Application of Wearable Monitoring System in Tourette Syndrome Assessment

Sofia Scataglini[1(✉)], Marcello Fusca[1], Giuseppe Andreoni[1], and Mauro Porta[2]

[1] Department of Design, Politecnico di Milano, Via Durando 38/a, 20158 Milan, Italy
{sofia.scataglini,giuseppe.andreoni}@polimi.it
[2] Department of Neurology, Tourette Centre, IRCCS "Galeazzi", Milan, Italy

Abstract. This study presents the application of a wearable monitoring system for the assessment of tic events in subjects affected by Tourette Syndrome (TS). A multifactorial analysis and validation of the proposed system is carried out collecting simultaneous and synchronized recordings of data from the wearable actigraph and from two video cameras that allowed two medical doctors with different expertise to classify the motor events as tics and their related severity scale. A dedicated software implements the algorithm for automatic tic detection and to compare this assessment with the standard video recording protocol used to discriminate and classify tic events of high intensity and tic event of low intensity (facial grimacing or vocal tics). Double blind analysis on a nine subjects allowed us to compare the variability between operators and wearable device, and conclude the system has good potential but algorithms refinement is still needed before its possible application in clinical practice. Currently it still requires the integration with a video analysis protocol if the tics are mild or are vowels giving a complete clinical frame.

Keywords: Wearable monitoring · Tourette Syndrome · Tics · Video monitoring · Automatic tic detection

1 Introduction

Tourette Syndrome (TS) is a chronic neurologic disorder characterized by the childhood onset of multiple motor and phonic tics that wax and wane over time [1]. A distinctive issue concerning the assessment of TS is the difficulty in quantifying and classifying objectively tics from clinical manifestations. Multiple variables such as frequency, number of tic-types, intensity, complexity, body distribution, supressibility, and interference with normal activities are commonly considered to assess the TS severity. The evaluation of TS is currently carried out either through a clinical examination or through patient reports based on self-assessment of tic disorder. The analysis of video-recordings is the standard reference for many TS studies but it is confined to a short and in-hospital visit: 10 min video recording in controlled environments (with and without the presence a clinical operator) and the related tic counting represents today the gold standard technique to define the choice between pharmacological or Deep Brain Stimulation (DBS) implant therapy. A longer and multifactorial monitoring should be adopted to support a

P. Perego et al. (Eds.): MobiHealth 2016, LNICST 192, pp. 100–106, 2017.
DOI: 10.1007/978-3-319-58877-3_13

more precise and fitting to the patient intervention. Furthermore, a multifactorial assessment, e.g. including also other biosignals related to psychophysiological states (e.g. heart rate – HR and/or breathing rate - BR) that could affect the frequency of tic events, could support the identification of a more complete clinical frame that is appreciated by clinicians. For this purpose, wearables technologies (WT) are very promising. Recently our research group proposed a novel method based on WT for monitoring and quantifying motor-tics caused by the TS whose validation with respect to the standard protocol is reported in [2]. This paper discusses the further analysis of the application of this novel approach for the reliability of the systems and as a clinical decision support tool for not expert operators because it offers the opportunity to have a reliable semi-automated computation of tics also over long monitoring period and capable also to implement the different severity scale adopted in clinical practice.

2 Materials and Methods

The experimental protocol consisted in a joint video and a biosignal acquisition. The videotape protocol proposed by Goetz et al. [3] involves four sessions of about 2.5 min.

The subject was placed in a quiet room in front of a video camera. Two body views were recorded: a full frontal body and head and shoulders, both with the presence of the examiner or the absence of the examiner. An wearable device (WD) was placed through a belt on the chest of the subject for the automatic tic detection. The commercial system (PROTHEO I, SXT - Sistemi per la Telemedicina, Lecco, Italy) [5] consists of a plastic case containing a 1 ECG-lead and a 3D-acceleration sensor (LIS3L06AL, STMicroelectronics, Geneva, Switzerland), a Bluetooth transmission module, and a rechargeable Li-ion battery. The wearable device is supported by a software (HIM, Sensibilab, Politecnico di Milano, Lecco, Italy) [6] designed for collecting and managing data in applications requiring real-time biosignal monitoring (Heart Rate) and movement (3D acceleration). Two characteristics of the motor-tics were evaluated: intensity and frequency related to a 5 class scale proposed in literature [4].

According to the severity assessment proposed in the Rush Videotaping Rating Scale and revised by Bernabei et al. for the accelerometric assessment [2], we classified the motor tics into a 5-level scale: (1) barely perceptible; (2) visible; (3) some problem; (4) impaired function; (5) no function. Video analysis was carried out through an open source software (Advene, CNRS, France) providing a model and a format to share annotations about digital video documents (movies, courses, conferences…), as well as tools to edit and visualize the hypervideos generated from both the annotations and the audiovisual documents [7]. It is an easy-to-use platform with intuitive interfaces and text file output for further analysis. Video analysis was conducted by two operators: one expert, (operator 1) who was responsible for the registration; the second (operator 2) inexperienced that elaborated the data in double blind. The two operators evaluated the presence or absence of tic events and the intensity giving them a level (from 1 to 5). The third dataset is composed by tics identified by the WT system through the peaks of accelerations associated to each event. No specific on subject or self-adaptive calibration was adopted to verify a raw approach based on a simple acceleration threshold that can

be related to tics. The three data were synchronized in time in the Advene platform together with their operator assessment or the classification provided by the wearable device Fig. 1).

