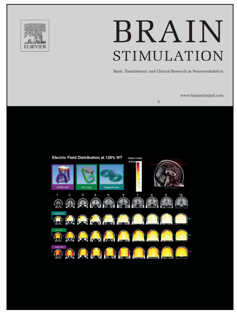
Accepted Manuscript



False positives associated with responder/non-responder analyses based on motor evoked potentials

Mark van de Ruit, Michael J. Grey

PII: S1935-861X(18)30416-9

DOI: https://doi.org/10.1016/j.brs.2018.11.015

Reference: BRS 1359

To appear in: *Brain Stimulation*

Received Date: 15 June 2018

Revised Date: 24 August 2018

Accepted Date: 29 November 2018

False positives associated with responder/non-responder analyses based on motor evoked potentials

Mark van de Ruit[1] *& Michael J. Grey*[2.]

[1] Department of Biomechanical Engineering, Delft University of Technology, Delft, the Netherlands

[2] Acquired Brain Injury Rehabilitation Alliance, School of Health Sciences, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, UK.

Keywords: variability, MEP, TMS, plasticity, corticospinal excitability, responders

Corresponding Author:

Dr. Michael J. Grey

Acquired Brain Injury Rehabilitation Alliance,

School of Health Sciences, University of East Anglia,

Norwich Research Park,

Norwich

NR4 7TJ

United Kingdom

Phone: +44 (0)1603 59 1682

E-mail: m.grey@uea.ac.uk

Abstract

Background: A trend in the non-invasive brain stimulation literature is to assess the outcome of an intervention using a responder analysis whereby participants are di- or trichotomised in order that they may be classified as either responders or non-responders.

Objective: Examine the extent of the Type I error in motor evoked potential (MEP) data subjected to responder analyses.

Methods: Seven sets of 30 MEPs were recorded from the first dorsal interosseous muscle in 52 healthy volunteers. Four classification techniques were used classify the participants as responders or non-responders: (1) the two-step cluster analysis, (2) Dichotomised thresholding, (3) relative method and (4) baseline variance method.

Results: Despite the lack of any intervention, a significant number of participants were classified as responders (21-71%).

Conclusion: This study highlights the very large Type I error associated with dichotomising continuous variables such as the TMS MEP.

Introduction

Similar to many other interventions, the efficacy of non-invasive brain stimulation (NIBS) is limited to a subset of the population and it is important to better understand what proportion of participants might respond. A recent trend in the NIBS literature is to use a responder analysis to classify participants as responders or non-responders following an intervention. This simplifies the statistical analysis, interpretation and presentation of results [1]. In the NIBS literature, this classification is typically performed by di- or trichotomising the motor evoked potential (MEP) produced in response to transcranial magnetic stimulation (TMS) as this is considered a surrogate marker of neuroplasticity [2].

Pellegrini, et al. 2018 [3] recently conducted a systematic review of responder analyses in NIBS and concluded that they can effectively identify subgroups based on response patterns, and used to estimate the proportion of participants who might respond to the intervention. However, they also noted a lack of consistency and consensus in the methods by which the data are quantified. Furthermore, they highlighted that many studies in the NIBS literature lack a control group. As a result, the effect of natural variability of the MEP is not accounted for with these analyses. The MEP magnitude has considerable trial-to-trial variability and drift over time, which arise due to controllable and uncontrollable factors of physiological (e.g. cortical rhythms, arousal, etc.) and non-physiological (e.g. TMS coil placement and/or movement) origin [4, 5].

Responder analyses methods gained popularity in the early 2000s in the clinical medicine and psychology literature primarily as a means to establish proportions of responders in drug trials and in marketing studies [6-8]. However, these methods were then criticised by methodologists who questioned the validity of dichotomising (or trichotomising) continuous variables. They noted in particular that inferences made from such analyses are susceptible to large Type I error (false positives) that can lead to erroneous conclusions [1, 6, 9-19]. The aim of the present study was to examine the extent of the Type I error in MEP data that are subjected to different types responder analyses.

