

# Translating Research on Myoelectric Control into Clinics – Are the Performance Assessment Methods Adequate?

1 Ivan Vujaklija<sup>1\*</sup>, Aidan D. Roche<sup>2</sup>, Timothy Hasenoehrl<sup>3</sup>, Agnes Sturma<sup>2,4</sup>, Sebastian Amsuess<sup>5</sup>,  
2 Dario Farina<sup>6</sup> & Oskar C. Aszmann<sup>2,7</sup>

3 <sup>1</sup> Clinic for Trauma Surgery, Orthopaedic Surgery and Plastic Surgery, Research Department for  
4 Neurorehabilitation Systems, University Medical Centre Göttingen, GERMANY

5 <sup>2</sup>Christian Doppler Laboratory for Restoration of Extremity Function, Medical University of Vienna,  
6 Vienna, AUSTRIA

7 <sup>3</sup>Department of Physical Medicine and Rehabilitation, Medical University of Vienna, Vienna,  
8 AUSTRIA

9 <sup>4</sup>Master Degree Program “Health Assisting Engineering”, University of Applied Sciences FH  
10 Campus Wien, Vienna, AUSTRIA

11 <sup>5</sup>Otto Bock Healthcare Products GmbH, Vienna, AUSTRIA

12 <sup>6</sup>Department of Bioengineering, Imperial College London, London, UK

13 <sup>7</sup>Division of Plastic and Reconstructive Surgery, Department of Surgery, Medical University of  
14 Vienna, Vienna, AUSTRIA

15 **\* Correspondence:**

16 Ivan Vujaklija

17 [ivan.vujaklija@bccn.uni-goettingen.de](mailto:ivan.vujaklija@bccn.uni-goettingen.de)

18 **Keywords:** myoelectric prosthesis, prosthetic assessment, myoelectric control, SHAP, box and  
19 blocks

20 **Funding:** This work was supported by the European Union’s Horizon 2020 research and innovation  
21 program under grant agreement number 687795 (project INPUT) and by the Christian Doppler  
22 Research Foundation of the Austrian Federal Ministry of Science, Research and Economy

23 **Competing interests:** The authors declare that the research was conducted in the absence of any  
24 commercial or financial relationships that could be construed as a potential conflict of interest.

## 25 Abstract

26 Missing an upper limb dramatically impairs daily-life activities. Significant efforts in overcoming the  
 27 issues arising from this disability have been made in both academia and industry, although their  
 28 clinical outcome is still limited. Translation of prosthetic research into clinics has been challenging  
 29 because of the difficulties in meeting the necessary requirements of the market. In this perspective,  
 30 we focus on myocontrol algorithms for upper limb prostheses and we emphasize that one relevant  
 31 factor determining the relatively small clinical impact of these methods is the limit of commonly  
 32 used laboratory performance metrics. The laboratory conditions, in which the majority of the  
 33 solutions are being evaluated, fail to sufficiently replicate real-life challenges. We qualitatively  
 34 substantiate this argument with data from seven transradial amputees. Their ability to control a  
 35 myoelectric prosthesis was tested by measuring the accuracy of offline EMG signal classification, as  
 36 a typical laboratory performance metrics, as well as by clinical scores when performing standard tests  
 37 of daily living. Despite all subjects reached relatively high classification accuracy offline, their  
 38 clinical scores were largely different and were not strongly predicted by classification accuracy. As  
 39 argued in previous reports, we reinforce the suggestion to test myocontrol systems using clinical tests  
 40 on amputees, fully fitted with sockets and prostheses highly resembling the systems they would use  
 41 in daily living, as evaluation benchmark. Agreement on this level of testing for systems developed in  
 42 research laboratories would facilitate clinically relevant progresses in this field.

## 43 1 Introduction

44 Recent progresses in active prosthesis control for the upper limb include the introduction of novel  
 45 control approaches (Amsuess et al., 2015; Jiang et al., 2014a; Scheme and Englehart, 2011), sensor  
 46 types and sensor fusion algorithms (Cipriani et al., 2014; Dosen et al., 2010; Nissler et al., 2016;  
 47 Ortenzi et al., 2015; Weir et al., 2003), surgical techniques (Aszmann et al., 2015; Kuiken et al.,  
 48 2004), as well as advanced hardware (Catalano et al., 2014; Cipriani et al., 2011; Grebenstein et al.,  
 49 2011). Nonetheless, the impact of these advances towards improving the experience of the everyday  
 50 end user is still limited. The discrepancy between myoelectric solutions that academia develops and  
 51 promotes, and the systems available on the market is indeed substantial. This issue has been  
 52 previously discussed (e.g. (Hill et al., 2009; Jiang et al., 2012)) and relates to the conditions in which  
 53 new methods are tested.

54 The necessity for testing prosthetic solutions in a greater number of amputees than currently done is a  
 55 widely recognized problem. Moreover, the tests used often fail to include clinically relevant metrics.  
 56 Performance metrics prevalent in laboratory research may be poorly associated to the clinical  
 57 outcome (Jiang et al., 2014b; Ortiz-Catalan et al., 2015; Simon et al., 2011). Here, we reinforce these  
 58 arguments to further substantiate the relevance of this problem.

