Automatic detection of cataplexy



*Published Version:*
Bartolini, I., Pizza, F., Di Luzio, A., Neccia, G., Antelmi, E., Vandi, S., et al. (2018). Automatic detection of cataplexy. SLEEP MEDICINE, 52, 7-13 [10.1016/j.sleep.2018.07.018].

*Availability:*
This version is available at: https://hdl.handle.net/11585/653402 since: 2018-12-27

*Published:*

DOI: http://doi.org/10.1016/j.sleep.2018.07.018

Manuscript Number: SLEEP-D-18-00229

Title: Automatic Detection of Cataplexy

Article Type: Original Article

Keywords: Narcolepsy, Type 1 narcolepsy, cataplexy, automatic recognition,

Corresponding Author: Dr. Giuseppe Plazzi,

Corresponding Author's Institution: Università di Bologna

First Author: Ilaria Bartolini

Order of Authors: Ilaria Bartolini; Fabio  Pizza; Andrea Di Luzio; Elena Antelmi; Stefano Vandi; Giuseppe Plazzi

Abstract: Objectives: Although being the most specific symptom of narcolepsy type 1 (NT1), cataplexy is currently investigated by clinical interview only, with potential diagnostic pitfalls. Our study aimed at testing the feasibility of an automatic video detection of cataplexy in NT1 patients versus controls undergoing a standardized test with emotional stimulation.
Methods: Fifteen drug naive NT1 patients and 15 age- and sex- balanced controls underwent a standardized video recording procedure including successful emotional stimulation. Video recordings were visually inspected by sleep experts to detect three typical cataplexy facial motor patterns (ptosis, mouth opening and head drop), and then analyzed by SHIATSU (Semantic-based HIearchical Automatic Tagging of videos by Segmentation using cUts). Expert-based and automatic attack detection was compared in NT1 patients and controls.
Results: All NT1 patients and no controls displayed cataplexy during emotional stimulation. Automatic detection well correlated with experts' assessment in NT1, showing higher sensitivity. In controls, automatic detection falsely identified cataplexy in two out of 15 (13.3%) subjects who showed active eyes closure during intense laughter as confounder with ptosis.
Conclusions: Automatic cataplexy detection by applying SHIATSU to a standardized test for video documentation of cataplexy is feasible and reliable. Further studies are warranted to enlarge the range of elementary motor patterns detected, analyze their temporal/spatial relations and quantify cataplexy for diagnostic and severity assessment purposes.

**Title Page**

# Title: Automatic Detection of Cataplexy

## Subtitle: Cataplexy Automatic Facial Expressions Recognition from Video Recorded Attacks

**Authors:** Ilaria Bartolini[1], Fabio Pizza[2,3], Andrea Di Luzio[1,2], Giulia Neccia[3], Elena Antelmi[2,3], Stefano Vandi[2,3], Giuseppe Plazzi[2,3].

**Affiliations:**

1 - Department of Computer Science and Engineering, University of Bologna, Bologna, Italy;

2 - Department of Biomedical and Neuromotor Sciences, University of Bologna, Bologna, Italy;

3 - IRCCS Institute of the Neurological Sciences, ASL di Bologna, Bologna, Italy.

**Institution where the work was performed:** Department of Computer Science and Engineering, University of Bologna, Bologna, Italy; and Department of Biomedical and Neuromotor Sciences, University of Bologna, Bologna, Italy.

**Corresponding Author**:

Giuseppe Plazzi

Department of Biomedical and Neuromotor Sciences, Ospedale Bellaria, Padiglione G

Via Altura 3, 40139 Bologna, Italy. Phone: +39 051 4966926 Fax: +39 051 4966176

Mail: giuseppe.plazzi@unibo.it

Word counts:

Abstract: 218

Paper: 3690

**Abstract**

*Objectives:* Although being the most specific symptom of narcolepsy type 1 (NT1), cataplexy is currently investigated by clinical interview only, with potential diagnostic pitfalls. Our study aimed at testing the feasibility of an automatic video detection of cataplexy in NT1 patients versus controls undergoing a standardized test with emotional stimulation.

*Methods:* Fifteen drug naive NT1 patients and 15 age- and sex- balanced controls underwent a standardized video recording procedure including successful emotional stimulation. Video recordings were visually inspected by sleep experts to detect three typical cataplexy facial motor patterns (ptosis, mouth opening and head drop), and then analyzed by SHIATSU (Semantic-based HIearchical Automatic Tagging of videos by Segmentation using cUts). Expert-based and automatic attack detection was compared in NT1 patients and controls.

*Results:* All NT1 patients and no controls displayed cataplexy during emotional stimulation. Automatic detection well correlated with experts' assessment in NT1, showing higher sensitivity. In controls, automatic detection falsely identified cataplexy in two out of 15 (13.3%) subjects who showed active eyes closure during intense laughter as confounder with ptosis.

*Conclusions:* Automatic cataplexy detection by applying SHIATSU to a standardized test for video documentation of cataplexy is feasible and reliable. Further studies are warranted to enlarge the range of elementary motor patterns detected, analyze their temporal/spatial relations and quantify cataplexy for diagnostic and severity assessment purposes.

**Keywords:** cataplexy, narcolepsy, video recording, automatic video analysis, laughter

**Highlights:**

- Cataplexy clinical assessment is a diagnostic challenge
- A standardized test can be optimized for video recording and automatic detection of cataplexy
- An objective cataplexy test can improve diagnosis
- An objective cataplexy detector is needed to better phenotype motor patterns and warning signs

## Introduction

Narcolepsy type 1 (NT1) is a rare, chronic and lifelong central disorder of hypersomnolence, arising in infancy or early adulthood. [1] NT1 is linked to the selective loss of hypothalamic hypocretin producing neurons of possible autoimmune origin, [2] and is clinically characterized by a pentad of symptoms, namely excessive daytime sleepiness (EDS) with sleep attacks typically showing rapid transition into REM sleep (sleep onset REM periods, SOREMPs), cataplexy (sudden loss of muscle tone during wakefulness triggered by emotions), sleep related paralysis and hallucinations, as well as disrupted nocturnal sleep. [1] The diagnosis requires the presence of clinical complaints (EDS and cataplexy) confirmed by the evidence of a pathological sleep propensity with multiple SOREMPs at the Multiple Sleep Latency Test (MSLT), or alternatively cerebrospinal hypocretin-1 deficiency in patients complaining EDS. [1] Cataplexy is the pathognomonic symptom of NT1 and it is currently assessed by clinical interview. [1,3]

