Università Campus Bio-Medico di Roma

Corso di dottorato di ricerca in
Scienze Biomediche Integrate e Bioetica

XXXII ciclo a.a. 2016-2017

OBJECTIVE HYPOMIMIA BIOMARKER FOR PARKINSON'S DISEASE DIAGNOSIS

Lazzaro di Biase

Coordinatore
Prof Paolo Pozzilli

Tutor
Prof. Vincenzo Di Lazzaro

09 Luglio 2020

SUMMARY

Summary 0

Objective hypomimia biomarker for Parkinson's disease diagnosis 2

Abstract 2

Background 4

Parkinson's disease diagnosis 4

Hypomimia 9

Facial Action Coding System (FACS) 11

Automatic facial movement analysis 14

Openface 18

Aim 21

Materials and Methods 21

Patients 21

Face recording and analysis 21

Video check and analysis 23

Statistical analysis 23

Results 24

Video check 24

MDS-UPDRS III item 3.2 (facial expression) 24

Automatic face analysis 25

Rest 25

Conversation 27

Combined parameter 29

Discussion 34

Conclusion 36

References 37

OBJECTIVE HYPOMIMIA BIOMARKER FOR PARKINSON'S DISEASE DIAGNOSIS

Abstract

Background.

In vivo misdiagnosis in Parkinson's disease is one of the biggest unmet needs in this disease. In the last decades, several studies attempt to improve diagnostic accuracy by means of quantitative evaluations. Hypomimia is one of the earlier motor symptoms in Parkinson's disease that starts 10 years before clinical diagnosis. The aim of the present study is to use a technology able to automatically extract face features, in order to verify the accuracy of objective hypomimia biomarker for Parkinson's disease diagnosis.

Materials and Methods

Nine healthy subjects (HS), (age 53.3 ± 8.9 years), and twelve patients affected by Parkinson's disease (PD), according to UK PD Society Brain Bank diagnostic criteria (age 68.2 ± 6.4 years), with and H&Y stage range 1.5-2.5, were enrolled in the study. All enrolled subjects, were filmed, with a standard camera, under two different test conditions for 1 min for each task: (1) rest, (2) conversation. Face features were extracted from digitally recorded video images, focusing on blinking intensity and lips distance, and a final index merging both parameters for rest task was created.

Results

The maximum intensity of blinking showed to be lower in Parkinson's disease patients compared to controls, during rest task. The absolute value of maximum lips distance showed to be higher in Parkinson's disease patients compared to controls, during both rest and conversation task. The combined index, created using only rest parameters, defined as Parkinson's disease hypomimia predictor (PHP), showed a high diagnostic accuracy (95%) in Parkinson's disease vs. healthy subjects discrimination,

with an ROC AUC of 0,949, a positive predictive value (PPV) of 92% and a negative predictive value (NPV) of 100%. The correlation between the MDS-UPRS III item 3.2 and PHP was $r_s(18) = 0{,}738$.

Discussion

Both literature and the present study data showed that hypomimia in Parkinson's disease patients, is a good candidate as a proxy symptoms for diagnosis of Parkinson's disease. Quantitative and objective assessments make assessments more accurate and reproducible. In the present study, the data were collected in the less intrusive way, by using only a standard camera, without markers placed on the face of subjects. Custom parameters were calculated from the extracted face features, focusing on the most relevant hypomimia features, in line to face features evaluated in clinical practice and MDS-UPDRS scale, i.e. blinking and lips movements. The final predictor (PHP) showed a high diagnostic accuracy and, in addition, the variable MDS-UPRS III item 3.2 and PHP were found to be strongly correlated, showing that PHP could be a useful objective tool to evaluate hypomimia in Parkinson's disease.

Conclusion

In conclusion, PHP is a new hypomimia measure, which can be an aid tool for the diagnosis of Parkinson's disease. This new tool has a high diagnostic accuracy, positive and negative predictive value. It can be derived from short, cheap, widely available and non-intrusive face recordings at rest, and it is correlated to standard clinical motor scale scoring system.

BACKGROUND

Parkinson's disease diagnosis

Parkinson's disease diagnosis relay on clinical diagnostic criteria, the UK Brain Bank criteria, which are the only criteria validated from post-mortem pathology examination. (Gibb and Lees, 1988b; Hughes *et al.*, 1992; Hughes *et al.*, 2002) (Figure 1)

STEP 1. Diagnosis of PARKINSONIAN SYNDROME.
BRADYKINESIA (slowness of initiation of voluntary movement with progressive reduction in speed and amplitude of repetitive actions).
And at least one of the following:
a. muscular rigidity
b. 4–6 Hz rest tremor
c. postural instability not caused by primary visual, vestibular, cerebellar or proprioceptive dysfunction.

STEP 2. Exclusion criteria for Parkinson's disease.
history of repeated strokes with stepwise progression of Parkinsonian features
history of repeated head injury
history of definite encephalitis
oculogyric crises
neuroleptic treatment at onset of symptoms
more than one affected relative
sustained remission
strictly unilateral features after three years
supranuclear gaze palsy
cerebellar signs
early severe autonomic involvement
early severe dementia with disturbances of memory, language and praxis
Babinski sign
presence of a cerebral tumour or communicating hydrocephalus on CT scan.
negative response to large doses of levodopa (if malabsorption excluded)
MPTP exposure

STEP 3. Supportive prospective positive criteria for PARKINSON'S DISEASE. Three or more required for diagnosis of definite Parkinson's disease.
unilateral onset
rest tremor present
progressive disorder
persistent asymmetry affecting the side of onset most
excellent response (70–100%) to levodopa
severe levodopa-induced chorea
levodopa response for 5 years or more
clinical course of 10 years or more

Figure 1 UK Parkinson's Disease Society Brain Bank clinical diagnostic criteria (Gibb and Lees, 1988a)

However, in vivo misdiagnosis in Parkinson's disease is one of the biggest unmet needs in this disease. Applying the UK Brain Bank criteria, misdiagnosis range from 20 to 30%, according to if the diagnosis is made by a movement disorders expert or by nonexperts. (Rizzo *et al.*, 2016)

Several attempts have been made to better define, and improve the accuracy of the diagnostic criteria.

Gelb *et al.* (1999) proposed three levels of diagnosis: possible or probable based on clinical diagnosis alone, and definite diagnosis, which requires histopathologic confirmation (Figure 2-4).

Group A features: characteristic of Parkinson disease
- Resting tremor
- Bradykinesia
- Rigidity
- Asymmetric onset

Group B features: suggestive of alternative diagnoses
- Features unusual early in the clinical course
  - Prominent postural instability in the first 3 years after symptom onset
  - Freezing phenomena in the first 3 years
  - Hallucinations unrelated to medications in the first 3 years
  - Dementia preceding motor symptoms or in the first year
- Supranuclear gaze palsy (other than restriction of upward gaze) or slowing of vertical saccades
- Severe, symptomatic dysautonomia unrelated to medications
- Documentation of a condition known to produce parkinsonism and plausibly connected to the patient's symptoms (such as suitably located focal brain lesions or neuroleptic use within the past 6 months)

Figure 2 Grouping of Clinical Features According to Diagnostic Utility (Gelb et al., 1999)

Criteria for POSSIBLE diagnosis of Parkinson disease:

At least 2 of the 4 features in Group A* are present; at least 1 of these is tremor or bradykinesia

and

Either None of the features in Group B* is present
Or Symptoms have been present for less than 3 years, and none of the features in Group B* is present to date

and

Either Substantial and sustained response to levodopa or a dopamine agonist has been documented
Or Patient has not had an adequate trial of levodopa or dopamine agonist

Criteria for PROBABLE diagnosis of Parkinson disease:

At least 3 of the 4 features in Group A* are present

and

None of the features in Group B* is present (note: symptom duration of at least 3 years is necessary to meet this requirement)

and

Substantial and sustained response to levodopa or a dopamine agonist has been documented

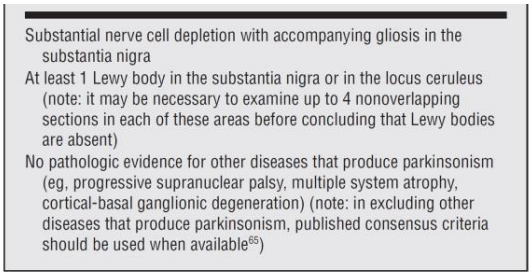
Criteria for DEFINITE diagnosis of Parkinson disease:

All criteria for POSSIBLE Parkinson disease are met

and

Histopathologic confirmation of the diagnosis is obtained at autopsy (see Table 3)

Figure 3 Proposed Diagnostic Criteria for Parkinson Disease (Gelb et al., 1999)



Substantial nerve cell depletion with accompanying gliosis in the substantia nigra

At least 1 Lewy body in the substantia nigra or in the locus ceruleus (note: it may be necessary to examine up to 4 nonoverlapping sections in each of these areas before concluding that Lewy bodies are absent)

No pathologic evidence for other diseases that produce parkinsonism (eg, progressive supranuclear palsy, multiple system atrophy, cortical-basal ganglionic degeneration) (note: in excluding other diseases that produce parkinsonism, published consensus criteria should be used when available[65])

Figure 4 Proposed Criteria for Histopathologic Confirmation (Gelb et al., 1999)

The three cardinal clinical motor features for the diagnosis of Parkinson's disease are bradykinesia, rigidity and resting tremor. Ward and Gibb (1990) found that among patients that in vivo show 2 of 3 cardinal motor features only 69-75% had a postmortem diagnosis of Parkinson's disease.

The International Parkinson and Movement Disorder Society, proposed in 2015 a revision of the diagnostic criteria (Postuma *et al.*, 2015) (Figure 5). The benchmark for these new criteria is expert clinical diagnosis, and therefore are not validated with postmortem anatomopathological diagnosis. These criteria define two level of clinical diagnosis: probable and established. Also for these new criteria, the core of the diagnosis is the motor parkinsonism defined as bradykinesia, in combination with at least one of rest tremor or rigidity. (Postuma *et al.*, 2015) For both probable and established diagnosis is needed the absence of absolute exclusion criteria. In addition, for established diagnosis, no red flags, and two supportive criteria are required, in order to increase the specificity of the diagnosis. For probable diagnosis, in order to balances sensitivity and specificity, no more than 2 red flags are allowed, with at least an equal number of red flags and supportive criteria.

