University of Alberta

Using Two- and Three-Dimensional Kinematic Analysis to Compare Functional Outcomes in Patients who have Undergone Facial Reanimation Surgery

by

Lisa Dunwald

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Masters of Science

Speech-Language Pathology and Audiology

Fall 2010
Edmonton, Alberta

Examining Committee

Carol Boliek, Speech-Language Pathology and Audiology

Jonathan Norton, Surgery

Jana Rieger, Speech-Language Pathology and Audiology

Abstract

The current study was designed to: (1) compare the sensitivity of a 2-dimensional video-based system with a 3-dimensional optical system, and (2) investigate movement on the affected and unaffected side of the face during the production of various functional movement tasks in 5 patients who had undergone facial reanimation surgery. The study showed that: (1) distance is the most valuable measure for evaluating facial paralysis, regardless of system; (2) movements associated with maximal contraction and running speech tasks are most informative when assessing facial paralysis; (3) area and volume ratios may be an appropriate measure for tracking changes in facial movement over time; (4) velocity and acceleration measures provide minimal information regarding facial movement; and (5) 2-dimensional analysis is most effective when distance is measured during maximal contraction and running speech tasks. Both systems were effective in tracking small movements of the face, but the 3-dimensional system was superior overall.

Table of Contents

List of Tables

Table Page

List of Figures

Figure Page

INTRODUCTION

Introduction

The incidence of residual facial paralysis is estimated to be approximately 127, 000 cases annually in the United States. Idiopathic types, such as Bell's palsy, account for approximately half of these. Infections (e.g. otitis media, Lyme disease), neoplastic etiologies (e.g. acoustic neuromas), neurologic etiologies (e.g. Guillain-Barre syndrome, cerebral vascular accidents), and traumatic injuries (e.g. during birth, temporal bone fractures) make up 15.3 percent, 13.5 percent, 13.5 percent and 8.2 percent of cases; respectively (Bleicher, Hamiel, & Gengler, 1996). Facial paralysis can result from metabolic, toxic, or iatrogenic etiologies as well.

The loss of facial nerve function creates a serious disfigurement as well as numerous functional defects. Facial paralysis is the loss or impairment of motor function of facial muscles due to damage of the facial nerve (Cranial Nerve VII), brainstem nuclei of the facial nerve, and/or the neuromuscular system innervated by this nerve. The facial nerve is formed by motor and sensory roots which exit the brain at the cerebellopontine angle to innervate the tissues of the face and neck. The branchial motor branch of the facial nerve is responsible for innervating the muscles of facial expression (e.g., risorius, platysma, and buccinator), the sphincter muscles of the eye (orbicularis oculi), the sphincter muscles of the mouth (orbicularis oris), stylohyoid, and the posterior belly of the digastric muscle. The visceral motor branch controls special functions such as lacrimation and salivation. Facial paralysis can occur from a lesion or lesions to

the central or peripheral nervous system. Unilateral upper motor neuron lesions of the facial nerve cause paralysis in the contralateral lower half of the face. Bilateral upper motor neuron lesions can result in bilateral facial paralysis. Unilateral facial paralysis is the result of damage to the facial nucleus or the peripheral nerve. The degree of paralysis ranging from minor weakness to complete paralysis depends on the site and extent of the lesion.

Issues resulting from facial paralysis include those associated with: (a) somatic function, (b) social function, (c) psychological wellbeing, and (d) physiological function (Devriese, 1998). Physical symptoms of facial paralysis in the upper face include the inability to close the eyes and sagging lower eyelid. Tears may run from the affected eye, or alternatively, tearing may be reduced. Failure of the eyelids to approximate, combined with the loss of the blink reflex may cause irritation, exposure keratopathy, corneal ulcerations, and blindness (Tate & Tollefson, 2006). The sagging of the corner of the mouth is the most notable characteristic in the lower half of the face. The lack of control over lip closure can result in difficulties with eating, drinking, speech (especially with sounds that require the lips such as /p/ and /b/) and control of drooling. The involvement of the chorda tympani, a branch of the facial nerve, impairs taste sensation. Also, the nostril on the affected side may collapse, causing nasal obstruction (Schaitkin & May, 2000). Additionally, patients with facial paralysis associated with inflammation, as in the case with Bell's palsy and herpes zoster oticus, may have pain in the mastoid process, neck, ear, or the face (Schaitkin & May, 2000). Patients with facial nerve palsy will develop synkinesis which is one

sequelae of facial paralysis (Nakamura, Toda, Sakamaki, Kashima, & Takeda, 2003). Common examples of synkinesis include eye closure with volitional contraction of mouth muscles and involuntary movement of the mouth during eye closure. These abnormal movement muscle contractions have massive cosmetic and functional implications.

One of the most socially devastating effects of facial paralysis is the inability to produce facial expression of emotion. Goldblatt & Williams (1986) have reported that the lack of facial reanimation leads to social perceptions of lower intelligence. Patients commonly report personal and work-related problems as well as limited social integration and interpersonal communication (Goldberg, DeLorie, Zuker, & Manktelow, 2003; McGregor, 1990; Verzijl, van der Zwaag, Cruysberg, & Padberg, 2003). These problems stem from facial disfigurement and difficulties eating, drinking and communicating effectively in social settings. Consequently, people with facial paralysis often are introverted and can become isolated (Gillberg, 1992). The face is central to characterizing an individual as a unique entity and is essential for interaction with those around us. Static and dynamic movements of the face are crucial to conveying conversational and emotional messages used daily. The face also is a crucial component of beauty, sexual attractiveness, and sexual interest (Ekman, 1986). Therefore, patients with facial paralysis experience pronounced psychological distress. Oftentimes, they must cope with feelings of shame, decreased self-esteem, anxiety, depression, guilt, anger, and/or fear (Devriese, 1998; Ekman, 1986, MacGregor, 1990; Twerski & Twerski, 1986). They also experience fatigue from managing social

interaction as well as the reactions from others (MacGregor, 1990). People with facial paralysis can suffer from first order psychological distress resulting in impaired psychosocial functioning and secondary social distress (VanSwearingen, Cohn, Turnbull, Mrzai, & Johnson, 1998).

Nutritional intake also may be impaired when facial paralysis exists. Routine and seemingly easy tasks such as eating can be quite challenging for patients with facial paralysis, who often demonstrate swallowing difficulties that occur in both the oral and/or pharyngeal phases of deglutition. More specifically, a survey conducted by Sjogreen and colleagues (Sjogreen, Andersson-Norinder, & Jacobsson, 2001) revealed that this group had difficulties in the oral phase including "getting food off spoon with lips," "food and liquid leak out of the corners of mouth," and "takes a long time to swallow bites of food." "Choking on food" and "coughs when receiving liquid" were some of the problems in the pharyngeal phase of the swallow. Another study by Secil, Aydogdu, and Ertekin (2002) showed that 79% of their patients with Bell's palsy had difficulty controlling the bolus and 39% had a decreased ability to taste (hypoagusia or agusia). They also discovered that there was decreased tactile sensation of the tongue and the mucosa of the cheek on the affected side (intraoral hypaesthesia). Impairment of the motor, sensory, taste and parasympathetic innervation interact to produce the swallowing disorder (Secil et al., 2002). Problems in the oral phase or the compensations used in this stage also may cause swallowing disorders by disrupting the later, more automatic swallowing processes (de Swart,Verheij, & Beurskens, 2003).

Verbal communication also may be impaired in this population of patients. Bilabial incompetence produces characteristic speech error patterns including substitutions, distortions, and omissions of bilabial sounds (/p/, /b/, /m/) and the alveolar sounds (/t/, /d/, /n/). Goldberg and colleagues (2003) noted that these patients have to compensate for /f/, /v/, /s/, /sh/, and /w/ sounds as well. Generally, the parameters of manner and voice are maintained while the place of articulation is compromised. (Goldberg, et al., 2003). People with facial paralysis often compensate for these sounds by varying the placement of the tongue in relationship to teeth to approximate the target sound. Furthermore, in patients with Moebius syndrome, where there may be multiple cranial nerve involvement, intelligibility of vowels is compromised due to the addition of velopharyngeal incompetence. Weakness or paralysis of the velopharynx introduces a hypernasal voice quality combined with increased glottal substitutions.

Treatment

Several treatment options are available to patients with facial paralysis. Generally, these procedures can be divided into two broad categories including static (cosmetic) reanimation and dynamic reanimation procedures. The choice of procedure is dependent on patient objectives, age, underlying medical condition, tolerance to undergo several surgical procedures, and potential for spontaneous recovery. Static reanimation procedures are used to restore symmetry to the face. The most popular static procedures include gold weighting the eyelids, brow and forehead lifts, alloplastic slings to make the face more symmetrical, and the use of

botulinum toxin to decrease hyperkinesis in the face. The use of corticosteroids treatments (Ramsey, DerSimonian, Holtel, & Burgess, 2000), facial neuromuscular re-education (Brach, VanSwearingen, Lennert, & Johnson, 1997), selective myectomy (Guerrissi, 1991), biofeedback rehabilitation (Nakamura et al., 2003) or mime therapy (Beurskens & Heymans, 2003) also have been noted as alternative static procedures in the literature. Static reanimation procedures are recommended to patients who are not interested in restoring facial motion or who are at high operative risks. Often, older patients with unilateral paralysis who have a noticeable drooping on the affected side of the face opt for these procedures (Harrison, 2005). The arbitrary age cutoff is approximately 55 years for determining whether static procedures are indicated. There is some evidence to suggest that nerve regeneration and functional results are poorer in older patients (Streppel, Angelov, Gutinas-Lichius, & Neiss, 1998, Guntinas-Lichius, Streppel, & Stennert, 2006).

Dynamic reanimation surgery is recommended for younger individuals with facial paralysis. The long-term goal for dynamic facial reanimation is to improve symmetry of the face at rest and restore volitional movement and expression of emotion to the lower face (retraction of the lips). Surgical options for these individuals are influenced by the amount of atrophy present in the face (i.e., the time since damage to the nerve and the degree of damage) and the possibility for spontaneous reinnervation. When paralysis has existed for longer than 6 months, surgeons recommend neural reinnervation procedures such as neurorrhaphy (direct repair of the nerve), interposition nerve grafting, or

crossover reinnervation. With paralysis longer than one year in duration, muscle transpositions and microneurovascular free flap muscle transfers are recommended because once a muscle has been denervated for a year, the mimetic musculature atrophies and its fibrosis becomes permanent (Guntinas-Lichius et al., 2006). Transferred muscle can be used for animation and also to add bulk to areas of the face that have atrophied. The choice of dynamic reanimation procedure depends on the status of the distal facial nerve fibers and motor endplates. If the distal nerve fibres are available, then reinnervation can be achieved through interposition grafts or crossover reinnervation. If the motor endplates are no longer viable due to long-term de-innervation, then a muscle transposition or a microneurovascular free muscle flap transfer is required (Shindo, 1999).

Interposition nerve grafting is attempted if the proximal and distal ends of the nerve are available. The great auricular nerve, sural nerve, and the ansa hypoglossi are commonly used as donors in this procedure. In the case where the distal facial nerve end is viable and the proximal facial nerve is not viable, crossover reinnervation is indicated (the distal end of the facial nerve is anastomosed to a donor nerve). The hypoglossal nerve, great auricular nerve, and the buccal branch of the contralateral facial nerve are generally used as donors in crossover reinnervation procedures. Of these donors, the hypoglossal nerve is the most commonly used. Although this procedure results in speech and swallowing difficulties, the similar cortical mapping of the facial and hypoglossal nerve allows facial movements to be easily after recovery (Danner, 2008). Despite the

"success" of crossover reinnervation, interposition nerve grafting should be used whenever possible since crossover innervation involves the sacrifice of a cranial nerve. Furthermore, spontaneous expression is lost after crossover reinnervation because the muscle is no longer controlled by the appropriate motor cortex (Shindo, 1999).

Muscle transpositions and microneurovascular free-muscle transfers are the procedures of choice for long-standing facial paralysis. Muscle transpositions are necessary when: 1) distal facial nerve fibre is not viable, 2) the motor endplates are not viable, 3) there is a soft tissue defect in the cheek and 4) a neuropathy or a crossover reinnervation procedure fails (Shindo, 1999). The temporalis muscle is used commonly in the procedure because of its "location, contractility, and vector of pull to the corner of the mouth to create a lateral smile" (May & Schaitkin, 2000, p.635). Not only do free flaps provide soft tissue augmentation, they have a definite advantage over muscle transpositions as they can be reinnervated by the proximal stump of the nerve to allow for spontaneous expression. If the proximal end of the nerve is unavailable, an interposition nerve graft can be used to connect the flap to a nerve source. The gracilis muscle and the inferior rectus abdominis muscle also are widely harvested as flaps as they provide adequate lip elevation, have reliable vascular and nerve anatomy, and cause little damage to the donor site. Also, because the flap is not restricted by a pedicle and can be resized and reshaped, it can be customized to the patient's needs.

Outcome

There are several studies that use patient report criteria as a method to determine functional outcome after facial reanimation surgery. O'Brien et al. (1990) reported that 68% of the 69 participants who underwent microneurovascular free-muscle transfers had good or excellent outcomes. Fifty-three percent indicated that they had improved speech and 64% believed that their eating had improved since the surgery. In another study by Yla¨-Kotola, Kauhanen, and Asko-Seljavaara (2004), 78% of the 40 participants reported satisfaction with the outcome of the surgery. Patient-report criteria can be useful to determine the outcome of a treatment. However, patient satisfaction can be influenced by factors other than treatment. Furthermore, a successful outcome may be judged differently by the physician than by the patient (Ikeda, Nakazato, Hiroshige, Abiko, & Sugiura, 2003). Although the value of patient report cannot be diminished, additional methods may be useful in the evaluation of functional outcomes.

Studies that investigate functional outcome in patients are extremely scarce. Goldberg et al. (2003) used standardized speech protocols to assess patients' bilabial competence, screened for compensatory phonemes to evaluate articulation, and completed a subjective rating to determine intelligibility. After evaluating videotapes from pre- and post-surgery, they report that patients with Moebius syndrome who underwent facial reanimation surgery had "decreased frequency of compensatory phonemes after surgery, improved bilabial competence, and significantly improved intelligibility." (p.689). Post-surgery, the

percentage of patients with bilabial incompetency decreased from 75%to 25%. Eighty-three percent of the subjects were judged to have improved intelligibility. However, this study was based on the subjective judgment of a single unblinded observer. Furthermore, the study only examined the outcomes of facial reanimation surgery in patients with Moebius syndrome; therefore, generalization is limited. There is no known research that probes the outcome of facial reanimation surgery in other populations. Moreover, there is no systematic documentation of changes in speech, feeding or swallowing in this patient population

Methods of Assessment

Ideally, a facial nerve grading system should: (1) be universal and reproducible with high inter-observer reliability, (2) have measures for static and dynamic facial muscle function, (3) recognize secondary defects of facial nerve dysfunction, (4) be cost-effective and easy to use (5) be continuous (as opposed to discrete), and (6) be sensitive to change across time (Kang, Vrabec, Giddings, & Terris, 2002; Ross, Fradet, & Nedzelski,1996; Sargent, Fadhli, & Randall, 1998).

The development of such a system has been hindered by the inherent complexity of facial physiology provides another obstacle to the development of a universal measurement system. Damage to the facial nerve can result in varying degrees of paralysis or paresis to any one or combination of functional units in the face. "Secondary defects," such as synkinesis, contractures, and spasms, are aberrant facial movements which are not a direct result of facial nerve damage

(i.e. muscle weakness). These abnormal facial movements add another obstacle the development of a measurement system.

Although an "ideal" system has not yet been developed, a large number of methods have been proposed to assess facial movement. Existing measurement systems can be broadly classified into subjective and objective measurement systems. Refer to *Appendix A* for operational definitions.

Subjective Measurement Systems

Traditionally, observation-based subjective grading systems have been employed to assess facial nerve dysfunction. Subjective assessments of facial nerve function are accomplished by inferring nerve function through the observation of voluntary facial movement. Maximal eye closure, puckering, smiling, and frowning are commonly used in the assessment of facial movement. These methods usually yield a single value which hypothetically corresponds to the severity of facial paralysis.

The most recognized of these subjective grading systems is the House-Brackmann grading scale (HBGS) (House & Brackmann, 1985). In 1984, the HBGS was adopted by the American Academy of Otolaryngology- Head and Neck Surgery as the North American standard for evaluating facial nerve palsy and it remains the clinical gold standard to this day (Kang et al., 2002). HBGS is a gross system that assigns a single grade to evaluate the degree of facial nerve palsy and secondary defects at once. The HBGS is a six-point grading scale ranging from I (normal) to VI (no movement). These grades are arbitrary numbers which cannot be mathematically manipulated. One of the biggest

advantages of the HBGS is that it yields a single numerical value that describes the severity of the facial nerve dysfunction. This advantage is one of its biggest drawbacks. A global number cannot account for the finer grades of dysfunction, such as regional differences within the face, synkinesis, compensatory movement, and other secondary defects (e.g. contractures and hemispasms) (Linstrom, Silverman, & Colson, 2002). Changes in the paralytic face that can be detected by the naked eye are not differentially described using a gross scaling system (Neely, Joaquin, Kohn, & Cheung, 1996). Not only is the HBGS unable to distinguish finer grades of dysfunction, it puts very little emphasis on secondary defects. Kang et al. (2002) also report that the HBGS has low inter-observer reliability, especially in the middle ranges of nerve dysfunction. The Sunnybrook Scale (Ross et al., 1996) was developed to overcome the shortcomings of the HBGS. It was designed to distinguish finer levels of facial nerve function before and after treatments for facial paralysis by increasing intermediate grades. This scale yields a single composite score that is calculated by adding 3 scores: the resting symmetry score, voluntary movement score, and synkinesis score. Since the Sunnybrook Scale is based on the HBGS, it continues to be prone to observer disagreement and is unable to capture incremental changes of facial movement.

Objective Measurement Systems

To date, there are no subjective assessment methods that can reliably produce quantitative information regarding facial nerve function. Objective assessment methods generate quantitative data to describe facial nerve function. Over the last 20 years, researchers and clinicians have proposed a wide range of

methods to obtain objective data to quantify and determine the severity of facial paralysis. Some have suggested the use of calipers and rulers to take physical measurements of facial movement whereas others have relied on computer analysis. The incorporation of computers into the field of motion analysis has allowed for the development of innovative procedures to measure facial movement in two- and three-dimensions using photography and videography. Despite these developments in motion tracking, no quantitative method for objective description of facial nerve function is universally accepted as the gold standard.

Physical Measurement

One of the simplest ways to obtain an objective assessment is through physical measurement. These assessments are dependent upon the use of a measurement device such as a ruler or a caliper to measure the displacement of features on the face from repose to maximal contraction. Mantkelow, Zuker, and Tomat (2008) have endorsed the use of a hand-held ruler as a simple, accurate, and reliable method for measuring facial movement. Additionally, Frey and colleagues (Frey, Jenny, Giovanoli, & Stussi, 1994) have recommended the use of a digital caliper for more reliable and valid measurements of displacement.

The Burres-Fisch system (Burres & Fisch, 1986) was one of the first methods based on linear measurements to develop an index to describe facial nerve function. The linear measurement index is calculated using the percent displacement of various anatomical landmarks during facial movement. A comparison of the Burres-Fisch system to the criteria of an ideal measurement

system reveals all of its limitations. Even though the Burres-Fisch system is capable of capturing static facial movement (which subjective systems are unable to do), it does not include the assessment of secondary defects nor are the suggested anatomical landmarks consistently reproducible (Frey, Jenny, Giovanoli, & Stussi, 1994). In addition to this problem, the calculation of the linear measurement index can be complicated, time-consuming and the assessment of simultaneous facial movement over more than one region is not possible. All of these characteristics make the Burres-Fisch system clinically undesirable.

The Nottingham system (Murty, et al., 1994) is a modified version of the Burres-Fisch system developed to overcome the shortcomings previously outlined. The displacement from rest to maximum excursion is measured for two distances bilaterally (supraorbital point to infraorbital point, lateral canthus to angle of mouth). The total amount of movement between each side of the face is calculated and is expressed as a percentage of the opposite side. The absence or presence of hemifacial spasm, synkinesis, or contractures as well as gustatory tears, dry eyes, or dysguesia is noted. The Nottingham system proves to be more informative than the Burres-Fisch system as it incorporates secondary defects and is more clinically useful as it can be scored within three minutes. However, one of the limitations of the Nottingham system is its inability to accurately assess bilateral facial nerve dysfunction since the final score derived is a ratio of the movement in the affected and contralateral side. This means that the Nottingham system cannot give an accurate representation of either nerve. Furthermore, the

secondary defects which are incorporated into this newer version do not contribute to the overall numerical score to quantify the severity of the paralysis, and can only be used as a descriptor.

Computer-Assisted Analysis in Two-Dimensional Space

The advancement in technology has lead to increasingly sophisticated measurement systems. Compared to subjective observation-based systems, computer-assisted quantitative measurement systems have the advantage of producing accurate and reproducible measures while tracking multiple points on the face simultaneously. Researchers have taken two main approaches to tracking facial motion: (1) taking linear measurements of pre-determined anatomic landmarks and (2) capturing movement that is distributed globally across the face. These approaches have been applied to both digitized photographs and video recordings.

Photographs

In 1994, a method based on making linear measurements using photographs was developed. The Maximal Static Response Assay (MSRA) was established as a tool to assess simultaneous movement in multiple clinically relevant zones of the face, track motor recovery, determine the direction of the muscle movement, and detect synkinetic muscle movement (Johnson et al., 1994). Adhesive markers are place on pre-determined anatomic landmarks and a 2cm ruler is attached vertically to the nasal tip to calibrate displacement during the analysis. A final dot is placed behind and partially above the nasal ruler. This dot is used as a reference point to determine whether the head position has changed

between photographs (the horizontal and vertical position of the dot in relation to the ruler must be maintained). A baseline photograph is taken when the face is at rest and is looking straight forward. A subsequent photograph is taken when the required voluntary movement has reached its maximal displacement. The photographs are processed using a digitizer board, hand-held puck, and a projector. The projector is used to project the image onto the digitizer board. The projector is used to superimpose a grid onto the photograph. The centers of the markers are manually selected in the frames corresponding to baseline and maximal facial response. Linear *x* and *y* displacements of each marker are calculated by subtracting the maximal displacement position from their baseline position. Because each facial movement is tested by region with the remainder of the face relaxed, when the data are represented as a bar graph comparing the displacement of each landmark bilaterally for each task, facial nerve dysfunction can be detected (i.e., more movement will be observed on the unaffected side compared to the affected side). Also, information regarding expected movement of each landmark can be used to infer facial nerve function. For example, movement of the mouth may induce movement in the musculature around the eye, which may be interpreted as synkinesis. Similarly, the absence of movement can be explained as paralysis.

The MSRA was one of the first attempts made to capture quantitative information of facial movement by tracking multiple points on the face simultaneously. As with most early models, it has several limitations: (1) error can be introduced into measurement as the markers must be identified in multiple

frames, (2) the method is labour-intensive and time-consuming, (3) motion information is only in two-dimensions, (4) the physical markers present on the face may interfere with spontaneous facial movement, and (5) the method does not provide dynamic information.

The MSRA also has been used with video sequences rather than with photos (Bajaj-Luthra, Mueller, & Johnson, 1997; Bajaj-Luthra, VanSwearingen, Thornton, & Johnson,1998; Johnson, Bajaj-Luthra, Llull, & Johnson, 1997). As expected, video-based MSRA faces the same limitations as when photographs are used.

Video

In addition to video-based MSRA, numerous other methods that involve video analysis exist. Linstrom (2002) has used a video-based system using linear measurements between various anatomic landmarks to assess facial movement. In his study, each participant was video-taped while performing voluntary movements with reflective markers adhered to points of interest on the face. Once the video was digitized, *Peak Motus Motion System*, a video-computer interface, was used to coordinate 2D data. This software allows the user to track points of interest by manually identifying each marker by clicking on its center with a mouse. Once all of the markers have been identified, the software automatically tracks all of the points. To remove any confounds introduced by head movement, the static point on the nose is designated as (0,0) for each picture on a Cartesian coordinate system because the static point acts as a reference point to the moving structures. Linstrom (2002) used the displacement

of each marker to compare the side-to-side displacement to the total amount of displacement to derive a percent asymmetry score during eye closure or making a close-lipped smile. In this method, synkinesis is detected by measuring marker displacement at the remote site on the affected side and then comparing it to the amount of movement seen in the normal population. For example, if the participant smiles, then the movement at the remote site would be at the upper eyelid. The distance the upper eyelid moves is compared to the distance the upper eyelid would move in a normal subject. Synkinesis is deemed to be present if the displacement at the remote site exceeds the 95^{th} percentile, based on normal subjects. In facial movements that require eye closure, the oral commissure is considered to be the remote site.

Isono, Murata, Kawamoto, and Azuma (1996) used a video editing software program and a personal computer to determine the trajectory of various facial movements and their distances. White markers are attached to the pre-determined landmarks on the face. The white markers are necessary to ensure that only the markers are extracted after *thresholding*. In this method, ten frames in a movement from repose to maximum movement are selected to determine the trajectory during various facial expressions. The image-processing software employed is capable of taking the extracted markers from each frame and plotting them on a coordinate system which is centered around the static point on the nasal apex. In this way, the movement of all markers during a voluntary facial movement can be visualized on a coordinate system. The scoring system compares the ratio of mean facial movement on the left- and right-side of the face.

These ratios can be represented graphically to track recovery over time. The biggest advantage of this system is that the instrumentation is inexpensive and is all commercially available. However, it has been criticized to be "labour intensive, more expensive, and dependent on the skill of the practitioner in both applying the marks and keeping a uniform system" (Schaitken & May, 2000, p.279).

Automated Facial Analysis (AFA) is a recently-employed linear measurement method to quantify facial motion using computer vision approaches (Rogers et al., 2007; Wachtman, Cohn, VanSwearingen, & Manders, 2001). This program comes with several modules for *dense flow extraction*, *feature-based tracking*, and *high-gradient component detection* to detect, extract, and recognize facial movement using in conjunction with *optical flow* techniques. With these modules, physical markers are not required to track features on the face. Once the video sequences of the subjects performing the voluntary facial movements are digitized, the user manually marks the features to be tracked with the computer mouse. After the initial frame of the image sequence is analyzed, the program is able to track all of the features automatically. This method has several advantages over previous linear measurement methodologies. Because AFA does not required physical markers to be placed on the face, facial movement is not impeded. Furthermore, because the operator is able to designate his/her own markers, feature analysis is not limited to pre-selected points on the face.

Like AFA, *subtraction techniques* are not limited to certain features on the face, but some researchers believe that a more comprehensive examination of

facial movement is possible with these techniques as they do not rely on any markers (Neely et al., 1992; Meier-Gallati, Scriba, & Fisch,1998). Instead of markers, these techniques take advantage of differences in grey-scale intensities between adjacent frames of a video sequence. Neely and colleagues (Neely, Cheung, Wood, Byers, & Rogerson, 1992) developed a pixel subtraction method that relied on surface areas rather than dynamic points on the face to measure the degree of facial surface deformation. The authors videotaped the subjects with a single camera under controlled conditions of lighting, head stabilization, and camera position while engaging them in specific voluntary facial expressions. During the analysis, each pixel in an image was assigned a value depending on its light intensity (i.e., a pixel with greater light intensity would be assigned a bigger value). The computer system employed allowed the subtraction of the pixels in the image of the face in motion from an image of the face at rest. *Thresholding* was used to maximize the difference between the areas of the face that had changed and the areas that had remained the same. For each area of interest (lower, mid, and upper face), all of the white pixels were summed for the frame of the face at rest and in motion (the subtraction pair) for each pair of captured images (the left and the right side). The number of white pixels against the subtraction pairs were plotted to create a computer-generated intensity-duration curve that is used to compare the paretic and unaffected side of the face.

In 1998, Meier-Gallati and colleagues also presented a similar method. However, rather than subtracting pixels, they determined changes in the face by subtracting luminance values associated the face at rest from subsequent images

of the face in motion. The changes in luminance values were calculated via thresholding to produce an image characterized by binary data. These images are then divided in to three areas: the mouth, eyes, and forehead. Areas where movement is present are represented in white while areas that show no movement remain dark. For each area, a value for percent change is calculated and weighted. The overall clinical grading of facial function is determined by adding together the amount of weighted percent change in each area.

Both of these global assessment techniques claim that objective examination for the evaluation of facial paralysis is possible and show strong correlations with the House-Brackmann grading system. However, a substantial amount of time must be dedicated to digitizing and processing the videos (Hontanilla & Auba, 2008). Also, maintaining consistency in the processing procedures over many patients is difficult (Isono et al., 1996). Moreover, the software used in the Neely et al. (1998) study has not been disseminated to the public (Sargent et al., 1998).

Computer-Assisted Analysis in Three-Dimensional Space

Recently, there has been a shift in perfecting 3D measurement systems. Gross et al. (1996) found that 2D video analysis systems underestimate the 3D amplitudes by as much as 43% for some markers. Because important information and accuracy is lost when 3D data are projected in two-dimensions, the importance of a 3D analysis system has been stressed (Frey et al., 1994). The 3D study of 2D video has become a possibility with advancements in the field of

image analysis. Video-analysis and optical tracking have been the two most popular methods for tracking motion in the face.