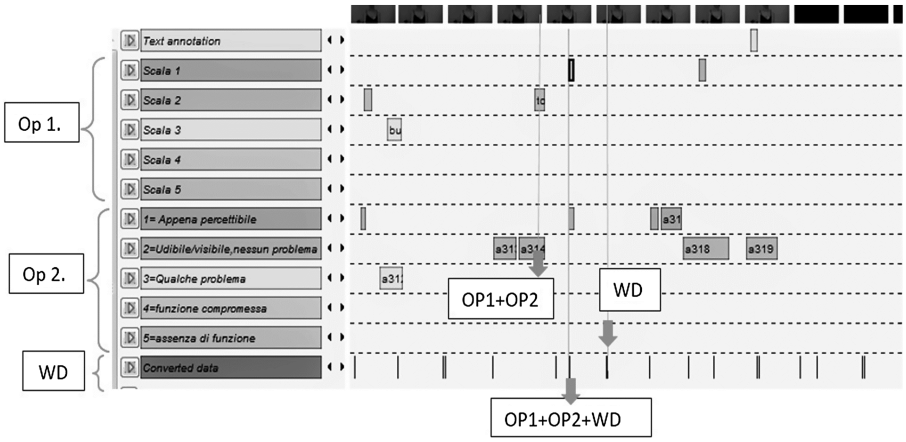

Fig. 1. The ADVENE Software suite for data synchronization and subjective assessment by the operators and correspondence analysis with automated tics computation by the WD system.

A video analysis comparison of the tics recognized by the wearable device with the tics recognized from the two operators was made. The tics can be noticed in three different ways: (a) by only the WD, (b) by the WD at least one operator, (c) by only one or two operators.

The procedure was repeated on 9 subjects without DBS implant [8] but under pharmacological treatment.

3 Results and Discussion

Table 1 shows the tics recognized and classified by both operators for typology and severity while the following columns report the tic counting carried out only by the operator no. 1 (expert neurologist). These data are assumed to be the reference in our study and they also represent the clinical gold standard.

The main goal of this study was to assess the reliability of the wearable system in measuring motor tics in TS. In Table 2 the occurrence of all tics detected by the operator no. 1 (expert neurologist) and by the WD are compared.

In these dataset motor, facial and vocal tics are included even if borderline with the goal of the study being very probably not detectable with a WD mounted onto the trunk of the subject. In any case we decided to use the raw data under the hypothesis that also facial tics produce a perturbation of the body equilibrium so that an actigraph could be suitable to measure also this kind of events. Instead, these data indicate that the WD approach produces an underestimation of events with respect to the clinical classification given by doctors. The percentage of correspondence for tics observed by the WD and

tics observed by Operator no. 1 (expert clinician) was only 44.5%, with an average value of 47,66% ± 16,49% (SD) among the 9 subjects.

Table 1. Tic classification in body positioning and in severity as carried out by both clinical operators for each subject, and tic count for each severity level according to the subjective assessment by Operator no.1.

Subject	Tic typology	Level 1	Level 2	Level 3	Level 4	Level 5
1	Eyes, trunk	6	8			
2	Eyes, facial grimacing	17	19			
3	Trunk	5	15	14	3	
4	Trunk, facial grimacing	2	14	10		
5	Grimace, eyes	8	11	6		
6	Trunk	3	9	8	3	1
7	Eyes	13	13			
8	Eyes, head, trunk	1	4	11	2	
9	Vocal, trunk	4	11	21		

Table 2. Comparison of tics recognized by the Operator no. 1 and the WD.

Subject	Op.1	WD	% Corresp.
1	14	12	85,71
2	36	15	41,67
3	40	12	30,00
4	26	14	53,85
5	25	8	32,00
6	24	12	50,00
7	26	10	38,46
8	18	9	50,00
9	36	17	47,22
Tot.	245	109	44,49

This low value could be also related to the absence of a specific single subject calibration but a simple thresholding approach, but more probably they could be due to a wrong (but mandatory) sensor positioning onto the subject (in fact in case of facial tics it is unacceptable for the subject to wear a sensing sensor onto the face). Therefore, we conducted a more detailed analysis in case of mild episodes or facial tics.

In the performed experiments, the main criticisms arise with low intensity events that occur in anatomical districts far from the sensing system, i.e. the WD that is usually placed onto the trunk, so that it is not able to detect them; from parallel video check, this seems to be the main cause determining underestimation. Excluding the subjects presenting only facial tics (i.e. S2, S5, and S7) and excluding the facial tics by the occurrences detected both by the operator and the WD (null in the analyzed data) the situation significantly improves as reported in the following Table 3.

