A STUDY TO ASSESS THE REPRODUCIBILITY OF FACIAL EXPRESSIONS USED IN THE SUNNYBROOK FACIAL GRADING SCALE USING 4D MOTION CAPTURE TECHNOLOGY.

by

ANANT BAKSHI

A thesis submitted to the University of Birmingham for the degree of

MASTER OF SCIENCE BY RESEARCH

School of Dentistry
College of Medical and Dental Sciences
University of Birmingham
December 2022

ACKNOWLEDGMENTS

This project could not have been possible without the support and guidance from my supervisor and mentor Professor Khambay. He has inspired me to be dedicated in all aspects of work that I do and has instilled in me a dedication for research which I am extremely grateful for. He has been there for me during the toughest of times and has always encouraged me to see the light at the end of the tunnel and persevere to achieve the highest outcome possible. I will be forever indebted for the teaching and wisdom he has provided me during my training, and I am privileged to be his student.

I would also like to thank the staff and students at the Birmingham Dental Hospital and School who readily volunteered and supported this project throughout. I am thankful for the time they gave up allowing this study to be carried out.

Finally, I would like to thank my family, my parents, brother, sister-in-law and my niece for their unwavering love and support. They have always been there for me with never ending encouragement throughout my life. Without them this would not have been possible, and I am truly thankful to them.

ABSTRACT

AIM

The aim of this study was to assess the reproducibility of facial expressions used in the Sunnybrook scale using 4D motion capture technology. This was determined by how well the participants could perform six facial expressions at a 6-month interval (T_1 and T_2).

STUDY DESIGN

Prospective longitudinal cohort study.

MATERIALS AND METHODS

Twenty healthy participants (10 male & 10 female) meeting the inclusion criteria, were invited to voluntarily participate in the study. They were imaged performing the six facial expressions stated in the Sunnybrook scale (eyebrow lift, eye closed gentle, maximum smile, snarl, lower teeth show and lip pucker) at T_1 and T_2 (6 months later). All facial expressions were carried out from repose and the volunteers were imaged using a 4D facial motion capture system on both occasions. Upto 8 landmarks were digitally placed on the images, depending on the facial expression, which were then automatically tracked throughout the recording.

The Euclidian distance of each landmark, between T_1 and T_2, for each facial expression were analysed to determine the magnitude of displacement of the landmarks and determine their reproducibility over the two time points. In addition, for each facial expression the mean absolute displacement of each landmark, in the x, y and z direction between T_1 and T_2 was determined. A threshold of 2mm was set as being clinically significant different.

RESULTS

Based on the Euclidian distances, the results of this study showed four of the six facial expressions i.e., snarl, pucker, maximum smile, and eyebrow lift were reproducible. The other two expressions, eye closed gentle and lower teeth show, had differences in Euclidian distance which were statistically significantly different between the two-time intervals.

On further assessment of the x, y and z directions the eyes closed gentle expression was associated with the least reproducibility error, with all distances being statistically significantly less than 2.0mm ($p>0.05$). None of the six facial expressions had mean absolute differences statistically significantly greater than 2.0mm. The greatest errors in reproducibility were seen with lip pucker were in some directions the upper 95% confidence interval limit exceeded 3.0mm.

CONCLUSIONS

Four of the facial expressions showed acceptable reproducibility over a 6-month interval, snarl, eyes closed gentle, maximum smile, and eyebrow lift. Lip pucker and lower teeth show were the least reproducible in the x, y and z directions, with the majority of the landmarks being clinically significantly greater than the 2.0mm threshold. The clinical impact of the study is that any changes in facial movement, because of an intervention, need to be greater than 3.0mm to be a true change and not a result of decreased reproducibly of the facial expression over time.

LIST OF FIGURES

LIST OF TABLES

Contents

CHAPTER 1

LITERATURE REVIEW

1.1 MUSCLES OF FACIAL EXPRESSION

The human face functions through the harmonious interaction between the hard and soft tissues and the dentition to provide form and function. Certain muscle groups are specifically tailored to perform expressions portraying emotions. These muscles of facial expression originate from the hyoid branchiomeric arch and are innervated by the seventh cranial nerve (Young, 1962; Carlson, 1981). Facial expression muscles are formed solely from somatic muscles and in the majority cases, attach directly to the skin from the underlying skeleton. However, in some cases, such as the orbicularis oris muscle, there is no attachment to bone at all. They also differ from other muscles as they have no fascial sheath, thereby providing it with the flexibility to perform fluid motions (Martone, 1962). The attachments of the facial expression muscles form distinct facial features such as the eyes, cheeks and lips and directly control voluntary facial movements used in everyday life such as mastication. In conjunction with these everyday functions, the muscles of facial expressions also play a major role in non-verbal communications such as emotional expressions in the form of happiness or fear (Burrows and Smith, 2003).

Early investigators looking at facial expression muscles included Darwin (1898) and Lightoller (1925), who looked at the anatomy of the muscles during function, such frowning and smiling. They noted that the muscles of facial expression were amongst the strongest of muscles in relation to their size (Mosher, 1951). Their early research through dissections allowed for a better understanding of the functioning of these groups of muscles. Facial expressions involve a precise co-ordination of complex neuromotor and psychomotor processes to move facial muscles (Garcia et al., 2015). For instance, maximum smile involves the simultaneous contraction and relaxation of the zygomaticus major, zygomaticus minor, depressor anguli oris, levator labii superioris and risorius muscles, Table 1.1. All these muscles take their innervations from branches of the facial nerve. Thus, the neuromuscular component is an integral part of ensuring the facial muscle movement performs appropriately. Facial muscle movements can be affected and altered either by conditions altering the facial nerve innervating the muscles such as central nervous system diseases (Renault and Quijano-Roy, 2015), or by affecting the muscles directly such as through cancer resection (Terzis and Konofaos, 2012).

1.2 FACIAL PALSY

The facial nerve innervates the muscles of facial expression which are responsible for complex facial movements such as speech, mastication, and non-verbal communication (Vaughan et al., 2020). If the facial nerve is damaged, either temporarily or permanently, it can cause weakness affecting the facial muscles, termed “facial palsy”. Damage to the facial nerve prevents the required signals being sent from the to the muscles to function normally, resulting in paralysis of part of the face (Facial Palsy UK, 2002), Table 1.2. The symptoms of facial palsy can vary and have both a physical and emotional impact, affecting the quality of life of the sufferer. Symptoms can include (Stew and Williams, 2013):

- Loss of motor function on the affected side
- Difficulty closing eyes
- Dysarthria
- Facial asymmetry

Table 1.1 Anatomy of the muscles of facial expression

Muscle	Origin	Insertion	Action	Innervation
Brow				
Occipito-frontalis, frontal belly	Epicraneal aponeurosis	Underneath skin of forehead	Furrowing brow	Facial nerve
Occipito-frontalis, occipital belly	Occipital bone; mastoid process (temporal bone)	Epicraneal aponeurosis	Unfurrowing brow	Facial nerve
Corrugator supercilii	Frontal bone	Skin underneath eyebrow	Draws eyebrows medially and downward; frowning	Facial nerve
Nose				
Nasalis	Maxilla	Nasal bone	Widens nostrils	Facial nerve
Mouth				
Levator labii superioris	Maxilla	Underneath skin at corners of the mouth; orbicularis oris	Elevates upper lip	Facial nerve
Depressor labii inferioris	Mandible	Underneath skin of lower lip	Draws lower lip downward	Facial nerve
Depressor angulus oris	Mandible	Underneath skin at corners of mouth	Opening mouth and sliding lower jaw left and right	Facial nerve
Zygomaticus major	Zygomatic bone	Underneath skin at corners of mouth (dimple area); orbicularis oris	Draws angle of mouth upward and laterally; smiling	Facial nerve
Orbicularis oris	Tissue surrounding lips	Underneath skin at corners of the mouth	Shaping of lips (as during speech)	Facial nerve
Buccinator	Maxilla, mandible; sphenoid bone (via pterygomandibular raphae)	Orbicularis oris	Lateral movement of cheeks (e.g., sucking on a straw; also used to compress air in mouth while blowing)	Facial nerve
Risorius	Fascia of parotid salivary gland	Underneath skin at corners of the mouth	Draws angle of mouth laterally.	Facial nerve
Mentalis	Mandible	Underneath skin of chin	Elevates and protrudes lower lip and skin of the chin	Facial nerve

Table 1.2 Causes of facial palsy

Trauma	Facial injuries, Birth trauma, Basal skull fractures, Otitic barotrauma
Infection	Mumps, Mastoiditis, Syphilis, Leprosy, Herpes Zoster oticus, Lyme disease, Otitis media
Metabolic	Diabetes mellitus, Acute porphyria, Pre-eclampsia
Neurologic	Multiple Sclerosis, Brainstem infarction, Myasthenia gravis
Autoimmune	Sarcoidosis, Temporal arteritis, Systemic lupus erythematosus, Thrombotic thrombocytopenic purpura
Neoplastic	Tumour of the parotid gland, Lymphoma Tumour of the middle ear, Tumour of the petrosal bone, Bell’s palsy
Iatrogenic	Mastoid surgery, Parotid surgery

There are a variety of causes of unilateral facial palsy, as seen in Table 1, with 70% being diagnosed as Bell's Palsy (Pieterson, 2002).

Various facial paralysis scales have been used to assess facial nerve paralysis, either bilaterally or unilaterally. To be able to accurately image and record reproducible facial expressions, with the aid of a valid facial paralysis scale, is key to determining the site, severity and improvement of facial nerve paralysis following a rehabilitation intervention.

1.3 FACIAL NERVE PARALYSIS GRADING SYSTEMS

Several scales have been developed to assess facial nerve paralysis. For objective measurements the patient's progress is monitored against specific criteria, depending on the scale selected (Burns and Grove, 1997). The grading scales can be categorised into either gross or regional scales. The gross scales consider the overall function of the face and provides a score which represents the extent of facial paralysis. Regional scales compartmentalise different areas of the face into separate entities, which are individually assessed and scored, with the sum of these scores giving a gross score of facial paralysis (House, 1983). For the purpose of this literature review, the main indirect grading scales will be discussed, these include House Brackmann Facial Grading Scale (HBFGS) (House and Brackmann, 1985), House Brackmann Facial Grading Scale (HBFGS 2.0) (Vrabec et al., 2009), Nottingham Scale (Murty et al., 1994), Sunnybrook Facial Grading System (Ross et al., 1996) and computer assisted grading systems.

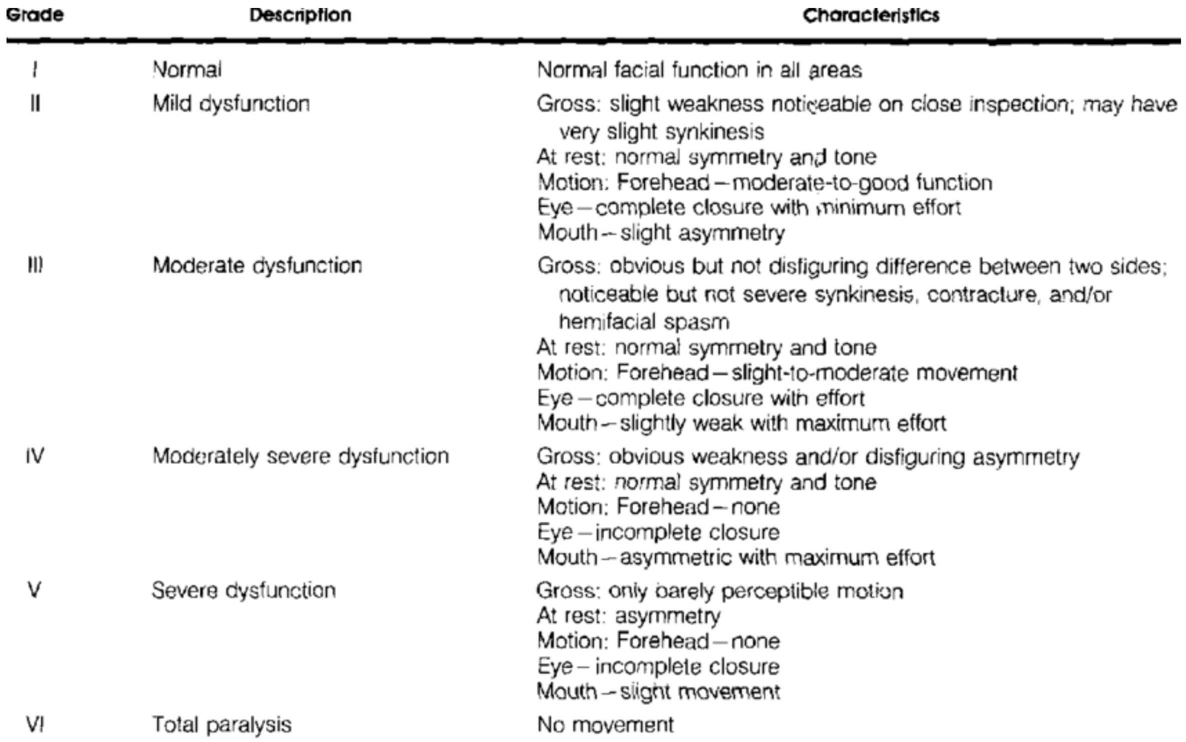
1.3.1 House Brackmann Facial Grading Scale (HBFGS)

This is a gross grading scale that is a combination of two independent facial assessments (House and Brackmann, 1985). The scale is based on a 6-grade score (I-VI) with the severity of paralysis worsening with increasing score, Figure 1.1. This scale was adopted by the Facial Disorders Committee of the American Academy of Otolaryngology in 1984 as the international standard method of reporting (House and Brackmann, 1985).

The HBFGS scale is internationally accepted and is helpful in assessing facial nerve function (Vrabec et a., 2009). It predominantly considers the functionality of the eyes, however not the specific muscle deficit, through the upward movement of midportion of the top of the eyebrow. The motility of the mouth is assessed by a single expression, the outwards movement of angle of mouth. The grading scale has been thought to describe a patient's facial function accurately and quickly, as well as being easy to use by clinicians (Reitzen et al., 2009). Evans et al, (1989) assessed the inter-rater reliability of the HBFGS. The study reviewed 40 patients suffering with a range of unilateral facial palsy by three observers. The outcome was that there was an inter-observer agreement in 93% of assessments, showing it was a simple and robust scale for clinicians to use.

Despite having good inter-operator reliability, there are well known limitations with the use of the HBFGS. A patient is given an overall number which represents the global functioning of the face, despite there being possible varying levels of function in various regions of the face such as the forehead or eye. This singular measurement may not accurately represent the degree of facial paralysis to either the patient or other clinicians (Yen et al., 2003). In addition to this, with advances

Figure 1.1 House Brackmann facial grading scale (House and Brackmann, 1985)



Grade	Description	Characteristics
I	Normal	Normal facial function in all areas
II	Mild dysfunction	Gross: slight weakness noticeable on close inspection; may have very slight synkinesis At rest: normal symmetry and tone Motion: Forehead – moderate-to-good function Eye – complete closure with minimum effort Mouth – slight asymmetry
III	Moderate dysfunction	Gross: obvious but not disfiguring difference between two sides; noticeable but not severe synkinesis, contracture, and/or hemifacial spasm At rest: normal symmetry and tone Motion: Forehead – slight-to-moderate movement Eye – complete closure with effort Mouth – slightly weak with maximum effort
IV	Moderately severe dysfunction	Gross: obvious weakness and/or disfiguring asymmetry At rest: normal symmetry and tone Motion: Forehead – none Eye – incomplete closure Mouth – asymmetric with maximum effort
V	Severe dysfunction	Gross: only barely perceptible motion At rest: asymmetry Motion: Forehead – none Eye – incomplete closure Mouth – slight movement
VI	Total paralysis	No movement

in surgical procedures and technology, the HBFGS does not identify subtle changes which may have occurred through the rehabilitation process, leading to an underestimation of treatments performed. Finally, the scale has the inability to evaluate synkinesis and contracture (Lee et al., 2013).

1.3.2 House Brackmann Facial Grading Scale 2.0

Due to the limitations of the initial HBFGS, it has undergone numerous modifications throughout the years (Yen et al., 2003; Reitzen et al., 2009; Vrabec et al., 2009). The new scale categorises the face into four sections (brow, eye, nasolabial fold and mouth). The four regions are assessed on a 5-point scale depending on the severity of the paralysis, ranging from normal to no movement. The four regional scores are summated and assigned a Grade I to IV. In conjunction to this, a separate synkinesis score was also included. The interobserver reliability for the modified scale was assessed against the original scale (Vrabec et al., 2009). Fourteen physicians scored 21 patients with facial nerve injury using the original HBFGS and the HBGFS 2.0 version, and it was seen that the interobserver reliability was essentially identical. The HBFGS 2.0 version was able to provide additional information without compromising its reliability. The flaw with the modified HBFGS 2.0 was that it was seen to reduce the weight of the secondary defects on the overall grades compared to the original version. A more pressing issue, however with use of the modified scale, was that it encompassed too many issues into each grade, with some issue being incorporated in multiple different grades. This overlap between different grades lead to ambiguity on the final overall scoring (Kang et al., 2002).