Video

Several studies have explored 3D computer analysis utilizing various computer software systems and digital video cameras (Frey et al., 1994; Mishima, Yamada, Fujiwara, & Sugahara, 2004; Trotman, Stolhler, & Johnston, 1997). 3D video-analysis requires the extraction of an additional spatial dimension from a 2D image. Linear measurement systems usually accomplish this by using multiple cameras strategically placed at different angles from the participants face. The investigator uses 2D information captured by each camera to create a 3D picture. Computer vision techniques can also be used to track facial movements in three dimensions.

A group of researchers from Vienna were among the first to experiment with 3D video analysis (Frey et al., 1994). To provide a comprehensive evaluation of the face, they combine a qualitative assessment based on the static function of the face augmented by quantitative information based on the excursion of numerous points on the face. They use four cameras to trace the reflective markers on the face from four different locations. A calibration grid is placed in the room at the beginning of each session to calibrate the system. The reflective markers are attached to the participants' face and they are asked to perform a series of voluntary actions. Once the videos are digitized, the VICON system is able to track the centers of the reflective markers during each facial movement from the digitized video. By using the static markers as reference

points, the system is able to calculate 3D coordinates from the 2D data captured by the cameras by using a *direct linear transformation*.

Several years later, the same researchers describe another system involving a digital video camera, an optically precise custom mirror system, a complex calibration device, and the VICON system for 3D motion analysis (Frey, Giovanoli, Gerber, Slamecza, & Stussi, 1999). In this sophisticated procedure, two mirrors are place on either side of the subject to reflect the profile view of the face. A single digital camera captures the subject performing a set of facial movements that is videotaped in standardized position, lighting, room, and time. The video sequence that are digitized and edited are analyzed by the VICON system to determine the 3D coordinates of the pre-defined points on the face and calculates each of their trajectories in two- and three- dimensions. Displacement in the z-axis is determined by facial movements reflected in the mirrors. The operator marks each landmark manually for every frame in a consistent order as indicated by the program. The 3D analysis can provide visualization of precise coordinates and can calculate the distance and velocity of each trajectory. This method has been used clinically to assess and quantify the degree of asymmetry of patients with facial asymmetry before and after surgery (Giovanoli, Tzou, Ploner, & Frey, 2003). The technique provides accurate 3D analysis of the movement of selected facial landmarks and provides useful graphic representations of these movements. However, reports on the accuracy of the system were limited to static measurements only and did not include the accuracy information on dynamic measurements. Frey et al. (1999) listed the advantages of

their system as follows: (1) taking the video does not interfere with facial motion, (2) there is little or no fatigue associated with producing movements of the patients, and (3) an international registry may be more easily recognized. Despite these advantages, the analysis and report generation is complicated and takes hours to perform, making immediate client follow-up impossible (Mantkelow et al., 2008). Furthermore, the complex data analysis can only be carried out in Vienna and thus limits its clinical feasibility (Hontanilla & Auba, 2008).

Because the two analysis methods applied to the VICON system are so demanding of the investigator, other research groups have focused on software systems with increased automation. Trotman et al. (1997) describe a measuring system involving four cameras, reflective markers, special lighting, and *Motion Analysis*, a commercially-available motion analysis system. Their methodology has been used to measure facial movement in healthy subjects (Trotman et al. 1997) and in repaired cleft-lip and palate patients (Trotman, Faraway, & Essick, 2000). This same motion analysis system is also used by Coulson, Croxson, and Gilleard (1999, 2000).

Mishima and colleagues (2004) also have developed an automated video-analysis system employing computer vision techniques. This system is similar to the AFA system described earlier. The features that are designated for the first frame in the image sequence are automatically tracked through an *optical flow technique, a template-matching technique*, and the *Lucas-Kanade algorithm*. The 3D coordinates of the tracked features are calculated using a stereo technique. This study showed reproducibility and accuracy of their system, but it was limited

to the investigation of lip movement in three patients with bilateral cleft lip and palate. Although this method has not been used in a population with facial paralysis, clinical application to this population is a possibility.

Tomat and Mantkelow (2004) have proposed an alternative video analysis system. Rather than tracking individual points on the face with computer software, a video editing program is used to overlay the image frames with the patient at rest and when smiling. The image is then imported into a photographic imaging software system where measurements are obtained using the "ruler" function. Distances in the z-axis are accounted for by using images taken in the semi-profile view. Like many of the previous systems, this system is time-consuming and is unable to provide information on dynamic movement.

Optical Tracking

Recently, a 3D motion capture technique, FACIAL CLIMA, was used to measure facial movements in healthy patients (Hontanilla & Auba, 2008). Optical tracking systems are able to calculate the position and orientation of specific points on the face within a pre-defined coordinate system. This optical tracking system relies on reflective markers affixed to specific landmarks on the face. Based on the information received from the three infrared cameras, the accompanying software program is able to define the location of every facial marker through triangulation. The angles, areas, vectors, and velocities between facial markers can be also calculated with this software program. According to Hontanilla and Auba (2008, p 28), "its main advantage is its ability to automatically capture and reliably quantify 3D small excursions of movement in

the whole face and to describe both the spatial and temporal features of this movement." Additionally, this method is able to capture static and dynamic facial motion, is reliable and easy to use.

Other

Laser scanning (Vannier, Pilgram, Bhatia, Brunsden, & Commean, 1991; O'Grady & Antonyshyn, 1999), mesh analysis (Ferrario, Sforza, Schmitz, Miani, & Serrao, 1998) and moiré topography (Yuen, Inokuchi, Maeta, Kawakami, & Masuda, 1997) also have been used to measure 3D facial motion.

Comparison of Facial Motion Tracking in Two- and Three-Dimensional Space

Over the years, numerous video-based methods have been developed to track facial movement in 2D and 3D space. As mentioned earlier, 2D video-based methods have inherent errors due to projection errors that arise from its inability to accurately capture movement in the anterior-posterior dimension. However, any approach using 2D image sequences to measure 3D facial kinematics must take the problems introduced by *image motion estimation* into consideration. These problems include perspective distortion, poor spatial resolution, and non-linear lighting effects (Kroos, Kuratate, & Vatikiotis-Bateson, 2002). The use and creation of new video-based methods has resulted in the comparison of 2D and 3D video-based systems.

Gross et al. (1996) compared 2D amplitudes captured by a single camera to 2D and 3D amplitudes captured by three cameras. They found that all 2D motion amplitudes correlated with 3D amplitudes, but 2D amplitudes were found to be significantly less than the 3D amplitudes across all animations and

landmarks. The difference was the greatest in the lower face, where 2D amplitudes underestimated the 3D amplitudes by up to forty-three percent. Furthermore, they found that as the differences in the two- and 3D amplitudes increased, the 3D amplitudes increased. These results were replicated by Lin and colleagues (Lin, Chiu, Ho, Su, & Chou, 2000). Consequently, the authors suggest that "2D analysis may not be adequate to assess facial motion during maximum animations and that a 3D analysis is more appropriate for detecting differences in facial function due to disfigurement or surgical interventions" (Gross et al., 1996, p.193).

Purpose

There are no known studies that have used objective measurement tools to assess the functional outcome of patients who have undergone facial reanimation surgery. In the present study, the amount of facial movement in these patients was compared using a 2D video-based system and a 3D optical tracking while performing a range of functional tasks. Previous studies on groups other than post-surgical facial reanimation have indicated that the 3D systems, whether they are video-based or optical, are superior to 2D systems. These assessments are required to describe the patients' facial function, to monitor changes over time, and to document the effectiveness of various treatment options. Therefore, it is important determine if there are significant clinical advantages to using one system over the other on a variety of movement-related variables.

The purpose of this study was two-fold: (1) to compare the amount of movement captured in the 2D video-based system and optical tracking system and

(2) to determine whether there was a difference in the amount of movement captured across different functional tasks such as speech and emotional expression and the affected and unaffected side of the face.

Specific questions were: (1) Will the 3D measurement system capture more movement (in maximum distance, velocity, and acceleration and area-volume ratio) across all of the tasks and variables than the 2D measurement system? (2) Will the 3D system capture more movement than the 2D system on the affected side (i.e. smaller movements) of the face than the unaffected side (i.e. larger movements) of the face? (3) Will the 3D measurement system capture more movement than the 2D system for all tasks? (4) Will the 3D measurement system capture more movement than the 2D system on the affected side of the face than the unaffected side of the face for all tasks?

METHODS

Subjects

Five patients who had undergone facial reanimation surgery at the University of Alberta Hospital were recruited to participate in this study. These 5 patients, who presented with a number of different etiologies, included 3 females and 2 males ranging in age from 11 to 49. Etiologies included Bell's Palsy, Mobius Syndrome, and trauma. All of the patients underwent a microneurovascular transfer involving the gracilis muscle in the thigh, which was either innervated by the mandibular division of the trigeminal nerve or the

contralateral facial nerve. The surgical procedure was dependent upon their clinical situation.

All of the procedures were approved by the University of Alberta Ethics Committee and all of the subjects provided informed consent. Refer to *Appendix 2* for demographic information of all subjects.

Archival Data

The present study was based on archival data. The following procedures were used to collect the two- and 3D kinematic data.

Equipment

For collecting 2D facial kinematics data, a video camera was used. The digital videos were edited using Pinnacle Studio Plus version 10.70 (Pinnacle Systems, Mountainview, CA) to create individual .avi files for each trial and then were converted into Windows.avi files using Adobe Premier version 3.0 (Adobe Systems Incorporated, San Jose, CA). These .avi files were analyzed using Motion Tracker 2D (Dr. D. Webber, University of Pittsburgh, Pittsburgh, PA) in Matlab version 7.1.0 (The Mathworks, Inc., Natick, MA).

3D kinematic data were obtained using an optical tracking system (VZ 3000, VisualEyez, Phoenix Technologies Incorporated, Vancouver, BC). This system consisted of a tracking unit, target control modules (TCMs), light emitting diodes (LEDs), and computer interface software, VZ Server v2.70 and VZ Soft v2.70. The user specified all of the required settings and parameters according to his/her needs with the VZSoft interface. These settings and parameters were passed to the tracking unit, which in turn passed them to the TCMs (e.g., marker

names, sampling period, sensor exposure, etc.) via wireless link or tether cables. The TCM was responsible for designating which and when each LED marker should be fired as the system captured one LED position at a time in the user-specified order and sampling rate. The 3D coordinates of an LED target were calculated by the tracking unit through triangulation. Then, the coordinate position data were transferred back to VZSoft on the host computer to be saved.

Procedure

Two-Dimensional Data Capture

To track facial movement, circular markers approximately 1 cm in diameter at 6 reproducible anatomical landmarks were drawn onto the face using a water soluble eyeliner pencil. The following points were drawn by the investigator: (1) above the midpoint of each eyebrow (directly above the pupil), (2) on each of the cheeks (directly below the pupil at the horizontal projection from the ala of the nose), (3) on each side of the upper lip (between the philtrum and lip corner along the upper lip border), (4) on each lip corner, (5) on each side of the lower lip (between the edge of the chin and the corner of the lip), and (6) on the midpoint of the lower lip. The central nose also was marked as a static point to act as a reference point to measure the excursion of the other facial markers. To maximize the contrast between the markers and skin, the subjects were positioned in front of a white screen and a spotlight was directed onto the face. For the exam, the subjects sat in front of a digital camera and were asked to perform various maximal contraction and speech tasks.

For maximal contraction tasks, the participants were asked to produce the following facial expressions: smiling and puckering. The participants were provided with verbal instructions as well as a model at the start of each trial. In between trials, verbal reminders to relax the face also were given.

In the speech tasks, the participants were asked to repeat eleven consonants (/pi/, /bi/, /mi/, /ni/, /ti/, /di/, /fi/, /vi/, /si/, /shi/, and /wi/) after they were modeled by the investigator. After producing all speech tokens twice, the participants spontaneously produced the familiar nursery rhyme "Jack and Jill" in their regular conversational voice. Each task was repeated twice with time to relax between activities (approximately 5 seconds). The order in which the tasks were performed was randomized for each subject.

The trials with the greatest amount of facial movement and least amount of extraneous head movement were selected for further analysis. The digital video was edited using Pinnacle Studio Plus version 10.70 (Pinnacle Systems, Mountainview, CA) to create individual .avi files for each trial. Approximately one second of the face at rest at the beginning and at the end of the task was saved. Then, these edited video clips were converted to Windows .avi files using Adobe Premier version 3.0 (Adobe Systems Incorporated, San Jose, CA).

Each video clip (video format of 60 fps (DV): resolution, 720 x 480) was analyzed using Motion Tracker 2D (Dr. D. Webber, University of Pittsburgh, Pittsburgh, PA), a Matlab application that is capable of tracking up to 18 markers in a Cartesian coordinate system. The program allows the position of each marker in relation to the static point to be calculated in pixels. The position of each

landmark in the video clip was designated by clicking the center of each marker with the cursor. Once the center of the marker was specified, a “search box”, a program capable of detecting contrast differences between pixels (i.e., between the marker and the face), automatically tracked each marker on the face within the Cartesian plane. If the program was unable to detect a marker or begins to track an inappropriate target, the user re-specified the marker location. The Motion Tracker 2D program used the distance between each subject’s pupils, which was measured by the investigator with a ruler at the time of recording, as a reference to determine the number of pixels in one millimeter. The Matlab files were saved for analysis.

Three-Dimensional Data Capture

3D kinematic information was obtained using an optical tracking system (VZ 3000, VisualEyez, Phoenix Technologies Incorporated, Vancouver, BC). Small light emitting diodes (LEDs) were attached to 6 anatomically reproducible landmarks (lip corners, left and right midlateral upper lip, left and right depressor angular oris, left and right zygomatic arch, left and right midpoint above eyebrow, and central lower lip) using double-sided tape and were reinforced with surgical tape. Three LEDs were placed on static points (the left and right tragus, and the central nose) to define the 3D space. There was a calibration LED on the table beside the patient.

The data were collected using two VisualEyez motion measurement and tracking systems at a sampling frequency of 200 frames per second. The commercial software provided by the manufacturer (VZ Server v2.70 and VZ Soft

v2.70) was used to capture and digitize the information being streamed from the LEDs to the tracking system. Also, each functional task was recorded using a digital video camera that was placed in front of the patient. The video data were taken so that they could be used for post-hoc comparisons of kinematics data and perceptual judgment of function.

Each participant was seated in a comfortable chair that was placed at a specified distance between the sensors of the tracking system and the digital video camera. Kinematic and video recordings were taken while the participant engaged in the speech and maximal contraction tasks described already. Once the examination was over, all of the LEDs were carefully removed from the face.

The tracking system yielded information on distance traveled (mm), the velocity (cm/s), the angle of the movement (degrees), and the acceleration (m/s^2) values for every frame recorded for each dynamic point in the x, y, and z axes in .vzp files. These .vzp files were saved to be analyzed at a later date.

The Present Study

Procedure

Sensitivity of the Measurement Systems

Sensitivity measurements of each measurement system were required to make meaningful comparisons between the two measurement systems. The accompanying manual to the VisualEyez system stated that this system was accurate to 0.015mm at a 1.2 meter distance. At 4x zoom, the 2D video-based system was accurate to 0.5mm for measuring displacements between 1mm and 20mm at a 1.5m distance. The sensitivity of the 2D measurement system was

determined by recording a caliper moving a known distance and comparing this known distance to the distance captured by the 2D video-based system under the same conditions described above.

Two-Dimensional Data

The 2D kinematic data that were previously collected were analyzed using a custom-written Matlab application designed to calculate the maximum excursions of each marker and the total area each marker traveled during the excursion. The Matlab application set the static point as the origin by subtracting the position of static marker from the dynamic ones. Then, the following formula was used to calculate the distance (in pixels) each marker traveled:

$$D = \sqrt{((x_2 - x_1)^2 + (y_2 - y_1)^2)}$$

where D is the distance a marker has traveled from the starting point at a designated point in time, x_1 is the horizontal coordinate of the starting point, x_2 is the horizontal coordinate of the marker at a designated point in time, y_1 is the vertical coordinate of the starting point, and y_2 is the vertical coordinate of the marker at a designated point in time. Using this formula, the distance the marker traveled from its starting point was calculated for all time frames. The largest D value was determined to be the maximum displacement of that marker. All distances traveled by each marker in the x-axis were converted into millimeters using the distance between each of the subject's pupils as a referent for all x values. All vertical values were multiplied by 1.080264 to eliminate horizontal compression, which was a product of the video capture process.

The velocity of each marker was calculated using the values which had been converted into millimeters using pixel ratios, as described above. Average velocity was calculated using the following formula:

$$V= (\sqrt{\ } ((x_2 - x_1)^2 + (y_2 - y_1)^2))/ (t_2 - t_1)$$

where V is the velocity of a marker between two points in time, x_1 is the horizontal coordinate at t_1, x_2 is the horizontal coordinate of the marker at t_2, y_1 is the vertical coordinate at t_1, y_2 is the vertical coordinate of the marker at t_2, t_1 is the time at which the movement begins, and t_2 is the time at which the movement ends. The velocity of movement was calculated for each frame interval. The largest *V* value across all time frames was determined to be the maximum velocity achieved for that marker.

Acceleration of each marker was calculated using the following formula:

$$A= (V_2 - V_1)^2 / (t_2 - t_1)$$

where *A* is the acceleration of a marker between two points in time, V_1 is the velocity at t_1, and V_2 is the velocity at t_2. The acceleration of each marker was calculated for each frame interval. The largest *A* value across all time frames was determined to be the maximum acceleration achieved for that marker.

In order to calculate the area of movement of each marker, the "convhull" function in Matlab was used. This function defined the area to be calculated by identifying all data points lying furthest from the center of the scatter of data points. Once these points were identified, Delaunay triangulation was used to determine area within these points by segmenting it into multiple triangles. The

area of each triangle was calculated and the areas of all triangles were summed to determine the area of movement for each marker.

Three-Dimensional Data

The raw data in the archived .vzp files were extracted and saved into Excel files. The raw data were analyzed by another in-house Matlab program which was capable of calculating distance traveled, velocity, acceleration, and volume much in the same way that the 2D data were analyzed.

Before any calculations were made, a coordinate axis was set within the frame of the face to eliminate the effects of head or body movement. This was achieved by creating a triangular plane using the nose point and the left and right tragus points. All of the markers were rotated to be aligned with the tragus-nose plane. The nose point was used as the origin and was the common reference between frames. Setting the nose point as the origin prevented translation of the coordinate axis. Three rotations of all data points were required to align the coordinate axis with the triangular plane formed by the nose point and tragus points. First, a vector running between the midpoint of the tragus points and the nose were aligned to the ZY plane. Then the nose point and the midpoint between the tragus points were aligned with the y-axis. Finally, the nose to left tragus vector was aligned to the positive XY plane, which also created a z-axis that is normal to the triangular plane.

Distance values were obtained by using the following formula:

$$D = \sqrt{((x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2)}$$

where *D* is the distance a marker has traveled from the starting point at a designated point in time, x_1 is the horizontal coordinate of the starting point, x_2 is the horizontal coordinate of the marker at a designated point in time, y_1 is the anterior-posterior coordinate of the starting point, y_2 is the anterior-posterior coordinate of the marker at a designated point in time, z_1 is the vertical coordinate of the starting point, and z_2 is the vertical coordinate of the marker at a designated point in time. The largest *D* value was used to determine the coordinates for the maximal excursion of each marker.

The average velocity of each marker was calculated using the following formula:

$$V= (\sqrt{\ } ((x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2))/ (t_2 - t_1)$$

where *V* is the velocity of a marker between two points in time, x_1 is the horizontal coordinate at t_1, x_2 is the horizontal coordinate of the marker at t_2, y_1 is the anterior-posterior coordinate at t_1, y_2 is the anterior-posterior coordinate of the marker at t_2, z_1 is the vertical coordinate at t_1, and z_2 is the vertical coordinate of the marker at t_2. The velocity of movement was calculated for each frame interval. The largest *V* value across all time frames was determined to be the maximum velocity achieved for that marker.

Acceleration for 3D data was calculated using the same formula used when calculating acceleration for 2D data.

Variables

The first degree variable in this study was the maximum distance that each marker traveled in millimeters. The following second degree variables were also

measured: area and volume ratios, velocity, and acceleration. These variables were considered to be secondary as there were few studies that had investigated them previously. Velocity and acceleration variables have not been widely used in kinematic studies as each distance-time derivation propagates error. Standard mathematical formulas were available to derive all variables. The definition of each dependent variable is listed in Table 1.

Table 1

Definitions of dependent variables measured in two and three dimensions

Variables	Definition
Maximum distance	The largest distance that each marker traveled from its starting point (in mm).
Area ratio	Area was defined as the 2D space (in mm^2) that each marker covered during a task. The ratio was obtained by comparing the area for each bilateral marker on the unaffected face to the affected face. When the paralysis was bilateral, the operated side was considered to be the unaffected side. The area ratio obtained was contrasted with the volume ratio data collected from the 3D measurement system.
Volume ratio	Volume is defined as the 3D space (in mm^3) that each marker covered during a task. The ratio was obtained by comparing the volume for each bilateral marker on the unaffected face to the affected face. When the paralysis was bilateral, the operated side was considered to be the unaffected side. The volume ratio obtained was contrasted with the area ratio data collected from the 2D measurement system.
Maximum Velocity	The maximum speed at which each marker traveled during a task (in mm/s).
Maximum Acceleration	The maximum rate at which the velocity of each marker changes during a task (mm/s^2).

Design

The current study followed a 2 x 2 x 5 within-subjects quasi-experimental design. There were three independent variables in this study: (1) type of measurement system, (2) side of face, and (3) task. The first independent variable had two levels: 2D video-based system and 3D optical tracking. The second independent variable also had two levels: affected and unaffected side of face. The type of task had five levels: smile, pucker, /wi/, /ni/, and the nursery rhyme "Jack and Jill." There were four dependent variables with which to measure facial movement: the maximal distance, area/volume ratio, velocity, and acceleration for each marker. All of the dependent variables were ratio level data.

DATA ANALYSIS

Data were collected for each facial marker for each task for all subjects. The movement data for each marker during each task averaged across trials for each subject was used in the analysis. Because the research questions in the present study were related to the performance of 2D vs. 3D tracking, data from facial marker placements were examined for task and method rather than averaged for each subject. When outliers were identified as belonging to a single subject, this was clearly noted in the result section to follow. To answer the first research question, all data from 2D and 3D systems were compared by collapsing tasks and markers for each dependent variable. The same analysis method was used for comparing the affected and unaffected side of the face to answer the second research question. To answer the third research question, the 2D and 3D

measurement systems were compared by collapsing the data across markers within each task. The data from the third question were further subdivided into the affected and unaffected side of the face to answer the fourth question. When comparing the affected and unaffected side of the face, the data for all central markers were excluded. Also, by necessity, the area-volume ratio data were excluded from this comparison as theses values were inclusive of affected and unaffected side of face.

Box plots were used for visual inspection of the data in terms of distribution, central tendency and outliers. The central box represents data that fall between the 25^{th} and 75^{th} percentile, or 50 percent of the data, while the whiskers extending from the box represent the range into which the data fall. Outliers and extreme outliers are pictured by open circles and asterisks, respectively. The line within the box represents the median value. If the descriptive statistics for each data set met the criteria for a normal distribution, then paired *t*-tests were used to compare the two measurement systems. If the skewness value was greater than ±2, the *Wilcoxon* signed-rank test was used to make within group comparisons. Predictions were in one direction and all comparisons were one-tailed with an *a priori* significance of $p < .05$.

RESULTS

Question #1: *Will the 3D measurement system be more sensitive to degree of movement than the 2D measurement system, independent of task type and measurement variable?*

Question 1a. Will the 3D measurement system be more sensitive to distance than the 2D measurement system, independent of task type and measurement variable?

Table 2
Descriptive Statistics for Distance (in mm) Collapsed for All Markers and All Tasks for 2D and 3D Measurement Systems

	Minimum	Maximum	Mean	SD	Variance	Skewness
2D[a]	.728	14.693	4.686	2.898	8.396	1.082
3D[b]	1.057	23.359	7.151	4.520	20.430	.920

[a]n=211. [b]n=225.

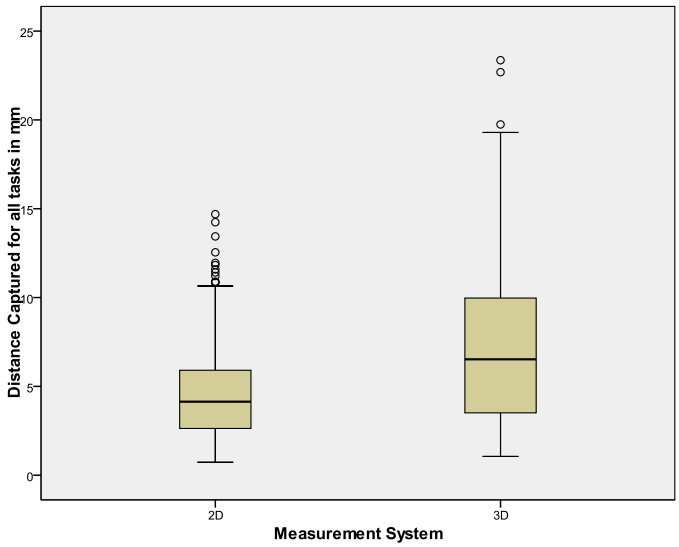

Figure 1. Box plots showing the distribution of the distance measurement (in mm) collapsed for markers and tasks for 2D and 3D measurement systems

Visual inspection of Figure 1 revealed that the distance values for both measurement systems were normally distributed and that there were more outliers associated with the 2D data, but the 3D data had a larger median and range than the 2D data. The distance captured by the 3D system was significantly greater than distance captured by the 2D system [t (210) = -11.34, $p \leq .0001$, one-tailed].

Question 1b. Will the 3D measurement system be more sensitive to maximum velocity than the 2D measurement system, independent of task type and measurement variable?

Table 3
Descriptive Statistics for Maximum Velocity (mm/s) Collapsed Across All Markers and All Tasks for 2D and 3D Measurement Systems

	Minimum	Maximum	Mean	SD	Variance	Skewness
2D[a]	22.377	307.191	90.349	52.172	2721.943	1.778
3D[b]	156.899	1599.346	323.892	217.107	47135.484	2.597

[a]n=211. [b]n=225.

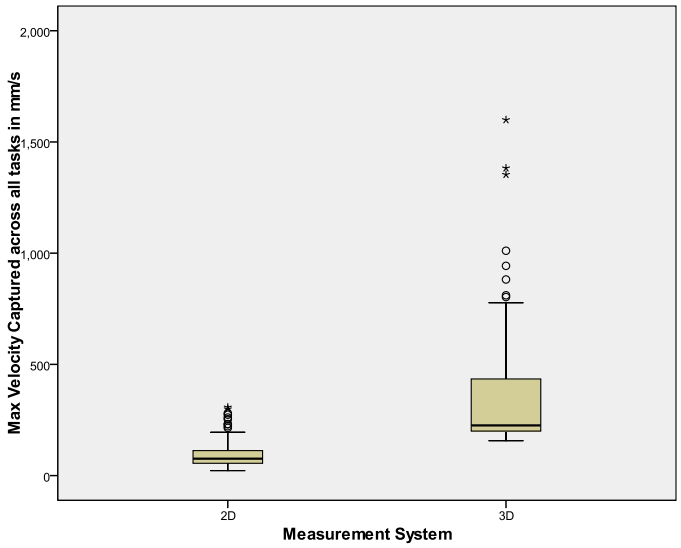

Figure 2. Box plots of maximum velocity (in mm/s) collasped across all markers and tasks for 2D and 3D measurement systems

As shown in Figure 2, the 2D data appeared to be normally distributed whereas the 3D data appeared positively skewed. Visual inspection showed that the median value was greater and that the box was larger for the 3D data than the 2D data. All outlying data points in the 3D data are the result of a single subject, S5. Not surprisingly, there was a significant difference between the 2D and 3D measurement systems for maximum velocity measures (*Wilcoxon Z*= -12.533, $p \leq$ .0001, one-tailed). Analysis of data excluding S5, the most extreme outlier, showed that differences between 2D and 3D continued to be significant (*Wilcoxon Z*=-11.086, $p \leq .0001$, one-tailed).

Question 1c. Will the 3D measurement system be more sensitive to maximum acceleration than the 2D measurement system, independent of task type and measurement variable?

Table 4
Descriptive Statistics for Maximum Acceleration (in mm/s^2) Collapsed Across All Markers and All Tasks for 2D and 3D Measurement Systems

	Minimum	Maximum	Mean	SD	Variance	Skewness
2D[a]	1.95×10^3	2.68×10^4	6.65×10^3	3.99×10^3	1.59×10^7	2.083
3D[b]	3.49×10^4	4.22×10^5	7.61×10^4	5.24×10^4	2.74×10^9	3.158

[a]n=211. [b]n=225.