Methods

Experimental procedures

Fifty-two healthy participants, without contraindication to TMS and no history of neurological psychiatric disorder, participated in the study (20 ± 2 y, range 18-25, 35 female). Participants visited the laboratory once for ~1 h, during which MEPs were recorded from the first dorsal interosseus (FDI). Participants sat comfortably and were instructed to relax both the hand and arm, and to keep their eyes open for the duration of the experiment. To facilitate this instruction throughout the experiment, interactive feedback of FDI muscle activity was provided on a computer monitor. TMS was delivered through a 90 mm figure-of-8 coil (type: batwing; type no. 15411) using a Magstim Rapid^2 stimulator (Magstim Ltd, Dyfed, United Kingdom). Coil position and orientation were monitored with frameless stereotaxy (BrainSight 2, Rogue Research Inc, Montreal, Canada). The stimulation intensity required to evoke 1 mV (SI_{1mV}) peak-to-peak MEPs (MEP_{pp}) was determined by adjusting the intensity until the mean of 30 stimuli produced a 1 mV MEP_{pp} (calibration data set in Figure 1A). Next, seven sets of 30 MEPs were recorded with a 4 s inter-stimulus interval and 2 min rest between sets. The first set was deemed a baseline to which the remaining 6 data sets would be compared. Figure 1A summarises the experimental protocol.

Statistical Analysis

The MEP_{pp} amplitude was extracted between 20-50 ms after stimulation and averaged across all stimuli within a set. The mean MEP_{pp} for each set was then used for statistical analysis and classification either: (1) without any further processing; or (2) after normalisation to the mean MEP_{pp} of the baseline set (B), the ‘grand average (GA) method’. Therefore, each classification method was performed twice on the same data, either the absolute mean MEP_{pp} amplitudes for each set, or the normalised GA data.

Before classification, the continuous data was analysed using a repeated measures analysis of variance (RM-ANOVA) across sets for the mean absolute MEP_{pp} values. Subsequently, the participants were classified using the four common methods found in the NIBS literature. Following classification, a mixed RM-ANOVA was performed on the absolute MEP_{pp} data with the within-factor ‘set’ and between-subjects factor 'group' (i.e. the result of the classification method). In addition, a one-way RM-ANOVA was performed for each group individually on the absolute MEP_{pp} data to classify groups of participants as either:

- (+) responders: significant increase in MEP_{pp} across set
- (−) responders: significant decrease in MEP_{pp} across set
- (0) responders or non-responders: no significant change in MEP_{pp} across set

If Mauchly's Test of Sphericity indicated that the assumption of sphericity had been violated, a Greenhouse-Geisser correction (GG) was performed. All statistical tests were performed using SPSS, with significance accepted at $p<0.05$.

Responder Analysis Methods

1) *Two-step cluster analysis:* This SPSS method uses a two-step clustering approach that allows automatic detection of the optimal number of clusters. In the first step all cases are scanned an pre-clustered based on a predefined distance criterion (e.g. squared Euclidian distance or log-likelihood) that specifies either the difference or similarity between cases. In the second step, the algorithm uses agglomerative hierarchical clustering to merge the sub clusters resulting from the first step into a smaller number of clusters. In the present study we allowed the algorithm to automatically determine the number of clusters rather than specifying two or three clusters. This is a commonly used method in NIBS literature [20-26].
2) *Dichotomised thresholding:* This method separates data into two groups based on a predefined threshold. For GA data, participants were categorised using the mean GA of sets (in our case sets T1-T6). Participants were then classified as negative responders for mean GA < 1 and positive responders for mean GA > 1. This analysis was also performed on absolute MEP_{pp} data. With absolute MEP_{pp} data this method can be applied either on a group level or individually. For the group level analysis, the mean MEP_{pp} amplitude across all participants was chosen as the threshold (1.35 mV in this study). For the individual analysis, the threshold is set to the mean MEP_{pp} of the baseline set for each participant individually. Next, each participant is classified as a positive responder if the mean MEP_{pp} across T1-T6 is greater than

the threshold and a negative responder if the mean MEP_{pp} across T1-T6 is less than the threshold. Dichotomised thresholding is a common method of subgrouping normalised MEP data [22, 24-33].