59 Transferring myoelectrical systems developed in the laboratory to clinical settings is a challenge that  
 60 requires multidisciplinary efforts. Clinical tests, although not ideal, offer the most realistic prediction  
 61 of the system performance in the daily use. These tests account for several of the challenges that  
 62 laboratory-based assessment methodologies tend to neglect. For example, noiseless laboratory-based  
 63 evaluation platforms fail to account for the end effector loads, poor socket fitting, and sweating.

64 In this perspective, we briefly introduce evaluation methods regularly applied for prosthetics use,  
 65 with a focus on offline approaches and some selected clinical measures. Moreover, we provide  
 66 experimental data on seven conventional myoelectric users. The literature review and the

experimental data are limited to the primary aim of providing our view on assessment procedures for myocontrol and suggestions for their improvement.

## 2 Performance evaluation

Laboratory-based techniques and tests for measuring the performance in controlling a myoelectric interface are numerous and, in case of offline techniques, have been mainly derived or adapted from the machine learning literature. On the other hand, initially, clinicians have mostly adapted established hand and arm impairment assessment tools to the evaluation of functional recovery with prostheses. In recent years, clinical measures have however been introduced to target specifically the amputee patient population.

### 2.1 Laboratory metrics

Evaluation and assessment techniques for myocontrol in strictly laboratory conditions can be broadly divided in two groups – those quantifying the system performance through offline metrics and those based on online assessments using virtual prostheses or games.

Depending on the type of the evaluated control algorithm, offline performance is most commonly assessed using either classification accuracy (Ortiz-Catalan et al., 2013) or the  $R^2$  error with respect to a given prompt (Ameri et al., 2014). The first approach relies on the number of correct estimates that the tested classifier makes, given the new, unseen data. The second compares the estimated command with respect to a reference cue. It has been shown that offline analysis fails to reflect the performance exhibited in online scenarios (Jiang et al., 2014b; Ortiz-Catalan et al., 2015). This is classically attributed to the fact that offline analyses do not account for adaptation of the user to non-stationary signal features.

Several virtual reality (VR) based assessment benches have been proposed in recent years. These systems simulate the online use of the prosthesis, at various levels of abstraction, while still being research-based settings. They offer the advantage of not dealing with the full implementation of the system, avoiding the challenges of socket design and hardware implementations. These VR systems are sometimes abstract with respect to the intended control (Ison et al., 2015) and commonly consist in steering a computer avatar in multiple directions to assess the performance when controlling specific DoFs. Alternatively, computer games can be presented to the users, e.g. controlling a cursor to hit targets on a computer screen (Ameri et al., 2014; Jiang et al., 2014a). Finally, users can also be instructed to move a virtual arm into a target posture (Simon et al., 2011), as a part of an elaborate VR test bench.

The online systems are superior to the offline evaluations since they include the user in the loop and therefore account for his/her adaptation to the system. Parameters such as completion rate, path efficiency, number of overshoots or throughput, provide a solid quantitative evaluation of online performance. Further, (Fimbel et al., 2006) introduced the Fitts' law (Fitts, 1954) in evaluating myocontrol. Through some iterations (Jiang et al., 2014b; Park et al., 2008; Scheme and Englehart, 2013), a single statistical measure has been proposed to characterize a myocontroller online. Nonetheless, even if some of these test benches offer realistic testing scenarios, they have limitations. For example, weight bearing by the prosthesis and stump dynamics causing pressure changes within the socket fitting are important realistic factors of influence (Daly et al., 2014), not included in these tests. On the other hand, VR systems have found relevant applications in patient training (Roche et al., 2015; Sturma et al., 2015) and can be combined with table-top prosthetics (Stubblefield et al., 2011).

## 2.2 Clinical metrics

Clinical and rehabilitation specialists rely on a set of tests as well as questioners for assessing the user performance in myoelectric control. These tests prompt users to manipulate a variety of objects and to execute tasks mimicking those of daily living. The majority of the clinical scores validate the capability of executing certain tasks by quantifying the completion time. A battery of clinical tests requires the presence of certified examiners.

The box and blocks (B&B) test is one of the simplest and most commonly used clinical tests for evaluating the severity of upper limb deficiency. It consists of transporting, one by one, a number of square wooden blocks over a barrier using the prosthesis. The quantitative performance index for this test is the number of blocks that are successfully moved in a fixed time interval (usually 1 min). This test is simple to implement but only focuses on a limited number of DoFs and requires a minimal skill by the user.

The Clothes Pin Relocation Test (CPRT) requires the user to move a set of clothes pins of various resistances from a horizontal to a vertical bar. Since this is primarily a rehabilitation tool, the exact evaluation procedure has not been defined yet. However, most therapists use four clothespins of different resistances (1, 2, 4 and 8 lbs.) and request the subjects to relocate them from the lowest horizontal bar to the most convenient position on the vertical bar. The time of execution is then recorded from the starting neutral position to the final neutral position. The CPRT requires activation of several DoFs, although it often promotes compensatory movements which are not accounted for in the final outcome score.