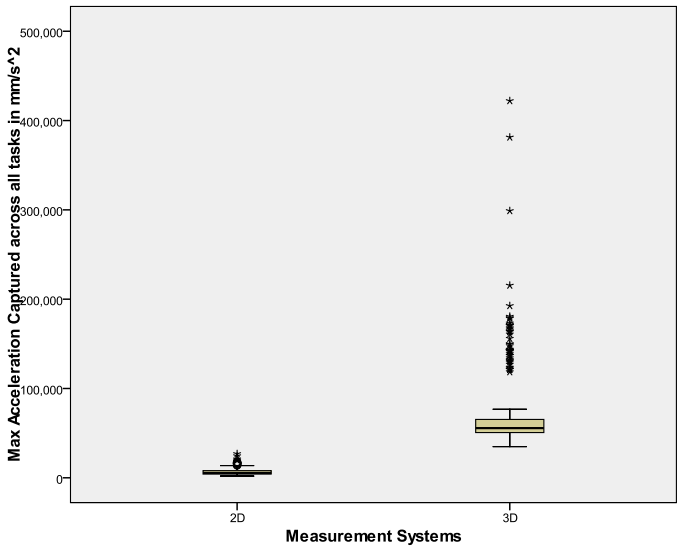

Figure 3. Box plots of maximum acceleration (in mm/s^2) collapsed across all markers and tasks for 2D and 3D measurement systems

Similar to the two previous figures, the 3D system captured more movement and a greater range of movement than the 2D system. It also was evident that the 3D system captured far more extreme outliers than the 2D system. However, it should be noted that all outlying acceleration values measured by the 3D system belong to S5. The 2D data were distributed at the lower end of the scale. The median of the 3D data was much larger than the 2D median and there was no overlap between the two sets of data. A *Wilcoxon* signed-rank test

revealed a significant difference between the 2D and 3D system when measuring maximum acceleration (*Wilcoxon Z*=-12.595, $p \leq .0001$, one-tailed). When the data were re-analyzed without S5, differences between the two systems remained significant (*Wilcoxon Z*=-11.175, $p \leq .0001$, one-tailed).

Question 1d. Will the 3D measurement system be more sensitive to area-volume ratio than the 2D measurement system, independent of task type and measurement variable?

Table 5
Descriptive Statistics for Area-volume Ratio Collapsed Across All Markers and All Tasks for 2D and 3D Measurement Systems

	Minimum	Maximum	Mean	SD	Variance	Skewness
2D[a]	.464	6.888	1.799	1.119	1.253	2.552
3D[b]	.164	30.104	4.196	4.915	24.160	2.847

[a]n=211. [b]n=225.

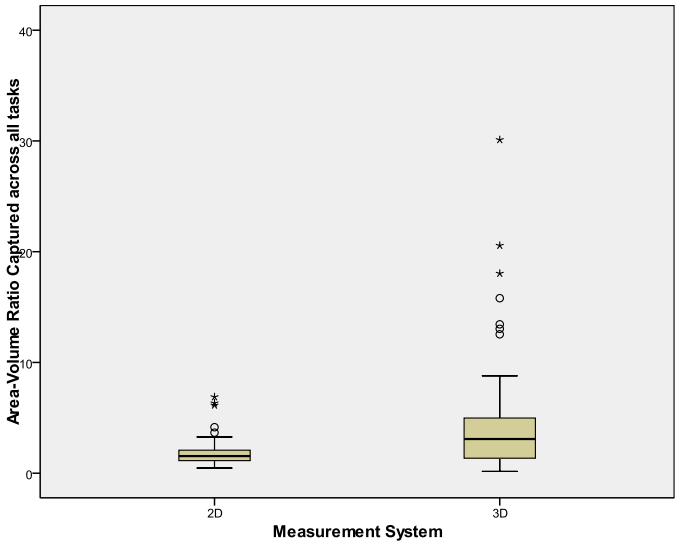

Figure 4. Box plots of maximum acceleration (in mm/s^2) collapsed across all markers and tasks for 2D and 3D measurement systems

Figure 4 shows that both 2D and 3D data sets were positively skewed. Although the values at the lower end of the range were similar, the upper range was visibly larger for 3D ratios. The median and the range of data captured was larger in the 3D measurement system than 2D (*Wilcoxon Z* = -5.359, $p \leq .0001$, one-tailed).

Question 2: *Will the 3D measurement system be more sensitive than the 2D system to degree of movement (i.e., larger movements on the unaffected side of the face versus smaller movement on the affected side of the face)?*

Question 2a. Will the 3D measurement system be more sensitive to maximum distance than the 2D system on the affected side of the face than the unaffected side of the face?

Table 6
Descriptive Statistics for Distance (in mm) Collapsed on the Affected and Unaffected Side of the Face for 2D and 3D Measurement Systems

	Minimum	Maximum	Mean	SD	Variance	Skewness
2D Affected[a]	.728	11.415	3.580	1.994	3.978	1.169
2D Unaffected[b]	.8555	12.547	5.091	2.800	7.838	.583
3D Affected[c]	1.057	16.091	4.948	3.131	9.805	1.127
3D Unaffected[c]	21.609	1.078	22.688	4.399	19.349	.631

[a]n=91. [b]n=96. [c]n=100.

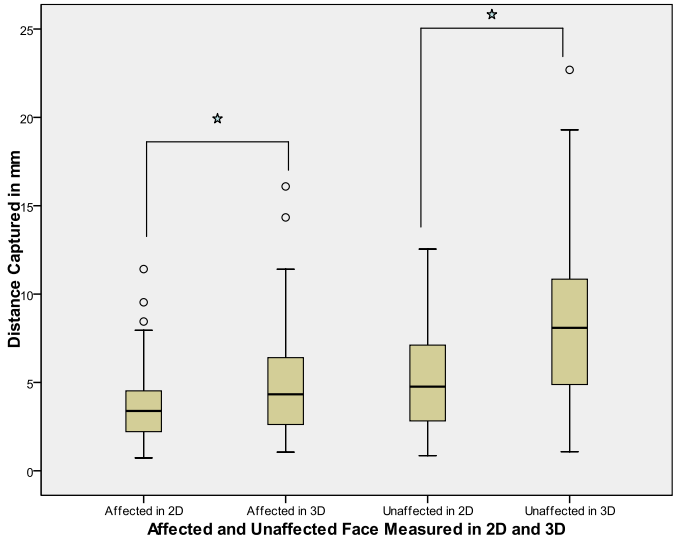

Figure 5. Box plots for maximum distance (in mm) collapsed for the affected and unaffected side of the face in 2D and 3D. *Significant differences ($p < .05$) between 2D and 3D measurement systems

Figure 5 shows normally distributed data for all conditions. More movement was captured on the unaffected side of the face compared to the affected side of the face. The low-end of the range for each condition was similar for both measurement techniques; however, maximum distance captured was greatest for the unaffected face measured in 3D. The affected side of the face measured in 2D had the most reduced range. More outliers were observed in the affected side face measured in 2D and 3D than in the unaffected side of the face. Also, more movement was captured by the 3D system compared to the 2D system. A paired-samples *t*-test revealed a significant difference between the 2D and 3D systems when measuring the affected side of the face [$t(90)= -5.759$, $p \leq .0001$, one-tailed] and the unaffected side of the face [$t(95)= -8.900$, $p \leq .0001$, one-tailed].

Question 2b. Will the 3D measurement system be more sensitive to maximum velocity than the 2D system on the unaffected side of the face than the affected side of the face?

Table 7
Descriptive Statistics for Maximum Velocity (in mm/s) Collapsed on the Affected and Unaffected Side of the Face for 2D and 3D Measurement Systems

	Minimum	Maximum	Mean	SD	Variance	Skewness
2D Affected[a]	22.377	231.357	76.340	38.457	1478.953	1.595
2D Unaffected[b]	30.219	280.833	95.931	53.769	2891.110	1.549
3D Affected[c]	178.070	942.958	298.246	177.631	31552.82	1.898
3D Unaffected[c]	156.899	1599.346	354.412	260.287	67749.69	2.594

[a]n=91. [b]n=96. [c]n=100.

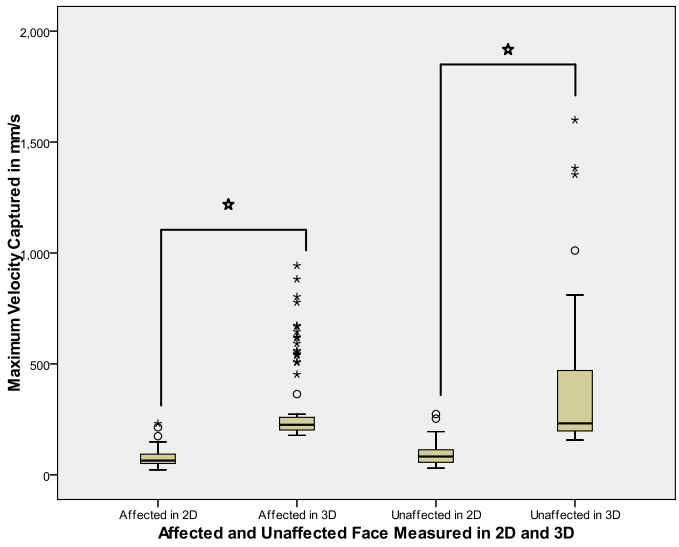

Figure 6. Box plots for maximum velocity (in mm/s) collapsed for the affected and unaffected side of the face in 2D and 3D. *Significant differences ($p < .05$) between 2D and 3D measurement systems.

As shown in *Figure 6*, the 2D data appeared to be normally distributed whereas the 3D data appeared to be positively skewed. It was evident that the 3D

system detected a greater range of marker velocity than the 2D system. The box plots showed little difference between the maximum velocity of the affected and unaffected face when measured by the 2D system. However, there was more movement and a greater range of movement in the unaffected face compared to the affected face when measured by the 3D system. The 3D measurement system appeared to capture more outliers than the 2D system. All outliers perceived in 3D belonged to a single subject, S5. Figure 6 also shows the apparent difference between the amount of movement captured between the 2D and 3D systems. The 3D system detected faster marker movement on the affected (*Wilcoxon Z* = -8.284, $p \leq .0001$, one-tailed) and the unaffected side of the face (*Wilcoxon Z* = -8.420, $p \leq .0001$, one-tailed) than the 2D system. Differences between the two systems remained significant for the affected side of the face (*Wilcoxon Z*=-7.323, $p \leq .0001$, one-tailed) and the unaffected side of the face [$t(75)= -12.186$, $p \leq .0001$, one-tailed] when data from S5 were excluded.

Question 2c. Will the 3D system be more sensitive to maximum acceleration than the 2D system on the unaffected side of the face than the affected side of the face?

Table 8
Descriptive Statistics for Maximum Acceleration (in mm/s ^{2}) Collapsed on the Affected and Unaffected Side of the Face for 2D and 3D Measurement Systems

	Minimum	Maximum	Mean	SD	Variance	Skewness
2D Affected[a]	1.95×10^3	1.62×10^4	5.93×10^3	2.80×10^3	7.87×10^6	1.553
2D Unaffected[b]	2.17×10^3	2.68×10^4	7.23×10^3	4.80×10^3	2.31×10^7	1.892
3D Affected[c]	4.61×10^4	1.79×10^5	7.47×10^4	3.93×10^4	1.55×10^9	1.585
3D Unaffected[c]	4.10×10^4	4.22×10^5	7.83×10^4	6.52×10^4	4.25×10^9	3.245

[a]n=91. [b]n=96. [c]n=100.

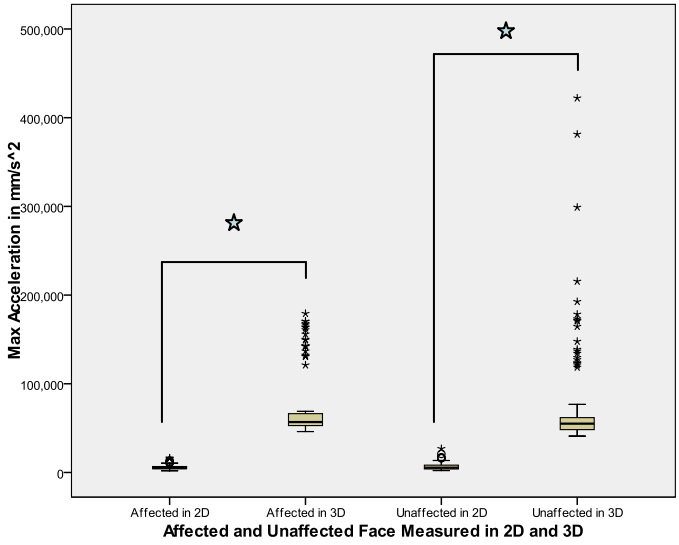

Figure 7. Box plots for Maximum Acceleration (in mm/s ^{2}) collasped for the Affected and Unaffected Side of the Face in 2D and 3D. * Significant differences ($p < .05$) between 2D and 3D measurement systems

Figure 7 shows the difference in the amount of movement captured between 2D and 3D. There was a significant difference in the maximum acceleration detected by the 2D system and 3D on the affected (*Wilcoxon Z* = -8.284, $p \leq .0001$, one-tailed) and unaffected side of the face (*Wilcoxon Z* = -8.284, $p \leq .0001$, one-tailed). Not only did the 3D system capture more movement, it

also detected a greater range of movement. Interestingly, there did not appear to be a difference in maximum acceleration when comparing the affected and unaffected face, regardless of the measurement system. Again, all extreme outliers found in the 3D condition belonged to S5. Analysis completed without data from S5 showed a significant difference between the two measurement system on the affected side of the face [$t(75)= -12.186$, $p \leq .0001$, one-tailed] and the unaffected side of the face (*Wilcoxon* $Z = -7.574$, $p \leq .0001$, one-tailed).

Question #3: *Will the 3D measurement system be more sensitive to degree of movement than the 2D system for all tasks?*

Question 3a. Will the 3D measurement system be more sensitive to distance than the 2D measurement system for all tasks?

Table 9
Descriptive Statistics for Distance (in mm) Collapsed for All Markers for 2D and 3D Measurement Systems by Task

	Minimum	Maximum	Mean	SD	Variance	Skewness
Smile						
2D[a]	1.396	11.585	4.762	2.551	6.509	1.044
3D[b]	1.057	22.688	7.402	5.465	29.867	.997
Pucker						
2D[c]	.728	10.841	5.578	2.277	5.184	.592
3D[b]	1.138	16.091	6.811	3.582	12.832	.377
Wi						
2D[c]	.930	11.827	3.744	2.467	6.086	1.631
3D[b]	1.078	18.450	6.763	4.125	17.019	.681
Ni						
2D[c]	.855	13.442	3.262	2.828	7.997	2.041
3D[b]	1.121	19.744	6.664	4.529	20.508	.929
Jack & Jill						
2D[c]	1.017	14.693	6.101	3.277	10.739	.930
3D[b]	1.352	23.359	8.118	4.717	22.246	1.036

[a]n=35. [b]n=45. [c]n=44.

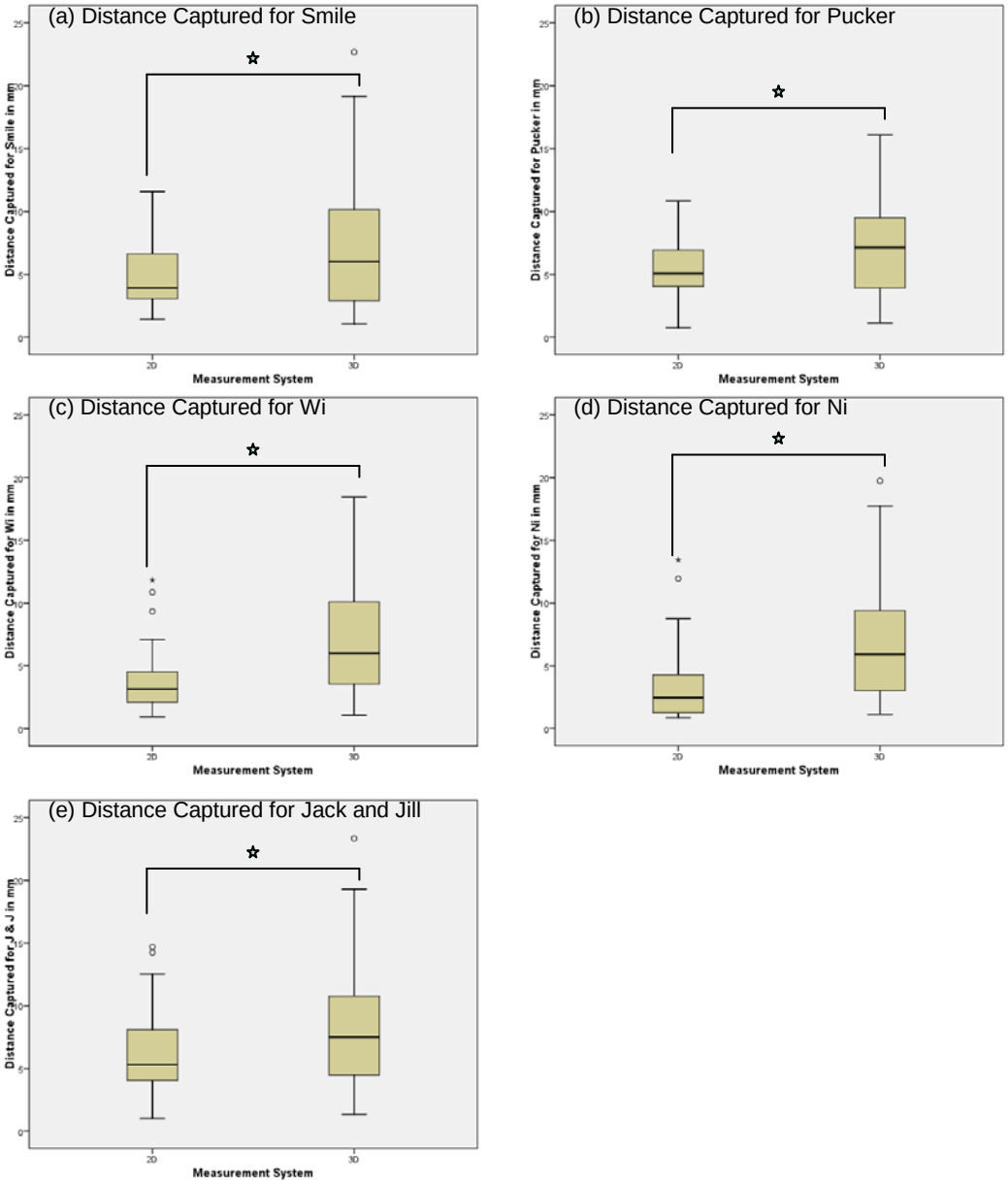

Figure 8. Distance (in mm) captured by 2D and 3D measurement systems for (a) Smile, (b) Pucker, (c) Wi, (d) Ni, and (e) Jack and Jill. *Significant differences ($p < .05$) between 2D and 3D measurement systems

Figure 8 showed larger values and greater range for all 3D data compared to 2D data. There was a significant difference between 2D and 3D measurement system for all tasks. Paired *t*-tests revealed significant differences between the two measurement systems for smile [$t(34)= -3.878$, $p < .0001$, one-tailed], pucker

[$t(43)= -3.168$, $p = .0015$, one-tailed], wi [$t(43)= -6.840$, $p \leq .0001$, one-tailed], and Jack and Jill [$t(43)= -4.988$, $p \leq .0001$, one-tailed]. A Wilcoxon signed-rank test indicated a significant difference between the 2D and 3D system for ni (*Wilcoxon Z* = -5.613, $p \leq .0001$, one-tailed).

Question 3b. Will the 3D measurement system be more sensitive to maximum velocity than the 2D measurement system for all tasks?

Table 10
Descriptive Statistics for Maximum Velocity (in mm/s) Collapsed for All Markers for 2D and 3D Measurement Systems by Task

	Min.	Max.	Mean	SD	Variance	Skewness
Smile						
2D[a]	41.255	231.277	85.198	39.369	1549.904	1.892
3D[b]	167.853	1353.159	324.839	251.840	63423.322	2.553
Pucker						
2D[c]	33.042	260.959	86.414	38.157	1455.958	2.340
3D[b]	178.070	1382.372	331.382	231.078	53397.208	2.613
Wi						
2D[c]	25.887	280.833	87.864	46.207	2135.116	1.803
3D[b]	179.953	810.652	306.777	163.977	26888.568	.354
Ni						
2D[c]	22.377	184.871	65.483	36.497	1332.036	1.646
3D[b]	156.899	668.150	304.594	163.655	26783.037	1.063
Jack & Jill						
2D[c]	35.272	307.191	125.733	72.021	5186.956	.945
3D[b]	165.983	1599.346	351.867	260.636	67931.188	2.935

[a]n=35. [b]n=45. [c]n=44.

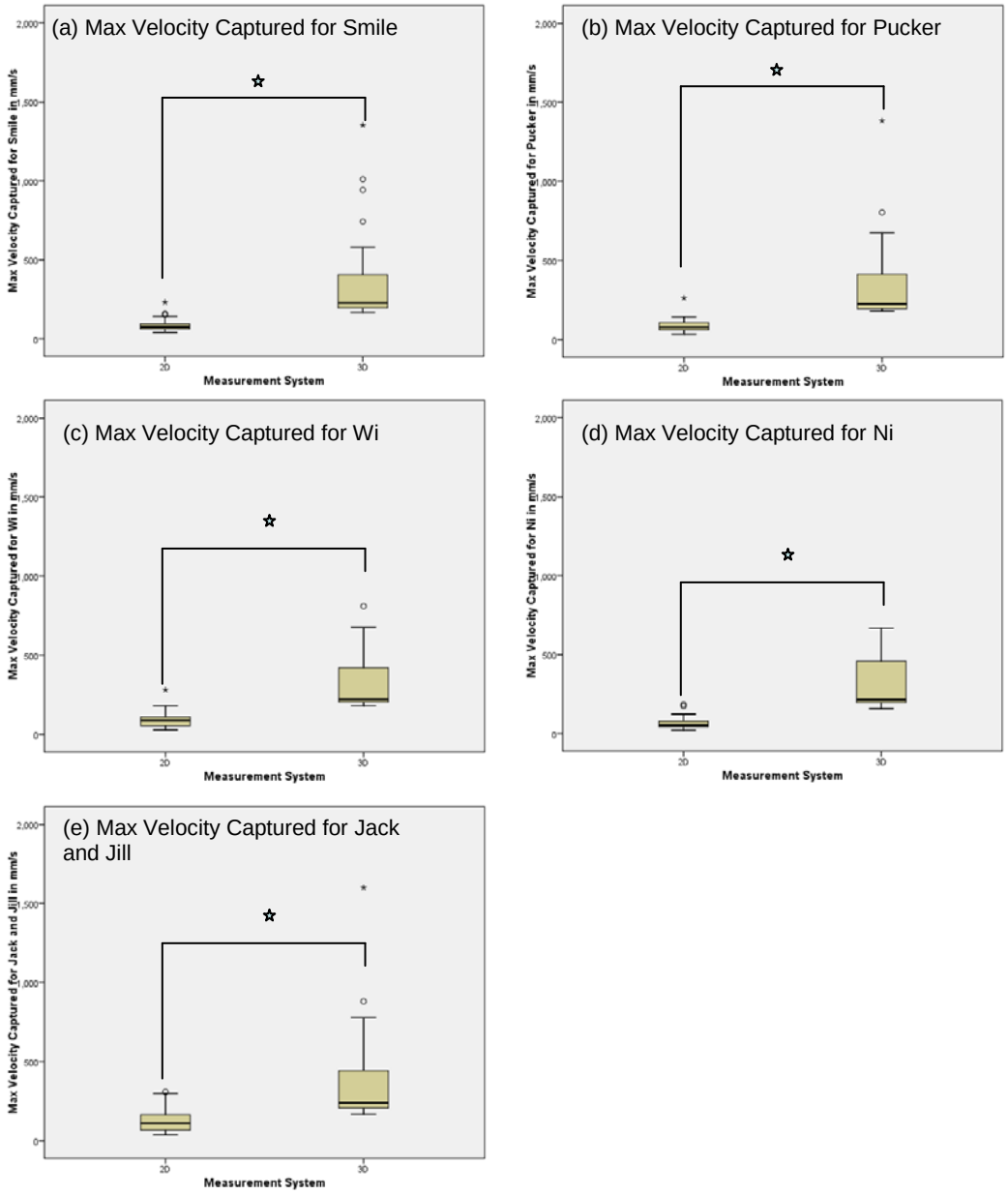

Figure 9. Maximum velocity (in mm/s) captured by 2D and 3D measurement systems for (a) Smile, (b) Pucker, (c) Wi, (d) Ni, and (e) Jack and Jill. *Significant differences ($p < .05$) between 2D and 3D measurement systems

The box plots in Figure 9 look relatively similar. They all followed the general trend of greater movement and greater range of movement captured by 3D compared to the 2D system. Furthermore, the 2D values appeared to be approximating a floor effect. 2D data for all tasks were normally distributed

except the data derived for pucker. 3D data were normally distributed for wi and ni. Significant differences between 2D and 3D measurement systems were found for smile (*Wilcoxon Z* = -5.143, $p \leq .0001$, one-tailed), pucker (*Wilcoxon Z* = -5.70, $p \leq .0001$, one-tailed), wi [$t(43) = -8.857$, $p \leq .0001$, one-tailed], ni (*Wilcoxon Z* = -5.765, $p \leq .0001$, one-tailed), and Jack and Jill (*Wilcoxon Z* = -5.718, $p \leq .0001$, one-tailed). Significant differences between the two measurement systems remained for all tasks when data from S5 were excluded.

Question 3c. Will the 3D measurement system be more sensitive to maximum acceleration than the 2D measurement system for all tasks?

Table 11
Descriptive Statistics for Maximum Acceleration (in mm/s^2) Collapsed for All Markers for 2D and 3D Measurement Systems by Task

	Minimum	Maximum	Mean	SD	Variance	Skewness
Smile						
2D[a]	3.44×10^3	1.65×10^4	6.31×10^3	3.32×10^3	1.10×10^7	2.043
3D[b]	4.25×10^4	2.99×10^5	7.95×10^4	6.13×10^4	3.76×10^9	2.536
Pucker						
2D[c]	2.20×10^3	1.78×10^4	6.23×10^4	3.21×10^3	1.03×10^7	1.727
3D[b]	3.49×10^4	3.81×10^5	8.09×10^4	6.20×10^4	3.84×10^9	3.081
Wi						
2D[c]	1.95×10^3	2.45×10^4	6.70×10^3	4.09×10^3	1.68×10^7	2.381
3D[b]	4.19×10^4	1.93×10^5	7.16×10^4	4.00×10^4	1.60×10^9	1.781
Ni						
2D[c]	2.13×10^3	1.16×10^4	5.18×10^3	2.37×10^3	5.61×10^6	1.126
3D[b]	4.10×10^4	1.78×10^5	7.07×10^4	3.80×10^4	1.44×10^9	1.075
Jack & Jill						
2D[c]	2.83×10^3	2.68×10^4	8.78×10^3	5.41×10^3	2.93×10^7	1.401

3D[b]	4.342 x 10^4	4.22 x 10^5	8.08 x 10^4	6.52 x 10^4	4.25 x 10^9	3.657

[a]n=35. [b]n=45. [c]n=44.

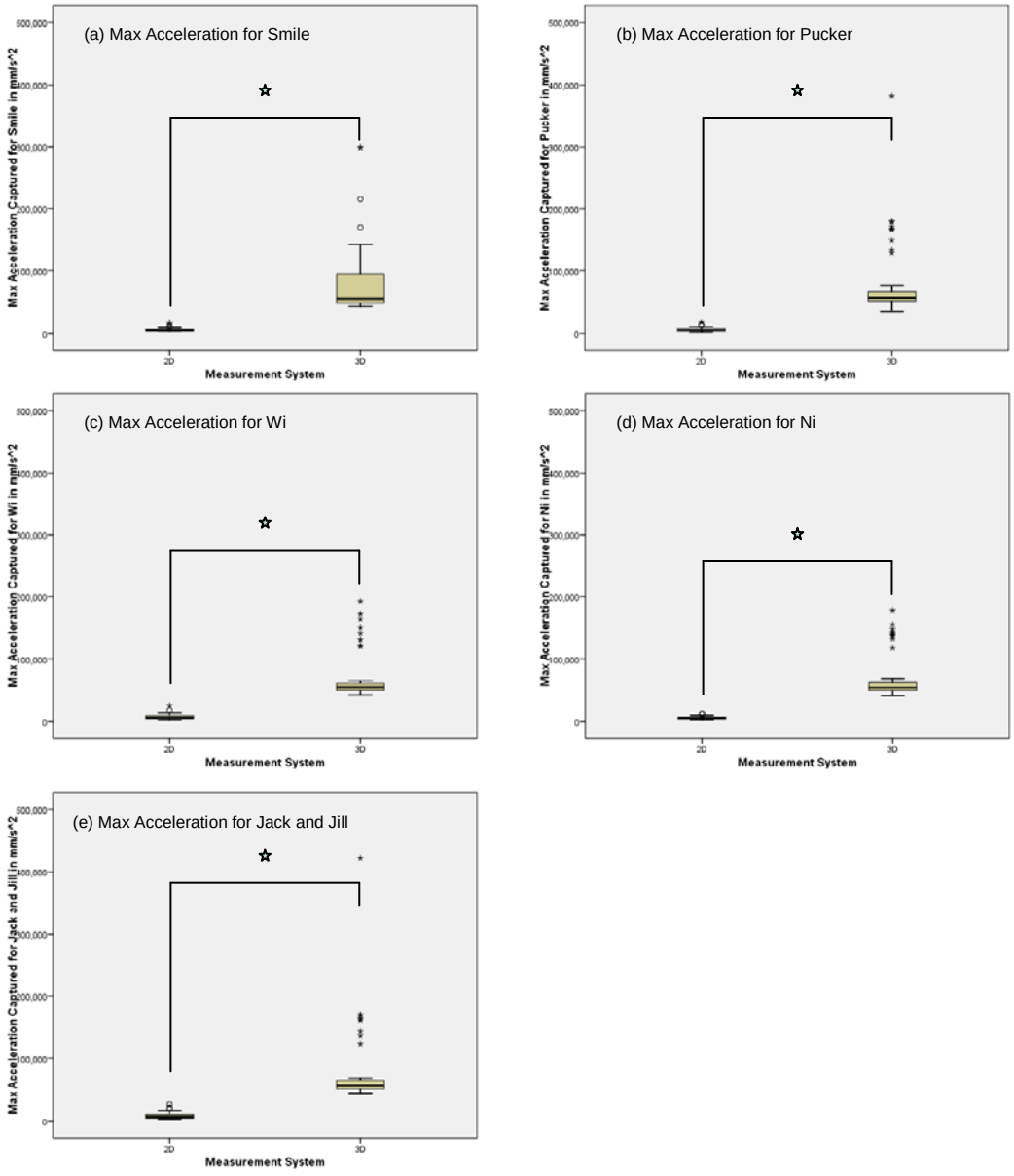

Figure 10. Maximum acceleration (in mm/s^2) captured by 2D and 3D measurement systems for (a) Smile, (b) Pucker, (c) Wi, (d) Ni, and (e) Jack and Jill. *Significant differences ($p < .05$) between 2D and 3D measurement systems

The box plots in Figure 10 indicated that when measuring maximum acceleration, the 3D system captured more acceleration and a greater range of

acceleration than the 2D system. The data measured by the 2D system may have indicated a floor effect. The data appeared to be normally distributed for "ni" only. The numerous outliers in the 3D data were also apparent on visual inspection of the box plots. All 3D outliers were for S5. Additionally, a significant difference between the 2D and 3D measurement systems was apparent for all tasks. A significant difference was found between the 2D and 3D systems when measuring the maximum acceleration for smile (*Wilcoxon Z* = -5.159, $p \leq$ .0001, one-tailed), pucker (*Wilcoxon Z* = -5.777, $p \leq .0001$, one-tailed), wi (*Wilcoxon Z* = -5.777, $p \leq .0001$, one-tailed), ni [$t(43) = -11.647$, $p < .0001$, one-tailed], and Jack and Jill (*Wilcoxon Z* = -5.777, $p \leq .0001$, one-tailed). Differences between the two measurement systems continued to be significant for all tasks when data for S5 was excluded.

