y=-63$, $z=32$) for the left inferior parietal cortex (LIPC) and ($x=48$, $y=-69$, $z=32$) for the right inferior parietal cortex (RIPC).

2.2 | Effective connectivity: First level analysis with spectral DCM

The causal interactions between sub-regions forming the core DMN can be examined using DCM (Friston et al., 2003). DCM consists of two main components: a hemodynamic forward model that describes the transformation of synaptic activity to the hemodynamic response and a bilinear model that describes how activity changes as a function of inputs, connections, and modulations (Stephan et al., 2010). The bilinear model consists of the following equation:

$$\frac{dz}{dt} = \left[A + \sum_{j=1}^m u_j B_j \right] z + Cu, \quad (1)$$

where z is the state of the system describing activity level of each region in the modeled system and u denotes inputs into the system,

which are the external inputs applied to the system. Given the values of z and u , the dynamics of the system (1) is estimated via calculation of the matrix A , which describes the fixed EC between sub-regions, the matrix C describing the sensitivity of different regions to external inputs, and the matrix B describing the modulations of connectivity as a function of experimental manipulation j .

Here, we use spectral DCM, which models the FC (statistical dependencies) among the time series using cross-spectral density as a measure for fitting data features in the frequency domain (Friston et al., 2014). Estimation in the frequency domain is significantly quicker, more efficient, and more sensitive to group differences (Razi et al., 2015; Seguin et al., 2019). All the four sessions of the HCP dataset were used in this study. Four fully connected DCM models were defined and inverted for every subject using spectral DCM to obtain posterior estimates of the neuronal parameters using the same standard priors (Friston et al., 2016). All four session-specific individual DCMs fitted for every subject were used jointly to maximize the dimensionality of the model space and make more accurate model selection and connectivity parameter estimates (Friston et al., 2016; Zeidman, Jafarian, Corbin, et al., 2019).

2.3 | Effective connectivity: Second level group analysis

Next, we used the Parametric Empirical Bayes (PEB) approach to model how individual (within-subject) connections are related to group means (Zeidman, Jafarian, Seghier, et al., 2019). In accordance with this approach, intrinsic connectivity is treated as a random (between-subject) effect, which is modeled by adding a random Gaussian component to subject-specific parameters. This random-effects modeling allows the use of the full posterior density (the expected strength of each connection and the associated covariance) over the parameters from each subject's DCM for the estimation of posterior expectation and uncertainty at the group level (Friston et al., 2016; Zeidman, Jafarian, Seghier, et al., 2019).

To evaluate the interaction of the nodes forming the core DMN, we used Bayesian model comparison to explore the space of possible hypotheses (or models), where each hypothesis assumed that a different combination of connectivity parameters could characterize all subjects in the group. A candidate model was obtained by removing one or more connections to produce nested or reduced forms of a full model. The obtained reduced model y_i is retained if its log evidence satisfies the following condition:

$$F_i = \ln(p|y_i) - \ln(p|y_0) > -20, \quad (2)$$

where y_0 is the full model. This procedure was performed using Bayesian model reduction (BMR), which enables analytical calculation of the log evidence of reduced models from a full model (Friston et al., 2016; Zeidman, Kazan, et al., 2019).

Then, the posterior probability P_i for each of the best 256 models was calculated as the softmax function of the log model evidence:

$$P_i = \exp(F_i) / \sum_j \exp(F_j). \quad (3)$$

All individual DCM models at a given sample size were re-fitted using empirical priors calculated as group means from the original individual DCMs (Friston et al., 2015; Friston et al., 2016; Litvak et al., 2015). First, the use of empirical priors informed by all subjects from the analyzed group enabled us to obtain a joint model space formed by the nested models derived from the fully connected model by pruning one or more connections (Friston et al., 2016; Zeidman, Kazan, et al., 2019). Second, the obtained joint model space was reduced to a smaller set of 256 plausible models by means of BMR, based on scoring of a model log-evidence (Friston et al., 2015; Friston et al., 2016; Litvak et al., 2015; Zeidman, Jafarian, Seghier, et al., 2019; Zeidman, Kazan, et al., 2019). The above computations at the group level were performed 100 times for randomly sampled subjects at every considered sample size. Accordingly, all measures used in our study were computed as the mean of 100 independent computations for different sample sizes. In particular, we computed the mean probabilities for each of the reduced models and explored the variability of the log-evidence estimates by calculating the confidence intervals for the Bayes factor of the most probable model. The mean coupling strengths and the corresponding confidence intervals were computed for each connection. We also report effect sizes and mean probability of parameters observation during the BMR. A subsequent focus on weaker connections allowed for the determination of the minimal sample size required for stable model selection and robust estimation of connectivity parameters.

2.4 | Tissue-specific signals regression

In principle, data without GSR provides more information to estimate EC compared to data after GSR, which encourages the use of data without GSR in DCM studies. Moreover, the effect of the averaged GSR on DCM was found to be minor and manifested itself mainly in the case of small networks (Almgren et al., 2020).

The confound signals originating from non-neuronal compartments, such as WM and CSF also affect voxel signals and may be computed using different approaches and used in different combinations. Tissue-specific WM and CSF signals can be computed using the same or subject-specific whole-brain segmentation mask for every subject. The amount of noise also depends on the probability threshold used during the creation of each specific mask. To address these two issues, we considered three different combinations of tissue-specific signals that were computed as the first principal components of brain signals using different whole-brain masks during BOLD signal extraction.

The first combination WM-CSF-95%INDV comprised WM and CSF signals, computed separately for every subject by using an individual (subject-specific) whole-brain mask with fixed high probability threshold equal to 0.95; the second one WM-CSF-50%TPM comprised the global WM and CSF signals obtained using the standard

SPM tissue probability map (TPM) for the whole-brain segmentation with fixed probability threshold equal to 0.5; the third one GM-WM-CSF-70%TPM comprised WM, CSF and GM tissue-specific signals computed using the standard TPM whole-brain mask at the fixed threshold equal to 0.7.

3 | RESULTS

3.1 | Stable connectivity pattern within the core DMN

The stability of the observed connectivity patterns is understood here as complete reproducibility at the group level for sample sizes exceeding a threshold value. The stable connectivity pattern within the core DMN revealed in our study is presented schematically in Figure 1. The mean values of posterior expectations for all connectivity parameters computed for the entire population of subjects and the three different TSSR combinations are shown in Table 1. In all three cases, the revealed stable EC pattern comprised nine connections. The connections included in the stable pattern were classified as weak if their mean coupling strengths were lower than the heuristic level of 0.1 or strong, otherwise.

