SCIENTIFIC DATA

Motor Evoked Potentials Database

Jan Yperman[1], Thijs Becker[2*], Liesbet Peeters

November 27, 2020

1. Theoretical Physics, Hasselt University, Diepenbeek, Belgium 2. I-Biostat, Data Science Institute, Hasselt University, Diepenbeek, Belgium 3. BIOMED, Hasselt University, Diepenbeek, Belgium 4. Revalidation and MS Center Pelt, Pelt, Belgium *corresponding author(s): Liesbet Peeters (liesbet.peeters@uhasselt.be)

Abstract

Multiple Sclerosis (MS) is a chronic disease affecting millions of people worldwide. The signal conduction through the nervous system of MS patients deteriorates, leading to a wide range of symptoms, e.g., balance impairment and muscle weakness. Evoked potential measurements allow clinicians to monitor the degree of deterioration and are used for decision support. In particular, this dataset contains motor evoked potential measurements, in which the brain is stimulated and the resulting signal is measured by electrodes in the hands and feet. This results in time series of 100 milliseconds long, mostly consisting of 1920 samples. The dataset consists of roughly 100 000 such measurements, performed in day-to-day clinical care over a period of 6 years. Alongside these measurements, clinical metadata is also available such as EDSS (a score indicating disease severity), as well as other patient metadata. This dataset can be used for medical research to explore the role of evoked potentials in MS research, but may also be used as a real-world benchmark for machine learning techniques for time series analysis.

Background & Summary

Multiple sclerosis (MS) is an incurable, chronic disease characterized by the disruption of electrical signal conduction over axons in the central nervous system. One of the main reasons for this disruption is the loss of the myelin sheath [1]. To assess the impact on normal conduction in people with MS clinicians rely on, among other techniques, evoked potential measurements (EP) [2]. These types of measurements provide insight in neural conduction by stimulating the nervous system in one place, and measuring the resulting signal at some other point. The different EP modalities correspond to different sites of stimulus and measurement. In the case of motor evoked potentials (MEP), which is the modality this dataset contains, the motor cortex (M1) is stimulated using transcranial magnetic stimulation (TMS) and the resulting signal is measured in the hands or feet.

The current dataset contains MEPs obtained during about 6 years of follow-up in the Rehabilitation and MS center at Pelt, Belgium. For most of the patients the Expanded Disability Status Score (EDSS) is also available. Some exploratory work has already been performed on this dataset [3] and [4]) and it is currently one of the datasets being used in the AI Challenge Flanders initiative as a proof of concept. Its size and the fact that it contains the full measurement time series (as opposed to just some summary values such as latency and amplitude) make this a rich dataset for both AI and clinical research. Possible use cases for this dataset include time series analysis, disease progression prediction tasks, and matrix completion tasks (for e.g. patient trajectory analyses).

Methods

Motor evoked potentials were recorded from the abductor pollicis brevis (APB) and abductor hallucis (AH) muscles bilaterally. Magnetic stimuli were delivered to the hand and leg areas of the motor cortex with a

Magstim 200^2 device (The Magstim Company Ltd., Whitland, UK) via a round coil with an inner diameter of 9 cm with maximal output of the stimulator (2.2 T). The signal is recorded for 100 ms. The dataset contains measurements from two separate machines (cfr. Section Data Records, the machine field). The acquisition rate for machine A is 20 kHz, for B this is 19.2 kHz. Signals from machine A are filtered between 0.6 Hz and 10 kHz, while machine B has a high-pass filter of 100 Hz. The measurements are not averaged across multiple trials.

The measurements are performed in a standardized way to minimize variations due to factors such as coil orientation, stimulus intensity etc. For the hands, electrodes are placed at three places: on top of the hand (ground), the APB muscle, and the proximal phalanx of the thumb. The first excitation is at 45% of the maximal stimulator output. New stimuli are presented with an increase of 5 percentage points. The measurement ends if the amplitude reaches 1 millivolt, or if the amplitude stops increasing for stronger stimuli. If the signal is of bad quality, as judged by the nurse, it is discarded.

For the feet, electrodes are placed at three places: on top of the foot (ground), the big toe, and the AH muscle. The first excitation is at 50% of the maximal stimulator output. New stimuli are presented with an increase of 5 percentage points. The measurements end if the amplitude reaches 1 millivolt, or if the amplitude stops increasing for stronger stimuli. If the signal is of bad quality, as judged by the nurse, it is discarded.

An example of all the EPTS of the MEP for a single visit is shown in Figure 1. For each limb, each excitation strength results in one EPTS.