Question 3d. Will the 3D measurement system be more sensitive to area-volume ratio than the 2D measurement system for all tasks?

Table 12
Descriptive Statistics for Area-Volume Ratio Collapsed for All Markers for 2D and 3D Measurement Systems by Task

	Minimum	Maximum	Mean	SD	Variance	Skewness
Smile						
2D[a]	1.187	6.329	3.044	1.540	2.370	1.137
3D[b]	.889	30.104	7.740	7.957	63.313	1.419
Pucker						
2D[c]	.464	6.888	1.737	1.395	1.946	2.971
3D[b]	.164	21.471	4.550	5.714	32.646	2.182
Wi						
2D[c]	.564	3.019	1.613	.611	.374	.762
3D[b]	.448	6.168	2.350	1.456	2.120	.954
Ni						
2D[c]	.496	2.454	1.397	.545	.297	.157
3D[b]	.395	8.258	3.243	2.405	5.785	.560
Jack & Jill						
2D[c]	.925	2.375	1.468	.433	.187	.743
3D[b]	.339	5.519	3.099	1.691	2.858	-.206

[a]n=15. [b]n=20. [c]n=19.

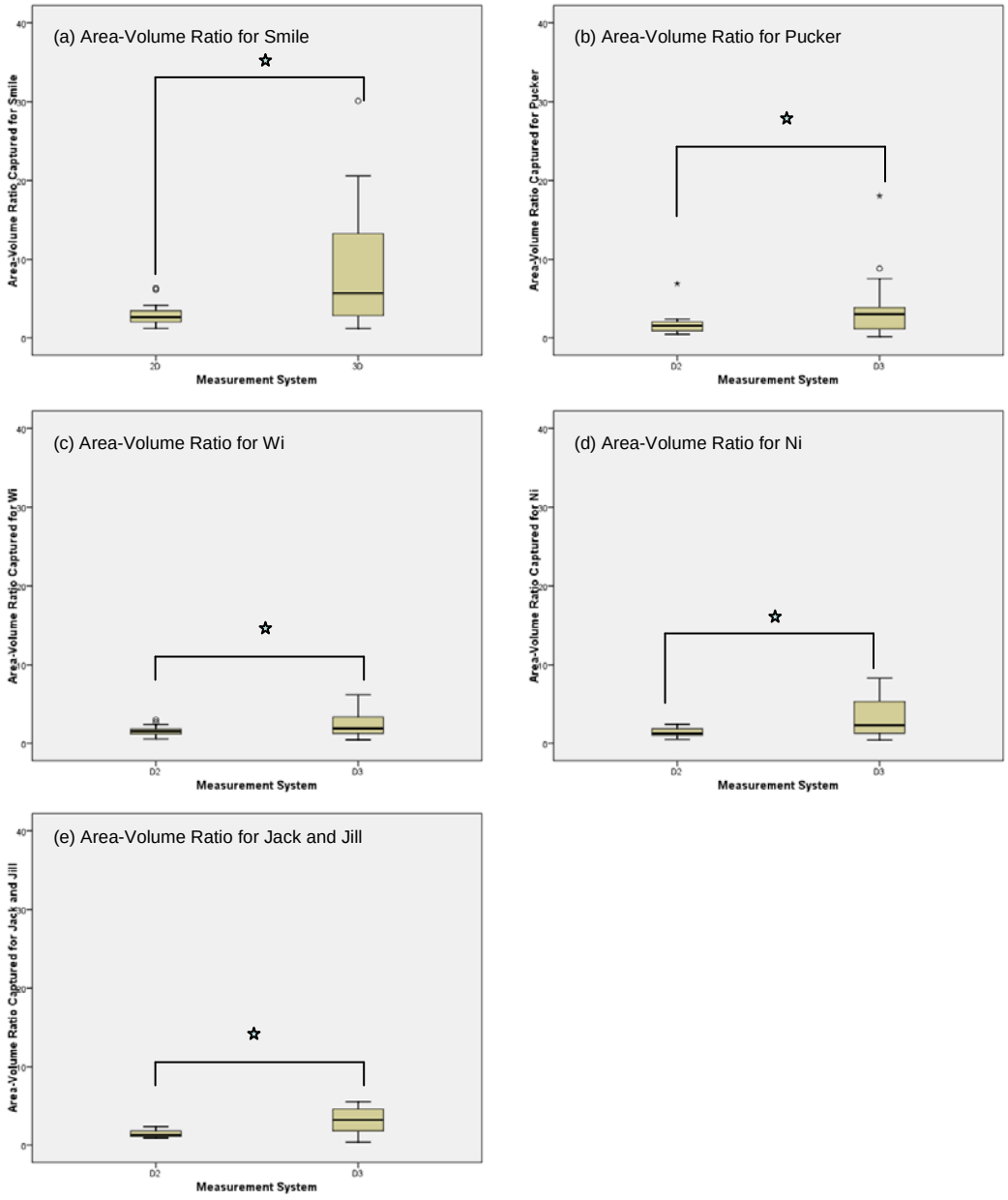

Figure 11. Area and volume ratios captured by 2D and 3D measurement systems for (a) Smile, (b) Pucker, (c) Wi, (d) Ni, and (e) Jack and Jill. *Significant differences ($p < .05$) between 2D and 3D measurement systems

The box plots shown in Figure 11 revealed that although the lowest values for the 2D and 3D data were similar, the 3D system captured a greater range of area-volume ratios. The medians of the data captured by the 3D system were greater than the 2D system for all tasks. The 2D data appeared to be normally

distributed except for "pucker" and the 3D data for "wi." There was a significant difference between the 2D and 3D system when measuring area-volume ratios for smile ($t(14)= -2.769$, $p = .0075$, one-tailed), pucker (*Wilcoxon Z* = -2.059, $p = .02$, one-tailed), wi ($t(18)= -1.881$, $p = .038$, one-tailed), ni ($t(18)= -2.769$, $p = .001$, one-tailed), and Jack and Jill ($t(18)= -4.200$, $p = .0005$, one-tailed).

Question #4: *Will the 3D measurement system be more sensitive to degree of movement than the 2D measurement system on the unaffected side of the face than the affected side of the face for all tasks?*

Question 4a. Will the 3D system be more sensitive to distance than the 2D system on the unaffected side of the face than the affected side of the face for all tasks?

Table 13
Descriptive Statistics for Distance (in mm) Collapsed for the Affected and Unaffected Side of the Face for 2D and 3D Measurement Systems by Task

	Minimum	Maximum	Mean	SD	Variance	Skewness
Smile						
2D						
Affected[a]	1.396	3.924	2.837	.799	.639	-.332
Unaffected[b]	3.290	11.585	6.517	2.440	5.955	.745
3D						
Affected[c]	1.057	11.579	4.161	2.728	7.440	1.321
Unaffected[c]	2.401	22.688	10.354	5.879	34.556	.432
Pucker						
2D						

Affected[d]	.728	7.953	4.653	1.600	2.559	-.420
Unaffected[c]	2.747	10.841	6.311	2.366	5.598	.331
3D						
Affected[c]	1.139	16.091	5.330	3.465	12.004	1.612
Unaffected[c]	2.016	12.691	7.955	3.176	10.084	-.452
Wi						
2D						
Affected[d]	.930	4.540	2.705	1.125	1.265	.155
Unaffected[c]	.966	9.342	3.879	2.198	4.832	.809
3D						
Affected[c]	1.116	11.407	5.059	3.097	9.592	.707
Unaffected[c]	1.078	14.601	7.071	3.858	14.888	.268
Ni						
2D						
Affected[d]	.880	5.960	2.320	1.454	2.114	1.026
Unaffected[c]	.855	8.734	3.126	2.236	5.002	1.465
3D						
Affected[c]	1.121	11.367	4.318	2.893	8.368	1.171
Unaffected[c]	1.169	17.743	7.884	4.317	18.638	.394
Jack & Jill						
2D						
Affected[d]	1.017	11.415	5.227	2.536	6.430	.925
Unaffected[c]	1.544	12.547	5.906	3.076	9.472	.645
3D						
Affected[c]	1.352	14.339	5.874	3.398	11.545	.911
Unaffected[c]	2.006	19.293	8.699	4.023	16.183	.733

[a]n=15. [b]n=16. [c]n=20. [d]n=19.

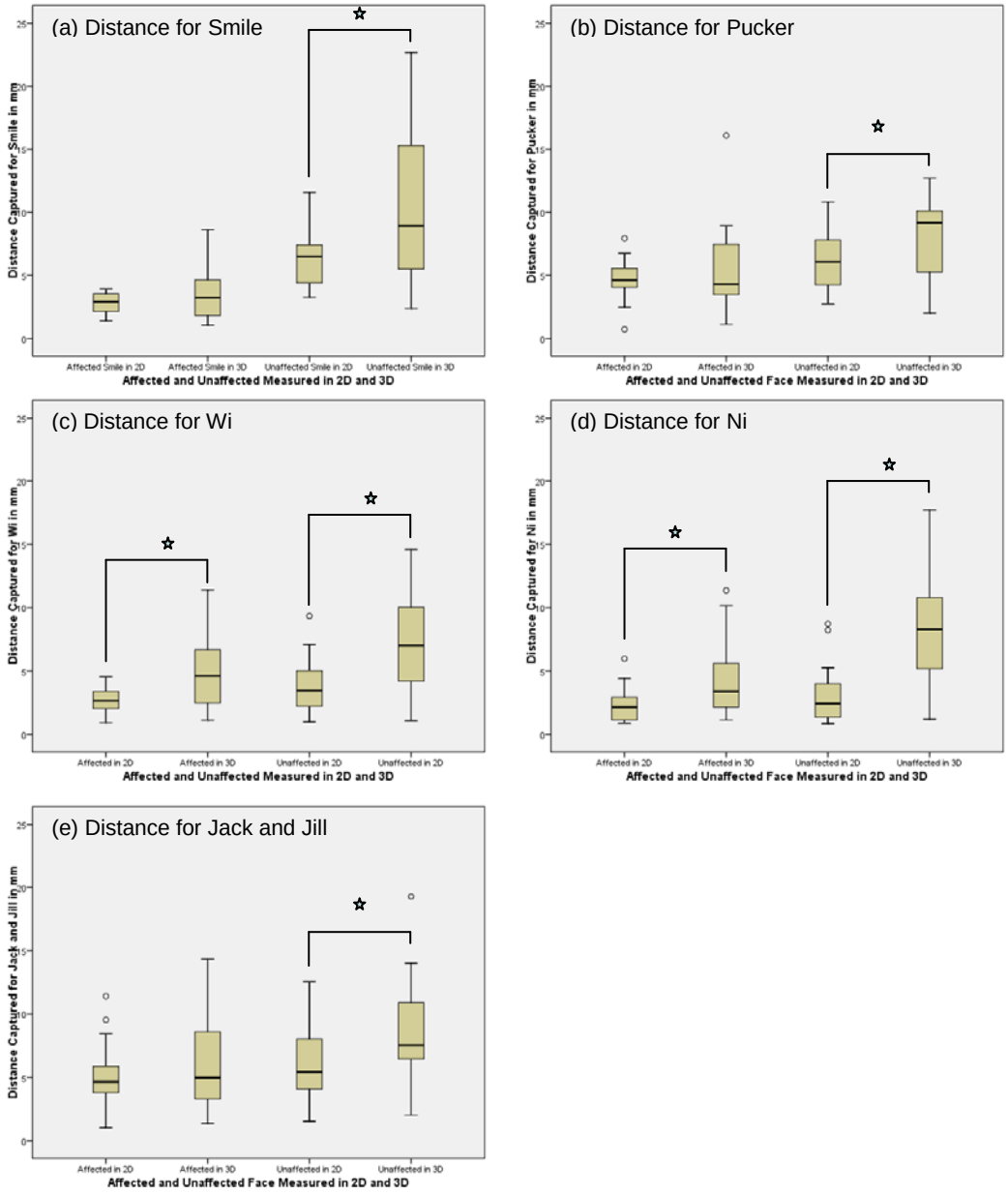

Figure 12. Box plots for distance (in mm) captured for the affected and unaffected Side of the face in 2D and 3D for all tasks: (a) Smile, (b) Pucker, (c) Wi, (d) Ni, and (e) Jack and Jill. *Significant differences ($p < .05$) between 2D and 3D measurement systems

The box plots in Figure 12 are similar to the box plots in Figure 5. The 3D measurement system captured greater displacement than the 2D system on the unaffected face compared to the affected face. The 3D system also captured a

greater range of distance for each task than the 2D system. All data in each condition were normally distributed. Although a significant difference was found between the 2D and 3D systems on the affected and unaffected face when each data set was collapsed across tasks, when the two systems were compared by task, a significant difference was found for all comparisons between the unaffected face, but not the affected face. Paired *t*-tests revealed a significant difference between the 2D and 3D systems on the unaffected face for smile [$t(15)= -3.240$, $p = .0025$, one-tailed], pucker [$t(19)= -2.598$, $p = .009$, one-tailed], wi [$t(19)= -4.017$, $p = .0005$, one-tailed], ni [$t(19)= -5.734$, $p \leq .0001$, one-tailed], and Jack and Jill [$t(19)= -5.785$, $p \leq .0001$, one-tailed]. For the affected side of the face, a significant difference between the 2D and 3D systems for wi [$t(18)= -4.330$, $p \leq .0001$, one-tailed], and ni [$t(18)= -4.825$, $p < .0001$, one-tailed] was found. There was no difference between the two systems on the affected face when smile, pucker, and Jack and Jill was performed; however, pucker [$t(19)= -2.598$, $p = .08$, one-tailed] showed a statistical trend.

Question 4b. Will the 3D measurement system be more sensitive to maximum velocity than the 2D measurement system on the unaffected side of the face than the affected side of the face for all tasks?

Table 14
Descriptive Statistics for Maximum Velocity (in mm/s) Collapsed for the Affected and Unaffected Side of the Face for 2D and 3D Measurement Systems by Task

	Min.	Max.	Mean	SD	Variance	Skewness
Smile						
2D	41.255	93.646	60.454	13.375	178.896	.884
Affected[a]	177.130	231.277	107.932	45.879	2104.925	1.370
Unaffected[b]						
3D	191.702	942.958	297.859	196.047	38434.52	2.359
Affected[c]	167.853	1353.159	362.253	316.0739	99902.05	2.309
Unaffected[c]						
Pucker						
2D						
Affected[d]	33.042	136.017	83.541	27.629	763.389	.420
Unaffected[c]	39.503	260.959	90.382	50.088	2508.835	2.210
3D						
Affected[c]	178.070	802.692	303.977	189.420	35879.79	1.774
Unaffected[c]	181.660	1382.372	356.875	280.281	78557.52	2.834
Wi						
2D						
Affected[d]	25.887	115.341	66.363	26.963	726.995	.412
Unaffected[c]	37.631	280.833	103.082	55.0135	3026.484	1.824
3D						
Affected[c]	179.953	591.845	271.282	128.504	16513.52	1.668
Unaffected[c]	183.213	810.652	342.196	192.412	37022.55	1.199
Ni						

2D						
Affected[d]	22.377	174.121	61.301	36.802	1354.385	2.068
Unaffected[c]	30.219	119.206	63.198	28.168	793.424	.605
3D						
Affected[c]	180.218	668.150	294.535	160.494	25758.56	1.535
Unaffected[c]	477.448	156.899	634.348	326.552	174.459	30435.99
Jack & Jill						
2D						
Affected[d]	35.272	231.357	106.696	53.191	2829.295	.952
Unaffected[c]	39.819	273.897	117.459	68.131	4641.779	.956
3D						
Affected[c]	184.879	881.600	323.575	215.245	46330.56	1.795
Unaffected[c]	165.983	1599.346	384.183	324.215	105115.2	3.017

[a]n=15. [b]n=16. [c]n=20. [d]n=19.

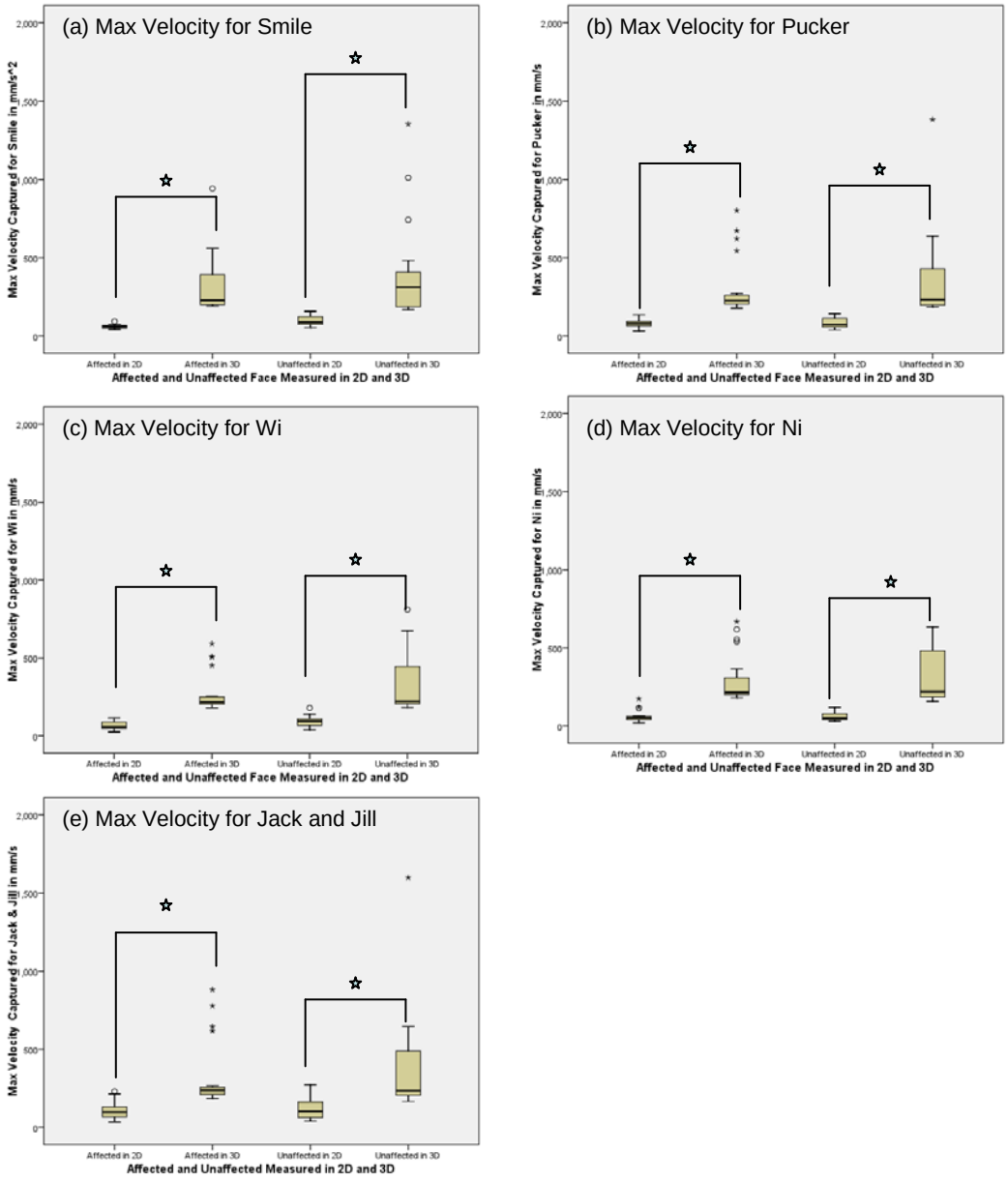

Figure 13. Box plots for maximum velocity (in mm/s) captured for the affected and unaffected side of the face in 2D and 3D for all tasks: (a) Smile, (b) Pucker, (c) Wi, (d) Ni, and (e) Jack and Jill. *Significant differences ($p < .05$) between 2D and 3D measurement systems

As can be seen by examining the box plots in Figure 13, the data followed a similar to pattern to the velocity comparisons. There was greater movement and range of movement captured by the 3D system compared to the 2D system. There did not seem to be a noticeable difference when comparing the affected and

unaffected side of the face with the 2D system. However, it was apparent that there was more movement captured on the affected face compared to the unaffected face when using the 3D system. All comparisons of the 2D and 3D measurement system for the affected and unaffected side of the face were significant ($p < .05$, one-tailed). Significant differences were still found when S5 was excluded from the data analysis.

Question 4c. Will the 3D measurement system be more sensitive to maximum acceleration than the 2D measurement system on the unaffected side of the face than the affected side of the face for all tasks?

Table 15
Descriptive Statistics for Acceleration (in mm/s^2) Collapsed for the Affected and Unaffected Side of the Face for 2D and 3D Measurement Systems by Task

	Minimum	Maximum	Mean	SD	Variance	Skewness
Smile						
2D						
Affected[a]	3.44×10^3	6.51×10^3	4.88×10^3	1.02×10^3	1.04×10^6	.448
Unaffected[b]	3.67×10^3	1.65×10^4	7.95×10^3	4.31×10^3	1.85×10^7	1.022
3D						
Affected[c]	4.49×10^4	2.99×10^5	8.15×10^4	6.27×10^6	3.93×10^9	2.637
Unaffected[c]	4.25×10^4	2.99×10^5	7.95×10^4	6.65×10^4	4.42×10^9	2.558
Pucker						
2D						
Affected[d]	3.09×10^3	1.37×10^4	6.13×10^3	2.68×10^3	7.18×10^6	1.591
Unaffected[c]	2.20×10^3	1.78×10^4	6.55×10^3	3.91×10^3	1.53×10^7	1.606
3D						
Affected[c]	4.81×10^4	1.79×10^5	7.83×10^4	4.55×10^4	2.07×10^9	1.611
Unaffected[c]	4.53×10^4	3.81×10^5	8.47×10^4	7.80×10^4	6.07×10^9	3.271

Wi						
2D						
Affected[d]	1.95×10^3	1.02×10^4	5.69×10^3	2.51×10^3	6.32×10^6	.452
Unaffected[c]	2.79×10^3	2.45×10^4	7.77×10^3	5.37×10^3	2.88×10^7	1.968
3D						
Affected[c]	4.90×10^4	1.50×10^5	7.09×10^4	3.26×10^4	1.07×10^9	1.644
Unaffected[c]	4.19×10^4	1.93×10^5	7.27×10^4	4.80×10^4	2.30×10^9	.512
Ni						
2D						
Affected[d]	2.13×10^3	1.16×10^4	5.15×10^3	2.52×10^3	6.35×10^6	1.572
Unaffected[c]	2.17×10^3	1.10×10^4	5.22×10^3	2.45×10^3	6.00×10^6	.769
3D						
Affected[c]	4.61×10^4	1.56×10^5	7.20×10^4	3.63×10^4	1.32×10^9	1.606
Unaffected[c]	4.10×10^4	1.78×10^5	6.97×10^4	4.07×10^4	1.65×10^9	1.798
Jack & Jill						
2D						
Affected[d]	2.83×10^3	1.62×10^4	7.57×10^3	3.76×10^3	1.42×10^7	1.161
Unaffected[c]	3.17×10^3	2.68×10^4	8.81×10^3	6.52×10^3	4.25×10^7	1.571
3D						
Affected[c]	4.84×10^4	1.70×10^5	7.81×10^4	4.48×10^4	2.00×10^9	1.560
Unaffected[c]	4.34×10^4	4.22×10^5	8.50×10^4	8.65×10^4	7.48×10^9	3.484

[a]n=15. [b]n=16. [c]n=20. [d]n=19.

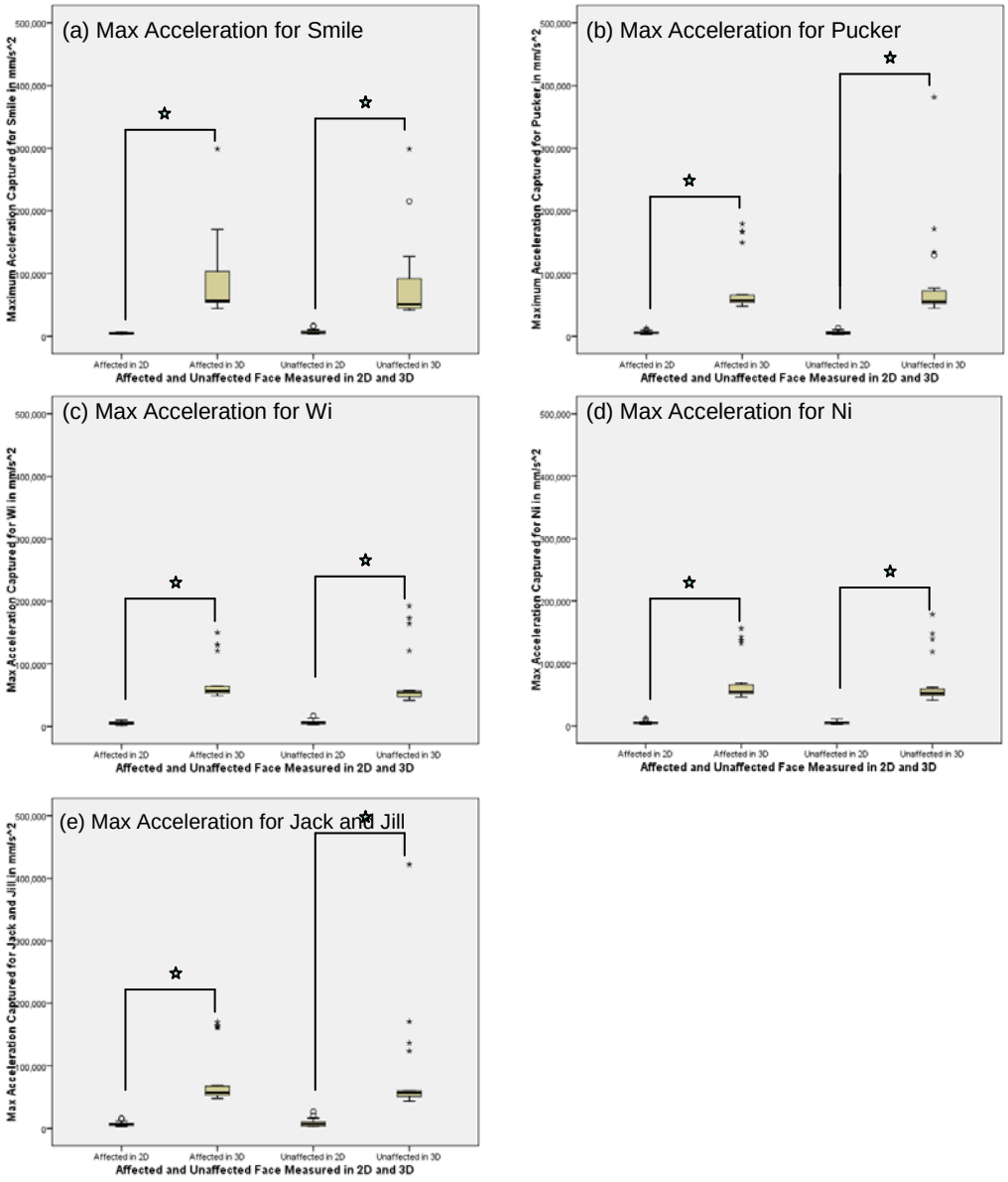

Figure 14. Box plots for maximum acceleration (in mm/s^2) captured for the affected and unaffected side of the face in 2D and 3D for all tasks: (a) Smile, (b) Pucker, (c) Wi, (d) Ni, and (e) Jack and Jill. *Significant differences ($p < .05$) between 2D and 3D measurement systems

The pattern displayed by each condition in Figure 14 was similar to that of the same variable shown in Figure 7. The box plots depicted in Figure 14 showed that, although there was a difference in the amount of acceleration measured between the 2D and 3D systems, there did not appear to be a difference in the amount of movement captured from the affected and unaffected side of the face. The acceleration values measured by the 2D system approached the lower limits of the scale, indicating a possible floor effect. The amount of movement captured appeared to be uniform for all tasks except for "smile." Figure 14a showed a large amount of movement captured by the 3D system on both sides of the face. Despite being able to capture greater amounts of acceleration, the 3D system also captured many more outliers. It is important to note that all outliers in the 3D data set were from a single subject. The 3D data were also positively skewed whereas the 2D data were normally distributed. A significant difference was found between the 2D and 3D measurement systems on the affected and unaffected side of the face ($p < .05$). A significant difference continued to exist after data from S5 were excluded in the analysis.