Table 3. Comparison of tics recognized by the operator 1 and the WD (related to tics only involving the trunk).

Subject	Op.1	WD	%Corresp.
1	13	12	92,31
3	37	12	32,43
4	18	14	77,78
6	24	11	45,83
8	15	9	60,00
9	24	17	70,83
Tot.	129	75	57,25

The total correspondence percentage increases at 57,25% and the average agreement level raises at 63,20% ± 21,80% (SD). This is a good result but still not optimal for introduction into the clinical practice.

A similar analysis is performed between the experienced operator (Op.1) and the non-expert neurologist (Op.2): data are reported in Table 4.

Table 4. Comparison of tics recognized by the two clinical operators.

Subject	Op.1	Op.2	%Corresp.
1	14	4	28,60
2	36	12	33,30
3	40	16	40
4	26	1	3,80
5	25	13	52
6	24	19	79
7	26	15	57,70
8	18	1	6
9	36	21	58,30
Tot.	245	102	41,60

The unexpert operator (Op.2) had a tic recognition percentage of 41.6% compared to the expert operator. It is also noticeable the high variability in not-experienced evaluation: the average correspondence percentage is 39,86% ± 25,56% (SD).

Comparing the results of the expert operator with the results of the inexperienced operator and the WD recognition capability, we can conclude that the WD behaves as an inexperienced operator presenting a small difference and a lower variability in judgements. This makes the developed system a useful tool for preliminary analysis in clinical practice.

For further application in long term and remote monitoring, the algorithm could and should be optimized to recognize severe events that are discarded now, because they are considered as voluntary movements (e.g. walking). This is a limitation of the present method and has to be considered in the future developments. Similarly, but at the opposite, another source of error in the method are other rapid movements but not related to the pathology could happen that can overcome the imposed acceleration threshold so

that they are classified as tics by the software. This would have produced an overestimation but the experiments did not elicit this occurrence.

4 Conclusions

Improving the quality of life and social distress are factors that affect TS. The repeated visits, the pharmacological changes, anxiety and depression are entrusted to the treating physician for evaluation of the symptoms that afflicts them and their care. In the absence of an objective quantification of the "tics" associated with the syndrome, it becomes difficult to assess the evolution of the disease, monitor the changes and rate any improvements. The proposed systems aims at offering a novel solution to intervene in these cases and to support clinician in their decisions.

The proposed approach based on an actigraph and simple two-level thresholding in low and high trunk acceleration shows good discriminatory capacity for tic event classification in subjects with high intensity scales (scale 3–4), but it is underperforming in recognizing low severity tics as facial grimacing or vocal tics. The WD approach has a smaller increase in reliability to detect motor tics with respect to an unexperienced operator. The experienced operator, i.e. the skilled neurologist, is the gold standard in our study. Data analysis revealed that the WD is unable to detect the facial or voice tics given its positioning on the body. Considering the inexperienced operator, the difference is due to the separation capacity of the tic events and their scale. For example, the operator 2 recognizes one tic even twenty minutes in duration. This is not realistic and consistent with the clinical significance and interpretation. From a global point of view, the unskilled operator and the WD showed a similar overall evaluation.

The objective measures provided by the WD system can be a valuable addition to the medical evaluation, in order to offer a more detailed up-to-date clinical picture of the patient to neurologists, neurosurgeons and psychiatrists.

Acknowledgments. The authors would like to thank you all the subjects that took part to the study for their patients, commitments and motivated collaboration.

References

1. Shapiro, A.K., Shapiro, E.S., Young, J.G., Feinberg, T.E.: Gilles de la Tourette Syndrome, 2nd edn. Raven Press, New York (1998)
2. Bernabei, M., Preatoni, E., Mendez, M., Piccini, L., Porta, M., Andreoni, G.: A novel automatic method for monitoring Tourette motor tics through a wearable device. Mov. Disord. **25**(12), 1967–1972 (2010)
3. Goetz, C.G., Tanner, C.M., Wilson, R.S., Shannon, K.M.: A rating scale for Gilles de la Tourette's Syndrome: description, reliability, and validity data. Neurology **37**, 1542–1544 (1987)
4. Goetz, C.G., Pappert, E.J., Louis, E.D., Raman, R., Leurgans, S.: Advantages of a modified scoring method for the rush video-based tic rating scale. Mov. Disord. **14**, 502–506 (1999)
5. Sistemi per Telemedicina. http://www.sxt-telemed.it/
6. Laboratorio Sensibilab, Politecnico di Milano. http://www.sensibilab.lecco.polimi.it

7. Aubert, O., Prié, Y.: Advene: active reading through hypervideo. In: Proceedings of ACM Hypertext 2005, pp. 235–244 (2005)
8. Bajwa, R.J., de Lotbinière, A.J., King, R.A., Jabbari, B., Quatrano, S., Kimberly, K., Scahill, L., Leckman, J.F.: Deep brain stimulation in Tourette's Syndrome. Mov. Disord. **22**(9) (2007)