Code availability

To be added once the dataset is online.

Data Records

Cohort description

For the patients who have clinical data and for whom the MS type was entered (263 patients), we have the following distribution of MS type: Relapsing-remitting MS (RRMS) (76.0%), Secondary progressive MS (SPMS) (19.4%), Primary progressive MS (PPMS) (3.4%), Clinically Isolated Syndrome (CIS) (1.1%) The average age at the time of a visit is 49 ± 14 years. The average of the average EDSS score per patient is 3.5.

Tables overview

In this section we provide some general numbers of the dataset. The names of the tables are self-explanatory. The relations between patients, visits, tests, and measurements are one-to-many, in that order. That is, each patient can have multiple visits. Each visit can have multiple tests. Each test can have multiple measurements. Finally, each patient can have multiple EDSS-measurements.

- patient (963 records)
- visit (5586 records)
- test (20844 records)
- measurement (96290 records)
- edss (7414 records)

SCIENTIFIC DATA

Field descriptions

patient

This table is contained in the file patient.csv.
Key:

- **patient_uid**: A unique identifier for a patient. Range: 0 - 964

Other fields:

- **date_of_birth**: The date of birth of the patient, accurate to 5 years. Format: YYYY-MM-DD. Missing values for this field: 1.87%. Average age at time of visit: 49 ± 14 years.
- **has_clinical_data**: Whether there is are EDSS-measurements available for this patient. Missing values for this field: 0.00%. Possible values: True (60.44%), False (39.56%)
- **sex**: The sex of the patient. Missing values for this field: 3.12%. Possible values: Female (70.53%), Male (29.47%)

visit

This table is contained in the file visit.csv.
Key:

- **patient_uid** - cfr. patient_uid in Table patient.
- **visit_uid**: A unique identifier for a visit. It is only unique in conjunction with the patient uid. Range: 0 - 25

Other fields:

- **machine**: Indicates which of the two machines the measurement was performed on. Missing values for this field: 0.00%. Possible values: B (74.81%), A (25.19%)
- **team**: Indicates which team performed the measurements. Missing values for this field: 0.00%. Possible values: B (52.22%), A (47.78%)
- **visit_date**: The date of a particular visit. Note that these were shifted per patient to preserve privacy. The relative time between visits of a patient are preserved. Format: YYYY-MM-DD. Missing values for this field: 0.00%. Range: 12/03/1984 - 26/12/2014

test

This table is contained in the file test.csv.
Key:

- **patient_uid** - cfr. patient_uid in Table patient.
- **visit_uid** - cfr. visit_uid in Table visit.
- **test_uid**: A unique identifier for a test. It is only unique in conjunction with the patient, and visit uid. Range: 0 - 6

Other fields:

- **anatomy**: The muscle on which the measurement is performed. Missing values for this field: 0.00%. Possible values: APB (50.71%), AH (49.29%)
- **side**: The side on which the measurement was performed, e.g., left or right arm. This value is generated based on the measurement protocol which says the tests must be performed on the right limb first, then the left limb. The accuracy of this field depends on how well the nurses stuck to this protocol. Missing values for this field: 0.00%. Possible values: R (50.45%), L (49.55%)

measurement

This table is contained in the file measurement.csv.
Key:

- **patient_uid** - cfr. patient_uid in Table patient.
- **visit_uid** - cfr. visit_uid in Table visit.
- **test_uid** - cfr. test_uid in Table test.
- **measurement_uid**: A unique identifier for a measurement. It is only unique in conjunction with the patient, visit, and test uid. Range: 0 - 49

Other fields:

- **marker_N_amplitude(mv)**: Float field which indicates the amplitude of a certain marker, which were placed by nurses in the clinic. The placement of the markers are illustrated in Figure 2. This should be used with the corresponding marker_N_latency(ms) field. Refer to Table 1 for some statistics of this key.
- **marker_N_latency(ms)**: Float field which indicates the amplitude of a certain marker, which were placed by nurses in the clinic. The placement of the markers are illustrated in Figure 2. This should be used with the corresponding marker_N_amplitude(ms) field. Refer to Table 1 for some statistics of this key.
- **notch_filter**: A boolean field indicating whether a 50Hz notch filter was applied during the measurement. Missing values for this field: 0.00%. Possible values: False (81.42%), True (18.58%)
- **timeseries**: The raw EP timeseries. Missing values for this field: 0.00%. This field holds an integer identifier for the time series belonging to the measurement. The timeseries themselves are stored separately, with the filename corresponding to the identifier stored in this field. These timeseries measure the electrical signal received from the motor cortex in the muscle indicated in the **anatomy** field. The measurement always takes 100 milliseconds. Due to the different sample rate of the machines, the number of samples per timeseries differs slightly (1920 or 2000).

edss

This table is contained in the file edss.csv.
Key:

- **patient_uid** - cfr. patient_uid in Table patient.
- **edss_uid** - See below.