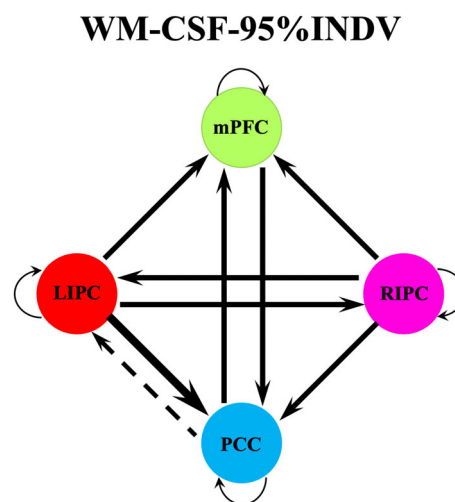


FIGURE 1 The nodes of the core default mode network (DMN) and connections between them are schematically indicated by the circles filled up with the different colors: medial prefrontal cortex (mPFC) (green), posterior cingulate cortex (PCC) (blue), left inferior parietal cortex (LIPC) (red), right inferior parietal cortex (RIPC) (pink). A stable pattern of effective connectivity was formed by connections that were observed with a high mean probability during the Bayesian model reduction (BMR). The stable pattern presented in the plot was computed for the WM-CSF-95%INDV combination of tissue-specific signals regression (TSSR) using the entire population of subjects. Stronger connections between DMN nodes are shown by the thicker arrows. Weak connection is shown by the dashed line arrow. The mean values of the coupling strengths for all connections within the core DMN are listed in Table 1.

TABLE 1 Mean posterior expectations (Hz).

Nodes	From mPFC	From PCC	From LIPC	From RIPC
To mPFC	(−0.46, −0.47, −0.49)	(0.17, 0.18, 0.19)	(0.13, 0.09, 0.06)	(0.16, 0.12, 0.11)
To PCC	(0.18, 0.18, 0.17)	(−0.08, −0.05, −0.02)	(0.41, 0.37, 0.34)	(0.16, 0.15, 0.15)
To LIPC	(−0.02, −0.02, −0.01)	(0.06, 0.06, 0.07)	(−0.44, −0.47, −0.5)	(0.16, 0.14, 0.15)
To RIPC	(0.03, 0.03, 0.02)	(0.00, 0.01, 0.02)	(0.17, 0.14, 0.11)	(−0.46, −0.48, −0.5)

Note: The mean values of the posterior expectations for connections from the three stable connectivity patterns are shown in Figure 1. All mean connectivity strengths were computed for the entire population of participants using three different combinations of TSSR (WM-CSF-95%INDV, WM-CSF-50%TPM and GM-WM-CSF-70%TPM). All connections comprising the model with the highest log evidence and those detected with high probability during the BMR search are shown in bold. The self-connections were parameterized on a log-scale (relative to the prior mean of −0.5). Abbreviations: LIPC, left inferior parietal cortex; mPFC, medial prefrontal cortex; PCC, precuneus/posterior cingulate cortex; RIPC, right inferior parietal cortex.

3.2 | Minimal sample size for stable model detection and impact of different TSSR combinations

To reveal a stable connectivity pattern within the core DMN, we compared multiple reduced models that encoded different hypotheses to determine the best model by using the Bayesian model comparison implemented in the PEB framework. An efficient search for the 256 most plausible models was performed over the joint space of all the reduced PEB models by scoring their free energy. The resulting log evidence for each of the 256 most plausible reduced models was computed as the difference between the free energy and the free energy of the fully connected model. This difference, called the Bayes factor, was used to quantify the supremacy of a specific reduced model with respect to the fully connected model. The reduced models with positive log evidence had the largest weights during Bayesian model averaging (BMA).

It should be noted here that most of the 256 plausible reduced models demonstrated similar connectivity patterns and can be characterized by the mean posterior probabilities (3), which were computed for each plausible model at different sample sizes and different TSSR combinations (see Figure 2a). The reduced model that was observed with the largest probability can be considered the best model (model 256 in Figure 2a). The best model represents the optimal balance between accuracy and complexity and is characterized by the largest positive Bayes factor. The presence of non-zero probabilities for other reduced models is caused by the small quantitative differences in connectivity parameters with respect to the best model, which led to lower Bayes factor estimates for these models. Sets of 256 plausible reduced models were detected for different sample sizes and different TSSR combinations. All probability estimates were bootstrapped 100 times and averaged for any given sample size and each TSSR combination.

The best model was detected with approximately the same probability for each considered TSSR combinations (Figure 2a). Moreover, the mean probability computed for the best model gradually increased for $n \geq 50$ (Figure 2b). However, the obtained dependencies of the mean probabilities on the sample size computed for the different TSSR combinations were slightly different (Figure 2b). In particular,

the best model was detected with a relatively higher probability and for smaller sample sizes in the case of WM-CSF-95%INDV (Figure 2b). Since the free energy is the main measure used for model selection in DCM, the analysis of its variability at different sample sizes may allow the estimation of a minimal sample size required for stable DCM analysis. Let us consider the impact of sample size on the variability of the free energy and the best model selection.

3.3 | Impact of sample size and TSSR on variability of log-evidence estimates

The main measure used in the DCM for the reduced model selection is the log-evidence (2) or Bayes factor, computed for a specific model comparison to the full model (Friston et al., 2003). The bootstrapped mean log-evidence computed for each of the reduced plausible models may be useful for quantifying the quality and stability of the model-selection procedure. In particular, the sensitivity of reduced model selection to inter-subject variability can be quantified by the mean variance of the log-evidence computed for the best model for different sample sizes and TSSR combinations (Figure 2c). The observed mean variance was relatively large at the small sample sizes while decreasing as sample size grows. The obtained dependencies of the mean variance on the sample size demonstrated fluctuating behavior and started to rapidly decrease at $n \geq 50$ for the combinations WM-CSF-95%INDV and WM-CSF-50%TPM. By contrast, in the case of combination GM-WM-CSF-70%TPM, the gradual decline started at the larger sample sizes $n \geq 100$ (Figure 2c).

To further clarify the impact of the sample size and TSSR on model selection, we computed the mean values of the log evidence for the best model, together with the bounds of corresponding confidence intervals at different sample sizes and for the different TSSR combinations. The obtained means fluctuate at approximately 3, which corresponds to the “strong” evidence for the best model (Zeidman, Jafarian, Seghier, et al., 2019). The sample size at which the lower bound of the confidence interval becomes larger than zero can be considered an approximate minimal threshold value required for stable estimates of the free energy and stable model selection. The

