Int. J. Oral Maxillofac. Surg. 2017; xxx: xxx–xxx
http://dx.doi.org/10.1016/j.ijom.2017.08.005, available online at http://www.sciencedirect.com

International Journal of Oral & Maxillofacial Surgery

Research Paper
Imaging

Reproducibility of the dynamics of facial expressions in unilateral facial palsy

M. A. Alagha[1], X. Ju[2], S. Morley[3], A. Ayoub[4]

[1]Scottish Craniofacial Research Group, Glasgow Dental Hospital & School, University of Glasgow, Glasgow, UK; [2]Medical Device Unit, Department of Clinical Physics and Bioengineering, NHS Greater Glasgow and Clyde, Glasgow, UK; [3]Canniesburn Plastic Surgery Unit at the Royal Infirmary of Glasgow, NHS Greater Glasgow and Clyde, Glasgow, UK; [4]Scottish Craniofacial Research Group, School of Medicine, Dentistry and Nursing, College of Medical, Veterinary and Life Sciences, the University of Glasgow, Glasgow, UK

M.A. Alagha, X. Ju, S. Morley, A. Ayoub: Reproducibility of the dynamics of facial expressions in unilateral facial palsy. Int. J. Oral Maxillofac. Surg. 2017; xxx: xxx–xxx.

Abstract. The aim of this study was to assess the reproducibility of non-verbal facial expressions in unilateral facial paralysis using dynamic four-dimensional (4D) imaging. The Di4D system was used to record five facial expressions of 20 adult patients. The system captured 60 three-dimensional (3D) images per second; each facial expression took 3–4 seconds which was recorded in real time. Thus a set of 180 3D facial images was generated for each expression. The procedure was repeated after 30 min to assess the reproducibility of the expressions. A mathematical facial mesh consisting of thousands of quasi-point 'vertices' was conformed to the face in order to determine the morphological characteristics in a comprehensive manner. The vertices were tracked throughout the sequence of the 180 images. Five key 3D facial frames from each sequence of images were analyzed. Comparisons were made between the first and second capture of each facial expression to assess the reproducibility of facial movements. Corresponding images were aligned using partial Procrustes analysis, and the root mean square distance between them was calculated and analyzed statistically (paired Student *t*-test, $P < 0.05$). Facial expressions of lip purse, cheek puff, and raising of eyebrows were reproducible. Facial expressions of maximum smile and forceful eye closure were not reproducible. The limited coordination of various groups of facial muscles contributed to the lack of reproducibility of these facial expressions. 4D imaging is a useful clinical tool for the assessment of facial expressions.

Key words: 3D; 4D; facial paralysis; imaging.

Accepted for publication

The assessment of the functional and morphological deficits of facial muscle movements and/or recovery is crucial for the diagnosis and management of facial paralysis and for the follow-up of various treatment modalities. Measuring facial nerve function is challenging due to the inherent complexity of the nerve physiology, the multi-regional motor function, and the complex autonomic control of blinking, lacrimation, and salivation. Facial nerve paralysis/weakness results in varying degrees of dysfunction. The ideal method for the assessment of the facial nerve should be sensitive, specific, and reliable[1].

Over the years, various methods have been proposed for the assessment of facial palsy[1–9]; however each has its own limita-

Please cite this article in press as: Alagha MA, et al. Reproducibility of the dynamics of facial expressions in unilateral facial palsy, *Int J Oral Maxillofac Surg* (2017), http://dx.doi.org/10.1016/j.ijom.2017.08.005

tions[10]. The House–Brackmann system[11,12] and the Sunnybrook facial grading system[13] are the two main scales used for the evaluation of facial palsy. Both are subjective and have limited value in quantifying facial paralysis and its related asymmetry. The introduction of two-dimensional (2D) facial imaging based on photographs and/or video images to quantify facial movement represented a positive step towards the standardized evaluation of facial expressions. However, in 1996, Gross et al. demonstrated that 2D recording of facial expressions underestimated muscle movements by about 43%[14].

The assessment of facial expressions using static three-dimensional (3D) imaging is more comprehensive than using 2D photographs. This has, however, fallen out of favour due to the inadequate recording of the path of facial muscle movements. This shortcoming has been addressed by the development of four-dimensional (4D) facial imaging systems, which are now available to capture and reconstruct the dynamics of 3D facial morphology with satisfactory clinical accuracy[15].