DISCUSSION

The evaluation of facial movement tracking systems is an important step towards establishing an evaluation protocol designed to assess outcomes following facial reanimation surgery. Over the years various techniques to measure facial movement have been developed including physical measurement tools (Mantkelow et al., 2008), 2D video-based systems (Isono et al., 1996; Linstrom, 2002; Neely et al., 1992; Rogers et al., 2007), 3D video-based systems (Coulson et al., 1999; Frey et al., 1999; Mishima et al., 2004; Trotman et al., 1997), laser scanning (O'Grady & Antonyshyn, 1999; Vannier et al., 1991), and optical tracking systems (Hontanilla & Auba, 2008; Kroo et al., 2002). The present study compared the use of a 2D video-based system with a 3D optical tracking system for measuring facial movements in patients who had undergone facial reanimation surgery. This is the first known study to have compared a 2D video-based system and a 3D optical tracking system on a clinical population. The purpose of this study was to (1) compare the amount of overall movement captured in the video-based system and optical tracking system and (2) compare facial movement captured across various tasks and (3) to compare movement on the affected and unaffected sides of face.

Results of the study showed that: (1) the 3D system was able to capture more movement than the 2D system for all variables (maximum distance, velocity, acceleration and area-volume ratios); (2) the 3D system consistently detected more movement than the 2D system across all tasks; (3) the 3D system was more sensitive than the 2D system for detecting movement on all dependent

variables assessed on the unaffected side of the face; and (4) the 2D and 3D systems were most similar for comparisons on the affected side of the face when maximum contraction and running speech tasks were performed. The clinical implication of these findings are as follows: (1) distance is the most valuable measure for evaluating facial paralysis, (2) maximal contraction and running speech tasks produce large facial movement; therefore, are recommended for the assessment of patients with facial paralysis, (3) area and volume ratios may be an appropriate measure for tracking changes in facial movement over time, (4) velocity and acceleration measures provide minimal information with regard to facial movement, and (5) the capabilities of the 2D system for analyzing facial paralysis is maximized when distance is measured during maximal contraction tasks and running speech.

Similar studies which have used 2D and 3D video-based systems in healthy subjects corroborate the current findings. Lin et al. (2000) and Gross et al. (1996) also found that their respective 3D systems were able to capture significantly more movement than their 2D systems. They reported that differences between their 2D and 3D systems were greatest for tasks that required the most movement in the anterior-posterior dimension and were the most similar for movements that produced the least amount of movement. Whereas the current study did not measure anterior-posterior dimensions directly, the patterns for detecting smaller and larger movements were similar to previous work (Lin et al. 2000; Gross et al., 1996). As shown with previous studies, the 3D tracking system used in the current study evaluated facial movement in a more adequate

manner perhaps because information about and accuracy of movement is enhanced when three dimensions are available rather than only two.

Comparing 2D and 3D

As in previous studies (Gross et al., 1996; Lin et al., 2000), the current comparisons of the 2D and 3D systems revealed that the 3D system captured significantly greater amounts of movement for all dependent variables than the 2D system when the data were collapsed across all tasks and when they were analyzed by task. Gross et al. (1996) discovered that their 3D video-based system was able to capture an upwards of 43% more amplitude than their 2D video-based system. The difference between the two systems is likely due to the "projection error," which is inherent to the 2D system. Projection error arises because 2D systems are unable to detect anterior-posterior movement of the facial landmarks (Gross et al., 1996). Furthermore, projection errors can also be produced when the frontal plane of the face and image plane of the camera sensors are misaligned (Giovanoli et al., 2003). Perspective distortion, poor spatial resolution, and non-linear lighting effects have also been listed as potential sources of error in 2D video-based systems (Kroos et al., 2002).

Comparing Dependent Variables in 2D and 3D

Four dependent variables were investigated in this study. Distance was the primary variable, whereas velocity, acceleration, and area and volume ratios were considered secondary variables due to the fact that most kinematic studies using a linear measurement system have reported distance measures only. Hontanilla and Auba (2008) noted the importance of using velocity in addition to

distance measures to obtain more information on facial movement and outcome after surgical repair. Acceleration was investigated in the present study as it may be an indicator of neuromuscular function, even though each distance-time derivation is susceptible to error propagation. Ratios were used in the current study to reflect facial asymmetry in a single value which can be easily understood. Additionally, these measures provide a method of evaluating facial movement independently of time. The secondary dependent variables added a more comprehensive understanding of facial movement in patients who have undergone facial reanimation surgery and served to document how 2D and 3D capture systems compared on measurements other than distance.

There was a notable absence in the use of velocity and acceleration measures in previous 3D kinematic studies (Fager, Green, & Nip, 2006; Hontanilla & Auba, 2008; Sawyer, See, & Nduka, 2010). Like studies utilizing 2D systems, studies of facial kinematics using 3D systems primarily report distance measures. Hontanilla & Auba (2008) and Sawyer et al. (2010) reported on velocity measures using 3D video systems with healthy subjects. Fager et al. (2006) used maximum velocity measures in a single subject with velopharyngeal incompetency following brainstem impairment. None of these studies commented on the relation of marker velocity to facial movement. Previous studies reviewed have not used acceleration measurements in facial motion analyses in healthy subjects or in a clinical population using a 3D measurement system. Area values have been reported in 3D (Hontanilla & Auba, 2008), but the

method in which these values are derived are not mentioned. Again, the authors did not elaborate on the relationship between area values and facial movement.

The velocity and acceleration data from this study cannot be compared to data from other studies as velocity and acceleration have not been used to measure functioning clinical population. In the present study, one trend that readily became apparent when comparing the 2D and 3D measurement system was the difference in movement captured across distance, velocity, and acceleration measures. Interestingly, the differences between the two systems were most significant for velocity and acceleration measurements. It is likely that the differences between the two systems that were inherent when measuring distance were magnified with each distance-time derivation.

Unlike the velocity and acceleration variables which are functions of time, area and volume ratios compare bilateral landmarks on the unaffected side of the face to the affected side of the face. A group of researchers from Japan calculated areas ratio by comparing area at maximum movement to the area at rest on both the affected and unaffected sides of the face (Isono et al., 1994). The authors found that while normal subjects showed little asymmetry, marked asymmetry was found in patients with facial palsy.

In the present study, each ratio value inherently reflected the relationship between the affected and unaffected side of the face. A visual representation of the maximum area traveled by each marker can be found in Figure 15 in *Appendix 3*. However, the primary focus of this study was to document the ability of each measurement system to capture different variables. Because area and volume

values are derived from absolute distances, the variability is reduced. The variability is further reduced when the area and volume values are converted into ratios. Since there is naturally less variability in the ratio scores than the distance values, the ratio values may be a more sensitive measure to catalogue change in muscle movement over time.

The outlying data points in the 3D velocity and acceleration data were another salient result which arose from comparing the 2D and 3D measurement systems. Analysis of the outlying points showed that they belonged to a single subject, S5. Interestingly, outliers in the distance and volume ratio data measured by the 3D system were unremarkable. S5 data may be unique as this subject had the longest time between the last surgery and study participation. Whereas S5 had 2 years and 11 months for recovery before participating in the study, the other subjects had between 3 months and 1 year and 5 months. Refer to *Appendix 2* for further demographic information of the subjects. The sizeable difference in the efficiency of movement between S5 and the others may be due to better neuromotor connections which have been established through repetitive muscle use. The ability of the 3D system to capture this difference may indicate that maximum velocity and acceleration measures can be used as indices of neuromotor efficiency.

Comparing Tasks in 2D and 3D

Most studies which have looked at surgical outcomes in patients with facial paralysis have used patient report or subjective rating scales. Outcome studies using quantitative measurement systems are notably lacking. Interviews

and questionnaires revealed that patients rate the outcome of their facial reanimation surgery as good or excellent, feel the surgery was worthwhile, and report improved appearance and self-esteem (O'Brien, Pederson, Khazanchi, Morrison, MacLeod, & Kumar, 1990). In another study using subjective methods, Goldberg et al. (2003) reported that overall, patients had a decreased frequency of compensatory errors, but 83% of patients who underwent facial reanimation surgery still reported occasional or frequent speech problems following their surgery.

Besides interviews, functional and esthetic results also have been evaluated using subjective scoring systems, such as the House-Brackmann Facial Nerve Grading Scale (HBGS). The HBGS grades global facial function from a score of I (normal) to VI (no movement) based on facial symmetry. Studies have found that patients improve to a grade of III or IV after surgery (Manni, Beurskens, van de Velde, & Stokroos, 2001; Yla-Kotola et al., 2004). Although these studies have shown that there is improvement in function and esthetics following facial reanimation surgery through subjective methods, the need for a quantitative evaluation of surgical outcomes using speech and non-speech tasks is evident.

The present study was the first to investigate the differences between the 2D and 3D measurement systems to compare movements across three different behavioural tasks including: maximal muscle contractions, speech tokens, and running speech. Traditionally, studies have used either maximal contraction tasks or speech tasks, but not both in kinematic studies. Maximal contraction tasks

have been used to establish reliability and sensitivity of various measurement systems (Trotman et al., 1997; Frey et al., 1999; Hontanilla & Auba, 2008) or to grade facial nerve function (Johnson et al., 1994; Isono et al., 1996; Meier-Gallati et al., 1998). On the other hand, speech tasks have been more frequently used in the field of speech animation synthesis and auditory-visual speech (Cao, Faloutsos, Kohler, & Pighin, 2004; Cao, Tien, Faloutsos, & Pighin, 2005; Kroos et al., 2002).

Tasks used to investigate the lower half of the face in other studies include: maximal smile, maximal frown, nose wrinkling, lip purse, grimace, and cheek puff (Meier-Gallati et al., 1998; Johnson et al., 1994, Trotman et al., 1998). Articulatory movement has been investigated in a patient fitted with a palatal lift appliance (Fager et al., 2006) and in patients with dysarthria (Yunusova, Weismer, Westbury, & Lindstrom, 2008). Fager et al. (2006) used an optical tracking system to measure movement of the upper and lower lip at midline using VCV stimuli. On the other hand, Yunusova et al. (2008), used a X-ray microbeam technique to measure tongue, lip, and jaw movement in patients with amyotrophic lateral sclerosis and Parkinson's disease. Speech tasks have not yet been used to assess facial motion in patients who have undergone facial reanimation surgery.

Recently, the utterance "puppy" was compared to a standard smile expression measure facial movement using the *3dMDface Dynamic System,* a 3D measurement system which utilizes infra-red speckle projection to capture both pattern-projected and non-pattern projected white-light images simultaneously

(Popat, Richmond, Playle, Marshall, Rosin, & Cosker, 2008). Unlike the current study, this study did not employ a linear measurement system nor were the participants from a clinical population. Nonetheless, it is valuable to note that the authors concluded that "puppy" was a more appropriate measure of facial movement when compared with the smile expression as it created more stable points on the face. Stated differently, their data could also be interpreted to mean that there were more stable points for "puppy" since it is unable to produce as much facial movement as a smile. Others kinematic studies have supported the use of maximal contraction tasks to measure nerve function as they allow more accurate reproductions compared to moderate expressions (Johnston, Millett, & Ayoub, 2003).

In the current study, the 2D system captured noticeably more maximum distance for smile, pucker, and running speech than the speech tokens. Although the differences between tasks were not quite as prominent in 3D, there appeared to be more movement captured for smile and running speech compared to movements associated with the production of CV tokens. This confirms previous findings from the pilot work conducted on the same participants using 2D kinematics (Harasem, 2008). Similarly, area and volume ratios for smile appeared larger when compared to other tasks. This is likely because smile animation produces the largest and most consistent movements in circumoral markers (Trotman et al., 1998) as well as the greatest excursion in the anterior-posterior direction (Gross et al., 1996). Taken together, the results of the current

and previous studies indicate that maximal contraction and running speech tasks are both important for the evaluation of facial paralysis.

Comparing the Affected and Unaffected Side of the Face

Several researchers have compared the affected and unaffected side of the face in kinematic studies to determine facial asymmetries by using ratio or percent values. In studies employing 2D methods, facial asymmetries have been investigated using linear measurement and subtraction techniques. Isono et al. (1996) summed the distance traveled by 10 markers on the affected side of the face and compared it to the sum of the same landmarks on the normal side of the face. Linstrom et al. (2002) proposed a method in which the difference in marker displacement on the affected and unaffected side was divided by total displacement of that marker to calculate asymmetry relative to the total amount of displacement. A group of researchers from Switzerland used a subtraction technique to develop various indices to compare facial asymmetry (Scriba, Stoeckli, Veraguth, Pollak, & Fisch, 1999). They developed a regional symmetry index (RSI), which was determined by comparing the percentage of change in luminance for each facial area on the weaker side to the better side, and a left-right index (LRI), which was calculated by "adding the changes of luminance of all three facial areas during maximal movement on the left side and relating it to the sum of the changes on the right side, taking the latter as 100%" (p.642). Facial asymmetries have yet to be studied in 3D as this technology has not yet been applied to this clinical population. The literature did not show wide-spread use of any of these methods.

Quantitative evaluation of the affected and unaffected side of the face is important to track changes in facial nerve function over time and to identify finer grades of nerve dysfunction such as contractures, hemispasms, and synkinesis. The best use of percentages and ratios would be for unilateral nerve repairs; however, these measures would be capable of reflecting improvement in bilateral repairs as well.

Comparing Dependent Variables on the Affected and Unaffected Face

In the present study, the movements of all markers on the affected and unaffected side of the face were used to investigate the difference in sensitivity of the 2D video-based system and 3D optical tracking system. Presumably, more movement would be present on the unaffected side of the face, which was confirmed by the results of this study. However, when velocity and acceleration were compared, there was little difference between the two measures on the affected and unaffected side of the face for both 2D and 3D measurement systems. The 2D measures of velocity and acceleration were less effective than the 3D measures for differentiating the affected and unaffected sides of the face. Clinically, this would imply that distance measures may be more appropriate for the evaluation of unilateral movements.

Comparing Tasks on the Affected and Unaffected Face

As mentioned earlier, both the 2D and 3D systems were capable of detecting more movement on the unaffected side of the face than the affected side of the face. The greatest amount of movement was captured for “smile” on the unaffected side of the face. Curiously, comparisons of the affected and unaffected

face showed some unexpected findings. First of all, differences between the two measurement systems on the affected side of the face were only significant for the isolated speech tasks, which produce less movement than maximal contraction and running speech tasks (Harasem, 2008). Based on findings from Gross et al. (1996), the difference between 2D and 3D systems should be greatest for maximal contraction tasks. As this was not the case, it is possible that the unaffected side of the face may be countering the movement on the affected side of the face for tasks which require more effort to produce. The extra effort placed on the muscles on the affected side of the face may have interfered with their functioning. Secondly, there were minimal differences in the amount of movement captured across tasks when maximum velocity and acceleration were used as measurement variables. This implies that these measures were not sensitive to differences between tasks because each marker reaches a similar maximum velocity or acceleration during a task. The clinical implications of these findings are: (1) even though the 3D system does not differ in its ability to detect movement on the affected side of the face, it is the preferred system of measurement because it is able to capture a greater amount of movement overall; (2) the 3D measurement system is preferred when measuring facial movements which require less anterior-posterior such as isolated speech tasks; and (3) distance measure derived from either 2D or 3D measurement systems are most sensitive for detecting asymmetries in facial movement.

Limitations of the Measurement Systems

There were several limitations associated with both measurement systems. Although both systems were capable of collecting kinematic data, neither provided information regarding the direction or angle of the trajectory of movement. This information would be useful for discriminating biomechanically correct movement from abnormal movement. However, graphing maximal displacement of each marker from rest may aid in distinguishing the two kinds of movement. Refer to Figure 16 in *Appendix 3*.

One of the biggest advantages of the 2D system was its ease of use and its cost effectiveness. The operator was only required to have basic knowledge of a home video camera and accompanying video editing and analyzing software. Along with these advantages, the 2D system employed in this study faced several limitations. First of all, although the 2D video system was capable of accounting for head movement in the frontal plane, it was unable to accommodate head rotations. Misalignment between the frontal plane of the face and the image plane of the camera sensor can introduce error into all subsequent calculations. Secondly, all calculations were made on the assumption that the subject was at resting position for the first video frame of each task. Error can also be introduced during the automatic and manual tracking of markers through each video frame.

The major advantage of the 3D optical system lies in its ability to track dynamic facial movement accurately and reliably. The data can be filmed at any time and analyzed at a later date. The downside of this system is the expense

associated with it. One of the main issues with facial tracking in 3D was the LED markers placed on the face. Physical markers placed on the face may have interfered with facial motion. Also, head movement may disrupt the signal given off by the LED, as it can be disturbed as it passes through soft tissue.

Conclusion

The current study investigated the ability of a 2D video-based system and a 3D optical tracking system to capture movement across various dependent variables and tasks on the affected and unaffected side of the face in patients who had undergone facial reanimation surgery. The primary finding was that distance measurements derived from either technique were more sensitive to detecting movement than acceleration, velocity or area/volume measurements. Additionally, this study showed that area and volume ratios may be a valuable measure for tracking changes in the paralyzed face across time. In terms of tasks, the results from the current study demonstrated the importance of using both running speech and maximal contraction tasks for assessing patients with facial nerve impairment pre- and post-surgery. Finally, the results indicated that 3D optical tracking was superior to the 2D video-based system for measuring distance, velocity, acceleration, and area and volume ratios across maximal contraction, isolated speech, and running speech tasks. Even though the 3D system was more sensitive than the 2D system overall, the 2D system was capable of detecting differences in facial movement, especially when capturing facial movement using the distance variable across maximal contraction tasks and running speech. For this reason, a 2D video-based system would be an

appropriate alternative for centres that do not have access to a 3D system. Further development of measurement variables and behavioural tasks are required to provide a comprehensive assessment protocol for patients who have undergone facial reanimation surgery.

Limitations of the Study

The number of participants (n = 5) in this study was its greatest limitation. The small sample had an effect on the statistical power of the study. In order to replicate and increase the statistical power of the results, more participants are required. In addition, the current sample was not homogeneous so the ability to generalize to subgroups of individuals with facial paralysis is limited. Comparisons of the affected and unaffected side of the face revealed important differences between the way in which the 2D and 3D systems captured motion. However, the definition for the affected and unaffected side of the face in this study also was limited. In theory, the division of affected and unaffected face would be most informative for patients with unilateral paralysis. However, in this study there was one subject, S3, who had bilateral facial paralysis. At the time of testing, he had completed the first stage of masseteric nerve surgery unilaterally. The side on which the operation was completed was considered to be the unaffected for the purpose of this study. A ratio value of the operated side and the paralyzed face may have confounded the results by producing larger values, as can be seen in Figure 28. However, because this was done consistently in the 2D and 3D data, it would not reflect in the measurement systems' ability to track

facial motion. Lastly, marker placement and measurement reliability studies are required.

Future Research

Currently, the House-Brackmann Grading System (HBGS), a subjective global rating system is the gold standard for measuring for evaluating facial nerve palsy (House & Brackmann, 1985). The HBGS system has been criticized as it cannot account for finer grades of dysfunction, detect small changes in the paralytic face, and has lower inter-rater reliability in the middle ranges of nerve dysfunction (Kang et al., 2002). The development and application of an objective and universally accepted measurement system is essential to document the degree of facial paralysis and to provide a standardized outcome measure of surgical success. Also, future research should be designed to identify speech and non-speech tasks that will best predict functional outcomes following facial reanimation surgery. Furthermore, the motion data could be used to establish 3D facial models which would aid in the assessment, treatment, and tracking outcomes of surgical reconstruction of the face in patients with facial nerve involvement but also for various other orthodontic and craniofacial procedures. The establishment of 3D facial models would also be invaluable in the field of facial animation and audio-visual speech processing.

REFERENCES

Bajaj-Luthra, A., Mueller, T., & Johnson, P. C. (1997). Quantitative analysis of facial motion components: Anatomic and nonanatomic motion in normal persons and in patients with complete facial paralysis. *Plastic & Reconstructive Surgery*, *99*(7), 1894-1902.

Bajaj-Luthra, A., VanSwearingen, J., Thornton, R. H., & Johnson, P. C. (1998). Quantitation of patterns of facial movement in patients with ocular to oral synkinesis. *Plastic & Reconstructive Surgery*, *101*(6), 1473-1480.

Beurskens, C.H. & Heymans, P.G. (2003). Positive effects of mime therapy on sequelae of facial paralysis: stiffness, lip mobility, and social and physical aspects of facial disability. *Otology & Neurotology*, 24, 677-81.

Bleicher, J.N., Hamiel, S., & Gengler, J.S. (1996). A survey of facial paralysis: etiology and incidence. *Ears Nose Throat Journal*, 75, 355-357.

Brach, J.S., VanSwearingen J.M., Lennert, J., & Johnson, P.C. (1997). Facial neuromuscular retraining for oral synkinesis. *Plastic Reconstructive Surgery*, 99,1922-1931.

Burres, S. & Fisch, U. (1986). The comparison of facial grading systems. *Archives of Otolaryngology -- Head & Neck Surgery, 112, 753-758.*

Coulson, S. E., Croxson, G. R., & Gilleard, W. L. (1999). Three-dimensional quantification of 'still' points during normal facial movement. *Annals of Otology, Rhinology and Laryngology*, *108*(3), 265-268.

Coulson, S. E., Croxson, G. R., & Gilleard, W. L. (2000). Quantification of the three-dimensional displacement of normal facial movement. *Annals of Otology, Rhinology and Laryngology, 109*(5 I), 478-483.

Danner, C.J. (2008). Facial nerve paralysis. *Otolaryngologic Clinics of North America*, 41, 619-632.

de Swart, B.J., Verheij, J.C., & Beurskens, C.H. (2003). Problems with eating and drinking in patients with unilateral peripheral facial paralysis. *Dysphagia, 18, 267-73.*

Devriese, P.P. (1998). Treatment of sequelae after facial paralysis: a global approach. *The Journal of Laryngology and Otology*, 112, 429-431.

Ekman, P. (1986). Psychosocial aspects of facial paralysis. In: M. May & B.M. Schaitkin (Eds.), *The Facial Nerve* (pp.781-787). New York: Thieme Inc., 781-7.

Fager, S.K., Green, J.R., Nip, I., & Hakel, M. (2006). Facial kinematics of a speaker with a palatal lift following brainstem impairment. *Journal of Medical Speech-Language Pathology, 14*(4), 235-240.

Ferrario, V. F., Sforza, C., Schmitz, J. H., Miani, A.,Jr, & Serrao, G. (1998). A three-dimensional computerized mesh diagram analysis and its application in soft tissue facial morphometry. *American Journal of Orthodontics & Dentofacial Orthopedics, 114*(4), 404-413.

Frey, M., Giovanoli, P., Gerber, H., Slameczka, M., & Stussi, E. (1999). Three-dimensional video analysis of facial movements: A new method to assess

the quantity and quality of the smile. *Plastic & Reconstructive Surgery, 104*(7), 2032-2039.

Frey, M., Jenny, A., Giovanoli, P., & Stussi, E. (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*(7), 1334-1349.

Frey, M., Giovanoli, P., Gerber, H., Slameczka, M., & Stussi, E. (1999). Three-dimensional video analysis of facial movements: a new method to assess the quantity and quality of smile. *Plastic and Reconstructive Surgery, 104*(7), 2032-2039.

Giovanoli, P., Tzou, C. H., Ploner, M., & Frey, M. (2003). Three-dimensional video-analysis of facial movements in healthy volunteers. *British Journal of Plastic Surgery, 56*(7), 644-652.

Goldblatt, D., & Williams, D. (1986). "I an sniling!" Mobius' syndrome inside and out. *Journal of Child Neurology*. 1, 71-78.

Gillberg, C. (1992). Subgroups in autism: there behavioral phenotypes typical of underlying medical conditions? *Journal of Intellectual Disability Research*, 36, 201-214.

Goldberg, C., DeLorie, R., Zuker, R.M., & Mantkelow, R.T. (2003). The effects of gracilis muscle transplantation on speech in children with Moebius syndrome. *The Journal of Craniofacial Surgery*, 14, 687-690.

Gross, M. M., Trotman, C. A., & Moffatt, K. S. (1996). A comparison of three-dimensional and two-dimensional analyses of facial motion. *Angle Orthodontist, 66*(3), 189-194.

Guntinas-Lichius, O., Streppel, M., & Stennert, E. (2006). Postoperative functional evaluation of different reanimation techniques for facial nerve repair. *The American Journal of Surgery*, 191, 61-67.

Guerrissi, J.O. (1991). Selective myectomy for postparetic facial synkinesis. *Plastic Reconstructive Surgery*, 87, 459-466.

Harasem, M.A. (2008). Physiological and functional outcomes in patients with facial paralysis following facial feanimation surgery. Unpublished master's thesis, University of Alberta, Edmonton, Canada.

Harrison, D.H. (2005). Surgical correction of unilateral and bilateral facial palsy. *Postgraduate Medical Journal*, 81, 562-567.

House, J.W. & Brackmann, D.E. (1985). Facial nerve grading system. *Otology & Neurotology*, *93*, 146-147.

Hontanilla, B. & Auba, C. (2008). Automatic 3D quantitative analysis for evaluation of facial movement. *Journal of Plastic, Reconstructive, and Aesthetic Surgery*, 61, 18-30.

Ikeda, M., Nakazato, H., Hiroshige, K., Abiko, Y., & Sugiura, M. (2003). To what extent do evaluations of facial paralysis by physicians coincide with self-evaluations by patients: comparison of the Yanagihara method, the House-

Brackmann method, and self-evaluation by patients. *Otology & Neurotology*, 24(2), 334-8.

Isono, M., Murata, K., Tanaka, H., Kawamoto, M., & Azuma, H. (1996). An objective evaluation method for facial mimic motion. *Otolaryngology - Head & Neck Surgery*, *114*(1), 27-31.

Isono, M., Murata, K., Tanaka, H., Minoyama, M., & Azuma, H. (1994). Computerized analysis of facial motions: Objective evaluation of facial palsy. *Otolaryngology Japan*, 97, 393-400.

Johnson, P.J., Bajaj-Luthra, A., Llull, R., & Johnson, P.C. (1997). Quantitative facial motion analysis after functional free muscle reanimation procedures. *Plastic & Reconstructive Surgery*, 100, 1710-1719.

Johnson, P.C., Brown, H., Kuzon, W.M., Balliet, R., Garrison, J.L., & Campbell, J. (1994). Simultaneous quantitation of facial movements: the maximal response assay of facial nerve function. *Annals of Plastic Surgery*, 32, 171-179.

Johnston, D., Millett, D.T., & Ayoub, A.F. (2003). Are facial expressions reproducible? *Cleft Palate- Craniofacial Journal, 40* (3), 291-296.

Kang, T. S., Vrabec, J. T., Giddings, N., & Terris, D. J. (2002). Facial nerve grading systems (1985-2002): Beyond the House-Brackmann scale. *Otology & Neurotology, 23*(5), 767-771.

Kroos, C., Kuratate, T., & Vatikiotis-Batesion, E. (2002). Video-based motion measurement. *Journal of Phonetics, 30,* 569-590.

Lin, S. C., Chiu, H. Y., Ho, C. S., Su, F. C., & Chou, Y. L. (2000). Comparison of two-dimensional and three-dimensional techniques for determination of facial motion--absolute movement in a local face frame. *Journal of the Formosan Medical Association, 99*(5), 393-401.

Linstrom, C.J. (2002). Objective facial motion analysis in patients with facial nerve dysfunction. *Laryngoscope, 112*, 1129-1147.

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

MacGregor F.C. (1990). Facial disfigurement: problems and managementof social interaction and implications for mental health. *Aesthetic Plastic Surgery*, 14, 249-57.

Manktelow R.T., Zuker, R.M., & Tomat, L.R. (2008). Facial paralysis measurement with a handheld ruler. *Plastic & Reconstructive Surgery, 121*(2):435-442.

Manni, J.J., Beurskens, C.H.G., van de Velde, C., & Stokroos, R.J. (2001). Reanimation of paralyzed face by indirect hypoglossal-facial nerve anastomosis. *The American Journal of Surgery, 182,* 268-273.

May, M. (2000). Muscle transposition techniques: temporalis, masseter, and digastric. In: M. May & B.M. Schaitkin (Eds.), *The Facial Nerve* (2nd ed., pp.635-664). New York: Thieme Inc.

Meier-Gallati, V., Scriba, H., & Fisch, U. (1998). Objective scaling of facial nerve function based on area analysis (OSCAR). *Otolaryngology - Head & Neck Surgery, 118*(4), 545-550.

Mishima, K., Yamada, T., Fujiwara, K., & Sugahara, T. (2004). Development and clinical usage of a motion analysis system for the face: Preliminary report. *Cleft Palate-Craniofacial Journal, 41*(5), 559-564.

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 & Neck Surgery*, 110, 156-161.

Nakamura, K., Toda, N., Sakamaki, K., Kashima, K., & Takeda, N. (2003). Biofeedback rehabilitation for prevention of synkinesis after facial palsy. *Otolaryngology - Head & Neck Surgery*, 128, 539-43.

