Developing Brain Vital Signs: Initial Framework for Monitoring Brain Function Changes Over Time

Sujoy Ghosh Hajra [1,2], Careesa C. Liu [1,2], Xiaowei Song [1,2,3], Shaun Fickling [1,2], Luke E. Liu [2], Gabriela Pawlowski [2,4], Janelle K. Jorgensen [5], Aynsley M. Smith [5], Michal Schnaider-Beeri [6,7], Rudi Van Den Broek [8], Rowena Rizzotti [8], Kirk Fisher [8] and Ryan C. N. D'Arcy [1,2,3,4,8]*

[1] Faculty of Applied Science, School of Engineering Science, Simon Fraser University, Burnaby, BC, Canada, [2] NeuroTech Lab, Simon Fraser University and Fraser Health Authority, Surrey, BC, Canada, [3] Health Sciences and Innovation, Surrey Memorial Hospital, Fraser Health Authority, Surrey, BC, Canada, [4] Biomedical Physiology and Kinesiology, Faculty of Science, Simon Fraser University, Burnaby, BC, Canada, [5] Sports Medicine Center, Mayo Clinic, Rochester, MN, USA, [6] Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY, USA, [7] Joseph Sagol Neuroscience Centre, Sheeba Medical Centre, Ramat Gan, Israel, [8] HealthTech Connex Inc., Surrey, BC, Canada

Clinical assessment of brain function relies heavily on indirect behavior-based tests. Unfortunately, behavior-based assessments are subjective and therefore susceptible to several confounding factors. Event-related brain potentials (ERPs), derived from electroencephalography (EEG), are often used to provide objective, physiological measures of brain function. Historically, ERPs have been characterized extensively within research settings, with limited but growing clinical applications. Over the past 20 years, we have developed clinical ERP applications for the evaluation of functional status following serious injury and/or disease. This work has identified an important gap: the need for a clinically accessible framework to evaluate ERP measures. Crucially, this enables baseline measures before brain dysfunction occurs, and might enable the routine collection of brain function metrics in the future much like blood pressure measures today. Here, we propose such a framework for extracting specific ERPs as potential "brain vital signs." This framework enabled the translation/transformation of complex ERP data into accessible metrics of brain function for wider clinical utilization. To formalize the framework, three essential ERPs were selected as initial indicators: (1) the auditory N100 (Auditory sensation); (2) the auditory oddball P300 (Basic attention); and (3) the auditory speech processing N400 (Cognitive processing). First step validation was conducted on healthy younger and older adults (age range: 22–82 years). Results confirmed specific ERPs at the individual level (86.81–98.96%), verified predictable age-related differences (P300 latency delays in older adults, $p < 0.05$), and demonstrated successful linear transformation into the proposed brain vital sign (BVS) framework (basic attention latency sub-component of BVS framework reflects delays in older adults, $p < 0.05$). The findings represent an initial critical step in developing, extracting, and characterizing ERPs as vital signs, critical for subsequent evaluation of dysfunction in conditions like concussion and/or dementia.

Keywords: clinical neuroscience, ERPs, evoked potentials, vital signs, neurology

INTRODUCTION

Vital signs such as heart rate, pulse oxygenation, and blood pressure are essential to monitoring and managing the health of various body systems. Yet there are no such vital signs identified for brain function—despite the clearly instrumental role such vital signs could play. Current clinical assessments for screening brain functional status relies largely on subjective, behavior-based measures, such as the Glasgow Coma Scale (GCS), to evaluate level of conscious awareness following brain injury (Teasdale and Jennett, 1974, pp. 81–84; Reith et al., 2016, pp. 89–94). However, subjective behavior-based tests of this nature have been reported to have misdiagnosis rates as high as 43% (Gawryluk et al., 2010, p. 11; Schnakers et al., 2009, p. 35). More detailed clinical evaluation of cognitive function and associated impairments are often reliant on neuropsychological assessment (Folstein et al., 1975, pp. 189–198; Lezak, 2004). These too are behavior-based measures, depending heavily on the patient's capacity to produce voluntary, on-demand motor and/or verbal responses to stimuli (Connolly and D'Arcy, 2000, pp. 31–47). Unfortunately, confounding factors, such as motoric and communicative limitations, often hamper greatly the clinical effectiveness for many of these measures.

Over the last 20 years, our group has demonstrated the critical need for a physiological, objective brain function assessment that utilizes event-related potentials or ERPs (Connolly et al., 1995, pp. 548–565; D'Arcy et al., 2003, pp. 662–672; Sculthorpe-Petley et al., 2015, pp. 64–72). ERPs are derived from long-standing electroencephalography (EEG; Pravdich-Neminsky, 1913, pp. 951–960). They can be recorded using minimal non-invasive scalp electrodes, combined with time-locked stimulation, to reflect target brain responses during information processing (Gawryluk and D'Arcy, 2010; Gawryluk et al., 2010, p. 11). EEG combines practical features of being accessible, available, low cost, and portable (Giacino et al., 2014, pp. 99–114), which makes the technology well suited for point-of-care applications (Gawryluk et al., 2010, p. 11). Work to date has demonstrated that ERPs can provide specific information across a spectrum of brain functioning, from low-level sensory to higher level cognitive processing (Luck, 2014). Moreover, ERPs have been shown to have robust diagnostic and prognostic capabilities (Morlet et al., 2000, pp. 198–206; D'Arcy et al., 2003, pp. 662–672; Kotchoubey et al., 2005, pp. 2441–2453; Wijnen et al., 2007, pp. 597–605; Vanhaudenhuyse et al., 2008, pp. 262–270; Daltrozzo et al., 2009, pp. 53–62).

In recent years, clinical ERP integration has focused on developing rapid, automated approaches in order to successfully utilize key ERPs that can be robustly recorded at the individual-level. The initial effort focused on developing a rapid evaluation framework for neurological status after severe acquired brain injuries, called the Halifax Consciousness Scanner (HCS; D'Arcy et al., 2011, pp. 750–754). The HCS was developed to examine the presence or absence of five key ERP responses linked to sensation (N100), perception (mismatch negativity, MMN), attention (P300), memory for one's own name (early negative enhancement, ENE), and semantic speech processing (N400). These ERPs were validated across a large sample of healthy individuals and clinically applied in neurological status assessment (Sculthorpe-Petley et al., 2015, pp. 64–72; Fleck-Prediger et al., 2014). However, it has become increasingly clear that mere assessments of presence or absence of a particular ERP does not fundamentally address the need to measure healthy individual brain function over time. Effective longitudinal monitoring of brain functional changes requires the establishment of individual functional "baselines" of brain vitality prior to conditions of dysfunction. Specifically, there is no framework for establishing and monitoring well established ERPs that can serve as indicators for an individual's healthy brain function, in spite of the evidence for relatively stable within-subject variance over time (Williams et al., 2005, pp. 1605–1630; Cassidy et al., 2012, pp. 659–664). This gap is essential to address in order to successfully assess the significance of any ERP-related change in which questions arise about possible dysfunction, which are increasingly arising as a potential application and common challenge in the evaluation of, for example, concussion and/or dementia.