3) *Relative method:* This method is used to classify participants into three groups based on a predefined percent change from baseline threshold. This method has been used in several studies to trichotomise participants using a threshold of 10% [23, 34], 15% [35], 20% [20] or 50% [36]. In the present study we used a conservative approach by choosing 20% change from baseline as the threshold. For the GA data, participants are classified as negative responders for mean GA across sets T1-T6 < 0.8, positive responders for mean GA > 1.2 and non-responders between 0.8-1.2. Likewise for the absolute MEP_{pp} data the threshold was 1.35 ± 0.27 mV as for the collected data the group mean of the baseline set B was 1.35 mV. This procedure was also performed on an individual level, in which case the threshold was individually determined based on the mean MEP_{pp} amplitude of set B.

4) *Baseline variance method:* In this method participants are trichotomised based on the variance of the baseline measure. For the GA data, the standard error (SE) of the GA of the baseline set was 0.14 across all participants. Therefore, a participant was classified as a (−) or (+) responder if the mean GA across sets T1-T6 was smaller or greater than 1.27 (95% confidence limit (CL) 1 ± 0.27) and a non-responder otherwise. Similarly, for MEP_{pp} data the SE of the baseline set was 0.17 across all participants (95% CL 1.35 ± 0.36 mV) and therefore a participant was a (+) responder when above this upper limit, a (−) when below the lower limit or a non-responder otherwise. The same analysis was also performed on the level of each individual, i.e. the CL of the baseline set was determined individually to assign the participant to the correct group. This method has been used in several studies [28, 33, 37-41].

Results

A one-way RM-ANOVA applied across all seven data sets (B-T6) before dichotomisation revealed neither a significant difference in mean MEP_{pp} amplitude across these data sets ($F_{(4.76,242.75)} = 1.27^{GG}$, p=0.28) nor in GA ($F_{(4.74,241.73)} = 1.31^{GG}$ p=0.26; Figure 1B).

The results for the subgrouping methods are presented in Table 1 and for the group level analysis visualized in Figure 1C. The SPSS two-step cluster analysis determined two clusters to best separate the data. For the MEP_{pp} data 11 participants (~21%) were classified as responders, showing a significant increase in MEP_{pp} (p<0.01) across time, and 41 participants (~79%) were classified as non-responders (p=0.96). The same groups were identified using the GA data but with 19 responders (p<0.01) and 33 non-responders (p=0.22). The MEP_{pp} and GA across time for each group is illustrated in Figure 1C.

Using the dichotomised thresholding method on MEP_{pp} data and a group level , 33 participants (63%) were classified as (+) responders (p<0.01) and 19 participants (37%) as non-responders (p=0.88). For the GA data, 28 participants (54%) were classified as (+) responders (GA > 1, p<0.01) and 24 participants (46%) were classified as (−) responders (GA < 1, p=0.01) (Figure 1D).

The relative and baseline variance methods produced similar proportions of responders when performed irrespective of the group or individual level analysis. Generally, more participants were classified as non-responders for the GA data (40-58%) than the MEP_{pp} data (29-52%). Moreover, the baseline variance method resulted in more non-responders (46-58%) than the relative method (29-40%).

Discussion

The present study followed a typical intervention design where TMS MEP data are collected at baseline and then again at pre-defined times following the intervention. However, in the present study the participants were not exposed to an intervention. Therefore, subject to normal MEP variability, the 'post-intervention' data sets would not be expected to be different from baseline. As expected, parametric statistics performed on this continuous data set revealed no significant difference with time. However, when the data were subjected to responder analyse between 21-71% of the participants were classified as responders, thus revealing a large number of false positives.

The responder analysis has been used throughout clinical medicine and psychology literature because it simplifies the analysis and interpretation of experimental results; with proponents of the analysis highlighting its usefulness in clinical decision making [7]. However, methodologists have argued for more than two decades that the dichotomisation of continuous variables is not valid for hypothesis testing [1, 9-14, 16-18]. The dichotomisation of continuous variables results in significant loss of information (~35-50% depending on the distribution of the data), reduced power of the statistical tests, high probability of Type I error, biased parameter estimates and erroneously small variances (for detailed discussion see: [1, 13, 16]).