Few studies attempted at characterizing cataplexy from a neurophysiological standpoint, and showed the non specific occurrence of phasic or prolonged EMG atonia during the attacks in the context of a complex picture with coexisting fragments of wakefulness and REM sleep. [4-7] In some NT1 children, moreover, it has been suggested that cataplexy may be associated with an EEG paroxysmal theta hypersynchronous activity (∼4 Hz), [8] as well as with hemodynamic changes detected at functional MRI in the frontal operculum–anterior insular and ventro-medial prefrontal cortices. [9] Despite these investigations contributed to the understanding of the phenomenon, a clear-cut biomarker of cataplexy is still missing, and the differential diagnosis of episodic muscle weakness during wakefulness stems on other objective evidences supporting NT1 against other disorders at different ages. [3] Recent data, however, well indicated that cataplexy can present with different phenotypes depending on age and disease duration, [10-14] further challenging the recognition of this singular phenomenon. The lack of an objective descriptor of cataplexy challenges its proper recognition and often leads to misdiagnosis: for example, on the one hand the sensation of "being weak with laughter" can be genuinely reported by healthy subjects, [15-16] and on the other hand the occurrence of episodes of muscle weakness during wakefulness with various emotions can be a functional symptom in patients with a conversion disorder (and normal sleep features) as well as in NT1 subjects. [17-19]. Bearing in mind these potential pitfalls in diagnosing cataplexy only by means of clinical interview, we recently developed and validated a systematic approach to document cataplexy by means of video recording of patients under subject-oriented emotional stimulation in the diagnostic work-up of suspected hypersomnia of central origin, that proved useful and feasible to document clinically suspected cataplexy. [20] By analyzing and contrasting the video documented attacks of patients with NT1 and with conversion disorder, we disclosed a set of useful clues evident at video recordings that allowed to distinguish cataplexy from pseudocataplexy in blind to other clinical and objective information. [19] In particular, we confirmed that a complex array of motor phenomena occurring in the facial district (i.e. ptosis, mouth opening, tongue protrusion - the "cataplectic facies") during laughter behavior and abruptly interrupting smile and facial expression showed to be tell-tales to correctly identify cataplexy. [10,19] Translating the above visual detection approach in clinical practice, however, would suffer from significant inter-observer reliability pitfalls. [19] Therefore, the development of an automatic approach to video-recorded attacks of suspected cataplexy is the avenue to establish new possible tools to identify NT1 for diagnostic and severity issue purposes.

The applications of automatic video recognition in the medical field has shown promising results ranging from systems for monitoring and detecting the level of pain in the patient during medical treatment, [21] to software systems capable of analyzing video endoscopies for the detection of gastrointestinal polyps, [22] up to a software architecture capable of recognizing the phases of the surgical workflow, [23] and to tools able to give quantitative support in the clinical assessment of different neurological diseases with movement impairment. [24]

After obtaining promising preliminary results by a proof of concept study on the potential application of SHIATSU (Semantic-based HIearchical Automatic Tagging of videos by Segmentation using cUts), [25,26] to detect cataplexy in 5 NT1 patients, [27] here we present the first systematic attempt for an automatic video-analysis of the characteristic facial motor phenomena of cataplexy. Automatic video analysis is performed exploiting and extending the features offered by SHIATSU, a pre-existing general and extensible software framework for video retrieval which is based on the semi-automatic hierarchical semantic annotation of videos exploiting the analysis of their visual content. [25,26,28] SHIATSU has been conceived and developed by the "Multimedia Database Group" (`http://www-db.disi.unibo.it/Shiatsu/`) at the Department of Computer Science and Engineering (DISI) of the University of Bologna.

## Methods

### *Population*

The population of subjects included in the study is a clinical population of patients referred to the Outpatient Clinic for Narcolepsy of the University of Bologna for suspected hypersomnia of central origin, who underwent a systematic clinical, neurophysiological and biological assessment including 48 hour continuous polysomnography, test for cataplexy, MSLT, blood drawn for HLA-DQB1*0602 determination and, whenever possible, lumbar puncture for cerebrospinal hypocretin-1 assay. [20,29] The study was approved by local ethical committee, and all subjects signed a written informed consent including the use of video recordings for research purposes on the differential diagnosis of central disorders of hypersomnolence. Among 84 consecutive patients hospitalized between June and December 2017, we selected 15 drug-naive patients with a final diagnosis of NT1 who also showed a minimum of five episodes of cataplexy at our standardized test, and who subjectively confirmed cataplexy occurrence during the test. [20] As clinical controls for cataplexy we selected a group of 15 sex- and age- balanced drug naive patients with clinical complaint of EDS, but with neurophysiological and biological assessments excluding NT1 (2 patients with final diagnosis of idiopathic hypersomnia, 13 patients with subjective EDS complaint and normal sleep features). Table 1 shows the clinical, neurophysiological and biological information of the studies population.

*Methods*

*Definition of the best recording setting for automatic video analysis*

In the proof-of-concept paper, [27] and in the subsequent preliminary experiments, automatic video-analysis showed some pitfalls due to the heterogeneity of the collected data. For this reason, starting from the laboratory setting described in Vandi et al [20] we tried to harmonize the video recording method in order to facilitate the automatic extraction of visual content.

Therefore, all the technical requirements of the new recordings (30 fps video speed, FULL-HD resolution, h264 video coding and AAC audio coding, both encapsulated in the mp4 container, that is today's standard for video distribution), the type of framing (Medium Close-Up) and the type of video camera (GoPro Hero4 Session™) have been carefully defined.

Starting from the laboratory setting including the baseline recording of the subjects without any stimulation followed by up to 30 minutes of recording during emotional stimulation (watching funny videos), we maintained the distance between the patient and the monitor as indicated in Vandi et al, [20] modifying the position of the camera, trying to find the minimum distance that represented the best compromise between visual detail of the face and the fact that the patient was not distracted by the presence of the camera: we placed the camera on an adaptable mobile support, at the same height of the patient's face (or slightly lower) and about 65 cm away from him/her. This new setting allows the recording of high-resolution movies where the facial expressions of subjects is well documented to allow in depth visualization of typical facial cataplexy features. [19]

*Human manual labeling of video recordings of cataplexy attacks and their automatic analysis for the identification of ptosis, head drop and mouth opening*

In order to proceed with the work of setting-up and testing automatic recognition methods, we created a labelling of the video recordings, i.e. the "Ground Truth" (GT), through a manual procedure: narcolepsy specialists examined each video using a freeware editing software, highlighting the start frame and the end frame of each motor phenomenon and of every cataplectic episode present in the video. The labeling was performed by two sleep technicians (SV, GN), and further validated for any disagreement by a neurologist expert in sleep medicine and narcolepsy (FP). Given the different temporal resolution of the human eye and of SHIATSU in detecting motor phenomena, a maximum tolerance of about 500 ms (corresponding to 15 frames, in our recordings) has been granted to the domain experts in labeling the beginning and end of each motor phenomenon. These labeling information, together with the reference videos, represent the GT of the project, that we used to set-up SHIATSU and test its accuracy.

We focused our framework on the recognition of the facial components of cataplexy, namely ptosis, dropping of the head and opening of the mouth, given that these have emerged as more specific and sensitive signs of cataplexy. [19,20]

In detail, the visual features characterizing the three motor phenomena together with the corresponding comparison metrics exposed by SHIATSU have been adapted in order to reflect the requirements of the new application domain.

To achieve this goal, we have used methods for extracting facial landmarks from every frame of the video in order to obtain information about the position of the patient's face, the shape and the size of the eyes, nose, mouth, cheekbones and jaw. The position of the facial landmarks is defined in terms of ($x$, $y$) coordinates within the video frame, with the origin point ($0,0$) fixed at the top left of the image; the coordinates of landmarks provide coordinates of the patient's face (for example, through the coordinates of the barycentre of all landmarks). Starting from these features, a mathematical formalization has been opportunely defined for each motor phenomenon.

Regarding ptosis, the Eye Aspect Ratio (EAR) was calculated, defined as the ratio of the eye height to the eye width. The semantics of EAR are as follows: when an eye is closing, EAR approaches zero, whereas when the eye is completely open, EAR attains its maximum value (which varies from person to person). We perform the EAR extraction process on every pair of the video recordings, at both baseline conditions and while the patients were undergoing emotional stimulation. By measuring, the median of EAR values for the baseline video recordings, we obtain the characterization of the patient opened eyes, while collecting every EAR values from video of the patients were undergoing emotional stimulation, the PERcentage of eye CLOSure (PERCLOS) was calculated, [30] using the median computed in the baseline videos as a reference.