The first essential criterion is parkinsonism, which is defined as bradykinesia, in combination with at least 1 of rest tremor or rigidity. Examination of all cardinal manifestations should be carried out as described in the MDS–Unified Parkinson Disease Rating Scale.[30] Once parkinsonism has been diagnosed:

Diagnosis of Clinically Established PD requires:
1. Absence of absolute exclusion criteria
2. At least two supportive criteria, and
3. No red flags

Diagnosis of Clinically Probable PD requires:
1. Absence of absolute exclusion criteria
2. Presence of red flags counterbalanced by supportive criteria
   If 1 red flag is present, there must also be at least 1 supportive criterion
   If 2 red flags, at least 2 supportive criteria are needed
   No more than 2 red flags are allowed for this category

Supportive criteria
(Check box if criteria met)

☐ 1. Clear and dramatic beneficial response to dopaminergic therapy. During initial treatment, patient returned to normal or near-normal level of function. In the absence of clear documentation of initial response a dramatic response can be classified as:
   a) Marked improvement with dose increases or marked worsening with dose decreases. Mild changes do not qualify. Document this either objectively (>30% in UPDRS III with change in treatment), or subjectively (clearly-documented history of marked changes from a reliable patient or caregiver).
   b) Unequivocal and marked on/off fluctuations, which must have at some point included predictable end-of-dose wearing off.
☐ 2. Presence of levodopa-induced dyskinesia
☐ 3. Rest tremor of a limb, documented on clinical examination (in past, or on current examination)
☐ 4. The presence of either olfactory loss or cardiac sympathetic denervation on MIBG scintigraphy

Absolute exclusion criteria: The presence of any of these features rules out PD:

☐ 1. Unequivocal cerebellar abnormalities, such as cerebellar gait, limb ataxia, or cerebellar oculomotor abnormalities (eg, sustained gaze evoked nystagmus, macro square wave jerks, hypermetric saccades)
☐ 2. Downward vertical supranuclear gaze palsy, or selective slowing of downward vertical saccades
☐ 3. Diagnosis of probable behavioral variant frontotemporal dementia or primary progressive aphasia, defined according to consensus criteria[31] within the first 5 y of disease
☐ 4. Parkinsonian features restricted to the lower limbs for more than 3 y
☐ 5. Treatment with a dopamine receptor blocker or a dopamine-depleting agent in a dose and time-course consistent with drug-induced parkinsonism
☐ 6. Absence of observable response to high-dose levodopa despite at least moderate severity of disease
☐ 7. Unequivocal cortical sensory loss (ie, graphesthesia, stereognosis with intact primary sensory modalities), clear limb ideomotor apraxia, or progressive aphasia
☐ 8. Normal functional neuroimaging of the presynaptic dopaminergic system
☐ 9. Documentation of an alternative condition known to produce parkinsonism and plausibly connected to the patient's symptoms, or, the expert evaluating physician, based on the full diagnostic assessment feels that an alternative syndrome is *more likely* than PD

Red flags

☐ 1. Rapid progression of gait impairment requiring regular use of wheelchair within 5 y of onset
☐ 2. A complete absence of progression of motor symptoms or signs over 5 or more y unless stability is related to treatment
☐ 3. Early bulbar dysfunction: severe dysphonia or dysarthria (speech unintelligible most of the time) or severe dysphagia (requiring soft food, NG tube, or gastrostomy feeding) within first 5 y
☐ 4. Inspiratory respiratory dysfunction: either diurnal or nocturnal inspiratory stridor or frequent inspiratory sighs
☐ 5. Severe autonomic failure in the first 5 y of disease. This can include:
   a) Orthostatic hypotension[32]—orthostatic decrease of blood pressure within 3 min of standing by at least 30 mm Hg systolic or 15 mm Hg diastolic, in the absence of dehydration, medication, or other diseases that could plausibly explain autonomic dysfunction, or
   b) Severe urinary retention or urinary incontinence in the first 5 y of disease (excluding long-standing or small amount stress incontinence in women), that is not simply functional incontinence. In men, urinary retention must not be attributable to prostate disease, and must be associated with erectile dysfunction
☐ 6. Recurrent (>1/y) falls because of impaired balance within 3 y of onset
☐ 7. Disproportionate anterocollis (dystonic) or contractures of hand or feet within the first 10 y
☐ 8. Absence of any of the common nonmotor features of disease despite 5 y disease duration. These include sleep dysfunction (sleep-maintenance insomnia, excessive daytime somnolence, symptoms of REM sleep behavior disorder), autonomic dysfunction (constipation, daytime urinary urgency, symptomatic orthostasis), hyposmia, or psychiatric dysfunction (depression, anxiety, or hallucinations)
☐ 9. Otherwise-unexplained pyramidal tract signs, defined as pyramidal weakness or clear pathologic hyperreflexia (excluding mild reflex asymmetry and isolated extensor plantar response)
☐ 10. Bilateral symmetric parkinsonism. The patient or caregiver reports bilateral symptom onset with no side predominance, and no side predominance is observed on objective examination

Criteria Application:

1. Does the patient have parkinsonism, as defined by the MDS criteria? Yes ☐ No ☐
   If no, *neither* probable PD nor clinically established PD can be diagnosed. *If yes:*
2. Are any absolute exclusion criteria present? Yes ☐ No ☐
   If "yes," *neither* probable PD nor clinically established PD can be diagnosed. *If no:*
3. Number of red flags present ____
4. Number of supportive criteria present ____
5. Are there at least 2 supportive criteria *and* no red flags? Yes ☐ No ☐
   If yes, patient meets criteira for clinically established PD. *If no:*
6. Are there more than 2 red flags? Yes ☐ No ☐
   If "yes," probable PD *cannot* be diagnosed. *If no:*
7. Is the number of red flags equal to, or less than, the number of supportive criteria? Yes ☐ No ☐
   If yes, patient meets criteria for probable PD

Figure 5 Movement Disorders Society Clinical Diagnostic Criteria for Parkinson's Disease .(Postuma et al., 2015)

All the proposed clinical diagnostic criteria, relay on qualitative clinical evaluation, with high variability regarding accuracy between expert and non-expert in movement disorders. (Rizzo *et al.*, 2016)

In the last decades, several studies attempt to improve diagnostic accuracy by means of quantitative evaluations. (Sánchez-Ferro *et al.*, 2016) Non-invasive motion analysis technologies, such as wearable sensors, can objectively extract the features of the cardinal motor symptoms in Parkinson's disease patients, like resting tremor (Deuschl *et al.*, 1996; Sánchez-Ferro *et al.*, 2016) (di Biase *et al.*, 2017), bradykinesia (di Biase *et al.*, 2018), rigidity (Sánchez-Ferro *et al.*, 2016) (di Biase *et al.*, 2018), in addittion to balance, gait and posture (Sánchez-Ferro *et al.*, 2016).

Another open issue is that, to date, the clinical diagnosis is made with delay. Diagnostic accuracy in early disease stage is even lower, in studies including early diagnosis of PD patients the percentage of patients with a confirmed diagnosis of PD at autopsy was between 38% and 65%. (Rajput *et al.*, 1991; Jankovic *et al.*, 2000; Adler *et al.*, 2014; Beach and Adler, 2018)

With quantitative motor tests, Parkinson's disease diagnosis could be made 3 years before the clinical diagnosis with 71–82% sensitivity and specificity. (Postuma *et al.*, 2012)

However, despite, a huge number of studies, focusing on different technologies, showed in the last decades, a high level of accuracy in Parkinson's disease diagnosis and motor function characterization, none of these technologies is routinely used in clinical practice.

Hypomimia

Hypomimia is one of the earlier motor symptoms in Parkinson's disease patients. It is characterized by an impairment of spontaneous, voluntary and emotional face movements. Very early diagnostic hypothesis sometimes relies only on this sign, the "masked face". The main signs characterizing hypomimia in Parkinson's disease are (Goetz *et al.*, 2007; Agostino *et al.*, 2008; Jankovic, 2008; Bologna *et al.*, 2013; Marsili *et al.*, 2014):

1. staring expression, with wider palpebral fissures
2. decreased frequency of blinking
3. slowed switching between closing and opening blinking phase
4. flattened nasolabial folds
5. reduced wrinkles on the orbicularis oris
6. fewer movements around the mouth, such as less spontaneous smiling
7. reduced amplitude and speed of posed smiling and grinning
8. lips parted when the mouth is at rest

It is widely described that PD patients show a low blink rate (Karson, 1983; Deuschl and Goddemeier, 1998; Altiparmak *et al.*, 2006; Korosec *et al.*, 2006), however with disease course, in the advanced phase, the blink rate can be abnormally high (Kimber and Thompson, 2000; Suppa *et al.*, 2017).

L-dopa does not modify voluntary facial movements kinematics (Marsili *et al.*, 2014; Suppa *et al.*, 2017), but shows a modulatory effect on spontaneous blinking, increasing the blink rate when abnormally decreased (Korosec *et al.*, 2006) and vice versa (Kimber and Thompson, 2000; Agostino *et al.*, 2008; Suppa *et al.*, 2017).

To date, during clinical practice, semiquantitative scoring system, like UPDRS (Fahn and Elton, 1987) or the MDS-UPDRS(Goetz *et al.*, 2008) are the most objective and standardized instrument widely available. Among the different features evaluated by the UPDRS, face impairment, seem to be a prodromal marker, which starts 10 years before clinical diagnosis. (Figure 6, E) (Postuma *et al.*, 2012)

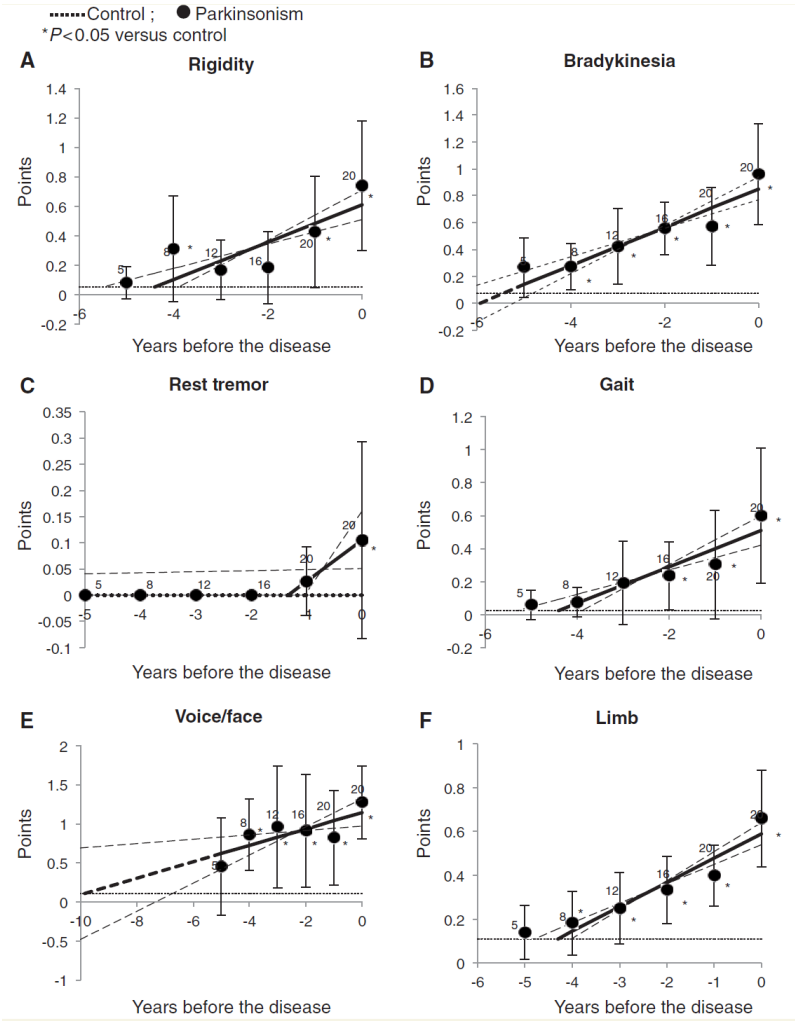


*Figure 6 Progression of UPDRS motor markers in the 5 years before diagnosis of parkinsonism. Progression of UPDRS measures are subdivided according to cardinal manifestation and location. Error bars represent SD. Confidence intervals (90%) of the slope are represented by the thin dashed lines. Reference value is the mean control value (horizontal line). The number beside each point represents the number of observations at this time interval. *Significantly different from control values on non-parametric the non-parametric Mann–Whitney U-test. (A) Rigidity, (B) bradykinesia, (C) rest tremor, (D) gait, (E) voice/face and (F) limb. (Postuma et al., 2012)*

However, a feature like hypomimia is not easily detected with clinical evaluation, which of course is not able to detect early impairment. The MDS-UPDRS item 3.2, evaluate facial expression in a semiquantitative way, with a score from 0 (normal face expression) to 5 (masked facies with lips parted most of the time when the mouth is at rest). This scoring system relies only on subjective evaluation, based on the eye of the examiner, which is variable in accuracy according to several factors,

among which the most important is the examiner experience in Parkinson's disease evaluation. Basically, we ask the examiner, in this task to make a pattern recognition of hypomimia. New technologies showed to be able to characterize these motor features, and therefore could be useful to improve the diagnostic accuracy of Parkinson's disease detection. (Hamm *et al.*, 2011; Wu *et al.*, 2014)

Facial Action Coding System (FACS)

The Facial Action Coding System (Ekman *et al.*, 1972; Ekman *et al.*, 2002) is an anatomically based tool which taxonomize facial movement, determined by a single or a group of facial muscles (Figure 7).