Neely, J. G., Cheung, J. Y., Wood, M., Byers, J., & Rogerson, A. (1992). Computerized quantitative dynamic analysis of facial motion in the paralyzed and synkinetic face. *American Journal of Otology, 13*(2), 97-107.

Neely, J. G., Joaquin, A. H., Kohn, L. A., & Cheung, J. Y. (1996). Quantitative assessment of the variation within grades of facial paralysis. *Laryngoscope, 106*(4), 438-442.

O'Brien, B.M., Pederson, W.C., Khazanchi, R.K., Morrison, W.A., MacLeod, A.M., & Kumar, V. (1990). Results of management of facial palsy with

microvascular free-muscle transfer. *Plastic and Reconstructive Surgery*, 86, 12-22.

O'Grady, K. F., & Antonyshyn, O. M. (1999). Facial asymmetry: Three-dimensional analysis using laser surface scanning. *Plastic & Reconstructive Surgery*, *104*(4), 928-937.

Popat, H., Richmond, S., Playle, R., Marshall, D., Rosin, P.L., & Cosker, D. (2008). Three-dimensional motion analysis- an exploratory study. Part 1: assessment of facial movement. *Orthodontic & Craniofacial Research, 11*, 216-223.

Ramsey M.J., DerSimonian R., Holtel M.R., & Burgess L.P. (2000). Corticosteroid treatment for idiopathic facial nerve paralysis: a meta-analysis. *Laryngoscope, 110*, 335-341.

Rogers, C.R., Schmidt, K.L., VanSwearingen, J.M., Cohn, J.F., Wachtman, G.S., Manders, E.K., & Deleyiannis, F.W. (2007). Automated facial image analysis: Detecting improvement in abnormal facial movement after treatment with Botulinum Toxin A. *Advances in Plastic Surgery, 58* (1), 39-47.

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

Sargent, E.W., Fadhli, O.A., & Cohen, R.A. (1998). Measurement of facial movement with computer software. *Archives of Otolaryngology -- Head & Neck Surgery, 124*(3), 313-318.

Sawyer, A.R., See, M., & Nduka, C. (2010). Quantitative analysis of normal smile with 3D stereophotogrammetry: an aid to facial reanimation. *Journal of Plastic, Reconstructive, & Aesthetic Surgery*, 63, 65-72.

Schaitkin, B.M. & May, M. (2000). Reporting recovery of facial function. In: M. May & B.M. Schaitkin (Eds.), *The Facial Nerve* (2nd ed., pp. 275-299). New York: Thieme Inc.

Scriba, H., Stoeckli, S.J., Veraguth, D., Pollak, A., & Fisch, U. (1999). Objective evaluation of normal facial function. *Annals of Otology, Rhinology and Laryngology, 108*, 641-644.

Secil, Y., Aydogdu, I., Ertekin, C. (2002). Peripheral facial palsy and dysfunction of the oropharynx. *Journal of Neurology, Neurosurgery & Psychiatry, 72*(3), 391-393.

Shindo, M. (1999). Management of facial nerve paralysis. *Savlivary Gland Diseases, 32, 945-964.*

Sjogreen, L., Andersson-Norinder, J., & Jacobsson, C. (2001). Development of speech, feeding, eating, and expression in Mobius sequence. *International Journal of Pediatric Otorhinolaryngology*, 60, 197-204.

Streppel, M., Angelov, D.N., Gutinas-Lichius, O., & Neiss, W. (1998). Nerve-plasty interventions on the facial nerve in the elderly patient-

morphological evaluation of disappointing functional results. *Laryngorhinootologie*, 77, 332-336.

Tate, J.R. & Tollefson, T.T. (2006). Advances in facial animation. *Current Opinion in Otolaryngology- Head and Neck Surgery*, 14, 242-248.

Tomat, L. R., & Manktelow, R. T. (2005). Evaluation of a new measurement tool for facial paralysis reconstruction. *Plastic & Reconstructive Surgery, 115*(3), 696-704.

Trotman, C.A., Stolhler, C.S., & Johnston L.E. (1997). Measurement of facial soft tissue mobility in man. *Cleft Palate- Craniofacial Journal, 35* (1), 16-25.

Trotman, C.A., Faraway, J.J., Silvester, K.T., Greenlee, G.M., & Johnston L.E. (1998). Sensitivity of a method for the analysis of facial mobility: I. Vector of displacement. *Cleft Palate- Craniofacial Journal, 35* (2), 132-141.

Trotman, C.A., Faraway, J.J., & Essick, G.K. (2000). Three-dimensional nasolabial displacement during movement in repaired cleft lip and palate patients. *Plastic and Reconstructive Surgery, 105 (4)*, 1273-1283.

Twerski AJ, & Twerski B. The emotional impact of facial paralysis. In: May M, editor. The facial nerve. New York: Thieme Inc.;1986. p. 788-94.

Vannier, M.W., Pigram, T.K., Bhatia, G., Brunsden, B., & Commean, P. Facial surface scanner. *IEEE Computer Graphics and Applications*, 11, 72-80.

VanSwearingen, J. M., Cohn, J.F., Turnbull, J., Mrzai, T., & Johnson, P. (1998). Psychological distress: linking impairment with disability in facial neuromotor disorders. *Otolaryngology - Head & Neck Surgery, 118*(6), 790-796.

Vatikiotis-Bateson, E. & Kelso, J.A.S. (1993). Rhythm type and articulatory dynamics in English, French, and Japanese. *Journal of Phonetics, 21*, 231-265.

Verzijl, H.T., van der Zwaag, B., Cruysberg, J.R., & Padberg, G.W. (2003). Mobius syndrome redefined: a syndrome of rhombencephalic maldevelopment. *Neurology,* 61, 327-333.

Wachtman, G. S., Cohn, J. F., VanSwearingen, J. M., & Manders, E. K. (2001). Automated tracking of facial features in patients with facial neuromuscular dysfunction. *Plastic & Reconstructive Surgery, 107*(5), 1124-1133.

Yla-Kotola, T.M., Kauhanen, M.S., & Asko-Seljavaara, S.L. (2004). Facial reanimation by transplantation of a microneurovascular muscle: long-term follow-up. *Scandinavian Journal of Plastic & Reconstructive Surgery & Hand Surgery*, 38(5), 272-276.

Yuen, K., Inokuchi, I., Maeta, M., Kawakami, S., & Masuda, Y. (1997). Evaluation of facial palsy by moiré topography index. *Otolaryngology - Head & Neck Surgery*, 117, 567-572.

Yunusova, Y., Weismer, G., Westbury, J.R., & Lindstrom, M.J. (2008). Articulatory movements during vowels in speakers with dysarthria and healthy controls. *Journal of Speech, Language, and Hearing Research*, 51, 596-611.

APPENDIX 1: OPERATIONAL DEFINITONS

Table 16
Operational Definitions

Term	Definition
Dense flow extraction	an algorithm for detecting motion of the overlying skin produced by muscle contractions
Direct linear transformation	an algorithm that solves a set of variables from a set of similarity relations
Feature-based tracking	a module in Automated Facial Analysis that is able to track markers using a Lucas-Kanade algorithm and optical flow techniques
High-gradient component detection	a means of detecting wrinkles and furrow features
Image motion estimation	recovering 3D motion information from a 2D image
Linear measuremet system	a measurement system that relies on the displacement of predetermined anatomical landmarks to track motion
Lucas-Kanade algorithm	a two-frame differential method to estimate the deformations between two image frames
Optical flow technique	a type of motion detection that uses the pattern of apparent motion of objects, surfaces, and edges in relation to the observer and the scene

Subtraction technique	a global assessment method of facial movement that relies on changes in either pixels or luminance from repose to maximal movement during a voluntary facial expression
Template-matching technique	an image processing technique that matches small parts of an image with a template image using brightness information
Thresholding	a type of image-processing used on a grey-scale image to create binary images. During this process, individual pixels with a value greater than the threshold value is given a value of "1" ("object" pixels) whereas pixels below the threshold value is given a value of "0" ("background" pixels), if the object is brighter than the background.

APPENDIX 2: DEMOGRAPHIC INFORMATION FOR ALL SUBJECTS

Table 17
Demographic information for all subjects

Subject	Age	Gender	Etiology	Duration of Paralysis Prior to Initial Surgical Intervention	Bilateral or Unilateral	Nerve Utilized	Time Elapsed Between 1st & 2nd Surgeries	Time between the Last Surgical Intervention Completed and Study Participation
1	47	F	Bell's Palsy	15 yrs 2 mo	Unilateral	Facial	SSFNS occurred 5 mo following FSFNS	9 mos
2	49	F	Pleomorphic Adenoma	1 yr 10 mos	Unilateral	Facial	FSFNS had occurred, but SSFNS had not	3 mos
3	19	M	Mobius Syndrome	19 yrs	Bilateral	Trigeminal	FMNS had occurred, but SMNS had not	1 yr
4	11	M	Trauma (R. occipital skull fracture)	7 yrs	Unilateral	Facial	SSFNS occurred 7 mo following FSFNS	1 yr 5 mos
5	17	F	Congenital	17 yrs	Unilateral	Facial	SSFNS occurred 10 mo following FSFNS	2 yrs 11 mos

FSFNS = First stage facial nerve surgery
SSFNS = Second stage facial nerve surgery
FMNS = First masseter nerve surgery
SMNS = Second masseter nerve surgery

From “Physiological and Functional Outcomes in Patients with Facial Paralysis Following Facial Reanimation Surgery” by M. Harasem, 2008, Unpublished Master’s thesis, p.20.

APPENDIX 3: EXAMPLE OF AREA AND MAXIMAL DISTANCE DATA

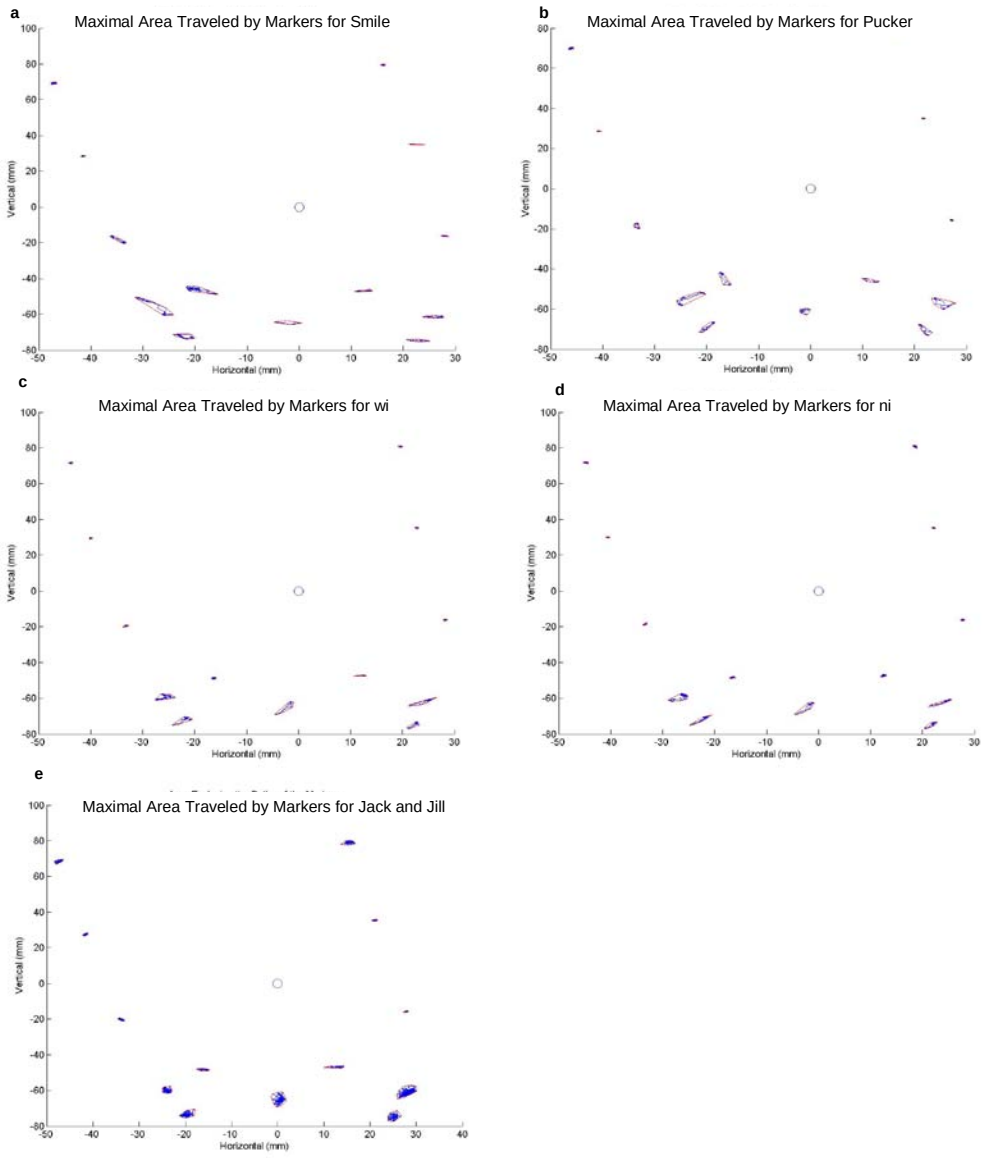

Figure 15. Graphical representation of maximal area (in mm2) traveled by all markers in a single trial by S2 for (a) Smile, (b) Pucker, (c) Wi, (d) Ni, and (e) Jack and Jill. *Note.* Open circle in center represents reference marker.

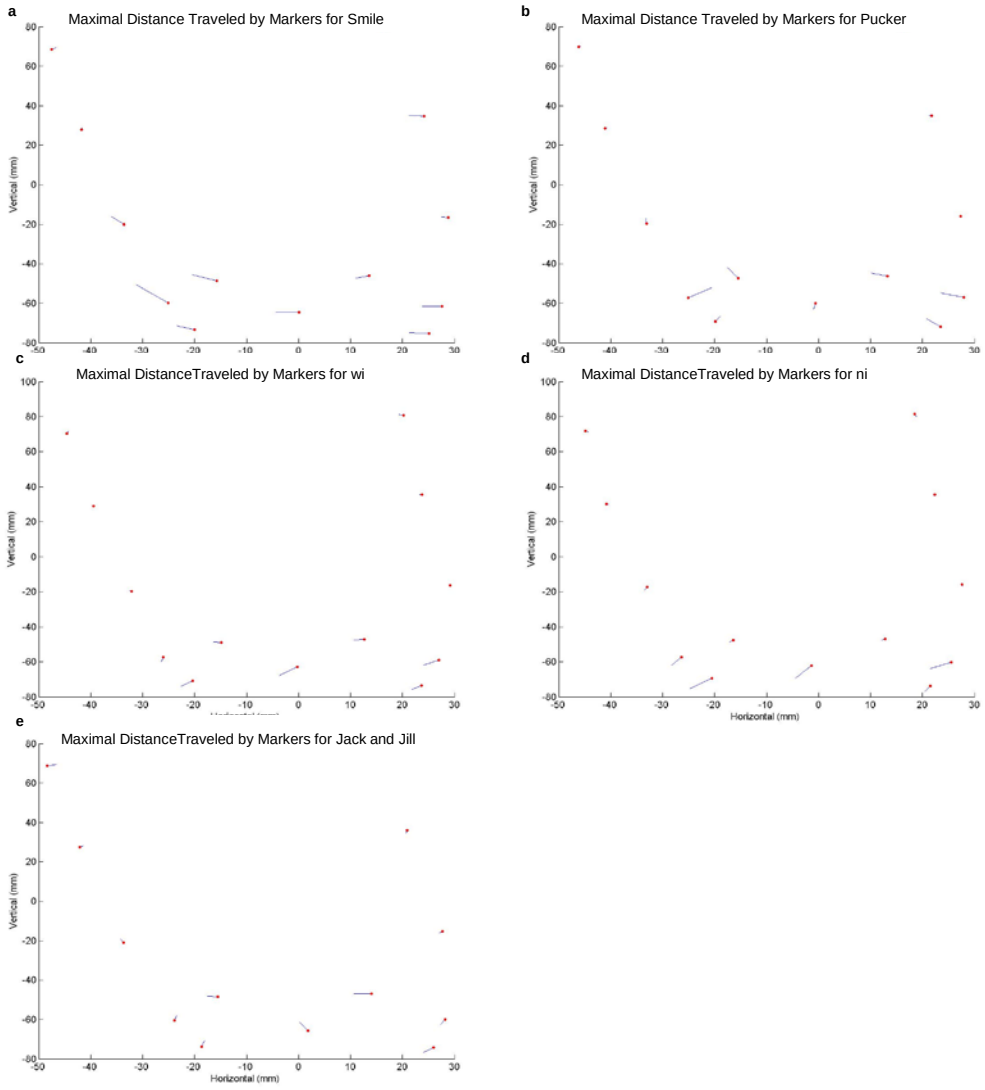

Figure 16. Graphical representation of maximal distance (in mm) traveled by all markers in a single trial by S2 for (a) Smile, (b) Pucker, (c) Wi, (d) Ni, and (e) Jack and Jill. *Note.* Red dot represents the starting position of the marker.

APPENDIX 4: FIGURES FOR S1

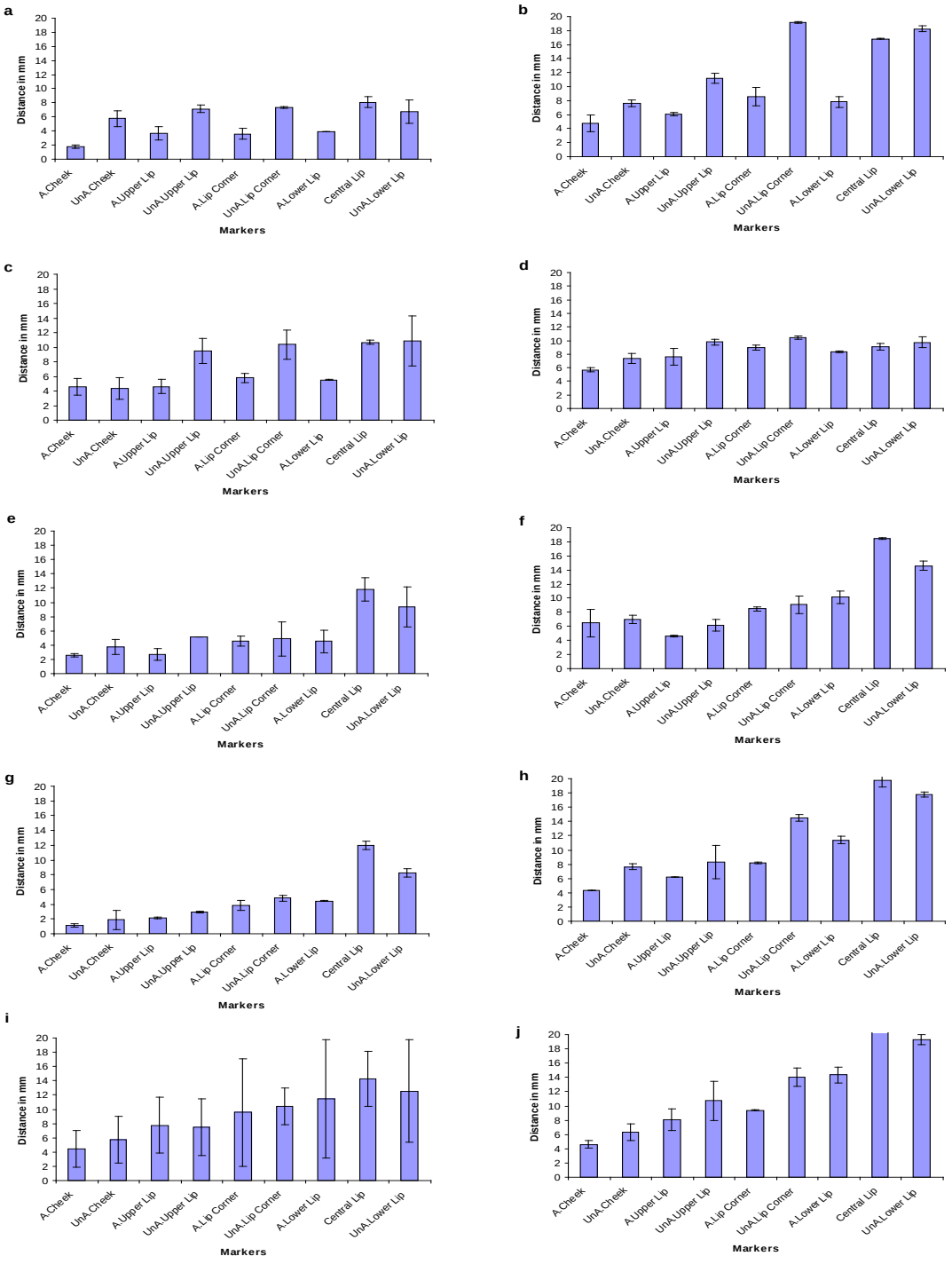

Figure 17. Error bar graphs of maximal distance (in mm) traveled by all markers for S1 for (a) Smile in 2D, (b) Smile in 3D, (c) Pucker in 2D, (d) Pucker in 3D, (e) Wi in 2D, (f) Wi in 3D, (g) Ni in 2D, (h) Ni in 3D, (i) Jack and Jill in 2D, and (j) Jack and Jill in 3D. *Note*. A = Affected side of face, UnA = Unaffected side of face.

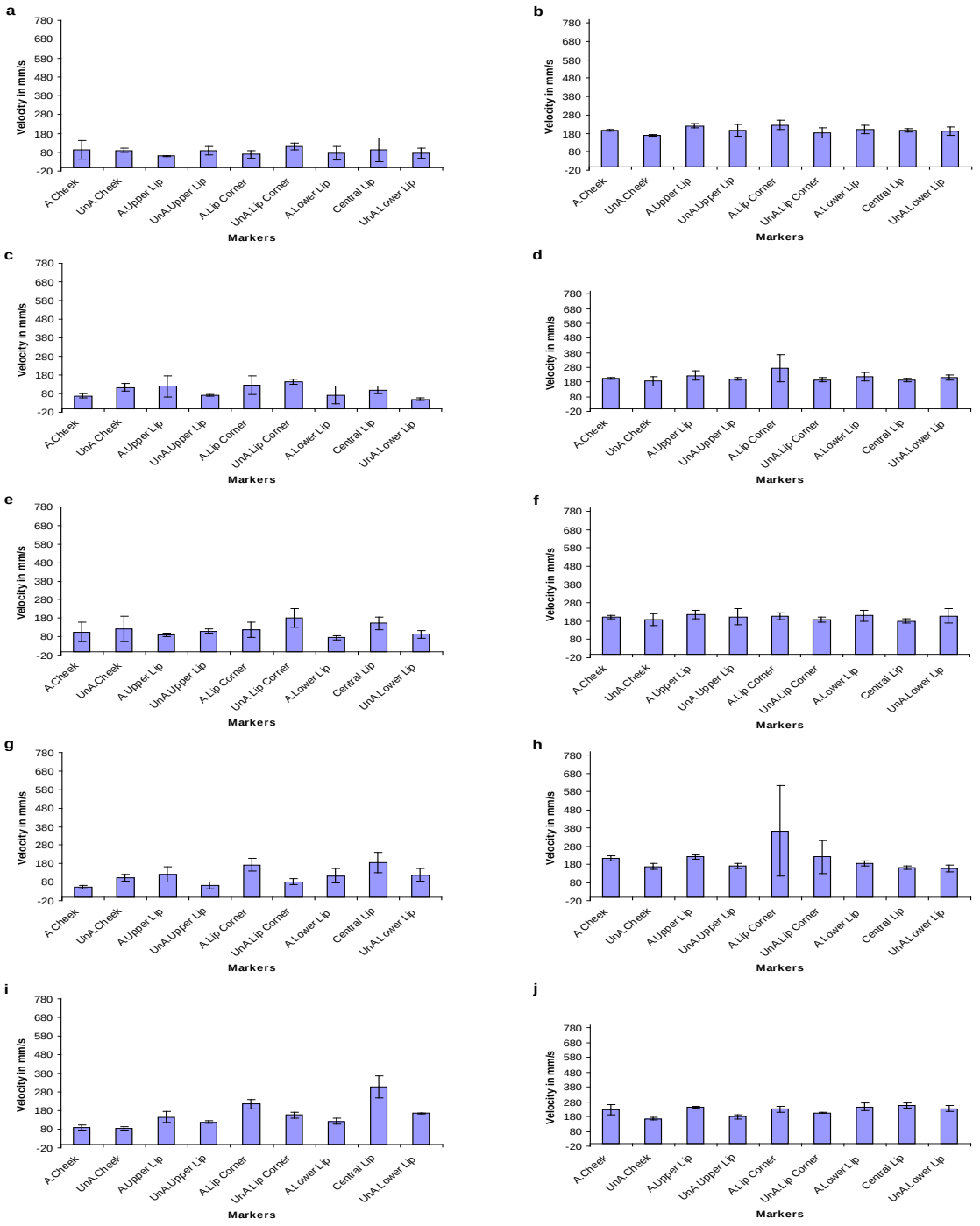

Figure 18. Error bars of velocity (in mm/s) for all markers for S1 for (a) Smile in 2D, (b) Smile in 3D, (c) Pucker in 2D, (d) Pucker in 3D, (e) Wi in 2D, (f) Wi in 3D, (g) Ni in 2D, (h) Ni in 3D, (i) Jack and Jill in 2D, and (j) Jack and Jill in 3D. *Note*. A = Affected side of face, UnA = Unaffected side of face.

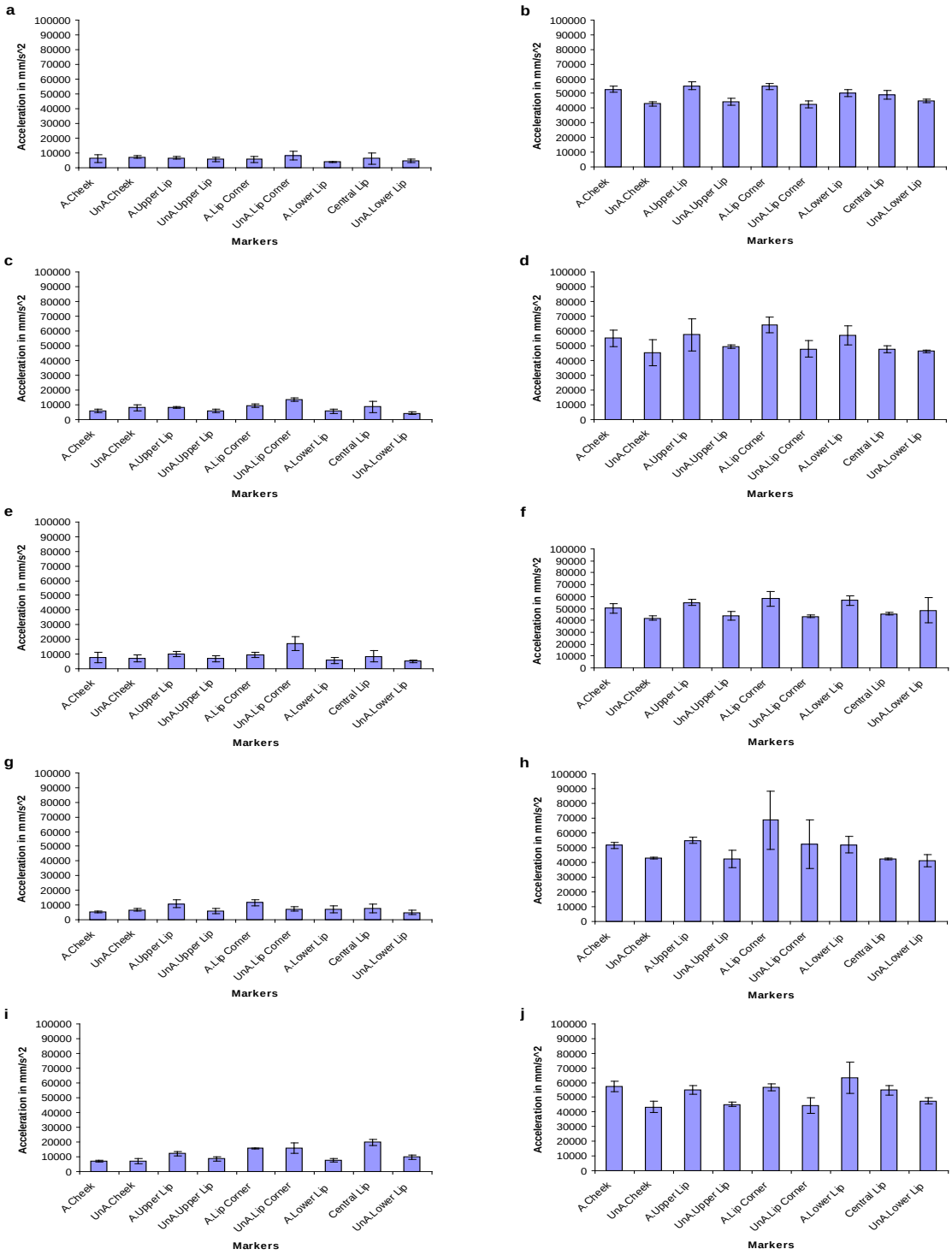

Figure 19. Error bars of acceleration (in mm/s^2) for all markers for S1 for (a) Smile in 2D, (b) Smile in 3D, (c) Pucker in 2D, (d) Pucker in 3D, (e) Wi in 2D, (f) Wi in 3D, (g) Ni in 2D, (h) Ni in 3D, (i) Jack and Jill in 2D, and (j) Jack and Jill in 3D. *Note.* A = Affected side of face, UnA = Unaffected side of face.