We defined ptosis as present if the amount of time when PERCLOS exceeds an experimentally defined threshold higher of the maximum duration of an eye blink (ie about 400 ms). [27]

Regarding mouth opening, the Mouth Aspect Ratio (MAR) was calculated, defined as the ratio of the mouth width to the mouth height. The semantics of MAR are as follows: when the mouth is closed, MAR attains its maximum value (which varies from person to person), whereas when the mouth is completely open, MAR reaches his lowest value; intermediate values characterize various types of smile. We perform the MAR extraction process only on the video recordings while the patients were undergoing emotional stimulation because it may happen that the patient never closes his mouth for the entire duration of the baseline video so the median value might not match a mouth while smiling widely and therefore it cannot be used as a reference: this problem was solved by normalizing values of the MAR time series to a range between 0 and 1.

We defined Mouth Opening as present if normalized MAR is lower than an experimentally defined threshold and PERCLOS exceeds an experimentally defined threshold or head drop is recognized as present for the considered frame.

Regarding the dropping of the head, the two values that represents the rotation of the head (in radiants) around X, Z axes and the center of gravity (CoG) of the two landmarks that characterize the external corner of each eye and the tip of the nose were calculated. As for the EAR values, we perform the CoG extraction process on the video recordings at both baseline and under emotional stimulation conditions. In this way, by measuring the median value of CoG values for the baseline video recordings, we obtain the position of the considered patient's face under normal condition.

We defined head drop as present if the CoG values are, respectively, greater or less two experimentally defined thresholds (computed starting from the median CoG of the baseline video) or the two values representing the rotation of the head are greater than an experimentally defined angle. Further, for head drop to be considered as present also ptosis has to be considered present for the entire duration of the sequence of frames in order to avoid the inclusion of voluntary head movements.

Figure 1 reports visual examples of the features involved in above described motor patterns.

Finally, we defined cataplexy as present if at least one of the motor phenomena above described was detected.

*Statistical analysis*

The performances of SHIATSU have been evaluated by comparison, at the level of a single video frame, with the GT produced by the domain specialists. To do this, 15-labeled videos of NT1 patients with documented cataplexy and 15 unlabeled videos of clinical controls were used, for a total amount of ~370000 frames and ~3 hours and 20 minutes of video recordings.

To objectively evaluate the performance of the automatic analyzer we used the classic sensitivity / specificity values, together with the accuracy value.

These measures are calculated starting from four values that describe the number of events correctly and incorrectly recognized for each of the two classes available (in our case, motor phenomenon actually present or absent in the considered video frame).

More in detail, we define as:

- true negative (tn): a frame that has been labeled as not containing the considered motor phenomenon (or cataplexy in general) by both the domain specialists and SHIATSU;
- false negative (fn): a frame that has been labeled as containing the considered motor phenomenon (or cataplexy in general) by the domain specialists but not by SHIATSU;
- true positive (tn): a frame that has been labeled as containing the considered motor phenomenon (or cataplexy in general) by both the domain specialists and SHIATSU;
- false positive (fp): a frame that has been labeled as containing the considered motor phenomenon (or cataplexy in general) by SHIATSU but not by the domain specialists.

Starting from these measures, we defined the sensitivity *sens* as the proportion of positives that are correctly identified as such.

$$sens = \frac{tp}{tp + fn}$$

Specificity *spec*, instead, is the proportion of negatives that are correctly identified as such.

$$spec = \frac{tn}{fp + tn}$$

Accuracy *a*, finally, is defined as the number of correct (positive and negative) classifications on the total number of frames of the considered video.

$$a = \frac{tp + tn}{tp + fp + tn + fn}$$

To evaluate the performances in the videos of the clinical controls only specificity measure will be used, because, since there are no cases of true positive or false negatives, the sensitivity cannot be calculated and the accuracy is equivalent to the specificity. Furthermore, for the phenomena not present in the individual recordings of patients showing NT1 manifestations only the specificity was calculated (for the reasons just described) and was not considered in the weighted average at the end of the table: this to avoid that they could positively affect the final result.

**Results**

Scores of motor features occurrence in patients' recordings are reported in Table 2. Noteworthy, SHIATSU detects a few episodes more than those observed and certified by the domain experts. Scores of motor features occurrence in healthy controls' recordings are reported in Table 3. The absence of cataplexy detected by sleep experts is confirmed by SHIATSU with the exception of two videos (ie #24 and #25) in which ptosis was detected.

A visual example of output produced by SHIATSU with regard to the detection head-drop within a video recording can be seen in Figure 2. As we can see from the figure, the time course of the automatic detection by SHIATSU is in line with the annotation performed by the domain experts. Table 4 shows the results concerning patients who presented cataplexy. It should be noted that SHIATSU turns to be less sensitive in identifying the phenomenon “head drop” when compared to the other phenomena, possibly for implementation choices of internal detection validation through ptosis presence. The rest of the results are in line with the general accuracy of the system, even if with a slight penalization for the detection of the mouth opening due to the fact that this phenomenon may sometimes occur with other mouth motor phenomena not yet implemented in SHIATSU.

Supplemental Table 1 shows the accuracy (in the form of specificity) of SHIATSU in recognizing cataplexy as absent in healthy controls. Specifically, in 13 videos out of 15 considered, cataplexy is not detected in any frame (specificity equal to 1), while ptosis was identified in about 45 seconds of the approximately 15 minutes which constitute the remaining 2 videos examined (31 sec in video 24, 16 sec in video 25). The parts of video identified by SHIATSU consist of a particularly pronounced laughter associated with almost complete eyes closure for several seconds.

Finally, supplemental Table 2 shows the difference, in terms of global presence or absence in the whole video, in the detection of cataplexy carried out by the domain experts and by SHIATSU. In this study, it was decided to implement a rule of simplified detection of cataplexy based on the presence of the detected motor phenomena (i.e. the presence of at least one of the three considered motor patterns), disregarding any information related to the position, the order and the temporal density of each individual motor phenomena.

Figure 3 shows real examples of visual results for the three motor patterns and their combination at the same time, as detected by SHIATSU.

## Discussion

This is the first systematic attempt to automatically detect cataplexy based on facial expressions recognition from video-recorded attacks occurring during emotional stimulation. Video recordings made on NT1 patients and controls were automatically analyzed by exploiting and extending the software SHIATSU, [25,26,28] and comparing its sensitivity in identifying three typical facial fragments of cataplexy (ptosis, mouth opening, head drops) against the detection provided by domain experts. SHIATSU proved to have higher performances in detecting episodes than domain experts in NT1 patients. On the other hand, only 2 of the 15 videos with clinical controls were erroneously recognized as containing cataplexy due to the presence of persistent eyes closure during laughter. This study paves the way to automatic video analysis of cataplexy for diagnostic and severity assessment purposes and to cataplexy automatic recognition overall, along with the increasing availability of video recording technologies.