Other fields:

- **date**: The date on which the EDSS measurement was performed. Format: YYYY-MM-DD. Missing values for this field: 0.00%. Range: 06/02/1963 - 21/03/2016
- **edss**: The result of the EDSS measurement. Missing values for this field: 0.00%. Average value: 3.28 ± 2.05

Privacy

A number of steps were taken to ensure the privacy of the patients on whom these measurements were performed. Sensitive fields that were removed include the patient id used in the clinic and the names of the patients. The dates of the measurements were shifted by a random period of time, though in such a way that the relative time between measurements of a given patient is preserved. The same shift was applied to the dates of the clinical measurements. The patient's date of birth was shifted in the same way the measurement dates were shifted to ensure the age at the time of measurement still matched. The resulting date of birth was then shifted randomly within a span of 10 years.

Technical Validation

The Rehabilitation and MS center, where these measurements were performed, is a respected MS treatment center under the supervision of Dr. Bart Van Wijmeersch. Measurements are performed in a standardized fashion and are used in day-to-day clinical followup meaning they were all at one point reviewed by the neurologist treating the patient.

In the software the patients were only identified through names and birthdates entered by the nurses during the patient's first visit. This was done separately for each machine. These combinations of names and birthdates were matched with the clinical database containing the results of the EDSS measurements. This process was done very conservatively, erring mainly on false negatives as opposed to false positives. The date of birth is matched first. If there is a match, a sequence matcher is then used to match the names. The sequence matcher ensures that small spelling mistakes in the names do not discard of the match entirely. Inexact matches were reviewed manually. Some patients also had an old patient identifier, which was constructed from their birthdate. If the birth date field was not entered, the birthdate was extracted from this identifier. Finally, if a patient's birthdate was not recovered during some visits (in which case that visit would be discarded), we look at other visits with the exact same first and last name which do have a birthdate. The clinic ID of those visits is then used for the visits without a match.
To check the validity of the final database, a number of sanity checks were automated. These checks include things like the ages being in a valid range, the sex of the patients being unique, the number of visits per patient to be reasonable, the average latency should coincide with that given in the literature etc. If any systematic error was made during the extraction of the data, these checks are likely to fail.

The timeseries were also compared to the reports that are generated by the software, which are the files the neurologists look at to assess the measurements. This was done for a random sample of the visits. These reports could be replicated exactly using our extracted database.

Usage Notes

The measurements almost always include a measurement artifact at the start of the timeseries. This is caused by the electric field generated by the coil which is also picked up by the electrodes. This part can be safely discarded. In our work, we usually discard the first 70 points of any time series.

There are multiple measurements per test. Usually the measurement with the highest peak-to-peak amplitude is considered to be the most informative, which is what we used in our previous work. One should be careful to discard the previously mentioned measurement artifact when calculating the peak-to-peak amplitude, as this artifact usually has a larger amplitude than the relevant peaks in the timeseries.

Since the measurements were performed on two separate machines, a researcher using this dataset has two options. Either preprocess the time series to match the settings of the two machines (e.g. by applying the same filters to the machines). Alternatively, the researcher could opt to use the measurements in their current form and use the differences between the machines as regularization for their model, ensuring it generalizes well to multi-center studies. For the former approach, note that the two machines on which the measurements were made have slightly different sampling rates. Therefore, we suggest downsampling the

time series from the machine with 2000 samples per time series to the 1920 samples of the other machine. The high-pass cutoff frequencies on both machines also differ. This could be rectified by applying a 100Hz filter to the machine which originally has a 0.6Hz filter applied to it.

In case no spontaneous response or MEP in rest position is obtainable, a light voluntary contraction of the muscle in question is asked in order to activate the motor cortex and increase the possibility of becoming a motor answer. This so-called facilitation method is usually very noisy due to baseline contraction of the muscle measured. Unfortunately, whether this method was used was not consistently indicated with each measurement. Facilitated measurements are characterized by a non-flat signal right from the start of the measurement. In our previous work we dropped any time series where the sum of absolute values in the interval between 5 milliseconds (to avoid the aforementioned measurement artifact) and 17 milliseconds (the earliest the signal can physiologically arrive for the hands) is above an empirically determined threshold of 10. This threshold was chosen by visually inspecting the time series, and should therefore only be considered as a guideline. This calculation was done on the resampled time series (cfr. previous paragraph).