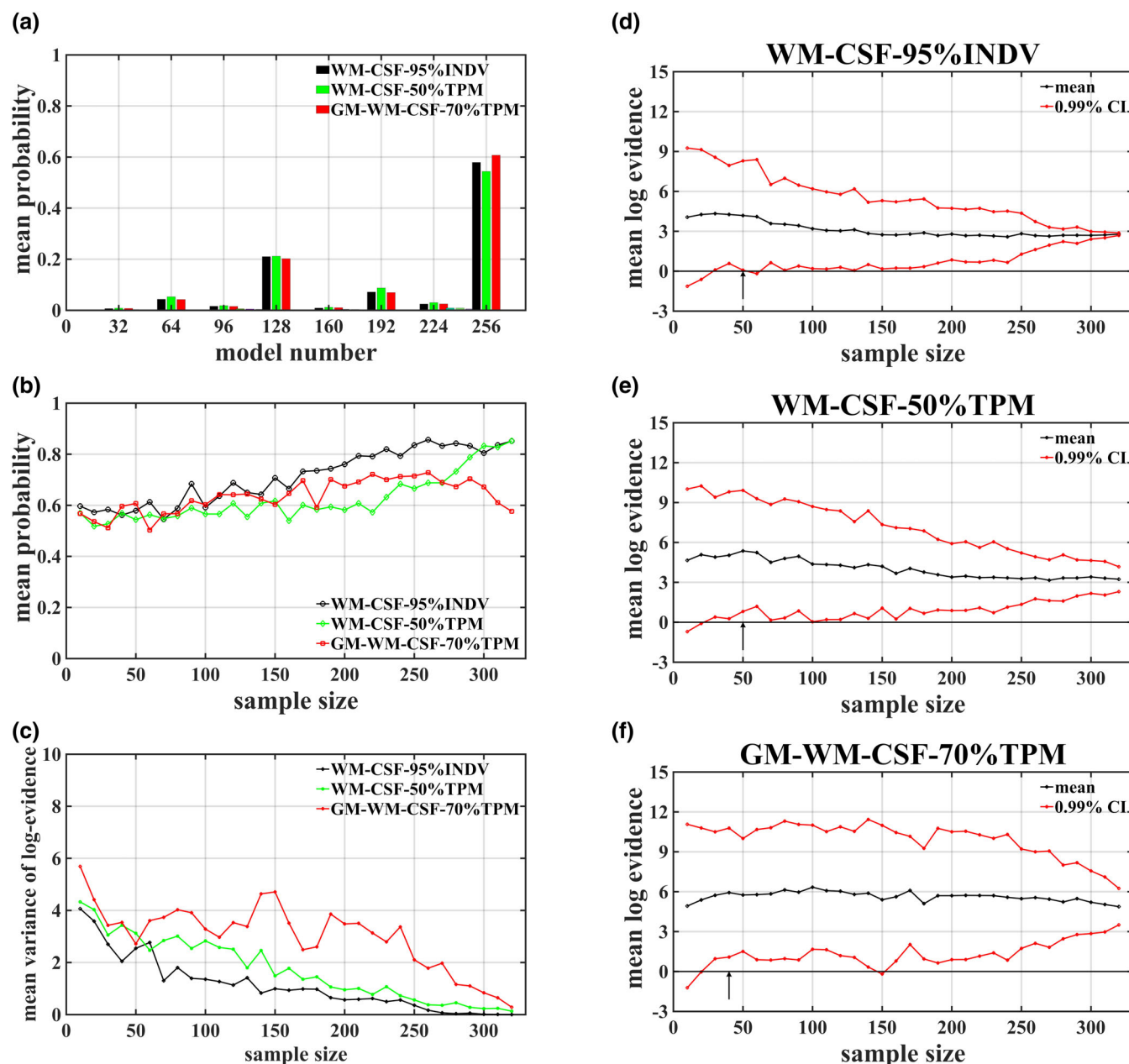


FIGURE 2 (a) Mean posterior probabilities for the set of 256 selected plausible reduced models. The mean probabilities were computed for subjects during the Bayesian model reduction (BMR) performed for different combinations of tissue-specific signals regression (TSSR). (b) Mean posterior probabilities for the most probable model (model 256 in a) computed for different TSSR combinations versus sample size. (c) Mean variance of log-evidence for the most probable model (model 256 in a) computed for different TSSR combinations versus sample size. (d–f) Mean values (black curves) of the log-evidence defined by Equation (2) for the most probable model and bounds of its confidence interval (red curves) computed at different sample sizes and for different TSSR combinations. The approximate values of the minimum sample size, required to make stable the mean log-evidence estimates, (lower bounds of the confidence intervals do not intersect with zero) are indicated by arrows. All values presented in the plots were obtained as a mean across 100 bootstrap computations at every sample size.

revealed dependencies of the mean log-evidence and bounds of the corresponding confidence intervals on the sample size were similar for all three TSSR combinations (Figure 2d–f). The log-evidence estimates for the best model became stable at the sample sizes $n \geq 50$ as estimated for the significance level 99% confidence interval (CI) (Figure 2d–f). Thus, the approximate minimal sample size required for stable model selection may be estimated as $n \approx 50$.

3.4 | Stability of connectivity parameters estimates at different TSSR

To study the impact of sample size and inter-individual variability on the stability of connectivity parameters estimates, we computed the mean values of their posterior expectations and variances for different sample sizes. This was accomplished by inverting all individual DCMs