The main relevance of our study is to extend the utility of video recordings in the recognition and in the differential diagnosis of cataplexy. Indeed, the investigation of cataplexy by clinical interview only can be challenging and misleading in patients with objectively confirmed EDS, [18] and also in NT1 subjects with coexistence of cataplectic and pseudo-cataplectic attacks. [17] Objective cataplexy documentation is the pre-requisite to allow clinicians to an in depth analysis of the motor patterns able to steer towards a correct diagnosis. [3,19]

Indeed, it should be born in mind that NT1 is a rare and life-long disorder most frequently arising in childhood that goes unrecognized for up to 15 years by physicians, [31,32] and therefore the development of objective tools to identify its most sensitive and specific signs are strongly needed and will increase overall ability to correctly screen, diagnose, and manage patients across disease course.

Moreover, accurate and automatic cataplexy detection can further improve the possibility to precisely identify the onset (in the time domain) and course (in the space domain) of attacks and to disclose features not perceived by the human eye, opening new avenues for the study of the cataplectic phenomena.

It is also interesting to discuss the detection of a single motor phenomenon (i.e. "ptosis") misinterpreted by SHIATSU for cataplexy in clinical controls. Indeed, weakness with laughter is a subjective sensation perceived by otherwise healthy subjects, [15,16] and is paralleled by neurophysiological changes (such as H-reflex amplitude reduction) qualitatively similar to those reported in cataplexy, albeit of lower intensity. [33-35] Previous data based on blinded video documented facial expression changes during our standardized test for cataplexy similarly disclosed ptosis-like patterns in both healthy subjects (4.8%) and non-cataplectic hypersomnia patients (9.1%). [20] The occurrence of partial eyes closure during intense laughter has also been observed during intense laughter also in conversion disorder patients within pseudocataplexy attacks. [19] This synkinetic contraction of the zygomatic muscles and the pars lateralis of the orbicularis oculi muscles (known as "the Duchenne smile") describes the subjective experience of enjoyment during pleasant positive emotions, [36] and could be regarded as the motor correlate of the successful trigger during our experimental procedure. [19,20] Future system implementation will easily address the distinction between the "Duchenne smile" and cataplexy motor patterns. Indeed, future implementation avenues will include the automatic recognition of other motor phenomena (already identified in the literature) to address the automatic detection of premonitory phenomena, of misidentified paucisymptomatic cases, and the study of cataplexy severity along disease course and treatment. Finally, it may be of considerable interest testing alternative detection methods based on artificial intelligence techniques such as "Deep Learning" with convolutional neural networks.

*Conclusions*

In this preliminary work, we tested the feasibility of an automatic detector for the identification of cataplexy, the pathognomonic symptom of NT1. We exploited the visual content of video recordings made on NT1 patients and controls based on SHIATSU versus human-based detection. Our study focused on ptosis, opening of the mouth and dropping of the head, three of the most recurrent cataplexy facial motor patterns. Experiments conducted on real data demonstrated the accuracy of SHIATSU in detecting cataplexy, encouraging the expansion of the work, through the enlargement of the database of labeled videos, the implementation of new characteristic facial patterns, the study of additional aspects of cataplexy, and the creation of new rules for combining different motor phenomena in order to provide a model as accurate as possible for cataplexy recognition in all its aspects.

**Acknowledgements**

Thanks to Fondazione del Monte di Bologna e Ravenna for a research grant in the year 2016 to Giuseppe Plazzi. We are indebted to all the patients and families participating to this study, most notably the Italian Association of Narcolepsy (AIN onlus) patients. Without their contributions, this study would not have been possible.

**Disclosure Statement**

**Financial Disclosures:** Giuseppe Plazzi participated in advisory board for Jazz Pharmaceuticals, UCB pharma, and Bioprojet outside the submitted work. Ilaria Bartolini, Fabio Pizza, Andrea Di Luzio, Giulia Neccia, Elena Antelmi, and Stefano Vandi have nothing to disclose.

**Funding:** This work was supported by a research grant from Fondazione del Monte di Bologna e Ravenna to Giuseppe Plazzi for the project “Riconoscimento automatico della cataplessia” in the year 2016.

## References

1. American Academy of Sleep Medicine. International Classification of Sleep Disorders. 3rd ed. Darien, IL: American Academy of Sleep Medicine; 2014.
2. Peyron C, Faraco J, Rogers W, et al. A mutation in a case of early onset narcolepsy and a generalized absence of hypocretin peptides in human narcoleptic brains. Nat Med 2000;6:991-7.
3. Pillen S, Pizza F, Dhondt K, Scammell TE, Overeem S. Cataplexy and Its Mimics: Clinical Recognition and Management. Curr Treat Options Neurol 2017;19:23.
4. Guilleminault C, Wilson RA, Dement WC. A study on cataplexy. Arch Neurol 1974;31:255e61.
5. Dyken ME, Yamada T, Lin-Dyken DC, et al. Diagnosing narcolepsy through the simultaneous clinical and electrophysiologic analysis of cataplexy. Arch Neurol 1996;53:456e60.
6. Krahn LE, Boeve BF, Olson EJ, et al. A standardized test for cataplexy. Sleep Med 2000;1:125e30
7. Vetrugno R, D'Angelo R, Moghadam KK, et al. Behavioural and neurophysiological correlates of human cataplexy: a video-polygraphic study. Clin Neurophysiol 2010;121:153-62.
8. Vassalli A, Dellepiane JM, Emmenegger Y, et al. Electroencephalogram paroxysmal θ characterizes cataplexy in mice and children. Brain 2013;136:1592-608.
9. Meletti S, Vaudano AE, et al. The Brain Correlates of Laugh and Cataplexy in Childhood Narcolepsy. J Neurosci. 2015;35:11583-94.
10. Serra L, Montagna P, Mignot E, Lugaresi E, Plazzi G. Cataplexy features in childhood narcolepsy. Mov Disord 2008;23:858-65.
11. Plazzi G, Pizza F, Palaia V, et al. Complex movement disorders at disease onset in childhood narcolepsy with cataplexy. Brain 2011;134:3477-89.
12. Pizza F, Franceschini C, Peltola H, et al. Clinical and polysomnographic course of childhood narcolepsy with cataplexy. Brain. 2013;136:3787-95.a
13. Antelmi E, Vandi S, Pizza F, Liguori R, Plazzi G. Parkinsonian tremor persisting during cataplexy. Sleep Med 206;17:174-6.
14. Antelmi E, Pizza F, Vandi S, Plazzi G. Stereotyped episodes of aphasia and immobility: how cataplexy mimics stroke in an elderly patient. Sleep Med 2017;36:122-4.
15. Anic-Labat S, Guilleminault C, Kraemer HC, Meehan J, Arrigoni J, Mignot E. Validation of a cataplexy questionnaire in 983 sleep-disorders patients. Sleep 1999;22(1):77-87.
16. Overeem S, Lammers GJ, van Dijk JG. Weak with laughter. Lancet. 1999;354:838.
17. Plazzi G, Khatami R, Serra L, Pizza F, Bassetti CL. Pseudocataplexy in narcolepsy with cataplexy. Sleep Med 2010;11:591-4.

