A Systematic Review and Meta-Analysis of the Relationship Between Brain Data and the Outcome in Disorders of Consciousness

Boris Kotchoubey[1]* and Yuri G. Pavlov[1,2]

[1] Institute of Medical Psychology, University of Tübingen, Tübingen, Germany, [2] Department of Psychology, Ural Federal University, Yekaterinburg, Russia

A systematic search revealed 68 empirical studies of neurophysiological [EEG, event-related brain potential (ERP), fMRI, PET] variables as potential outcome predictors in patients with Disorders of Consciousness (diagnoses Unresponsive Wakefulness Syndrome [UWS] and Minimally Conscious State [MCS]). Data of 47 publications could be presented in a quantitative manner and systematically reviewed. Insufficient power and the lack of an appropriate description of patient selection each characterized about a half of all publications. In more than 80% studies, neurologists who evaluated the patients' outcomes were familiar with the results of neurophysiological tests conducted before, and may, therefore, have been influenced by this knowledge. In most subsamples of datasets, effect size significantly correlated with its standard error, indicating publication bias toward positive results. Neurophysiological data predicted the transition from UWS to MCS substantially better than they predicted the recovery of consciousness (i.e., the transition from UWS or MCS to exit-MCS). A meta-analysis was carried out for predictor groups including at least three independent studies with N > 10 per predictor per improvement criterion (i.e., transition to MCS versus recovery). Oscillatory EEG responses were the only predictor group whose effect attained significance for both improvement criteria. Other perspective variables, whose true prognostic value should be explored in future studies, are sleep spindles in the EEG and the somatosensory cortical response N20. Contrary to what could be expected on the basis of neuroscience theory, the poorest prognostic effects were shown for fMRI responses to stimulation and for the ERP component P300. The meta-analytic results should be regarded as preliminary given the presence of numerous biases in the data.

Keywords: consciousness, improvement criteria, meta-analysis, minimally conscious state, neurophysiological markers, prognosis, publication bias, unresponsive wakefulness syndrome

…the quality of methodological reporting in the social and behavioral science research literature is poor. Reports are often silent or ambiguous on important methodological and procedural matters making it difficult for the analyst to determine what was done. The metaanalyst who develops elaborate and detailed methodological criteria for study selection, therefore, will most likely find that study reports do not provide sufficient information for those criteria to be confidently applied. [Lipsey and Wilson (1) (p. 22)]

A systematic analysis of several published datasets can yield substantial new knowledge as compared with the data of each single experiment (2). This insight, increasingly admitted during the last decades, underlies the use of meta-analyses and other kinds of quantitative reviews that can largely overcome the subjectivity and deliberateness of the "good old" narrative reviews. The domain of the severe Disorders of Consciousness (DoC) is, however, still dominated by the latter genre. Thus a brief overview of journal publications about brain imaging data in DoC for the last 5 years reveals that almost every fourth paper (exactly, 33 of the 137 papers) is a narrative review. Book publications further increase this number.

The notion DoC most usually includes two diagnostic entities: the vegetative state, or unresponsive wakefulness syndrome (UWS), and the minimally conscious state (MCS) (3, 4). To the best of our knowledge, the first systematic analysis of neurophysiological data in DoC was devoted to the question whether these data confirm the reality of the distinction between UWS and MCS (5). The authors came to the conclusion that there were no reliable differences in terms of neurophysiological variables (mainly EEG, PET, and fMRI) between the two diagnoses.

Hannawi et al. (6) concentrated on brain imaging studies and performed a voxel-based meta-analysis of 13 PET and fMRI studies in DoC patients, in which the corresponding data were reported. The authors identified a number of structures whose resting state activity was significantly decreased in patients as compared with healthy controls. On the other hand, they did not find convincing differences between UWS and MCS, which was in line with Liberati et al. (5). Kondziella et al. (7) came, however, to a different conclusion that brain connectivity data in rest and under passive stimulation (but not in active instruction conditions) reliably differ between UWS and MCS. Unfortunately, inclusion criteria in this study were not completely clear; thus the question still remains open whether late event-related brain potential (ERP) components (P300, N400) can be regarded as indicators of cortical connectivity, and therefore, the authors should either include all P300 and N400 studies in their analysis (if they answer this question positively), or exclude all of them (if they answer it negatively), but instead, they included only some of them. The data were not checked for publication bias, that is, the tendency for positive results or stronger effects to get published more readily than negative results or weaker effects (8). The simplest index of this bias is a negative correlation between the size of the obtained effect and its reliability (9). On the other hand, Kondziella et al. (7) indicated a bias in patient selection. The risk of this bias was estimated as "high" in 81.4% of the analyzed studies and as "uncertain" in further 11.6%.

Also, Bender et al. (10) were interested in the abilities of neurophysiological techniques to distinguish between UWS and MCS. Their meta-analysis aimed not at the presence and size of the effects, but at the parameters of sensitivity and specificity. The authors concluded "… that modern diagnostic techniques can already make a major contribution to the diagnostic assessment of MCS." The inspection of their empirical findings yields a modest support for this conclusion, because good sensitivity and specificity values were found only for the measures of quantitative EEG; ERP and fMRI measures revealed, to the contrary,

only moderate specificity and rather low sensitivity that did not significantly differ from chance.

Kotchoubey (11) carried out a quantitative analysis of 61 reports on ERPs in DoC. ERPs are the most frequently used neurophysiological technique in DoC, which, however, does not mean that they are also most useful. In general, the results of the analysis were rather disappointing. Most studies possessed such a low statistical power that their findings can at best be regarded as "preliminary results." In addition, there was strong evidence for a publication bias toward positive findings.

However, there were good news. The above-mentioned deficits mainly concerned the studies where ERPs were compared between UWS and MCS, which largely concurs with the conclusions of Liberati et al. (5). The negative tendencies were substantially less expressed in the literature about the relationship between ERP and the prognosis of DoC outcome. Furthermore, the power of the prognostic studies correlated positively, and the effect sizes (ESs) correlated negatively, with the rank of journals where the data were reported. This indicated that weaker but more reliable effects could be published more successfully in top-ranking journals than strong but less reliable ones.

Like all areas in which there is no golden diagnostic standard, meta-analyses of novel diagnostic tests in the domain of DoC have a strong circular component. The expensive neurophysiological techniques are developed to complement imperfect clinical methods and to increase the diagnostic precision; but in a meta-analysis, these novel techniques are evaluated on the basis of the same (presumably imprecise) diagnostic criteria that these methods should improve! The lack of the golden diagnostic standard makes another strategy more preferable, i.e., a search for the measurements most reliably related to prognosis. The above-cited findings, that prognostic studies in DoC appear to have a higher quality than diagnostic studies, are in line with this view.

The aim of the present study was a systematic analysis of all publications relating any functional brain data recorded in UWS and MCS patients to their outcome several weeks or months after the measurement. Data using only anatomical brain measurements were not included in the present analysis. Each of the analyzed publications will hereafter be designated as "record." The term "dataset" will, in contrast, refer to any individual comparison between a neurophysiological variable (e.g., fMRI activation in a specific task) and an outcome variable (e.g., Glasgow Outcome Scale—Extended, or GOSE (12)). One record can, therefore, contain many datasets.

METHODS

Literature Search and ES Calculation

A search in MEDLINE and SCOPUS was conducted on the 23 November 2017 by using search terms ((prognos* OR predict* OR outcome) AND (vegetative state OR minimally conscious state OR unresponsive wakefulness syndrome)) AND (eeg OR fmri OR event related potentials OR erp OR positron emission). No time limits were set for the search. In addition, the systematic reviews cited above in the Introduction as well as recent

informal reviews on neurophysiology of DoC, were consulted. Eight hundred ninety-seven peer-reviewed records published in English, German, or Russian, were identified. After reading abstracts and removing duplicates, 822 of them were rejected as irrelevant. Full test was sought for the other 76 records. Seven of them were rejected because they did not contain any outcome data on DoC patients, or contained outcome data already published elsewhere. Further exclusion criteria were (a) case studies or series of cases; (b) using patients' survival, and not the clinical improvement, as the only prognostic criterion; (c) presentation of the results in such a general form that the size of the observed effects cannot be calculated; and (d) reporting the data of UWS and MCS patients together with other diagnoses such as coma, exit-MCS, or locked-in syndrome: on the basis of these criteria 22 records were rejected. Regarding (d), we accepted studies in which UWS and MCS data were reported together, but not those in which the sample included more than these two diagnoses and the reported data did not give the reader a possibility to distinguish between the different diagnostic groups. The process of the selection of relevant records is shown in **Figure 1**.

The search resulted in 47 records containing a total of 381 datasets. These records are summarized in **Table 1**. Effect size was calculated for each dataset (i.e., for each predictor–outcome pair) on the basis of primary data on each individual patient presented in the tables of most studies, or chi-squared based on the same patient data, or the t-statistics. Only in one record, ES was calculated from a coefficient of correlation. All these parameters were converted into Cohen's d following the methods summarized by Lenhard and Lenhard (13). If a resulting 2×2 table contained a zero cell (e.g., all patients having a positive neurophysiological sign recovered), the blind

application of the corresponding formulas results in $d = $ infinite; to avoid this, we added 0.5 to all cells, as recommended by Nakagawa and Cuthill (14). When d-values were included in further operations (added, averaged, etc.), they were weighted by inverse standard error (SE).

A big and still underestimated problem of all quantitative reviews is the plenty of non-reported data. Several authors (63–65) indicated that measured but unreported variables constitute one of the main sources of false positive findings ubiquitous in biology and psychology and thus an important cause of the contemporary "replication crisis" (66). When, and only when, it was evident for both present authors from the text of a paper that a neurophysiological variable was measured but not reported in relation to the outcome (or, rarely, reported as "non-significant"), the ES of this variable was assumed to be 0, and the SE of ES was assumed to be equal to the median SE calculated for the reported variables in the same record.

When the results presented several strongly correlated predictor variables (e.g., the same EEG variable in several adjacent regions), they were regarded as representing the same "construct" (1), and the mean and standard deviation (SD) for the construct were calculated according to the formulas

$$\text{Mean} = (M_1 + M_2) / N;$$

$$SD = \sqrt{\frac{(N-1)(SD_1^2 + SD_2^2) + 0.5N(M_1 + M_2 - 2M_1M_2)}{2N - 1}}$$

where M_1 and M_2, SD_1 and SD_2 are mean and SD values for two to-be-combined variables, respectively. (Note that the formulas are so simple because both variables have the same N.)