Accordingly, the objective of the current study was to begin developing an initial framework to translate/transform ERPs into practical and accessible brain vital signs. The conceptual development of such a framework required a systematic process anchored to other existing vital sign frameworks. Therefore, the current paper is divided into two main sections: (1) a proposed conceptual framework for brain vital signs; and (2) a first step evaluation of practical implementation in a test sample of healthy adults across the lifespan.

I. Brain vital sign framework: As with other vital signs, a potential brain vital sign framework must satisfy some fundamental requirements: (1) the responses should be EEG hardware platform independent; (2) each response should be extensively characterized within the literature; (3) responses should be recorded reliably within healthy adult individuals[1]; (4) they should be accessible to normative data comparisons for essential response characteristics (e.g., amplitudes and latencies); and (5) importantly, the responses should be translated into a clinically accessible framework, which can be readily communicated.

To start, we selected well-established sensation-to-cognition ERPs[2] : the N100, P300, and N400[3] . A complex-to-basic pyramidal approach provided an overview of the translation from technical ERP nomenclature to easy to communicate brain vital signs (**Figure 1**). Sub-scores reflective of specific brain functions were derived from the mean and standard deviations (**Figure 2**). Lastly, linearly transformed scores normalized to the best possible results for each amplitude and latency measure were created and referred to as elemental brain scores (EBS).

---

[1] Most ERP responses are also well characterized across brain development, but require increasing reliance on age specific norms and caveats that are not discussed here.

[2] While the main focus is on auditory ERP responses, all three are present across both auditory and visual modalities.

[3] Note that the mismatch negativity (MMN; Duncan et al., 2009) may also be considered a potential ERP response for a brain vital sign framework. As the MMN can also be derived from the same auditory stimuli that elicit the N100/P300, it remains possible candidate for future expansion.
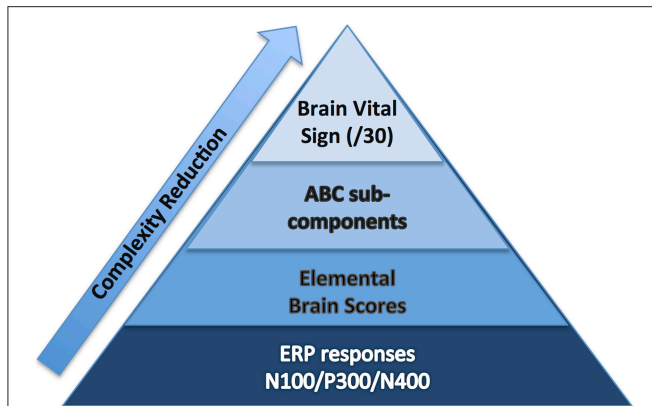
FIGURE 1 | Brain vital sign framework: (1) overall brain vital sign score: highest 30; (2) ABC break down into Auditory sensation, Basic attention, and Cognitive processing; and (3) Elemental Brain Scores linearly transformed from N100, P300, and N400 response amplitudes and latencies (3 responses[*]2 measures = 6 scores).
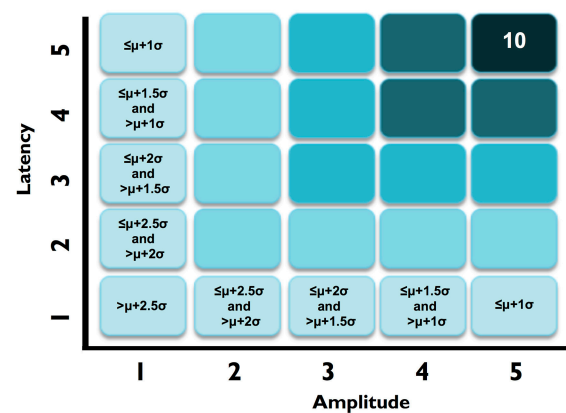

FIGURE 2 | ABC breakdown demonstrating graded measures. Calculation shown for BVS sub-components "A". Similar calculations undertaken for "B" and "C".

II. Practical implementation in healthy adults: To address implementation, the following steps were undertaken: (1) Hardware performance characterization was critical for platform independence. While some studies have compared hardware system performance for one time point analysis (Ries et al., 2014, pp. 10–20), analysis of performance over time was conducted to characterize instrument noise levels for longitudinal monitoring. (2) Stimulus sequence optimization was crucial to balance the trade-off between short testing times and highest possible response signal-to-noise ratio (SNR). (3) Response extraction and identification required an expert-independent, quality-checked approach. (4) Response results then required translation/transformation into the brain vital sign framework for interpretation and reporting. (5) A test sample of ERP data, with an age-related difference comparison embedded, provided initial validation across the adult life span (i.e., younger vs. older adults).

The current study utilized a healthy sample data set to test three hypotheses: (1) The three ERPs would be detectable at the individual level (Hypothesis 1); (2) Comparison between younger and older adults will show predictable age-related changes (Hypothesis 2); and (3) Because the brain vital sign framework is generated by applying a linear transformation to the raw ERP responses, we anticipate the pattern of age-related ERP changes would be preserved in the EBS results. With this initial step, it would then be possible to expand the brain vital sign framework into more extensive normative development along with applications in possible dysfunction related to conditions like concussion and dementia.

METHODS

Characterizing and Calibration of EEG Hardware Performance

Four candidate EEG systems (gNautilus and gMobiLab systems manufactured by g.Tec Medical Engineering, and two Enobio systems manufactured by Neuroelectrics) were evaluated in order to identify the most reliable hardware. Hardware evaluation used a 5-min known input calibration signal ("ground truth"), derived from a combination of sinusoidal waves with frequencies of 5, 10, 15, and 30 Hz (in MATLAB software). The test signal was delivered through the audio output port (at maximum volume setting) and recorded on 2 channels of the EEG systems as well as a Tektronix oscilloscope (model # 795-TBS1052B). Testing was conducted 2 times per day over 3 consecutive days (6 total).