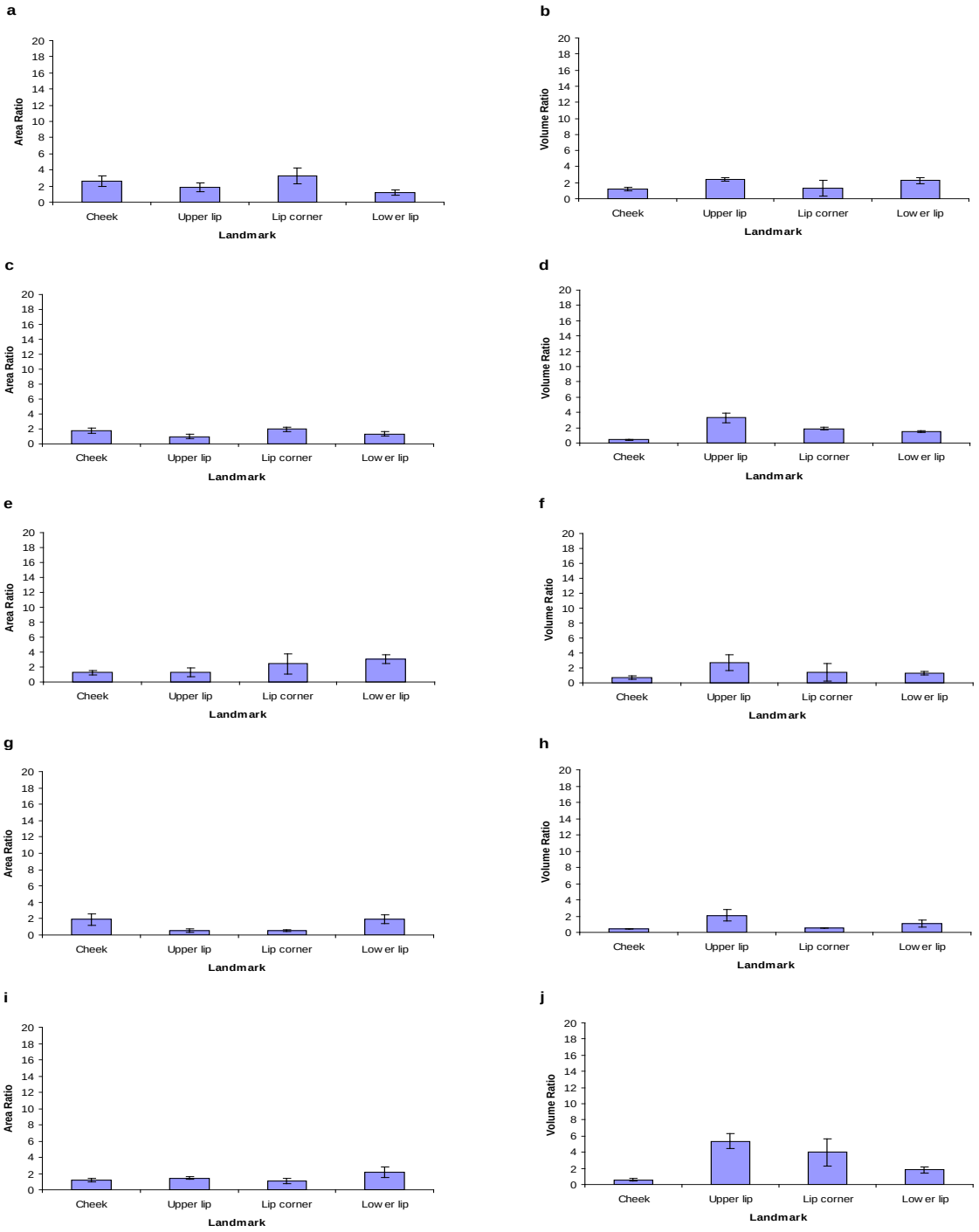

Figure 20. Error bars of area and volume ratios for all markers for S1 for (a) Smile in 2D, (b) Smile in 3D, (c) Pucker in 2D, (d) Pucker in 3D, (e) Wi in 2D, (f) Wi in 3D, (g) Ni in 2D, (h) Ni in 3D, (i) Jack and Jill in 2D, and (j) Jack and Jill in 3D. *Note*. A = Affected side of face, UnA = Unaffected side of face

APPENDIX 5: FIGURES FOR S2

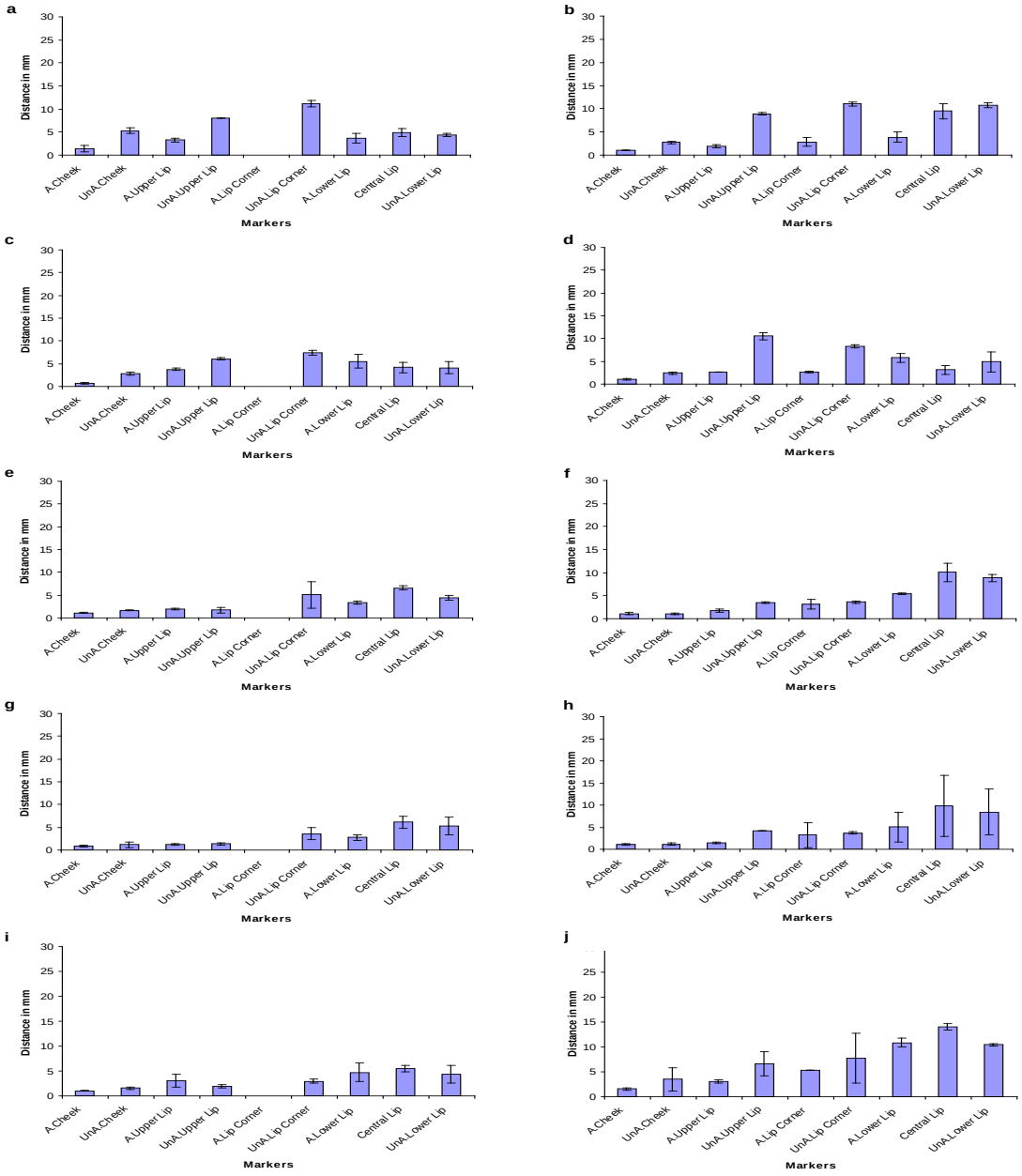

Figure 21. Error bar graphs of maximal distance (in mm) traveled by all markers for S2 for (a) Smile in 2D, (b) Smile in 3D, (c) Pucker in 2D, (d) Pucker in 3D, (e) Wi in 2D, (f) Wi in 3D, (g) Ni in 2D, (h) Ni in 3D, (i) Jack and Jill in 2D, and (j) Jack and Jill in 3D. *Note*. A = Affected side of face, UnA = Unaffected side of face.

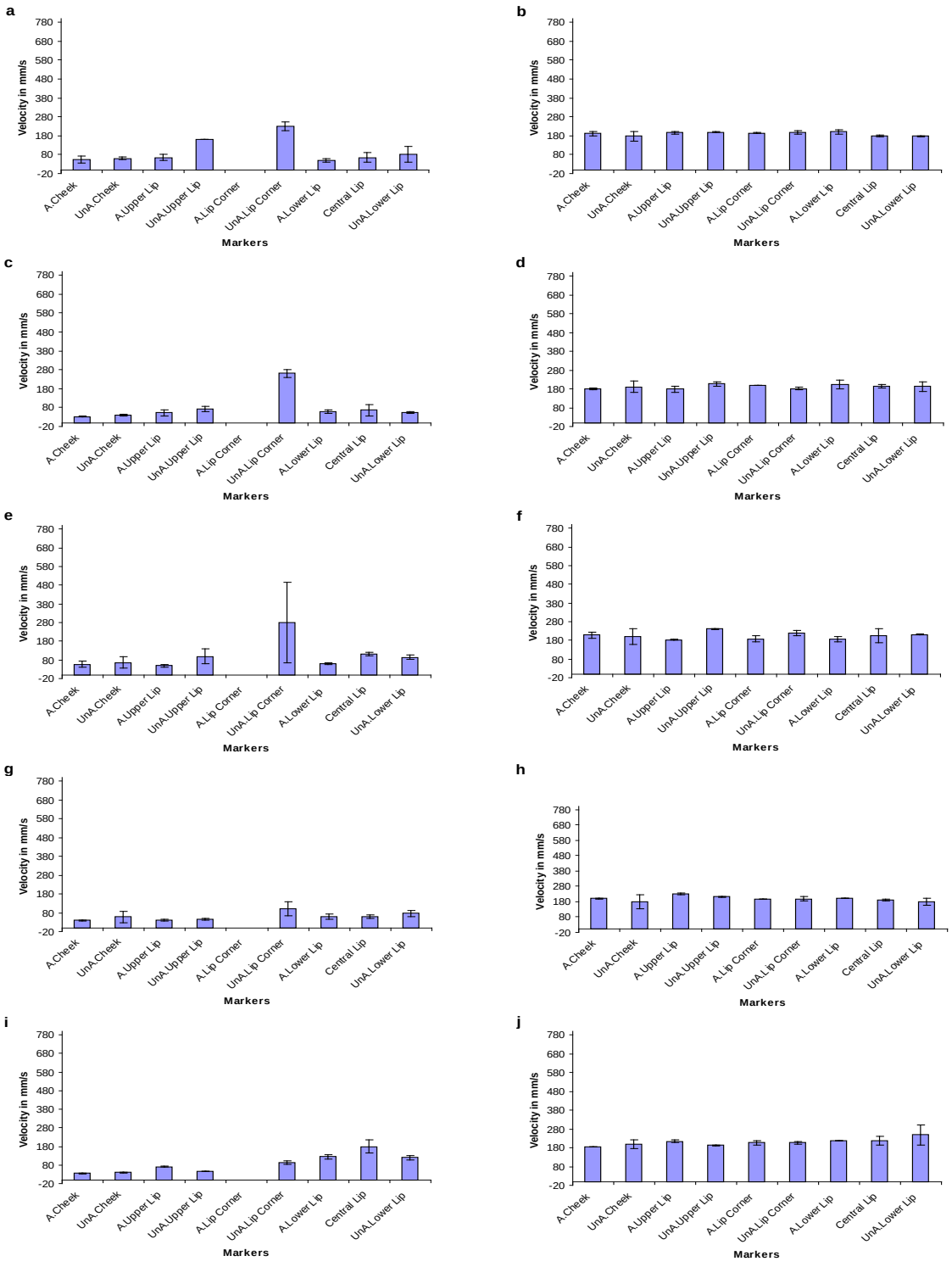

*Figure 22***.** Error bars of velocity (in mm/s) for all markers for S2 for (a) Smile in 2D, (b) Smile in 3D, (c) Pucker in 2D, (d) Pucker in 3D, (e) Wi in 2D, (f) Wi in 3D, (g) Ni in 2D, (h) Ni in 3D, (i) Jack and Jill in 2D, and (j) Jack and Jill in 3D. *Note*. A = Affected side of face, UnA = Unaffected side of face.

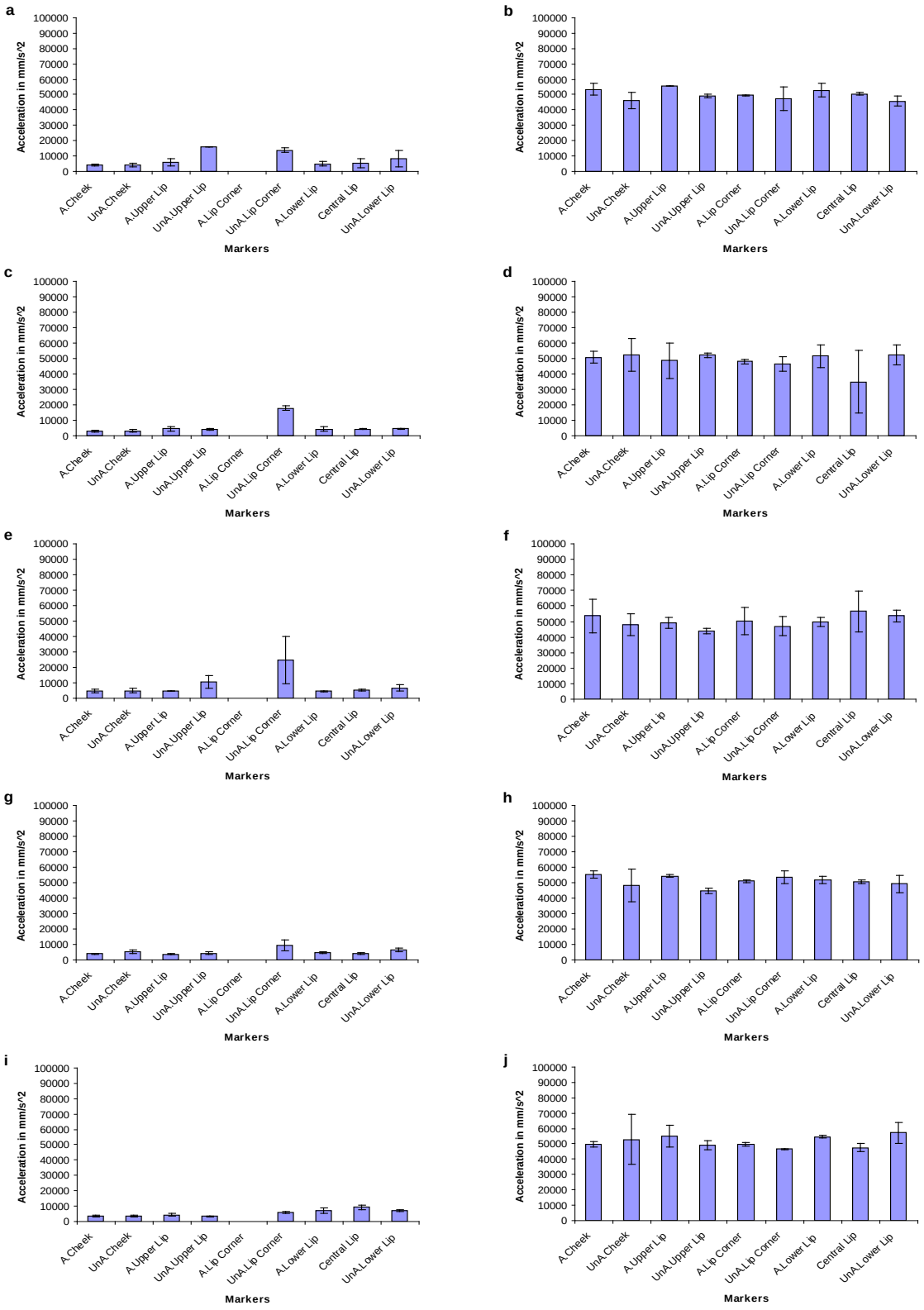

Figure 23. Error bars of acceleration (in mm/s^2) for all markers for S2 for (a) Smile in 2D, (b) Smile in 3D, (c) Pucker in 2D, (d) Pucker in 3D, (e) Wi in 2D, (f) Wi in 3D, (g) Ni in 2D, (h) Ni in 3D, (i) Jack and Jill in 2D, and (j) Jack and Jill in 3D. *Note*. A = Affected side of face, UnA = Unaffected side of face.

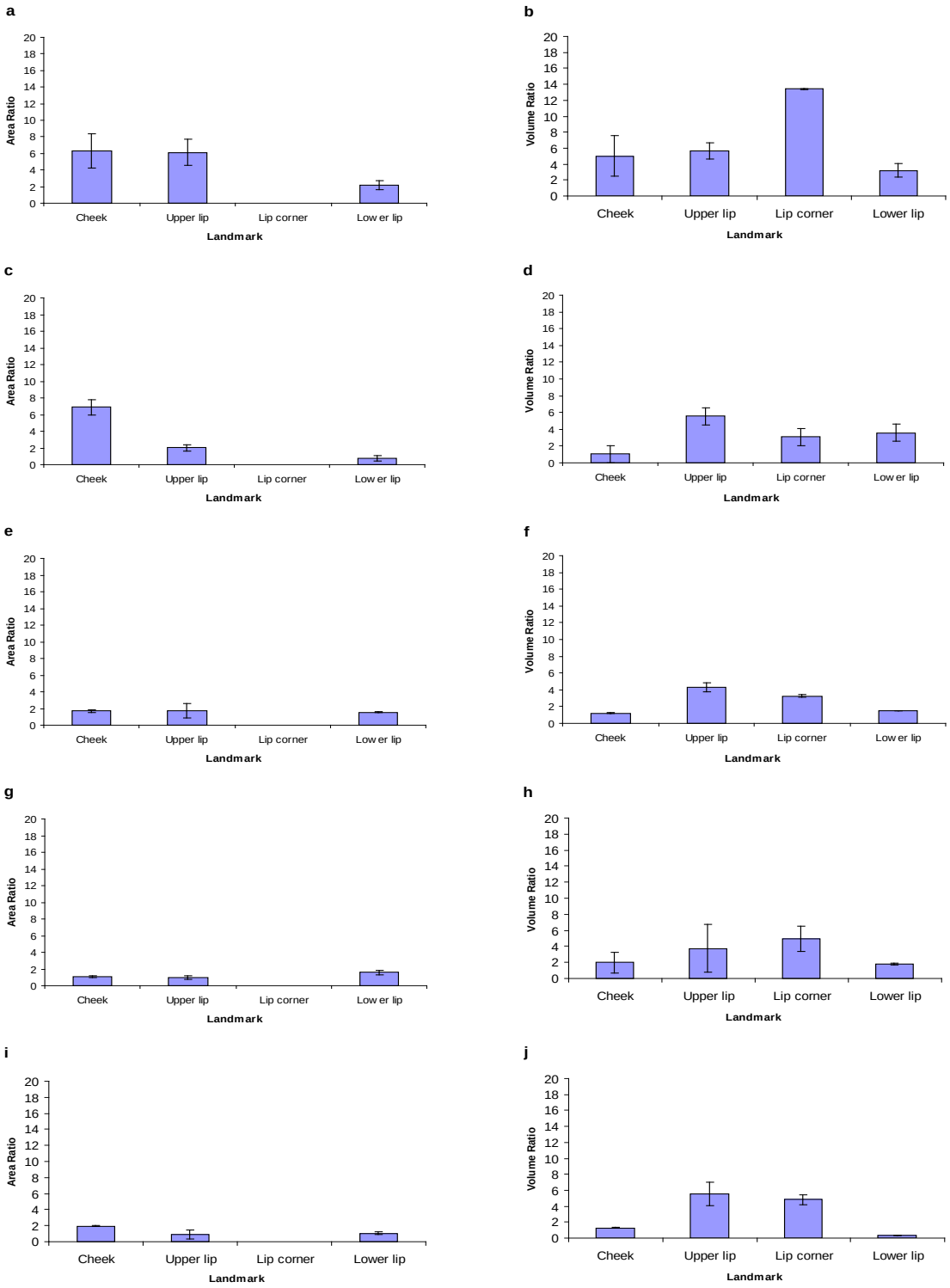

Figure 24. Error bars of area and volume ratios for all markers for S2 for (a) Smile in 2D, (b) Smile in 3D, (c) Pucker in 2D, (d) Pucker in 3D, (e) Wi in 2D, (f) Wi in 3D, (g) Ni in 2D, (h) Ni in 3D, (i) Jack and Jill in 2D, and (j) Jack and Jill in 3D. *Note*. A = Affected side of face, UnA = Unaffected side of face**.**

APPENDIX 6: FIGURES FOR S3

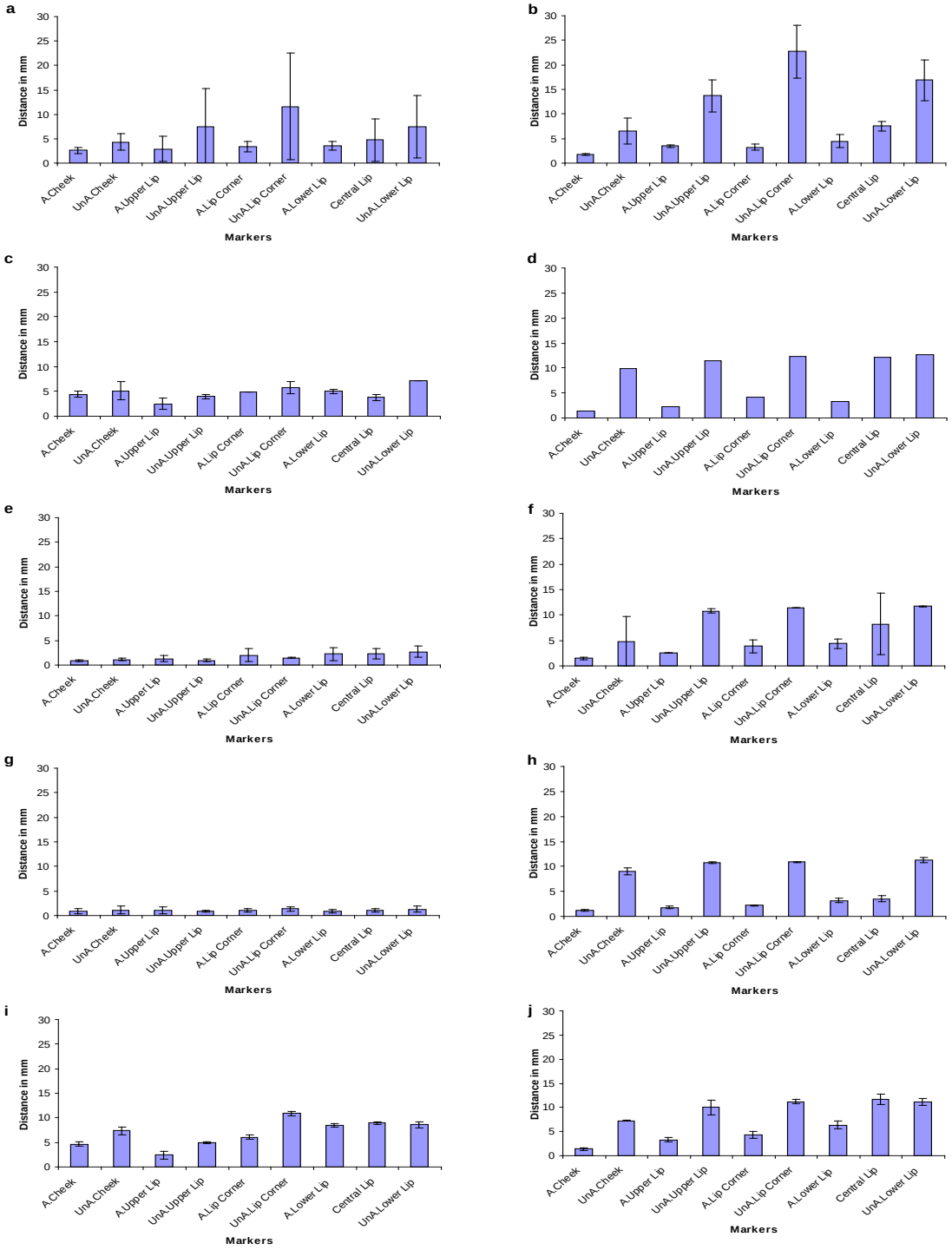

Figure 25. Error bar graphs of maximal distance (in mm) traveled by all markers for S3 for (a) Smile in 2D, (b) Smile in 3D, (c) Pucker in 2D, (d) Pucker in 3D, (e) Wi in 2D, (f) Wi in 3D, (g) Ni in 2D, (h) Ni in 3D, (i) Jack and Jill in 2D, and (j) Jack and Jill in 3D. *Note*. A = Affected side of face, UnA = Unaffected side of face.

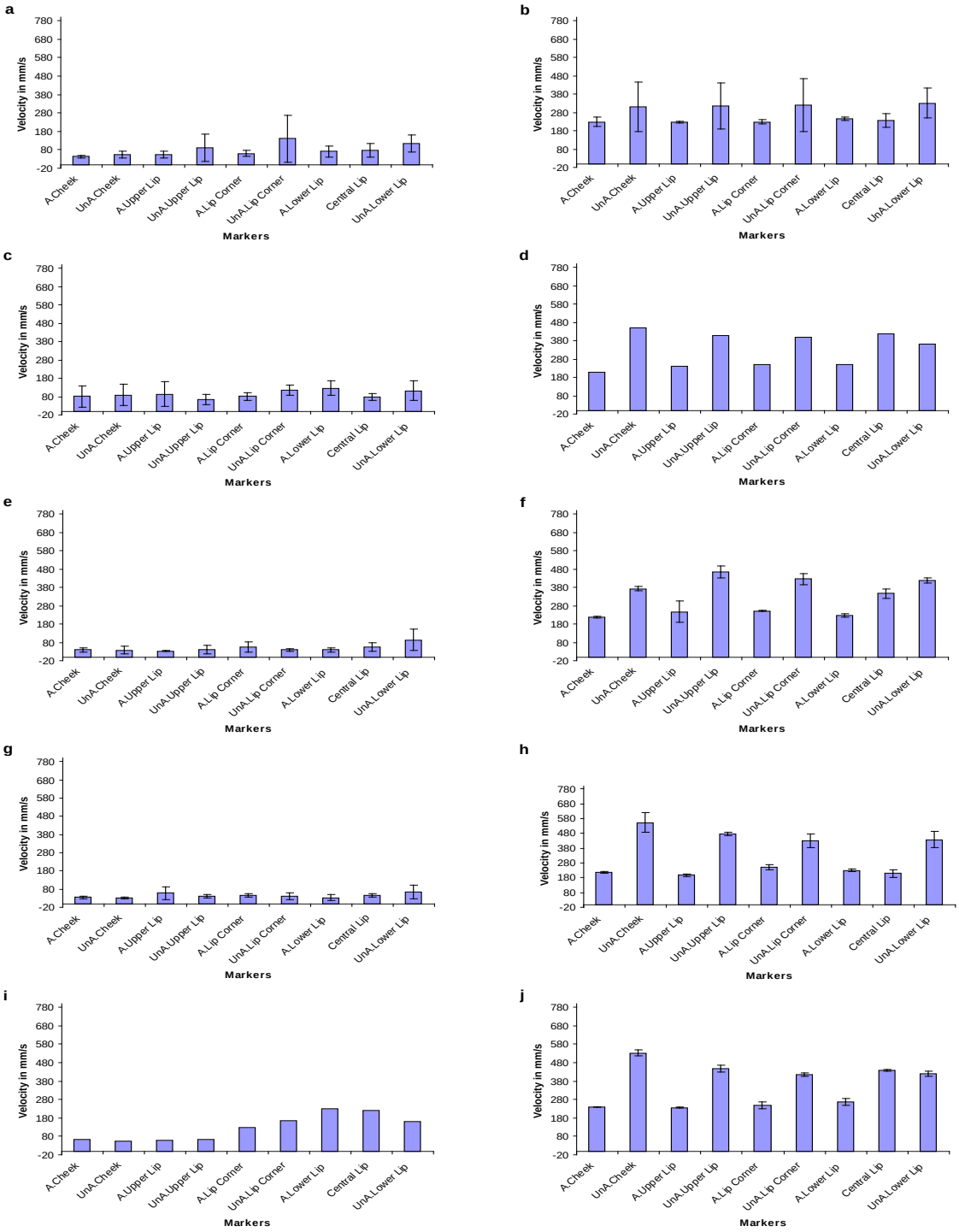

Figure 26. Error bar graphs of maximal velocity (in mm/s) traveled by all markers for S3 for (a) Smile in 2D, (b) Smile in 3D, (c) Pucker in 2D, (d) Pucker in 3D, (e) Wi in 2D, (f) Wi in 3D, (g) Ni in 2D, (h) Ni in 3D, (i) Jack and Jill in 2D, and (j) Jack and Jill in 3D. *Note*. A = Affected side of face, UnA = Unaffected side of face.