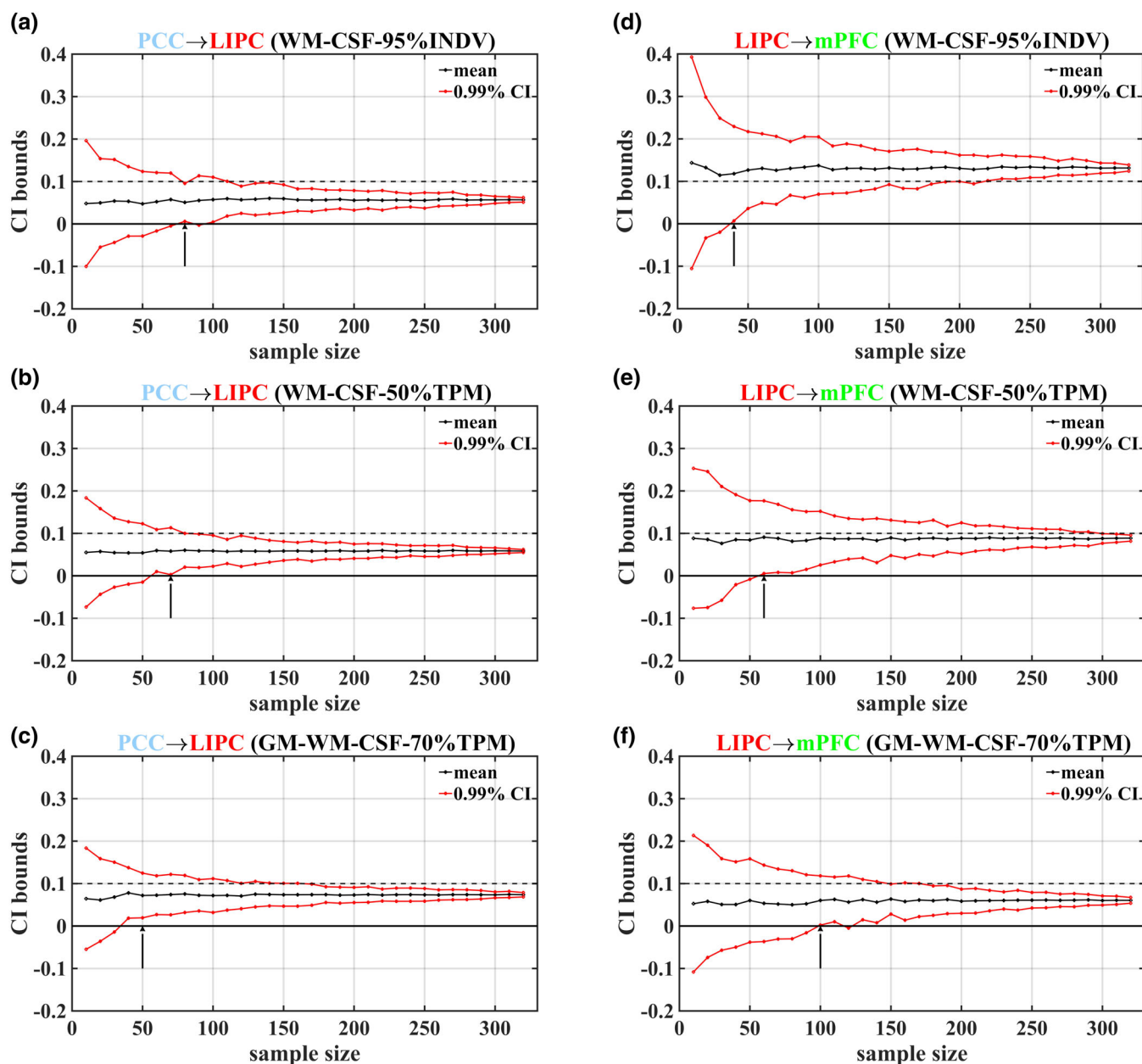


FIGURE 3 The mean posterior expectations (black solid lines) and bounds of the confidence intervals (red solid lines) were computed for the two weak connections at different sample sizes. The dependencies presented in plots (a) and (d), (b) and (e), and (c) and (f) were computed for the different combinations of tissue-specific signals regression (TSSR), as indicated in the titles of the plots. The approximate values of a minimum sample size required to obtain stable estimates of the coupling strength are indicated by arrows. The mean values and bounds of the confidence intervals presented in the plots were obtained for the distribution of results of 100 bootstrap computations at every sample size.

(four DCMs per subject) re-fitted using empirical priors. The results of the PEB group analysis and BMA, performed at the sample size varied between 10 and 320 with step 10 for the different TSSR combinations, are presented in Figures 3 and S5–S8. All strong connections attain the strong enough evidence at the sample size values $n \approx 30–40$ for each of the TSSR combinations used in our study (Figures 3 and S5–S8).

In contrast, a larger sample size $n \geq 50$ was needed to detect weak connections $PCC \rightarrow LIPC$ and $LIPC \rightarrow mPFC$, which demonstrated enhanced sensitivity to inter-subject variability (Figure 3). Mean values of connectivity strength for the $PCC \rightarrow LIPC$ connection

attained their strong enough evidence and saturated at the sample sizes from the interval $n \approx 50–80$ for each of the three TSSR combinations (Figure 3a–c). Another weak connection $LIPC \rightarrow mPFC$ required different threshold sample size values for each of the three TSSR combinations (Figure 3d–f).

The group estimates of all connectivity parameters computed at the threshold sample size value $n = 50$ are summarized in Figure S1. As seen, the strength of $PCC \rightarrow LIPC$ connection just to became detectable at $n \geq 50$ only for the combination GM-WM-CSF-70% TPM, whereas a larger sample size is required for the other two TSSR

combinations. The LIPC $\rightarrow$ mPFC connection appeared to be strong in the case of WM-CSF-95%INDV, whereas for the other two TSSR combinations, it was weak and became detectable only when $n > 70$ (Figures 3 and S1). This finding is in line with the above results and allows to consider the sample size $n \approx 50$ as the approximate threshold sample size value required for the stable DCM analysis. Of note, the impact of TSSR on strong connections was not substantial to even make them weak (Figures S1–S4).

Perhaps, the individual (subject-specific) whole-brain masking used for the combination WM-CSF-95%INDV allowed to slightly decrease noise at the subject level what caused strengthening of connection LIPC $\rightarrow$ mPFC and made it observable at smaller sample sizes (Figures 3d and S2). The impact of two other TSSR combinations was characterized by increased estimates of the minimal sample size to $n \geq 60$ for the combination WM-CSF-50%TPM and to $n \geq 100$ for GM-WM-CSF-70%TPM (Figures 3e, f and S3, S4).

3.5 | Effect size and parameters observation probabilities for the different combinations of TSSR

To quantify the deviation from the null hypothesis for strong and weak connections, we calculated the effect size for these connections for different sample sizes and TSSR combinations. The classical Cohen's d effect size was calculated at every given sample size using the averaged mean values and variances obtained as the result of inversion of individual DCM models re-fitted using empirical priors.

As follows from the results presented in Figure 4a–c, comparatively low mean values of the effect size were observed only for the 2 weak connections at the sample size $n \leq 50$. In the case of strong connections, the mean effect size was high for every connection, even for relatively small sample sizes. Moreover, the obtained mean effect size values were larger for stronger connections at any given sample size (Figure 4a–c). The same relationship between the mean effect size values obtained for the strong and weak connections was observed for each TSSR combination (Figure 4a–c).

Another basic measure used to detect statistically significant connections within the core DMN was the mean probability of observing a specific connection computed during the BMR search when all possible models were scored in accordance with their log evidence. The computed observation probabilities were bootstrapped across 100 independent computations, for each sample size. The mean observation probabilities demonstrated different dependencies on sample size for strong and weak connections (Figure 4d–f). All strong connections were observed more frequently during the BMR search and appeared with a probability close to one for sample sizes $n \geq 30$. This estimate is in line with the above estimates for the minimal threshold sample size for strong connections. In contrast, the minimal sample size sufficient to observe weak connections with high probability ($P \geq 0.99$) ranged from $n \approx 50$ in the case of PCC $\rightarrow$ LIPC till $n \approx 50$ –150 in the case of LIPC $\rightarrow$ mPFC (Figure 4d–f).

Different combinations of TSSR had different impact on the mean probabilities of observing weak connections. In particular, PCC $\rightarrow$

LIPC can be detected with smaller sample sizes for combinations WM-CSF-70%TPM and GM-WM-CSF-70%TPM ($n = 100$ and $n = 50$, respectively), while in the case of WM-CSF-95%INDV the required sample size was $n \approx 120$ (Figure 4d). The opposite was observed for connection LIPC $\rightarrow$ mPFC. Mean observation probability approaches to 0.99% threshold at $n \geq 40$ for the combination WM-CSF-95%INDV and requires larger sample sizes $n = 100$ and 150 for the two other combinations (Figure 4e, f). These results agree well with above estimates of the minimum sample size based on computations of the confidence intervals for the mean posterior expectations (Figure 3a–f).

4 | DISCUSSION

DCM analysis of connectivity in neuronal networks examines inter-regional communication in the brain far more realistic than commonly used correlation approaches to connectivity, as DCM models directionality of regional interactions. FC is known to be underpowered in the case of small sample sizes and to be highly sensitive to different types of noise (Marek et al., 2022), specifically within scanner movements. To shed light on these two aspects with regard to DCM, we explored to what extent the stability of EC within the core DMN, estimated by spectral DCM, is affected by inter-subject variability using systematic sample size variation and bootstrap analysis, respectively. Additionally, we investigated the impact of extracting different tissue-specific signal fluctuations, particularly with respect to partial volume effects and even global GM signals, on the stability of these connectivity estimates for comparatively noise-free BOLD time series. The later aims to clarify the influence of extensive/non-specific cleaning of resting-state data on DCM analysis by regressing out potentially meaningful BOLD fluctuations. The stable connectivity pattern within the core DMN revealed for the large sample sizes comprised nine connections (Figure 1). The mean posterior expectation for each of these connections had a high probability of being observed at the group level. The connections from bilateral IPC to the mPFC and PCC were rather strong and nonsymmetrical (Table 1). The revealed interhemispheric coupling strength asymmetry had an emphasis on the left side (Figure 1 and Table 1). The bidirectional connections between the mPFC and PCC had approximately the same strength and were slightly stronger than those originating from the RIPC and interhemispheric connections within the IPC (Table 1).