Stability and reliability was assessed using inter-channel stability (correlation between channels at each time point of test), day-over-day stability (percentage change in peak voltage over the 3 days), peak-to-peak voltage recorded, and SNR (defined as ratio of sum of spectral power surrounding the 5, 10, 15, 30 Hz ["signal"] and 60 Hz ["noise"]). The g.Nautilus device, provided maximal SNR and had the best recording stability over the 3 days (average change 1.45% over 3 days, see Supplementary Material for full results), and was therefore utilized for subsequent human data collection.

Scalp-recorded ERPs were recorded from 3 midline electrodes (Fz, Cz, & Pz, embedded within a cap), insert earphones, g.Nautilus EEG acquisition hardware (bandpass: dc-250 Hz, 500 Hz sampling, 3 axis head motion accelerometers), and a portable computing platform. Four additional electrodes provide ground (forehead), reference (ear lobe) and eye monitoring (electro-oculogram, EOG).

Following signal amplification, conditioning, and digitization, the data were transmitted over Bluetooth link to the portable host computer. Time-stamping signals were sent from the host computer using custom-designed USB-to-TTL converter subsystem to mark stimulus presentation events. These TTL pulses were logged by the amplifier along with the EEG data and later used for signal averaging to derive ERPs.
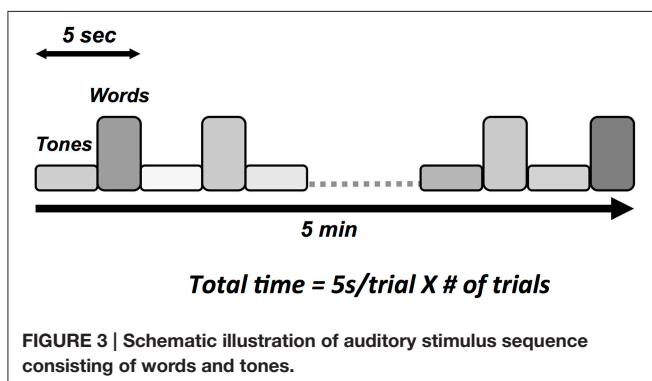
Stimulus Sequence Balancing SNR and Short Testing Time

Auditory tone and spoken word pair stimuli were presented through the insert earphones. Tone stimuli elicited the N100 and P300 responses and spoken word pairs elicited the N400 (**Figure 3**). Tones (100 ms duration) were divided into standard (75 dB, 80%) and deviant (100 dB, 20%) conditions, with the N100 and the P300 derived from the deviant condition. Paired spoken words were divided into congruent prime pairs (e.g., bread-butter, 50%) and incongruent prime pairs (Romeo-table, 50%). The N400 was derived from the incongruent prime word pairs. The interlacing of tones and word pair stimuli enabled full optimization of near maximum trials per unit time (e.g., 5 s/stimulus cycle × 60 cycles = 5 min).

ERP Response Elicitation, Extraction and Identification

EEG scans were conducted with minimal preparation compared to conventional EEG techniques (<5 min setup). Each participant was fitted with an elastic cap with embedded electrodes, and g.GAMMAsys electrode gel was injected at each location for conductivity. EOG channels were recorded using disposable Ag/AgCl electrodes on the supra-orbital ridge and outer canthus of the left eye. Skin-electrode impedances were maintained at <30kΩ impedance at each site. Acoustic stimuli were delivered binaurally through Etymotic ER4 insert earphones. Participants were instructed to pay attention to the auditory stimuli while maintaining visual fixation on a cross located 2.0 m away (black on white background). Three runs of the 5-min stimulus sequence were collected on each participant (approximately 15 min total run time).

Automated ERP pre-processing used established methods, including spectral filtering, segmentation, baseline correction, and conditional averaging (Luck, 2014). Signal-to-noise optimizations include ocular correction to remove eye artifact, jittered stimulus timing to minimize potential alpha contamination, and artifact de-noising using pattern recognition. ERP processing parameters were as follows: 1–20 Hz bandpass filter, 60 Hz notch filter, −100–900 ms epoch length for segmentation relative to stimulus onset. ERP response identification was undertaken through a template matching process in which N100, P300, and N400 peaks were identified

FIGURE 3 | Schematic illustration of auditory stimulus sequence consisting of words and tones.

by specifying expected polarity within expected temporal ranges (Marchand et al., 2002, pp. 1715–1722). Each ERP response value was measured as peak-to-peak measure relative to the preceding peak of opposite polarity.

Machine learning methods such as support vector machine (SVM), allow training of two-category classifiers to distinguish contrasting experimental conditions (see Parvar et al., 2014, pp. 1–12; Sculthorpe-Petley et al., 2015, pp. 64–72). The best results were obtained using single run, trial-averaged data from all three-electrode sites as inputs to the SVM with a radial kernel. Ninety Percent of the available data were randomly selected to train a two-category classifier to distinguish between the two stimulus conditions for each experiment (i.e., standard vs. deviant tones for N100 & P300, congruent vs. incongruent words for N400). The trained classifier was then applied to the remaining 10% of datasets to test the accuracy of classification. Under 10-fold cross-validation, this process is repeated 10 times such that the classifier is trained and tested on all available data. The total instances of correct group classification (ex. number of correct standard and deviant classifications) relative to the total number of classifications provide an accuracy number. Standard statistical measures including true positive (TP), false positive (FP), sensitivity, and specificity are derived from the confusion matrix. The SVM analyses were further verified using non-parametric permutation statistics to assess if the observed performance could be obtained by chance ((Golland and Fischl, 2003), pp. 330–341). This involved randomly redistributing the class labels in the training sets and observing the performance of the new SVM solution. After 1000 permutations, the observed classification accuracies were used as a null distribution against which the significance of the true SVM solution was determined.

Translation/Transformation of ERP Responses to Brain Vital Sign Framework

First, similar to neuropsychology assessment, a total brain vital sign score of 30 was defined to represent the most basic result: all responses fall with the normative range, bounded by standard deviation. The highest level of brain vital sign framework combines all three ERP peak amplitude (μV) and latency (ms) measures, ranked in terms of standard deviations from the mean (M/SDs), into one composite score of 30. The total brain vital sign score of 30 reflects overall healthy brain processing[4]. The total scores for each participant were generated by comparing the amplitude (X) and latency (L) measures of each of the 3 components to the normative database, with scoring criteria determined using the mean (μ) and standard deviation (σ) of the corresponding measures in the normative database. Details are shown in **Table 1**.

A standard clinical scheme of "ABC" was implemented for the breakdown of individual responses, with the N100 as an indicator for <u>A</u>uditory sensation (A) (Davis, 1939, pp. 494–499); the P300 as an indicator for Basic attention (B) (Sutton et al., 1967, pp. 1436–1439); and the N400 as an indicator for Cognitive processing during speech perception (C) (Kutas and

---

[4]Note that the initial SD cut-off can be customized and optimized in accordance with the enhanced development of normative databases.