FIGURE 1 | Flow chart of the selection of records.
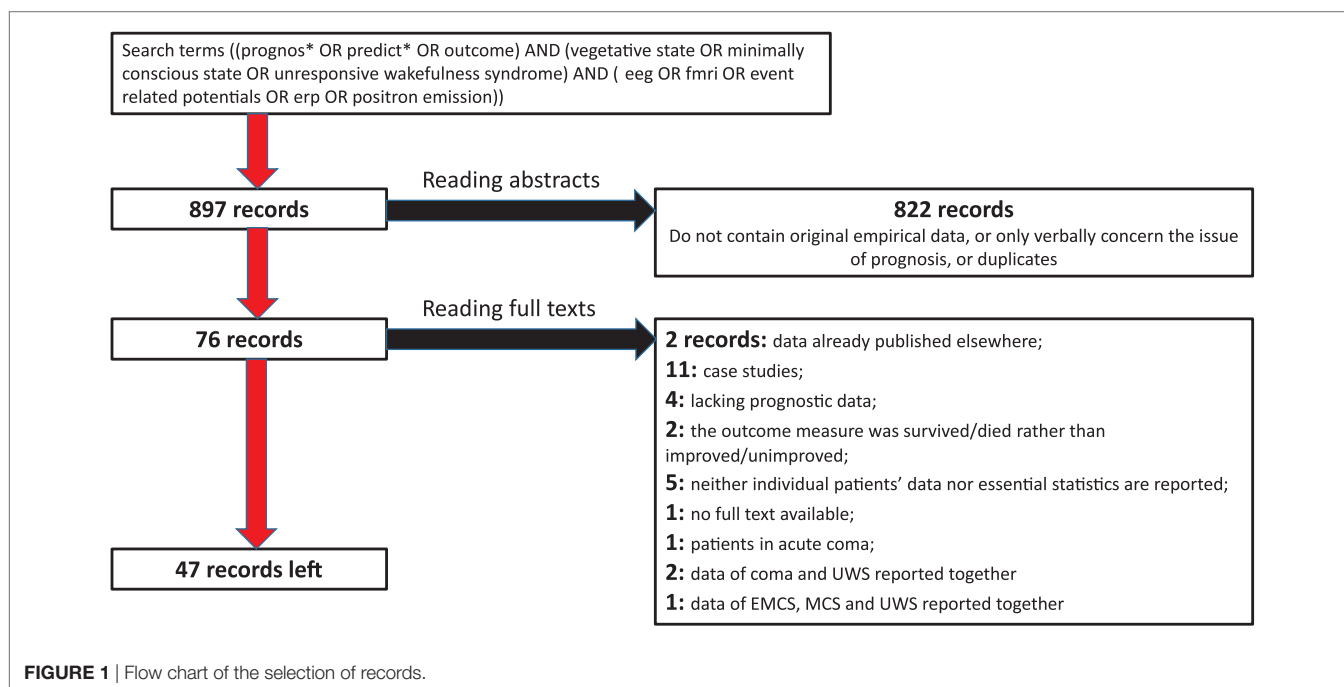
TABLE 1 | A summary of the 47 records included in the systematic review.

Reference	N (UWS/MCS)	Follow up, months	Methods	Relevant measures	Improvement criteria[a]
(15)	64/0	2	EEG, 24-h polysomnography	General sleep patterns Dominant EEG rhythm: beta, theta or delta	Diagnosis MCS, or GOS > 2
(16)	14/4	6–38	24-h polysomnography	Sleep complexity, presence of different sleep stages	CRS-R
(17)	10/0	36	~14-h polysomnography	Sleep complexity, number of sleep spindles, presence of different sleep stages	GOS
(18)	42/0	3	EEG	Spectral power of resting state EEG in delta, theta, alpha1, alpha2, beta1 and beta2 frequency bands	LCF > 5
(19)	59/47	3	EEG	EEG amplitude normality (>20 μV), dominant frequency, reactivity to forced eyes opening EEG combined AFR index	CRS-R; change from UWS to MCS or from MCS to EMCS
(20)	28/0	6	EEG	The same as in Ref. (19)	CRS-R; change from UWS to MCS or from MCS to EMCS
(21)	12/1	3	EEG	EEG normality according to Synek scale	LCF
(22)	4/5	6	Auditory ERP	MMN, N2 and/or P300 in control and after listening to music condition	CRS-R ≥ MCS
(23)	34/0	24	Short-latency EP, somatosensory EP, EEG	BAEPs grade, EEG reactivity to passive eyes opening, pain and acoustic stimuli, EEG Synek index, N20 SEP grade, P300 to a patient's own name	DRS < 22
(24)	17	12	PET	FDG-PET in the resting state	GOSE > 2
(25)	7/0	2–9	fMRI	BOLD response to speech and noise stimulation	CRS-R ≥ MCS
(26)	22/16	6	fMRI	fMRI BOLD response to speech and sound stimulation	CRS-R
(27)	7/4	3	fMRI	fMRI BOLD response to subject's own name	CRS = MCS
(28)	3/7	6	EEG, fMRI	EEG, BOLD response to language, music, active motor imagery instruction	GOSE > 2
(29)	43/0	24	EEG, somatosensory EP	EEG classified according to Synek scale N20 SEP grade	CRS-R ≥ MCS
(30)	14/0	3	EEG	Resting state EEG Index of Structural Synchrony (amplitude, length, instability, number of functional connections in Alpha, Beta1, Beta2 bands)	LCF = MCS
(31)	8/0	24	Auditory ERP	N2, P300	Recovery of awareness but no standardized assessment
(32)	20/0	NA	Auditory and visual ERP, SPECT	Auditory MMN, N100, N200, P300; visual EP present/absent; resting state brain metabolism assessed by SPECT	GOS > 2
(33)	75/38	4	Somatosensory EP	N20	CRS scores of ≥ 23
(34)	56/0	12	EEG, somatosensory EP, 24-h polysomnography	EEG reactivity to noxious stimulation, N20, sleep spindles in 24-h EEG All predictors scored as absent or present	GOS > 2 or transition UWS to MCS
(35)	10/0	3	Somatosensory EP	N20 grade and latency	Recovery of awareness but no standardized assessment
(36)	1/4	3	fMRI	fMRI default mode network normality	Level of consciousness according to the Multi-Society Task Force on PVS
(37, 38)	24/19	6	Auditory ERP, EEG	MMN, N400, EEG dominant background activity	DRS ≥ MCS
(39)	6/5	12	EEG, fMRI	EEG reactivity to warm water stimulation, fMRI activation to thermal stimulation	GOS > 2 or transition from UWS to MCS
(40)	12/10	1, 2, 3, 6, 9, 12	Auditory ERP	P300	CRS-R ≥ MCS
(41)	50/0	5	EEG	EEG normality according to Synek scale, EEG reactivity to pain stimulation	Regaining consciousness according to GOS, LCF
(42)	23/0	6	fMRI	fMRI BOLD response to speech (adapted affective speech)	GOS > 2 or transition from UWS to MCS
(43)	11/0	6	Polysomnography	REM sleep characteristics	Recovery of awareness but no standardized assessment
(44)	6/2	3	Auditory ERP	MMN to subject's own name stimuli, N100	CRS-R
(45)	6/5	3	PET	PET global GABA A receptor binding	CRS-R
(46)	52/0	3	fMRI	fMRI resting state connectivity	GOS > 2
(47)	5/0	0.5–2	EEG-TMS	TMS-evoked cortical responses	CRS-R

(Continued)

TABLE 1 | Continued

Reference	N (UWS/ MCS)	Follow up, months	Methods	Relevant measures	Improvement criteria[a]
(48)	38/0	6	EEG	Resting state EEG Approximate Entropy, EEG reactivity (stimulation protocol is not described)	GOSE > 2
(49)	56/0	12	EEG	Spectral power in Delta, Alpha, Theta, Beta, Gamma frequency bands	CRS-R = MCS
(50)	71/0	1.5	EEG, auditory ERP	92 measures including: CNV, MMN, P1, P3a, P3b; normalized and absolute spectral power of delta, theta, alpha, beta, gamma rhythms; permutation entropy, Komolgorov–Chaitin Complexity; phase lag index (PLI), spectral entropy, imaginary coherence and weighted symbolic mutual information (wSMI) in different frequency bands	CRS-R
(51)	18/51	12	PET, fMRI	Resting state FDG-PET, BOLD response to active motor and visuospatial imagery tasks	GOSE > 2
(52)	53/39	24	Auditory ERP	N400, P300	CRS-R = EMCS
(53)	9/0	2–54	PET	Resting state FDG-PET	Recovery of awareness but no standardized assessment
(54)	10/12	1–6	fMRI	BOLD response to active motor and visuospatial imagery tasks	CRS-R; change from UWS to MCS, or MCS to EMCS
(55)	39/25	12	fMRI	fMRI BOLD response to subject's own name	CRS-R
(56)	6/5	6	Auditory ERP	MMN, P300 to subject's own name	CRS-R
(57)	10/0	24	Auditory ERP	MMN	LoC > 6
(58)	10/0	24	Auditory ERP	N200, N350, P300 in active and passive paradigms	LoC > 6
(59)	11/0	26–36	Visual ERP	N2, N3, P2 amplitude and latency, P2–P3 peak to peak magnitudes of VEP	LoC > 6
(60)	10/8	1–150	EEG, 24-h polysomnography	Permutation entropy, alpha-to-theta ratio, density of slow waves, high-to-low frequencies ratio, density of sleep spindles	GOSE > 2 or CRS-R
(61)	21/0	6	Short-latency EP, EEG, somatosensory EP	BAEP, N20 SEP grade, EEG normality, approximate entropy (ApEn), cross-approximate entropy, Lempel–Ziv complexity to pain, auditory and music stimulation in comparison with eyes-closed condition	GOS > 2
(62)	36/0	12 (after injury)	Somatosensory ERP	N20, P25, N20–N25 SEP grade and amplitude	GCS ≥ MCS

N (UWS/MCS) means the number of UWS and MCS patients whose outcome and neurophysiological data were available. In the case of different number of patients available for different neurophysiological measurements, only the largest number is reported; it may be less than the total number of patients in the study.
[a]If GOS(E), CRS-R, LCF mentioned with no additional description, it was possible to calculate improvement criteria either way.
NA, not available, AFR index, Amplitude/Frequency/Reactivity index; UWS, unresponsive wakefulness syndrome; MCS, minimally conscious state; EMCS, Exit form MCS; GOS(E), Glasgow Outcome Scale (Extended), GCS, Glasgow Coma Scale; DRS, Disability Rating Scale; LCF, levels of cognitive functioning scale; LoC, Level of Consciousness Scale; SEP, somatosensory evoked potentials; BAEP, brain stem auditory evoked potentials; dwPLI, debiased weighted phase lag index; SPECT, single-photon emission computed tomography; CRS-R, Coma Recovery Scale-Revised.

Outcome Criteria

The category "bad outcome" for UWS and MCS patients presumed remaining in the same condition. Deaths were included in this category only if it was clear that a patient died as a direct consequence of the brain lesion, otherwise excluded from the analysis. The category "good outcome" has, on the other hand, two different definitions: minimal clinical improvement or regaining full consciousness.[1] For MCS, both criteria are the same, because their minimal improvement implies the transition to Exit-MCS. But this is not true for UWS, because their minimal improvement means only the transition to MCS. As shown in Section "Results," the two different improvement criteria of UWS yield different results.

Avantaggiato et al. (17) analyzed a group of DoC patients containing children and adolescents; because the authors presented individual data of each patient, we selected the results for patients >13 years only. The category "good prognosis" for MCS implied the recovery of consciousness. For UWS, however, it might include either the recovery of consciousness or the transition into MCS. All 381 datasets were included in the systematic review.

Quality Assessment

Quality of the records was estimated on the basis of the QUADAS criteria (67) that have been tailored, as recommended in the original publication, for the specific research field. Because Kondziella et al. (7) expressed concerns about possible bias in patient selection, we recorded whether a publication included a patient flow chart, and whether it described explicit exclusion criteria for UWS and MCS patients or simply mentioned that all patients admitted in the clinics for a particular time period were investigated. We also marked the records in which obviously more neurophysiological data were collected than reported in the analysis of outcome prediction. Another quality index was the use of the Coma Recovery Scale-Revised (CRS-R (68)) for DoC diagnostics, because this scale, though not being golden

---

[1]Practically, "full consciousness" is defined as reliable communication or functional use of objects. In the following, when we speak about "regaining of consciousness" we shall mean these two abilities (at least one of them) and not just minimal signs of consciousness presented also in the MCS.

standard, possesses substantially better psychometric qualities than all other instruments for the assessment of DoC (69). In addition, the impact factor (IF) of the publishing journal was included as an indirect quality criterion, because the data indicate high correlation between the IF and the informal reputation of the corresponding journal among neurologists (70).