The specific objective of the present study was to investigate the Type I error associated with responder analyses when MEP data are used to classify participants. In general, we observed substantial Type I errors with all of the responder analyses methods. Our results suggest that at best, 20% of the participants who have been classified as responders will have been classified erroneously. It may be valid to use a responder analysis to compare an intervention with a control group, but the specific response rates may be over-estimated.

Acknowledgements

We would like to thank Mr. Chris W. Wright for his assistance with the data collection, and Dr. Allan Clark for valuable discussions with respect to the responder analysis and statistical processing.

Conflict of Interest:

We have no conflicts of interest to declare.

Ethical approval:

The study was carried out in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki) and informed consent was obtained from all participants recruited to the study. Ethical approval for the study was granted from the University of Birmingham's Science, Technology, Engineering and Mathematics ethics committee (ERN_13-0701).

Funding:

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

[1] Altman DG, Royston P. The cost of dichotomising continuous variables. Bmj 2006;332(7549):1080.

[2] Rossini PM, Burke D, Chen R, Cohen LG, Daskalakis Z, Di Iorio R, et al. Non-invasive electrical and magnetic stimulation of the brain, spinal cord, roots and peripheral nerves: Basic principles and procedures for routine clinical and research application. An updated report from an I.F.C.N. Committee. Clin Neurophysiol 2015;126(6):1071-107.

[3] Pellegrini M, Zoghi M, Jaberzadeh S. Cluster analysis and subgrouping to investigate inter-individual variability to non-invasive brain stimulation: a systematic review. Reviews in the neurosciences 2018.

[4] Schmidt S, Bathe-Peters R, Fleischmann R, Ronnefarth M, Scholz M, Brandt SA. Nonphysiological factors in navigated TMS studies; confounding covariates and valid intracortical estimates. Hum Brain Mapp 2015;36(1):40-9.

[5] Kiers L, Cros D, Chiappa KH, Fang J. Variability of Motor Potentials-Evoked by Transcranial Magnetic Stimulation. Electroencephalography and Clinical Neurophysiology 1993;89(6):415-23.

[6] Senn S, Julious S. Measurement in clinical trials: A neglected issue for statisticians? Statistics in Medicine 2009;28(26):3189-209.

[7] Snapinn SM, Jiang Q. Responder analyses and the assessment of a clinically relevant treatment effect. Trials 2007;8(1):31.

[8] Iacobucci D, Popovich DL, Bakamitsos GA, Posavac SS, Kardes FR. Three Essential Analytical Techniques for the Behavioral Marketing Researcher: Median Splits, Mean-Centering, and Mediation Analysis. Foundations and Trends® in Marketing 2015;9(2):83-174.

[9] Weinberg CR. How bad is categorization? Epidemiology 1995;6(4):345-7.

[10] Senn S. Disappointing dichotomies. Pharm Stat 2003;2(4):239-40.

[11] Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. Stat Med 2006;25(1):127-41.

[12] Metze K. Dichotomization of continuous data--a pitfall in prognostic factor studies. Pathol Res Pract 2008;204(3):213-4.

[13] Maxwell SE, Delaney HD. Bivariate Median Splits and Spurious Statistical Significance. Psychol Bull 1993;113(1):181-90.

[14] MacCallum RC, Zhang S, Preacher KJ, Rucker DD. On the practice of dichotomization of quantitative variables. Psychol Methods 2002;7(1):19-40.

[15] Lewis JA. In defence of the dichotomy. Pharm Stat 2004;3(2):77-9.

[16] Fedorov V, Mannino F, Zhang R. Consequences of dichotomization. Pharm Stat 2009;8(1):50-61.

[17] DeCoster J, Iselin AM, Gallucci M. A conceptual and empirical examination of justifications for dichotomization. Psychol Methods 2009;14(4):349-66.

[18] Cohen J. The cost of dichotomization. Applied Psychological Measurement 1983;7(3):249-53.

[19] Julie R. Irwin, McClelland GH. Negative Consequences of Dichotomizing Continuous Predictor Variables. Journal of Marketing Research 2003;40(3):366-71.

[20] Chew T, Ho KA, Loo CK. Inter- and Intra-individual Variability in Response to Transcranial Direct Current Stimulation (tDCS) at Varying Current Intensities. Brain stimulation 2015.