TABLE 1 | BVS scoring criteria for the three ERP components.

P300		N400 & N100	
Amplitude (X)/ Latency (L)	BVS Score	Amplitude (X)/ Latency (L)	BVS Score
$X > \mu - \sigma$ $L < \mu + \sigma$	5	$X < \mu + \sigma$ $L < \mu + \sigma$	5
$\mu - 1.5\sigma < X < \mu - \sigma$ $\mu + 1.5\sigma > L > \mu + \sigma$	4	$\mu + 1.5\sigma > X > \mu + \sigma$ $\mu + 1.5\sigma > L > \mu + \sigma$	4
$\mu - 2\sigma < X < \mu - 1.5\sigma$ $\mu + 2\sigma > L > \mu + 1.5\sigma$	3	$\mu + 2\sigma > X > \mu + 1.5\sigma$ $\mu + 2\sigma > L > \mu + 1.5\sigma$	3
$\mu - 2.5\sigma < X < \mu - 2\sigma$ $\mu + 2.5\sigma > L > \mu + 2\sigma$	2	$\mu + 2.5\sigma > X > \mu + 2\sigma$ $\mu + 2.5\sigma > L > \mu + 2\sigma$	2
$X < \mu - 2.5\sigma$ $L > \mu + 2.5\sigma$	1	$X > \mu + 2.5\sigma$ $L > \mu + 2.5\sigma$	1

Hillyard, 1980, pp. 203–205). Within the ABC scheme 5 points for amplitude and latency each were awarded. In addition to establishing a healthy brain vital sign range ($A = 10$, $B = 10$, $C = 10$; Total 30), it was also possible to derive metrics for monitoring ABC amplitude and latency changes over time. Amplitude and latency metrics for ABC were used to calculate 6 elemental brain scores (EBS). Each EBS was normalized to the best possible response measurement. Therefore, for each EBS, it is possible to rank ABC amplitude and latency results relative to the largest normative ERP response amplitude and shortest normative ERP response latency, with scores ranging from 0 to 1. A score of 1 matched the outer bounds for best possible measurement.

Mathematically, EBS measures can be expressed as shown in equations 1 and 2 below:

$$Score = 1 - abs\left(\frac{M - best}{max - min}\right) \qquad (1)$$

$$Score = 1 - abs\left(\frac{best - M}{max - min}\right) \qquad (2)$$

Where, *M* is the mean value of the amplitude/latency, *max* is the maximum value and *min* is the minimum value and *best* is the "ideal" value that should be achieved. Best value can either be the max or the min value depending on whether the lowest or the highest value represents the ideal situation—generally for latency the lowest (smallest) value represents faster (better) processing, whereas for amplitude the highest positive value or lowest negative value is thought to represent ideal processing.

Equation (1) was utilized for N100 and N400 amplitude and latency as well as P300 latency, whereas Equation (2) was used for P300 amplitude. All EBS calculations were undertaken using an existing database of 100 healthy controls (Sculthorpe-Petley et al., 2015, pp. 64–72) containing information about N100, P300 and N400 components. To account for outliers in the normative database, all data values were ranked, and the interquartile range between 75th and 25th percentiles calculated. Extremity thresholds were determined by calculating the points

corresponding to 1.5-times the interquartile range above the 75th percentile, and 1.5-times the interquartile range below the 25th percentile. Data points beyond the boundary formed by these thresholds were excluded as outliers from the normative database prior to BVS and EBS extraction. This process resulted in the removal of 6 participants (out of 100) from the N100 amplitude database, 2 participants from the N100 latency database, 1 participant from the N400 amplitude, 6 participants from the P300 amplitude database and 1 participant from the P300 latency database. No participants were excluded from the N400 latency database.

Initial Validation across the Healthy Adult Lifespan

Sixteen (16) participants ranging in age from 22 to 82 years were recruited (46.81 ± 22.14, 8 females). A bimodally distributed sample was selected across lifespan, with 8 in the 20–35 year-old range (26.13 ± 4.00, 4 females) and 8 in the 50–85 year-old range (67.5 ± 11.25, 4 females). Participants had no history of neurological problems or psychoactive medications. All individuals were fluent in English and had normal or corrected-to-normal hearing. EEG noise compromised data from 2 participants. Furthermore, to ensure analyses used a minimum of a 20-year age separation with controlled age-appropriate cognitively matched samples, data from 2 participants were not included (1 from each sub-sample). The final matched analysis therefore included 6 in the 20–30 year range and 6 in the 50–85 year range. Research Ethics Boards at Simon Fraser University and Fraser Health Authority approved the study.

Each participant underwent neuropsychological screening along with EEG/ERP testing using Mini Mental State Exam (MMSE) (Folstein et al., 1975, pp. 189–198) and Montreal Cognitive Assessment (MoCA) (Nasreddine et al., 2005, pp. 695–699). MMSE examines 5 areas of cognition (orientation, registration, attention and calculation, recall, and language), with scores below 23 suggestive of cognitive impairment (maximum score 30; Folstein et al., 1975, pp. 189–198). MoCA examines a multitude of high-level cognitive functions (e.g., short-term memory recall, delayed recall, visuospatial abilities, working memory, and language etc.).

ERP results were divided into the two groups (20–30 and 50–85 age ranges). Quantitative group-level ERP response characteristics were compared using two-tailed independent samples *t*-test. Results are presented as mean ± SD. Moreover, to assess the performance of the expert-independent method (SVM) across the age ranges, a sub-analysis was undertaken for each group to compare and contrast any performance differences.

ERP responses were also transformed into the brain vital sign framework, generating an overall brain vital sign score and the 6 EBS scores each participant. EBS scores were compared at the group-level. Normality was assessed using the Shapiro-Wilk W test. Only the EBS measures for amplitude in the "C" component of the framework did not pass the normality test, and they were therefore compared using the Wilcoxon test. All others were compared using two-tailed independent-samples *t*-test. Results are presented as mean ± SD.

RESULTS

Participant Cognitive Status Evaluation

Participant characteristics and MMSE/MoCA scores are presented in **Table 2**. Both the younger (age 20–30) and older (age 50–85) groups scored in the healthy range. All individuals in the younger group obtained full scores (30/30) for both MMSE and MoCA. Participants in the older group scored 30 for MMSE and 29.3 ± 0.5 for MoCA.