Meta-Analysis

Two additional inclusion criteria for meta-analyses were (1) at least three independent records reporting the same predictor variable or variables related to the same construct and (2) each of these records includes at least 10 patients. The criteria can be regarded as very liberal because, first, only three records and only 10 patients (who further should be subdivided into at least two groups) are rather low numbers, and second, the notion of construct is rather vague and permits to include into one group, for example, studies of P300 to simple tones and to patients' own names, thus increasing heterogeneity. On the basis of these criteria, 319 datasets were excluded (of course, this number would be larger if the criteria were more conservative). The remaining 62 datasets finally entered the meta-analysis.

We used a random-effects meta-analysis with restricted maximum-likelihood (REML) estimator for pooling ESs. We assessed the level of heterogeneity between studies with a standard Q-test statistic as well as by I^2 calculation (71). Heterogeneity was regarded as significant when $p < 0.05$ or $I^2 > 50\%$. Potential publication bias for individual predictors was assessed with the Egger test for Funnel plot asymmetry and represented graphically with Begg's funnel plots of the ES versus its SE. Additionally, Rosenthal fail-safe test was also applied. All meta-analyses were performed using R package "metafor" (72) using inverted standard errors as weighting parameter.

RESULTS

Quality of Reporting

None of the 47 records presented a flow chart depicting patient selection. The authors of 20 records (42.6%) explicitly state that they included all patients within some exactly described time period. Nine reports (19.1%) depicted at least some inclusion and/or exclusion criteria. In the remaining 18 records, patient selection was not described.

Most selected records present their data either for each individual patient or as mean ± SD (or SE) for each relevant group (e.g., recovered versus non-recovered). Three records present data only in a general form (e.g., as correlations). Five records (10.6%) mention the size of some effects.

CRS-R was used for the diagnosis of UWS and MCS in 29 records (61.7%); other studies employed Disability Rating Scale, Glasgow Coma Scale, or other less powerful instruments.

Quite surprisingly, only two records explicitly state that the neurologists who assessed the outcome were blinded to the neurophysiological data collected before. In six records, blindness of the outcome might be assumed because neurophysiological examination and outcome diagnostics were performed in different institutions. These eight records (17%) were combined into one "blind" group. In the majority of the records (83%), the outcome was diagnosed with knowledge of the neurophysiological findings.

The median time between the measurement and the outcome assessment was 6 months, mean minimal time per record was 9 months (range 1–36 months), and the mean maximal time per record 16 months (range 1.5–150 months). In 42 records (89%), this interval was same for all the examined patients. Seven records used broad variable intervals for different patients (1–6, 1–30, 2–9, 6–38, 10–150, 26–36, and 24–144 months). One publication does not report the measurement–outcome interval.

The mean total sample was 31.11 ± 3.66 patients, with a median of 21 patients and a range of 5–123 patients. Eight records included <10 patients, 14 records had between 10 and 19 patients. Adding the case studies filtered out at the previous stage, we come to the result that about a half of all prognostic studies included <20 patients. The records that did not describe patient selection included significantly less patients (means 14.1 versus 43.0, $t = 5.11$, $p < 0.001$) than records describing their selection process.

The median IF of the publishing journals was 3.87, range from 0 to 44. IF did not differ between the records with correctly versus incorrectly described patient selection. We hypothesized that studies with more patients (thus having higher power) are published in more prestigious journals, but the corresponding correlation was not significant (Spearman's $\rho = 0.26$, $p = 0.078$). However, studies employing CRS-R were published in journals with higher IF than studies that did not use this scale: $p < 0.001$, Mann–Whitney test. A few studies with blind outcome assessment included larger sample sizes than studies without outcome blindness ($t = 2.35$, $p = 0.023$) and were also published in more prestigious journals ($p = 0.046$, Mann–Whitney test).

Notably, we did not find a significant relationship between any of the variables and the time elapsed from the neurophysiological measurement till the assessment of the outcome. Non-weighted ES correlated with the mean time between neurophysiological measurement and outcome assessment with Spearman's $\rho = -0.09$, with minimal time per study $\rho = -0.18$, with maximal time per study $\rho = -0.07$ (all nonsignificant). For weighted ES, the corresponding correlations were 0.03, −0.06, and −0.11, respectively (all nonsignificant). Likewise, correlations of the time interval with the SD of ES (as a measure of its reliability) were all between 0.00 and 0.03, and correlations of the time interval with sample size were between 0.00 and 0.06, all non-significant.

Also, the year of publication did not correlate with any other measures. The bibliographic literature gives a reason to expect that later publications might have large samples or smaller ESs than earlier (73). Although the corresponding correlations were in the expected direction, they did not reach significance (year/N: Spearman's $\rho = 0.12$; year/ES: $\rho = -0.08$). Also, the correlation between publication year and IF was close to 0 ($\rho = 0.03$).

Publication Bias

The inverted SE (1/SE) was taken as a measure of the reliability of an ES. Across all datasets, the rank-order correlation

between ES and its 1/SE was weak but significant ($\rho = -0.22$, $k = 381$, $p < 0.001$). The result might be biased because different records contribute disproportionally to the whole mass of data. However, the positive correlation between the ES and its SE became even stronger when calculated for the selected subset of datasets included in the meta-analysis ($\rho = -0.47$, $k = 62$, $p < 0.001$), as well as for the ESs averaged for each record ($\rho = -0.41$, $k = 47$, $p = 0.004$). The results are shown in **Figure 2**.

FIGURE 2 | Negative correlations between effect size (ES) and its reliability, estimated by the inverted standard error (1/SE), for all individual datasets [**(A)**: Spearman's $\rho = -0.22$, $p < 0.001$], datasets included in the meta-analysis [**(B)**: $\rho = -0.47$, $p < 0.001$], and for mean ESs per record [**(C)**: $\rho = -0.41$, $p = 0.004$]. The regression lines are presented for illustration only, not for quantitative analysis.
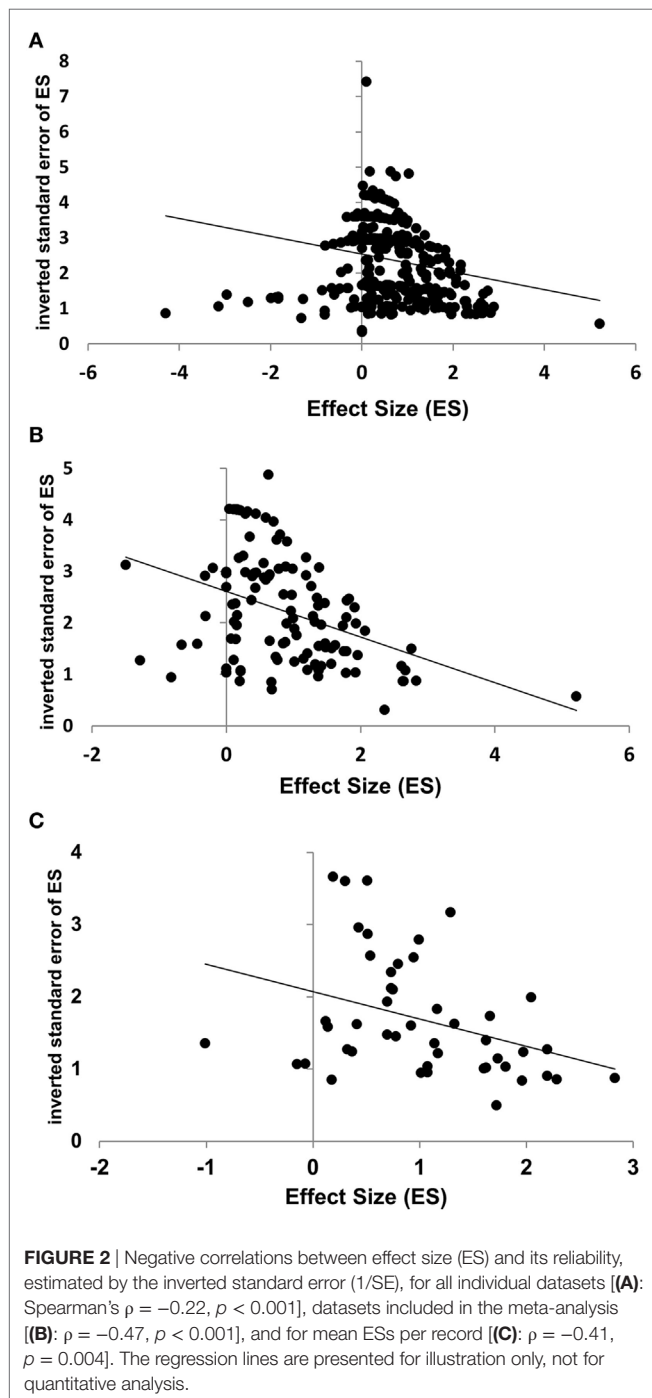
Note that the first analysis (across all datasets) overestimates the contribution of the records reporting many datasets. The last analysis (across averaged ES), to the contrary, may overestimate the contribution of the records presenting few or only one dataset. Despite this contrary bias, very similar results were obtained. To sum up, these data show a trend to selective publication of strong but unreliable effects. How serious the bias is in respect of each particular predictor variable will be discussed below.
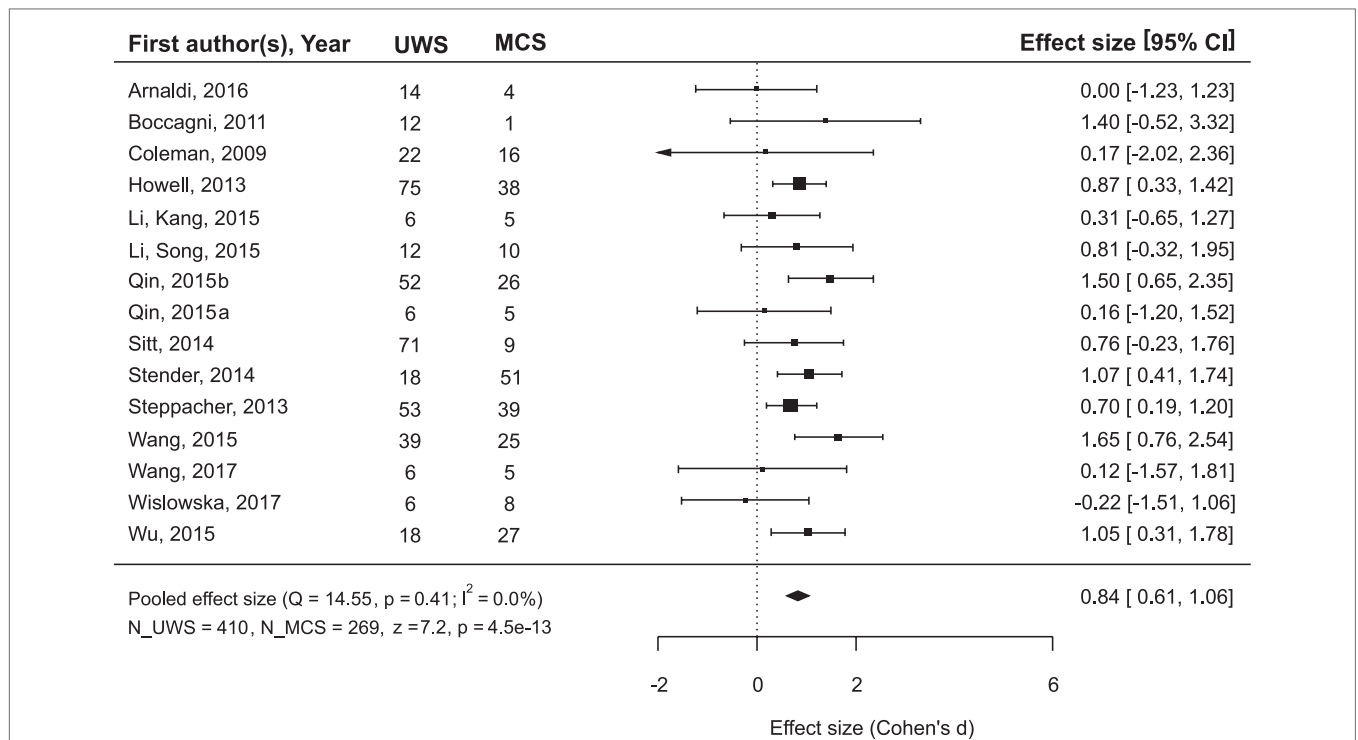
UWS and MCS

Thirteen of the 381 datasets included only MCS samples, 248 datasets included only UWS patients, and the remaining datasets included both diagnostic groups of DoC patients.