1.3.3 Nottingham System

Developed in 1994, the Nottingham system utilised the concept of objective measurements and was performed in three steps. The initial step involved measuring two distances bilaterally (supraorbital point to infraorbital point; lateral canthus to angle of mouth) when at rest and at maximum, whilst carrying out three facial movements: smiling, eyebrow raise and tight eye closure. The differences in measurements between rest and at maximum are summated for either side and the side suffering from palsy is expressed as a percentage of the non-affected side. Following this, the presence or absence of any of the following are noted: synkinesis, hemifacial spasm and contractures. The final step is documenting whether there is presence or absence of: dysgeusia, gustatory tears or dry eyes (Murty et al., 1994). The Nottingham system has the benefit of being quick to implement, provides an objective score that reduces subjective bias and interobserver variability (Zhai et al., 2008). Providing a percentage score of the affected side compared to the non-affected side, allows a quantitative score of the significance of the paralysis. The Nottingham Scale is limited as it is unable to be used for patients with bilateral facial paralysis, as there is no “non-affected” side for comparison. In addition to this, where secondary defects are present, they are noted as descriptive text and their effect is not included in the overall score. This therefore significantly reduces the impact the secondary defects on the extent of paralysis and is a limitation to its use.

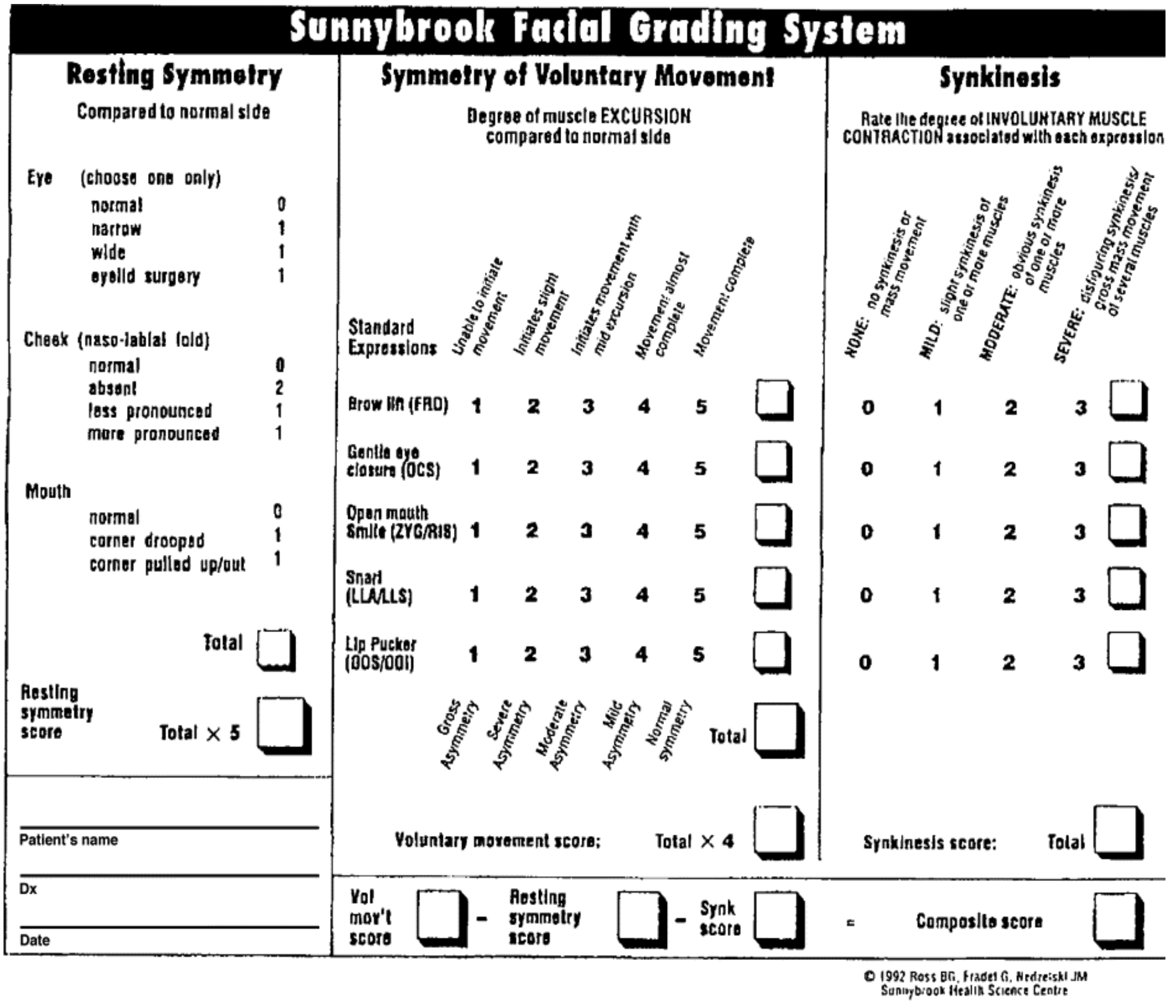
1.3.4 Sunnybrook Facial Grading System

Ross et al. (1996) developed a grading system called the Sunnybrook facial grading system (SFGS). This, unlike the HBFGS, is a regional weighted system which is based upon evaluating resting symmetry, the degree of voluntary excursion and the incorporation of synkinesis. The scale produces a score of 100. The SFGS allowed clinicians to develop a clear, concise, and well-defined understanding of the regional extent of the paralysis (Ross et al., 1996), Figure 1.2.

When carrying out the SFGS, the clinician initially assesses the symmetry of various facial features such as the eye, cheek (nasolabial fold) and mouth at rest and gives a score accordingly. Secondly the facial movements are scored during five standard facial expressions according to the amount of displacement achieved. In the final step, the clinician assesses and scores the severity of synkinesis during the same five standard facial expressions. An overall score is then provided which is the sum of the three parts, Figure 3 (Meng-Yao, 2008). This system enables the clinicians to evaluate facial function at rest and through quantitative scoring. By not simply describing facial paralysis by descriptive methods, it allows for finer changes of facial function to be distinguished.

Validity of the SFGS was assessed by Ross et al. (1996). Internal validity of the SFGS was determined through assessing correlation for interrelationship of the different aspects of the scale. Comparison of the different components to all other components showed that they were all fairly independent of each other. Each score however was equally significant to the overall score. External validity was determined by comparing the results of the SFGS on 19 patients to linear measurements of facial function and the HBFGS. The SFGS was shown to be more sensitive, having the ability to identify change in facial function following rehabilitation intervention (Kayhan et al., 1997).

Figure 1.2 Sunnybrook facial grading system (SFGS) (Ross et al., 1996)



Sunnybrook Facial Grading System

Resting Symmetry

Compared to normal side

Eye (choose one only)	
normal	0
narrow	1
wide	1
eyelid surgery	1
Cheek (naso-labial fold)	
normal	0
absent	2
less pronounced	1
more pronounced	1
Mouth	
normal	0
corner drooped	1
corner pulled up/out	1

Total ☐

Resting symmetry score Total × 5 ☐

Patient's name

Dx

Date

Symmetry of Voluntary Movement

Degree of muscle EXCURSION compared to normal side

Standard Expressions	Unable to initiate movement	Initiates slight movement	Initiates movement with mid excursion	Movement almost complete	Movement complete	
Brow lift (FRO)	1	2	3	4	5	☐
Gentle eye closure (OCS)	1	2	3	4	5	☐
Open mouth Smile (ZYG/RIS)	1	2	3	4	5	☐
Snarl (LLA/LLS)	1	2	3	4	5	☐
Lip Pucker (OOS/OOI)	1	2	3	4	5	☐
	Gross Asymmetry	Severe Asymmetry	Moderate Asymmetry	Mild Asymmetry	Normal symmetry	Total ☐

Voluntary movement score: Total × 4 ☐

Synkinesis

Rate the degree of INVOLUNTARY MUSCLE CONTRACTION associated with each expression

NONE: no synkinesis or mass movement	MILD: slight synkinesis of one or more muscles	MODERATE: obvious synkinesis of one or more muscles	SEVERE: disfiguring synkinesis/ gross mass movement of several muscles	
0	1	2	3	☐
0	1	2	3	☐
0	1	2	3	☐
0	1	2	3	☐
0	1	2	3	☐

Synkinesis score: Total ☐

Vol mov't score ☐ - Resting symmetry score ☐ - Synk score ☐ = Composite score ☐

Previous studies have been carried out assessing the reliability of the SFGS. (Kayhan et al., 1997; Ross and Nedzelski, 1998). Hu et al., (2001) investigated the inter and intra rater reliability of the SFGS by novice observers. Twenty-two patients were assessed who suffered from a wide spectrum of facial dysfunction by eight novice observers. The intrarater reliability coefficients ranged from 0.838 to 0.929. The inter-rater reliability for all eight raters at the first attempt was 0.982 and for the second attempt was 0.970. This is similar to the results found in previous studies, showing the SFGS as a reliable scale even when applied by novice users, proving its ease of use.

Assessments of the intra-rater repeatability and inter-rater agreement of the SFGS and HBFGS have been previously reported (Kanerva et al., 2006). Eight facial palsy patients were video recorded and graded. Repeatability of the SFGS ranged from good to excellent and for HBFGS from fair to good. Agreement between raters for SFGS was nearly perfect based on the intra-class correlation coefficient (ICC). For HBFGS, the generalised kappa indicated only fair agreement. The authors were able to conclude that SFGS was at least as good in repeatability as the HBFGS and showed more reliable results in agreement between raters and thus should be used instead of the HBFGS.

Whilst the HBFGS has historically been the scale used to assess facial nerve paralysis and function, there are many limitations, including the inability to detect subtle change or diagnose synkinesis. The SFGS is a scale which is easy to use and provides both regional and a gross score of facial function, giving clinicians a clear understanding of what is occurring. The SFGS has been shown throughout to be valid, as well as consistently have good inter- and intra-operator reliability, proving it to be a scale clinicians can rely on.

1.3.5 Computer assisted grading

1.3.5.1 Based on two-dimensional images

Technological advances have allowed implementation of computerised assisted grading to assess the severity of facial palsy. Initially, recording systems were trialled, by tracking dots which represented facial landmarks (Johnson et al., 1994). Two-dimensional photographs of patients at rest and with dynamic motion were compared using a digital grid placed over the photographs to quantify the displacements of the landmarks. The system allowed clinicians to assess a variety of landmarks, however it was seen to be extremely time consuming, but more importantly it was unable to assess images in motion (Brenner and Neely, 2004).

1.3.5.2 Based on two-dimensional video images

A further development was FACE (Facial Analysis Computerised Evaluation) which used the method of subtracted-image light reflectance to assess facial movement (Neely and Cheung, 1992). Seventeen patients with facial paralysis and five control patients were recorded and changes in light that occurred between captured images of various dynamic facial expressions were recorded based on grey-scale pixel values. The pixels during rest were subtracted from subsequent motion captures and the difference in grey scale further amplified; the greater the subtraction the greater the severity of paralysis. This method was seen to be an improvement as it allowed assessment of different regions of the face, providing clinicians with region-specific paralysis values. This system was further developed by Scriba et al. (1999) which subsequently showed strong correlation with the HBFGS scores and provided data of facial motion over a continuous period of time. The limitations of the software however were the costs required to implement the system and its inability to undertake vector analysis which is paramount in determining the severity of the paralysis.

1.3.5.3 Based on landmark tracking

A landmark based system was subsequently developed in which initially a total of 24 reflective marks were placed on the face at specific anatomical locations and recorded and expressed as a montage (Isono et al., 1996). A ratio system was used to demonstrate the velocity and amplitude of facial motions between a normal and abnormal site over a period of time. It was shown to be effective in measuring facial displacement to determine the extent of paralysis, but like previous software's, required a significant amount of time and investment, making it non-viable within a clinical setting.

The tracking of movements of facial markers is still the method of choice to objectively assess the extent of paralysis using a computer-based system. A landmark based system, the Peak Motus Motion Measurement System, was used to assess the dynamic asymmetry of facial expressions in a cohort of patients diagnosed with acoustic neuroma and a control group (Linstrom et al., 2002). The study looked at the components of the Euclidian distance of landmarks such as the commissures of the mouth during smile and showed that landmarks were tracked for 100% of subjects. The system was ideal for measuring speed, velocity, direction, and magnitude of landmarks. It also provided information on the spatial and temporal positions of facial landmarks during dynamic motion, something the previous systems were unable to undertake (Horta et al., 2014).

1.4 MEASURING FACIAL MOVEMENTS

The human face is a complex and dynamic three-dimensional (3D) collection of hard and soft tissues. The ability to accurately map the face at rest and during facial expressions can be a challenging task. The method of facial imaging has varied over time, and new methods of imaging are available as technology advances. A variety of methods have been adopted to objectively measure facial movements based on either direct or indirect measurements.

1.4.1 Direct measurements

A simple instrument developed by Frey et al. (1994) was described to measure distances on the face, Figure 1.3. The principle behind it was simple and directly measured the distance between two points on the face using callipers. The “Faciometer" was designed to enable it to reach all points of interest on the face. A study involving twenty “normal” subjects in conjunction with ten patients who suffered from unilateral paralysis were analysed. The study reported that there was a mean measurement error of 0.67 ± 0.66mm at two different time intervals on the same face. However, whilst the intra-operator reliability was acceptable, there was poor inter-operator reliability using the device. The Faciometer was designed to for static measurements and not measure the dynamics of facial animation. In addition, the callipers posed a potential risk to patients, as unexpected movement could lead to soft tissue trauma.

An alternative method of measurement, using handheld rulers to record the position and movement of points on the face was also reported (Manktelow et al., 2008). In this study, two operators independently measured 21 unilateral facial paralysis patients at two different time

Figure 1.3 Faciometer to assess landmark displacement (Frey et al., 1994)

points. The results of the study showed that this simple method had good inter-operator and intra-operator reliability with an average intraclass correlation coefficients exceeding 0.89. The mean difference shown between the measurements using the handheld ruler was 1.7mm. The study concluded that this was an accurate and reliable method of assessing facial nerve paralysis. However, there were some limitations, a major one being the method analysed the change in distance Euclidian distance between two points, which is the shortest straight-line distance in two dimensions (2D). Despite a similar displacement of the landmark in terms of Euclidian distance, the landmark displacements may have been different in x, y and z directions.

Both the use of callipers and rulers was seen as a rudimentary method of assessing facial nerve injury. It gave clinicians the ability to measure the distance between certain landmarks after movement but not during the movement. In addition, the measurements, were only in two dimensions and the patient had to b physically present during measurement recording.

1.4.2 Indirect Measurements

1.4.2.1 Two-dimensional (2D) measurements

Clinicians have previously evaluated and assessed soft tissues with conventional photographs, either from frontal view or profile photographs. Photographs have the benefit of allowing clinicians to assess the patient’s face a number of times without the need for the patient to be physically present. Unfortunately, 2D photographs are insufficient as, “facial depth and shape are not accounted for” (Da Silveira, 2003). Conventional facial photographs only allow clinicians to assess static images at certain time points, such as at rest or maximum smile, but are unable to capture dynamic motion, similar to direct 2D measurements. Photographs may also undergo distortions in magnification and perspective error. In conjunction with this, there is usually poor reproducibility for head positions and camera positions leading to poor identification of facial asymmetries (Sarver, 2001).

The measurement of anatomical landmark changes using photographs is difficult and is only possible in certain regions of the face (Johnson et al., 1994). In the study. landmarks were placed on seven healthy participants and three patients with facial asymmetries. Patients were then photographed, with a ruler to allow determination of scale. The participants carried out a variety of facial expressions including maximum smile, maximum frown, and maximum lip pucker. The study highlighted that landmark detection was only possible where there was a clear change from rest to a maximum facial expression. Unfortunately, the study did not quantify the minimum soft tissue change required to be clinically detectable; however, the authors suggested that any subtle changes would not be detectable.

The reliability of facial photography was investigated taking direct anthropometric measurements of the face (Farkas et al., 1980). The study assessed 104 facial measurements from 36 participants through measurements taken from both frontal and profile photographs. Only 62 of the 104 (60%) direct anthropometric measurements were able to be recorded from the photographs, with only 26 of the 62 (42%) being reliable and accurate. Errors due to incorrect head positioning in both the vertical and horizontal planes were reported as a possible cause of the inaccurate measurements.