[21] López-Alonso V, Cheeran B, Fernández-del-Olmo M. Relationship between non-invasive brain stimulation-induced plasticity and capacity for motor learning. Brain stimulation 2015;8(6):1209-19.

[22] Lopez-Alonso V, Cheeran B, Rio-Rodriguez D, Fernandez-Del-Olmo M. Inter-individual variability in response to non-invasive brain stimulation paradigms. Brain stimulation 2014;7(3):372-80.

[23] Puri R, Hinder MR, Canty AJ, Summers JJ. Facilitatory non-invasive brain stimulation in older adults: the effect of stimulation type and duration on the induction of motor cortex plasticity. Experimental brain research 2016;234(12):3411-23.

[24] Puri R, Hinder MR, Fujiyama H, Gomez R, Carson RG, Summers JJ. Duration-dependent effects of the BDNF Val66Met polymorphism on anodal tDCS induced motor cortex plasticity in older adults: a group and individual perspective. Frontiers in aging neuroscience 2015;7:107.

[25] Strube W, Bunse T, Nitsche MA, Nikolaeva A, Palm U, Padberg F, et al. Bidirectional variability in motor cortex excitability modulation following 1 mA transcranial direct current stimulation in healthy participants. Physiol Rep 2016;4(15).

[26] Wiethoff S, Hamada M, Rothwell JC. Variability in response to transcranial direct current stimulation of the motor cortex. Brain stimulation 2014;7(3):468-75.

[27] Goldsworthy MR, Vallence AM, Yang R, Pitcher JB, Ridding MC. Combined transcranial alternating current stimulation and continuous theta burst stimulation: a novel approach for neuroplasticity induction. Eur J Neurosci 2016;43(4):572-9.

[28] Hamada M, Murase N, Hasan A, Balaratnam M, Rothwell JC. The role of interneuron networks in driving human motor cortical plasticity. Cerebral cortex 2013;23(7):1593-605.

[29] Hinder MR, Goss EL, Fujiyama H, Canty AJ, Garry MI, Rodger J, et al. Inter- and Intra-individual variability following intermittent theta burst stimulation: implications for rehabilitation and recovery. Brain stimulation 2014;7(3):365-71.

[30] Labruna L, Jamil A, Fresnoza S, Batsikadze G, Kuo MF, Vanderschelden B, et al. Efficacy of Anodal Transcranial Direct Current Stimulation is Related to Sensitivity to Transcranial Magnetic Stimulation. Brain stimulation 2016;9(1):8-15.

[31] Lopez-Alonso V, Fernandez-Del-Olmo M, Costantini A, Gonzalez-Henriquez JJ, Cheeran B. Intra-individual variability in the response to anodal transcranial direct current stimulation. Clin Neurophysiol 2015.

[32] Muller-Dahlhaus JF, Orekhov Y, Liu Y, Ziemann U. Interindividual variability and age-dependency of motor cortical plasticity induced by paired associative stimulation. Experimental brain research 2008;187(3):467-75.

[33] Nakamura K, Groiss SJ, Hamada M, Enomoto H, Kadowaki S, Abe M, et al. Variability in Response to Quadripulse Stimulation of the Motor Cortex. Brain stimulation 2016;9(6):859-66.

[34] Muller-Dahlhaus F, Lucke C, Lu MK, Arai N, Fuhl A, Herrmann E, et al. Augmenting LTP-Like Plasticity in Human Motor Cortex by Spaced Paired Associative Stimulation. Plos One 2015;10(6):e0131020.

[35] Nettekoven C, Volz LJ, Leimbach M, Pool EM, Rehme AK, Eickhoff SB, et al. Inter-individual variability in cortical excitability and motor network connectivity following multiple blocks of rTMS. NeuroImage 2015;118:209-18.

[36] Strube W, Bunse T, Malchow B, Hasan A. Efficacy and interindividual variability in motor-cortex plasticity following anodal tDCS and paired-associative stimulation. Neural plasticity 2015;2015:530423.

[37] Ammann C, Lindquist MA, Celnik PA. Response variability of different anodal transcranial direct current stimulation intensities across multiple sessions. Brain stimulation 2017;10(4):757-63.