ERP Response Extraction and Expert-Independent Identification

Figure 4 shows the ERP components evoked using the stimulus sequence in representative individuals for the 20–30 and 50–85 age ranges. The N100 component was elicited during Auditory sensation ($-6.74 \pm 2.13\,\mu V$). The P300 was elicited during Basic attention to deviant tones ($10.72 \pm 2.66\,\mu V$). The N400 was elicited during Cognitive processing of semantically incongruent word pairs ($-5.09 \pm 2.67\,\mu V$). All components were present in all participants. **Figure 5** presents group-averaged ERP results.

ERP response results indicate that the trained SVM classifier successfully identified predicted response differences. For the P300, the SVM classification included deviant vs. standard tones. The individual-level accuracy of P300 classification is 98.96% across all ages, with 0.98 sensitivity and 1.00 specificity (**Table 3**). For the N400, the SVM classification included incongruent vs. congruent word pairs. The individual-level classification accuracy for N400 is 86.81% across all ages, with 0.84 sensitivity and 0.90 specificity (**Table 3**). Permutation analysis verified the accuracy of the SVM classification for P300 ($p < 0.001$) and N400 ($p = 0.05$).

Translation to the Brain Vital Sign Framework

The participant responses were successfully translated into the brain vital sign framework. Relative to norms, the representative participants in **Figure 5** both scored full 30, allocated the maximum 10 for each of the A, B, and C components. All individuals achieved scores of 30. EBS scores were also calculated: (A) amplitude (0.56 ± 0.17) and latency (0.41 ± 0.25) for the N100 using $N = 99$ norms; (B) amplitude (0.59 ± 0.11) and latency (0.59 ± 0.14) for the P300 using $N = 100$ norms; and (C) amplitude (0.50 ± 0.24) and latency (0.36 ± 0.16) for the N400 using $N = 100$ norms. As anticipated, each of the EBS measures straddled the 50th percentile mark (=0.5, representing average performance) within the mean ± 1 standard deviation segment.

Initial Validation Across the Healthy Adult Lifespan

Table 3 presents quantitative group-level component ERP response characteristics. **Table 4** shows that P300 latencies increased significantly between younger and older groups ($p < 0.05$), with a similar trend for the N400 latencies ($p = 0.07$). **Table 5** shows that SVM classification undertaken separately for the younger and older age groups, showed comparable results between the two groups. **Figure 6** and **Table 6** show group-level EBS scores between younger and older groups. Similar to P300 latency measures, the corresponding EBS score ("B" latency) demonstrated a significant group difference ($p < 0.05$) between old and young. Additionally, the "C" latency EBS also showed a trend ($p = 0.07$) for group differences, again in agreement with the corresponding N400 latency trends.

DISCUSSION

The current study had two objectives: (1) to describe a conceptual framework for brain vital signs, which can provide an objective physiological evaluation of healthy baseline brain function; and (2) to conduct an initial practical evaluation in a test sample of healthy adults across the lifespan. The results demonstrated

TABLE 2 | Sample characteristics and cognitive test scores.

	Age 20–30	Age 50–85
Sample Size (n)	6	6
Education (years)	18.3 ± 1.9	17.8 ± 5.6
MMSE (/30)	30	30
MoCA (/30)	30	29.3 ± 0.5
Sex (M:F)	1:1	1:2

TABLE 3 | SVM classification for P300 and N400.

	P300	N400
Accuracy	98.96%	86.81%
True positive	0.98	0.84
False positive	0.00	0.10
Sensitivity	0.98	0.84
Specificity	1.00	0.90

TABLE 4 | Quantitative measures for group-level ERP characteristics.

		Age 20–30	Age 50–85
P300	Amplitude (μV)	11.09 ± 3.39	10.36 ± 1.91
	Latency (ms)	276.00 ± 20.59	310.00 ± 15.02[*]
N400	Amplitude (μV)	5.93 ± 3.60	4.51 ± 1.00
	Latency (ms)	460.67 ± 65.11	516.67 ± 57.53

Mean ± SD. *p < 0.05 between groups.

TABLE 5 | SVM classification comparisons between the two age groups.

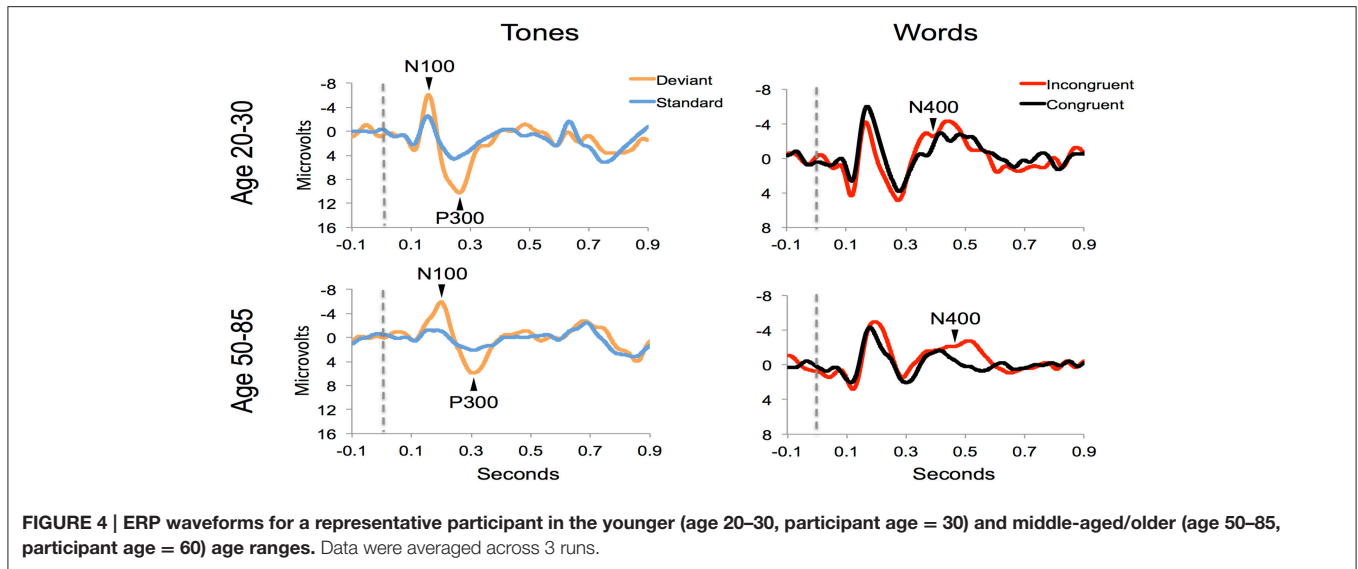|	P300, Younger	P300, Older	N400, Younger	N400, Older
Accuracy	99.31%	98.61%	86.11%	86.11%
TP	0.99	0.97	0.88	0.85
FP	0.00	0.00	0.15	0.13
Sensitivity	0.99	0.97	0.88	0.85
Specificity	1.00	1.00	0.85	0.88

TP, True Positive; FP, False Positive.