These studies have shown the limitations in assessing facial movements in 2D photographs. Clinicians are unable to assess dynamic movements, only the terminal points such as rest or maximum smile could be assessed. In conjunction, the way the photograph is taken and the patient's position play a vital role in the outcome measurements. There was also poor reliability of facial measurements between conventional photographs and the clinically derived measurements. As technology improved, there began a shift, from assessing patients in 2D to assessing them in 3D using digital cameras. This allowed clinicians to overcome the static nature of conventional photographs.

Comparisons have been made between three dimensional and two-dimensional analysis of facial motion (Gross et al., 1996). Four subjects were recruited, two were the control and two patients had repaired unilateral cleft lip and palate. They each carried out a set of five maximum facial expressions, such as smile or eye closure. Patients were then recorded using 3D imaging and the 3D coordinate data extracted, as well as using a conventional camera to obtain x and y coordinates for 2D analysis. The landmark displacement measurements were larger using 3D imaging analysis compared to the two-dimensional data. For example, during maximum smile, the two-dimensional images underestimated the same three-dimensional measurements by as much as 43%. The study had its limitations such as a small sample size but it was able to highlight that two-dimensional analysis failed to adequately assess facial motion.

1.4.2.2 Three-dimensional (3D) image capture

Three-dimensional imaging of the human facial anatomy has become more prevalent over the past few years with advancing technology (Blais, 2004). New imaging techniques can record facial movements in 3D over time in a non-invasive way, overcoming the limitations seen in 2D imaging analysis by providing information of all facial components (Souccar, 2012). In addition, 3D imaging is less susceptible to errors from patient positioning and camera positioning which was highlighted in previous 2D studies (Ferrario, 1996). No single imaging modality can record the facial structure in a single capture.; to overcome this multimodal image fusion is required. This involves capturing the skeletal hard tissue with conebeam CT, the teeth with intra-oral scans and the 3D colour photorealistic soft tissue with surface scanning. Following fusion with the appropriate software it is then possible to produce a digital copy of the patient's anatomy and carry out “virtual surgical planning,” allowing a range of diagnostic procedures and analysis to be performed.

Hajeer et al. (2004) classified soft tissue imaging in the following ways:

- Photogrammetry
- Laser scanning
- Stereophotogrammetry

1.4.2.2.1 Photogrammetry

Photogrammetry is the technique of measuring objects from standardised photographs. It was initially a method that was laborious and expensive, therefore rarely used (Berkowitz et al., 1977). However, over time it became a method that was cheap, easy to use and was able to be acquired in a short period of time (Galantucci et al., 2013). This method overcame the issues of normal photograph analysis as it was not susceptible to distortion and magnification issues. 3D images were obtained through the acquisition and comparison of multiple photographic images (D'Apuzzo, 2002).

1.4.2.2.2 Laser scanning

Laser scanning was seen for many years as the ideal method of recording and analysing 3D facial shape. The principle behind laser technology is triangulation, which refers to the process of determining a point in 3D space given its projections onto two, or more, images. The laser scanner records soft tissue form with the use of a light source and a receiver. The laser beam reflects scattered light off the face, which is then recorded by the receiver and processed to produce a 3D image along with its spatial location (Blais, 2004). Laser scanning has the ability of producing high resolution facial imaging, with studies showing its accuracy and clinical validity in capturing 3D images (Kau, 2005a; Kau, 2005b).

There are some limitations for the use of laser scanners as a data capture method of soft tissues. It is a time-dependent process and any changes in the patient's head position or movement in facial expressions distort the scan. This can lead to artifacts and reduces the accuracy of the images recorded (Kovacs et al., 2006). In conjunction with this, when using the early laser scanners the final images lacked the photorealistic appearance, which resulted in difficulties identifying certain landmarks which are dependent on these features (Khambay et al., 2008; Hajeer, 2003).

Laser scanners, although no longer the preferred method of 3D imaging the face, have continued to improve, and are still widely used. Scanning times have greatly reduced, from previously taking 1-15 minutes to now taking less than 3 seconds. This greatly reduces the number of artefacts and distortions present on the final images produced, as well as being less time consuming for the patient (Komazaki et al., 2011; Djordjevic et al., 2014). There have been a variety of studies which have used laser scanning to capture the face, including facial morphology studies among different populations (Stephen, 2009) and assessments of surgical outcomes (Guest, 2001; Kau, 2007).

1.4.2.2.3 Stereophotogrammetry

The human face is a dynamic three-dimensional structure. At present the face can be captured in two-dimensions using conventional photographs but this fails to record the 3D nature of the face. Three-dimensional imaging, whilst an improvement, captures the face as a static 3D image. Viewing images captured with three dimensional stereophotogrammetry have been shown to yield different visual clinical outcomes to 2D imaging (Zhu et al., 2016). Forty pre-orthognathic patients were captured using conventional photographs and using a 3D facial capture system. Raters were then asked to assess the severity of the facial deformity. The results showed that rating symmetry was more reliable using 3D imaging, with the intra-rater reliability being highest when assessing facial symmetry. The study also showed that all raters reported that 3D viewing provided more clinical information than conventional 2D photographs. The study highlights the point that the use of static 3D capture systems, to assess facial form, is more reliable and yields different clinically relevant information than conventional 2D photographs. Given the dynamic nature of the face, the ability to record a dynamic facial sequence will allow recording of the 3D temporal changes of the face over a short time interval.

Static and dynamic three-dimensional stereophotogrammetry is based on the principle of triangulation. This principle is the basis of the 3D visual perception of the human vision system. As a result of the horizontal separation of the right and left eye, each eye sees the same object slightly differently. The visual cortex in the brain then processes this disparity between the two images and triangulates depth with a high degree of accuracy. The camera system uses the same principle and pairs of cameras to create a horizontal disparity based on the left and right cameras pods diverging to a common volume and vertical disparity by angulating the cameras within each pod. Visual data is captured for each side of the face and using the appropriate software the two stereo-images are merged based on the calibration data to produce a complete 3D image of the face from ear to ear (Honrado, 2004; Souccar, 2012). The 3D static system has a capture time of around 1 millisecond, whilst the dynamic system has a rate of capture of 60 frames per second, in this case each frame is a 3D image. The dynamic system can capture at a slower and faster frame rate; a faster frame rate generates more data which may not be analysable whilst a slower capture rate may result in missing data. At present a rate of 60 frames per second is considered to be the optimum rate.

Three-dimensional imaging has been shown to be a valid and reliable method of 3D facial image capture (Ayoub et al., 2003). The validation study was based on capturing and landmarking plaster casts of 21 cleft lip patients who had undergone primary lip repair, and comparing the landmark positions obtained from a co-ordinate measuring machine The operator error was found to be within 0.2mm of the true CMM measurements. The overall accuracy of the 3D system used (C3D) was found to be accurate to within 0.4mm. The study concluded that the C3D system had a high degree of accuracy, and that the 3D stereophotogrammetry system was a reliable method of recording facial shape. Other 3D imaging systems based on the same technology have also been shown to be accurate and reliable (Khambay et al., 2008). The images produced were of high quality with a landmark error within 0.7mm, which was clinically acceptable. The results of the study were in agreement with previous studies (Ayoub et al., 1998; Ayoub et al., 2003).

The development of 3D motion capture (4D) technology has now allowed the evaluation of dynamic facial movements. Two main types of facial 4D image capture have been described; marker or makerless based technology. Marker based technology relies on the use of retroreflective markers which are physically secured to the individuals face prior to image capture. These retroreflective marker-based systems use a series of cameras to capture marker movement during the facial expression (Weeden et al., 2001). Even though the use of marker-based technology is well documented there are a few disadvantages, these include marker size and the type of data recorded. As the markers are physically secured to the individuals face, with adhesive, the size and location of the marker may inhibit or alter the facial movement which the individual is making. Secondly, the data that is captured are the changes in position of that specific marker, this means visually, a “facial surface” is not produced but a series of moving dots. This makes visualisation of complete facial changes impossible as there are a limited number of markers that can be placed on the face. These short comings are addressed with markerless capture systems, which do not rely on pre-placed facial markers but landmarks are placed on the facial surface after image sequence capture. The disadvantage of these systems includes the placement of landmarks on points that need manual palpation to determine their extract location, for example gonion and orbitale.

This form of assessment of facial expressions has been shown to be reliable and overcomes many of the pitfalls of previous methods (Shujaat et al., 2014; Trotman, 2013; Al-Rudainy et al., 2018). The process utilises skin features such as facial textures and pores to determine the depth. The capturing process for the Di4D system (Dimensional Imaging Ltd., Glasgow) occurs through the use of two pods which contain three video cameras which produce high resolution images at 60 frames per second. In order to enhance the facial features, a lighting system is used, and the entire capturing process occurs on a PC (Al-Anezi et al., 2013; Ju et al., 2016). Di4D software is then used to analyse the video sequence date by accurately pinpointing and tracking pixels through each frame via optical flow.

A potential disadvantage of 3D motion capture is the need to manually landmark each individual frame making up the entire 3D sequence. It would be time consuming to manually landmark each 3D image with the same landmarks. In addition, there would be a baseline level of error with landmark placement on each frame. This could be accumulative and by the end of the sequence could be significant. With advances in software, automatic landmark tracking has overcome this problem using “optical flow”. Optical flow is a computer vision algorithm which checks the pixel characteristic at time T to see if it is the same as the pixel characteristics in the next frame but at a different location, in this way the landmark placed on the first frame of the 3D sequence can be tracked in each frame throughout the remaining sequence. This is an automated process which allows fast tracking of anatomical landmarks (Borland et al., 2006; Ayoub et al., 2013). Al-Anezy et al. (2013) assessed the validity of automatic tracking process of facial landmarks on 3D motion captured images (4D) (Al-Anezy et al., 2013). The study was based on thirty-two subjects carry out three facial expressions from rest, with 23 landmarks pre-marked manually on the patients face and digitised on the day of the study. The landmarks were digitised one month after the study to assess the accuracy. The discrepancy between the manual landmarking and the digitised landmarks was 0.17mm and the mean distance between the two landmarking methods was within 0.55mm. This study showed that the automatic facial tracking landmark process showed good accuracy allowing it to be used in the clinical assessment of dynamic motion during facial expressions.

The use of dynamic motion capture has been used on different cohorts of patients, normal patients (Khambay et al., 2019), cleft lip and palate (Hallac et al., 2017) and oncology patients (Shujaat et al., 2014). Throughout these studies, it has been shown that the 3D motion capture (4D) technology has overcome the time-consuming issues its predecessors faced and has been shown to be both reliable and accurate in assessing dynamic facial expressions.

1.5 REPRODUCIBILITY OF FACIAL EXPRESSIONS

Reproducibility is defined as "the variation in measurements made on a subject under changing conditions" (National Institute for Standards and Technology, 2007). Providing the same measurement method is used, by the same observer, then the changing condition is time. This is important in clinical intervention studies where time may influence the ability of a patient to undertake the facial expression. Generally, with facial expressions, for example a smile, the expression will be reproducible to a great extent, a smile is a smile. However, the extent or magnitude the corners of the mouth move, as well as the direction, would be a more precise definition of reproducibility. For a smile to have a high degree of reproducibility the corners of the mouth would need to move the same amount, and in the same direction, on two separate occasions. In addition, they would need to move at the same speed and follow the same trajectory. This would probably be potentially impossible for a human being to achieve, so there must be a baseline level of reproducibility error. When assessing facial nerve paralysis, pre and post intervention, it is important for this baseline error to be less than the clinical change being measured because of the intervention. This ensures any change in the expression between the two time points, for example pre- and post-surgery, are a result of the intervention and not a measure of the inability of the patient to perform the expression reproducibly.

Several studies have investigated the reproducibility of facial expressions based on healthy participants using both static and dynamic imaging (Johnston et al., 2003; Sawyer et al., 2009; Ju et al., 2016; Özsoy et al., 2019; Popat et al., 2010; Qui et al., 2022). Based on static 3D images, the at rest facial image and the 3D maximum facial expression image of the same patient have been used to assess reproducibility expression. This is not strictly dynamic facial motion capture as the 3D images in between the two time points have not been captured. The advantage of 3D motion capture (4D) is that all the images from rest to maximum are captured and can be potentially analysed, Table 1.3.

Table 1.3 Summary of studies that have investigated the reproducibility of facial expressions based on healthy participants using both static and dynamic imaging.

Study	Capture system	Capture type	Number of subjects & type	Facial expressions	Time intervals	Outcome measure	Most reproducible expressions	Least reproducible expressions
Johnston *et al.*, 2003	Stereophotogrammetric camera system (C3D)	Static Stereophotography	30 adults (15 men, 15 women)	• Rest position • Natural smile • Maximum smile • Lip purse • Maximum cheek puff	$T_1 = 0$ $T_2 = 2$ weeks	Difference in Euclidian distance of landmarks at each time interval	Rest position	Cheek puff
Sawyer *et al.,* 2009	VECTRA-3D 2 dual module system (Canfield Scientific, Inc, Fairfield, NJ)	Static Stereophotography	39 adult volunteers (21 male, 18 female)	• Repose (resting position) • Maximum raised eyebrows • Close Eyes closed gently • Close eyes tightly • Maximum smile lips closed	$T_1 = 0$ $T_2 = 15$ mins $T_3 = 4$ weeks	Difference in Euclidian distance of landmarks at each time interval	Repose (rest position)	Smile-with-lips open Blow-out-the-cheeks lease

				• Maximum smile lips open • lips pursed • Show lower teeth • Blow out the cheeks				
Ju *et al.,* 2016	3D motion capture system (Dimensional Imaging Ltd, Glasgow, U.K.)	Dynamic Stereophotography	32 adults (16 females and 16 males)	• Maximum smile • Cheek puff • Lip purse	$T_1 = 0$ $T_2 = 15$ mins	Difference in Euclidian distance, speed and similarity of landmarks at each time interval	Maximal smile	Cheek puff Lip purse
Özsoy *et al.,* 2019	Light scanner (Artec$^{\text{TM}}$ Eva, Artec Group, Luxembourg)	Static 3D laser scan	30 adults (13 male and 17 female)	• Rest. • Maximum lifting of eyebrows • Maximum closure of the eyelids • Maximum showing of the teeth • Whistling	$T_1 = 0$ $T_2 = 3$ months	Root mean square value which provided a magnitude of facial movement Full face analysis based on captured facial mesh	All were found to be reproducible	None detected

				• Maximum compression of teeth				
Popat *et al.,* 2010	3dMDFace Dynamic System (3Q Technologies, Atlanta, GA)	Dynamic Stereophotography	25 adults (13 male, 11 female)	• Maximum smile (lips closed). • Maximum smile (lips open).	$T_1 = 0$ $T_2 = 1$ month	Principle Component Analysis (PCA)	Maximum smile (with lips closed)	Maximum smile (lips open)
Qui *et al.,* 2022	3dMD-Face Dynamic System	Static Stereophotography	27 adults, 12 males and 15 females,	• Smile (lips closed). • Smile (lips open). • Lip purse. • Cheek puff.	$T_1 = 0$ $T_2 = 15$ mins $T_3 = 1$ week	Root mean square between different pixels of key frames Full face analysis based on captured facial mesh	Smile (lips closed) Lip pure	Cheek puff

Four of the studies used a static 3D camera system to assess the reproducibility both at rest and at maximum expression, at similar time intervals (Johnston et al., 2003; Sawyer et al., 2009; Özsoy et al., 2019; Qui et al., 2022). Two studies used the difference in Euclidian distance of key anatomical landmarks at each time interval as a measure of facial expression reproducibility (Johnston et al., 2003; Sawyer et al., 2009). Landmark placement was shown to be time consuming and related to operator experience (Johnston et al., 2003). Individuals were imaged performing five facial expressions whilst maintaining a natural head position (T_1). The individuals were then asked to repeat the exact same expression on 2 further occasions: T_2 being 15 minutes after the initial imaging and then T_3 being 2 weeks later. This meant that the individual were required to be re-landmarked at T_3. This was similar to the other study where individuals were imaged at similar time intervals; initial capture (T_1), 15 minutes later (T_2) and one month later (T_3) (Sawyer et al., 2009). The last two studies based on 3D static images used the whole face for analysis and used the Root Mean Squared (RMS) distance between the at rest and maximum expression facial meshes (Özsoy et al., 2019; Qui et al., 2022). Two studies used a dynamic facial capture system to assess facial expression reproducibility (Ju et al., 2016; Popat et al., 2010). The outcome measures were based on differences in Euclidian distance, speed, and similarity (Ju et al., 2016) and Principal Component Analysis (PCA) (Popat et al., 2010).

The studies investigated a variety of different facial expressions and time interval. There was some common ground between all the studies, with all looking at maximum smile. Maximum smile was found to be the most reproducible facial expressions. The results showed that cheek puff was the least statistically significant reproducible facial expression. It was however noted that in some studies, the interval between the two captures times (T_1 and T_2) was only 15 minutes (Sawyer et al., 2009; Ju et al., 2016; Qui et al., 2022). Overall reproducibility of facial expressions was much higher where the time interval between T_1 and T_3 was a few weeks apart; suggesting that reproducibility reduces with time. The conclusion of the studies was that some expressions were more reproducible than others and that reproducibility reduced over time.