The Southampton Hand Assessment Protocol (SHAP) is one of the most elaborate hand impairment evaluation tests (Light et al., 2002). It consists of 26 individual tasks that include six grips and their combinations. It can be separated into abstract object handling and execution of activities of daily living (ADL). Its final outcome is a number in the range 0-100, where 0 corresponds to absence of hand function and 100 to a healthy hand function, which mainly reflects the time needed for completing the tasks. SHAP is a very elaborate hand assessment tool and therefore it is also lengthy and tiring for the patients, especially those with limited capabilities. Additionally, it mainly quantifies the time needed for execution and does not account for the way in which the tasks are completed.

The Action Research Arm Test (ARAT) is a global arm function assessment procedure. It is divided into four sub-scales – grasp, grip, pinch and gross movement – that evaluate abstract object manipulation strategies. The maximum ARAT score is 57, corresponding to normal upper limb function. This score is based on the opinion of certified examiners that rate the quality of execution of each task on a scale from 0 (cannot perform) to 3 (performs normally).

In addition to the above, several other clinical tests and questioners have been devised targeting different functions and ways of assessing upper limbs, such as the Assessment of Capacity for Myoelectric Control (ACMC) (Hermansson et al., 2005) and the Jebsen-Taylor Test of Hand Function (JTHF) (Davis Sears and Chung, 2010). The former is a clinical evaluation test specifically tailored for myocontrol and, although it suffers of a strong subjective component and it has not yet received wide recognition, may be a promising evaluation tool.

## 3 Experiments

We provide data on amputees that compare the accuracy estimated offline, for one of the classic control schemes developed over the past decades, with clinical scores. These data serve the purpose

of reinforcing the conclusion for the need of clinical tests in an exemplary way. Therefore, the experiment and results do not aim at providing general conclusions on all myocontrol schemes and evaluation methods but rather support the view presented in this perspective.

Seven male transradial myoelectric users agreed to participate. They were all fit with custom-made sockets and with the Michelangelo hand (Ottobock Healthcare GmbH, Austria) with additional wrist rotation and flexion/extension units. The study was performed in accordance with the recommendations of the local ethics board of the Medical University of Vienna (Ethics Commission number 1044/2015), with written informed consent from all subjects. All given consents are in accordance with the Declaration of Helsinki.

The control of the prosthesis was based on the common spatial pattern (CSP) based classifier, as described by (Amsuess et al., 2015). The EMG signals were recorded with 8 bipolar surface electrodes (Otto Bock raw signal electrodes 13E200=50AC). The control system allowed the subjects to access seven prosthetic functions – wrist flexion/extension, wrist pronation/supination, hand open, pinch, and key grip. All the motions were recorded in three arm positions (relaxed, fully extend arm in front of the ipsilateral shoulder and fully extended arm across the contralateral shoulder) and at three forces (30%, 60% and 90% of the EMG level at maximum voluntary contraction force, for each motion) while wearing the full prosthetic fitting. For offline accuracy assessment, the classifier was trained by data collected in only one arm position and tested against the remaining two data sub-sets. The average of the three scores was the reference performance of the subject. The entire data set was used for training the same CSP classifier that allowed execution of the B&B and SHAP tests. These particular clinical tests have been representatively chosen since they cover a wide range of assessment goals while being entirely objective. Additionally, these two tests have been widely recognized and familiar to academic and industry-based developers as well as clinical experts.

The performance scores in both offline and clinical tests are presented in Figure 1. The offline classification accuracies are slightly lower than in other studies (Ahsan et al., 2010; Liu et al., 2013) because of the different arm positions used for training and testing as well as the full prosthetic fitting which is not usual in offline evaluation studies. Although with these choices we have maximized the prediction capacity of offline indexes for clinical scores, still the clinical scores did not strongly correlate with the offline performance measures. For example, there were two patients who achieved a similar SHAP score just below 40 whereas they showed substantially different classification accuracies of < 70% and > 85% (Figure 1A). Similarly, two patients who had very similar classification accuracies of 70-75% had SHAP scores of 27 and 47 (Figure 1A). The B&B test requires less skill to be performed than the SHAP. However, the B&B score was even less associated to the offline classification than the SHAP (Figure 1B). For example, subjects with an offline accuracy >95% performed very differently in this test (Figure 1B). Furthermore, when considering strictly the hand movements – hand open, fine pinch and key grip - that are primarily used for this test, the mismatch between this test and offline performance was even more substantial. This was observed consistently in all patients but it is shown representatively for only two patients in Figure 2. For these patients, the average classification rate across the three hand motions was 89% and 79% whereas the transferred blocks (score of the B&B) were 5 and 12, respectively.

When the offline evaluation was performed by using data collected without wearing the prosthesis and tested on the same arm position as the training, as more commonly done in laboratory tests (e.g., (Englehart et al., 1999; Hargrove et al., 2009; Li et al., 2010; Ortiz-Catalan et al., 2014b)), the resulting offline classification rates were high and comparable to those reported in the literature (>90% on average). However, once fully fitted, the majority of patients were unable to successfully

conclude the clinical evaluations without retraining, indicating that the classic offline evaluation procedure performed in several research studies does not provide strongly relevant clinical information.