We showed that both stability of the model selection procedure and inference of connectivity parameters are dependent on the sample size and inter-individual variability. In particular, the impact of sample size on group DCM analyses manifests itself in the perturbed stability of the log-evidence estimates (Figure 2). Bootstraps over sample sizes revealed high variability of log-evidence for small sample sizes showing high sensitivity of DCM to the inter-individual variability in resting-state time series.

It should be noted that the resting-state HCP ICA-FIX datasets used in our study were cleaned from noise and physiological artifacts and were rather large compared with the groups of subjects used in other studies. Because of these two factors, the minimal threshold

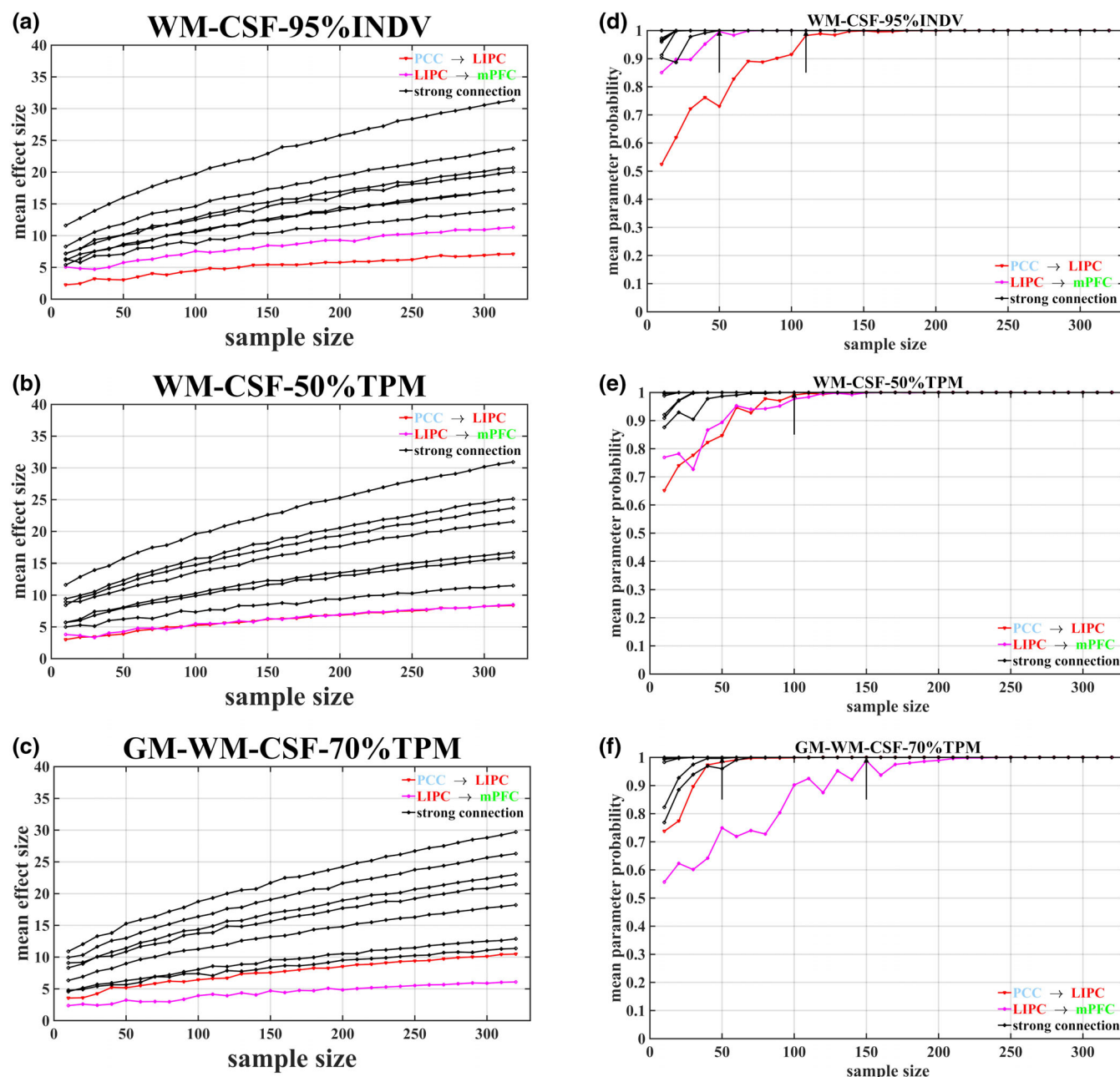


FIGURE 4 (a–c) The mean values of Cohen's effect size were computed for statistically significant strong and weak connections within the core default mode network (DMN) at different sample sizes and different combinations of the tissue-specific signals regression (TSSR). The dependencies obtained for the strong and weak connections are indicated by the black and colored (red and pink) solid lines, respectively. (d–f) Mean probabilities of parameters observation computed as the result of Bayesian model reduction (BMR) search for the best model in the model space for different sample sizes. The presented mean values were averaged across 100 bootstrap computations at every considered sample size.

value revealed in our study may be considered as a lower estimate of the sample size required for stable DCM analysis. Indeed, a large inter-subject variability, together with a high noise level and contamination of rs-fMRI datasets by the presence of different movement artifacts, may require sample sizes that are larger than the $n \approx 50$ reported in our study.

The first attempts to reveal a pattern of EC within the core DMN led to a discrepancy in the results published by several studies (Di & Biswal, 2014; Friston et al., 2014; Jiao et al., 2011; Li et al., 2012; Razi

et al., 2015; Sharaev et al., 2016). For instance, Li et al. (2012) showed a directed influence from the PCC to the mPFC using stochastic DCM, while other authors (Di & Biswal, 2014; Jiao et al., 2011) reported a causal influence from the mPFC to the PCC, but not vice versa. We found that both connections were present in a stable pattern (see Figure 1 and Table 1). The group-averaged connectivity pattern for the best model revealed by Di and Biswal (2014) is also quite different from the stable pattern revealed in our study. The authors used stochastic DCM approach for the resting-state fMRI

data to study the patterns of EC within the core DMN. Despite the fact that the sample size used in that study was equal to 64, most of the connections in a best model were weak and statistically nonsignificant at the group level with emphasis on the right side (Di & Biswal, 2014). In this particular case, the presence of such a huge discrepancy between best models may be explained by the lower accuracy of stochastic DCM in the case of the resting-state fMRI datasets compared to the spectral DCM (Friston et al., 2014).