1.6 SUMMARY

This literature review has shown that there are different methods of assessing facial movement using a variety of scales. With regards to facial nerve paralysis, quantifying the site and severity, prior to and post medical and / or surgical intervention is necessary to allow assessment of the success of treatment or disease progression. A valid and reliable scale allows planning of treatment need as well as the impact of any treatment carried out.

There are a variety of grading systems which have been developed to systematically assess facial nerve paralysis. One such system is the "Sunnybrook Facial Grading System" (SFGS) which evaluates movements of different regions of the face to provide a weighted score to determine the severity of the facial palsy (Ross et al., 1996). The scale assesses key facial structures during voluntary facial expressions carried out from repose. In conjunction to this, the scale also has the ability to quantify the severity of involuntary contraction of muscles from contraction of other muscles voluntarily (synkinesis). Through quantitative scoring, the scale is able to identify fine changes of facial function. Furthermore, it incorporates symmetry at rest and secondary defect of synkinesis into the scale. The SFGS has been reported in various studies to be superior to other scales in its comprehensiveness, sensitivity, and ease of use.

Before considering the impact of an intervention or quantifying the change in facial expression, multiple issues need to be taken into account. Firstly, the individual needs to have the ability to consistently reproduce the facial expression. Secondly, the method used to measure the change needs to have a high level of both intraobserver and interobserver reliability. Finally, any range of changes in the facial expression due to the impact of the intervention needs to be measurable. The SFGS has been shown to have intraclass correlation coefficients ranging from 0.838 to 0.98 for intraobserver variability and from 0.83 to 0.997 for interobserver variability. This shows that observers shown recordings of the same facial expression at two separate time points agree on the SFGS score. The glaring issue however is if the facial expressions themselves are reproducible, as this is currently unknown. For instance, if any change were to occur following the application of the SFGS pre and post intervention, was the change a result of the intervention or due to the individual not being able to reproducibly perform the facial expression? The concept of reproducibility is at both a patient-level, how well can they perform the expression on two-separate occasions and secondly on the observer and how well can the same image be scored on two-separate occasions.

At present, no studies have been carried out to assess the reproducibility of the facial expressions used in the Sunnybrook Facial Grading System. The literature review has highlighted that there have been contradicting studies on the reproducibility of facial expressions. Technology now exists to record, quantify, and analyse oro-facial expression. The most accurate and reliable method of imaging is using 3D motion capture technology (4D). This allows clinicians the ability to characterise, locate the site and severity of any residual movement deficit e.g., an asymmetrical smile. Using this technology, facial expressions can be assessed and analysed both at rest and in motion on numerous occasions.

CHAPTER 2

AIMS AND NULL HYPOTHESIS

2.1 AIMS

The aim of this study was to assess the reproducibility of facial expressions used in the Sunnybrook scale using 4D motion capture technology. This was determined by how well the participants could perform six facial expressions at a 6-month interval (T_1 and T_2).

The outcome measure was the mean absolute displacement of each landmark, in the x, y and z direction, and the Euclidian distance, between T_1 and $T_{2.}$

The x direction refers to the horizontal direction, the y direction refers to the vertical direction and the z direction refers to the anterior-posterior direction.

2.2 NULL HYPOTHESIS

There were no statistically significant differences ($p<0.05$) in the Euclidian distance of each landmark between T_1 and $T_{2,}$ for each facial expression.

For each facial expression the mean absolute displacement of each landmark, in the x, y and z direction between T_1 and T_2 was not statistically significantly different to 2.0mm ($p<0.05$).

Any differences 2.0mm or above would be deemed clinically significant.

CHAPTER 3

MATERIAL AND METHODS

3.1 STUDY DESIGN

The study was a prospective longitudinal cohort study assessing the reproducibility of facial expressions commonly used in the assessment of facial paralysis. The study was carried out on a group of volunteers at a 6-month interval.

3.2 SAMPLE SIZE CALCULATION

Based on a power of 80%, a statistical significance level of 0.05, with a standard deviation of 3.0mm (Lowney et al., 2018) a sample size of 20 individuals was required to detect a 2.0mm difference between the two-time interval (T_1 and T_2).

3.3 ETHICS

As the study involved volunteers, HRA ethical approval was not required. University ethical approval and sponsorship was granted by the University of Birmingham, Ethics Reference: ERN_20-0479 (Appendix I).

3.4 LOCATION

This study was a single centre study with volunteers selected from a convenience sample of staff and students working within the Birmingham Dental Hospital and School, Birmingham, United Kingdom. Three dimensional (4D) facial motion capture imaging of the volunteers was carried out at Birmingham Dental Hospital & School.

3.5 STUDY PARTICIPANTS

Twenty volunteers were recruited for the study, 10 male and 10 female volunteers. The volunteers were recruited from January 2021 to April 2021. The study was advertised by word of mouth from the staff and student body of the School of Dentistry, to include their friends and colleagues. No participants were patients currently being treated at the Dental Hospital. Participants were invited to voluntarily participate in the study. The first twenty (10 males and 10 females) that met the inclusion criteria were recruited.

3.5.1 Inclusion criteria

- Age 20-40 years of age.
- Willing to participate in the study and provide written consent.

3.5.2 Exclusion criteria

- Non English-speaking participants.
- Volunteers who have been diagnosed with an oro-facial deformity.
- Volunteers who have been diagnosed with oro-facial palsy.
- Volunteers who have been diagnosed with craniofacial syndromes.
- Volunteers who have been diagnosed with neurological condition which may affect their oro-facial movement e.g., Parkinson Disease.
- Males with excessive facial hair.
- Individual whose forehead is not visible.

- Individuals who had undergone any cosmetic adjunctive procedures, for example Botox or fillers.

3.6 RECRUITMENT

Volunteers interested in taking part in the study were given an information leaflet (Appendix II) which contained further details on the study. Researcher AB confirmed verbally with the volunteers whether they met the inclusion and exclusion criteria and were fit and well. Once confirmed, the volunteers were contacted to arrange an appointment for the consent process (Appendix III) and for the imaging to be carried out at T_1 and T_2 (6 months later). The consent and imaging process took approximately 30 minutes, this included capturing the volunteer and part-processing of the video sequence prior to analysis.

3.7 DI4D MOTION CAPTURE SYSTEM

The Di4D™ Pro passive stereophotogrammetric capture system (Dimensional Imaging Ltd., Glasgow, Scotland, U.K.) was used for facial image capture. The system used both high resolution colour and monochrome digital video cameras which allowed sequential 3D facial image capture. The system consisted of two pods with three video cameras in each pod: two monochrome and one colour (Model avA 1600-65km/kc, Kodak sensor model KAI02050, Basler, Germany). The two pods of cameras were held at a fixed distance apart from each other with a camera rig framework. The cameras were synchronised to a capture rate of 60 frames per second at a resolution of 2048 x 2048 pixels. To adequately illuminate the subjects during the capture sequences, two LED studio lights (NanGuang CN-600 HS Studio LED Lighting, China) were used alongside the camera pods. As part of the imaging setup, a blue screen was used as a backdrop. This is because the software used “blue screen subtraction”, separating the facial image in the foreground from the blue background through subtraction, Figure 3.1.

Each camera pod, made up of the three cameras, were connected to their own individual personal computers (PC). The two PC’s were connected together to allow fast transfer of information between the PC’s for image processing. Each PC used StreamPix (Norpix) software to simultaneously capture the two monochrome and one colour video camera sequence generated by each pod. StreamPix Remote (Norpix) was installed onto one PC which ensured the synchronisation of all six cameras at the start of capturing as well as throughout the process. Once captured, the data from each of the three cameras was saved immediately onto a Solid State Drive (SSD) to allow for future processing in proprietary .SQE format.

3.7.1 Di4D system calibration

Prior to the image capture session, calibration of the camera system was undertaken. The calibration process determined the internal and external camera parameters. The intrinsic parameters of the camera, such as focal length and lens distortions were used together with the extrinsic parameters i.e., orientation of each camera to one another, to calibrate the system. Calibration utilises a calibration target provided with the 3D imaging system. The calibration target was made of a series of dots with a known diameter and their centres a known distance apart. The calibration target was placed at a fixed distance from the cameras in a position

Figure 3.1 Di4D™ Pro passive stereophotogrammetric capture system.

where all six cameras were able to see the same points on the calibration target. Using Streampix Remote, six static images of the target were taken simultaneously. The calibration target was then repositioned, and captured, in six different horizontal and vertical positions; left and the right from the original position. Once captured, the 36 images were saved onto a single PC within a common folder. The calibration images were uploaded into DiHydra software (Dimensional Imaging Ltd., Glasgow, Scotland, U.K.), and processed to create a calibration file for that session. This calibration file was saved and, along with the volunteer's video sequence, would be used to create a 3D motion capture sequence.

3.7.2 Image capture

Prior to video sequence capture, participants were shown several short, pre-recorded video sequences that illustrated the facial expressions that were required. These included eyebrow raise, lip pucker, snarl, eye closed gentle, maximum smile with teeth together and lips apart and in addition lower teeth show. The participants were asked to practice the facial expressions before capture until they were comfortable performing them.

During facial imaging, each participant was seated on a chair directly in front of the imaging system in an upright position, in front of the blue background. They were asked to remove any glasses to prevent reflection or glare resulting from the lighting system. Surgical caps were provided if required to prevent loose strands of hair covering the forehead. The chair was adjusted to ensure the volunteers height was correct relative to the two camera pods as dictated by the live preview function in the StreamPix software. The lighting was adjusted to ensure appropriate exposure, and the volunteers were asked to focus on a pre-placed red marker positioned in the middle of the camera rig framework. Once focused on the red marker, final head adjustments were made to achieve correct alignment with the cameras.

Following correct positioning, each volunteer was asked to adopt the rest position with their teeth together and their soft tissues in repose. Video capture began using Streampix Remote and the volunteer proceeded to carry out the maximum smile with teeth together and lips apart facial expression. Once the maximum extent of the smile was reached the recording was stopped. The same process was carried out for each of the remaining five facial expressions. The video sequence was saved in a propriety format (.SQE) and reviewed through the playback function in StreamPix to check for any errors. If any errors were found the expression was re-captured and the erroneous data was discarded. Once all six expressions were correctly recorded the video sequence was reviewed in StreamPix. This was to allow the start frame (when the volunteer was at rest just before the facial expression started) and end frame (when the volunteer reached maximum facial expression) to be determined.

DiHydra software was then used to build the video capture sequence, along with the calibration file for that session, into a series of 3D images (at a rate of 60 fps, this produced 60 3D images per 1 second of video capture). Once constructed, the series of 3D images were saved and viewed for analysis as a continuous video sequence through Di4Dview or as individual image frames with Di3Dview. The total file size for each 3D video sequence of each facial expression was approximately 15GB.

3.7.3 Image landmarking

For each volunteer the T_1 maximum smile with teeth together and lips apart facial expression was imported into Di4DView. Di4DView allowed placement of the landmarks through manipulation of the image in terms of rotation and zoom features as well as viewing the image in three different planes simultaneously, which allowed more accurate landmark placement in all three dimensions. Four “stabilisation landmarks” were placed on the first frame of the image sequence depending on where there was minimal movement during the facial expression. For example, maximum smile with teeth together and lips apart facial expression, the stabilisation landmarks were generally placed on the forehead, as this did not undergo deformation. The “stabilisation landmarks” controlled any head movements that may have occurred when carrying out the facial expression. With regards to the landmarks, this were placed specifically depending on the facial expressions as follows,

- Eyebrow lift - left and right landmarks were placed at the junction of the eyebrow and skin perpendicular to the mid-pupil of each eye (2 landmarks in total), Figure 3.2
- Eye close gentle - left and right landmarks placed on the upper and lower eyelids, with each point being perpendicular to the mid-pupil of each eye (4 landmarks in total), Figure 3.3
- Pucker - left and right landmarks were placed on each commissure of the mouth (2 landmarks in total), Figure 3.4
- Snarl - left and right landmarks were placed on each nostril, at the junction of the facial bone and nasal fold (2 landmarks in total), Figure 3.5.

Figure 3.2 Eyebrow lift - left and right landmarks (red) were placed at the junction of the eyebrow and skin perpendicular to the mid-pupil of each eye (2 landmarks in total). Orange landmarks used for image stabilisation.

Figure 3.3 Eye closed gentle - left and right landmarks (red) placed on the upper and lower eyelids, with each point being perpendicular to the mid-pupil of each eye (4 landmarks in total). Orange landmarks used for image stabilisation.

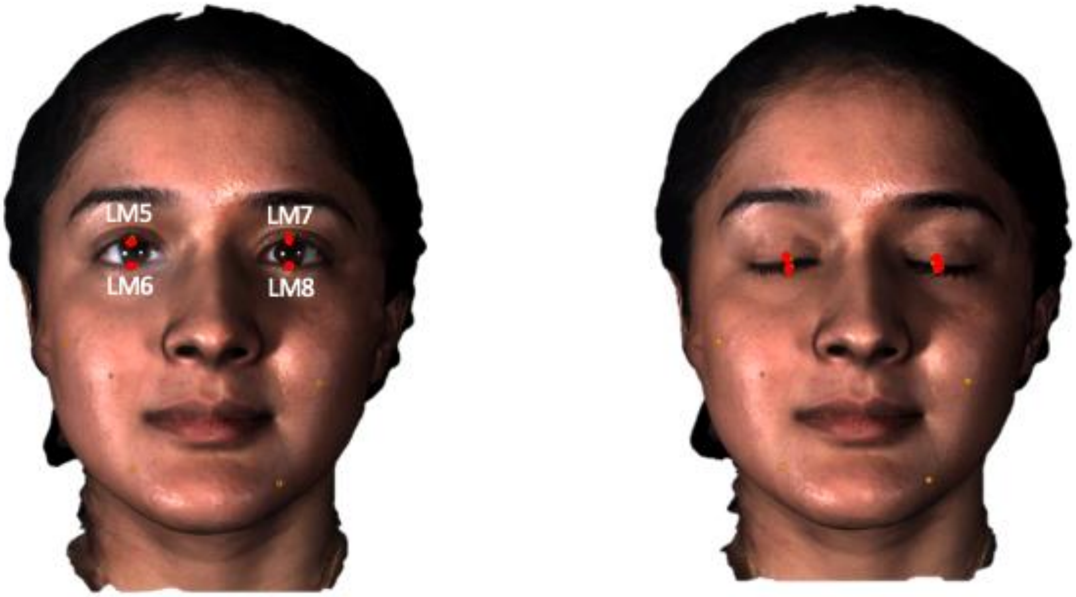


Figure 3.4 Pucker - left and right landmarks (red) were placed on each commissure of the mouth (2 landmarks in total). Orange landmarks used for image stabilisation.

Figure 3.5 Snarl - left and right landmarks (red) were placed on each nostril, at the junction of the facial bone and nasal fold (2 landmarks in total). Orange landmarks used for image stabilisation.

- Maximum smile with teeth together and lips apart (full smile) - left and right landmarks were placed on each commissure of the mouth (2 landmarks in total), Figure 3.6
- Lower teeth show - left and right landmarks were placed on each commissure of the mouth and in the middle of the vermillion border of the lower lip (3 landmarks in total), Figure 3.7

This was repeated for the T_2 maximum smile with teeth together and lips apart facial expression sequence and for all the remaining five sequences at T_1 and T_2. The landmarks were saved for future reference if necessary (.dilm).

After the landmarks were placed on the first frame for each volunteer, the automatic tracking function was enabled in the Di4Dview software. This tracked the placed landmarks for all fames within that recording sequence.

3.7.4 Landmark tracking

Following the placement of the set number of landmarks on the first frame, the automatic tracking function was initiated within the Di4DView software. This automatically and accurately tracked the landmark set, to within 0.5mm, from the first frame through the whole capture sequence using the process of “optical flow” (Al-Anezi et al., 2013).

Once the landmarks were fully tracked throughout the whole video sequence the data was exported from the Di4Dview software as a .pc2 format. The data contained the co-ordinates of each tracked landmark in the x, y and z planes for each frame of the video sequence. With the

Figure 3.6 Maximum smile with teeth together and lips apart (full smile) - left and right landmarks (red) were placed on each commissure of the mouth (2 landmarks in total). Orange landmarks used for image stabilisation.