## 4 Discussion

Abandonment rates among upper limb myoelectric prosthetic users are still very high (Burrough and Brook, 1985; Glynn et al., 1986; Østlie et al., 2012). At the same time, research efforts have provided several new solutions for myocontrol that have been proven to be highly functional strictly under laboratory conditions. The negligible transfer from research to real world applications likely depends, as one of the most relevant factors, on an insufficient level of evaluation procedures.

Using novel prototypes of myoelectric systems in daily life would provide the ultimate assessment but this strategy would neither be safe nor always legal. Daily-usage tests often require a full development with proper certifications. The COAPT (Coapt LLC, 2016) is one of the first systems that has reached this level of testing. Clinical evaluations at earlier stages are a compromise between laboratory conditions and real-life tests. Although not perfect, clinical tests are closer to the conditions of interest for the users than offline assessments or online tests using virtual prostheses. Here we have presented an example of this dissociation on a small sample of amputee and focusing on offline metrics, for demonstration purposes. We have compared clinical scores with offline indexes of performance extracted in the most realistic offline conditions (patients wearing their own prosthesis, training and test sets on different arm postures). Despite these conditions rarely being met in the offline studies, the prediction capacity for clinical outcome was not strong. On the other hand, when the offline indexes were obtained in more common laboratory conditions without the prosthesis and for the same arm posture for test and training, the clinical information they provided was almost null. Further extrapolating, it is obvious that an offline analysis performed in these simple conditions and, in addition, on able-bodied individuals instead of patients, cannot be of strong clinical value. While we are fully aware that in the initial evaluation of a new myocontrol scheme the strict laboratory tests on healthy individuals is extremely valuable and needed for assessing the basic working principles, there is also the need to make efforts in continuing the evaluations of promising algorithms in clinically-relevant settings. We believe that the evaluation stages after the laboratory level have had a slower progress, and less academic interest, in the past with respect to the proposal of new algorithms.

Considering the discrepancy presented in the literature and supported here with the representative data shown, it is necessary that novel myoelectric systems that passed laboratory testing are then fully clinically evaluated for assessing their performance. For this purpose, researchers and clinicians should jointly devise a standardized testing framework for quantitatively and qualitatively assessing the performance of upper limb prosthetic devices and their users to boost the process of commercialization and availability for the patients. This need does not only relate to the feed-forward control aspects, on which we focused here, but also fully closed-loop systems that include sensory feedback integration (González and Yu, 2009; Jorgovanovic et al., 2014; Ortiz-Catalan et al., 2014a).

## 5 References

- Ahsan, M. R., Ibrahimy, M. I., and Khalifa, O. O. (2010). Advances in Electromyogram Signal Classification to Improve the Quality of Life for the Disabled and Aged People. *J. Comput. Sci.* 6, 706–715. doi:10.3844/jcssp.2010.706.715.
- Ameri, A., Scheme, E. J., Kamavuako, E. N., Englehart, K. B., and Parker, P. a. (2014). Real-time,

- 240 simultaneous myoelectric control using force and position-based training paradigms. *IEEE*  
241 *Trans. Biomed. Eng.* 61, 279–287. doi:10.1109/TBME.2013.2281595.
- 242 Amsuess, S., Vujaklija, I., Gobel, P., Roche, A., Graimann, B., Aszmann, O., et al. (2015). Context-  
243 Dependent Upper Limb Prosthesis Control for Natural and Robust Use. *IEEE Trans. Neural*  
244 *Syst. Rehabil. Eng.*, 1–1. doi:10.1109/TNSRE.2015.2454240.
- 245 Aszmann, O. C., Roche, A. D., Salminger, S., Paternostro-sluga, T., Herceg, M., Sturma, A., et al.  
246 (2015). Bionic reconstruction to restore hand function after brachial plexus injury : a case series  
247 of three patients. *Lancet* 6736, 1–7. doi:10.1016/S0140-6736(14)61776-1.
- 248 Burrough, S. F., and Brook, J. A. (1985). Patterns of Acceptance and Rejection of Upper Limb  
249 Prostheses. *Orthot. Prosthetics* 39, 40–47.
- 250 Catalano, M. G., Grioli, G., Farnioli, E., Serio, a., Piazza, C., and Bicchi, a. (2014). Adaptive  
251 synergies for the design and control of the Pisa/IIT SoftHand. *Int. J. Robot. Res.* 33, 768–782.  
252 doi:10.1177/0278364913518998.
- 253 Cipriani, C., Controzzi, M., and Carrozza, M. C. (2011). The SmartHand transradial prosthesis. *J.*  
254 *Neuroeng. Rehabil.* 8, 29. doi:10.1186/1743-0003-8-29.
- 255 Cipriani, C., Segil, J. L., Birdwell, J. A., and Weir, R. F. (2014). Dexterous control of a prosthetic  
256 hand using fine-wire intramuscular electrodes in targeted extrinsic muscles. *IEEE Trans. Neural*  
257 *Syst. Rehabil. Eng.* 22, 828–836. doi:10.1109/TNSRE.2014.2301234.
- 258 Coapt LLC (2016). Coapt engineering. Available at: <http://www.coaptengineering.com/> [Accessed  
259 April 29, 2016].
- 260 Daly, W., Voo, L., Rosenbaum-Chou, T., Arabian, A., and Boone, D. (2014). Socket Pressure and  
261 Discomfort in Upper-Limb Prostheses: A Preliminary Study. *JPO J. Prosthetics Orthot.* 26, 99–  
262 106. doi:10.1097/JPO.0000000000000021.
- 263 Davis Sears, E., and Chung, K. C. (2010). Validity and Responsiveness of the Jebsen–Taylor Hand  
264 Function Test. *J. Hand Surg. Am.* 35, 30–37. doi:10.1016/j.jhsa.2009.09.008.
- 265 Dosen, S., Cipriani, C., Kostić, M., Controzzi, M., Carrozza, M. C., and Popović, D. B. (2010).  
266 Cognitive vision system for control of dexterous prosthetic hands: experimental evaluation. *J.*  
267 *Neuroeng. Rehabil.* 7, 42. doi:10.1186/1743-0003-7-42.
- 268 Englehart, K., Hudgins, B., Parker, P. a, and Stevenson, M. (1999). Classification of the myoelectric  
269 signal using time-frequency based representations. *Med. Eng. Phys.* 21, 431–8. Available at:  
270 <http://www.ncbi.nlm.nih.gov/pubmed/10624739>.
- 271 Fimbel, E. J., Lemay, M., and Arguin, M. (2006). Speed-accuracy trade-offs in myocontrol. *Hum.*  
272 *Mov. Sci.* 25, 165–180. doi:10.1016/j.humov.2005.12.001.
- 273 Fitts, P. M. (1954). The Information Capacity of the Human Motor. *J. Exp. Biol.* 47, 381–391.  
274 doi:10.1037/h0055392.
- 275 Glynn, M. K., Galway, H. R., Hunter, G., and Sauter, W. F. (1986). Management of the upper-limb-