18. Pizza F, Vandi S, Poli F, et al Narcolepsy with cataplexy mimicry: the strange case of two sisters. J Clin Sleep Med 2013;9:611-2.b
19. Pizza F, Antelmi E, Vandi S, et al. The Distinguishing Motor Features of Cataplexy: A Study from Video Recorded Attacks. Sleep 2018. doi: 10.1093/sleep/zsy026.
20. Vandi S, Pizza F, Antelmi E, et al. A standardized test to document cataplexy. Sleep Med 2017; doi: 10.1016/j.sleep.2017.08.021.
21. Werner P., Al-Hamadi A., Niese R., Walter S., Gruss S. and Traue H. C. "Automatic Pain Recognition from Video and Biomedical Signals". 22nd International Conference on Pattern Recognition, Stockholm, 2014, pp. 4582-4587.
22. Billah M, Waheed S, and Rahman MM. An Automatic Gastrointestinal Polyp Detection System in Video Endoscopy Using Fusion of Color Wavelet and Convolutional Neural Network Features. International Journal of Biomedical Imaging, vol. 2017, Article ID 9545920; 2017 p 9.
23. Twinanda AP, Shehata S, Mutter D, Marescaux J, De Mathelin M, Padoy N. EndoNet: A Deep Architecture for Recognition Tasks on Laparoscopic Videos. arXiv 2016;1602.03012
24. Bandini A, Orlandi S, Escalante HJ, et al. Analysis of facial expressions in parkinson's disease through video-based automatic methods. J Neurosci Methods 2017;281:7-20.
25. Bartolini I, Patella M, Romani C. SHIATSU: Semantic-Based Hierarchical Automatic Tagging of Videos by Segmentation using Cuts. In Proceedings of the 3rd International ACM Multimedia Workshop on Automated Information Extraction in Media Production (AIEMPro10); 2010
26. Bartolini I, Patella M, Romani C. SHIATSU: tagging and retrieving videos without worries. Multimedia Tools and Applications 2013; 63: 357-85.
27. Bartolini I, Di Luzio A. Towards Automatic Recognition of Narcolepsy with Cataplexy. MoMM2017 Proceedings of the 15th International Conference on Advances in Mobile Computing & Multimedia. 2017, p 97-101.
28.Bartolini I, Patella M. Multimedia Queries in Digital Libraries. In Data Management in Pervasive Systems, Series: Data-Centric Systems and Applications, ISBN: 978-3-319-20062-0, Springer, November 2015.
29. Pizza F, Moghadam KK, Vandi S, et al. Daytime continuous polysomnography predicts MSLT results in hypersomnias of central origin. J Sleep Res. 2013;22:32-40. C
30. Knipling R, Rau P. PERCLOS: a valid psychophysiological measure of alertness as assessed by psychomotor vigilance. Washington: Office of Motor Carriers, 1998.
31. Luca G, Haba-Rubio J, Dauvilliers Y, Lammers GJ, et al; European Narcolepsy Network. Clinical, polysomnographic and genome-wide association analyses of narcolepsy with cataplexy: a European Narcolepsy Network study. J Sleep Res 2013;22:482-95.

32. Thorpy MJ, Krieger AC. Delayed diagnosis of narcolepsy: characterization and impact. Sleep Med 2014;15:502-7.
33. Guilleminault, C. and Gelb, M. Clinical aspects and features of cataplexy. Adv Neurol;1995:67: 65–77.
34. Lammers GJ, Overeem S, Tijssen MA, van Dijk JG. Effects of startle and laughter in cataplectic subjects: a neurophysiological study between attacks. Clin Neurophysiol. 2000;111:1276-81.
35. Overeem S, Reijntjes R, Huyser W, Lammers GJ, van Dijk JG. Corticospinal excitability during laughter: implications for cataplexy and the comparison with REM sleep atonia. J Sleep Res 2004;13:257-64.
36. Ekman P, Davidson RJ, Friesen WV. The Duchenne smile: emotional expression and brain physiology. II. J Pers Soc Psychol 1990;58:342-53.

## Tables

**Table 1:** clinical, neurophysiological and biological data of studies subjects

| ID | Sex | Age | Condition | Age Disease Onset | ESS | BMI | HLA DQB1*0602 | CSF hypocretin-1 | Cataplexy Frequency | Sleep Paralyses | Sleep Hallucinations | MSLT - Sleep Latency | MSLT - SOREMPs |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | F | 14 | NT1 | 8 | 20 | 24.22 | pos | 10.66 | 4 | No | Yes | 2.2 | 5 |
| 2 | F | 16 | NT1 | 15 | 19 | 24.97 | pos | 2.78 | 4 | No | No | 2.2 | 5 |
| 3 | F | 22 | NT1 | 20 | 21 | 40.40 | pos | 0.00 | 4 | No | No | 1.3 | 5 |
| 4 | M | 11 | NT1 | 9 | 16 | 17.84 | neg | 0.00 | 5 | No | Yes | 6.3 | 2 |
| 5 | M | 14 | NT1 | 7 | 13 | 24.46 | pos | 11.88 | 5 | No | No | 0.6 | 5 |
| 6 | M | 11 | NT1 | 9 | 20 | 17.78 | pos | 0.00 | 5 | No | No | 2.2 | 3 |
| 7 | M | 6 | NT1 | 5 | 18 | 16.53 | pos | 84.90 | 5 | No | No | 3.2 | 5 |
| 8 | M | 13 | NT1 | 9 | 11 | 18.66 | pos | 15.67 | 4 | No | No | 13.6 | 3 |
| 9 | M | 9 | NT1 | 6 | 12 | 22.50 | pos | 34.65 | 4 | No | No | 6.6 | 5 |
| 10 | M | 45 | NT1 | 45 | 19 | 28.69 | pos | 5.32 | 5 | Yes | Yes | 1.2 | 4 |
| 11 | F | 11 | NT1 | 4 | 17 | 21.51 | pos | 49.51 | 5 | No | Yes | 7.4 | 3 |
| 12 | M | 13 | NT1 | 12 | 14 | 34.22 | pos | 33.07 | 5 | No | No | 1.1 | 5 |
| 13 | F | 42 | NT1 | 38 | 21 | 31.91 | pos | 0.00 | 5 | No | No | 4.4 | 3 |
| 14 | F | 7 | NT1 | 6 | 11 | 22.00 | pos | 0.00 | 5 | No | No | 2.3 | 4 |
| 15 | M | 11 | NT1 | 9 | 11 | 18.78 | pos | 72.58 | 3 | Yes | Yes | 6.5 | 4 |
| 16 | F | 16 | IH | 13 | 4 | 17.69 | neg | 309.80 | 0 | No | No | 8 | 1 |
| 17 | F | 17 | sEDS | 14 | 16 | 20.58 | neg | NA | 0 | No | No | 20 | 0 |
| 18 | F | 20 | sEDs | 18 | 13 | 18.80 | neg | 288.12 | 0 | No | Yes | 10.3 | 2 |
| 19 | M | 11 | sEDS | 6 | 4 | 17.30 | neg | NA | 0 | No | No | 20 | 0 |
| 20 | M | 17 | sEDS | 17 | 15 | 20.50 | neg | 330.09 | 0 | No | No | 15.8 | 0 |
| 21 | M | 9 | sEDS | 9 | 1 | 15.31 | neg | NA | 0 | No | No | 16.3 | 0 |
| 22 | M | 4.2 | sEDS | 2 | 15 | 15.38 | neg | 324.04 | 0 | No | No | 12.4 | 0 |
| 23 | M | 17 | IH | 15 | 15 | 19.70 | pos | 300.60 | 0 | No | No | 3 | 0 |
| 24 | M | 9 | sEDS | 3 | 10 | 21.00 | neg | 335.70 | 0 | No | No | 18.7 | 0 |
| 25 | F | 47 | sEDS | 9 | 16 | 27.40 | neg | 356.20 | 0 | No | No | 10.3 | 0 |
| 26 | F | 16 | sEDS | 15 | 11 | 24.35 | neg | 345.50 | 0 | Yes | Yes | 11.9 | 0 |
| 27 | M | 14 | sEDS | 11 | 7 | 34.00 | neg | 318.90 | 0 | No | No | 14.5 | 0 |
| 28 | F | 39.4 | sEDS | 22 | 8 | 19.00 | neg | 351.70 | 0 | Yes | No | 8.9 | 0 |
| 29 | F | 12 | sEDS | 11 | 7 | 19.50 | neg | 361.17 | 0 | No | No | 16 | 0 |
| 30 | M | 21.1 | sEDS | 16 | 10 | 23.10 | neg | 356.60 | 0 | No | No | 15.5 | 0 |

Table Legend: ESS, Epworth Sleepiness Scale; BMI, Body Mass Index; CSF, cerebrospinal fluid; M, Male; F, Female; NT1, Narcolepsy Type 1; sEDS, subjective excessive daytime sleepiness; IH, Idiopathic Hypersomnia.