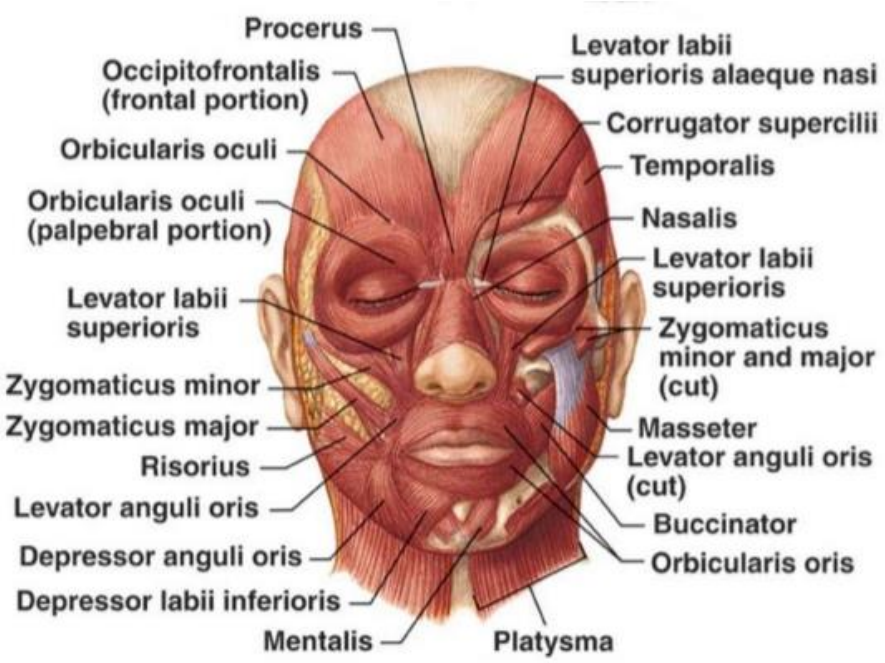


Figure 7 Facial muscles

It categorizes 44 unique Action Units (AUs), in addition to categories of head and eye movements. (Table 1-2) (Ekman *et al.*, 1972; Ekman *et al.*, 2002).

A single muscle or a muscle groups can lead to different AUs according to the way how the muscle contract, in which region and with a determined intensity. Each AUs can show different intensities on a five point scale. (Ekman *et al.*, 1972; Ekman *et al.*, 2002)

Table 1 Single action units (AU) in the Facial Action Coding System (Ekman et al., 1972; Ekman et al., 2002)

AU number	Descriptor	Muscular Basis
1.	Inner Brow Raiser	Frontalis, Pars Medialis
2.	Outer Brow Raiser	Frontalis, Pars Lateralis
4.	Brow Lowerer	Depressor Glabellae, Depressor Supercilli; Corrugator
5.	Upper Lid Raiser	Levator Palpebrae Superioris
6.	Cheek Raiser	Orbicularis Oculi, Pars Orbitalis
7.	Lid Tightener	Orbicularis Oculi, Pars Palebralis
9.	Nose Wrinkler	Levator Labii Superioris, Alaeque Nasi
10.	Upper Lip Raiser	Levator Labii Superioris, Caput Infraorbitalis
11.	Nasolabial Fold Deepener	Zygomatic Minor
12.	Lip Corner Puller	Zygomatic Major
13.	Cheek Puffer	Levator anguli oris
14.	Dimpler	Buccinator
15.	Lip Corner Depressor	Depressor anguli oris
16.	Lower Lip Depressor	Depressor Labii
17.	Chin Raiser	Mentalis
18.	Lip Puckerer	Incisivii Labii Superioris; Incisivii Labii Inferioris
20.	Lip Stretcher	Risorius
22.	Lip Funneler	Orbicularis Oris
23.	Lip Tightener	Orbicularis Oris
24.	Lip Pressor	Orbicularis Oris
25.	Lips Part	Depressor Labii, or Relaxation of Mentalis or Orbicularis Oris
26.	Jaw Drop	Masetter; Temporal and Internal Pterygoid Relaxed
27.	Mouth Stretch	Pterygoids; Digastric
28.	Lip Suck	Orbicularis Oris

Table 2 More grossly defined AUs in the Facial Action Coding System (Ekman et al., 1972; Ekman et al., 2002)

AU number	FACS name
8.	Lips Toward Each Other
19.	Tongue Out
21.	Neck Tightener
29.	Jaw Thrust
30.	Jaw Sideways
31.	Jaw Clencher
32.	Lip Bite
33.	Blow
34.	Puff
35.	Cheek Suck
36.	Tongue Bulge
37.	Lip Wipe
38.	Nostril Dilator
39.	Nostril Compressor
43.	Eyes Closure
45.	Blink
46.	Wink

The seven cardinal emotions (surprise, happiness, contempt, fear, sadness, disgust, anger) are encoded by facial muscle in a spontaneous, stereotyped and universal way, since different people will react with the same facial muscular pattern to the same emotion (Figure 8).

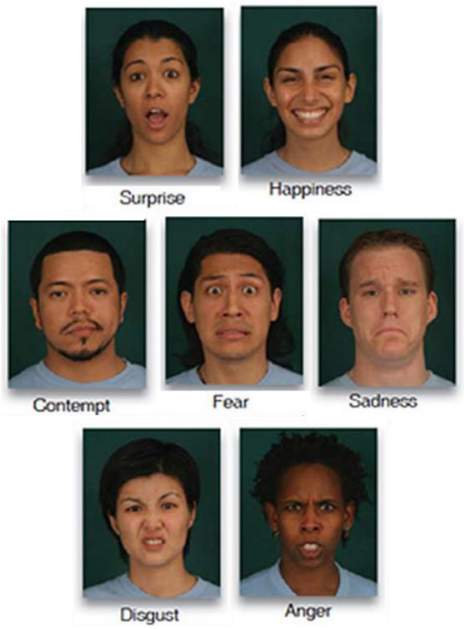


Figure 8 The seven cardinal emotions (surprise, happiness, contempt, fear, sadness, disgust, anger)

This seven cardinal emotions can be recognized and encoded according to FACSAID (Facial Action Coding System Affect Interpretation Dictionary) (Ekman *et al.*, 1998) and EMFACS (Emotional Facial Action Coding System) (Friesen and Ekman, 1983) (Table 3).

Table 3 AUs which codify the seven cardinal emotions.

Emotion	AUs number
Surprise	1; 2; 5; 26
Happiness	6; 12
Contempt	R12A; R14A
Fear	1; 2; 4; 5; 7; 20; 26
Sadness	1; 4;15
Disgust	9; 15; 16
Anger	4; 5; 7; 23

Automatic facial movement analysis

A taxonomic system like the Facial Action Coding System is a very useful tool to analyze facial movement, but the way how the system is used to extract facial features can be manual, made by a human, or automatic (computer aided). The recognition of face movement and expression, historically has been made by humans, however modern computer vision techniques combined with state of the art machine learning algorithms can automatically recognize, define and categorize single face segments. (Amos *et al.*, 2016)

In order to perform a full automatic facial movement analysis, the combined software and hardware used need to be able to perform the following task (Bettadapura, 2012; Sariyanidi *et al.*, 2014):

1. Face detection and tracking.
2. Dynamic face feature extraction.

In addition to these tasks, the software can perform expression recognition by combining features extracted at AUs level (Bandini *et al.*, 2017).

Few studies tried to create an automatic system to detect hypomimia in different conditions in Parkinson's disease patients; however, to date no standard technique has been defined and approved for the use in clinical practice.

Automatic face detection and tracking, can be performed with different algorithms, like Active Appearance Model (AAM) (based on Principal Component Analysis - PCA) (Cootes *et al.*, 2001; Matthews and Baker, 2004; Sariyanidi *et al.*, 2014), the constrained local model (CLM) (Cristinacce and Cootes, 2006), the Supervised Descending Method (SDM) (Xiong and De la Torre, 2013), and the particle filtering (Patras and Pantic, 2004; Bandini *et al.*, 2017).

Wu *et al.* (2014) used facial surface EMG signals and videotape analysis, through AdaBoost feature selection and binary SVM classification to detect the facial action units (AUs) (Figure 9-10), in order to detect automatically the AUs and their intensity, comparing 7 PD patients with 8 healthy subjects. The data were recorded during an emotion induction task, through videoclips able to elicit six classic emotions (amusement, sadness, anger, disgust, surprise, and fear). Authors used the total facial activity measure combining the total number of displayed AUs in response to the stimuli, and showed an attenuation of Parkinson's disease facial activities compared to controls. (Wu *et al.*, 2014)

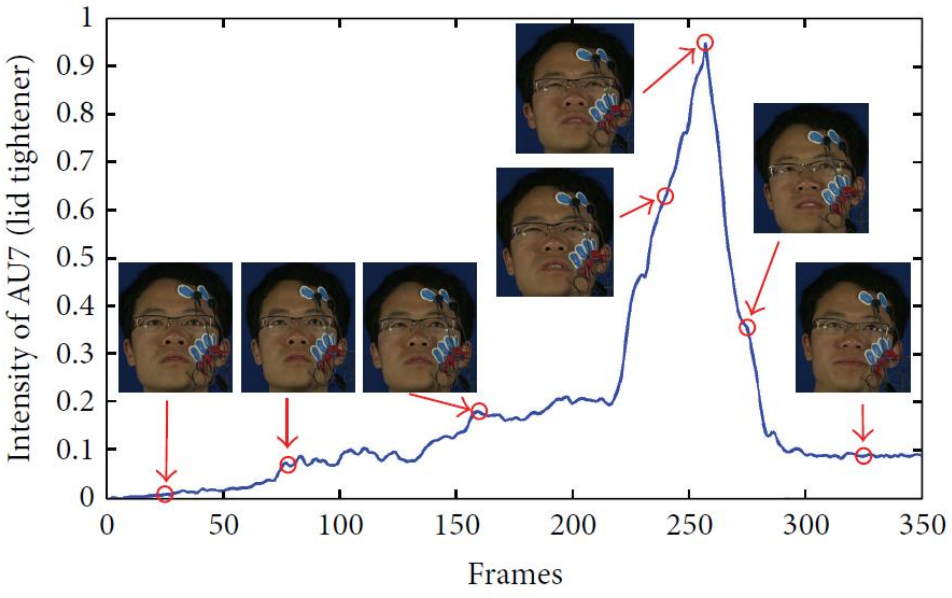


Figure 9 Wu et al. (2014) AUs intensity detection from EMG signal.