Possible use cases of this dataset include: Cross-sectional studies to determine the EDSS-value from the time series. From there the full time series could be used to define a more stable score for disability in MS than EDSS based on MEPs

Acknowledgements

The authors would like to thank the personnel working at the MS clinic for their help while collecting the data. In particular, we would like to thank Henny Strackx for his help with extracting the data from the local servers, Daisy Mobers for her input on how the measurements were performed, and Filip Van Elsen for his input on the technical aspects of the measurement devices used. Furthermore, we thank Patricia Tielens, Dieter Flament, An Voets, Jorina Nickmans and Benny Daems for their legal counsel and Wim Lamotte for answering our questions about the security measures that had to be taken for data transfer.

Author contributions

JY extracted the data from the clinic, wrangled the data, cleaned and deidentified the resulting dataset and wrote the manuscript. TB provided feedback in developing the data processing pipeline and wrote the manuscript. LMP provided feedback in developing the data processing pipeline, set up the collaboration between the hospital in Pelt and Hasselt University and handled the collaboration agreements and ethics approvals required for this work. BVW and VP performed the measurements in the hospital in Pelt.

Competing interests

None of the authors have anything to declare regarding conflicts of interest.

Figures and figures legends

SCIENTIFIC DATA

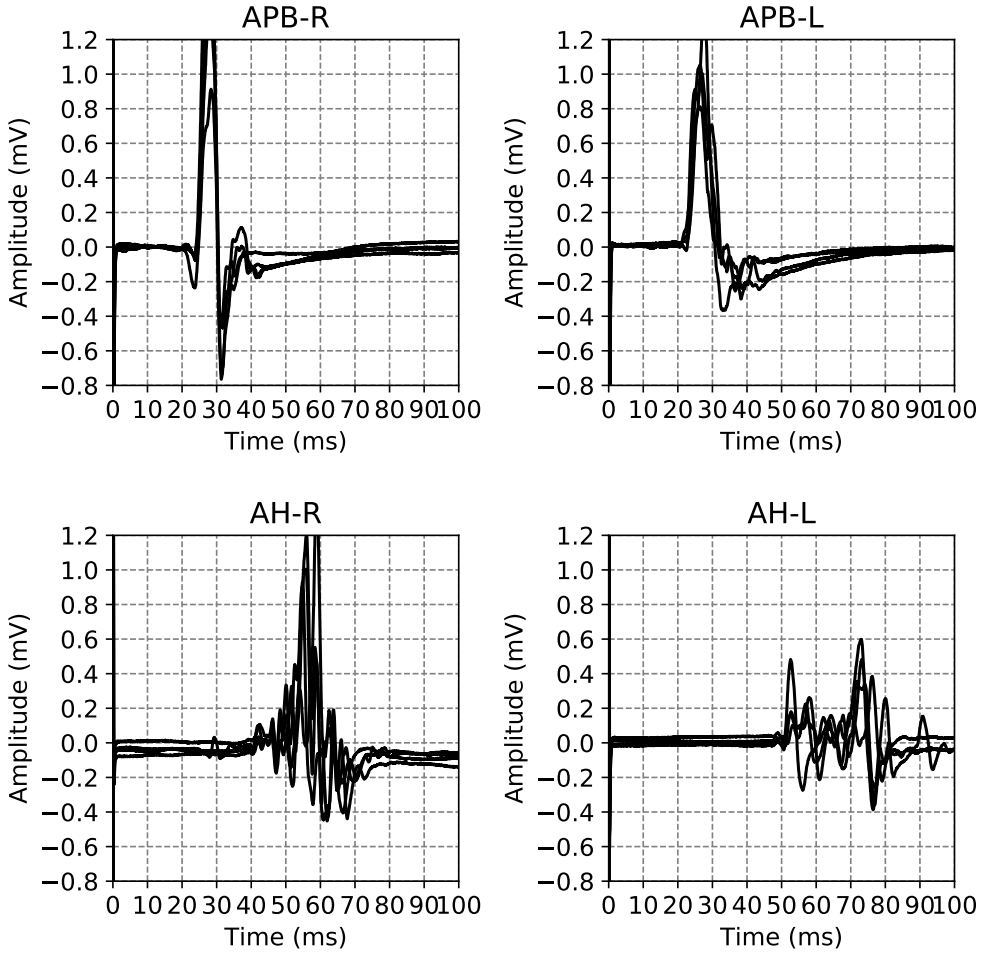

Figure 1: Example of the motor evoked potential time series recorded at a single hospital visit. The labels on the plot indicate the limb and the side on which the measurement was performed, M. Abductor Pollicis Brevis (APB) for the hands, M. Abductor Hallucis (AH) for the feet. The sides are indicated using R and L for right and left respectively. The time series for the same limb are the result of different magnetic excitation strengths.