**Table 2: Scores of motor features occurrence in NT1 patients' recordings.**

| | | GT | | | SHIATSU | | | GT confirmed by SHIATSU | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Video #** | **Duration** | **# PT** | **# MO** | **# HD** | **# PT** | **# MO** | **# HD** | **# PT** | **# MO** | **# HD** |
| **1** | 485 s | 5 | 8 | 0 | 12 | 19 | 0 | 5 | 8 | 0 |
| **2** | 479 s | 6 | 2 | 1 | 10 | 10 | 2 | 6 | 2 | 1 |
| **3** | 393 s | 10 | 2 | 4 | 17 | 11 | 4 | 10 | 2 | 4 |
| **4** | 487 s | 8 | 3 | 0 | 11 | 7 | 0 | 6 | 3 | 0 |
| **5** | 237 s | 4 | 5 | 0 | 11 | 9 | 0 | 4 | 5 | 0 |
| **6** | 476 s | 12 | 10 | 0 | 16 | 14 | 0 | 12 | 9 | 0 |
| **7** | 286 s | 4 | 2 | 4 | 3 | 5 | 4 | 3 | 2 | 3 |
| **8** | 407 s | 6 | 6 | 0 | 10 | 10 | 0 | 4 | 6 | 0 |
| **9** | 517 s | 2 | 10 | 0 | 11 | 15 | 0 | 2 | 10 | 0 |
| **10** | 346 s | 11 | 19 | 2 | 10 | 20 | 6 | 10 | 19 | 2 |
| **11** | 225 s | 4 | 6 | 0 | 6 | 10 | 0 | 4 | 6 | 0 |
| **12** | 507 s | 0 | 7 | 0 | 11 | 15 | 0 | 0 | 7 | 0 |
| **13** | 401 s | 28 | 9 | 0 | 31 | 21 | 1 | 26 | 9 | 0 |
| **14** | 485 s | 46 | 13 | 0 | 43 | 18 | 0 | 39 | 12 | 0 |
| **15** | 476 s | 7 | 3 | 0 | 11 | 9 | 0 | 5 | 3 | 0 |

Table Legend: GT, Ground truth; Duration, video duration in seconds; # PT, number of Ptosis episodes; # MO, number of Mouth Opening episodes; # HD, number of Head Drop episodes.

**Table 3: Scores of motor features occurrence in healthy controls' recordings.**

| | | SHIATSU | | |
|---|---|---|---|---|
| **Video #** | **Duration** | **# PT** | **# MO** | **# HD** |
| **16** | 470 s | -- | -- | -- |
| **17** | 486 s | -- | -- | -- |
| **18** | 315 s | -- | -- | -- |
| **19** | 245 s | -- | -- | -- |
| **20** | 484 s | -- | -- | -- |
| **21** | 256 s | -- | -- | -- |
| **22** | 487 s | -- | -- | -- |
| **23** | 281 s | -- | -- | -- |
| **24** | 245 s | 4 | 6 | -- |
| **25** | 530 s | 2 | 4 | 1 |
| **26** | 486 s | -- | -- | -- |
| **27** | 542 s | -- | -- | -- |
| **28** | 428 s | -- | -- | -- |
| **29** | 474 s | -- | -- | -- |
| **30** | 509 s | **--** | **--** | **--** |

Table Legend: GT, Ground truth; Duration, video duration in seconds; # PT, number of ptosis episodes; # MO, number of mouth opening episodes; # HD, number of head drop episodes.

**Table 4: Effectiveness of SHIATSU for NT1 patients in showing cataplectic attacks.**

| | PTOSIS | | | HEAD DROP | | | MOUTH OPENING | | | CATAPLEXY | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Video #** | **Sens.** | **Spec.** | **Acc.** | **Sens.** | **Spec.** | **Acc.** | **Sens.** | **Spec.** | **Acc.** | **Sens.** | **Spec.** | **Acc.** |
| **1** | 0.71 | 0.67 | 0.68 | -- | 1 | -- | 0.77 | 0.72 | 0.73 | 0.56 | 0.69 | 0.64 |
| **2** | 0.73 | 0.97 | 0.95 | 1 | 1 | 1 | 1 | 0.81 | 0.82 | 0.84 | 0.96 | 0.95 |
| **3** | 0.41 | 0.77 | 0.76 | 0.34 | 0.99 | 0.99 | 1 | 0.64 | 0.65 | 0.64 | 0.78 | 0.77 |
| **4** | 0.7 | 0.82 | 0.82 | -- | 1 | -- | 0.75 | 0.78 | 0.78 | 0.7 | 0.83 | 0.82 |
| **5** | 0.88 | 0.85 | 0.85 | -- | 1 | -- | 0.93 | 0.61 | 0.62 | 0.76 | 0.86 | 0.85 |
| **6** | 0.85 | 0.92 | 0.92 | -- | 1 | -- | 0.71 | 0.88 | 0.88 | 0.8 | 0.92 | 0.91 |
| **7** | 0.58 | 0.9 | 0.6 | 0.36 | 0.89 | 0.48 | 0.73 | 0.3 | 0.56 | 0.6 | 0.55 | 0.59 |
| **8** | 0.78 | 0.88 | 0.88 | -- | 1 | -- | 0.7 | 0.75 | 0.75 | 0.63 | 0.88 | 0.87 |
| **9** | 0.93 | 0.86 | 0.86 | -- | 0.99 | -- | 0.68 | 0.8 | 0.8 | 0.68 | 0.86 | 0.87 |
| **10** | 0.82 | 0.82 | 0.82 | 0.54 | 0.82 | 0.81 | 0.87 | 0.66 | 0.75 | 0.81 | 0.83 | 0.82 |
| **11** | 0.83 | 0.95 | 0.95 | -- | 1 | -- | 0.69 | 0.9 | 0.9 | 0.56 | 0.95 | 0.93 |
| **12** | -- | 0.68 | -- | -- | 1 | -- | 0.73 | 0.57 | 0.57 | 0.73 | 0.68 | 0.68 |
| **13** | 0.73 | 0.73 | 0.73 | -- | 0.97 | -- | 0.93 | 0.46 | 0.47 | 0.74 | 0.73 | 0.73 |
| **14** | 0.6 | 0.75 | 0.73 | -- | 0.99 | -- | 0.7 | 0.5 | 0.5 | 0.6 | 0.75 | 0.73 |
| **15** | 0.55 | 0.95 | 0.94 | -- | 1 | -- | 1 | 0.91 | 0.91 | 0.56 | 0.95 | 0.94 |
| **AVG** | **0.72** | **0.84** | **0.82** | **0.60** | **0.94** | **0.85** | **0.81** | **0.70** | **0.72** | **0.68** | **0.82** | **0.81** |

Legend: SENS, sensitivity; SPEC, specificity; ACC, accuracy; AVG, weighted average based on the video frame number.