- 276 deficient child with a powered prosthetic device. *Clin. Orthop. Relat. Res.* 209, 202–5.  
277 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/3731596>.
- 278 González, J., and Yu, W. (2009). Multichannel audio aided dynamical perception for prosthetic hand  
279 biofeedback. *2009 IEEE Int. Conf. Rehabil. Robot. ICORR 2009*, 240–245.  
280 doi:10.1109/ICORR.2009.5209521.
- 281 Grebenstein, M., Albu-Schäffer, A., Bahls, T., Chalon, M., Eiberger, O., Friedl, W., et al. (2011).  
282 The DLR hand arm system. *Proc. - IEEE Int. Conf. Robot. Autom.*, 3175–3182.  
283 doi:10.1109/ICRA.2011.5980371.
- 284 Hargrove, L., Member, S., Li, G., Englehart, K., and Hargrove, L. (2009). Principal Components  
285 Analysis Preprocessing to improve Classification Accuracies in Pattern Recognition Based  
286 Myoelectric Control Corresponding author. 56, 1–28. doi:10.1109/TBME.2008.2008171.
- 287 Hermansson, L. N., Fisher, A. G., Bernspang, B., and Eliasson, A. C. (2005). Assessment of capacity  
288 for myoelectric control: a new Rasch-built measure of prosthetic hand control. *J. Rehabil. Med.*  
289 37, 166–71.
- 290 Hill, W., Stavadahl, Ø., Hermansson, L. N., Kyberd, P., Swanson, S., and Hubbard, S. (2009).  
291 Functional Outcomes in the WHO-ICF Model: Establishment of the Upper Limb Prosthetic  
292 Outcome Measures Group. *J. Prosthetics Orthot.* 21, 115–119.  
293 doi:10.1097/JPO.0b013e3181a1d2dc.
- 294 Ison, M., Vujaklija, I., Whitsell, B., Farina, D., and Artemiadis, P. (2015). High-Density  
295 Electromyography and Motor Skill Learning for Robust Long-Term Control of a 7-DoF Robot  
296 Arm. *IEEE Trans. Neural Syst. Rehabil. Eng.* to appear, 1–10.
- 297 Jiang, N., Dosen, S., and Farina, D. (2012). Myoelectric control of artificial limbs: is there the need  
298 for a change of focus? *IEEE Signal Process. Mag.*, 12–15. Available at:  
299 [http://www.lifesciences.ieee.org/articles/174-myoelectric-control-of-artificial-limbs-is-there-a-](http://www.lifesciences.ieee.org/articles/174-myoelectric-control-of-artificial-limbs-is-there-a-need-to-change-focus)  
300 [need-to-change-focus](http://www.lifesciences.ieee.org/articles/174-myoelectric-control-of-artificial-limbs-is-there-a-need-to-change-focus) [Accessed June 27, 2013].
- 301 Jiang, N., Rehbaum, H., Vujaklija, I., Graimann, B., and Farina, D. (2014a). Intuitive, Online,  
302 Simultaneous, and Proportional Myoelectric Control Over Two Degrees-of-Freedom in Upper  
303 Limb Amputees. *IEEE Trans. Neural Syst. Rehabil. Eng.* 22, 501–10.  
304 doi:10.1109/TNSRE.2013.2278411.
- 305 Jiang, N., Vujaklija, I., Rehbaum, H., Graimann, B., and Farina, D. (2014b). Is accurate mapping of  
306 EMG signals on kinematics needed for precise online myoelectric control? *IEEE Trans. Neural*  
307 *Syst. Rehabil. Eng.* 22, 549–58. doi:10.1109/TNSRE.2013.2287383.
- 308 Jorgovanovic, N., Dosen, S., Djovic, D. J., Krajoski, G., and Farina, D. (2014). Virtual Grasping:  
309 Closed-Loop Force Control Using Electrotactile Feedback. *Comput. Math. Methods Med.* 2014,  
310 1–13. doi:10.1155/2014/120357.
- 311 Kuiken, T. a, Dumanian, G. a, Lipschutz, R. D., Miller, L. a, and Stubblefield, K. a (2004). The use  
312 of targeted muscle reinnervation for improved myoelectric prosthesis control in a bilateral  
313 shoulder disarticulation amputee. *Prosthet. Orthot. Int.* 28, 245–53. Available at:  
314 <http://www.ncbi.nlm.nih.gov/pubmed/15658637>.