[38] Hanajima R, Tanaka N, Tsutsumi R, Enomoto H, Abe M, Nakamura K, et al. The effect of age on the homotopic motor cortical long-term potentiation-like effect induced by quadripulse stimulation. Experimental brain research 2017;235(7):2103-8.

[39] Simeoni S, Hannah R, Sato D, Kawakami M, Rothwell J, Simeoni S, et al. Effects of Quadripulse Stimulation on Human Motor Cortex Excitability: A Replication Study. Brain stimulation 2016;9(1):148-50.

[40] Tremblay S, Hannah R, Rawji V, Rothwell JC. Modulation of iTBS after-effects via concurrent directional TDCS: A proof of principle study. Brain stimulation 2017;10(4):744-7.

[41] Tremblay S, Larochelle-Brunet F, Lafleur LP, El Mouderrib S, Lepage JF, Theoret H. Systematic assessment of duration and intensity of anodal transcranial direct current stimulation on primary motor cortex excitability. Eur J Neurosci 2016;44(5):2184-90.

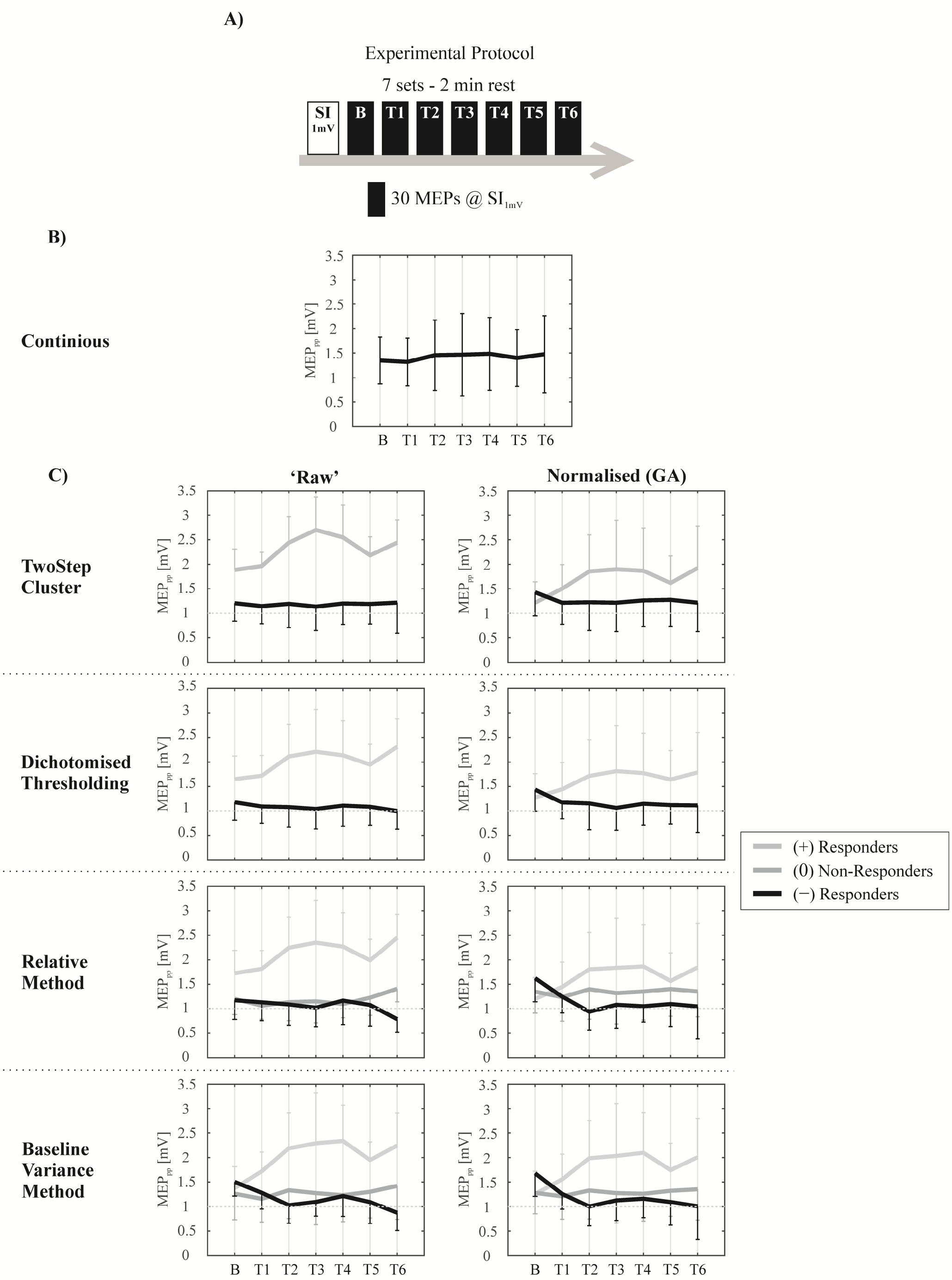
Figure/Table Legends

Figure 1: Responder/non-responder analysis across TMS MEP testing sets. (A) Seven sets of 30 MEPs were acquired at a stimulation intensity selected to producing a mean 1 mV peak-to-peak MEP amplitude (mean SI_{1mV}: 56 ± 10% of maximum stimulator output). The first set was considered the baseline to which the remaining six sets would be compared. (B) MEP_{pp} amplitude across all participants and all sets. No effect of set on MEP_{pp} amplitude observed for these data. (C) MEP_{pp} amplitude is shown across each of the seven data sets, with the participants di- or tricotomised using a two-step cluster analysis, dichotomised thresholding, relative