While the main issue of the present study was outcome prediction on the basis of neurophysiological data, we also asked the question whether the outcome can be predicted simply from the diagnosis. Many authors of the reviewed articles also asked this question and answered it negatively. However, a meta-analysis of the combined data from the records where both diagnosis and prognosis could be followed revealed that MCS patients recovered consciousness significantly more frequently than UWS patients (**Figure 3**): mean ES = 0.84, 95% CI from 0.61 to 1.06. On the other hand, if the positive outcome of UWS patients is defined as any minimal improvement, i.e., the transition to the MCS, the diagnosis loses its predictory value (**Figure 4**).

Because we found that the improvement criterion for UWS (transition to the MCS versus recovery of full consciousness, that is, exit-MCS) can play a role in the calculation of prediction effects, we compared weighted mean ES for the neurophysiological variables in three conditions: (i) prediction of the recovery of consciousness for MCS patients; (ii) prediction of the recovery of consciousness for UWS patients; and (iii) prediction of the transition to MCS for UWS patients. A one-way ANOVA across these three groups resulted in a highly significant effect: $F(2,294) = 23.11$, $p < 0.001$. The result does not change when we limit the analysis by only those datasets that will later enter the meta-analysis [$F(2,60) = 19.88$, $p < 0.001$], or when we exclude all mixed datasets [$F(2,86) = 12.08$, $p < 0.001$]. Independently of the method of calculation, the mean weighted ES for the groups (i) and (ii) (i.e., different diagnoses, the same improvement criterion) were very similar and varied—dependent on the selected data—between 0.41 and 0.48. The mean weighted ES for the group (iii) was about three times larger (between 1.40 and 1.68) and differed significantly from both of them, although the groups (iii) and (ii) included patients with the same diagnosis and some datasets involved in these two groups might even include some of the same patients. To sum up, neurophysiological methods are significantly more successful in prediction of the transition from UWS to MCS than in prediction of the recovery of full consciousness.

Meta-Analysis of Predictory Constructs

According to the above results, we performed the meta-analysis separately for (a) prediction of any clinical improvement (for which UWS patients means at least transition to the MCS), and

First author(s), Year	UWS	MCS	Effect size [95% CI]
Arnaldi, 2016	14	4	0.00 [-1.23, 1.23]
Boccagni, 2011	12	1	1.40 [-0.52, 3.32]
Coleman, 2009	22	16	0.17 [-2.02, 2.36]
Howell, 2013	75	38	0.87 [0.33, 1.42]
Li, Kang, 2015	6	5	0.31 [-0.65, 1.27]
Li, Song, 2015	12	10	0.81 [-0.32, 1.95]
Qin, 2015b	52	26	1.50 [0.65, 2.35]
Qin, 2015a	6	5	0.16 [-1.20, 1.52]
Sitt, 2014	71	9	0.76 [-0.23, 1.76]
Stender, 2014	18	51	1.07 [0.41, 1.74]
Steppacher, 2013	53	39	0.70 [0.19, 1.20]
Wang, 2015	39	25	1.65 [0.76, 2.54]
Wang, 2017	6	5	0.12 [-1.57, 1.81]
Wislowska, 2017	6	8	-0.22 [-1.51, 1.06]
Wu, 2015	18	27	1.05 [0.31, 1.78]

Pooled effect size (Q = 14.55, p = 0.41; I^2 = 0.0%)
N_UWS = 410, N_MCS = 269, z = 7.2, p = 4.5e-13   0.84 [0.61, 1.06]

Effect size (Cohen's d)

FIGURE 3 | The results of the meta-analysis for prediction of the outcome from the diagnosis. The criterion of improvement for all patients was recovery of full consciousness. *Q*, the corresponding *p*-value and *I²* are estimates of between-study heterogeneity; symbols ■ stay for the estimates of effect size (ES) in each single study, with the size of the symbol being proportional to the precision of the estimate. Error bars indicate the 95% confidence intervals of ES. The diamond ♦ is the estimate of the overall effect, the edges of the diamond represent the 95% confidence interval limits; CI, confidence interval; UWS and MCS, sample size of UWS and MCS patients in individual studies; N_UWS and N_MCS, overall sample size of the two patient groups. The resulting ES was tested for significance using *z*-criterion; the values of *z* and the corresponding *p* are given at the end of the lower left line.

First author(s), Year	UWS	MCS	Effect size [95% CI]
Arnaldi, 2016	14	4	0.99 [-0.37, 2.35]
Boccagni, 2011	12	1	0.96 [-0.92, 2.83]
Coleman, 2009	22	16	-1.83 [-3.45, -0.22]
Edlow, 2017	2	7	0.61 [-1.69, 2.91]
Kotchoubey, 2005	24	19	0.24 [-0.44, 0.93]
Qin, 2015a	6	5	0.16 [-1.20, 1.52]
Sitt, 2014	71	9	-0.37 [-1.27, 0.52]
Vogel, 2013	10	12	0.19 [-0.75, 1.12]
Wang, 2015	39	25	0.18 [-0.36, 0.73]
Wang, 2017	6	5	-0.76 [-2.26, 0.73]
Wislowska, 2017	6	8	-0.61 [-1.85, 0.64]

Pooled effect size (Q = 11.85, p = 0.29; I^2 = 0.0%)
N_UWS = 212, N_MCS = 111, z = 0.27, p = 0.79   0.04 [-0.26, 0.34]

Effect size (Cohen's d)

FIGURE 4 | The results of the meta-analysis for prediction of the outcome from the diagnosis. The criterion of improvement was "minimal improvement," that is for UWS patients, it was the transition to MCS, and for MCS patients, at least the transition to Exit-MCS. As can be seen, with this improvement criterion the diagnosis does not predict outcome. The rest is the same as in **Figure 3**.

(b) prediction of the recovery of full consciousness. Sixty-two datasets comprising a total of 1,919 patients were analyzed. They involved the following potential predictors:

1. EEG reactivity to "passive" stimulation (i.e., without an active instruction). Hypothesis: stronger EEG oscillatory responses = > better prognosis.
2. EEG entropy indices. Hypothesis: higher EEG entropy = > better prognosis.
3. EEG dominant oscillatory activity. Hypothesis: background activity closer to the alpha frequency = > better prognosis.
4. EEG Synek score (74) is frequently used in the intensive care medicine for the prognosis of the outcome of acute coma. Hypothesis: higher score = > better prognosis.
5. fMRI BOLD response to passive (auditory or nociceptive) stimulation. Hypothesis: stronger response = > better prognosis.
6. Resting state PET or SPECT metabolism. Hypothesis: closer to normal brain metabolism = > better prognosis.
7. N20 component of somatosensory evoked potentials (SSEP). Hypothesis: normal N20 = > better prognosis.
8. Auditory Mismatch Negativity (MMN) to a change in ongoing acoustic stimulation. Hypothesis: larger MMN = > better prognosis.
9. Auditory P300 as an index of complex processing in cortico-subcortical networks. Hypothesis: larger amplitude or shorter latency = > better prognosis.
10. Spindle activity as an index of information processing in sleep (75). Hypothesis: presence of sleep spindles = > better prognosis.

The findings are summarized in **Figures 5** and **6** and presented in more detail in Figures S1 and S2 and Tables S2 and S3 in Supplementary Material.

Only oscillatory EEG responses to passive stimulation appear to be reliably related to both prognostic criteria (i.e., minimal clinical improvement and the recovery of full consciousness). The included datasets are highly homogenous and yield a highly significant mean ES of 1.45 on a sample of 99 patients. Another prediction variable that significantly predicted the recovery of consciousness was brain metabolism assessed by PET/SPECT. It attained a mean ES of 1.40 on a sample of 106 patients. The prognostic value of the MMN, P300, EEG entropy variables, and fMRI responses to passive stimulation was not significant and characterized by strong heterogeneity of the primary datasets.[2]

More promising results have been obtained in relation to the minimal improvement criterion. In addition to the EEG reactivity, significant effects are found for the MMN and sleep spindles. Formally positive results are obtained for the SSEP component N20 and the background EEG frequency, but the data are too heterogeneous to make a conclusion.