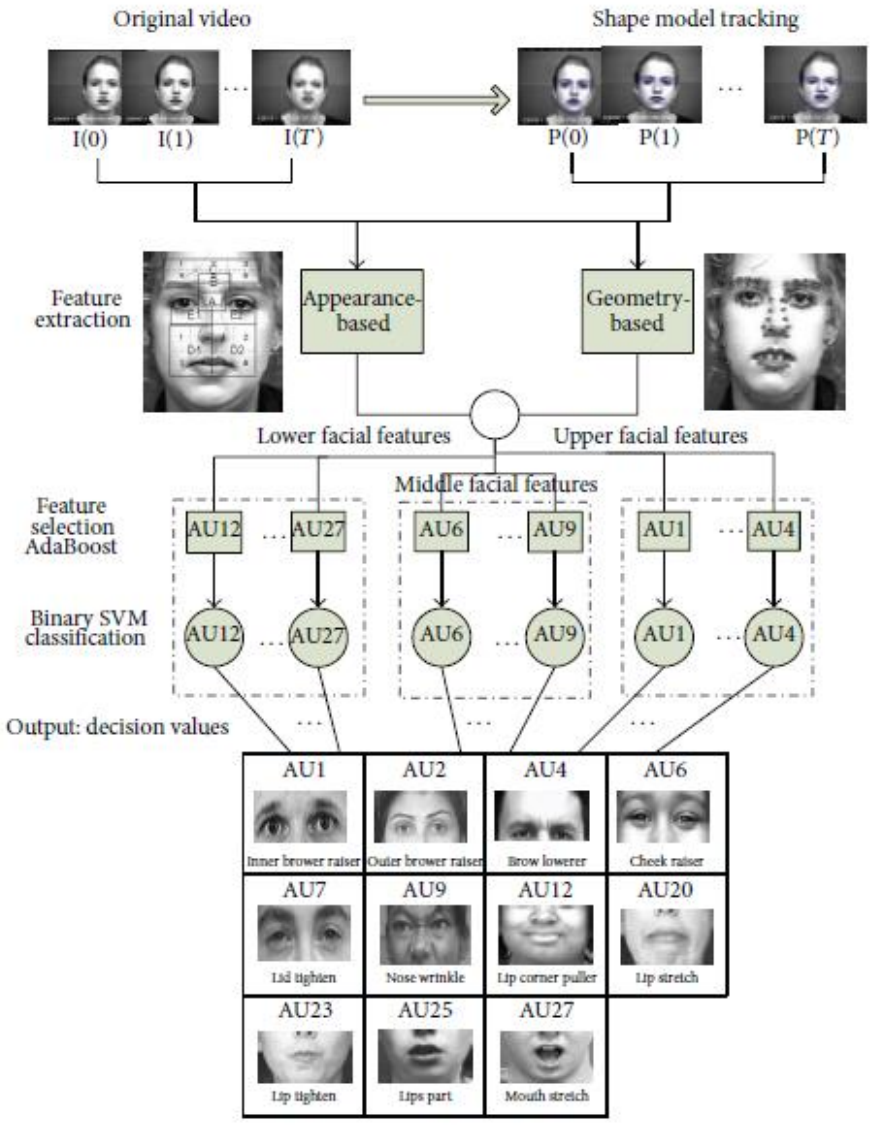


Figure 10 Wu et al. (2014) automatic AUs detection from videotape

Vinokurov N (2015) described a method to automatically detect hypomimia, through a commercial software (Faceshift) which is able to generate facial features with real-time face tracking, and a depth camera to collect data. Compared to the evaluation of hypomimia made by a neurologist the algorithm showed a correlation that range from 0,69 to 0,84.

Another study made by Marsili *et al.* (2014) analyzed posed smiling and voluntary grinning, in PD patients and control subjects, through a 3D optoelectronic

system, composed by three infrared cameras and six reflective markers placed on the subjects face, this study showed an abnormal reduction of these movements in PD patients compared to controls. Bologna *et al.* (2016) analyzed the kinematics of facial expression of six emotions (anger, disgust, fear, happiness, sadness, surprise), by means of a 3D optoelectronic system, composed by three infrared cameras and 21 reflective markers placed on face, results showed that PD patients movements, for all emotions, had slower velocity and lower amplitude in comparison to controls.

A study focusing on hypokinetic dysarthria in Parkinson's disease patients, analyzed lower lip peak velocities and accelerations during a syllable repetition task, using a depth camera (Microsoft Kinect) for face detection and tracking, this study showed that PD patients have reduced peak velocities and acceleration of the lower lip respect to control subjects (Bandini *et al.*, 2016). Bandini *et al.* (2017) through Microsoft Kinect sensor using only the color stream, have analyzed the average distance of the face model of Parkinsons's disease patients and control subjects from a neutral baseline, in addition to emotion recognition through an automatic facial expressions recognition algorithm, during three task: neutral expression, posed and imitated basic facial expressions (happiness, anger, disgust and sadness). Results showed that PD patients have on average lower distance from the face model neutral baseline respect to controls, and that anger and disgust facial expressions are the most impaired in PD patients.

Katsikitis and Pilowsky (1988) used a facial mathematical model to analyze face expression, results of this study showed a decrease of mouth opening and of the frequency of smiling in PD patients compared to controls. Bowers *et al.* (2006) analyzed, through custom software to extract face features and black and white camera to collect data, voluntary emotional expressions (happy, disgust, fear, sad, angry, and surprise). Authors computed expression entropy, starting from the frame-by-frame intensity change during the course of expression, showing that PD patients had lower facial movement (entropy) and were slower in reaching a peak expression compared to controls.

Langevin *et al.* (2019) using Openface algorithm for video analysis, at rest and during hands motor task, found an increase in AU4 at rest and a decrease in AU12

during the pronation-supination movements of hands task in Parkinson's disease patients respect to controls.

Openface

Openface 2.0 algorithm (Amos *et al.*, 2016) recognize the face landmarks (Baltrusaitis *et al.*, 2013; Zadeh *et al.*, 2017) (Figure 11), subdivided in 68 points, from 0 to 67, in 2D space (x,y pixels) and in 3D space (X,Y,Z millimeters) (Figure 12). This algorithm has been validated with machine-learning techniques, on more than 3500 images to detect face landmarks, in addition to 8000 images validation on poor illuminations condition (Baltrusaitis *et al.*, 2013).

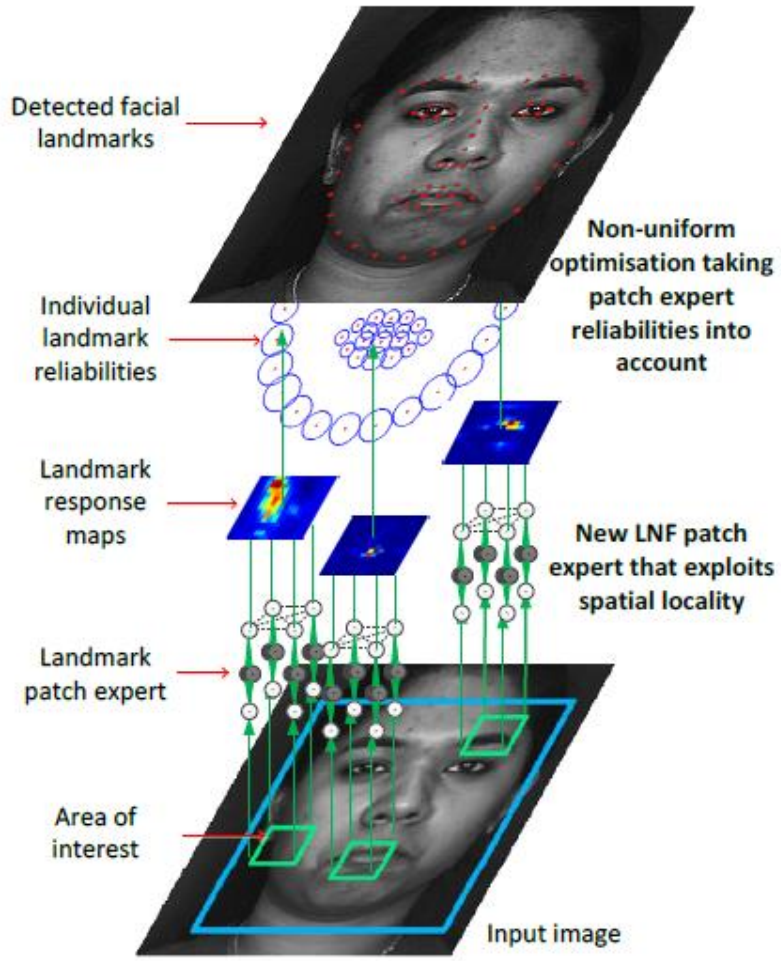


Figure 11 Face landmarks detection system in Openface (Baltrusaitis et al.*, 2013).*

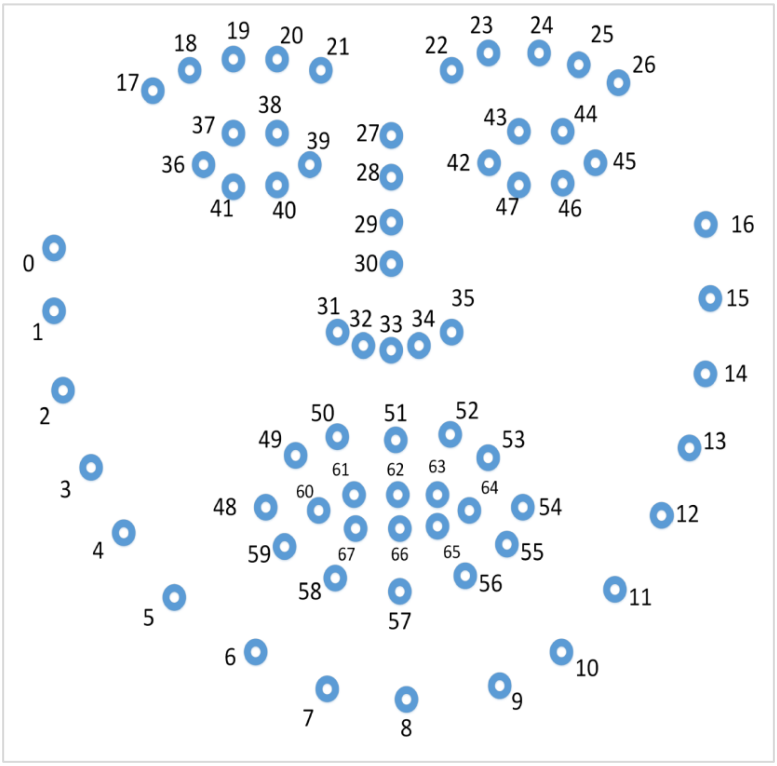


Figure 12 Face landmarks (Amos et al., *2016)*

The algorithm automatically recognize eye landmarks, subdivided in 56 points, from 0 to 55, in 2D space (x,y pixels) and in 3D space (X,Y,Z millimeters) (Figure 13).

Figure 13 Eye landmarks (Amos et al., *2016)*

Gaze is encoded in 3D world coordinates (x, y, z) for leftmost eye in the image (eye 0) and for the rightmost eye in the image (eye 1), in addition gaze direction is encoded in radians in 2D world coordinates (x, y) averaged for both eyes (x: form left

to right = angle from positive to negative; y: from up to down angle from negative to positive) (Wood et al., 2015).

The algorithm can recognize in automatic Action Units, in terms of occurrence (c: 0 absent, 1 present) and intensity (r: from 0 to 5) of 17 AUs (Baltrušaitis et al., 2015) (Table 4). The algorithm has been validated with machine-learning techniques for AUs detection, on three datasets respectively composed by: 150.000; 93.000; 130.000 AU labeled images (Baltrušaitis et al., 2015).

Table 4 AUs identified by openface algorithm

AU number	Descriptor	Muscular Basis
1.	Inner Brow Raiser	Frontalis, Pars Medialis
2.	Outer Brow Raiser	Frontalis, Pars Lateralis
4.	Brow Lowerer	Depressor Glabellae, Depressor Supercilli; Corrugator
5.	Upper Lid Raiser	Levator Palpebrae Superioris
6.	Cheek Raiser	Orbicularis Oculi, Pars Orbitalis
7.	Lid Tightener	Orbicularis Oculi, Pars Palebralis
9.	Nose Wrinkler	Levator Labii Superioris, Alaeque Nasi
10.	Upper Lip Raiser	Levator Labii Superioris, Caput Infraorbitalis
12.	Lip Corner Puller	Zygomatic Major
14.	Dimpler	Buccinator
15.	Lip Corner Depressor	Depressor anguli oris
17.	Chin Raiser	Mentalis
20.	Lip Stretcher	Risorius
23.	Lip Tightener	Orbicularis Oris
25.	Lips Part	Depressor Labii, or Relaxation of Mentalis or Orbicularis Oris
26.	Jaw Drop	Masetter; Temporal and Internal Pterygoid Relaxed
28.*	Lip Suck	Orbicularis Oris
45.	Blink	Relaxation of *Levator Palpebrae and Contraction of Orbicularis Oculi, Pars Palpebralis.*

*Defined only the occurrence

AIM

The aim of the present study is to use a technology able to automatically extract face features, in order to verify the accuracy of objective hypomimia biomarker for Parkinson's disease diagnosis.