FIGURE 4 | ERP waveforms for a representative participant in the younger (age 20–30, participant age = 30) and middle-aged/older (age 50–85, participant age = 60) age ranges. Data were averaged across 3 runs.
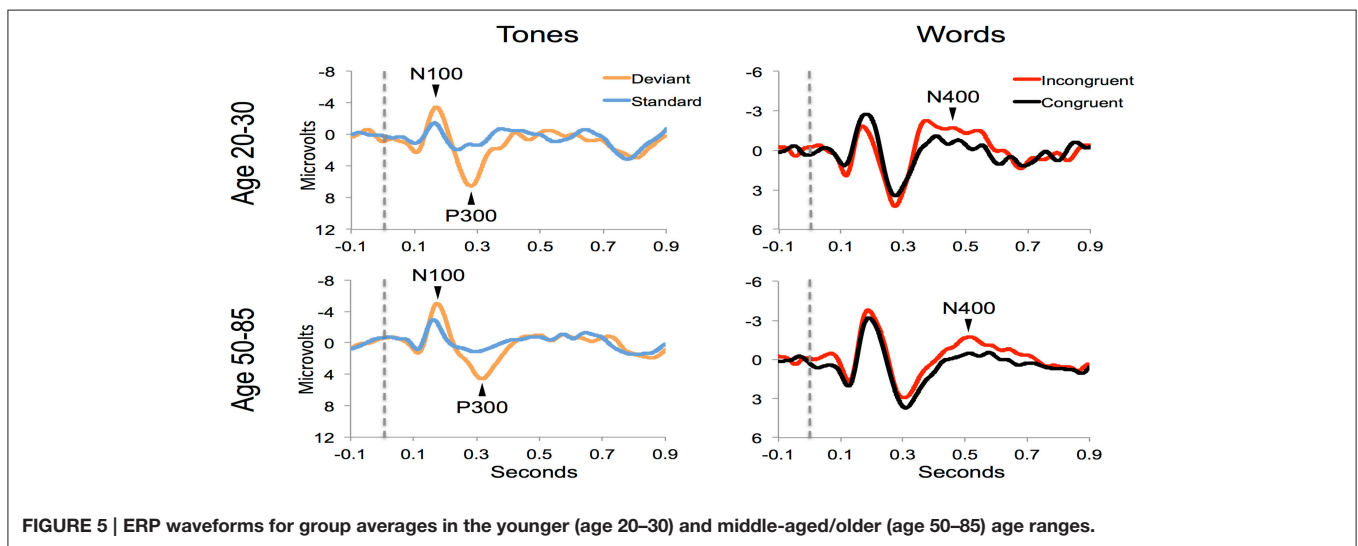

FIGURE 5 | ERP waveforms for group averages in the younger (age 20–30) and middle-aged/older (age 50–85) age ranges.

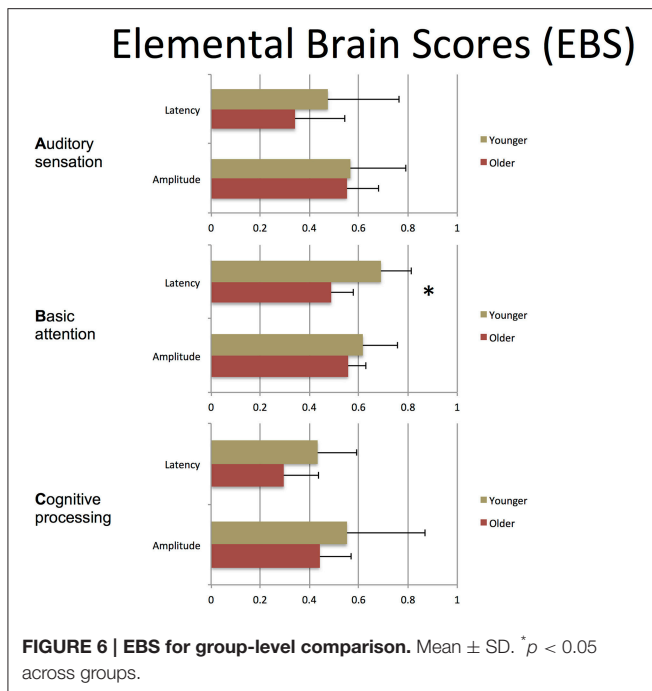**TABLE 6 | EBS values for group-level characteristics.**

		Age 20-30	Age 50-85
N100	Amplitude	0.57 ± 0.22	0.55 ± 0.13
	Latency	0.47 ± 0.29	0.34 ± 0.20
P300	Amplitude	0.62 ± 0.14	0.56 ± 0.07
	Latency	0.69 ± 0.12	0.49 ± 0.09*
N400	Amplitude	0.55 ± 0.32	0.44 ± 0.13
	Latency	0.43 ± 0.16	0.30 ± 0.14

Mean ± SD. $^*p < 0.05$ between groups.

the successful detection of the three key ERPs at the individual level (Hypothesis 1), confirmed the expected pattern of age-related ERP changes (Hypothesis 2), and enabled the translation of ERPs into the brain vital sign framework (Hypothesis 3). Importantly, this provided the initial step toward a brain vital sign approach that preserves and simplifies the essential valuable ERP results, but enables practical, accessible "vital sign" attributes.

Robust individual level detection of ERPs like the N100, P300, and N400 has become possible through machine learning advances (Parvar et al., 2014; Sculthorpe-Petley et al., 2015). Indeed, even within the current small initial validation sample, the ERPs were successfully detected for individuals across the life span (**Figures 4**, **5**). SVM-based analysis allowed expert-independent validation with high accuracy, sensitivity, and specificity. SVM-based methods are generally considered extremely well suited for use in biomedical data due to their ability to deal with sparse learning scenarios (Yeo et al., 2009, pp. 115–124). Traditionally, SVM-based techniques for ERP are restricted to within-subject training and classification for brain machine interface applications (Parvar et al., 2014, pp. 1–12). By contrast, in the current application

Elemental Brain Scores (EBS)

FIGURE 6 | EBS for group-level comparison. Mean ± SD. $^*p < 0.05$ across groups.

the SVM-based methods involved between-subject training and classification, further demonstrating potential for robust clinical applications. Moreover, permutation analysis based verification of performance provides further confidence regarding the robustness of these approaches.