---

[2]The ES in Wijnen et al. (57) strongly differs from all other effects in the MMN group. After removing this result, the MMN data become completely homogenous ($I^2 = 0$).
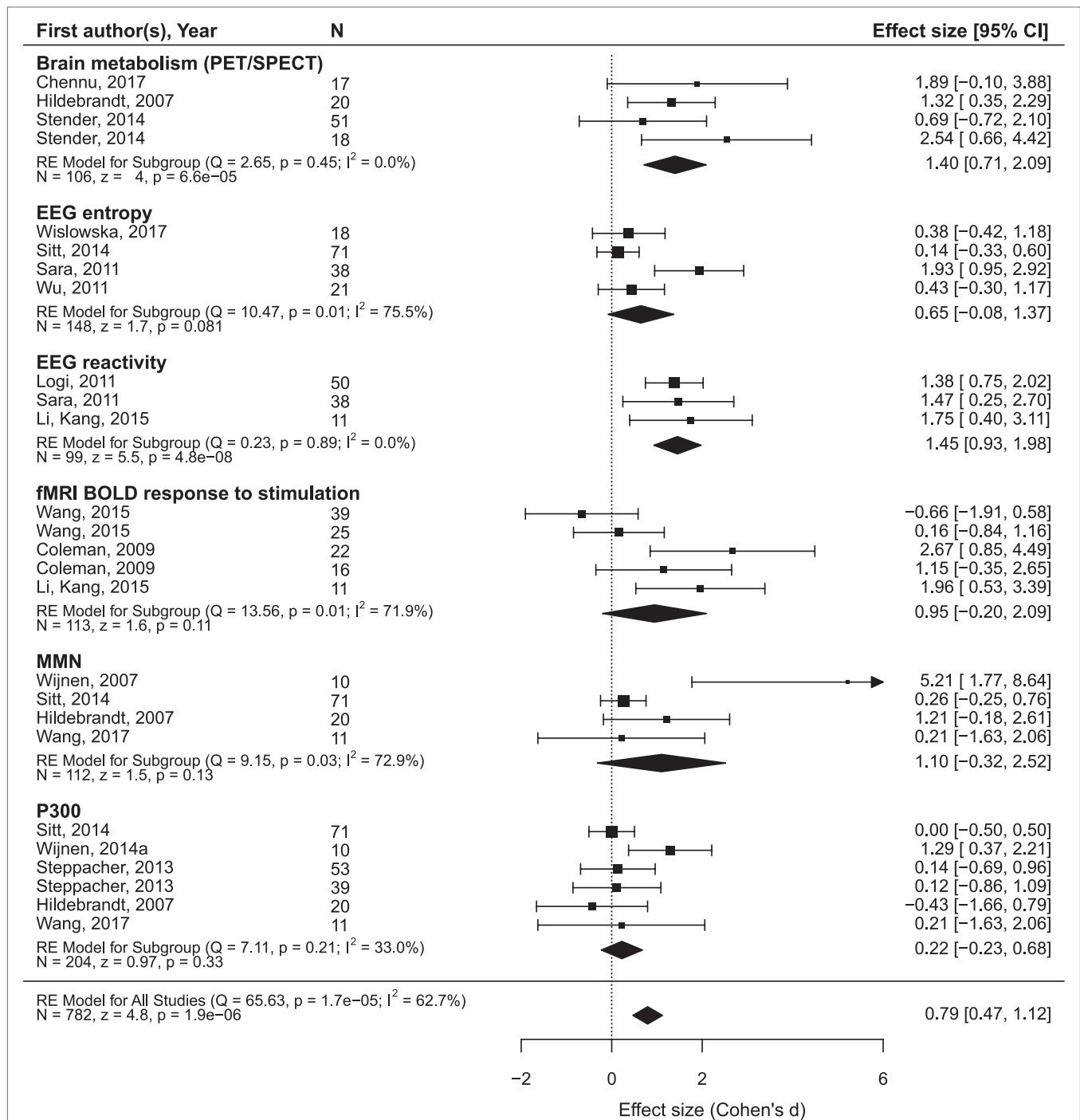
DISCUSSION

Although the Section "Limitation" is frequently placed at the end of Discussion, we believe that particularly the discussion of meta-analytic data is useful to begin with limitations. One important limitation is that of the present work as such. As the manuscript was prepared for a special issue, we did not systematically address the authors of the original publications but relied solely on the published data including supplementary information. Although we believe that personal contact with the authors may have enhanced our knowledge, at the first step we did not use this strategy because it might have caused considerable delays.

Other limitations of the meta-analysis are rather related to the limitations of the primary literature. A quantitative review can overcome some limitations of the reviewed studies, e.g., their small size (thus, it can reveal a consistent and significant effect on the basis of several inconsistent and non-significant ones), but it cannot remove the biases implied in its empirical basis.

To begin with the least, the quality of data reporting is far from the present-day standard. While Fritz et al. (76) bristled that only 42% of empirical psychological studies report the size of their effects, in the present sample ES was mentioned only in five records (10.6%). Presenting a patient flow chart is already a standard in many fields of clinical research but fully unknown in the domain of DoC. About a half of the reviewed records neither describe inclusion and exclusion criteria nor even make a simple statement that all patients admitted to the hospital during some period were included. Thus, the concern of Kondziella et al. (7) about possible bias in patient selection seems to be justified.

The majority of the reviewed studies employed a univariate approach, i.e., each predictor was separately compared with the target variable. Of course, this is a serious limitation because we know that the values of a regression strongly depend on the other predictor variables included or excluded in/from the equation. Thus, a neurophysiological variable (e.g., P300), which appears useless as a single predictor, might reveal its effect in a particular combination with other predictors. A few groups have recently attempted to overcome this limitation and employed a multivariate approach to outcome prediction in DoC (24, 50). This seems to be a perspective line of research, but now the number of such records is still too low to undertake a separate meta-analysis of these data.

The negative correlation between the description of patient selection and the number of patients suggests that selection bias might be particularly strong in small-size studies. Although the sample size in prognostic studies is on average larger than in studies comparing UWS and MCS, we found, together with single case studies, 33 records with less than 20 patients, which implies that even in the case of the equal distribution at least one of the outcome groups (recovered or non-recovered) includes <10 patients. Particularly, the studies with the total $N < 10$ result in huge confidence intervals making any reliable conclusion impossible. Of course, small or even single case studies may have sense at the very beginning of the research process, when nothing is known about an investigated phenomenon whatsoever. However, in such a case, one can expect an increase

First author(s), Year	N	Effect size [95% CI]

Brain metabolism (PET/SPECT)
Chennu, 2017	17	1.89 [−0.10, 3.88]
Hildebrandt, 2007	20	1.32 [0.35, 2.29]
Stender, 2014	51	0.69 [−0.72, 2.10]
Stender, 2014	18	2.54 [0.66, 4.42]

RE Model for Subgroup (Q = 2.65, p = 0.45; I² = 0.0%)
N = 106, z = 4, p = 6.6e−05 — 1.40 [0.71, 2.09]

EEG entropy
Wislowska, 2017	18	0.38 [−0.42, 1.18]
Sitt, 2014	71	0.14 [−0.33, 0.60]
Sara, 2011	38	1.93 [0.95, 2.92]
Wu, 2011	21	0.43 [−0.30, 1.17]

RE Model for Subgroup (Q = 10.47, p = 0.01; I² = 75.5%)
N = 148, z = 1.7, p = 0.081 — 0.65 [−0.08, 1.37]

EEG reactivity
Logi, 2011	50	1.38 [0.75, 2.02]
Sara, 2011	38	1.47 [0.25, 2.70]
Li, Kang, 2015	11	1.75 [0.40, 3.11]

RE Model for Subgroup (Q = 0.23, p = 0.89; I² = 0.0%)
N = 99, z = 5.5, p = 4.8e−08 — 1.45 [0.93, 1.98]

fMRI BOLD response to stimulation
Wang, 2015	39	−0.66 [−1.91, 0.58]
Wang, 2015	25	0.16 [−0.84, 1.16]
Coleman, 2009	22	2.67 [0.85, 4.49]
Coleman, 2009	16	1.15 [−0.35, 2.65]
Li, Kang, 2015	11	1.96 [0.53, 3.39]

RE Model for Subgroup (Q = 13.56, p = 0.01; I² = 71.9%)
N = 113, z = 1.6, p = 0.11 — 0.95 [−0.20, 2.09]

MMN
Wijnen, 2007	10	5.21 [1.77, 8.64]
Sitt, 2014	71	0.26 [−0.25, 0.76]
Hildebrandt, 2007	20	1.21 [−0.18, 2.61]
Wang, 2017	11	0.21 [−1.63, 2.06]

RE Model for Subgroup (Q = 9.15, p = 0.03; I² = 72.9%)
N = 112, z = 1.5, p = 0.13 — 1.10 [−0.32, 2.52]

P300
Sitt, 2014	71	0.00 [−0.50, 0.50]
Wijnen, 2014a	10	1.29 [0.37, 2.21]
Steppacher, 2013	53	0.14 [−0.69, 0.96]
Steppacher, 2013	39	0.12 [−0.86, 1.09]
Hildebrandt, 2007	20	−0.43 [−1.66, 0.79]
Wang, 2017	11	0.21 [−1.63, 2.06]

RE Model for Subgroup (Q = 7.11, p = 0.21; I² = 33.0%)
N = 204, z = 0.97, p = 0.33 — 0.22 [−0.23, 0.68]

RE Model for All Studies (Q = 65.63, p = 1.7e−05; I² = 62.7%)
N = 782, z = 4.8, p = 1.9e−06 — 0.79 [0.47, 1.12]
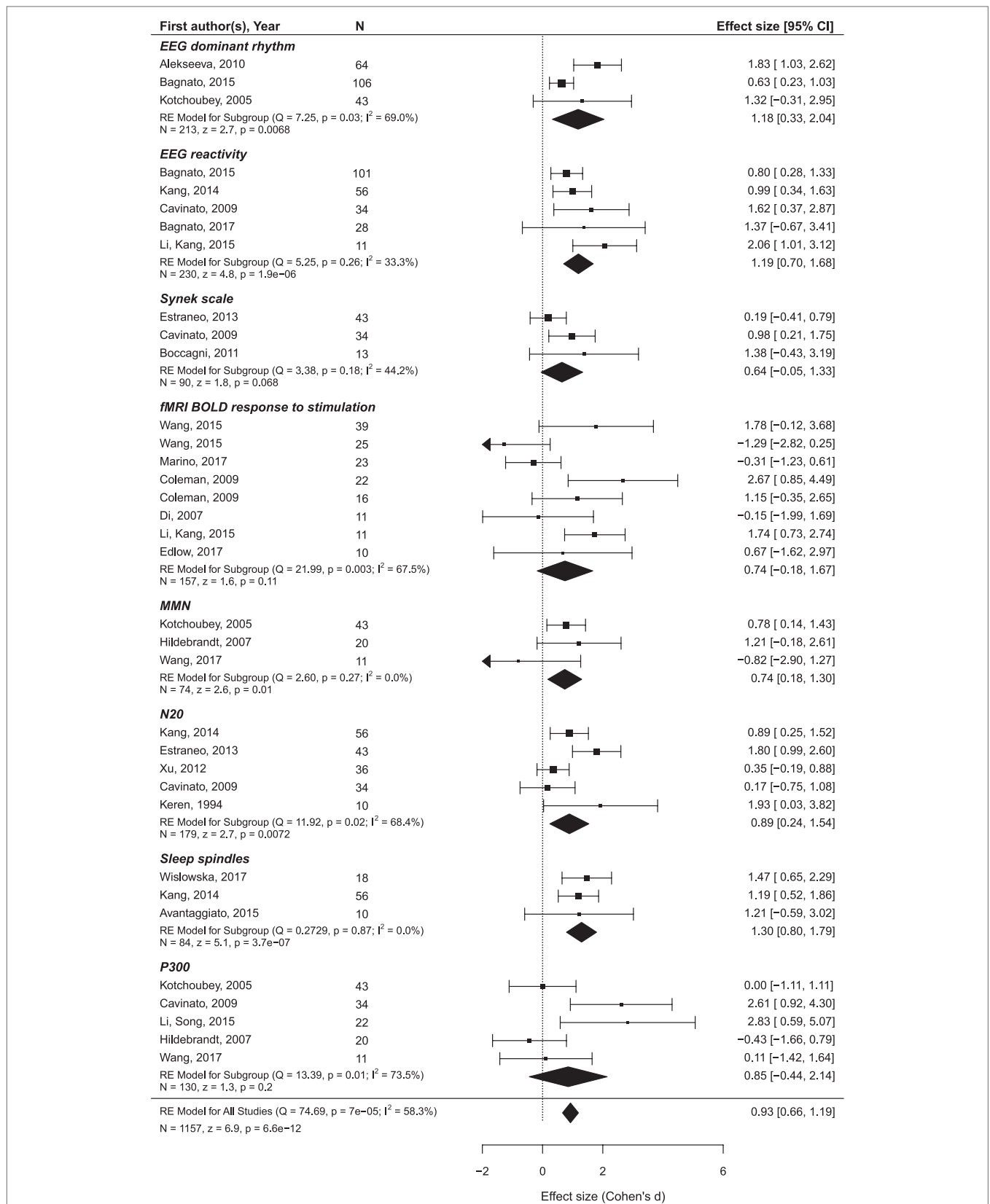
Effect size (Cohen's d): −2   0   2   6

FIGURE 5 | The results of the meta-analysis for prediction of the outcome from neurophysiological variables. The criterion of improvement for all patients was the recovery of consciousness. RE, random effects. The rest is the same as in **Figure 3**.

of sample sizes with the year of publication, but this correlation was not significant.