**Supplemental Material**

**Table 1: Effectiveness of SHIATSU for healthy controls.**

| | **PTOSIS** | | | **HEAD DROP** | | | **MOUTH OPENING** | | | **CATAPLEXY** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Video #** | **Sens.** | **Spec.** | **Acc.** | **Sens.** | **Spec.** | **Acc.** | **Sens.** | **Spec.** | **Acc.** | **Sens.** | **Spec.** | **Acc.** |
| **16** | -- | 1 | -- | -- | 1 | -- | -- | 1 | -- | -- | 1 | -- |
| **17** | -- | 1 | -- | -- | 1 | -- | -- | 1 | -- | -- | 1 | -- |
| **18** | -- | 1 | -- | -- | 1 | -- | -- | 0.99 | -- | -- | 1 | -- |
| **19** | -- | 1 | -- | -- | 1 | -- | -- | 1 | -- | -- | 1 | -- |
| **20** | -- | 1 | -- | -- | 1 | -- | -- | 1 | -- | -- | 1 | -- |
| **21** | -- | 1 | -- | -- | 1 | -- | -- | 0.98 | -- | -- | 1 | -- |
| **22** | -- | 1 | -- | -- | 1 | -- | -- | 1 | -- | -- | 1 | -- |
| **23** | -- | 1 | -- | -- | 1 | -- | -- | 0.99 | -- | -- | 1 | -- |
| **24** | -- | 0.87 | -- | -- | 1 | -- | -- | 0.79 | -- | -- | 0.87 | -- |
| **25** | -- | 0.96 | -- | -- | 0.97 | -- | -- | 0.94 | -- | -- | 0.97 | -- |
| **26** | -- | 1 | -- | -- | 1 | -- | -- | 1 | -- | -- | 1 | -- |
| **27** | -- | 1 | -- | -- | 1 | -- | -- | 0.98 | -- | -- | 1 | -- |
| **28** | -- | 1 | -- | -- | 1 | -- | -- | 0.98 | -- | -- | 1 | -- |
| **29** | -- | 1 | -- | -- | 1 | -- | -- | 0.99 | -- | -- | 1 | -- |
| **30** | -- | 1 | -- | -- | 1 | -- | -- | 1 | -- | -- | 1 | -- |
| **AVG** | **--** | **0.99** | **--** | **--** | **1** | **--** | **--** | **0.98** | **--** | **--** | **0.99** | **--** |

Legend: SENS, sensitivity; SPEC, specificity; ACC, accuracy; AVG, weighted average based on the video frame number.

**Supplemental material**

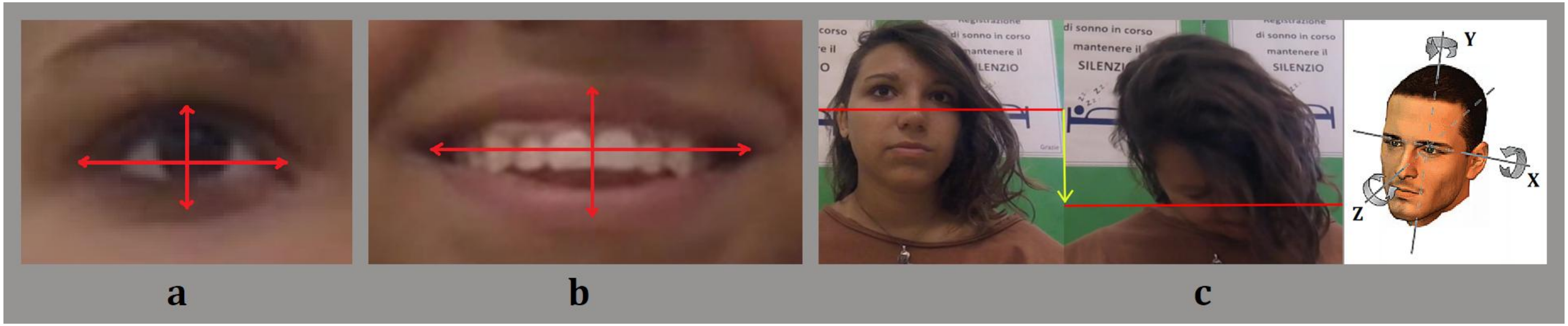
**Table 2: Comparison between the classifications performed by the domain specialist and SHIATSU.**

| | CATAPLEXY RECOGNIZED AS PRESENT | |
|---|---|---|
| **Video #** | **By the domain specialist** | **By SHIATSU** |
| 1 | Yes | Yes |
| 2 | Yes | Yes |
| 3 | Yes | Yes |
| 4 | Yes | Yes |
| 5 | Yes | Yes |
| 6 | Yes | Yes |
| 7 | Yes | Yes |
| 8 | Yes | Yes |
| 9 | Yes | Yes |
| 10 | Yes | Yes |
| 11 | Yes | Yes |
| 12 | Yes | Yes |
| 13 | Yes | Yes |
| 14 | Yes | Yes |
| 15 | Yes | Yes |
| 16 | No | No |
| 17 | No | No |
| 18 | No | No |
| 19 | No | No |
| 20 | No | No |
| 21 | No | No |
| 22 | No | No |
| 23 | No | No |
| 24 | No | Yes |
| 25 | No | Yes |
| 26 | No | No |
| 27 | No | No |
| 28 | No | No |
| 29 | No | No |
| 30 | No | No |

Legend: table summarizing the detection of the disease at the level of the entire video; Yes: at least one cataplectic episode detected; No: cataplectic episode detected.