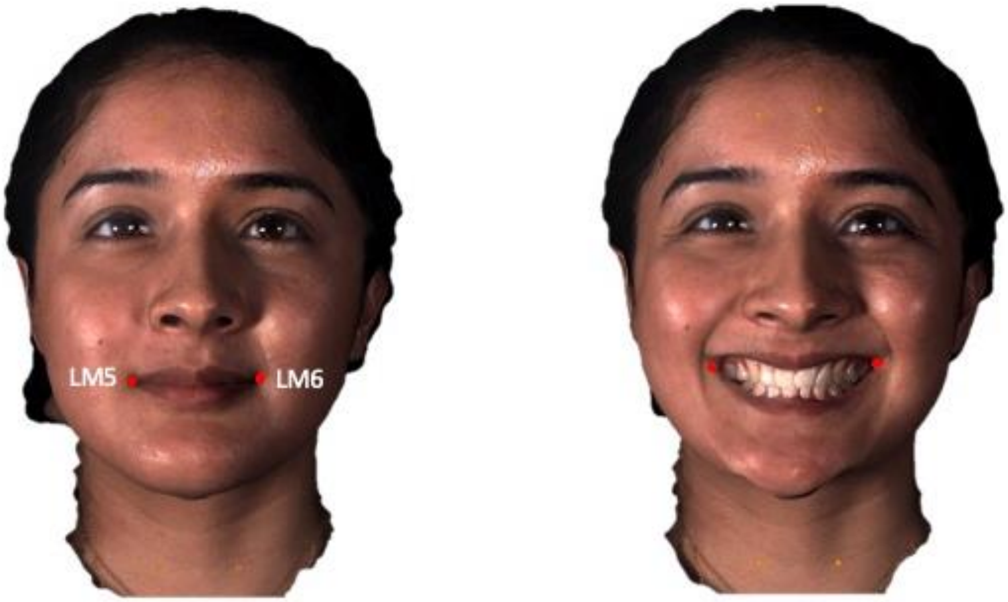


Figure 3.7 Lower teeth show - left and right landmarks (red) were placed on each commissure of the mouth and in the middle of the vermillion border of the lower lip (3 landmarks in total). Orange landmarks used for image stabilisation.

maximum number of landmarks being for the facial expression eye closed gentle, 8, and a capture rate of 60 frames per second, this produced 480 landmarks for each second of video sequence.

3.8 LANDMARK DISPLACEMENT ANALYSIS

To determine the displacement of landmark pairs from rest to the maximum facial expression, in-house developed MATLAB software (MATLAB and Statistics Toolbox Release 2017b, The Math-Works, Inc., Natick, Massachusetts, USA) was used. The analysis involved the subtraction of the x, y and z co-ordinates of each landmark at rest (first frame) to the maximum expression frame (last frame). This equated to the landmark displacement from rest to maximum expression, in the x, y and z direction, which in turn allowed the calculation of the Euclidian distances.

The mean absolute displacement of each landmark, in the x, y and z direction, as well as the Euclidian distance, were compared between T_1 and T_2 to determine reproducibility using a Students paired t-test. In addition, a one sample *t*-test was used determine if the mean absolute difference in the x. y and z directions was statistically significantly different to 2.0mm, as differences greater than this would be clinically significant.

3.9 INTRA-OPERATOR ERROR

After 4 weeks, to assess the reliability of the methodology, the 3D sequences of 20 subjects were taken at random and re-landmarked and tracked. The differences in Euclidian distance between the digitisations was assessed.

CHAPTER 4

RESULTS

4.1 SAMPLE DEMOGRAPHICS

In total 20 volunteers were recruited for the study, 10 male and 10 female. The age ranged between 20 years, 8 months and 43 years, 8 months. The mean age was 28 years, 8 months (± 3 years, 8 months).

4.2 ERROR STUDY

After 4 weeks, to assess the reliability of the methodology, the 3D sequences of 20 subjects were taken at random and re-landmarked and tracked. The differences in Euclidian distance between the digitisations was assessed. All coefficients of reliability were greater than 90% and no systematic errors were recorded. The mean overall measurement error was 0.4 ± 0.4mm.

4.3 MAXIMUM SMILE

Table 4.1 shows the mean displacement (Euclidian distance) of the landmarks for each of the 6 different facial expressions at T_1 and T_2. In addition, the mean signed and mean absolute differences in displacement is shown, as well as the 95% confidence intervals for the mean absolute differences.

The mean displacement of right cheilion (LM5) and left cheilion (LM6) at T_1 was 15.2 ± 2.8mm and 15.6 ± 3.8mm respectively. At T_2 right cheilion (LM5) and left cheilion (LM6) had a mean displacement of 15.2 ± 3.4mm and 16.0 ± 3.9mm. The mean absolute difference in right cheilion between T_1 and T_2 was 1.9mm (95% CI -1.7mm to 0.9mm). The mean absolute difference in

Table 4.1 Mean displacement (Euclidian distance) of the landmarks for each facial expressions at T_1 and T_2.

Landmark	T_1		T_2		Signed difference (T_2 - T_1)		Absolute difference (T_2 - T_1)		95% CI for the mean absolute difference (mm)		p-value
	Mean (mm)	SD (mm)	Mean (mm)	SD (mm)	Mean (mm)	SD (mm)	Mean (mm)	SD (mm)	Lower limit	Upper limit	
Lower teeth show											
LM5	8.5	4.0	9.0	3.6	0.5	2.7	1.9	1.9	-1.8	0.8	0.415
LM6	8.7	4.3	8.5	4.0	-0.2	2.5	1.7	1.7	-1.0	1.4	0.738
LM7	8.4	4.6	9.4	3.8	1.0	1.6	1.5	1.1	-1.8	-0.3	**0.010***
Snarl											
LM5	9.2	4.1	9.5	3.2	0.3	1.8	1.2	1.3	-1.2	0.5	0.440
LM6	9.2	4.4	9.4	3.8	0.2	2.2	1.4	1.6	-1.2	0.8	0.700
Pucker											
LM5	16.2	2.2	16.2	2.0	-0.2	2.4	1.9	1.5	-1.0	1.3	0.756
LM6	16.1	2.8	16.6	2.9	0.4	2.4	1.8	1.6	-1.5	0.7	0.477
Maximum smile											
LM5	15.2	2.8	15.2	3.4	0.4	2.7	1.9	1.6	-1.7	0.9	0.528
LM6	15.6	3.8	16.0	3.9	0.7	2.4	1.8	1.6	-1.8	0.4	0.195
Eyebrow lift											
LM5	11.6	4.0	11.4	3.6	-0.2	2.1	1.4	1.5	-0.8	1.1	0.729
LM6	11.3	4.0	11.3	3.4	0.0	1.1	1.4	1.6	-1.0	1.0	0.999
Eye closed gentle											
LM5	11.1	2.1	10.3	2.5	-0.8	1.5	1.3	1.1	0.2	1.6	**0.021***
LM6	3.3	1.3	3.0	1.4	-0.3	0.8	0.6	0.6	-0.1	0.6	0.154
LM7	11.2	2.2	10.3	2.4	-0.9	1.3	1.3	0.9	0.3	1.5	**0.005***
LM8	2.9	1.2	2.9	1.3	0.1	0.6	0.5	0.3	-0.3	0.2	0.723

displacement of left cheilion, between T_1 and T_2 was 1.8mm (95% CI -1.8mm to 0.4mm). Following a paired sample *t*-test, there was no statistically significant difference in the mean absolute difference of right (p=0.528) or left cheilion (p=0.195) between the two-time intervals, Table 4.2.

For right cheilion (LM5), decomposing the Euclidean distance into the respective x, y and z components, the mean absolute difference in the three directions was between 1.4mm and 1.7mm. A one sample *t*-test, with a hypothesised mean of 2.0mm, showed the differences in the x and y directions were not statistically significantly different to 2.0mm. The upper limit of the 95% confidence intervals were both above 2.0mm and could therefore be clinically significantly different in the population. In the z direction however, the mean absolute differences were statistically significantly less than 2.0mm as was the upper limit of the 95% confidence interval.

For left cheilion (LM6), the mean absolute difference between T_1 and T_2 was similar to right cheilion (LM5). The results of a one sample *t*-test showed the mean absolute differences between T_1 and T_2, in the x, y and z direction, were all statistically significantly less than 2.0mm ($p<0.05$) and were not clinically significant.

4.4 LOWER TEETH SHOW

During lower teeth show, the mean displacement of right cheilion (LM5) and left cheilion (LM6) at T_1 was 8.5 ± 4.0mm and 8.7 ± 4.3mm respectively whilst the mid-point of the lower lip (LM7) had a mean displacement of 8.4 ± 4.6mm. At T_2, right cheilion (LM5) and left cheilion (LM6) had a mean displacement of 9.0 ± 3.6mm and 8.5 ± 4.0mm respectively. The mid-point of the lower lip (LM7)

Table 4.2 The mean absolute difference of LM5 and LM6 in the x, y and z direction between T_1 and T_2 during maximum smile.

Landmark & Direction	**Mean absolute difference (T2-T1) (mm)**	**SD (mm)**	**95% CI for the mean absolute difference (mm)**		**p-value**
			Lower	Upper	
LM5					
x	1.5	1.3	0.9	2.1	0.103
y	1.7	1.0	1.2	2.1	0.158
z	1.4	1.2	0.8	1.9	**0.031***
LM6					
x	1.3	0.9	0.8	1.7	**0.002***
y	1.1	0.8	0.8	1.5	**0.001***
z	1.5	1.0	1.0	2.0	**0.037***

Results of a sample one t-test with a hypothesised mean of 2mm ($p < 0.05^*$)

had a greater mean displacement than T_1 with a mean displacement of 9.4 ± 3.8mm. For right cheilion, the mean absolute difference between T_1 and T_2 was 1.9mm (95% CI -1.8mm to 0.8mm) which was greater than the mean absolute difference for left cheilion at 1.7mm (95% CI -1.0mm to 1.4mm). The mid-point of the lower lip had a mean absolute difference between T_1 and T_2 of 1.5mm (95% CI -1.8mm to -0.3mm). Between the two-time intervals, the mean absolute difference of the displacement of the mid-point of the lower lip was statistically significantly different (p=0.010), whilst right cheilion (p=0.415) and left cheilion (p=0.738) were not statistically significantly different following a paired *t*-test, Table 4.3.

Assessing the x, y and z distances for each of the three landmarks, right cheilion (LM5) had a mean absolute difference ranging from 1.3mm to 1.7mm. Following a one sample *t*-test, only the absolute differences in the y direction was statistically significantly less than the 2mm (p=0.019). However, in the x and z direction the difference was not statistically significantly different to 2.0mm. The upper confidence interval was greater than 2.0mm in the x and y direction, indicating this difference could also potentially be clinically significantly.

For left cheilion (LM6), the mean absolute difference in all three directions ranged from 0.9mm to 2.1mm, with the least difference being in the vertical direction (0.9mm). The one sample *t*-test showed the mean absolute differences between T_1 and T_2, in the x and y directions were all statistically significantly less than 2mm (p<0.05). The largest difference was in the z-direction with an upper 95% confidence interval limit of 3.1mm and a mean difference greater than 2.0mm.

Table 4.3 The mean absolute difference of LM5, LM6 and LM7 in the x, y and z direction between T_1 and T_2 during lower teeth show.

Landmark & Direction	**Mean absolute difference (T2-T1) (mm)**	**SD (mm)**	**95% CI for the mean absolute difference (mm)**		**p-value**
			Lower	Upper	
LM5					
x	1.7	1.3	1.0	2.3	0.272
y	1.3	1.3	0.6	1.9	**0.019***
z	1.9	2.0	0.9	2.8	0.818
LM6					
x	1.3	1.0	0.8	1.8	**0.005***
y	0.9	0.7	0.6	1.3	**0.001***
z	2.1	2.1	1.1	3.1	0.882
LM7					
x	0.7	0.4	0.5	0.9	**0.001***
y	1.4	1.0	0.9	1.9	**0.022***
z	1.5	1.2	0.9	2.0	0.063

Results of a sample one t-test with a hypothesised mean of 2mm (**$p < 0.05^*$**)

For the mid-point of the lower lip (LM7), a one sample *t*-test showed the mean absolute differences, between T_1 and T_2, in z direction was not statistically significantly different ($p>0.05$) to 2mm. In the x and y directions the differences were statistically significantly less than 2.0mm and the upper limit of the 95% confidence intervals for the mean differences less than 2mm.

4.5 SNARL

Both landmarks on the right nasal fold (LM5) and left nasal fold (LM6) underwent similar mean displacements at T_1 of 9.2 ± 4.1mm and 9.2 ± 4.4mm respectively. At T_2, right nasal fold had a mean displacement of 9.5 ± 3.2mm whilst left nasal fold underwent a mean displacement of 9.4 ± 3.8mm. This consistency between the landmarks was also reflected in the mean absolute differences between T_1 and T_2. Right nasal fold underwent a mean absolute difference of 1.2mm (95% CI -1.2mm to 0.5mm) whilst left nasal fold experienced a mean absolute difference of 1.4mm (95% CI -1.2mm to 0.8mm). Following a paired *t*-test, neither displacement of right nasal fold ($p=0.440$) nor left nasal fold ($p=0.700$) were statistically significantly different between the two-time intervals, Table 4.4.

In the x, y and z directions, between T_1 and T_2, the mean absolute difference of right nasal fold (LM5) was between 1.1mm and 1.4mm. The difference in the x direction was statistically significantly less than 2.0mm, as were the upper limits of the 95% confidence interval. In the y and z directions the differences were not statistically significantly different; the upper limit of the 95% confidence intervals were 2.0mm, suggesting the differences were bordering on being clinically significant in the population.

Table 4.4 The mean absolute difference of LM5 and LM6 in the x, y and z direction between T$_1$ and T$_2$ during snarl.

Landmark & Direction	Mean absolute difference (T2-T1) (mm)	SD (mm)	95% CI for the mean absolute difference (mm)		p-value
			Lower	Upper	
LM5					
x	1.1	1.0	0.6	1.5	**0.001***
y	1.4	1.3	0.8	2.0	0.051
z	1.3	1.5	0.6	2.0	0.057
LM6					
x	1.0	0.9	0.6	1.4	**0.001***
y	1.5	1.5	0.8	2.2	0.154
z	1.2	1.4	0.6	1.9	**0.020***

Results of a sample one t-test with a hypothesised mean of 2mm ($p < 0.05^{*}$)

The left nasal fold (LM6) had a similar mean absolute difference as its counterpart, ranging from 1.0mm to 1.5mm, between T_1 and T_2. However, in the x and z direction the difference was statistically significantly less 2.0mm. The upper limit of the 95% confidence intervals were less than 2.0mm and therefore being clinically insignificant. The in the y direction, the mean absolute difference was not statistically significantly less than 2.0mm with a 95% confidence interval upper limit over 2.0mm suggesting a clinically significant difference.

4.6 PUCKER

Based on the mean Euclidean distances, the displacement of right cheilion (LM5) was 16.2 ± 2.2mm and 16.2 ± 2.0mm at T_1 and T_2 respectively. For left cheilion (LM6), the mean displacement was 16.1 ± 2.8mm at T_1 and 16.6 ± 2.9mm at $T_{2.}$ The mean absolute differences in displacement of right cheilion during lip pucker between T_1 and T_2 was found to be 1.9mm (95% CI -1mm to 1.3mm) whilst left cheilion was 1.8 (95% CI -1.5mm to 0.7mm). A paired *t*-test showed that the mean absolute difference, for both right cheilion (p=0.756) and left cheilion (0.477), were not statistically significantly different between T_1 and T_2, Table 4.5.

Breaking down the Euclidean distances to the x, y and z directions, right cheilion (LM5) had a mean absolute difference that was quite varied for all three planes of space, with a difference of 1.2mm. A one sample *t*-test showed that displacement in the x direction only had a mean absolute difference that was statistically significantly less than 2.0mm. In the y and z direction the mean absolute difference were both not statistically significantly different to 2.0mm but were clinically

Table 4.5 The mean absolute difference of LM5 and LM6 in the x, y and z direction between T_1 and T_2 during pucker.

Landmark & Direction	Mean absolute difference (T2-T1) (mm)	SD (mm)	95% CI for the mean absolute difference (mm)		p-value
			Lower	Upper	
LM5					
x	1.0	0.8	0.7	1.4	**0.001***
y	2.2	1.8	1.3	3.0	0.689
z	2.2	1.7	1.4	3.0	0.590
LM6					
x	1.9	1.5	1.2	2.6	0.791
y	2.3	1.7	1.5	3.1	0.480
z	2.0	1.6	1.2	2.8	0.991

Results of a sample one t-test with a hypothesised mean of 2mm ($p < 0.05^*$)

significantly different due to the upper limit of the 95% confidence levels being greater than 2mm (3.0mm and 3.0mm respectively).

Displacement of left cheilion (LM6) in the x, y and z direction had a mean absolute difference that was less varied, with a difference of 0.4mm. However, following a one sample *t*-test the mean absolute difference in all three directions were not statistically significantly different ($p<0.05$) between T_1 and T_2. In addition, all were clinically significantly different with the upper limits 95% confidence intervals greater than 2.0mm.