Strong negative effects of underpowered studies on the quality of the reported data have thoroughly been discussed in the literature in general (63, 77, 78) and specifically in neuroscience (79, 80). Positive findings of tiny studies can only result from chance or selective data report. Our data show a consistent and significant correlation between the size of a prognostic effect and its standard error, indicating that stronger effects are less reliable. The correlation even withstood the removal of all datasets with <10 patients. As these small samples yield particularly large SE, the variability of SE was severely restricted, which might be

First author(s), Year	N		Effect size [95% CI]
EEG dominant rhythm			
Alekseeva, 2010	64		1.83 [1.03, 2.62]
Bagnato, 2015	106		0.63 [0.23, 1.03]
Kotchoubey, 2005	43		1.32 [−0.31, 2.95]
RE Model for Subgroup (Q = 7.25, p = 0.03; I² = 69.0%) N = 213, z = 2.7, p = 0.0068			1.18 [0.33, 2.04]
EEG reactivity			
Bagnato, 2015	101		0.80 [0.28, 1.33]
Kang, 2014	56		0.99 [0.34, 1.63]
Cavinato, 2009	34		1.62 [0.37, 2.87]
Bagnato, 2017	28		1.37 [−0.67, 3.41]
Li, Kang, 2015	11		2.06 [1.01, 3.12]
RE Model for Subgroup (Q = 5.25, p = 0.26; I² = 33.3%) N = 230, z = 4.8, p = 1.9e−06			1.19 [0.70, 1.68]
Synek scale			
Estraneo, 2013	43		0.19 [−0.41, 0.79]
Cavinato, 2009	34		0.98 [0.21, 1.75]
Boccagni, 2011	13		1.38 [−0.43, 3.19]
RE Model for Subgroup (Q = 3.38, p = 0.18; I² = 44.2%) N = 90, z = 1.8, p = 0.068			0.64 [−0.05, 1.33]
fMRI BOLD response to stimulation			
Wang, 2015	39		1.78 [−0.12, 3.68]
Wang, 2015	25		−1.29 [−2.82, 0.25]
Marino, 2017	23		−0.31 [−1.23, 0.61]
Coleman, 2009	22		2.67 [0.85, 4.49]
Coleman, 2009	16		1.15 [−0.35, 2.65]
Di, 2007	11		−0.15 [−1.99, 1.69]
Li, Kang, 2015	11		1.74 [0.73, 2.74]
Edlow, 2017	10		0.67 [−1.62, 2.97]
RE Model for Subgroup (Q = 21.99, p = 0.003; I² = 67.5%) N = 157, z = 1.6, p = 0.11			0.74 [−0.18, 1.67]
MMN			
Kotchoubey, 2005	43		0.78 [0.14, 1.43]
Hildebrandt, 2007	20		1.21 [−0.18, 2.61]
Wang, 2017	11		−0.82 [−2.90, 1.27]
RE Model for Subgroup (Q = 2.60, p = 0.27; I² = 0.0%) N = 74, z = 2.6, p = 0.01			0.74 [0.18, 1.30]
N20			
Kang, 2014	56		0.89 [0.25, 1.52]
Estraneo, 2013	43		1.80 [0.99, 2.60]
Xu, 2012	36		0.35 [−0.19, 0.88]
Cavinato, 2009	34		0.17 [−0.75, 1.08]
Keren, 1994	10		1.93 [0.03, 3.82]
RE Model for Subgroup (Q = 11.92, p = 0.02; I² = 68.4%) N = 179, z = 2.7, p = 0.0072			0.89 [0.24, 1.54]
Sleep spindles			
Wislowska, 2017	18		1.47 [0.65, 2.29]
Kang, 2014	56		1.19 [0.52, 1.86]
Avantaggiato, 2015	10		1.21 [−0.59, 3.02]
RE Model for Subgroup (Q = 0.2729, p = 0.87; I² = 0.0%) N = 84, z = 5.1, p = 3.7e−07			1.30 [0.80, 1.79]
P300			
Kotchoubey, 2005	43		0.00 [−1.11, 1.11]
Cavinato, 2009	34		2.61 [0.92, 4.30]
Li, Song, 2015	22		2.83 [0.59, 5.07]
Hildebrandt, 2007	20		−0.43 [−1.66, 0.79]
Wang, 2017	11		0.11 [−1.42, 1.64]
RE Model for Subgroup (Q = 13.39, p = 0.01; I² = 73.5%) N = 130, z = 1.3, p = 0.2			0.85 [−0.44, 2.14]
RE Model for All Studies (Q = 74.69, p = 7e−05; I² = 58.3%) N = 1157, z = 6.9, p = 6.6e−12			0.93 [0.66, 1.19]

Effect size (Cohen's d)

−2   0   2   6

FIGURE 6 | The results of the meta-analysis for prediction of the outcome from neurophysiological variables. The criterion of improvement for UWS patients was the transition to MCS, and for MCS patients, the recovery of consciousness. The rest is the same as in **Figure 3**.

expected to reduce the correlation coefficient. This was, however, not the case (**Figure 2**).

Another important limitation of nearly all studies was the lack of outcome blindness. Only two groups of authors clearly indicated that the diagnosis of the outcome was performed by neurologists without the knowledge of predictor values. One might argue that the diagnosis of recovery of consciousness (based on the criteria of consistent communication and functional use of objects) is quite easy and can hardly be biased by neurophysiological data. Even if this argument is true for exit-MCS, it is obviously false for the other broadly used improvement criterion of UWS patients, namely the transition to MCS. As the differential diagnostics between UWS and MCS is notoriously difficult (81), any information about positive or negative neurophysiological findings could influence the diagnostic decision. If this influence really takes place, we can expect much stronger correlation of neurophysiological indices with the transition to MCS than with the transition to exit-MCS, because the latter diagnosis is easier and thus less affected by additional information.

Exactly this was true. When the improvement criterion for both diagnostic groups (UWS and MCS) was the transition to exit-MCS, the weighted average ES (in terms of Cohen's *d*) was rather moderate, in any case slightly smaller than 0.5. But the transition of UWS patients to MCS (which is quite difficult from the diagnostic point of view) was strongly related to the neurophysiological findings, with the weighted average ES being >1.4.

A potentially strong but still underestimated bias is related to unreported predictor variables. In most of the reviewed records, outcome prediction was not the main aim of the study, but rather a by-product of other analyses. Particularly, in such studies (though not only in these), many variables could be measured but not really reported. Sometimes, many variables are used in a UWS/MCS comparison but not even mentioned in relation to prognosis, although one may suppose that they were also compared with the follow-up data. Less frequently prognostic effects are referred to as "lacking" or "non-significant" without further quantification. Simmons et al. (65) suggest a very simple solution of the problem: whenever authors list their variables, they should add a short word "only." We tried to counteract this false positive effect by assigning the value of 0 to the effects of obviously omitted variables (with its SE being assumed as the median SE of the reported variables). However, this method is not only imprecise but can also be biased, first, because the real number of such omitted variables may be much larger than a reader can guess, and second, because negative effects (i.e., those which run against the starting hypothesis, such as better neurophysiological responses in non-recovered patients) can be omitted more frequently than positive effects.

With this in mind, we understand that the data of meta-analyses should be taken with great caution. Nevertheless, we believe that a glance on the meta-analytic results can be of interest. First, the analysis was strongly complicated by the high variability of the primary records. Very small number of studies using exactly the same predictor and the same improvement criterion enforced us to combine similar methods, which resulted in high heterogeneity indices such as I^2. Poor prognostic features of the characteristics of fMRI reactivity might partially

be attributed to this group of studies including fMRI responses to very different stimulations from pain to music. For the same reason, we excluded some possible predictors (e.g., responses to active behavioral instructions; ERP N400 component) that were employed in two records only.

Second, the empirical contribution of predictors does not necessarily follow their general theoretical value defined by basic neuroscience. This is quite demonstrative in the case of P300, one of the most useful and most widely employed indices in neuroscience whose effect in the prediction of the outcome turned out to be virtually 0. One might speculate that P300 is not immediately related to consciousness (37), but, rather, to a more specific function such as working memory. Another possible reason may be the extreme difficulty of the separation between different P300 subcomponents (P3a and P3b) in the target population (38). The subcomponents are usually distinguished by topography and responses to active instruction, but most DoC patients have changed ERP topography and do not respond to instruction. If the results of this preliminary analysis should be used to determine which lines of research should not be recommended for future studies, P300 is the first candidate for such a negative recommendation. Also, the importance of the (highly expensive) fMRI predictors might similarly be overestimated on the basis of their theoretical importance.

Oscillatory EEG responses to stimulation showed, to the contrary, most promising effects, which agree well with the results of Bender et al. (10) obtained on the basis of different data.

Although EEG reactivity was measured to very different stimuli [e.g., to passive (23) or forced eyes opening (19, 20), to pain (34, 41), to warm water (39), and no description was given by Sarà et al. (48)] and the definition of reactivity substantially varied, the results are very homogenous across the reports. Moreover, this was the only group of predictor variables whose predictive value was significant for both improvement criteria (transition from UWS to MCS and the recovery of consciousness). Publication bias was also presented in these data, but it was less strong than for many other predictors (see Tables S2 and S3 in Supplementary Material).

Other perspective variables are brain metabolism (estimated by means of PET or SPECT) and the presence of sleep spindles in the EEG. Recent sleep data indicate a vital importance of spindles in information processing during sleep, which affects numerous cognitive processes in the subsequent wakefulness (75, 82, 83). DoC belongs to rare medical conditions characterized by severe deficits, or even complete absence, of sleep spindles, also in patients with relatively preserved sleep structure (84, 85). We believe that the role of sleep spindles in the outcome prediction in DoC should be explored in future work.

Both MMN and SSEP are proven outcome predictors for acute coma (86, 87). However, their value for the chronic DoC remains unclear. The present findings indicate their prognostic effects for the transition from UWS to MCS but not for recovery of consciousness.

The data further show that the predictive value of the auxiliary (e.g., neurophysiological) variables should be compared with the values of clinical variables. In the currently reviewed data, MCS patients had about 4.5 times better chances (if we take the

lower limit of the 95% CI, three times better chances) to regain consciousness than UWS patients. It is true that we analyzed only records implementing the neurophysiological approach and missed similar data in the other publications not using neurophysiological variables. Nevertheless, the effect is very strong and, as far as we can judge, not strongly biased (as it was not a desired effect). This indicates the necessity to integrate neurophysiological and clinical predictors within a multivariate approach in further studies. This integration could be more productive in both diagnostic and prognostic respects than the attempts to oppose different classes of variables to each other.