threshold method or baseline variance method on a group level. In this way participants are classified as either (+) responders (light grey lines), showing an increase in MEP_{pp} amplitude compared to baseline, (0)- or non-responders (grey lines), no change in MEP_{pp} amplitude across SET, or (−) responders (black lines), a decrease in absolute MEP_{pp} across SET. The left column presents results when the classification was based on absolute MEP_{pp} data, the right column when based on GA data. All data are presented as Mean ± S.D. The number of participants for each group can be found in Table 1.

Table 1: Overview of results for subgrouping participants according to four methods for both normalised grand average (GA) data as well as non-normalised 'raw' MEP_{pp} data: (1). SPSS Two-Step Cluster analysis; (2) Relative % change with respect to baseline; (3) Dichotomised thresholding: a predefined fixed threshold; and (4) Change relative to the variance of the baseline set. A subgroup of participants is classified as positive responders (+) or negative responders (−), when there is a significant increase or decrease across SET respectively. Non-responders (0) are those participants in the group with no significant change in MEP_{pp} amplitude across SET. For some methods participants were subgrouped both on a threshold defined on an individual (Indv) basis as well as on a group (Gr) level. The %0 column highlights the proportion of non-responders. Results are shown with analysis performed on normalised grand average (GA) data and non-normalised absolute MEP_{pp} data.