MATERIALS AND METHODS

Patients

All patients gave informed consent and the study was approved by local research ethics committees in accordance with the Declaration of Helsinki. Nine healthy subjects (HS), with an age of 53.3 ± 8.9 years, six female and three male, and twelve patients affected by Parkinson's disease (PD), according to UK Parkinson's Disease Society Brain Bank diagnostic criteria (Gibb and Lees, 1988a), with an age of 68.2 ± 6.4 years, three female and nine male, with and Hoehn and Yahr stage range 1.5-2.5, were enrolled in the study. PD patients and healthy subjects had no cognitive impairment and no history of depression, anhedonia, anxiety or other major psychiatric illness, or facial structural alterations which can lead to facial expressivity impairment. All the PD patients were evaluated on their chronic antiparkinsonian therapy, in ON motor state.

Face recording and analysis

All enrolled subjects of both goups, HS and PD, were filmed, with a standard camera, under two different test conditions for 1 min for each task. The two test conditions were:

1. Rest, the subject was asked to look in front of camera without speaking
2. Conversation, the subject was asked to answer examiner questions about the lifestyle and the disease

A blinded examiner evaluated videotapes, and scored facial expression according to MDS-UPDRS III item 3.2 (score from 0 to 4). Face landmarks and AUs occurrence and intensity were automatically extracted from digitally recorded video images, with the Openface software (Amos *et al.*, 2016).

Starting from face landmarks and AUs intensity data, two face features, in different topographic face section in order to cover both the upper and lower face, form two relevant face indices (blinking intensity and lips distance) were calculated and analyzed:

1. **Blinking intensity**
   1.1. max.AU45.r: Maximum value for each patient, during the timeseries of a single task (rest, conversation), of the intensity (r: from 0 to 5) of the AU45 (blink: relaxation of *levator palpebrae and contraction of orbicularis oculi, pars palpebralis*)
2. **Lips distance**
   2.1. max.lips.2D.1: Maximum value for each patient, during the timeseries of a single task (rest, conversation), of the distance on the y axis (in pixels), between face landmarks 51 and 57 (Figure 14).

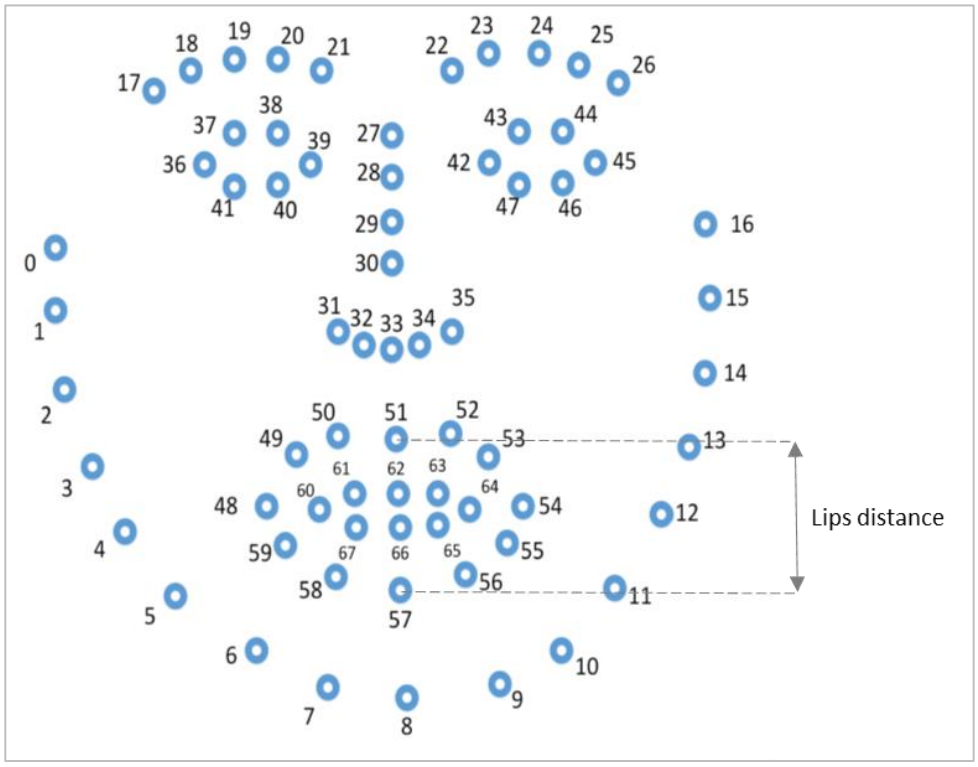


Figure 14 Lips distance of the right eye, based on distance of face landmarks 51 and 57.

Video check and analysis

Before the face features extraction, videos quality has been checked and videos with low quality that would make it difficult to recognize the face landmarks were not selected for the automatic face features extraction analysis.

Statistical analysis

In order to test the normal distribution of the extracted parameters, a Kolmogorov-Smirnov test of normality was performed. For each parameter, for PD vs. HS comparison, according to the normality test, a parametric independent sample t-test or non-parametric Mann-Whitney U test was performed.

For each single task (rest and conversation), in order to describe the difference of each single parameter, of upper and lower face, between Parkinson's disease patients and healthy subjects, an independent samples t-test was performed. Bonferroni correction was applied, considering a correction factor of 4 ([parameters] x [tasks] = 2x2 = 4), therefore the statistically significant value (p) threshold is equal to 0,0125 (0,05/4).

In addition, to describe the diagnostic performance of each parameter a ROC curve analysis considering as target a diagnosis of Parkinson's disease over healthy subject, was performed.

For rest task, that is the most easy and reproducible one, the two predictors were selected for binary logistic regression analysis, in order to evaluate the diagnostic performance for the binary diagnosis of Parkinson's disease versus healthy subject. According to the resulting binary logistic regression equation, these two predictors were combined, in order to create a new predictor identified as Parkinson's disease hypomimia predictor (PHP).

In order to describe the diagnostic performance of this new predictor, a Mann-Whitney U test between PD and HS groups, ROC curve analysis and binary logistic regression were performed. In addition, since the small sample size of the two groups, to evaluate the robustness of the combined parameter in differentiating PD from HS, we applied a

bootstrapping technique, with random sampling with replacement, to the dataset, for the binary logistic regression, 1000 iterations were performed with 95% C.I.

A Spearman correlation coefficient was computed to assess the relationship between the MDS-UPRS III item 3.2 (facial expression) score and Parkinson's disease hypomimia parameter (PHP). Also for Spearman correlation coefficient a bootstrapping technique, with 1000 iterations was performed with 95% C.I.

RESULTS

Video check

After videos quality check, videos with low quality were not selected for the automatic face features extraction analysis. The final number of videos analyzed for each task was the following:

1. Rest: 11 Parkinson's disease patients and 9 healthy subjects
2. Conversations: 12 Parkinson's disease patients and 9 healthy subjects

Normality test

For rest condition Kolmogorov-Smirnov test showed that max.AU45.r, $D(20) = 0{,}138$, $p > 0{,}05$; max.lips.2D.1, $D(20) = 0{,}138$, $p > 0{,}05$ and MDS-UPDRS III item 3.2, $D(20) = 0{,}215$, $p > 0{,}05$, were all significantly normal and that the derived parameter PHP, $D(20) = 0{,}246$, $p < 0{,}05$; was significantly non-normal.

For conversation condition Kolmogorov-Smirnov test showed that max.AU45.r, $D(21) = 0{,}146$, $p > 0{,}05$ and max.lips.2D.1, $D(21) = 0{,}124$, $p > 0{,}05$, were both significantly normal.

MDS-UPDRS III item 3.2 (facial expression)

For the predictor MDS-UPDRS III item 3.2 t-test showed a significant difference in HS and PD, M= 0,3 ± (SD) 0,7 in HS and M = 2,2 ± (SD) 0,9 in PD patients, t(18)= -5,115, ($p < 0,0001$).

Automatic face analysis

Rest

PD vs. HS independent sample t-test and ROC analysis

1. Blinking

For the predictor max.AU45.r t-test confirmed a significant difference in HS and PD, M= 3,072 ± (SD) 0,834 in HS and M = 1,873 ± (SD) 0,71 in PD patients, t(18)= 3,460, ($p = 0,002$) (Table 5). ROC curve analysis of the max.AU45.r considering as target a diagnosis of PD over HS afforded an AUC of 0,151 (95% C.I. 0 – 0,320) with a standard error of 0,086 ($p = 0,009$). (Figure 15)

2. Lips distance

For the predictor max.lips.2D.1 t-test confirmed a significant difference in the in HS and PD, M= -25,589 ± (SD) 8,620 in HS and M = -37,373 ± (SD) 9,611 in PD patients, t(18)= 2,855, ($p = 0,010$) (Table 5). ROC curve analysis of the max.lips.2D.1 considering as target a diagnosis of PD over HS afforded an AUC of 0,151 (95% C.I. 0 – 0,322) with a standard error of 0,087 ($p = 0,009$). (Figure 16)

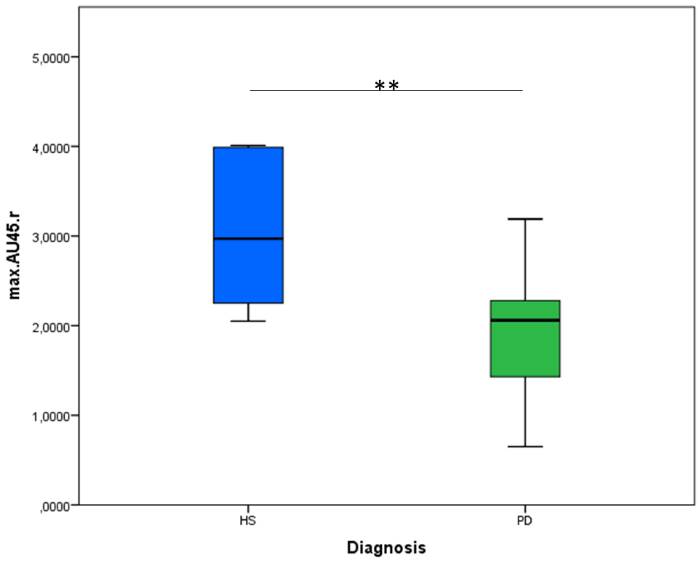


*Figure 15 Boxplot comparison between healthy subjects (HS) and Parkinson's disease patients (PD) of max.AU45.r, during rest task. Legend= ** : t-test p<0,01.*

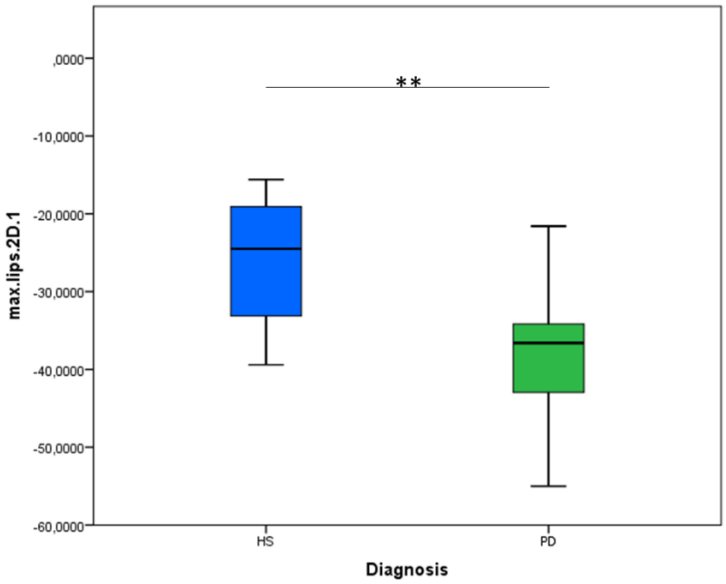


*Figure 16 Boxplot comparison between healthy subjects (HS) and Parkinson's disease patients (PD) of max.lips.2D.1, during rest task. Legend= ** : t-test p<0,01.*