- 315 Li, G., Schultz, A. E., and Kuiken, T. A. (2010). Quantifying Pattern Recognition - Based  
316 Myoelectric Control of Multifunctional Transradial Prostheses. *IEEE Trans. Neural Syst.*  
317 *Rehabil. Eng.* 18, 185–192. doi:10.1109/TNSRE.2009.2039619.
- 318 Light, C. M., Chappell, P. H., and Kyberd, P. J. (2002). Establishing a Standardized Clinical  
319 Assessment Tool of Pathologic and Prosthetic Hand Function: Normative Data, Reliability, and  
320 Validity. *Arch. Phys. Med. Rehabil.* 83, 776–783. doi:10.1053/apmr.2002.32737.
- 321 Liu, L., Liu, P., Clancy, E. a., Scheme, E., and Englehart, K. B. (2013). Electromyogram whitening  
322 for improved classification accuracy in upper limb prosthesis control. *IEEE Trans. Neural Syst.*  
323 *Rehabil. Eng.* 21, 767–774. doi:10.1109/TNSRE.2013.2243470.
- 324 Nissler, C., Mouriki, N., and Castellini, C. (2016). Optical Myography: Detecting Finger Movements  
325 by Looking at the Forearm. *Front. Neurobot.* 10, 1–10. doi:10.3389/fnbot.2016.00003.
- 326 Ortenzi, V., Tarantino, S., Castellini, C., and Cipriani, C. (2015). Ultrasound imaging for hand  
327 prosthesis control: A comparative study of features and classification methods. *IEEE Int. Conf.*  
328 *Rehabil. Robot.* 2015–Sept, 1–6. doi:10.1109/ICORR.2015.7281166.
- 329 Ortiz-Catalan, M., Brånemark, R., and Håkansson, B. (2013). BioPatRec: A modular research  
330 platform for the control of artificial limbs based on pattern recognition algorithms. *Source Code*  
331 *Biol. Med.* 8, 11. doi:10.1186/1751-0473-8-11.
- 332 Ortiz-Catalan, M., Hakansson, B., and Branemark, R. (2014a). An osseointegrated human-machine  
333 gateway for long-term sensory feedback and motor control of artificial limbs. *Sci. Transl. Med.*  
334 6, 257re6-257re6. doi:10.1126/scitranslmed.3008933.
- 335 Ortiz-Catalan, M., Håkansson, B., and Brånemark, R. (2014b). Real-time and simultaneous control of  
336 artificial limbs based on pattern recognition algorithms. *IEEE Trans. Neural Syst. Rehabil. Eng.*  
337 22, 756–764. doi:10.1109/TNSRE.2014.2305097.
- 338 Ortiz-Catalan, M., Rouhani, F., Branemark, R., and Hakansson, B. (2015). Offline accuracy: A  
339 potentially misleading metric in myoelectric pattern recognition for prosthetic control. *Proc.*  
340 *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS* 2015–Novem, 1140–1143.  
341 doi:10.1109/EMBC.2015.7318567.
- 342 Østlie, K., Lesjø, I. M., Franklin, R. J., Garfelt, B., Skjeldal, O. H., and Magnus, P. (2012). Prosthesis  
343 rejection in acquired major upper-limb amputees: a population-based survey. *Disabil. Rehabil.*  
344 *Assist. Technol.* 7, 294–303. doi:10.3109/17483107.2011.635405.
- 345 Park, J., Bae, W., Kim, H., and Park, S. (2008). EMG — force correlation considering Fitts’ law.  
346 *2008 IEEE Int. Conf. Multisens. Fusion Integr. Intell. Syst.*, 644–649.  
347 doi:10.1109/MFI.2008.4648017.
- 348 Roche, A. D., Vujaklija, I., Amsüss, S., Sturma, A., Göbel, P., Farina, D., et al. (2015). A Structured  
349 Rehabilitation Protocol for Improved Multifunctional Prosthetic Control: A Case Study. *J. Vis.*  
350 *Exp.*, e52968. doi:10.3791/52968.
- 351 Scheme, E., and Englehart, K. (2011). Electromyogram pattern recognition for control of powered  
352 upper-limb prostheses: State of the art and challenges for clinical use. *J. Rehabil. Res. Dev.* 48,

643. doi:10.1682/JRRD.2010.09.0177.