Subgrouping Method		# Participants +	0	−	%0	Mixed RM-ANOVA			OneWay RM-ANOVA +		0		−	
Normalised GA data														
Two Step Cluster		19	33	-	63%	SET: SET×GROUP :	$F_{(4.83,241.71)} = 3.43^{GG}$ $F_{(4.83,241.71)} = 8.40^{GG}$	**p<0.01** **p<0.01**	$F_{(3.66,65.93)} = 5.97^{GG}$	**p<0.01**	$F_{(5.02,160.76)} = 1.65^{GG}$	p=0.15	-	-
Threshold Dichotomisation		28	-	24	-	SET: SET×GROUP:	$F_{(4.88,243.73)} = 1.05^{GG}$ $F_{(4.88,243.73)} = 8.14^{GG}$	p=0.39 **p<0.01**	$F_{(3.96,106.90)} = 6.33^{GG}$	**p<0.01**	-	-	$F_{(6,138)} = 2.78$	**p=0.01**
Relative		20	21	11	40%	SET: SET×GROUP:	$F_{(4.66,228.43)} = 0.49^{GG}$ $F_{(9.32,228.43)} = 5.63^{GG}$	p=0.77 **p<0.01**	$F_{(3.69,70.22)} = 5.91^{GG}$	**p<0.01**	$F_{(4.41,88.25)} = 0.64^{GG}$	p=0.65	$F_{(6,60)} = 4.59$	**p<0.01**
Baseline Variance	Gr	15	27	10	52%	SET: SET×GROUP:	$F_{(4.63,226.73)} = 0.97^{GG}$ $F_{(9.25,226.73)} = 6.08^{GG}$	p=0.43 **p<0.01**	$F_{(6,84)} = 6.59$	**p<0.01**	$F_{(4.59,119.21)} = 0.52^{GG}$	p=0.74	$F_{(6,54)} = 4.29$	**p<0.01**
	Indv	13	30	9	58%	SET: SET×GROUP:	$F_{(4.57,223.80)} = 1.24^{GG}$ $F_{(9.14,223.80)} = 6.59^{GG}$	p=0.29 **p<0.01**	$F_{(3.11,37.37)} = 6.68^{GG}$	**p<0.01**	$F_{(4.56,132.17)} = 0.48^{GG}$	p=0.77	$F_{(6,48)} = 4.58$	**p=0.01**
Non-normalised MEP_{pp} data														
Two Step Cluster		11	41	-	79%	SET: SET×GROUP:	$F_{(6,300)} = 4.74$ $F_{(6,300)} = 4.96$	**p<0.01** **p<0.01**	$F_{(6,60)} = 4.50$	**p<0.01**	$F_{(6,240)} = 0.26$	p=0.96	-	-
Threshold Dichotomisation	Gr	33	19	-	37%	SET: SET×GROUP:	$F_{(6,300)} = 3.23$ $F_{(6,300)} = 6.69$	**p<0.01** **p<0.01**	$F_{(3.65,65.65)} = 5.80^{GG}$	**p<0.01**	$F_{(6,192)} = 0.88$	p=0.51	-	-
	Indv	24	-	28	-	SET: SET×GROUP:	$F_{(4.87,243.27)} = 1.06^{GG}$ $F_{(4.87,243.27)} = 7.44^{GG}$	p=0.38 **p<0.01**	$F_{(3.81,102.80)} = 5.80^{GG}$	**p<0.01**	-	-	$F_{(6,138)} = 2.57$	**p=0.02**
Relative	Gr	16	15	21	29%	SET: SET×GROUP:	$F_{(6,294)} = 2.12$ $F_{(12,294)} = 5.23$	p=0.05 **p<0.01**	$F_{(3.62,52.85)} = 5.00^{GG}$	**p<0.01**	$F_{(6,84)} = 2.43$	**p=0.03**	$F_{(6,120)} = 2.91$	**p=0.01**
	Indv	17	19	16	37%	SET: SET×GROUP:	$F_{(4.73,231.96)} = 1.47^{GG}$ $F_{(9.47,231.96)} = 6.63^{GG}$	p=0.20 **p<0.01**	$F_{(3.41,54.60)} = 6.44^{GG}$	**p<0.01**	$F_{(6,108)} = 1.70$	p=0.13	$F_{(6,90)} = 4.13$	**p<0.01**
Baseline Variance	Gr	12	27	13	52%	SET: SET×GROUP:	$F_{(4.75,232.84)} = 2.16^{GG}$ $F_{(9.50,232.84)} = 6.95^{GG}$	p=0.06 **p<0.01**	$F_{(3.10,34.09)} = 6.32^{GG}$	**p<0.01**	$F_{(6,156)} = 1.51$	p=0.18	$F_{(6,72)} = 4.77$	**p<0.01**
	Indv	13	24	15	46%	SET: SET×GROUP:	$F_{(4.79,234.73)} = 2.49^{GG}$ $F_{(9.58,234.73)} = 6.08^{GG}$	**p=0.03** **p<0.01**	$F_{(3.11,37.37)} = 6.68^{GG}$	**p<0.01**	$F_{(6,138)} = 0.95$	p=0.36	$F_{(6,84)} = 3.41$	**p<0.01**

Table 1


A)
Experimental Protocol
7 sets - 2 min rest
SI 1mV
B
T1
T2
T3
T4
T5
T6
30 MEPs @ SI1mV
B)
Continious
MEPpp [mV]
C)
'Raw'
Normalised (GA)
TwoStep Cluster
Dichotomised Thresholding
Relative Method
Baseline Variance Method
(+) Responders
(0) Non-Responders
(−) Responders