Conversation

PD vs. HS independent sample t-test and ROC analysis

1. Blinking

For the predictor max.AU45.r t-test showed a non significant difference in HS and PD, M= 3,340 ± (SD) 0,973 in HS and M = 2,834 ± (SD) 0,890 in PD patients, t(19)= 1,238, (p = 0,230) (Table 5). ROC curve analysis of the max.AU45.r considering as target a diagnosis of PD over HS afforded an AUC of 0,370 (95% C.I. 0,113 – 0,627) with a standard error of 0,130 (p = 0,319). (Figure 17)

2. Lips distance

For the predictor max.lips.2D.1 t-test confirmed a significant difference in HS and PD, M= -29,922 ± (SD) 11,165 in HS and M = -46,3 ± (SD) 7,228 in PD patients, t(19)= 4,083, (p < 0,001) (Table 5). ROC curve analysis of the max.lips.2D.1 considering as target a diagnosis of PD over HS afforded an AUC of 0,074 (95% C.I. 0 – 0,203) with a standard error of 0,066 (p = 0,001). (Figure 18)

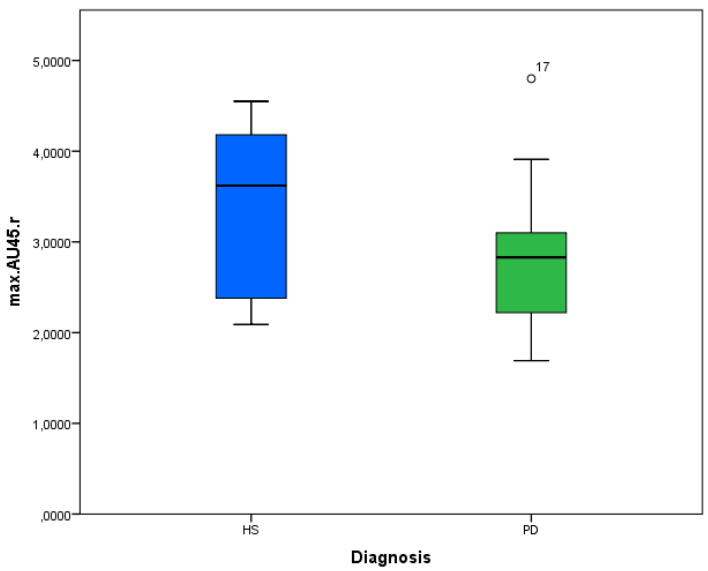


Figure 17 Boxplot comparison between healthy subjects (HS) and Parkinson's disease patients (PD) of max.AU45.r, during conversation task.

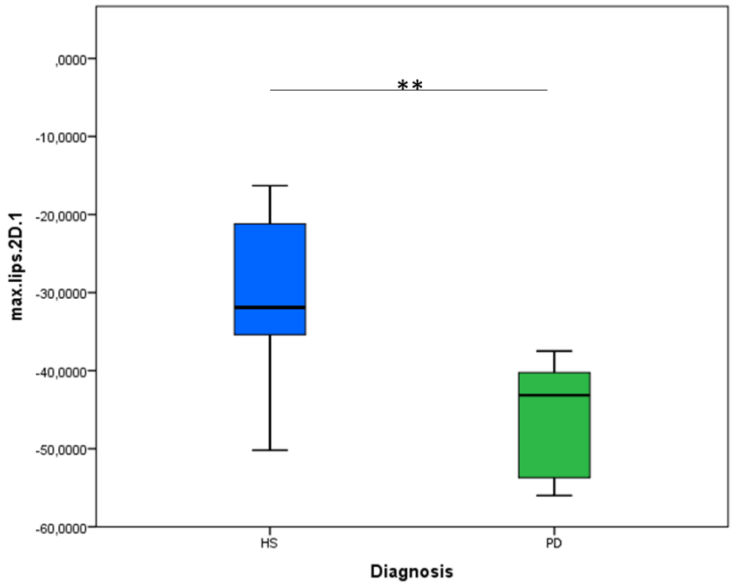


*Figure 18 Boxplot comparison between healthy subjects (HS) and Parkinson's disease patients (PD) of max.lips.2D.1, during conversation task. Legend= ** : t-test p<0,01.*

Table 5 summarizes the PD vs. HS comparison in both rest and conversation tasks, showing a statistically significant difference during both tasks for max.AU45.r parameter and only during rest task for max.lips.2D.1 parameter.

Table 5 Independent samples t-test summary between Parkinson's disease patients and healthy subjects groups.

	rest					conversation				
	HS		PD			HS		PD		
	M	SD	M	SD	p	M	SD	M	SD	p
max.AU45.r	3,072	0,834	1,874	0,716	0,00279	3,340	0,973	2,834	0,891	0,23070
max.lips.2D.1	-25,589	8,621	-37,373	9,611	0,01053	-29,922	11,165	-46,300	7,228	0,00063

Legend= HS: healthy subjects; PD: Parkinson's disease patients; M: mean; SD: standard deviation; p: t-test p value.

Combined parameter

Blink intensity and lips distance Binary logistic regression

The two predictors from rest task, showed both a statistically significant ROC AUC:

- max.AU45.r: ROC AUC of 0,151 (95% C.I. 0 – 0,320) with a standard error of 0,086 (p = 0,009).
- max.lips.2D.1: AUC of 0,151 (95% C.I. 0 – 0,322) with a standard error of 0,087 (p = 0,009).

These two predictors were selected for binary logistic regression analysis. The analysis showed that for every unit decrease of max.AU45.r, the odds (Exp(B)) of a subject having a diagnosis of PD, increased 11,6 times (95% C.I. for Exp(B) 1,3 – 100.0; p = 0,028). In addition, showed that for every unit increase of max.lips.2D.1 absolute value, the odds (Exp(B)) of a subject having a diagnosis of PD, increased 1,25 times (95% C.I. for Exp(B) 1,003 – 1,557; p = 0,047). The bootstrapping analysis made on 955 samples, showed a distortion of -23,124 for max.AU45.r predictor with a standard error of 66,773, and confirmed that max.AU45.r is a statistically significant predictor for the binary diagnosis of PD vs. HS (p = 0,008); for max.lips.2D.1 predictors showed a distortion of -1,634 with a standard error of 3,414, and confirmed that max.lips.2D.1 is a statistically significant predictor for the binary diagnosis of PD vs. HS (p = 0,006).

Combined parameter generation

Following the binary logistic regression equation (Equation 1), the two predictors max.AU45.r and max.lips.2D.1 were combined.

$$\text{logit}(P) = -0{,}806 - 2{,}449 * \text{max.AU45.r} - 0{,}224 * \text{max.lips.2D.1}$$

Equation 1. Binary logistic regression equation, for the binary diagnosis of PD vs. HS, with max.AU45.r and max.lips.2D.1 as predictors.

The resulting predictors was defined as Parkinson's disease hypomimia predictor (PHP).

PD vs. HS Mann–Whitney U test and ROC analysis

For PHP Mann–Whitney U test confirmed a significant difference in the PHP in HS and PD, Mdn = 0,043 in HS and Mdn = 0,955 in PD patients, U = 5,00, z = -3,381, p = 0,000226 , r = -0,756. (Figure 19)

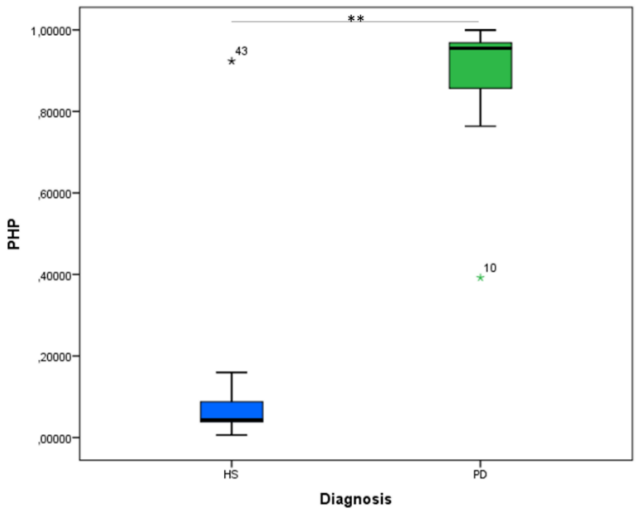


*Figure 19 Boxplot comparison between healthy subjects (HS) and Parkinson's disease patients (PD) of PHP, during rest task. Legend= ** : Mann–Whitney U test p<0,01*

ROC curve analysis of the PHP considering as target a diagnosis of PD over HS afforded an AUC of 0,949 (95% C.I. 0,846 – 1,000) with a standard error of 0,53 (p = 0,001). (Figure 20)

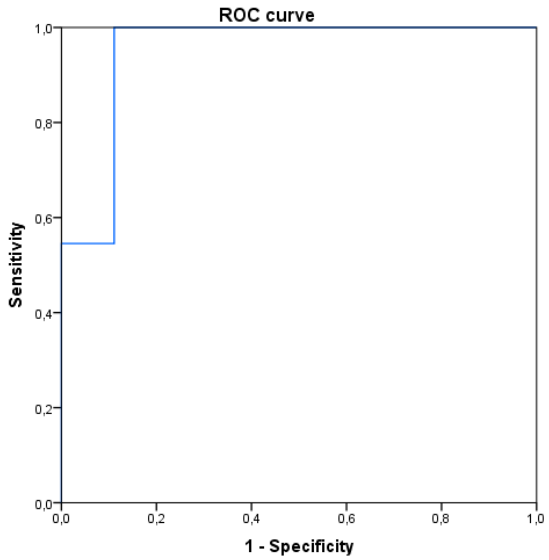


Figure 20 ROC curve of the PHP as a diagnostic test applied for differential diagnosis for PD and HS. AUC is equal to 0,949 (95% C.I. 0,846 – 1,000) with a standard error of 0,53 (p = 0,001).

PHP Binary logistic regression

Binary logistic regression analysis of PHP, for the binary diagnosis of PD vs. HS, showed that for every unit increase of PHP, the odds (Exp(B)) of a subject having a diagnosis of PD, increased 580,93 times (95% C.I. for Exp(B) 6,69 – 50398,24; p = 0,005). The bootstrapping analysis made on 930 samples, showed a distortion of 38,19 with a standard error of 63,90, and confirmed that PHP is a statistically significant predictor for the binary diagnosis of PD vs. HS (p = 0,001).

The binary logistic regression equation (Equation 2) for PHP as predictor for the binary diagnosis of PD vs. HS is the following:

$$\text{logit(P)} = -3{,}215 + 6{,}365 * \text{PHP}$$

Equation 2. Binary logistic regression equation, for the binary diagnosis of PD vs. HS, with PHP as predictor.

PHP cutoff selection

To find the optimal PHP threshold for differentiating HS and PD we selected the cutoff value of PHP which maximized the distance between sensitivity and (1-specificity) (the Youden Index), the inputs to the ROC curve above (Figure 21-22).