Scheme, E. J., and Englehart, K. B. (2013). Validation of a selective ensemble-based classification scheme for myoelectric control using a three-dimensional fitts' law test. *IEEE Trans. Neural Syst. Rehabil. Eng.* 21, 616–623. doi:10.1109/TNSRE.2012.2226189.

Simon, A. M., Hargrove, L. J., Lock, B. a., and Kuiken, T. a. (2011). Target Achievement Control Test: Evaluating real-time myoelectric pattern-recognition control of multifunctional upper-limb prostheses. *J. Rehabil. Res. Dev.* 48, 619. doi:10.1682/JRRD.2010.08.0149.

Stubblefield, K., Finucane, S. B., Miller, L. A., and Lock, B. A. (2011). Training Individuals to Use Pattern Recognition to Control an Upper Limb Prosthesis. *Proc. 2011 MyoElectric Control. Prosthetics Symp.*, 1–4.

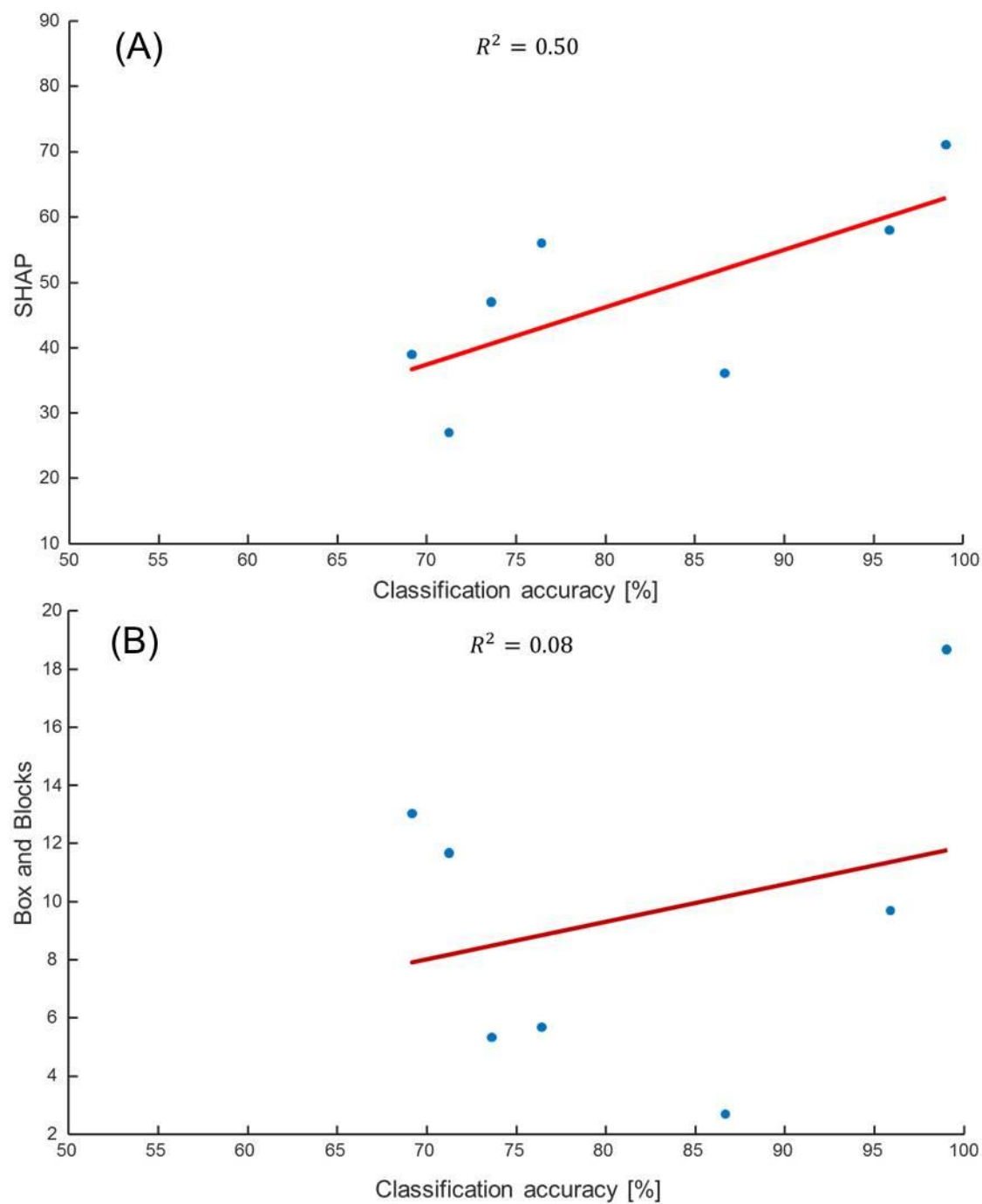
Sturma, A., Roche, A. D., Göbel, P., Herceg, M., Ge, N., Fialka-Moser, V., et al. (2015). A surface EMG test tool to measure proportional prosthetic control. *Biomed. Tech.* 60, 207–213. doi:10.1515/bmt-2014-0022.