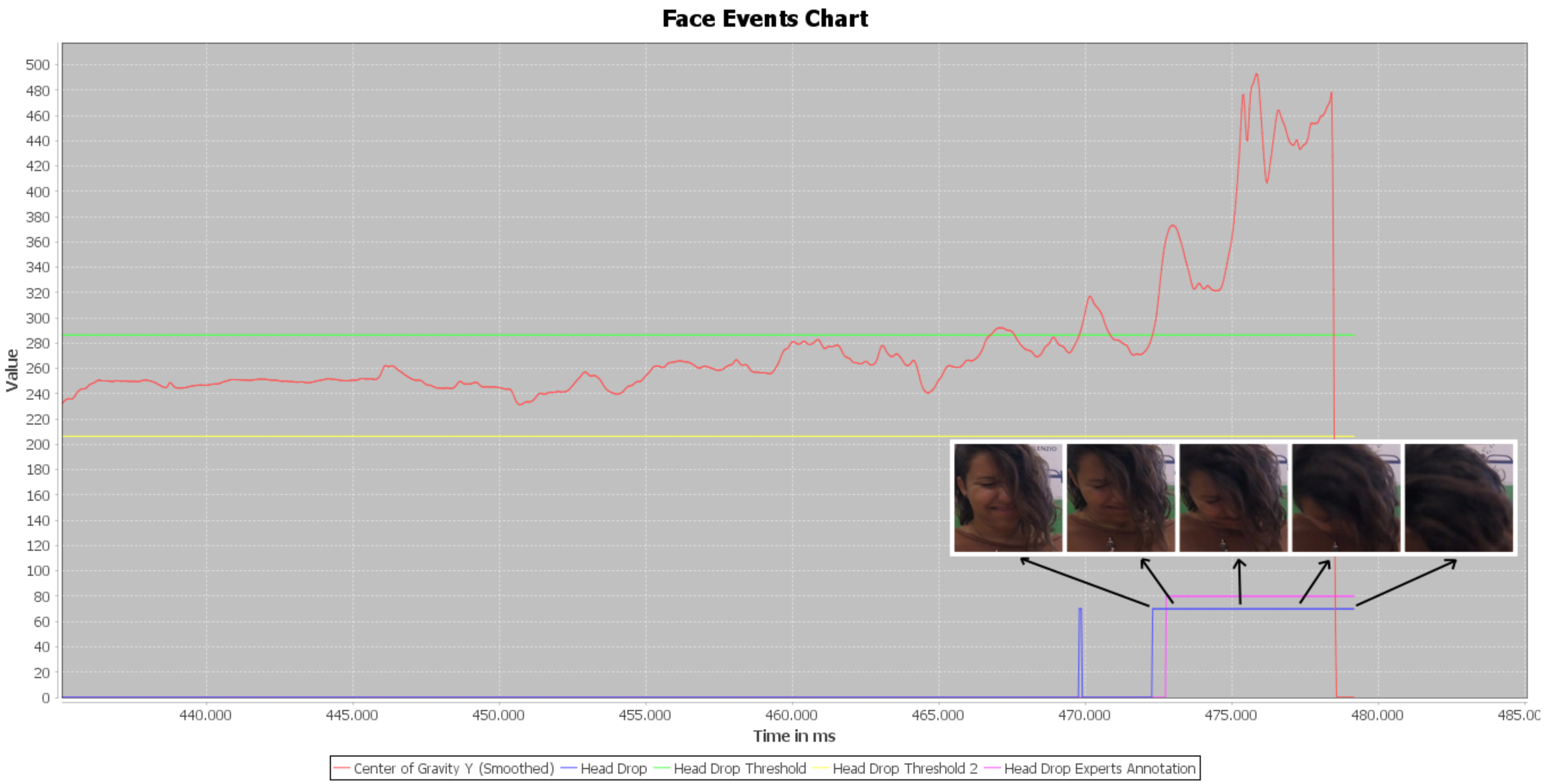
## Figures

**Figure 1: Features used for the detection of motor phenomena.**



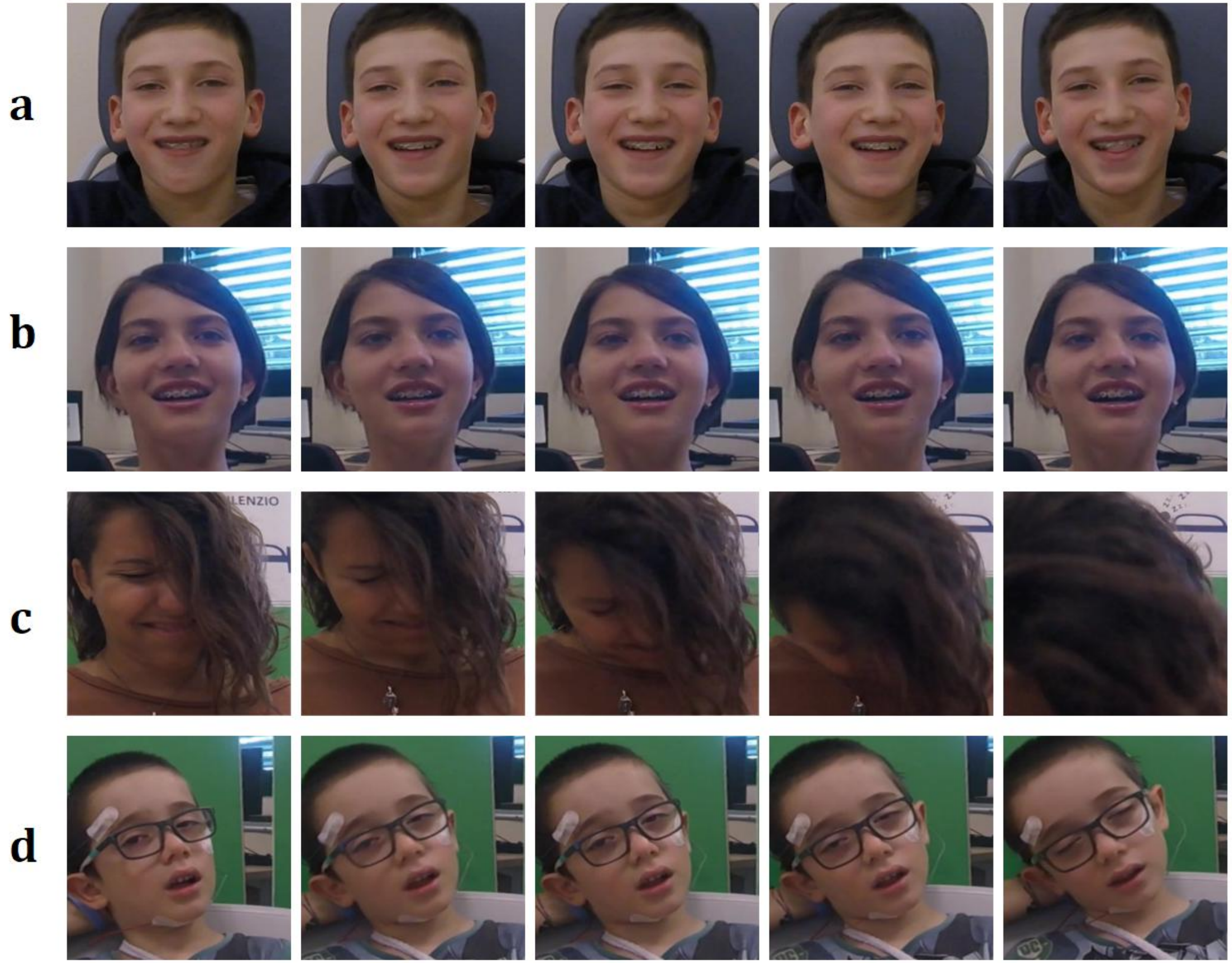

Legend: a) Eye Aspect Ratio (Ptosis); b) Mouth Aspect Ratio (Mouth opening); c) Position of the center of gravity and head rotation in Euler angles (Head drop).

**Figure 2: SHIATSU detection for motor patterns: an example of head drop detection.**




Legend: The horizontal axis represents the time coordinates (the time position within the video), the vertical axis represents the values of the center of gravity position with respect to its y coordinate. The red line represents the time series of the position of the center of gravity, while green and yellow lines are the maximum and minimum thresholds. The pink line indicates the presence of head drop detected by the domain experts (only two values, 0 or a positive number, chosen appropriately each time on a case by case basis so that the chart is as readable as possible), while the blue line indicates the presence of head drop detected by SHIATSU (two values: 0 and positive number slightly less than that chosen for the annotation of the expert of the domain, always to promote maximum readability). We decided not to show the part of the graph before 7 min and 20 seconds because neither the software nor the domain expert reported anything relevant. Please note that patient's images are only a visual example that we added to the graph to show the content of the frames detected as containing the motor phenomenon.

**Figure 3: Sequences of motor phenomena.**




Legend: a) Ptosis; b) Mouth Opening; c) Head Drop; d) all three motor phenomena at the same time.