The invention of spectral DCM allowed a decrease in variability in connectivity parameter estimates for resting-state fMRI data but was insufficient to reveal all connections within the core DMN at small sample sizes (Razi et al., 2015). Most of the group-averaged connectivity parameters obtained for the group of 24 subjects were weak, with emphasis on the left side, and interactions between the bilateral IPC and mPFC were not symmetric (Razi et al., 2015). A slightly different best model and connectivity pattern were revealed for a group of 30 subjects (Sharaev et al., 2016). The best model, reported by Sharaev et al. (2016), was characterized by the presence of symmetric interactions between the IPC, PCC, and mPFC with slight interhemispheric asymmetry with emphasis on the right side, and was similar to the model reported by Razi et al. (2015). Most of the connections reported by Sharaev et al. (2016) were stronger than those reported by Razi et al. (2015), which might be caused by the slightly increased sample size. The best model and pattern of EC reported by Sharaev et al. (2016) is also the closest but not identical to the best model and stable connectivity pattern revealed in our study. In particular, the pattern observed in our study was characterized by the interhemispheric asymmetry with emphasis on the left side. Moreover, there was a missing weak connection $PCC \rightarrow LIPC$ that could not be detected at the sample size $n = 30$. Thus, the sample size is one of the key factors defining the accuracy and stability of EC estimates.

The stable connectivity pattern revealed in our study verified the central role of the PCC, which is known to be a hub in the DMN, through which all other nodes interact (Buckner et al., 2008). The pivotal status of the PCC is also justified from a metabolic and mechanistic perspective in that previous positron emission tomography studies have shown that metabolic activity is higher in the PCC than in all other regions during rest (Gusnard & Raichle, 2001). The strong interconnectivity between the PCC node and the rest of the DMN, as revealed by the partial correlation network analysis, further supports the hypothesis that the PCC node in the DMN acts as a convergence node, where information processing in the two subsystems is integrated (Fransson & Marrelec, 2008). Moreover, PCC is not only driven by all DMN nodes but may also project weak backward connections to other nodes and networks, which can be detected only at relatively large sample sizes but nevertheless play an important role in communication between subnetworks (Di & Biswal, 2014; Frässle et al., 2021; Razi et al., 2017).

To examine the possible impact of regression of the tissue-specific signals on the estimates of EC, we used different combinations of tissue-specific signals. These signals were computed as the first principal components of the brain signals obtained by using different whole-brain masks. Two considered combinations, WM-

CSF-95%INDV and WM-CSF-50%TPM, utilize the whole-brain mask including all voxels belonging to WM and CSF only with probability thresholds 0.95 and 0.5, respectively. By using these combinations, we can estimate the extent to which the precision of confound signals detection is able to affect the estimates of EC. Furthermore, the results obtained for the first TSSR combination (WM-CSF-95%INDV) based on the individual (subject-specific) whole-brain mask can be compared with the results for the second combination (WM-CSF-50%TPM) based on the standard whole-brain mask from SPM. In addition, another GSR combination including the WM, CSF, and GM signals (GM-WM-CSF-70%TPM), was computed using the standard whole-brain mask from SPM. Including a GM signal to the commonly used combination of WM and CSF confound signals is similar to the global signal removal and allowed to decrease the amount of physiological noise. Regressing any of the three combinations of the tissue-specific signals had the same minor impact on strong connections. The mean posterior expectations for the strong connections attained higher values for the combination WM-CSF-95%INDV, involving more precise detection of WM and CSF voxels. The additional regression of the global GM signal, as reflected by the third combination, might lead to a decrease in the mean posterior expectations for all connections, which may be explained by a common weakening of neuronal signals originating from the GM. The connection $LIPC \rightarrow mPFC$ appears to be more sensitive to the GM regression and can be stably detected only for a rather large sample size ($n \geq 100$).

The connectivity parameters in DCM are measured in Hz and mean the rates of interaction between the chosen brain areas (Zeidman, Jafarian, Corbin, et al., 2019). Furthermore, there is an inverse relationship between the rate constants and time constants that characterize the decay of neuronal responses. When the values of the connectivity parameters are small or close to zero, the neuronal responses of the chosen brain regions would decay back for a long time, that makes it impossible to detect the causality of such interactions (Zeidman, Jafarian, Corbin, et al., 2019). The heuristic threshold value for the rate constants separating strong and weak connections was defined as 0.1 Hz (Razi et al., 2015). However, it remains unclear whether all connections with strengths less than 0.1 Hz must be removed from the consideration or some of them, having a non-zero probability to be observed during the BRM, can be included in the observed connectivity pattern. Moreover, the presence of inter-subject variability makes it difficult or even impossible to detect weak connections in the case of small sample sizes. Thus, the presence of weak connections and their ability to become detectable at large sample sizes play an important role in revealing the structure of interacting brain circuits. The observed weak connections demonstrated different sensitivities to the sample size for the different combinations of TSSR, as indicated in Figure 3. The first weak connection, $PCC \rightarrow LIPC$, was slightly more stable to the variation in the sample size in the case of the combination (GM-WM-CSF-70%TPM) (Figure 3a–c). The second weak connection, the $LIPC \rightarrow mPFC$, had different sensitivities to the combinations of TSSR. In particular, a noise reduction at the individual subject level obtained for the combination (WM-CSF-95%INDV) allowed the detection of that connection, even at rather

small sample sizes, as was observed for the strong connections (Figures 3d and S2). Thus, the individual (subject-specific) masking of WM and CSF during the preprocessing stage may allow to enhance the detection of weak connections. Two others TSSR combinations had opposite effects on the LIPC $\rightarrow$ mPFC and resulted in increased values of the minimum sample size required for the stable estimation of connectivity parameters (Figure 3e, f).

The mean effect size computed for the strong and weak connections supports the above findings and indicates that strong connections attain larger values even at smaller sample sizes, whereas weak connections can be stably detected at a larger sample size (Figure 4a-c). Moreover, the results for mean probabilities of parameters observation, which were computed during the course of BRM best model search, agree well with all findings reported above. The lowest probabilities to observe were also found for the weakest connections (Figure 4d-f). That supports our conclusion about weak connections as an important factor defining a minimum sample size required for the stable DCM analysis.

The generalizability of the results presented above is limited owing to the use of only one dataset and one brain network in our analysis. The reported estimates of the minimal sample size for the stable DCM analysis must be considered as the lowest estimates, which are likely to be increased in the case of datasets comprising shorter time series and larger noise levels.

5 | CONCLUSION

We analyzed the stability of the spectral DCM to sample size variation and the level of physiological noise using a large HCP dataset of 330 unrelated subjects. The estimates of EC within the core DMN demonstrated sensitivity to both sample size and noise level. In particular, the increased sample sizes allowed for the detection of weak connections, which were not detectable for small sample sizes. This made it possible to reveal a stable and reproducible pattern of EC within the core DMN. The sample size had a moderate impact on the strong connections within the core DMN at the group level. The subject-specific whole-brain masks used during the WM and CSF signals computations reduced the amount of noise and also enhanced the ability to detect weak connections. Thus, the presence of weak connections in brain networks and their sensitivity to inter-subject variability and physiological noise limits the sample size needed for stable and reproducible DCM analysis. As observed in our study, the minimum sample size required for stable estimates of EC must not be less than 50, even in the case of clean HCP datasets.

ACKNOWLEDGMENTS

This work was supported by the Forschungszentrum Jülich GmbH (Alexander Silchenko), Simon B. Eickhoff acknowledges funding by the European Union's Horizon 2020 Research and Innovation Program (grant agreements 945539 [HBP SGA3] and 826421 [VBC]), the Deutsche Forschungsgemeinschaft (DFG, SFB 1451 and IRTG 2150)

and the National Institute of Health (R01 MH074457). Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST STATEMENT

The authors declare no potential conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in DMN_DCM_HCP at https://gin.g-node.org/felixh/DMN_DCM_HCP.git.