Weir, R. F., Troyk, P. R., DeMichele, G., Kuiken, T., and Ku, T. (2003). Implantable myoelectric sensors (IMES) for upper-extremity prosthesis control- preliminary work. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* 25, 1562–1565. doi:10.1109/IEMBS.2003.1279658.

Figure 1 – (A) Correlation between the clinical SHAP score and offline classification accuracy. The offline scores have been obtained in realistic conditions with the patients wearing their prostheses and training and testing performed on sets of data obtained in different arm positions. Despite the realistic conditions, the associations shown here are not strong. For example, a SHAP score of approximately 40 may correspond to classification accuracy lower than 70% or greater than 85% depending on the user. The SHAP requires precise manipulation over short periods of time which is not captured by this offline metrics. (B) The correlation between the clinical Box&Blocks test and the offline classification accuracy shows almost complete absence of association between the two. For instance, the two patients who achieved classification accuracies >95% were radically different for the number of blocks they could transfer. When computed in less realistic conditions (without prosthesis and testing on the same arm posture as training) the offline scores were greater than in the presented conditions but showed almost no correlation with clinical tests, since the majority of the patients were not able to conclude the clinical evaluation without substantial retraining.

Figure 2 – Classification output for two patients with substantially different outcome of the Box&Blocks test but very similar classification accuracies over all motions. The focus here is on the three hand motions that are most relevant for the Box&Blocks task – hand open, key grip and fine pinch. The offline accuracy for these motions is lower for the subject with the higher clinical score.

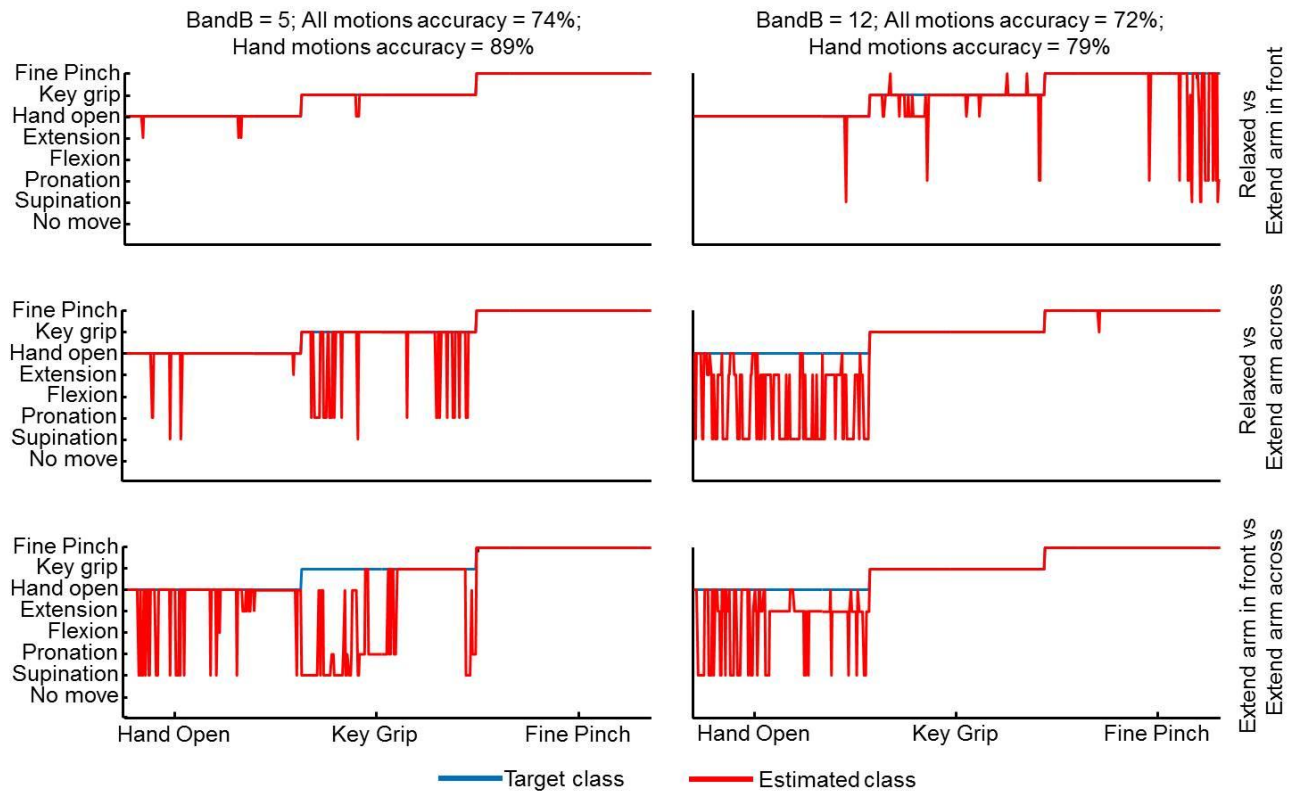
391 Figure 1



392

393

394 Figure 2



395