4.7 EYEBROW LIFT

The midpoint of the right eyebrow (LM5) had a mean displacement at T_1 of 11.6 ± 4.0mm whilst the midpoint of the left eyebrow (LM6) recorded a mean displacement of 11.3 ± 4.0mm. At T_2 the mean displacement of the right eyebrow reduced to 11.4 ± 3.6mm whilst the left eyebrow at T_2 showed a similar mean displacement to T_1, 11.3 ± 3.4mm. The mean absolute difference between T_1 and $T_{2,}$ for the right eyebrow was 1.4mm (95% CI -0.8mm to 1.1mm) and left eyebrow was 1.4mm (95% CI -1.0mm to 1.0mm). There was no statistically significant difference in mean absolute difference for right and left eyebrow midpoint between T_1 and $T_{2,}$ Table 4.6.

Assessment in the x, y and z direction for LM5 showed that the mean absolute difference for x and y were 0.7mm and 1.4mm respectively, whilst in the z direction was 0.9mm. A one sample *t*-test showed that in the y direction the mean absolute distance between T_1 and T_2 was not statistically significant. An upper limit of 2.0mm in the 95% confidence interval of 2.0mm suggests this

.

Table 4.6 The mean absolute difference of LM5 and LM6 in the x, y and z direction between T_1 and T_2 during eyebrow lift.

Landmark & Direction	**Mean absolute difference (T2-T1) (mm)**	**SD (mm)**	**95% CI for the mean absolute difference (mm)**		**p-value**
			Lower	Upper	
LM5					
x	0.7	0.5	0.4	0.9	**0.001***
y	1.4	1.4	0.7	2.0	0.058
z	0.9	1.0	0.5	1.4	**0.001***
LM6					
x	0.7	0.8	0.4	1.1	**0.001***
y	1.3	1.5	0.6	2.0	0.052
z	1.1	1.0	0.6	1.5	**0.001***

Results of a sample one t-test with a hypothesised mean of 2mm (**$p < 0.05^*$**)

difference maybe clinically significant. In the x and z directions the differences were statistically significantly less than 2.0mm.

For the left mid eyebrow point, LM6, in the y direction, following a one sample *t*-test, the difference was clinically significant but not statistically significant. Whilst in the x and z directions the differences were clinically and statistically significantly less than 2.0mm.

4.8 EYE CLOSED GENTLE

Eye closed gentle utilised four landmarks, two associated with the right eye and two with the left eye. The greatest mean displacement at T_1 and T_2 were the landmarks on the upper eyelids on the right (LM5) and left eye (LM7). At $T_{1,}$ the mean was 11.1 ± 2.1mm for the right upper eye lid (LM5) and 11.2 ± 2.2mm for the left upper eye lid (LM7). The lower eye lids underwent minimal movement with a of mean displacement of 3.3 ± 1.3mm for the right lower eye lid (LM6) and 2.9 ± 1.2mm for left lower eye lid (LM8). At T_2 both right (LM5) and left (LM7) upper eye lids underwent near symmetrical amounts of displacements, 10.3 ± 2.5mm and 10.3mm ± 2.4mm respectively. The lower eye lids were also almost symmetrical, with the lower right eyelid (LM6) undergoing a mean displacement of 3.0 ± 1.4mm and lower left eyelid (LM8) having a mean displacement of 2.9 ± 1.3mm. The near perfect symmetry of the pair of contralateral landmarks is reflected in the absolute mean difference. The upper right eyelid (LM5) had a mean absolute difference between T_1 and T_2 of 1.3mm (95% CI 0.2mm to 1.6mm) and the left upper eyelid was 1.3mm (95% CI 0.3mm to 1.5mm). Both were statistically significantly different ($p<0.05$) between the T_1 and T_2. With regards to the lower eyelid landmarks, the lower right eyelid had an absolute mean displacement of 0.6mm (95% CI -0.1mm to 0.6mm) and the left lower eye lid 0.5mm (95% CI -0.3mm to 0.2mm). Neither of these were statistically significant ($p>0.05$), Table 4.7.

A one sample *t*-test showed that the x, y and z directions for all for landmarks were statistically significantly less than 2.0mm as was their upper limit of their respective confidence intervals, showing none were clinically significant.

Table 4.7 The mean absolute difference of LM5 and LM6 in the x, y and z direction between T_1 and T_2 during eyes closed gentle.

Landmark & Direction	**Mean absolute difference (T2-T1) (mm)**	**SD (mm)**	**95% CI for the mean absolute difference (mm)**		**p-value**
			Lower	Upper	
LM5					
x	1.0	0.8	0.6	1.4	**0.001***
y	1.2	1.0	0.8	1.7	**0.004***
z	0.8	0.5	0.5	1.0	**0.001***
LM6					
x	0.8	0.5	0.5	1.0	**0.001***
y	0.7	0.6	0.4	0.9	**0.001***
z	0.5	0.4	0.3	0.7	**0.001***
LM7					
x	1.0	0.9	0.5	1.4	**0.001***
y	1.1	0.7	0.8	1.5	**0.001***
z	0.7	0.5	0.5	0.9	**0.001***
LM8					
x	0.6	0.3	0.4	0.7	**0.001***
y	0.7	0.5	0.4	0.9	**0.001***
z	0.5	0.4	0.4	0.7	**0.001***

Results of a sample one t-test with a hypothesised mean of 2mm ($p < 0.05^{*}$)

CHAPTER 5

DISCUSSION

5.1 DISCUSSION

The site and severity of the facial palsy determines its management, which can range from conservative management to medical intervention with steroids, or surgical intervention (Stew and Williams, 2013). The aim of treatment of facial palsy would be to allow the patient to regain facial muscle movement and perform facial expressions to the same magnitude as prior to the palsy. To determine which intervention is appropriate, the severity of the facial palsy needs first to be determined through accurate clinical assessment with the use of a validated scale. Following the intervention, the same scale should be used to determine the improvement the patient has experienced due to the intervention and whether further treatment is necessary. The premise of using the scales, such as the Sunnybrook Scale, are that all facial expressions assessed are reproducible and therefore any variation between pre- and post-intervention are due to the intervention and not due to natural human variation or the inability of an individual to be able to reproducibly undertake that facial expression. The drawback with using scales is that most involve the subjective measurement of displacement of the anatomical landmarks or facial features by a clinician. Objective measurements through the use of technology have previously not been adopted due to the cost of the equipment, lack of technology available to clinicians, technical expertise and time. Outcomes based on subjective assessment could mean some changes may not be noticeable by the clinician. The purpose of this study was to objectively assess the reproducibility of facial expressions used in the Sunnybrook scale. This study specifically studied the magnitude, based on the Euclidian distance and x, y and z directions, of anatomical landmarks used in the various facial expressions over two time points 6-months apart. Previous studies to date have used a much shorter interval ranging from 15 minutes (Sawyer et al., 2009; Ju et al., 2016) to 12 weeks (Ozsoy et al., 2019).

The landmarks chosen in the present study were based on those which would undergo most displacement whilst undertaking the facial expression, such as commissures of the mouth for lip pucker and maximum smile. The landmarks were also chosen as they are anatomically well defined and were associated with the least landmarking error (Johnston et al., 2003; Gwilliam et al., 2006). In the present study, lack of reproducibility was defined if there were differences in magnitude of landmarks between the two time points in either the x, y, z direction or in the Euclidian distance. Lack of reproducibility can be defined as either clinically or statistically significant. Discrepancies of 2.0mm or more have been shown to be the visual threshold, which a lay person is able to identify (Haraguchi et al., 2002; Chebib and Chamma, 1981; Silva et al., 2013). Therefore, in the present study 2.0mm was determined to be the clinically significant value. This meant that should there be more than a 2mm difference between the landmarks in each facial expression, in the x, y or z directions, between T_1 and T_2, the facial expression may not be clinically reproducible.

5.2 RECRUITMENT

Volunteers recruited for the study were aged between 21-40 years old and only those that were medically fit and well with no oro-facial palsy or syndromes affecting facial movements or neurological disease were recruited. The sample was a convenience sample based on staff members and students at the Birmingham Dental Hospital and School and were recruited using a standardised protocol to ensure a homogenous sample was collected.

Certain groups of individuals, for example, those that had botulinum toxin (Botox) injections between T_1 and T_2 were excluded from the study. Botox injections are carried out to cause specific facial muscle paralysis and temporarily eliminate wrinkles (Chang et al., 2016). This can impact a variety of facial expressions, such as eyebrow lift. Eyebrow lift involves the primary contraction of the frontal head of the occipitofrontalis muscle. Eyebrow lift can also simultaneously cause the development of horizontal wrinkles on the forehead (Fujimura and Hotta, 2011). The occiptofrontalis muscle, in conjunction with the glabellar complex, also has a role in carrying out the facial expression of a frown, forming frown lines. These dynamic frown lines and forehead wrinkles are a concern to those who are more aesthetically driven and therefore seek Botox treatment to improve their appearance (Nestor et al., 2020). Botulinum toxin functions by blocking the release of the chemical acetylcholine, which in turns leads to paralysis of the muscles within the occipitofrontalis muscle and glabellar complex areas which can last for a period of 3-6 months but can be longer depending on how quickly the body takes to break the components down (Satriyasa, 2019). The paralysis of these muscles, in terms of facial expressions, would result in the reduced ability to raise one's eyebrows thereby changing its maximum between pre- and post-Botox injections.

5.3 LANDMARKING

The majority of anatomical landmarks chosen were similar to those used in previous studies (Al-Anezi et al., 2013; Hallac et al., 2017; Gattani et al., 2020). The exceptions were “eye closed gentle” where previous studies have used exocanthion and endocantion as landmarks for assessment of “eye closure tight”. However, those landmarks would undergo minimal, if any, displacement during “eye closed gentle” movement. The problem with this specific facial expression in terms of landmarks was that, unlike facial landmark digitisation at rest where most of the face is clearly visible, much of the anatomy of the eye disappears when the eye is open. The landmark chosen was therefore based on previous studies that assessed gentle eye closure (Zhao et al., 2020). With regards to the stabilisation landmarks, the facial expression “eyebrow lift” actively involved the movement of the forehead. Therefore, for any facial expression involving the mid or lower portion of the face, the stabilisation points on the forehead were utilised (Al-Anezi et al., 2013; Hallac et al., 2017; Gattani et al., 2020). However, for “eyebrow lift,” the stabilisation points were placed around the lower third of the face which were static during the expression.

5.4 TIME INTERVAL BETWEEN T_1 AND T_2

This study is novel as no other study was identified in assessing the reproducibility of facial expressions over a 6-month period, with others capturing over a shorter time frame, with the maximum being 4 weeks (Sawyer *et al.,* 2009; Johnston *et al.,* 2003). A more clinically relevant and therefore valid assumption is the degree of reproducibility in facial expressions over a greater time interval as this would be more realistic. This prevents any memory bias and thereby the expression is carried out to what the volunteer feels is the maximum on that day.

In addition, the present study has collected normative data that can be used as a control group to assess whether patients have a facial palsy. The normative data provides baseline data for a cohort of healthy individuals, for example, the mean magnitude of displacement of the eyebrows during eyebrow lift is 11.0-11.5mm. For patients attending with possible facial nerve palsy, the maximum displacement of the patient can be measured and compared with this baseline data to give an indication on whether a facial palsy may be possibly diagnosed.

5.5 REPRODUCIBILITY OF FACIAL EXPRESSIONS

Based on the Euclidian distances, the results of this study, Table 4.1, showed four of the six facial expressions i.e., snarl, pucker, maximum smile, and eyebrow lift were reproducible. The other two expressions, eye closed gentle and lower teeth show, had differences in Euclidian distance which were statistically significantly different between the two-time intervals. However, decomposing the Euclidian distance into the x, y and z direction for each expression provided a more detailed explanation.

A novel approach of the present study addresses the potential shortcomings of previous studies which have used the differences in Euclidian distances, which measures the magnitude but not direction of the landmark displacement. For example, based on the Euclidian distance lip “pucker” was a reproducible facial expression but when in decomposing the Euclidian distance into its component x, y and z directions, some of the differences in direction were clinically significantly greater than the 2.0mm threshold. It is difficult to directly compare the results of the present study with previous studies as each study has different outcome measures. The more recent studies use the whole facial mesh for analysis and the Root Mean Square (RMS) error between facial meshes (Özsoy *et al.,* 2019; Qui *et al.,* 2022). This method of analysis takes into account regions of the face that have not moved and will therefore bias the results, as these regions will be reproducible on two separate occasions. Another study uses Principle Component Analysis (PCA) which is a statical method for describing patterns in large amounts of data but may be difficult to interpret clinically (Popat et al., 2010).

5.5.1 Eyebrow lift

The frontalis muscle is the muscle responsible for elevating the eyebrows during eyebrow lift. It is comprised of different components, with occipitofrontalis being the key muscle of interest during eyebrow lift (Costin *et al.,* 2015). The muscle is made up of vertical fibres only and when contracting, the frontalis portion of the muscle pulls the skin of the eyebrows vertically upwards. It is therefore understandable that, as the muscle consists of vertical fibres only, the largest difference in displacement between T_1 and T_2 would occur in the y direction. This may explain why eyebrow lift is least reproducible in the y direction.

5.5.2 Maximum smile

Based on the Euclidian distances, the results shows that maximum smile is reproducible, which coincides with other work (Peck and Peck, 1995; Trotman *et al.,* 2000, Johnston *et al.,* 2003; Ju

2016), but contradicts some studies which showed that “smile with lips open” was one of the least reproducible (Sawyer *et al.,* 2009). In this study, for maximum smile, the Euclidean mean distance of the commissures of the mouth at T_1 and T_2 were between 15.2mm to 16mm. This is in agreement with Lowney *et al* (2018) which had a maximum displacement of the left and right cheilions between 15.8mm to 17.5mm. It is common for there to be some mild variations during maximum smile both within and between individuals (Khambay *et al.,* 2019)

Maximum smile is formed by the static and dynamic relationship between the soft tissue and dento-skeletal components of the face (Mercado-Garcia *et al.,* 2021). It involves a number of different muscles including zygomaticus major, zygomaticus minor, depressor anguli oris, levator labii superioris and risorius. The contraction of these muscles occurs at different stages during the facial expression to raise the upper lip and commissures of the mouth in all three planes of space (Ackerman and Ackerman, 2002). The results show that whilst reproducible, the landmarks undergo significant displacement in the x, y and z directions and border on being clinically significant (>2.0mm). This is to be expected considering the number of muscles involved in the facial expression and the different orientations of each muscle. It was seen though that, despite the complexities involved in carrying out this facial expression, it was more reproducible than other facial expressions, such as pucker. The possibility of this may be that, unlike pucker, smiling is a facial expression carried out on a daily basis. This repetitive nature has allowed the muscles to become accustomed to synchronising with each other and carry out the facial expression to a reproducible manner. In conjunction with this, the dental-skeletal components provide a boundary, limiting the direction the landmarks on the commissures of the mouth can displace towards, thereby increasing its reproducibility.

A reason for the differences seen in other studies in terms of reproducibility (Sawyer *et al.,* 2009) could have been the ambiguity in the instructions given. During recording the facial expression, volunteers were specifically asked to perform it to the maximum. If not clearly stated, it can lead to variation in the facial expression (Burres, 1985).

5.5.3 Eye closed gentle

Simply looking at the Euclidean distances in Table 4.1, it would appear that eye closed gentle had poor reproducibility, due to the fact that the landmarks on the upper right and left eye lids were statistically significantly different between T_1 and T_2. When decomposing the Euclidean distances into the x, y and z directions, it can be seen that in fact eye closed gentle was reproducible with all of the x, y and z distances being statistically significantly less than the 2.0mm threshold. The results from this align with the results from Sawyer *et al,* (2009), which showed that apart from repose, eyes closed gentle was a reproducible facial expression. A reason for the reproducibility of this facial expression is possibly due to the boundary the upper and lower eyelids provide each other, restricting further movement from occurring. Unlike the other facial expressions such as lower teeth show, eye closed gentle is restricted and will stop once the upper and lower eyelids touch. This definitive end point allows for reproducibility to occur whereas for lower teeth show, the lower lip has no physical restrictions in terms of how much it can displace in the y direction.

The participant will have reached their maximum when they feel that muscle has reached its limit, however that can introduce variations.

5.5.4 Snarl

It was seen that apart from one component (left alar base in the y direction), none of the mean absolute difference values exceeded the 2.0mm threshold and the facial expression were therefore reproducible to an extent.