These considerations can only be conceived of as preliminary. It should be stressed that all biasing factors discussed above act in the same direction, potentially increasing the number of *false positive* results. The general critique of Brok et al. (64) remains valid also for the current study: as long as the original studies do not present all information, meta-analyses can only try to diminish, but not abolish the positive bias. If we want to eliminate the bias, (i) small-size studies should be avoided (for prediction studies, groups of recovered/non-recovered should include at least 20 patients each); (ii) a flow chart should make evident the procedure of patient selection; (iii) neurologists assessing the target variable (i.e., change of the diagnosis) should be completely blinded regardless the values of neurophysiological predictors; (iv) the full list of measured variables including all potential predictors should be presented from the beginning of a report (in a Methods section); (v) the intervals (a) between the accident and the neurophysiological measurement, and (b) between this measurement and the follow-up assessment should be specified;

finally, (vi) all positive *and negative* (e.g., non-significant relationships) results should be described in the same quantitative manner, either including the size of all effects, or permitting to calculate this size (e.g., mean and SD for the recovered and non-recovered groups).

Therefore, the numbers presented in **Figures 5** and **6** and in the Supplementary Materials can now be regarded, not as estimates of real effects, but rather, as upper limits of these effects. The current state of affairs is yet far away from the level at which any practical recommendation can be given except the recommendation to be highly careful with interpretations.

AUTHOR CONTRIBUTIONS

Both authors performed data search and analyses. The manuscript was initiated by the second author and finalized by the first author.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at https://www.frontiersin.org/articles/10.3389/fneur.2018.00315/full#supplementary-material.

REFERENCES

1. Lipsey MW, Wilson DB. *Practical Meta-Analysis (Applied Social Research Methods Ser., V. 49)*. Thousand Oaks, CA: SAGE (2001).

2. Gurevitch J, Koricheva J, Nakagawa S, Stewart G. Meta-analysis and the science of research synthesis. *Nature* (2018) 555:175–82. doi:10.1038/nature25753

3. Giacino JT, Ashwal S, Childs N, Cranford R, Jennett B, Katz DI, et al. The minimally conscious state definition and diagnostic criteria. *Neurology* (2002) 58(3):349–53. doi:10.1212/WNL.58.3.349

4. Laureys S, Celesia GG, Cohadon F, Lavrijsen J, León-Carrión J, Sannita WG, et al. Unresponsive wakefulness syndrome: a new name for the vegetative state or apallic syndrome. *BMC Med* (2010) 8(1):68. doi:10.1186/1741-7015-8-68

5. Liberati G, Hünefeldt T, Belardinelli MO. Questioning the dichotomy between vegetative state and minimally conscious state: a review of the statistical evidence. *Front Hum Neurosci* (2014) 8:865. doi:10.3389/fnhum.2014.00865

6. Hannawi Y, Lindquist MA, Caffo BS, Sair HI, Stevens RD. Resting brain activity in disorders of consciousness. *Neurology* (2015) 84:1272–80. doi:10.1212/WNL.0000000000001404

7. Kondziella D, Friberg CK, Frokjaer VG, Fabricius M, Moeller K. Preserved consciousness in vegetative and minimal conscious states: systematic review and meta-analysis. *J Neurol Neurosurg Psychiatry* (2016) 87:485–92. doi:10.1136/jnnp-2015-310958

8. Rosenthal R. The "file drawer problem" and tolerance for null results. *Psychol Bull* (1979) 86:638–41. doi:10.1037/0033-2909.86.3.638

9. Rothstein H, Sutton AJ, Borenstein M. *Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments*. Hoboken, NJ: Wiley (2005).

10. Bender A, Jox RE, Grill E, Straube A, Lulé D. Persistent Vegetative State and Minimally Conscious State: a systematic review and meta-analysis of diag-

nostic procedures. *Deutsches Ärzteblatt International* (2015) 112(14):235–42. doi:10.3238/arztebl.2015.0235

11. Kotchoubey B. Evoked and event-related potentials in disorders of consciousness: a quantitative review. *Conscious Cogn* (2017) 54:155–67. doi:10.1016/j.concog.2017.05.002

12. Wilson JTL, Pettigrew LEL, Teasdale GM. Structured interview for the Glasgow Outcome Scale and the Extended Glasgow Outcome Scale: guidelines for their use. *J Neurotrauma* (1998) 15:573–85. doi:10.1089/neu.1998.15.573

13. Lenhard W, Lenhard A. *Calculation of Effect Sizes*. Dettelbach, Germany: Psychometrica (2016). Available from: https://www.psychometrica.de/effect_size.html (Accessed: February 28, 2018).

14. Nakagawa S, Cuthill IC. Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biol Rev Camb Philos Soc* (2007) 87:591–605. doi:10.1111/j.1469-185X.2007.00027.x

15. Alekseeva EV, Alasheev AM, Belkin AA, Kudrinskikh NV, Nikov PN. Prognostic evaluation of sleep in patients in a vegetative state. *Anesteziol Reanimatol* (2010) 55(4):38–42.

16. Arnaldi D, Terzaghi M, Cremascoli R, De Carli F, Maggioni G, Pistarini C, et al. The prognostic value of sleep patterns in disorders of consciousness in the sub-acute phase. *Clin Neurophysiol* (2016) 127(2):1445–51. doi:10.1016/j.clinph.2015.10.042

17. Avantaggiato P, Molteni E, Formica F, Gigli GL, Valente M, Lorenzut S, et al. Polysomnographic sleep patterns in children and adolescents in Unresponsive Wakefulness Syndrome. *J Head Trauma Rehabil* (2015) 30(5):334–46. doi:10.1097/HTR.0000000000000122

18. Babiloni C, Sarà M, Vecchio F, Pistoia F, Sebastiano F, Onorati P, et al. Cortical sources of resting-state alpha rhythms are abnormal in persistent vegetative state patients. *Neurophysiol Clin* (2009) 120(4):719–29. doi:10.1016/j.clinph.2009.02.157

19. Bagnato S, Boccagni C, Sant'Angelo A, Prestandrea C, Mazzilli R, Galardi G. EEG predictors of outcome in patients with disorders of consciousness

admitted for intensive rehabilitation. *Neurophysiol Clin* (2015) 126(5):959–66. doi:10.1016/j.clinph.2014.08.005

20. Bagnato S, Boccagni C, Prestandrea C, Fingelkurts AA, Fingelkurts AA, Galardi G. Changes in standard electroencephalograms parallel consciousness improvements in patients with Unresponsive Wakefulness Syndrome. *Arch Phys Med Rehabil* (2017) 98(4):665–72. doi:10.1016/j.apmr.2016.09.132

21. Boccagni C, Bagnato S, Sant'Angelo A, Prestandrea C, Galardi G. Usefulness of standard EEG in predicting the outcome of patients with disorders of consciousness after anoxic coma. *J Clin Neurophysiol* (2011) 28(5):489–92. doi:10.1097/WNP.0b013e318231c8c8

22. Castro M, Tillmann B, Luauté J, Corneyllie A, Dailler F, André-Obadia N, et al. Boosting cognition with music in patients with disorders of consciousness. *Neurorehabil Neural Repair* (2015) 29(8):734–42. doi:10.1177/1545968314565464

23. Cavinato M, Freo U, Ori C, Zorzi M, Tonin P, Piccione F, et al. Post-acute P300 predicts recovery of consciousness from traumatic vegetative state. *Brain Inj* (2009) 23(12):973–80. doi:10.3109/02699050903373493

24. Chennu S, Annen J, Wannez S, Thibaut A, Chatelle C, Cassol H, et al. Brain networks predict metabolism, diagnosis and prognosis at the bedside in disorders of consciousness. *Brain* (2017) 140(8):2120–213210. doi:10.1093/brain/awx163

25. Coleman MR, Rodd JM, Davis MH, Johnsrude IS, Menon DK, Pickard JD, et al. Do vegetative patients retain aspects of language comprehension? Evidence from fMRI. *Brain* (2007) 130(10):2494–507. doi:10.1093/brain/awm170

26. Coleman MR, Davis MH, Rodd JM, Robson T, Ali A, Owen AM, et al. Towards the routine use of brain imaging to aid the clinical diagnosis of disorders of consciousness. *Brain* (2009) 132(9):2541–52. doi:10.1093/brain/awp183

27. Di HB, Yu SM, Weng XC, Laureys S, Yu D, Li JQ, et al. Cerebral response to patient's own name in the vegetative and minimally conscious states. *Neurology* (2007) 68(12):895–9. doi:10.1212/01.wnl.0000258544.79024.d0

28. Edlow BL, Chatelle C, Spencer CA, Chu CJ, Bodien YG, O'Connor KL, et al. Early detection of consciousness in patients with acute severe traumatic brain injury. *Brain* (2017) 140(9):2399–414. doi:10.1093/brain/awx176

29. Estraneo A, Moretta P, Loreto V, Lanzillo B, Cozzolino A, Saltalamacchia A, et al. Predictors of recovery of responsiveness in prolonged anoxic vegetative state. *Neurology* (2013) 80(5):464–70. doi:10.1212/WNL.0b013e31827f0f31

30. Fingelkurts AA, Fingelkurts AA, Bagnato S, Boccagni C, Galardi G. Prognostic value of resting-state electroencephalography structure in disentangling vegetative and minimally conscious states: a preliminary study. *Neurorehabil Neural Repair* (2013) 27(4):345–54. doi:10.1177/1545968312469836

31. Glass I, Sazbon L, Groswasser Z. Mapping "cognitive" event-related potentials in prolonged postcoma unawareness state. *Clin Electroencephal* (1998) 29(1):19–30. doi:10.1177/155005949802900109

32. Hildebrandt H, Happe S, Deutschmann A, Basar-Eroglu C, Eling P, Brunhöber J. Brain perfusion and VEP reactivity in occipital and parietal areas are associated to recovery from hypoxic vegetative state. *J Neurol Sci* (2007) 260(1):150–8. doi:10.1016/j.jns.2007.04.035

33. Howell K, Grill E, Klein A-M, Straube A, Bender A. Rehabilitation outcome of anoxic-ischaemic encephalopathy survivors with prolonged disorders of consciousness. *Resuscitation* (2013) 84(10):1409–15. doi:10.1016/j.resuscitation.2013.05.015

34. Kang XG, Li L, Wei D, Xu XX, Zhao R, Jing YY, et al. Development of a simple score to predict outcome for unresponsive wakefulness syndrome. *Crit Care* (2014) 18:R37. doi:10.1186/cc13745

35. Keren O, Sazbon L, Groswasser Z, Shmuel M. Follow-up studies of somatosensory evoked potentials and auditory brainstem evoked potentials in patients with post-coma unawareness (PCU) of traumatic brain injury. *Brain Inj* (1994) 8(3):239–47. doi:10.3109/02699059409150976

36. Kondziella D, Fisher PM, Larsen VA, Hauerberg J, Fabricius M, Møller K, et al. Functional MRI for assessment of the default mode network in acute brain injury. *Neurocrit Care* (2017) 27(3):401–6. doi:10.1007/s12028-017-0407-6

37. Kotchoubey B. Event-related potential measures of consciousness: two equations with three unknowns. *Prog Brain Res* (2005) 150:427–44. doi:10.1016/S0079-6123(05)50030-X

38. Kotchoubey B. Apallic syndrome is not apallic – is vegetative state vegetative? *Neuropsychol Rehabil* (2005) 15:333–56. doi:10.1080/09602010443000416

39. Li L, Kang XG, Qi S, Xu XX, Xiong LZ, Zhao G, et al. Brain response to thermal stimulation predicts outcome of patients with chronic disorders of consciousness. *Neurophysiol Clin* (2015) 126(8):1539–47. doi:10.1016/j.clinph.2014.10.148