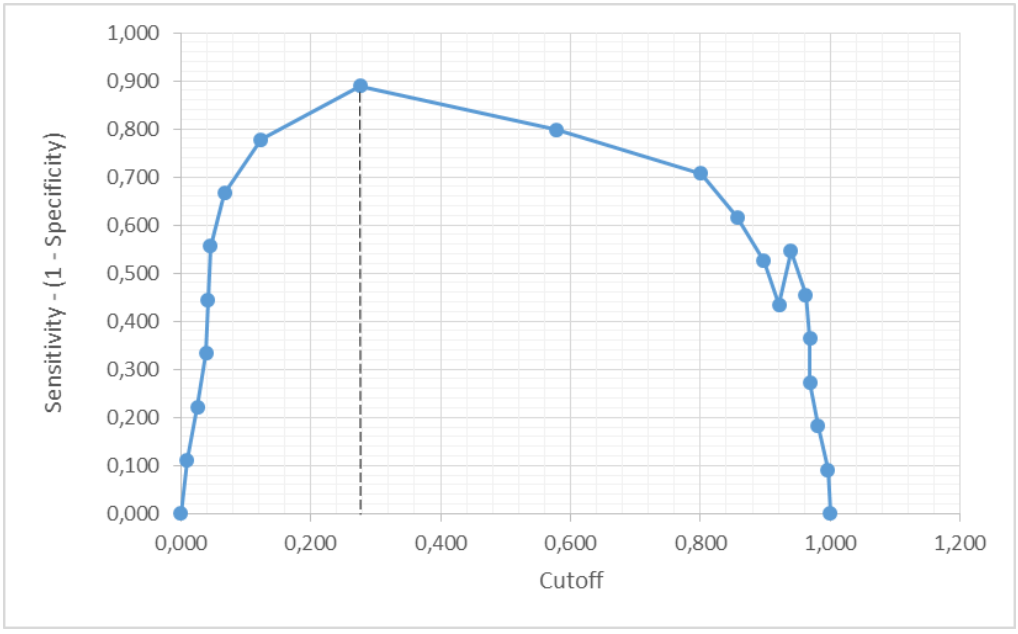


Figure 21 The difference between [sensitivity and (1-specificity)] (Youden Index), for each PHP cutoff value are plotted against the corresponding PHP cutoff value. The vertex of the resulting curve corresponds to the PHP cutoff that maximizes the combination of sensitivity and specificity.

This distance (Youden Index) corresponds to the threshold with the highest combination of sensitivity and specificity values (Figure 22).

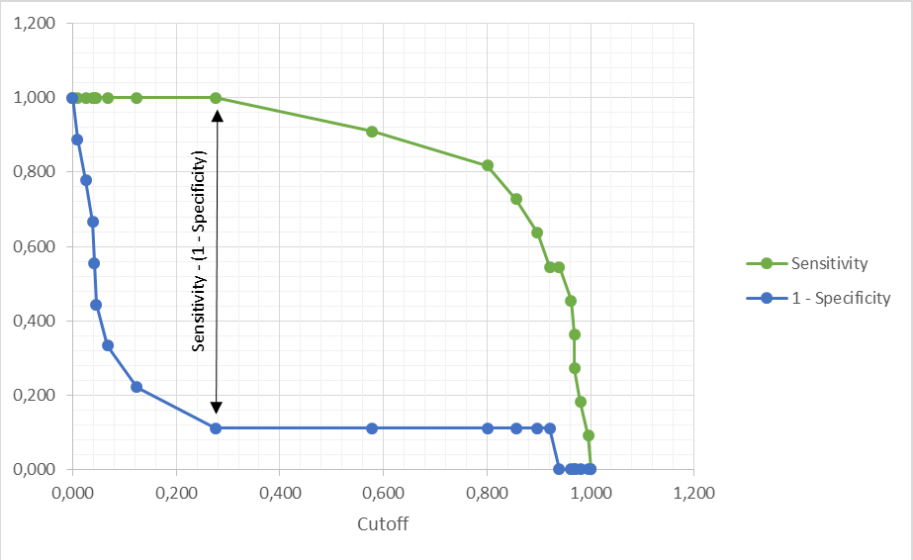


Figure 22 In the graph are reported the values of sensitivity and (1-specificity) for each cutoff of PHP value. The maximum distance between the sensitivity curve (green) and (1-specificity) curve (blue) (Youden Index) define the highest combination of sensitivity and specificity values, and the corresponding best PHP cutoff.

The optimal PHP threshold was 0,276 (Youden Index = 0,889). PHP values <0,276 and ≥0,276 suggested a diagnosis of HS and PD, respectively. (Figure 23)

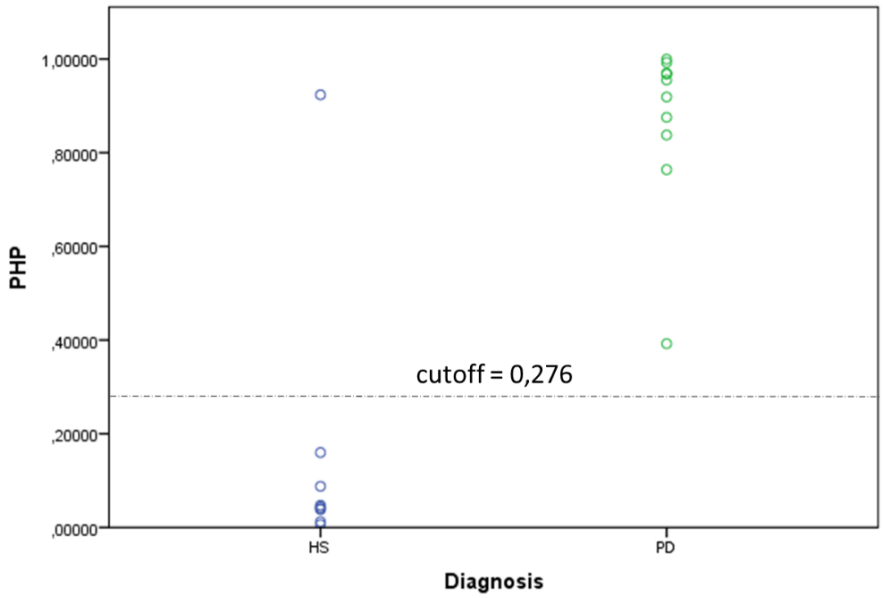


Figure 23 Dispersion graph comparison between healthy subjects (HS) and Parkinson's disease patients (PD) of PHP, during rest task, with a cutoff equal to 0,276 that subdivide the two groups.

PHP diagnostic performance

The 2x2 table for PHP as diagnostic test (Table 6) showed 0 false negative, and only 1 false positive, with a diagnostic performance of this PHP threshold (Table 7) that showed 100% of sensitivity, 89% of specificity, 95% of accuracy, 92% of positive predictive value (PPV) and 100% of negative predictive value (NPV).

Table 6 2x2 table for PHP as diagnostic test

		Real diagnosis		
		HS	**PD**	**tot**
Test diagnosis	**PHP<0,276**	8	0	8
	PHP>=0,276	1	11	12
	tot	9	11	20

Table 7 Diagnostic performance of PHP for the diagnosis of PD vs. HS

Diagnosis	**PD**
Sensitivity	100%
Specificity	89%
Accuracy	95%
PPV	92%
NPV	100%

PHP and MDS-UPRS III item 3.2 correlation

A Spearman correlation coefficient was computed to assess the relationship between the MDS-UPRS III item 3.2 (facial expression) score and Parkinson's disease hypomimia parameter (PHP). The variable MDS-UPRS III item 3.2 and PHP were found to be strongly correlated, $r_s(18) = 0{,}738$, $p = 0{,}0002$. Bootstrapping analysis made on 1000 samples, showed a distortion of -0,023 with a standard error of 0,133 (95% C.I. for r 0,358 – 0,876).

DISCUSSION

In vivo misdiagnosis in Parkinson's disease is one of the biggest unmet needs in this disease, in early, moderate or advanced stage. (Rajput *et al.*, 1991; Jankovic *et al.*, 2000; Adler *et al.*, 2014; Rizzo *et al.*, 2016; Beach and Adler, 2018)

Quantitative and objective assessments improve the inter-rater variability, and make assessments more accurate and reproducible. In the last decades, several studies attempt to improve diagnostic accuracy by means of quantitative evaluations. (Sánchez-Ferro *et al.*, 2016)

Hypomimia is one of the earlier motor symptoms in Parkinson's disease patients, which starts 10 years before clinical diagnosis. (Postuma *et al.*, 2012) Modern computer vision techniques combined with state of the art machine learning algorithms can automatically recognize, define and categorize single face segments, and therefore could be useful to improve the diagnostic accuracy of hypomimia detection in Parkinson's disease. (Hamm *et al.*, 2011; Wu *et al.*, 2014)

Few studies tried to create an automatic system to detect hypomimia in different conditions in Parkinson's disease patients, however, to date no standard technique has been defined and approved for the use in clinical practice.

The main difference in literature studies was the technology used to collect and analyze the data. Some studies collected data with more intrusive technique through markers (Marsili *et al.*, 2014; Bologna *et al.*, 2016) or EMG patches (Wu *et al.*, 2014) placed on face, other studies rely only on videotape markerless analysis. The type of camera used to collect data is variable among studies: standard camera (Katsikitis and Pilowsky, 1988; Wu *et al.*, 2014; Langevin *et al.*, 2019), black and white camera (Bowers *et al.*, 2006), depth camera (Vinokurov N, 2015; Bandini *et al.*, 2016), infrared camera for 3D optoelectronic system (Marsili *et al.*, 2014; Bologna *et al.*, 2016).

Globally all the reported studies showed an impairment of spontaneous, voluntary and emotional face movements for PD patients, compared to healthy controls. (Katsikitis and Pilowsky, 1988; Marsili *et al.*, 2014; Wu *et al.*, 2014; Bandini *et al.*, 2016; Bologna *et al.*, 2016; Bandini *et al.*, 2017)

With the aim to translate results, from lab to clinical practice, preserving the precision and accuracy of the collected data, less intrusive markerless recording techniques and standard camera, which are widely available respect to depth camera or optoelectronic systems, are the preferred setting.

In the present study the data were collected in the less intrusive way, by using only a standard camera, without markers placed on the face of subjects, the face features were extracted through the Openface algorithm (Amos *et al.*, 2016), and custom parameters were calculated from the extracted features, focusing on the most relevant hypomimia features, in line to face features evaluated in clinical practice and MDS-UPDRS scale, i.e. blinking and lips movements.

The maximum intensity of blinking (max.AU45.r) showed to be lower in Parkinson's disease patients compared to controls, during rest task, but this difference was not statistically significant for conversation condition. The non significant difference of maximum intensity of blinking between PD and HS, only during the conversation task could be explained by different factors. Conversation task increase the blink rate in healthy subjects (Bentivoglio *et al.*, 1997), in this study the delta of the increase of blink intensity in PD patients has been more intense respect to HS, decreasing the difference between the two groups. This enhanced blink intensity variation can be a feature of Parkinson's disease or is probably related to antiparkinsonian therapy effect, since patients were evaluated in ON state.

The absolute value of maximum lips distance (max.lips.2D.1) showed to be higher in Parkinson's disease patients compared to controls, during both rest and conversation task. These results are in line with literature data on blinking and lips movements. The combined index, created using only rest parameters, which is the most simple and reproducible task, was defined as Parkinson's disease hypomimia predictor (PHP), showed a high diagnostic accuracy (95%) in Parkinson's disease vs. healthy subjects discrimination, with an ROC AUC of 0,949, a positive predictive value (PPV) of 92% and a negative predictive value (NPV) of 100%. In addition, the variable MDS-UPRS III item 3.2 and PHP were found to be strongly correlated, $r_s(18) = 0{,}738$, showing that PHP could be an useful objective tool to evaluate hypomimia in Parkinson's disease.

The main limitations of this study are the small sample size, the lack of early and late-stage PD patients' subgroups, since the H&Y stage of the selected patients was between 1.5 and 2.5. These subgroups could add in future studies, important information, about the evolution of hypomimia objective markers during the disease course. In addition, in line with other face analysis studies in PD patients (Bandini *et al.*, 2016; Bologna *et al.*, 2016; Bandini *et al.*, 2017), all patients were evaluated on their chronic antiparkinsonian therapy, in ON motor state. An evaluation also of patients in OFF motor state or of drug-naive patients could clarify the role of antiparkinsonian therapy on hypomimia features.

CONCLUSION

In conclusion, the PHP is a new hypomimia measure, which can be an aid tool for the diagnosis of Parkinson's disease. This new tool has a high diagnostic accuracy, positive and negative predictive value. It can be derived from short, cheap, widely available and non-intrusive face recordings at rest, and it is correlated to standard clinical motor scale scoring system.