The levator labii superioris consists of three different heads with different points of origins and insertions. The levator labii superioris alaeque nasi is the specific branch of the muscle involved in “snarl,” lifting both the upper lip and wings of the nose superomedially (Hur *et al.,* 2010). Other muscles surrounding the nose and the mouth, such as the zygomatic major and minor muscles are involved in pulling the mouth and nasal wings superolaterally. The common theme with all these muscles is the elevation of the nasal wings in a vertical direction, thereby giving the greatest displacement of landmarks in this region in a y direction. This coincides with the results of the study. Both right and left alar base show the greatest mean absolute difference in the y direction (1.4mm and 1.5mm respectively).

5.5.5 Lower teeth show

Lower teeth show was seen to have poor reproducibility. It was seen that all three landmarks had the greatest error in the anterior-posterior direction (z direction). There are different depressor muscles involved in lowering the lower lip, such as the depressor anguli oris and the depressor labii inferioris. The depressor anguli oris's fibres are orientated in various directions due to its triangular shape. Depressor labii inferioris is orientated obliquely and joins with the platysma muscles. Both these muscles have different orientations and both contract to move the lower lip vertically but also anterior-posteriorly. The different orientations of these muscles may provide some explanation into why the lower lip has such poor reproducibility during lower teeth show, with each muscle pulling the lip in a different direction. The mentalis muscles also has a large role to play in lower teeth show, with its innervation causing everting of the lower lip. It was seen that many volunteers were unable to depress the lower lip vertically and instead everted the lower lip through contraction of the mentalis muscle to show their lower teeth. Through contraction of this muscle, the landmark of the lower lip would have its greatest displacement in the z direction.

It was also noted during lower teeth show, there was variation in contraction of the platysma muscle. The platysma muscles originates from the skin of the neck and upper chest and inserts into the inferior border of the mandible as well as the skin of the lower face. The contraction of the platysma muscle results in the depression of the lower lip and angle of the mouth as well as impacts the commissures of the mouth in a z direction (Hoerter and Patel, 2019). It was seen that there was wide variability in between T_1 and T_2 and between volunteers in lower teeth show due to the varying use and contraction of the platysma muscle.

5.5.6 Lip pucker

Simply looking at the Euclidian distances, it appears to show that the facial expression is reproducible, with right and left cheilion undergoing very similar mean displacements both at T_1

and T_2. In conjunction to this, neither landmarks were statistically significant between the two time points (p=0.756 and p=0.477 respectively). In the x, y and z directions however apart from the x direction for right cheilion, differences in all other directions for right and left cheilion were clinically significant as their upper limit of the confidence intervals were above 2.0mm, with some even being greater than 3.0mm. This shows that the facial expression had a poor reproducibility between the two-time intervals.

Anatomically, lip pucker is controlled by obicularis oris muscle. This muscle, whilst considered to be two muscles, functionally consists of both deep and superficial layers which attach to the dermis of the upper and lower lips and surround the commissures of the mouth to provide function (Ghassemi *et al.,* 2003). The muscles fibres orientate in different directions but work harmoniously to carry out functions such as sucking or pucker.

The uniqueness of the orbicularis oris muscle is that it has no connection to any bony structures. It may be for this reason, there is a large degrees of natural variation, as everyone will have a different tone of muscle and be capable of exerting that muscle a different amount, leading to a lack of reproducibility between participants. Lip pucker not only involves the contraction of orbicularis oris, but also involves the relaxation of all surrounding muscles such as the buccinator muscles. The extent of the relaxation of the surrounding muscles also determines the maximum extent of lip pucker, introducing further variation.

5.6 LANDMARK DIGITISATION

The assessment of facial soft tissues in an x, y and z direction requires the accurate identification and placement of these landmarks. The development of technology has allowed the placement of landmarks digitally on to the captured image. The landmarks can be placed in the researchers own time and placed where required, such as the commissures of the mouth for lip pucker. The use of a marker free system was initially flawed, with issues regarding resolution and frame rates. The advancement in in camera technology however has overcome these issues and subsequently markerless systems have since been used to capture facial expressions in different cohorts (Al-Rudainy *et al*., 2018; Shujaat *et al*., 2014; Alagha *et al.*, 2018; Lowney *et al*., 2018; Popat *et al.*, 2010). There are multiple benefits of digital landmarking including it being operator and participant friendly, with the participant not required to attend prior the scanning to have their face landmarked. The use of the digital landmarking allows magnification and easy placement of landmarks in three planes of space accurately, with the use of magnification and rotation of the images on Di4D software allowing the landmarks to be placed reproducibly to within 0.2mm (Khambay *et al.,* 2008). It is important to note though that the marker free system is still reliant on the researcher reliably and accurately landmarking the specific structures required on the first frame.

5.7 LIMITATIONS OF THE STUDY

The study provides information on a control group to which subsequent research on facial palsy patients can be compared. The drawbacks seen with the use of the 4D motion capture system were the bright lights required to illuminate the individuals, as some found it uncomfortable.

Another disadvantage was the large amount of data that was generated which can be both difficult to analyse and to store.

Finally in the male cohort, the presence of facial hair made it difficult on occasions to carry out landmarking and tracking. The 3D tracking of the facial skin surface has advanced immensely, a major limitation to the system though is the presence of facial hair. The presence of thin structures such as hair results in the capturing of a “rough shrink-wrapped” 3D surface (Winberg *et al.,* 2022). Whilst it was advised that individuals should be clean shaven prior to scanning, this could not be enforced. It also increases the complexity of the process, and many males are unwilling to shave as it can result in an undesired change in their appearance (Winberg *et al.,* 2022).

CHAPTER 6

CONCLUSIONS

6.1 CONCLUSIONS

The objective assessment of the facial expressions used in Sunnybrook Facial Grading System (SBFGS) using 4D motion capture technology has shown that, based on the Euclidian distance, four of the six facial expressions i.e., snarl, pucker, maximum smile, and eyebrow lift were reproducible. The other two expressions, eye closure gentle and lower teeth show, had differences in Euclidian distance which were statistically significant over a 6-month period.

Based on the Euclidian distance the null hypothesis was rejected for eye closure gentle and lower teeth show.

None of the differences in Euclidian distance were statistically significantly greater than 2.0mm for any of the six facial expressions.

With respect to the x, y and z directions none of the six facial expressions had mean absolute differences statistically significantly greater than 2.0mm. However, the majority were around 2.0mm, but both lip pucker and lower teeth show were associated with the greatest difference. In both facial expressions the upper 95% confidence interval limit exceeded 3.0mm.

This study has shown that even after a 6-month interval the facial expressions used in the Sunnybrook scale are reproducible to within 2.0mm to 3.0mm. Lip pucker and lower teeth show appeared to have the lowest reproducibility. The clinical impact of the study is that any changes in facial movement, because of an intervention, need to be greater than 3.0mm to be a true change and not a result of decreased reproducibly of the facial expression over time.

CHAPTER 7

REFERENCES

7.1 REFERENCES

Ackerman, M.B. and Ackerman, J.L. (2002) Smile analysis and design in the digital era. *Journal of Clinical Orthodontics*, 36 (4): 221–236.

Alagha, M.A., Ju, X., Morley, S., et al. (2018) Reproducibility of the dynamics of facial expressions in unilateral facial palsy. *International Journal of Oral and Maxillofacial Surgery,* 47 (2): 268–275.

Al-Anezi, T., Khambay, B.S., Peng, M.J., *et al.* (2013) A new method for automatic tracking of facial landmarks in 3D motion captured images (4D). *International Journal of Oral and Maxillofacial Surgery, 42 (1): 9–18.*

Al-Rudainy, D., Ju, X., Mehendale, F., *et al.* (2018) Assessment of facial asymmetry before and after the surgical repair of cleft lip in unilateral cleft lip and palate cases. *International Journal of Oral and Maxillofacial Surgery,* 47 (3): 411–419.

Ayoub, A., Garrahy, A., Hood C., *et al.* (2003) Validation of a vision-based, three-dimensional facial imaging system. *The Cleft Palate Craniofacial Journal, 40: 523-529.*

Ayoub, A., Siebert, P., Moos, K., *et al.* (1998) A vision-based three-dimensional capture system for maxillofacial assessment and surgical planning. *British Journal of Oral and Maxillofacial Surgery, 36: 353-357.*

Berkowitz, S. and Cuzzi, J. (1977) Bio Stereometric analysis of surgically corrected abnormal faces. *American Journal of Orthodontics, 72: 526-538.*

Blais, F. (2004) Review of 20 years of range sensor development. *Journal of Electronic Imaging, 13: 231-240.*

Brenner, M.J. and Neely, J.G. (2004) Approaches to grading facial nerve function. *Seminar in Plastic Surgery*, 18 (1): 13-21.

Burns, N. and Grove, S. (1997) *The practice of nursing research: conduct critique & utilization*. 3rd edition. Philadelphia, Pennsylvania: W.B. Saunders Company.

Burres, S.A. (1985) Facial biomechanics: the standards of normal. *Laryngoscope,* 95(6): 708- 714.

Burrows, A.M. and Smith, T.D (2003) Muscles of facial expressions in Otolemur, with a comparison to Lemuroidea. *The Anatomical Record Part A,* 274A: 827-836.

Carlson, B.M. (1981) *Patten's foundations of Embryology*. 4th Edition. New York: McGraw-Hill Book Company.

Chang, B.L., Wilson, A.J., Taglienti, A.J., *et al.* (2016) Patient perceived benefit in facial aesthetic procedures: FACE-Q as a tool to study botulinum toxin injection outcomes. *Aesthetic Surgery Journal,* 36 (7): 810-820.

Chebib, F.S. & Chamma, A.M. (1981) Indices of craniofacial asymmetry. *The Angle Orthodontist,* 51: 214–226.

Costin, B.R., Plesec, T.P., Sakolsatayadorn, N., *et al.* (02015) Anatomy and histology of the frontalis muscle. *Ophthalmic Plastic and Reconstructive Surgery*, 31 (1): 66-72.

D'Apuzzo, N. (2002) Surface measurement and surface tracking of human body parts from multi-image video sequences. *ISPRS Journal of Photogrammetry and Remote Sensing, 56 (4): 360-375.*

Da Silveira, A.C., Daw, J.L., Kusnoto, B., *et al.* (2003) Craniofacial applications of three-dimensional laser surface scanning. *Journal of Craniofacial Surgery,* 14: 449–456.

Darwin, C. (1989) *Expression of Emotion in Man and Animals. From Selected Works of Charles Darwin, authorized edition.* New York, D. Appleton and Company.

Djordjevic, J., Lewis, B.M., Donaghy, C.E., *et al.* (2014) Facial shape and asymmetry in 5-year-old children with repaired unilateral cleft lip and/or palate: an exploratory study using laser scanning. *European Journal of Orthodontics*, 36 (5): 497-505.

Evans, R.A., Harries, M.L., Baguley, D.M., *et al*. (1989) Reliability of the House and Brackmann grading system for facial palsy. *The Journal of Laryngology and Otology,* 103 (11): 1045-1046.

Facial Palsy UK (2002) *Symptoms of facial palsy.* Available from: https://www.facialpalsy.org.uk/inform/what-is-facial-palsy/symptoms-of-facial-palsy/. [Accessed on 8th July 2022].

Farkas, L.G., Bryson, W. and Klotz, J. (1980). Is photogrammetry of the face reliable? *Plastic and Reconstructive Surgery,* 66: 346.

Ferrario, V.F., Sforza, C., Poggio, C.E., *et al.* (1996) Facial three-dimensional morphometry. *American Journal of Orthodontics and Dentofacial Orthopedics, 109: 86-93.*

Frey, M., Jenny, A., Giovanoli, P., *et al.* (1994) Development of a new documentation system for facial movements as a basis for the international registry for neuromuscular reconstruction in the face. *Plastic and Reconstructive Surgery, 93: 1334-1349.*

Fujimura, T. and Hotta, M. (2011) The preliminary study of the relationship between facial movements and wrinkle formation. *Skin and Research Technology,* 18 (2); 219-224.

Galantucci, L., Percoco, G. and Lavecchia, F. (2013) A new three-dimensional photogrammetric face scanner for the morpho-biometric 3D feature extraction applied to a massive field analysis of Italian attractive women. *Procedia CIRP, 5: 259-264.*

Garcia, R.M., Hadlock, T.A., Klebuc, M.J., *et al.* (2015). Contemporary solutions for the treatment of facial nerve paralysis. *Plastic and Reconstructive Surgery,* 135 (6): 1025-1046.

Gattani, S., Ju, X., Gillgrass, T., *et al.* (2020) An Innovative Assessment of the Dynamics of Facial Movements in Surgically Managed Unilateral Cleft Lip and Palate Using 4D Imaging. *Cleft Palate-Craniofacial Journal,* 57 (9): 1125–1133.

Ghassemi, A., Prescher, A., Riediger, D., *et al* (2003) Anatomy of the SMAS revisited. *Aesthetic Plastic Surgery,* 27 (4): 258-264.

Gross, M.M, Trotman, C.A. and Moffatt, K.S. (1996) A comparison of three-dimensional and two-dimensional analyses of facial motion. *The Angle Orthodontist 66: 189-194.*

Guest, E., Berry, E. and Morris, D. (2001) Novel methods for quantifying soft tissue changes after orthognathic surgery. *International Journal of Oral and Maxillofacial Surgery, 30: 484–9.*

Gwilliam, J.R., Cunningham, S.J. and Hutton, T. (2006) Reproducibility of soft tissue landmarks on three-dimensional facial scans. *The European Journal of Orthodontics,* 28 (5): 408-415.

Hajeer, M.Y. (2003*) 3D soft-tissue, 2D hard-tissue and psychosocial changes following orthognathic surgery.* Thesis submitted to the University of Glasgow for the degree of PhD in Orthodontics Faculty

Hajeer, M.Y., Millett, D.T., Ayoub. A.F, *et al.* (2004) Applications of 3D imaging in orthodontics: Part I. *Journal of Orthodontics, 31: 62-70.*

Hallac, R.R., Feng, J., Kane, A.A., *et al.* (2017) Dynamic facial asymmetry in patients with repaired cleft lip using 4D imaging (video stereophotogrammetry). *Journal of Cranio-Maxillofacial Surgery*, 45 (1): 8–12.

Haraguchi, S., Takada, K. and Yasuda Y. (2002) Facial asymmetry in subjects with skeletal Class III deformity. *The Angle Orthodontist,* 72: 28–35.

Hoerter, J.E. & Patel, B.C. (2019) *Anatomy, Head and Neck, Platysma* [online]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK545294/. [Accessed 8th June 2022]

Honrado, C.O. and Larrabee, W.F. (2004) Update in three-dimensional imaging in facial plastic surgery. *Current Opinion in Otolaryngology and Head and Neck Surgery, 12 : 327-331.*

Horta, R., Aguiar, P., Monteiro, D., *et al.* (2014) A facegram for spatial–temporal analysis of facial excursion: Applicability in the microsurgical reanimation of long-standing paralysis and pretransplantation. *Journal of Cranio-Maxillofacial Surgery,* 42 (7): 1250-1259.

House, J.W. (1983) Facial nerve grading systems. *Laryngoscope,* 93: 1056- 1069.

House, J.W. and Brackmann, D.E. (1985) Facial nerve grading system. *Otolaryngology Head and Neck Surgery,* 93: 146–147.

Hu, W., Ross, B. and Nedzelski, H. (2001) Reliability of the Sunnybrook facial grading system by novice users. *Journal of Otolaryngology,* 30 (4): 208-211.

Hur, M.S., Hu, K.S., Park, J.T. *et al.* (2010) New anatomical insight of the levator labii superioris alaeque nasi and the transverse part of the nasalis. *Surgical and Radiologic Anatomy,* 32: 753–756.

Isono, M., Murata, K., Tanaka, H., *et al.* (1996) An objective evaluation method for facial mimic motion*. Otolaryngology Head Neck Surgery,* 114: 27–31.

Johnson, P.C., Brown, H., Kuzon, W.M., *et al.* (1994) Simultaneous quantitation of facial movements: the maximal static response assay of facial nerve function. *Annals of Plastic Surgery,* 32: 171-179.

Johnston, D.J., Millett, D.T., Ayoub, A.F., *et al.* (2003) Are facial expressions reproducible. *Cleft Palate-Craniofacial Journal, 40 (3): 291–296.*

Ju, X., O'Leary, E., Peng, M., *et al.* (2016) Evaluation of the reproducibility of nonverbal facial expressions using a 3D motion capture system. *Cleft Palate-Craniofacial Journal, 53* (1): 22–29.

Kanerva, M., Poussa, T. and Pitkaranta, A. (2006) Sunnybrook and House-Brackmann facial grading systems: intrarater repeatability and interrater agreement. *Otolaryngology Head and Neck Surgery,* 135: 865–871.

Kang, T.S., Vrabec, J.T., Giddings, N., *et al.* (2002) Facial grading systems (1985-2002): beyond the House-Brackmann Scale. *Otology and Neurotology*, 23 (5): 767-771.