ORCID

Alexander N. Silchenko  <https://orcid.org/0000-0001-5591-238X>

Felix Hoffstaedter  <https://orcid.org/0000-0001-7163-3110>

Simon B. Eickhoff  <https://orcid.org/0000-0001-6363-2759>

REFERENCES

- Almgren, H., Van de Steen, F., Kühn, S., Razi, A., Friston, K., & Marinazzo, D. (2018). Variability and reliability of effective connectivity within the core default mode network: A multi-site longitudinal spectral DCM study. *NeuroImage*, 183, 757–768. <https://doi.org/10.1016/j.neuroimage.2018.08.053>
- Almgren, H., Van de Steen, F., Razi, A., Friston, K., & Marinazzo, D. (2020). The effect of global signal regression on DCM estimates of noise and effective connectivity from resting state fMRI. *NeuroImage*, 208, 116435. <https://doi.org/10.1016/j.neuroimage.2019.116435>
- Bartoň, M., Mareček, R., Krajčovičová, L., Slaviček, T., Kašpárek, T., Zemánková, P., Říha, P., & Mikl, M. (2019). Evaluation of different cerebrospinal fluid and white matter fMRI filtering strategies—quantifying noise removal and neural signal preservation. *Human Brain Mapping*, 40(4), 1114–1138. <https://doi.org/10.1002/HBM.24433>
- Behzadi, Y., Restom, K., Liau, J., & Liu, T. T. (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage*, 37(1), 90–101. <https://doi.org/10.1016/j.neuroimage.2007.04.042>
- Biswal, B., Zerrin Yetkin, F., Haughton, V. M., & Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magnetic Resonance in Medicine*, 34(4), 537–541. <https://doi.org/10.1002/MRM.1910340409>
- Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network: Anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences*, 1124, 1–38. <https://doi.org/10.1196/ANNALS.1440.011>
- Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, 14(5), 365–376. <https://doi.org/10.1038/nrn3475>
- Di, X., & Biswal, B. B. (2014). Identifying the default mode network structure using dynamic causal modeling on resting-state functional magnetic resonance imaging. *NeuroImage*, 86, 53–59. <https://doi.org/10.1016/j.neuroimage.2013.07.071>
- Fox, M. D., Zhang, D., Snyder, A. Z., & Raichle, M. E. (2009). The global signal and observed anticorrelated resting state brain networks. *Journal of Neurophysiology*, 101(6), 3270–3283. <https://doi.org/10.1152/JN.90777.2008>
- Fransson, P., & Marrelec, G. (2008). The precuneus/posterior cingulate cortex plays a pivotal role in the default mode network: Evidence from a partial correlation network analysis. *NeuroImage*, 42(3), 1178–1184. <https://doi.org/10.1016/j.neuroimage.2008.05.059>
- Frässle, S., Harrison, S. J., Heinze, J., Clementz, B. A., Tamminga, C. A., Sweeney, J. A., Gershon, E. S., Keshavan, M. S., Pearlson, G. D.,