REFERENCES

Adler CH, Beach TG, Hentz JG, Shill HA, Caviness JN, Driver-Dunckley E, *et al.* Low clinical diagnostic accuracy of early vs advanced Parkinson disease: Clinicopathologic study. Neurology 2014; 83(5): 406-12.

Agostino R, Bologna M, Dinapoli L, Gregori B, Fabbrini G, Accornero N, *et al.* Voluntary, spontaneous, and reflex blinking in Parkinson's disease. Movement disorders : official journal of the Movement Disorder Society 2008; 23(5): 669-75.

Altiparmak UE, Eggenberger E, Coleman A, Condon K. The ratio of square wave jerk rates to blink rates distinguishes progressive supranuclear palsy from Parkinson disease. Journal of neuro-ophthalmology : the official journal of the North American Neuro-Ophthalmology Society 2006; 26(4): 257-9.

Amos B, Ludwiczuk B, Satyanarayanan M. Openface: A general-purpose face recognition library with mobile applications. CMU School of Computer Science 2016; 6: 2.

Baltrušaitis T, Mahmoud M, Robinson P. Cross-dataset learning and person-specific normalisation for automatic action unit detection. 2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG); 2015: IEEE; 2015. p. 1-6.

Baltrusaitis T, Robinson P, Morency L-P. Constrained local neural fields for robust facial landmark detection in the wild. Proceedings of the IEEE international conference on computer vision workshops; 2013; 2013. p. 354-61.

Bandini A, Orlandi S, Escalante HJ, Giovannelli F, Cincotta M, Reyes-Garcia CA, *et al.* Analysis of facial expressions in parkinson's disease through video-based automatic methods. Journal of neuroscience methods 2017; 281: 7-20.

Bandini A, Orlandi S, Giovannelli F, Felici A, Cincotta M, Clemente D, *et al.* Markerless Analysis of Articulatory Movements in Patients With Parkinson's Disease. Journal of voice : official journal of the Voice Foundation 2016; 30(6): 766 e1- e11.

Beach TG, Adler CH. Importance of low diagnostic Accuracy for early Parkinson's disease. Movement disorders : official journal of the Movement Disorder Society 2018; 33(10): 1551-4.

Bentivoglio AR, Bressman SB, Cassetta E, Carretta D, Tonali P, Albanese A. Analysis of blink rate patterns in normal subjects. Movement disorders 1997; 12(6): 1028-34.

Bettadapura V. Face expression recognition and analysis: the state of the art. arXiv preprint arXiv:12036722 2012.

Bologna M, Berardelli I, Paparella G, Marsili L, Ricciardi L, Fabbrini G, *et al.* Altered Kinematics of Facial Emotion Expression and Emotion Recognition Deficits Are Unrelated in Parkinson's Disease. Frontiers in neurology 2016; 7: 230.

Bologna M, Fabbrini G, Marsili L, Defazio G, Thompson PD, Berardelli A. Facial bradykinesia. Journal of neurology, neurosurgery, and psychiatry 2013; 84(6): 681-5.

Bowers D, Miller K, Bosch W, Gokcay D, Pedraza O, Springer U, *et al.* Faces of emotion in Parkinsons disease: micro-expressivity and bradykinesia during voluntary facial expressions. Journal of the International Neuropsychological Society : JINS 2006; 12(6): 765-73.

Cootes TF, Edwards GJ, Taylor CJ. Active appearance models. IEEE transactions on pattern analysis and machine intelligence 2001; 23(6): 681-5.

Cristinacce D, Cootes TF. Feature detection and tracking with constrained local models. Bmvc; 2006: Citeseer; 2006. p. 3.

Deuschl G, Goddemeier C. Spontaneous and reflex activity of facial muscles in dystonia, Parkinson's disease, and in normal subjects. Journal of neurology, neurosurgery, and psychiatry 1998; 64(3): 320-4.

Deuschl G, Krack P, Lauk M, Timmer J. Clinical neurophysiology of tremor. Journal of clinical neurophysiology 1996; 13(2): 110-21.

di Biase L, Brittain JS, Shah SA, Pedrosa DJ, Cagnan H, Mathy A, *et al.* Tremor stability index: a new tool for differential diagnosis in tremor syndromes. Brain : a journal of neurology 2017.

di Biase L, Summa S, Tosi J, Taffoni F, Marano M, Cascio Rizzo A, *et al.* Quantitative analysis of bradykinesia and rigidity in Parkinson's disease. Frontiers in neurology 2018; 9: 121.

Ekman P, Friesen WV, Ellsworth P. Emotion in the Human Face: Guide-lines for Research and an Integration of Findings: Guidelines for Research and an Integration of Findings: Pergamon; 1972.

Ekman P, Friesen WV, Hager JC. The Facial Action Coding System: The Manual on CD-ROM & Investigator's Guide. Salt Lake City, UT: Research Nexus 2002.

Ekman P, Rosenberg E, Hager J. Facial action coding system affect interpretation dictionary (FACSAID). 1998.

Fahn S, Elton RL. Unified Parkinson's disease rating scale. Recent Developments in Parkinson's Disease 1987; 2: 153-63.

Friesen WV, Ekman P. EMFACS-7: Emotional facial action coding system. Unpublished manuscript, University of California at San Francisco 1983; 2(36): 1.

Gelb DJ, Oliver E, Gilman S. Diagnostic criteria for Parkinson disease. Arch Neurol 1999; 56(1): 33-9.

Gibb WR, Lees AJ. The relevance of the Lewy body to the pathogenesis of idiopathic Parkinson's disease. Journal of neurology, neurosurgery, and psychiatry 1988a; 51(6): 745-52.

Gibb WRG, Lees AJ. The relevance of the Lewy body to the pathogenesis of idiopathic Parkinson's disease. Journal of Neurology, Neurosurgery and Psychiatry 1988b; 51(6): 745-52.

Goetz CG, Fahn S, Martinez-Martin P, Poewe W, Sampaio C, Stebbins GT, *et al.* Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Process, format, and clinimetric testing plan. Movement disorders : official journal of the Movement Disorder Society 2007; 22(1): 41-7.

Goetz CG, Tilley BC, Shaftman SR, Stebbins GT, Fahn S, Martinez-Martin P, *et al.* Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results. Movement disorders 2008; 23(15): 2129-70.

Hamm J, Kohler CG, Gur RC, Verma R. Automated Facial Action Coding System for dynamic analysis of facial expressions in neuropsychiatric disorders. Journal of neuroscience methods 2011; 200(2): 237-56.

Hughes AJ, Ben-Shlomo Y, Daniel SE, Lees AJ. What features improve the accuracy of clinical diagnosis in parkinson's disease: A clinicopathologic study. Neurology 1992; 42(6): 1142-6.

Hughes AJ, Daniel SE, Ben-Shlomo Y, Lees AJ. The accuracy of diagnosis of parkinsonian syndromes in a specialist movement disorder service. Brain : a journal of neurology 2002; 125(4): 861-70.

Jankovic J. Parkinson's disease: clinical features and diagnosis. Journal of neurology, neurosurgery, and psychiatry 2008; 79(4): 368-76.

Jankovic J, Rajput AH, McDermott MP, Perl DP. The evolution of diagnosis in early Parkinson disease. Parkinson Study Group. Arch Neurol 2000; 57(3): 369-72.

Karson CN. Spontaneous eye-blink rates and dopaminergic systems. Brain : a journal of neurology 1983; 106(3): 643-53.

Katsikitis M, Pilowsky I. A study of facial expression in Parkinson's disease using a novel microcomputer-based method. Journal of neurology, neurosurgery, and psychiatry 1988; 51(3): 362-6.

Kimber TE, Thompson PD. Increased blink rate in advanced Parkinson's disease: a form of 'off'-period dystonia? Movement disorders : official journal of the Movement Disorder Society 2000; 15(5): 982-5.

Korosec M, Zidar I, Reits D, Evinger C, Vanderwerf F. Eyelid movements during blinking in patients with Parkinson's disease. Movement disorders : official journal of the Movement Disorder Society 2006; 21(8): 1248-51.

Langevin R, Ali MR, Sen T, Snyder C, Myers T, Dorsey ER, *et al.* The PARK Framework for Automated Analysis of Parkinson's Disease Characteristics. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 2019; 3(2): 1-22.

Marsili L, Agostino R, Bologna M, Belvisi D, Palma A, Fabbrini G, *et al.* Bradykinesia of posed smiling and voluntary movement of the lower face in Parkinson's disease. Parkinsonism & related disorders 2014; 20(4): 370-5.

Matthews I, Baker S. Active appearance models revisited. International journal of computer vision 2004; 60(2): 135-64.

Patras I, Pantic M. Particle filtering with factorized likelihoods for tracking facial features. Sixth IEEE International Conference on Automatic Face and Gesture Recognition, 2004 Proceedings; 2004: IEEE; 2004. p. 97-102.

Postuma RB, Berg D, Stern M, Poewe W, Olanow CW, Oertel W, *et al.* MDS clinical diagnostic criteria for Parkinson's disease. Movement Disorders 2015; 30(12): 1591-601.

Postuma RB, Lang AE, Gagnon JF, Pelletier A, Montplaisir JY. How does parkinsonism start? Prodromal parkinsonism motor changes in idiopathic REM sleep behaviour disorder. Brain : a journal of neurology 2012; 135(Pt 6): 1860-70.

Rajput AH, Rozdilsky B, Rajput A. Accuracy of clinical diagnosis in Parkinsonism - A prospective study. Canadian Journal of Neurological Sciences 1991; 18(3): 275-8.

Rizzo G, Copetti M, Arcuti S, Martino D, Fontana A, Logroscino G. Accuracy of clinical diagnosis of Parkinson disease A systematic review and meta-analysis. Neurology 2016: 10.1212/WNL. 0000000000002350.

Sánchez-Ferro Á, Elshehabi M, Godinho C, Salkovic D, Hobert MA, Domingos J, *et al.* New methods for the assessment of Parkinson's disease (2005 to 2015): A systematic review. Movement Disorders 2016; 31(9): 1283-92.

Sariyanidi E, Gunes H, Cavallaro A. Automatic analysis of facial affect: A survey of registration, representation, and recognition. IEEE transactions on pattern analysis and machine intelligence 2014; 37(6): 1113-33.

Suppa A, Bologna M, Conte A, Berardelli A, Fabbrini G. The effect of L-dopa in Parkinson's disease as revealed by neurophysiological studies of motor and sensory functions. Expert review of neurotherapeutics 2017; 17(2): 181-92.

Vinokurov N WD, Arkadir D, Bergman H, Linetsky E. Quantifying Hypomimia in Parkinson Patients Using a Depth Camera. the 5th EAI International Symposium on Pervasive Computing Paradigms for Mental

Health; 2015 Sep 24-25, 2015; Milan, Italy.; 2015.

Ward CD, Gibb WR. Research diagnostic criteria for Parkinson's disease. Adv Neurol 1990; 53: 245-9.

Wood E, Baltrusaitis T, Zhang X, Sugano Y, Robinson P, Bulling A. Rendering of eyes for eye-shape registration and gaze estimation. Proceedings of the IEEE International Conference on Computer Vision; 2015; 2015. p. 3756-64.

Wu P, Gonzalez I, Patsis G, Jiang D, Sahli H, Kerckhofs E, *et al.* Objectifying facial expressivity assessment of Parkinson's patients: preliminary study. Computational and mathematical methods in medicine 2014; 2014: 427826.

Xiong X, De la Torre F. Supervised descent method and its applications to face alignment. Proceedings of the IEEE conference on computer vision and pattern recognition; 2013; 2013. p. 532-9.

Zadeh A, Chong Lim Y, Baltrusaitis T, Morency L-P. Convolutional experts constrained local model for 3d facial landmark detection. Proceedings of the IEEE International Conference on Computer Vision Workshops; 2017; 2017. p. 2519-28.