As an initial validity check, predictable age-related changes in ERPs were examined and verified within the brain vital sign framework. To demonstrate the relative sensitivity differences between subjective behavioral tests and objective physiological measures, standard mental status assessments were compared to that of the ERP results. Results from MMSE and MoCA were both in the healthy range for the younger (age 20–30) and older (age 50–85) groups. While the ERP results generally matched this pattern, it was possible to show subtle age-related P300 latency delays ($p = 0.008$) in older adults, consistent with previous studies (Braverman and Blum, 2003, pp. 124–139). A similar trend was observed for N400 latency delays ($p = 0.07$). Thus, while both behavior and brain—based testing showed intact cognitive status, only the ERP evidence showed enhanced sensitivity to age-related changes in healthy brain function. Future work will further characterize standard factors in larger normative samples. These include characterization of aging related confounds, effects of education levels, impact of concurrent changes in other vital signs such as heart rate and blood pressure, and correlations between specific EBS components and traditional behavioral measures. Similarly, planned future work will also explore the opportunity to include both resting state as well as other stimulus-related brain response measures (such as event related spectral perturbations) into the BVS framework.

To translate ERP results into the brain vital sign framework, we applied a linear transformation to reduce complexity and create a standard clinical schematic of ABC: (A) N100 =

Auditory sensation; (B) P300 = Basic attention; and (C) N400 = Cognitive processing (**Figure 1**). Brain vital sign scores were then derived through comparison to the mean and standard deviation of the normative data. All participants showed an overall brain vital sign score of 30, derived from perfect 10-point ABC sub-scores (**Figure 2**). This component provided a normative evaluation for healthy brain function. As an initial development and to retain applicability over a wide range of potential dysfunction, all components were weighted equally in this framework. Future work may create variations/improvements that weigh the components differently for applications in specific disorders.

To transform ERP results into measurements of individual changes over time, the amplitude and latency measurements for all three responses were converted into 6 elemental brain scores (EBS: 3 responses × 2 measurements). Importantly, the EBS transformation involved a normative comparison against the best possible measurement, resulting in scores ranging from 0 to 1. During initial validation, EBS transformation preserved the pattern of age-related changes, with significant change in the "B" component latency ($p = 0.004$) and a similar trend in "C" component latency ($p = 0.07$).

The justification for a brain vital sign framework is strongly within the need for a practical and objective physiological measure of healthy brain function, combined with the capability for portable EEG/ERPs to meet the practical requirements and utilize well-established neural responses (i.e., studied extensively for 35–70 years). The challenge has related to translating/transforming ERPs to begin addressing the clinical requirements for vital signs.

Accordingly, the current study represents only an initial development effort, with a number of steps and caveats remaining: (1) the initial validation used a relatively small sample size, with further validation work currently being conducted; (2) the critical need for hardware platform independence remains to be systematically examined in order to understand differences between EEG acquisition systems; (3) the development of standardized normative databases represents an on-going improvement and refinement; (4) the continuing development of analyses to characterize sensitivity, specificity, reliability, and other standard metrics are needed; and (5) more comprehensive evaluations anchored to standard vital sign developmental approaches must also be conducted. Nonetheless, the ability to move beyond the traditional and heavily expert-dependent ERP research setting to a more clinically-oriented brain vital sign framework allows for a systematic method of assessing healthy brain function. The current study provides an initial demonstration of the framework, but the small sample size necessitates that the results should be further validated in a larger sample. Furthermore, it should be noted that there are several approaches available for eliciting the ERP components. We have demonstrated one approach that we believe makes the oddball discrimination task easier in order to maximize applicability across age groups and brain functional status. Establishing a baseline measurement approach for healthy brain function is critical, particularly when questions of dysfunction arise due to conditions such as concussion and

dementia. This study represents the initial steps toward such an approach.

CONCLUSION

Clinical evaluations of healthy brain functioning is moving from indirect subjective behavior-based tests, to objective, physiological measures of brain function, such as those derived from ERPs. We have previously demonstrated the essential role for clinical ERPs to evaluate functional status following serious injury and/or disease. The current study addressed an important gap: the need for a clinical-accessible brain vital sign framework that utilizes well-established ERPs. As an initial step, the framework was used to evaluate healthy brain function across the life span. The findings confirmed the ERPs at the individual level, verified predictable age-related differences, and demonstrated successful linear transformation to create the brain vital sign framework.

AUTHOR CONTRIBUTIONS

Conceptualization and study design: SG, CL, XS, and RD. Literature search: SG, CL, and RD. Data collection: SG, CL, SF, LL, and GP. Analysis planning: SG, CL, XS, RV, RR, KF, and RD. Data analysis: SG, CL, and LL. Result presentation: SG, CL, XS, and RD. Analysis outcome verification: SG, CL, XS, and RD. Result interpretation: All authors. Manuscript preparation: SG, CL, and RD. Critical editing and approval of submission: All authors.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fnins.2016.00211

REFERENCES

Braverman, E. R., and Blum, K. (2003). P300 (Latency) event-related potential: an accurate predictor of memory impairment. *Clin. EEG Neurosci.* 34, 124–139. doi: 10.1177/155005940303400306

Cassidy, S. M., Robertson, I. H., and O'Connell, R. G. (2012). Retest reliability of event related potentials: evidence from a variety of paradigms. *Psychophysiology* 49, 659–664. doi: 10.1111/j.1469-8986.2011.01349.x

Connolly, J. F., Byrne, J. M., and Dywan., C. A. (1995). Assessing adult receptive vocabulary with event-related potentials: an investigation of cross-modal and cross-form priming. *J. Clin. Exp. Neuropsychol.* 17, 548–565. doi: 10.1080/01688639508405145

Connolly, J. F., and D'Arcy, R. C. N. (2000). Innovations in neuropsychological assessment using event-related brain potentials. *Int. J. Psychophysiol.* 37, 31–47. doi: 10.1016/S0167-8760(00)00093-3

Daltrozzo, J., Wioland, N., Mutschler, V., Lutun, P., Calon, B., Meyer, A., et al. (2009). Cortical information processing in coma. *Cogn. Behav. Neurol.* 22, 53–62. doi: 10.1097/WNN.0b013e318192ccc8

D'Arcy, R. C. N., Hajra, S. G., Liu, C., Sculthorpe, L. D., and Weaver, D. F. (2011). Towards Brain First-Aid: a diagnostic device for conscious awareness. *IEEE Trans. Biomed. Eng.* 58, 750–754. doi: 10.1109/TBME.2010.2090880

D'Arcy, R. C. N., Marchand, Y., Eskes, G. A., Harrison, E. R., Phillips, S. J., Major, A., et al. (2003). Electrophysiological assessment of language function following stroke. *Clin. Neurophysiol.* 114, 662–672. doi: 10.1016/S1388-2457(03)00007-5

Davis, P. A. (1939). Effects of acoustic stimuli on the Waking Human Brain. *J. Neurophysiol.* 2, 494–499.