Kau, C. and Richmond, S. (2007) Three-dimensional surface acquisition systems for the study of facial morphology and their application to maxillofacial surgery. *International Journal of Medical Robots and Computer Assisted Surgery, 3 (2): 97–110.*

Kau, C.H., Richmond, S., Zhurov, A.I., *et al.* (2005b) Reliability of measuring facial morphology with a 3-dimensional laser scanning system. *American Journal of Orthodontics and Dentofacial Orthopedics, 128 (4): 424–30.*

Kau, C.H., Zhurov, A.I., Knox, J., *et al.* (2005a) Reliability of measuring facial morphology. *American Journal of Orthodontics and Dentofacial Orthopedics,* 128 (3): 424-430.

Kayhan, F.T., Zurakowski, D. and Rauch, S.D. (2000) Toronto Facial Grading System: interobserver reliability. *Otolaryngology Head and Neck Surgery,* 122: 212–215.

Khambay, B.S., Lowney, C.J., Hsung, T.C., *et al.* (2019) Fluctuating asymmetry of dynamic smiles in normal individuals. *International Journal of Oral and Maxillofacial Surgery,* 48 (10): 1372–1379.

Khambay, B.S., Nairn, N., Bell, A., *et al.* (2008) Validation and reproducibility of a high-resolution three-dimensional facial imaging system. *British Journal of Oral and Maxillofacial Surgery,* 46: 27-32.

Komazaki, Y., Fujiwara, T., Kosaki, R., *et al.* (2011) The Methods of Three-Dimensional Anthropometric Measurements for Infants / Toddlers Face with Minor Anomalies. *Pediatric Research, 70: 415.*

Kovacs, L., Zimmermann, A., Brockmann, G., *et al.* (2006) Three-dimensional recording of the human face with a 3D laser scanner. *Plastic and Reconstructive Surgery, 59: 1193-1202.*

Lee, H.Y., Park, M.S., Byun, J.Y., *et al.* (2013) Agreement between the facial nerve grading system 2.0 and the House-Brackmann grading system in patients with Bells Palsy. *Clinical and Experimental Otorhinolaryngology,* 6 (3): 135-139.

Lightoller, G.S. (1925) Facial muscles-the modiolus and muscles surrounding the rima oris with some remarks about the Panniculus Adiposus. *Journal of Anatomy,* 60: 1-85.

Linstrom, C. J., Silverman, C. A. and Colson, D. (2002) Facial motion analysis with a video and computer system after treatment of acoustic neuroma. *Otology & Neurotology,* 23 (4): 572-579.

Lowney, C.J., Hsung, T., Morris, D.O., *et al.* (2018) Quantitative dynamic analysis of the nasolabial complex using 3D motion capture: A normative data set. *Journal of Plastic, Reconstructive and Aesthetic Surgery,* 71 (9): 1332–1345.

Manktelow, R.T., Zuker, R.M. and Tomat, L.R. (2008) Facial Paralysis Measurement with a Handheld Ruler. *Plastic and Reconstructive Surgery, 121: 435-442.*

Martone, A.L. (1962) Anatomy of facial expression and its prosthodontic significance. *The Journal of Prosthetic Dentistry,* 12 (6): 1020-1042.

Meng-Yao, Z., Guo-dong, F. and Zhi-qiang, G. (2008) Facial grading system: physical and psychological impairments to be considered. *Journal of Otolaryngology,* 3 (2): 61-67.

Mercado-Garcıa, J., Rosso, P., Gonzalvez-Garcıa, M., *et al.* (2021) Gummy Smile: Mercado-Rosso Classification System and Dynamic Restructuring with Hyaluronic Acid. *Aesthetic Plastic Surgery,* 45 (3): 1-12.

Mosher, H.D. (1951) The Expression of the Face and Man’s Type of Body as Indicators of His Character. *Laryngoscope,* 61: 1-38.

Murty, G.E., Diver, J.P., Kelly, P.J., *et al.* (1994) The Nottingham system: Objective assessment of facial nerve function in the clinic*. Otolaryngology Head and Neck Surgery,* 110 (2): 156-161.

National Institute for Standards and Technology (2007) *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*. Available from: http://physics.nist.gov/Pubs/guidelines/contents.html. [Accessed on 18th July 2022].

Neely, J.G., Cheung, J.Y., Wood, M., *et al.* (1992) Computerized quantitative dynamic analysis of facial motion in the paralyzed and synkinetic face. *American Journal of Otolaryngology & Head and Neck Surgery,* 13: 97– 107.

Nestor, M.S, Fischer, D. and Arnold, D. (2020) "Masking" our emotions: Botulinium toxin, facial expression, and well-being in the age of COVID-19. *Journal of Cosmetic Dermatology,* 19 (9): 2154-2160.

Özsoy, U., Sekerci, R., Hizay, A., *et al.* (2019) Assessment of reproducibility and reliability of facial expressions using 3D handheld scanner. *Journal of Cranio-Maxillo-Facial Surgery,* 47: 895-901.

Peck, S. & Peck, L. (1995) Selected aspects of the art and science of facial esthetics. *Seminars in Orthodontics,* 1: 105–126.

Peiterson, E. (2002) Bell's palsy: the spontaneous course of 2500 peripheral facial nerve palsies of different etiologies. *Acta Oto-laryngologica,* 549: 4–30.

Popat, H., Henley, E., Richmond, S., *et al.* (2010) A comparison of the reproducibility of verbal and nonverbal facial gestures using three-dimensional motion analysis. *Otolaryngology- Head and Neck Surgery,* 142: 867-872.

Qui, T., Xue, Z., Wu, L., *et al.* (2022) Three-dimensional dynamic analysis of the reproducibility of verbal and nonverbal facial expressions. *The Cleft Palate-Craniofacial Journal,* 59 (1): 22-31.

Reitzen, S.D., Babb, J.S. and Lalwani, A.K. (2009) Significance and reliability of the House-Brackmann grading system for regional facial nerve function. *Otolaryngology Head and Neck Surgery,* 140: 154-158.

Renault, F. and Quijano-Roy, S. (2015*) Chapter 13 - Congenital and Acquired Facial Palsies. Neuromuscular Disorders of Infancy, Childhood, and Adolescence.* 2nd Edition: 225-242. Academic Press.

Ross, B.G. and Nedzelski, J.M. (1998) Reliability and validity of the Sunnybrook Facial Grading System. *The Hague: Kugler,* 563–567.

Ross, B.G., Fradet, G. and Nedzelski, J.M. (1996) Development of a sensitive clinical facial grading system. *Otolaryngology Head and Neck Surgery,* 114 (3): 380-386.

Sarver, D.M. (2001) The face as the determinant of treatment choice. *Craniofacial Growth Series, 38: 19-54.*

Satriyasa, B.K. (2019) Botulinum toxin (Botox) A for reducing the appearance of facial wrinkles: a literature review of clinical use and pharmacological aspect. *Clinical Cosmetic and Investigational Dermatology*, 12: 223-228.

Sawyer, A., See, M. and Nduka, C. (2009) Assessment of the reproducibility of facial expressions with 3-D stereophotogrammetry. *Otolaryngology Head and Neck Surgery,* 140 (1): 76-81.

Scriba, H., Stoeckli, S.J., Veraguth, D., *et al*. (1999) Objective evaluation of normal facial function. *The Annals of Otololgy, Rhinology & Laryngology,* 108: 641–644.

Shujaat, S., Khambay, B.S., Ju, X., et al. (2014) The clinical application of three-dimensional motion capture (4D): A novel approach to quantify the dynamics of facial animations. *International Journal of Oral and Maxillofacial Surgery,* 43 (7): 907–916.

Silva, B.P., Jimenez-Castellanos, E., Martinez-de-Fuentes, R., *et al.* (2013) Laypersons' perceptions of facial and dental asymmetries. *The International Journal of Periodontics & Restorative Dentistry,* 33: 162-171.

Souccar, N.M. & Kau, C.H. (2012) Methods of measuring the three-dimensional Face. *Seminars in Orthodontics, 18 (3): 187-192.*

Stew, B. and Williams, H. (2013) Modern management of facial palsy: a review of current literature. *British Journal of General Practice*, 63 (607): 109-110.

Terzis, J.K. and Konofaos, P. (2012) Experience with 60 adult patients with facial paralysis secondary to tumor extirpation. *Plastic and Reconstructive Surgery,* 130 (1): 51-66.

Toma, A.M., Zhurov, A., Playle, R., *et al.* (2008) Reproducibility of facial soft tissue landmarks on 3D laser-scanned facial images. *Orthodontics and Craniofacial Research*, 12 (1): 33-42.

Trotman, C.A., Faraway, J.J. and Essick, G.K. (2000) Three-dimensional nasolabial displacement during movement in repaired cleft lip and palate patients. *Plastic and Reconstructive Surgery,* 105: 1273–1283.

Trotman, C.A., Phillips, C., Faraway, J.J., *et al.* (2013) Influence of objective three-dimensional measures and movement images on surgeon treatment planning for lip revision surgery. *Cleft Palate-Craniofacial Journal,* 50 (6): 684–695.

Urquhart, C.W, Green, D.S. and Borland, E.D. (2006) *4D capture using passive stereophotogrammetry*. The 3rd European Conference on Visual Media Production - Part of the 2nd Multimedia Conference London: p196.

Vaughan, A., Gardner, D., Miles, A., *et al.* (2020) A systematic review of physical rehabilitation of Facial Palsy. *Frontiers in Neurology* 11 (22): 1-13.

Vrabec, J.T., Backous, D.D., Djalilian, H.R., *et al.* (2009) Facial nerve grading system 2.0. *Otolaryngology Head and Neck Surgery*, 140 (4): 445-450.

Weeden, J.C., Trotman, C.A. and Faraway, J.J. (2001) Three-dimensional analysis of facial movement in normal adults: influence of sex and facial shape. *The Angle Orthodontist*, 71 (2): 132-140.

Winberg, S., Zoss, G., Chandran, P., *et al.* (2022) Facial hair tracking for high fidelity performance capture. *Association for Computing Machinery,* 41 (4): 1-12.

Yen, T.L., Driscoll, C.L.W. and Lalwani, A.K. (2003) Significance of House-Brackmann facial nerve grading global score in the setting of differential facial nerve function. *Otology and Neurotology,* 24: 118–122.

Young, J.Z. (1962) *The life of vertebrates*. New York: Oxford University Press.

Zhai, M., Feng, G. and Gao, Z. (2008) Facial grading system: Physical and psychological impairment to be considered. *Journal of Otolaryngology,* 3 (2): 61-67.

Zhao, Y., Feng, G., Wu, H., *et al.* (2020) Prognostic value of a three-dimensional dynamic quantitative analysis system to measure facial motion in acute facial paralysis patients. *Head Face Medicine***,** 16 (1): 15.

Zhu, S., Yang, Y. and Khambay, B.S. (2017) A study to evaluate the reliability of using two dimensional photographs, three dimensional images and stereoscopic projected three-dimensional images for patient assessment. *International Journal of Oral and Maxillofacial Surgery, 46: 394-400.*

CHAPTER 8

APPENDICES

8.1 Appendix I

Dear Professor Khambay,

Re: "The reproducibility of facial expressions used in the Sunnybrook scale using 4D motion capture technology."
Application for Ethical Review ERN_20-0479

Thank you for your application for ethical review for the above project, which was reviewed by the Science, Technology, Engineering and Mathematics Ethical Review Committee.

On behalf of the Committee, I confirm that this study now has full ethical approval.

I would like to remind you that any substantive changes to the nature of the study as described in the Application for Ethical Review, and/or any adverse events occurring during the study should be promptly brought to the Committee's attention by the Principal Investigator and may necessitate further ethical review.

Please also ensure that the relevant requirements within the University's Code of Practice for Research and the information and guidance provided on the University's ethics webpages (available at https://intranet.birmingham.ac.uk/finance/accounting/Research-Support-Group/Research-Ethics/Links-and-Resources.aspx) are adhered to and referred to in any future applications for ethical review. It is now a requirement on the revised application form (https://intranet.birmingham.ac.uk/finance/accounting/Research-Support-Group/Research-Ethics/Ethical-Review-Forms.aspx) to confirm that this guidance has been consulted and is understood, and that it has been taken into account when completing your application for ethical review.

Please be aware that whilst Health and Safety (H&S) issues may be considered during the ethical review process, you are still required to follow the University's guidance on H&S and to ensure that H&S risk assessments have been carried out as appropriate. For further information about this, please contact your School H&S representative or the University's H&S Unit at healthandsafety@contacts.bham.ac.uk.

Kind regards

Research Ethics Manager
Research Support Group
C Block Dome
Aston Webb Building
University of Birmingham
Edgbaston B15 2TT
Tel: 0121 414 8825
Web: https://intranet.birmingham.ac.uk/finance/RSS/Research-Support-Group/integrity-ethics-governance/Research-Ethics/index.aspx

8.2 Appendix II

UNIVERSITY OF
BIRMINGHAM

Participant Information Sheet

Oro-facial Motion Capture to assess reproducibility of the Sunnybrook Facial Grading System

Version 1.1 27th November 2020

Ethics Reference: ERN_20-0479

You can speak to a member of the research team for more information using the details below

You can change your mind about participating in the study at any time; you do not need to give a reason for your decision.

Enquiries & Correspondence:

The Chief Investigator of this study is Professor Balvinder Khambay

If you want to discuss this study further, please call 0121 466 5522 or email
[redacted]

Professor Balvinder Khambay
Birmingham Dental Hospital
5 Mill Pool Way
Birmingham
B5 7EG

Thank you for considering volunteering.

We do a lot of research at the University of Birmingham that needs volunteers to look at how the mouth, and face moves. We need all types of volunteers to be able to compare differences between them.

Your decision to participate is completely voluntary; volunteering will not affect any care you may receive in Birmingham Dental Hospital in any way. If you decide to withdraw consent we cannot make changes to images already taken. Just tell either researcher or email us (see bottom of page). You have 2 weeks are the rating session to contact the researcher and ask them to remove your data, after this time it will not be possible.

Information

1. Introduction

We would like to invite you to take part in a research study. You are free to decide whether or not to take part. Please take time to read the following information carefully.

2. What is the research about?

The purpose of this research project is to assess whether you move your face the same way each time you make a facial expression, such as a smile or a frown. We hope this research will add information to the current methods of assessing the amount of facial nerve paralysis individuals have pre and post treatment. The project will use motion capture cameras to assess these facial expressions, both at rest and in motion, over two different time points at least 3 months apart.

3. Why have we been asked?

We are looking for healthy volunteers aged 18 - 60 to assess how they move their faces.

4. What will happen if I decide to take part?

We want to give you time to consider the information in this leaflet. We will contact you at least two weeks after you receive the leaflet (providing you have given us permission to do so) to see if you wish to take part. If you make a decision sooner than two weeks from receiving this leaflet, you can contact us, using the details on the front of the leaflet, to let us know.

If you decide you would like to participate you will be seen at the Birmingham Dental Hospital and sign the consent form and have your facial image recorded. You will be shown a video of which facial expressions you will be asked to perform. You will be given time to practice before making the seven facial expressions in front of the camera. This will take around 10 minutes in total. You will be asked to return 3 months later to repeat the facial expressions.

5. What will happen with the videos?

The original videos will be stored securely on our research servers at the University of Birmingham. Using specialised software we will use the videos to make a computerised animation of your facial movements, which would provide mathematical data points for our research. Your image would not be identifiable on this animation as it does not have colour or texture. These images will be labelled with a code rather than your name to protect their identity.

6. What are the advantages of taking part?

You will be helping with research which may lead to better treatment for patients with facial palsy in the future.

7. What are the disadvantages of taking part?

You would need to give up some of your time. There are no known risks to taking part in the study.

8. Has this research been reviewed?

This project has been reviewed and approved by the University of Birmingham's Research Ethics Committee.

9. How can I obtain more information about this study?

You can speak to the researcher when they contact you or if you attend an appointment. Alternatively you can contact the research team using the details on the cover of this information sheet.

8.3 Appendix III

CONSENT FORM

Please initial each box you agree with and sign and date the form at the bottom.

		Initials
1	I confirm that I have read the information sheet dated 27th November 2020 (version 1.1) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.	
2	I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.	
3	I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason, without my medical care or legal rights being affected.	
4	The data from this research might be useful to future research; by initialling this box you give us permission to use it as part of other ethically approved research projects. These projects may be in collaboration with other institutes outside the university, but all the information used will be anonymised by research teams at the University of Birmingham.	
5	I agree to be contacted by the researcher to seek permission to publish my facial image in a scientific journal.	
6	I agree to take part in the above study.	

________________________	________________	_______________
Name of participant	Signed	Date

I have discussed this study with this participant who has agreed to give informed consent.

________________________	________________	_______________
Name of person consent	Signed	Date taking