40. Li R, Song W, Du J, Huo S, Shan G. Connecting the P300 to the diagnosis and prognosis of unconscious patients. *Neural Regen Res* (2015) 10(3):473–80. doi:10.4103/1673-5374.153699

41. Logi F, Pasqualetti P, Tomaiuolo F. Predict recovery of consciousness in post-acute severe brain injury: the role of EEG reactivity. *Brain Inj* (2011) 25(10):972–9. doi:10.3109/02699052.2011.589795

42. Marino S, Bonanno L, Ciurleo R, Baglieri A, Morabito R, Guerrera S, et al. Functional evaluation of awareness in vegetative and minimally conscious state. *Open Neuroimag J* (2017) 11:17–25. doi:10.2174/1874440001711010017

43. Oksenberg A, Gordon C, Arons E, Sazbon L. Phasic activities of rapid eye movement sleep in vegetative state patients. *Sleep* (2001) 24(6):703. doi:10.1093/sleep/24.6.703

44. Qin P, Di H, Yan X, Yu S, Laureys S, Weng X. Mismatch negativity to the patient's own name in chronic disorders of consciousness. *Neurosci Lett* (2008) 448(1):24–8. doi:10.1016/j.neulet.2008.10.029

45. Qin P, Wu X, Duncan NW, Bao W, Tang W, Zhang Z, et al. GABAA receptor deficits predict recovery in patients with disorders of consciousness: a preliminary multimodal [11C]Flumazenil PET and fMRI study. *Hum Brain Mapp* (2015) 36(10):3867–77. doi:10.1002/hbm.22883

46. Qin P, Wu X, Huang Z, Duncan NW, Tang W, Wolff A, et al. How are different neural networks related to consciousness? *Ann Neurol* (2015) 78(4):594–605. doi:10.1002/ana.24479

47. Rosanova M, Gosseries O, Casarotto S, Boly M, Casali AG, Bruno MA, et al. Recovery of cortical effective connectivity and recovery of consciousness in vegetative patients. *Brain* (2012) 135(4):1308–20. doi:10.1093/brain/awr340

48. Sarà M, Pistoia F, Pasqualetti P, Sebastiano F, Onorati P, Rossini PM. Functional isolation within the cerebral cortex in the vegetative state: a nonlinear method to predict clinical outcomes. *Neurorehabil Neural Repair* (2011) 25(1):35–42. doi:10.1177/1545968310378508

49. Schorr B, Schlee W, Arndt M, Bender A. Coherence in resting-state EEG as a predictor for the recovery from unresponsive wakefulness syndrome. *J Neurol* (2016) 263(5):937–53. doi:10.1007/s00415-016-8084-5

50. Sitt JD, King JR, El Karoui I, Rohaut B, Faugeras F, Gramfort A, et al. Large scale screening of neural signatures of consciousness in patients in a vegetative or minimally conscious state. *Brain* (2014) 137(8):2258–70. doi:10.1093/brain/awu141

51. Stender J, Gosseries O, Bruno MA, Charland-Verville V, Vanhaudenhuyse A, Demertzi A, et al. Diagnostic precision of PET imaging and functional MRI in disorders of consciousness: a clinical validation study. *Lancet* (2014) 384(9942):514–22. doi:10.1016/S0140-6736(14)60042-8

52. Steppacher I, Eickhoff S, Jordanov T, Kaps M, Witzke W, Kissler J. N400 predicts recovery from disorders of consciousness. *Ann Neurol* (2013) 73(5):594–602. doi:10.1002/ana.23835

53. Tommasino C, Grana C, Lucignani G, Torri G, Fazio F. Regional cerebral metabolism of glucose in comatose and vegetative state patients. *J Neurosurg Anesthesiol* (1995) 7:109. doi:10.1097/00008506-199504000-00006

54. Vogel D, Markl A, Yu T, Kotchoubey B, Lang S, Müller F. Can mental imagery functional magnetic resonance imaging predict recovery in patients with disorders of consciousness? *Arch Phys Med Rehabil* (2013) 94(10):1891–8. doi:10.1016/j.apmr.2012.11.053

55. Wang F, Di H, Hu X, Jing S, Thibaut A, Di Perri C, et al. Cerebral response to subject's own name showed high prognostic value in traumatic vegetative state. *BMC Med* (2015) 13:83. doi:10.1186/s12916-015-0330-7

56. Wang XY, Wu HY, Lu HT, Huang TT, Zhang H, Zhang T. Assessment of mismatch negativity and P300 response in patients with disorders of consciousness. *Eur Rev Med Pharmacol Sci* (2017) 21(21):4896–906.

57. Wijnen VJM, van Boxtel GJM, Eilander HJ, de Gelder B. Mismatch negativity predicts recovery from the vegetative state. *Clin Neurophysiol* (2007) 118(3):597–605. doi:10.1016/j.clinph.2006.11.020

58. Wijnen VJM, Eilander HJ, de Gelder B, van Boxtel GJM. Repeated measurements of the auditory oddball paradigm is related to recovery from the vegetative state. *J Clin Neurophysiol* (2014) 31(1):65–80. doi:10.1097/01.wnp.0000436894.17749.0c

59. Wijnen VJM, Eilander HJ, de Gelder B, van Boxtel GJM. Visual processing during recovery from vegetative state to consciousness: comparing behavioral indices to brain responses. *Neurophysiol Clin* (2014) 44(5):457–69. doi:10.1016/j.neucli.2014.08.008

60. Wislowska M, Del Giudice R, Lechinger J, Wielek T, Heib DPJ, Pitiot A, et al. Night and day variations of sleep in patients with disorders of consciousness. *Sci Rep* (2017) 7:266. doi:10.1038/s41598-017-00323-4

61. Wu DY, Cai G, Yuan Y, Liu L, Li GQ, Song WQ, et al. Application of nonlinear dynamics analysis in assessing unconsciousness: a preliminary study. *Clin Neurophysiol* (2011) 122(3):490–8. doi:10.1016/j.clinph.2010.05.036

62. Xu W, Jiang G, Chen Y, Wang X, Jiang X. Prediction of minimally conscious state with somatosensory evoked potentials in long-term unconscious patients after traumatic brain injury. *J Trauma Acute Care Surg* (2012) 72(4):1024–30. doi:10.1097/TA.0b013e31824475cc

63. Ioannidis JPA. Why most published research findings are false. *PLoS Med* (2005) 2(8):e124. doi:10.1371/journal.pmed.0020124

64. Brok J, Thorlund K, Gluud C, Wetterslev J. Trial sequential analysis reveals insufficient information size and potentially false positive results in many meta-analyses. *J Clin Epidemiol* (2008) 61(8):763–9. doi:10.1016/j.jclinepi.2007.10.007

65. Simmons JP, Nelson LD, Simonsohn U. False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychol Sci* (2011) 22(11):1359–66. doi:10.1177/0956797611417632

66. Begley CG, Ioannidis JP. Reproducibility in science: improving the standard for basic and preclinical research. *Circ Res* (2015) 116:116–26. doi:10.1161/CIRCRESAHA.114.303819

67. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* (2011) 155:529–36. doi:10.7326/0003-4819-155-8-201110180-00009

68. Giacino JT, Kalmar K, White J. The JFK Coma Recovery Scale-Revised: measurement characteristics and diagnostic utility. *Arch Phys Med Rehabil* (2004) 85:2020–9. doi:10.1016/j.apmr.2004.02.033

69. American Congress of Rehabilitation Medicine, Brain Injury-Interdisciplinary Special Interest Group, Disorders of Consciousness Task Force, Seel RT, Sherer M, Whyte J, Katz DI, Giacino JT, et al. Assessment scales for disorders of consciousness: evidence-based recommendations for clinical practice and research. *Arch Phys Med Rehabil* (2010) 91:1795–813. doi:10.1016/j.apmr.2010.07.218

70. Yue W, Wilson CS, Boller E. Peer assessment of journal quality in clinical neurology. *J Med Libr Assoc* (2007) 95:70–6.

71. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* (2002) 21(11):1539–58. doi:10.1002/sim.1186

72. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Softw* (2010) 36(3):1–48. doi:10.18637/jss.v036.i03

73. Brembs B, Button K, Munafò M. Deep impact: unintended consequences of journal rank. *Front Hum Neurosci* (2013) 7:291. doi:10.3389/fnhum.2013.00291

74. Synek VM. Value of a revised EEG Coma Scale for prognosis after cerebral anoxia and diffuse head injury. *Clin EEG Neurosci* (1990) 21(1):25–30.

75. Rasch B, Born J. About sleep's role in memory. *Physiol Rev* (2013) 93:681–766. doi:10.1152/physrev.00032.2012

76. Fritz CO, Morris PE, Richler JJ. Effect size estimates: current use, calculations, and interpretation. *J Exp Psychol* (2012) 141(1):2–18. doi:10.1037/a0024338

77. Sedlmeier P, Gigerenzer G. Do studies of statistical power have an effect on the power of studies? *Psychol Bull* (1989) 105(2):309–16. doi:10.1037/0033-2909.105.2.309

78. Ingre M. Why small low-powered studies are worse than large high-powered studies and how to protect against "trivial" findings in research: comment on Friston (2012). *Neuroimage* (2013) 81:496–8. doi:10.1016/j.neuroimage.2013.03.030

79. Vul E, Harris C, Winkielman P, Paschler H. Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. *Perspect Psychol Sci* (2009) 4:274–90. doi:10.1111/j.1745-6924.2009.01132.x

80. Button KS, Ioannidis JP, Mokrysz C, Nosek BA, Flint J, Robinson ES, et al. Power failure: why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci* (2013) 14:365–76. doi:10.1038/nrn3475

81. Schnakers C, Vanhaudenhuyse A, Giacino J, Ventura M, Boly M, Majerus S, et al. Diagnostic accuracy of the vegetative and minimally conscious state: clinical consensus versus standardized neurobehavioral assessment. *BMC Neurol* (2009) 9:Article35. doi:10.1186/1471-2377-9-35

82. Gais S, Molle M, Helms K, Born J. Learning-dependent increases in sleep spindle density. *J Neurosci* (2002) 22:6830–4. doi:10.1523/JNEUROSCI.22-15-06830.2002

83. Tamminen J, Ralph MAL, Lewis PA. The role of sleep spindles and slow-wave activity in integrating new information in semantic memory. *J Neurosci* (2013) 33:15376–81. doi:10.1523/JNEUROSCI.5093-12.2013

84. Pavlov YG, Gais S, Müller F, Schönauer M, Schäpers B, Born J, et al. Night sleep in patients with vegetative state. *J Sleep Res* (2017) 26:629–40. doi:10.1111/jsr.12524

85. Sebastiano RD, Visani E, Panzica F, Sattin D, Bersano A, Nigri A, et al. Sleep patterns associated with the severity of impairment in a large cohort of patients with chronic disorders of consciousness. *Clin Neurophysiol* (2018) 129(3):687–93. doi:10.1016/j.clinph.2017.12.012

86. Daltrozzo J, Wioland N, Mutschler V, Kotchoubey B. Predicting outcome of coma using event-related brain potentials: a meta-analytic approach. *Neurophysiol Clin* (2007) 118:606–14. doi:10.1016/j.clinph.2006.11.019

87. Zandbergen EGJ, Koelman JHTM, de Haan RJ, Hijdra A. SSEPs and prognosis in postanoxic coma: only short or also long latency responses? *Neurology* (2006) 67:583–6. doi:10.1212/01.wnl.0000230162.35249.7f

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer AP and handling Editor declared their shared affiliation.