- Powers, A., & Stephan, K. E. (2021). Regression dynamic causal modeling for resting-state fMRI. *Human Brain Mapping*, 42(7), 2159–2180. <https://doi.org/10.1002/HBM.25357>
- Friston, K. J., Zeidman, P., & Litvak, V. (2015). Empirical bayes for DCM: A group inversion scheme. *Frontiers in Systems Neuroscience*, 9, 164. <https://doi.org/10.3389/FNSYS.2015.00164/BIBTEX>
- Friston, K. J. (2011). Functional and effective connectivity: A review. *Brain Connectivity*, 1(1), 13–36. <https://doi.org/10.1089/BRAIN.2011.0008>
- Friston, K. J., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. *NeuroImage*, 19(4), 1273–1302. [https://doi.org/10.1016/S1053-8119\(03\)00202-7](https://doi.org/10.1016/S1053-8119(03)00202-7)
- Friston, K. J., Kahan, J., Biswal, B., & Razi, A. (2014). A DCM for resting state fMRI. *NeuroImage*, 94(100), 396–407. <https://doi.org/10.1016/J.NEUROIMAGE.2013.12.009>
- Friston, K. J., Litvak, V., Oswal, A., Razi, A., Stephan, K. E., Van Wijk, B. C. M., Ziegler, G., & Zeidman, P. (2016). Bayesian model reduction and empirical Bayes for group (DCM) studies. *NeuroImage*, 128, 413–431. <https://doi.org/10.1016/J.NEUROIMAGE.2015.11.015>
- Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J. R., Van Essen, D. C., & Jenkinson, M. (2013). The minimal preprocessing pipelines for the human connectome project. *NeuroImage*, 80, 105–124. <https://doi.org/10.1016/J.NEUROIMAGE.2013.04.127>
- Gusnard, D. A., & Raichle, M. E. (2001). Searching for a baseline: Functional imaging and the resting human brain. *Nature Reviews. Neuroscience*, 2(10), 685–694. <https://doi.org/10.1038/35094500>
- Hillebrandt, H., Friston, K. J., & Blakemore, S. J. (2014). Effective connectivity during animacy perception – Dynamic causal modelling of human connectome project data. *Scientific Reports*, 4(1), 1–9. <https://doi.org/10.1038/srep06240>
- Jiao, Q., Lu, G., Zhang, Z., et al., (2011). Granger causal influence predicts BOLD activity levels in the default mode network. *Human Brain Mapping*, 32(1), 154–161. <https://doi.org/10.1002/HBM.21065>
- Li, B., Wang, X., Yao, S., Hu, D., & Friston, K. (2012). Task-dependent modulation of effective connectivity within the default mode network. *Frontiers in Psychology*, 3, 206. <https://doi.org/10.3389/FPSYG.2012.00206/BIBTEX>
- Lindquist, M. A., Caffo, B., & Crainiceanu, C. (2013). Ironing out the statistical wrinkles in “ten ironic rules”. *NeuroImage*, 81, 499–502. <https://doi.org/10.1016/J.NEUROIMAGE.2013.02.056>
- Litvak, V., Garrido, M., Zeidman, P., & Friston, K. (2015). Empirical bayes for group (Dcm) studies: A reproducibility study. *Frontiers in Human Neuroscience*, 9, 670. <https://doi.org/10.3389/FNHUM.2015.00670/BIBTEX>
- Liu, T. T. (2016). Noise contributions to the fMRI signal: An overview. *NeuroImage*, 143, 141–151. <https://doi.org/10.1016/J.NEUROIMAGE.2016.09.008>
- Liu, T. T., Nalci, A., & Falahpour, M. (2017). The global signal in fMRI: Nuisance or information? *NeuroImage*, 150, 213–229. <https://doi.org/10.1016/j.neuroimage.2017.02.036>
- Mak, L. E., Minuzzi, L., MacQueen, G., Hall, G., Kennedy, S. H., & Milev, R. (2017). The default mode network in healthy individuals: A systematic review and meta-analysis. *Brain Connectivity*, 7(1), 25–33. <https://doi.org/10.1089/BRAIN.2016.0438>
- Marek, S., Tervo-Clemmens, B., Calabro, F. J., Montez, D. F., Kay, B. P., Hatoum, A. S., Donohue, M. R., Foran, W., Miller, R. L., Hendrickson, T. J., Malone, S. M., Kandala, S., Feczko, E., Miranda-Dominguez, O., Graham, A. M., Earl, E. A., Perrone, A. J., Cordova, M., Doyle, O., ... Dosenbach, N. U. F. (2022). Reproducible brain-wide association studies require thousands of individuals. *Nature*, 603(7902), 654–660. <https://doi.org/10.1038/s41586-022-04492-9>
- Mason, M. F., Norton, M. I., Van Horn, J. D., Wegner, D. M., Grafton, S. T., & Macrae, C. N. (2007). Wandering minds: The default network and stimulus-independent thought. *Science*, 315(5810), 393. <https://doi.org/10.1126/SCIENCE.1131295>
- Murphy, K., Birn, R. M., & Bandettini, P. A. (2013). Resting-state fMRI confounds and cleanup. *NeuroImage*, 80, 349–359. <https://doi.org/10.1016/J.NEUROIMAGE.2013.04.001>
- Poldrack, R. A., Baker, C. I., Durnez, J., Gorgolewski, K. J., Matthews, P. M., Munafò, M. R., Nichols, T. E., Poline, J. B., Vul, E., & Yarkoni, T. (2017). Scanning the horizon: Towards transparent and reproducible neuroimaging research. *Nature Reviews Neuroscience*, 18(2), 115–126. <https://doi.org/10.1038/nrn.2016.167>
- Raichle, M. E. (2015). The brain's default mode network. *Annual Review of Neuroscience*, 38, 433–447. <https://doi.org/10.1146/ANNUREV-NEURO-071013-014030>
- Razi, A., Kahan, J., Rees, G., & Friston, K. J. (2015). Construct validation of a DCM for resting state fMRI. *NeuroImage*, 106, 1–14. <https://doi.org/10.1016/J.NEUROIMAGE.2014.11.027>
- Razi, A., Seghier, M. L., Zhou, Y., McColgan, P., Zeidman, P., Park, H. J., Sporns, O., Rees, G., & Friston, K. J. (2017). Large-scale DCMs for resting-state fMRI. *Network Neuroscience (Cambridge, Mass.)*, 1(3), 222–241. https://doi.org/10.1162/NETN_A_00015
- Salimi-Khorshidi, G., Douaud, G., Beckmann, C. F., Glasser, M. F., Griffanti, L., & Smith, S. M. (2014). Automatic denoising of functional MRI data: Combining independent component analysis and hierarchical fusion of classifiers. *NeuroImage*, 90, 449–468. <https://doi.org/10.1016/J.NEUROIMAGE.2013.11.046>
- Seguin, C., Razi, A., & Zalesky, A. (2019). Inferring neural signalling directionality from undirected structural connectomes. *Nature Communications*, 10(1), 1–13. <https://doi.org/10.1038/s41467-019-12201-w>
- Sharaev, M. G., Zavyalova, V. V., Ushakov, V. L., Kartashov, S. I., & Velichkovsky, B. M. (2016). Effective connectivity within the default mode network: Dynamic causal modeling of resting-state fMRI data. *Frontiers in Human Neuroscience*, 10, 14. <https://doi.org/10.3389/FNHUM.2016.00014>
- Smith, S. M., Beckmann, C. F., Andersson, J., Auerbach, E. J., Bijsterbosch, J., Douaud, G., Duff, E., Feinberg, D. A., Griffanti, L., Harms, M. P., Kelly, M., Laumann, T., Miller, K. L., Moeller, S., Petersen, S., Power, J., Salimi-Khorshidi, G., Snyder, A. Z., Vu, A. T., ... Glasser, M. F. (2013). Resting-state fMRI in the human connectome project. *NeuroImage*, 80, 144–168. <https://doi.org/10.1016/J.NEUROIMAGE.2013.05.039>
- Song, M., Zhou, Y., Li, J., Liu, Y., Tian, L., Yu, C., & Jiang, T. (2008). Brain spontaneous functional connectivity and intelligence. *NeuroImage*, 41(3), 1168–1176. <https://doi.org/10.1016/J.NEUROIMAGE.2008.02.036>
- Stephan, K. E., Penny, W. D., Moran, R. J., den Ouden, H. E. M., Daunizeau, J., & Friston, K. J. (2010). Ten simple rules for dynamic causal modeling. *NeuroImage*, 49(4), 3099–3109. <https://doi.org/10.1016/J.NEUROIMAGE.2009.11.015>
- Szucs, D., & Ioannidis, J. P. (2020). Sample size evolution in neuroimaging research: An evaluation of highly-cited studies (1990–2012) and of latest practices (2017–2018) in high-impact journals. *NeuroImage*, 221, 117164. <https://doi.org/10.1016/J.NEUROIMAGE.2020.117164>
- Van Den Heuvel, M. P., Stam, C. J., Kahn, R. S., & Hulshoff Pol, H. E. (2009). Efficiency of functional brain networks and intellectual performance. *Journal of Neuroscience*, 29(23), 7619–7624. <https://doi.org/10.1523/JNEUROSCI.1443-09.2009>
- Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E. J., Yacoub, E., & Ugurbil, K. (2013). The WU-Minn human connectome project: An overview. *NeuroImage*, 80, 62–79. <https://doi.org/10.1016/J.NEUROIMAGE.2013.05.041>
- Van Essen, D. C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T. E. J., Bucholz, R., Chang, A., Chen, L., Corbetta, M., Curtiss, S. W., Della Penna, S., Feinberg, D., Glasser, M. F., Harel, N., Heath, A. C., Larson-Prior, L., Marcus, D., Michalareas, G., Moeller, S., ... Yacoub, E. (2012). The

- human connectome project: A data acquisition perspective. *NeuroImage*, 62(4), 2222–2231. <https://doi.org/10.1016/J.NEUROIMAGE.2012.02.018>
- Zeidman, P., Jafarian, A., Corbin, N., Seghier, M. L., Razi, A., Price, C. J., & Friston, K. J. (2019). A guide to group effective connectivity analysis, part 1: First level analysis with DCM for fMRI. *NeuroImage*, 200, 174–190. <https://doi.org/10.1016/J.NEUROIMAGE.2019.06.031>
- Zeidman, P., Jafarian, A., Seghier, M. L., Litvak, V., Cagnan, H., Price, C. J., & Friston, K. J. (2019). A guide to group effective connectivity analysis, part 2: Second level analysis with PEB. *NeuroImage*, 200, 12–25. <https://doi.org/10.1016/J.NEUROIMAGE.2019.06.032>
- Zeidman, P., Kazan, S. M., Todd, N., Weiskopf, N., Friston, K. J., & Callaghan, M. F. (2019). Optimizing data for modeling neuronal responses. *Frontiers in Neuroscience*, 13, 986. <https://doi.org/10.3389/FNINS.2018.00986/BIBTEX>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Silchenko, A. N., Hoffstaedter, F., & Eickhoff, S. B. (2023). Impact of sample size and regression of tissue-specific signals on effective connectivity within the core default mode network. *Human Brain Mapping*, 44(17), 5858–5870. <https://doi.org/10.1002/hbm.26481>