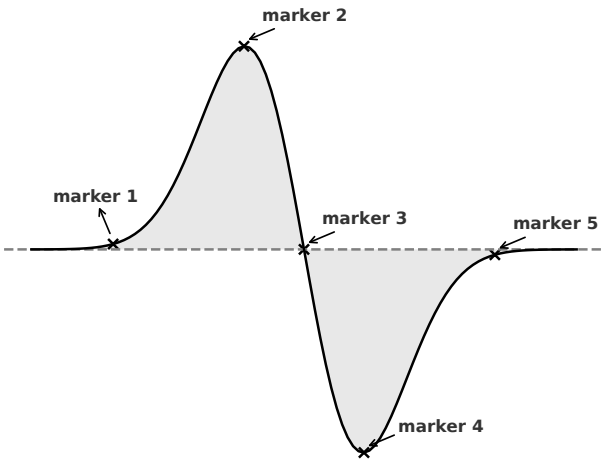

Figure 2: Illustration of the placement of the markers

Figure 3: A schematic overview of the tables in the dataset. Per table, all the bold fields combine to make a unique identifier for a record in that table (composite key). Section Data Records contains descriptions of each of the fields.

SCIENTIFIC DATA

Tables

	AH	APB
marker_1_amplitude(mv)	$-0.01 \pm 0.06(7.75\%)$	$-0.00 \pm 0.06(2.71\%)$
marker_1_latency(ms)	$43.15 \pm 7.57(7.75\%)$	$22.21 \pm 4.92(2.71\%)$
marker_2_amplitude(mv)	$0.78 \pm 0.72(17.40\%)$	$1.20 \pm 1.09(4.80\%)$
marker_2_latency(ms)	$53.08 \pm 8.96(17.40\%)$	$30.37 \pm 6.63(4.80\%)$
marker_3_amplitude(mv)	$0.04 \pm 0.16(18.70\%)$	$0.02 \pm 0.17(5.13\%)$
marker_3_latency(ms)	$55.41 \pm 9.03(18.70\%)$	$33.40 \pm 7.28(5.13\%)$
marker_4_amplitude(mv)	$-0.50 \pm 0.42(18.77\%)$	$-0.77 \pm 0.66(5.14\%)$
marker_4_latency(ms)	$59.22 \pm 9.86(18.77\%)$	$36.81 \pm 9.32(5.14\%)$
marker_5_amplitude(mv)	$-0.03 \pm 0.10(18.89\%)$	$-0.08 \pm 0.15(5.36\%)$
marker_5_latency(ms)	$76.35 \pm 11.93(18.89\%)$	$51.57 \pm 10.59(5.36\%)$

Table 1: Summary statistics of the available markers. Refer to Figure 2 for the placement of the markers. The formatting is as follows: mean $\pm$ standard deviation (percentage missing). The statistics are shown separately for each of the possible anatomies. ms and mv indicate milliseconds and millivolts respectively.

References

[1] R. G. Emerson. Evoked potentials in clinical trials for multiple sclerosis. *J Clin Neurophysiol*, 15(2):109–16, 1998.

[2] Martin Hardmeier, Letizia Leocani, and Peter Fuhr. A new role for evoked potentials in ms? repurposing evoked potentials as biomarkers for clinical trials in ms. *Multiple Sclerosis Journal*, 23(10):1309–1319, 2017.

[3] Jan Yperman, Thijs Becker, Dirk Valkenborg, Niels Hellings, Melissa Cambron, Dominique Dive, Guy Laureys, Veronica Popescu, Bart Van Wijmeersch, and Liesbet M Peeters. Deciphering the morphology of motor evoked potentials. *Frontiers in neuroinformatics*, 14:28, 2020.

[4] Jan Yperman, Thijs Becker, Dirk Valkenborg, Veronica Popescu, Niels Hellings, Bart Van Wijmeersch, and Liesbet M Peeters. Machine learning analysis of motor evoked potential time series to predict disability progression in multiple sclerosis. *BMC Neurol*, 20(1):1–15, 2020.

Citing Data

In line with emerging industry-wide standards for data citation, references to all datasets described or used in the manuscript should be cited in the text with a superscript number and listed in the 'References' section in the same manner as a conventional literature reference. See the examples above.