One method used to record the 3D dynamics of facial expressions involves the application of reflective markers to the patient's face to track the movements of the muscles[16]. With this system a sequence of 3D images representing the movements of the facial muscles is generated for each facial expression. Tracking the changes in position of the reflective markers throughout the course of a facial expression provides an accurate recording of muscle movements in real time (4D). One of the limitations of this method is the error associated with the direct placement of markers on the patient's face. In addition, this approach is labour-intensive and time-consuming, which limits the practicality of the method for regular clinical use[17].

An innovation of stereophotogrammetry has facilitated the development of markerless recording of facial expressions. This method tracks the optical flow of the recorded 3D facial images. Therefore, the facial landmarks, which are digitized on the first 3D frame of the set of images, are tracked automatically through a video image sequence. This allows the analysis of motion patterns and the evaluation of the magnitude and direction of facial movements[18].

However, the analysis of the video sequence of 3D facial images is complex. Landmark-based analysis of the 4D images provides a limited representation of the comprehensive morphology of the 3D surface during facial expressions. The maximum number of reproducible landmarks that can be utilized for 4D facial analysis is insufficient to express the complete morphological characteristics of facial expressions.

One way to overcome this obstacle is the use of dense correspondence analysis[19]. This method is based on the application of a generic facial mesh, i.e. a mathematical face mask that consists of a fixed number of thousands of quasi landmarks (vertices)[20]. The mesh is conformed, or 'wrapped', on the 3D image of the face for comprehensive representation of the facial morphology. This method has been used for the analysis of static 3D facial images, but has not yet been applied to the dynamic assessment of facial expressions.

This study was performed to investigate the reproducibility of facial expressions in patients with unilateral facial paralysis using advanced morphometric methods. The rationale for this investigation was the identification of expressions that could be used reliably to quantify facial muscle movements for clinical analysis and to measure the outcomes of various treatment modalities.

Materials and methods

Ethical approval was obtained from the South Central Oxford C Research Ethics Committee and the Research and Development National Health Services Greater Glasgow and Clyde Health Board, UK.

The sample size was determined according to Johnston et al.[21]. With the standard assumption of 70% power and significance set at $P < 0.05$, a numerator value of 8 was indicated[22]. To be able to detect a 0.5-mm change between similar expressions, 20 subjects were needed for this study[21].

This cross-sectional study involved 20 adult patients who suffered from unilateral facial paralysis (Table 1).

4D facial capture system

The Di4D system (Dimensional Imaging Ltd, Glasgow, UK) was used to record the dynamics of facial expressions. This method of 3D imaging is based on passive stereophotogrammetry. The system consists of two grey-scale cameras (model avA 1600–65 km/kc, resolution 1600×1200 pixels), a sensor (model KAI-02050; ON Semiconductor, Phoenix, AZ, USA), a colour camera (Kodak sensor model KAI-02050, Basler, Germany), and a lighting system (model DIV-401 Diva-Lite; Kino Flo Corporation, Burbank, CA, USA). The grey-scale cameras capture the video sequences of facial movements at a rate of 60 3D facial frames per second, and the colour camera captures the surface texture. The system was connected to a personal computer to build the sequence of the 3D facial images for each facial expression.

Imaging protocol

Before each capture session, the system was calibrated according to an established protocol to configure the 3D orientation of the cameras to the face[18]. Participants were then shown photographic cue cards illustrating each of the facial expressions to avoid the emotional impact of facial muscle movements. Each participant was seated in front of the 4D imaging system such that the full face could be captured and the head could be freely adjusted, as demonstrated in a previous study[23]. A 5-min training session was given in all cases, in which the patient was asked to perform each of the recorded facial expressions starting from the rest position, reaching the maximum magnitude of muscle movement, and then returning to the resting position. Each facial expression took 3–4 seconds and was recorded in real time at a rate of 60 3D frames per second, thereby generating a sequence of 180 3D images for each facial expression. Thirty minutes after this session the participants were invited back for the second capture session to record the same expressions following the same imaging protocol.

Five facial expressions were recorded: maximum smile (E1), lip purse (E2), cheek puff (E3), maximum raising of eyebrows (E4), and forceful eye closure (E5) (Fig. 1).

Data processing and statistical analysis

A set of 23 landmarks, which have been shown to be reliable by this research team, were digitized manually on the first frame of the 3D image sequence of each facial expression[18,23]. A description of the landmarks is given in Table 2.

These landmarks were used only to conform the generic mesh (Fig. 2) to the first 3D facial frame in each set of sequential 3D facial images for each expression, in order to generate the 'conformed mesh' (Fig. 3). The vertices of the mesh were tracked throughout the sequence of 3D images for each facial expression.

The reproducibility of the landmark digitization was assessed on 10 randomly selected facial expressions. The landmarks were digitized twice and those associated