Duncan, C. C., Barry, R. J., Connolly, J. F., Fischer, C., Michie, P. T., Naatanen, R., et al. (2009). Event-related potentials in clinical research: guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400. *Clin. Neurophysiol.* 120, 1883–1908. doi: 10.1016/j.clinph.2009.07.045

Fleck-Prediger, C., Hajra, S. G., Dick B. D., Gray, D. S., Liu, C. C., Petley, L., et al. (2014). *Clinical Applications of the Halifax Consciousness Scanner: Tracking Recovery in a Severely Brain Injured Patient.* International Brain Injury Association, International NeuroTrauma Letter, Retrieved from: http://www.internationalbrain.org/articles/int-issue-37/

Folstein, M. F., Folstein, S. E., and McHugh, P. R. (1975). "Mini-Mental State": a practical method for grading the cognitive state of patients for the clinician. *J. Psychiat. Res.* 12, 189–198. doi: 10.1016/0022-3956(75)90026-6

Gawryluk, J. R., and D'Arcy, R. C. N. (2010). "Electroencephalography: basic concepts and brain applications," in *Handbook of Physics in Medicine and Biology, 1st Edn.,* ed R. Splinter (New York, NY: CRC Press), 24–1–24–14

Gawryluk, J. R., D'Arcy, R. C., Connolly, J. F., and Weaver, D. F. (2010). Improving the clinical assessment of consciousness with advances in electrophysiological and neuroimaging techniques. *BMC Neurol.* 10:11. doi: 10.1186/1471-2377-10-11

Giacino, J. T., Fins, J. J., Laureys, S., and Schiff, N. D. (2014). Disorders of consciousness after acquired brain injury: the state of the science. *Nat. Rev. Neurol.* 10, 99–114. doi: 10.1038/nrneurol.2013.279

Golland, P., and Fischl, B. (2003). *Permutation Tests for Classification: Towards Statistical Significance in Image-Based Studies.* Berlin: Springer. doi: 10.1007/978-3-540-45087-0_28

Kotchoubey, B., Lang, S., Mezger, G., Schmalohr, D., Schneck, M., Semmler, A., et al. (2005). Information processing in severe disorders of consciousness: vegetative state and minimally conscious state. *Clin. Neurophysiol.* 116, 2441–2453. doi: 10.1016/j.clinph.2005.03.028

Kutas, M., and Hillyard, S. A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science* 207, 203–205. doi: 10.1126/science.7350657

Lezak, M. D. (2004). Neuropsychological Assessment. New York, NY: Oxford University Press.

Luck, S. J. (2014). *An Introduction to the Event-Related Potential Technique.* Cambridge: MIT Press.

Marchand, Y., D'Arcy, R. C. N., and Connolly, J. F. (2002). Linking neurophysiological and neuropsychological measures for aphasia assessment. *Clin. Neurophysiol.* 113, 1715–1722. doi: 10.1016/S1388-2457(02)00224-9

Morlet, D., Bouchet, P., and Fischer, C. (2000). Mismatch negativity and n100 monitoring: potential clinical value and methodological advances. *Audiol. Neuro Otol.* 5, 198–206. doi: 10.1159/000013881

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., et al. (2005). The montreal cognitive assessment, MoCA: a brief screening tool for mild cognitive impairment. *J. Am. Geriatr. Soc.* 53, 695–699. doi: 10.1111/j.1532-5415.2005.53221.x

Parvar, H., Sculthorpe-Petley, L., Satel, J., Boshra, R., D'Arcy, R. C. N., and Trappenberg, T. P. (2014). Detection of event-related potentials in individual subjects using support vector machines. *Brain Informat.* 2, 1–12. doi: 10.1007/s40708-014-0006-7

Pravdich-Neminsky, V. V. (1913). Experiments on the registration of the Electrical Phenomena of the Mammalian Brain. *Zbl. Physiol* 27, 951–960.

Reith, F. C., Brennan, P. M., Maas, A. I., and Teasdale, G. M. (2016). Lack of standardization in the use of the glasgow coma scale: results of international surveys. *J. Neurotrauma* 33, 89–94. doi: 10.1089/neu.2014.3843

Ries, A. J., Touryan, J., Vettel, J., McDowell, K., and Hairston, W. D. (2014). A comparison of electroencephalography signals acquired from conventional and mobile systems. *J. Neurosci. Neuroeng.* 3, 10–20. doi: 10.1166/jnsne.2014.1092

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

Sculthorpe-Petley, L., Liu, C., Hajra, S. G., Parvar, H., Satel, J., Trappenberg, T. P., et al. (2015). A rapid event-related potential (ERP) method for point-of-care evaluation of brain function: development of the halifax consciousness scanner. *J. Neurosci. Methods* 245, 64–72. doi: 10.1016/j.jneumeth.2015.02.008

Sutton, S., Tueting, P., Zubin, J., and John, E. R. (1967). Information delivery and the sensory evoked potential. *Science* 155, 1436–1439.

Teasdale, G., and Jennett, B. (1974). Assessment of coma and impaired consciousness: a practical scale. *Lancet* 304, 81–84. doi: 10.1016/S0140-6736(74)91639-0

Vanhaudenhuyse, A., Laureys, S., and Perrin, F. (2008). Cognitive event-related potentials in comatose and post-comatose states. *Neurocrit. Care* 8, 262–270. doi: 10.1007/s12028-007-9016-0

Wijnen, V. J. M., Van Boxtel, G. J. M., Eilander, H. J., and De Gelder, B. (2007). Mismatch negativity predicts recovery from the vegetative state. *Clin. Neurophysiol.* 118, 597–605. doi: 10.1016/j.clinph.2006.11.020

Williams, L. M., Simms, E., Clark, C. R., Paul, R. H., Rowe, D., and Gordon, E. (2005). The test-retest reliability of a standardized neurocognitive and neurophysiological test battery: "neuromarker". *Int. J. Neurosc.* 115, 1605–1630. doi: 10.1080/00207450590958475

Yeo, M. V. M., Li, X., Shen, K., and Wilder-Smith, E. P. V. (2009). Can SVM be used for automatic EEG detection of drowsiness during car driving? *Safety Sci.* 47, 115–124. doi: 10.1016/j.ssci.2008.01.007