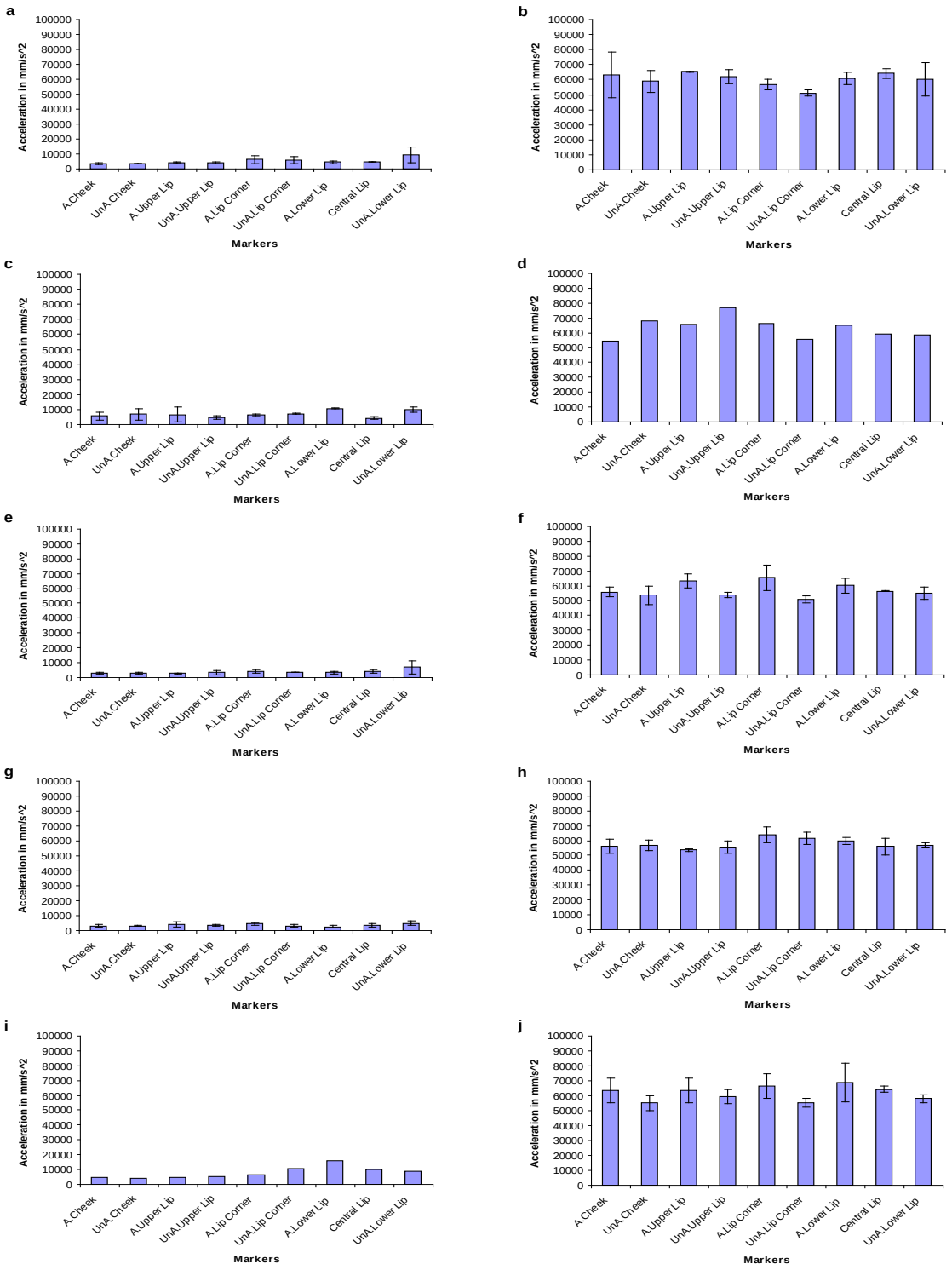

Figure 27. Error bar graphs of maximal acceleration (in mm/s^2) traveled by all markers for S3 for (a) Smile in 2D, (b) Smile in 3D, (c) Pucker in 2D, (d) Pucker in 3D, (e) Wi in 2D, (f) Wi in 3D, (g) Ni in 2D, (h) Ni in 3D, (i) Jack and Jill in 2D, and (j) Jack and Jill in 3D. *Note*. A = Affected side of face, UnA = Unaffected side of face.

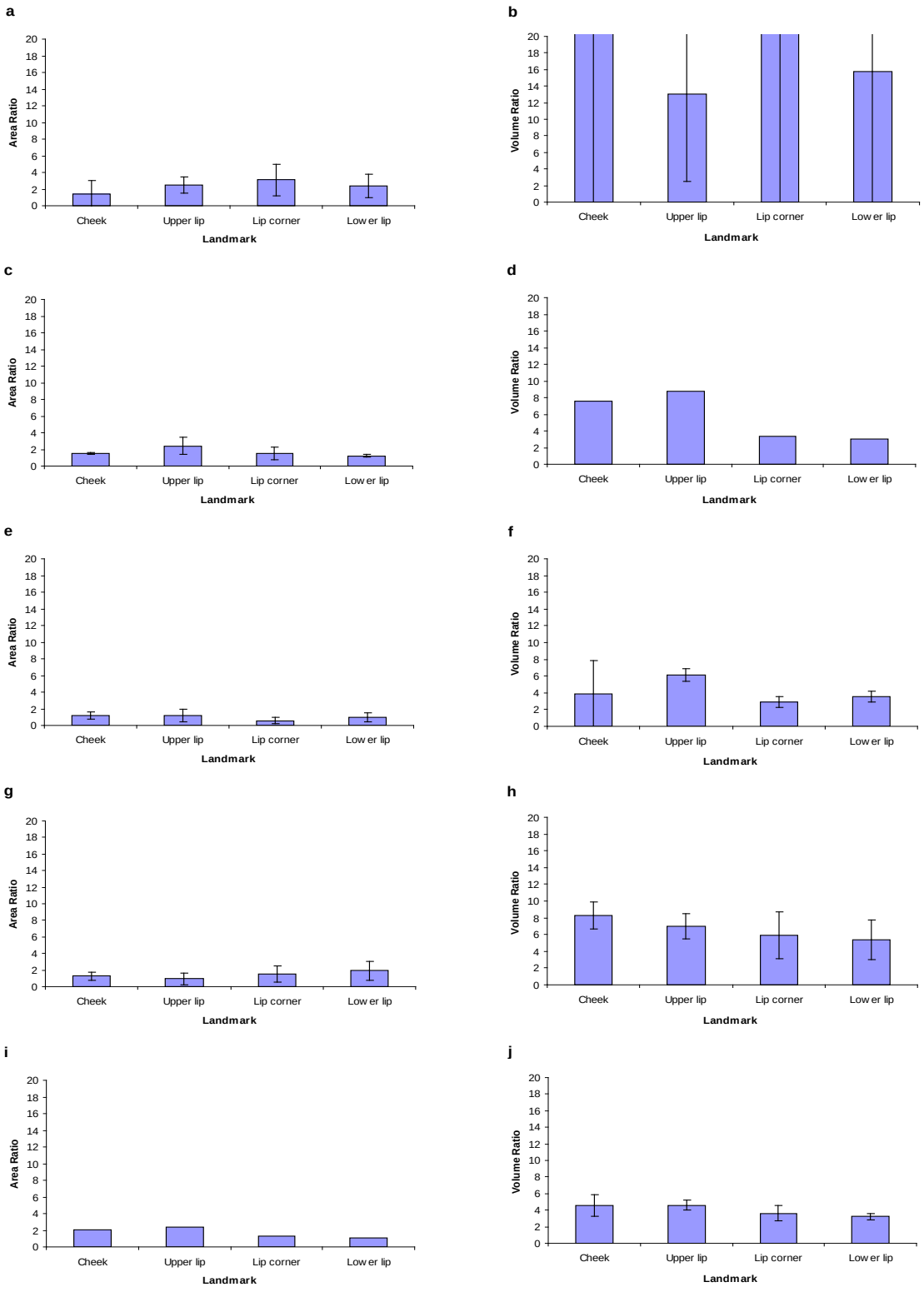

Figure 28. Error bars of area and volume ratios for all markers for S3 for (a) Smile in 2D, (b) Smile in 3D, (c) Pucker in 2D, (d) Pucker in 3D, (e) Wi in 2D, (f) Wi in 3D, (g) Ni in 2D, (h) Ni in 3D, (i) Jack and Jill in 2D, and (j) Jack and Jill in 3D. *Note*. A = Affected side of face, UnA = Unaffected side of face**.**

APPENDIX 7: FIGURES FOR S4

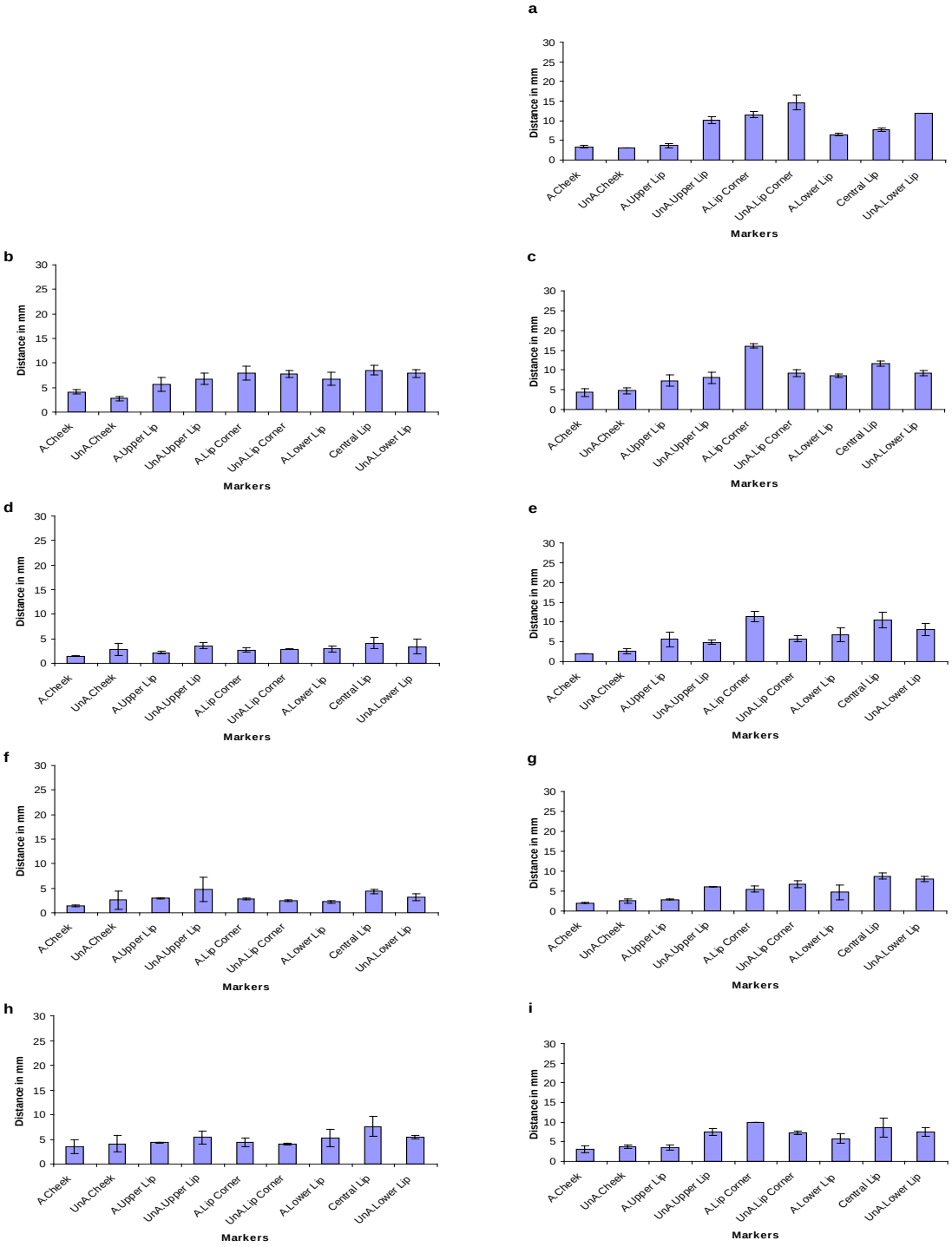

Figure 29. Error bar graphs of maximal distance (in mm) traveled by all markers for S4 for (a) Smile in 3D, (b) Pucker in 2D, (c) Pucker in 3D, (d) Wi in 2D, (e) Wi in 3D, (f) Ni in 2D, (g) Ni in 3D, (h) Jack and Jill in 2D, and (i) Jack and Jill in 3D. *Note*. A = Affected side of face, UnA = Unaffected side of face.

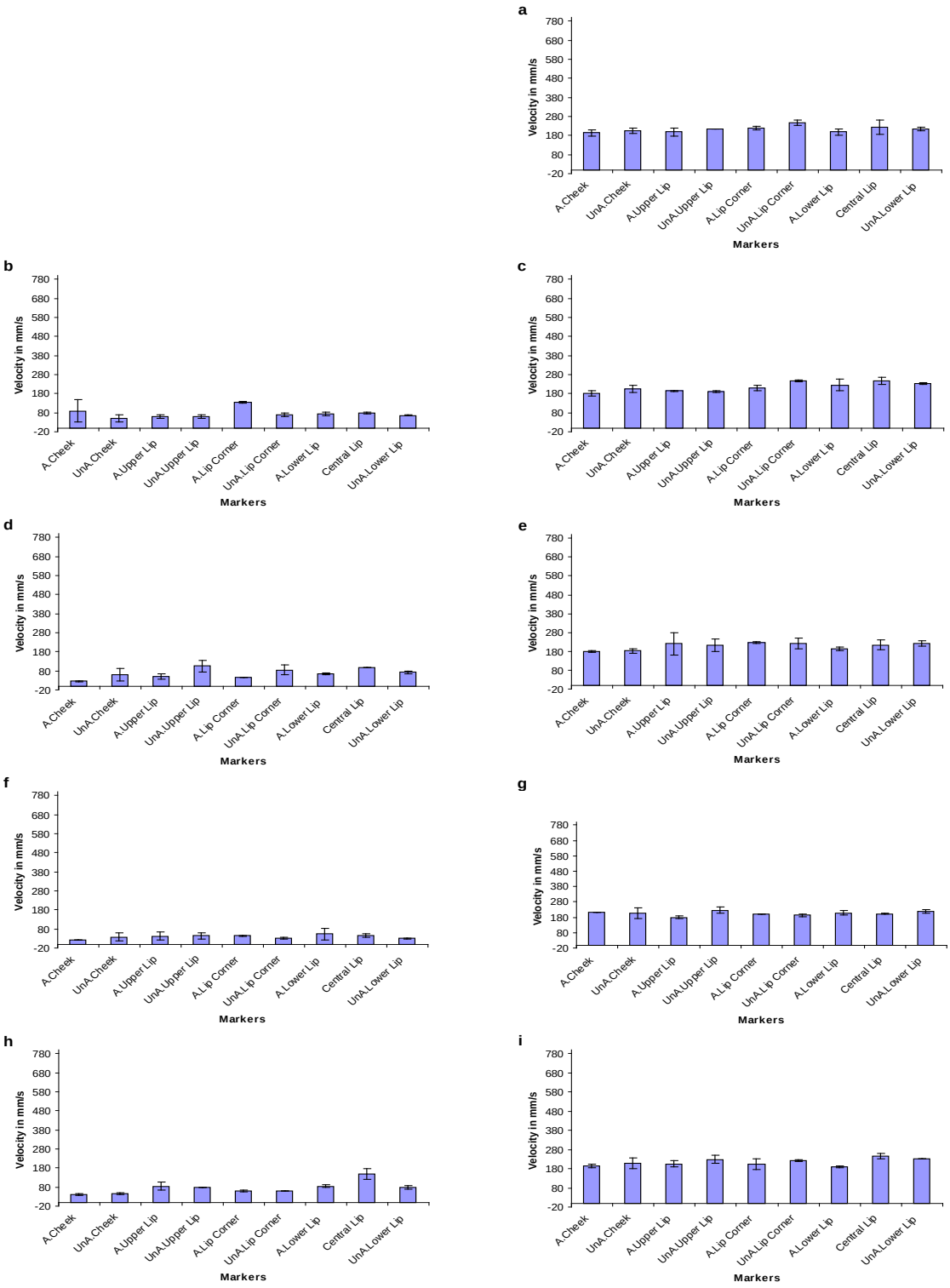

Figure 30. Error bar graphs of maximal velocity (in mm/s) traveled by all markers for S4 for (a) Smile in 3D, (b) Pucker in 2D, (c) Pucker in 3D, (d) Wi in 2D, (e) Wi in 3D, (f) Ni in 2D, (g) Ni in 3D, (h) Jack and Jill in 2D, and (i) Jack and Jill in 3D. *Note.* A = Affected side of face, UnA = Unaffected side of face.

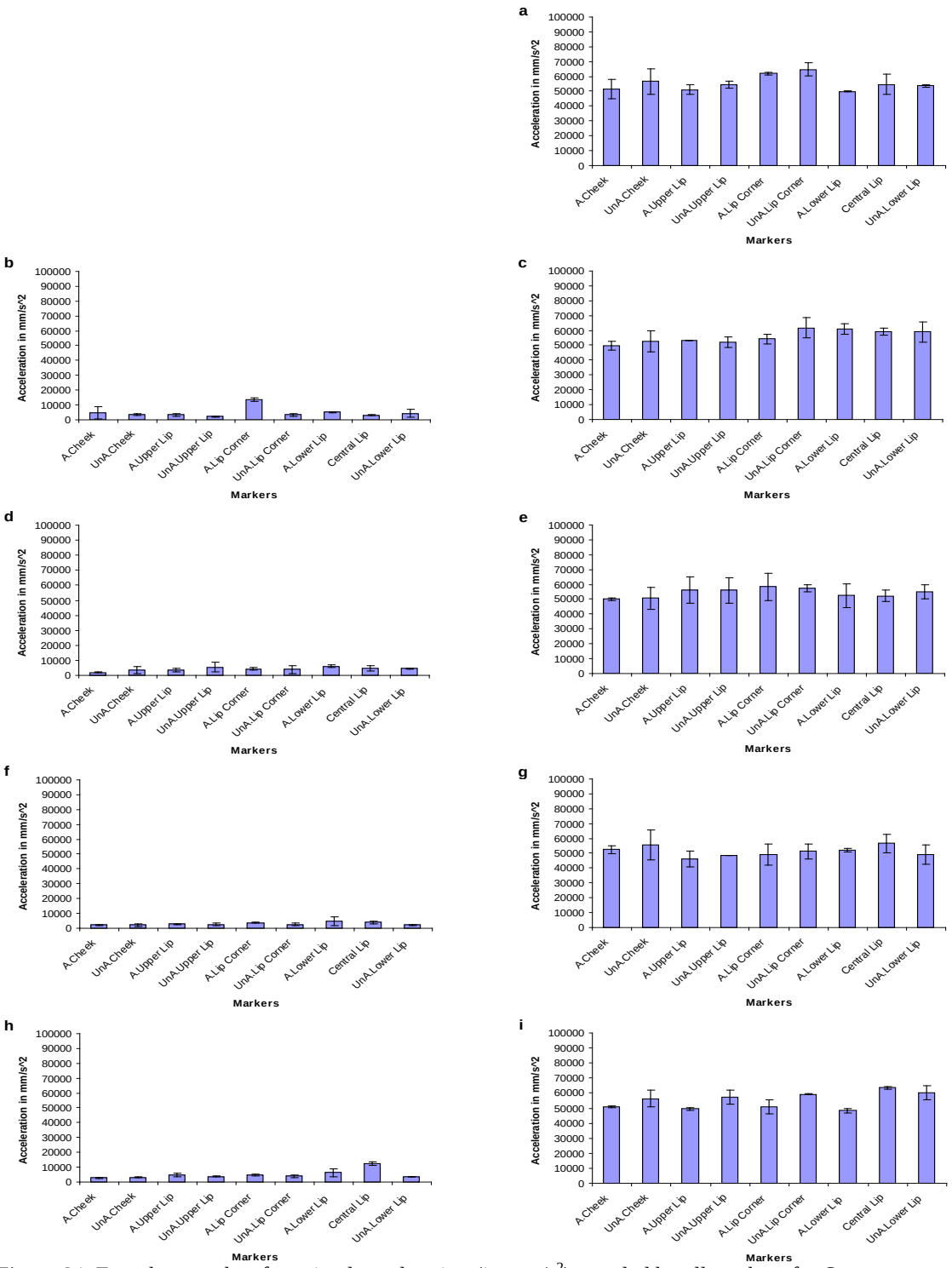

Figure 31. Error bar graphs of maximal acceleration (in mm/s^2) traveled by all markers for S 4 for (a) Smile in 3D, (b) Pucker in 2D, (c) Pucker in 3D, (d) Wi in 2D, (e) Wi in 3D, (f) Ni in 2D, (g) Ni in 3D, (h) Jack and Jill in 2D, and (i) Jack and Jill in 3D. *Note*. A = Affected side of face, UnA = Unaffected side of face.

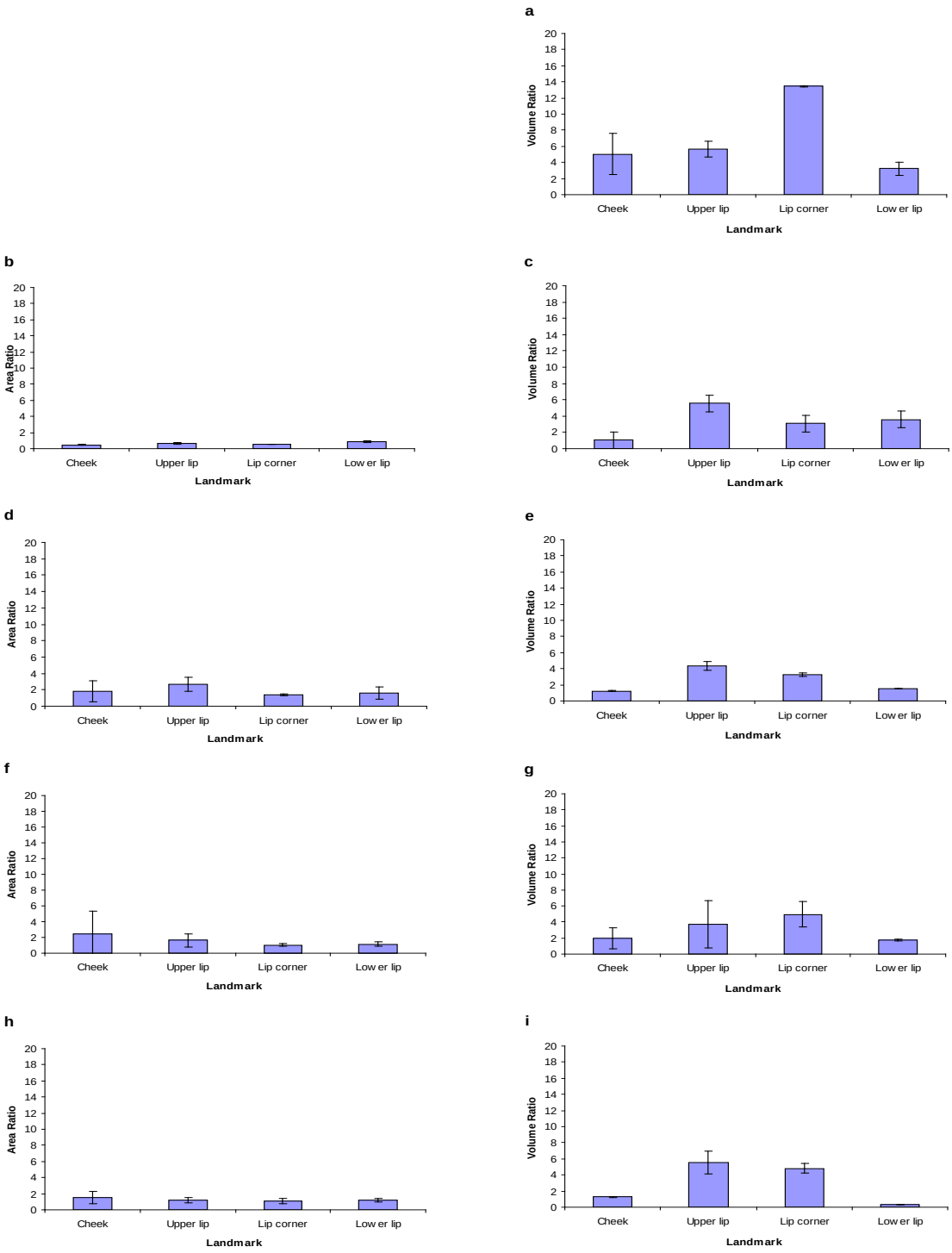

Figure 32. Error bar graphs of area and volume ratios for all markers for S4 for (a) Smile in 3D, (b) Pucker in 2D, (c) Pucker in 3D, (d) Wi in 2D, (e) Wi in 3D, (f) Ni in 2D, (g) Ni in 3D, (h) Jack and Jill in 2D, and (i) Jack and Jill in 3D. *Note.* A = Affected side of face, UnA = Unaffected side of face.

APPENDIX 8: FIGURES FOR S5

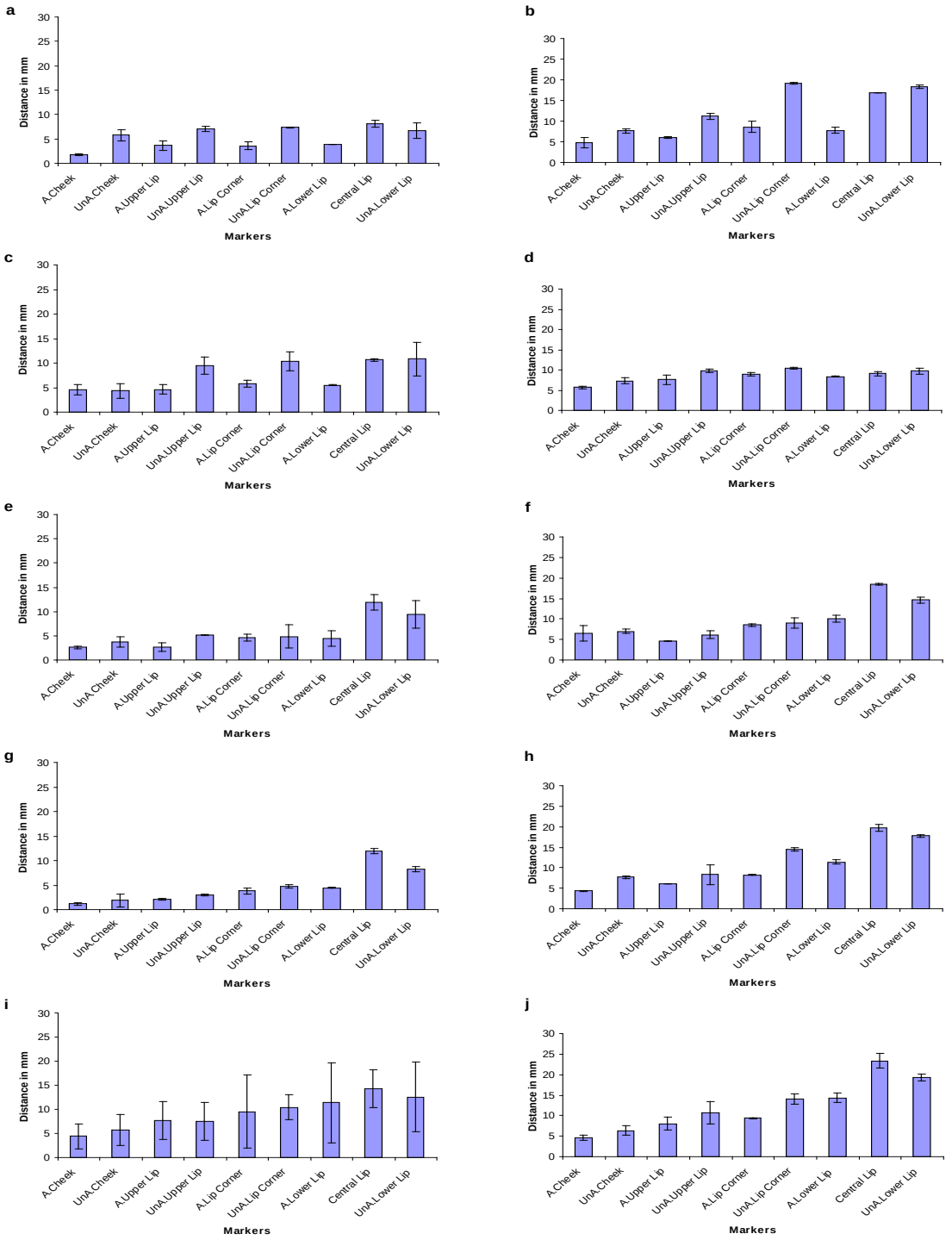

Figure 33. Error bar graphs of maximal distance (in mm) traveled by all markers for S5 for (a) Smile in 2D, (b) Smile in 3D, (c) Pucker in 2D, (d) Pucker in 3D, (e) Wi in 2D, (f) Wi in 3D, (g) Ni in 2D, (h) Ni in 3D, (i) Jack and Jill in 2D, and (j) Jack and Jill in 3D. *Note.* A = Affected side of face, UnA = Unaffected side of face.

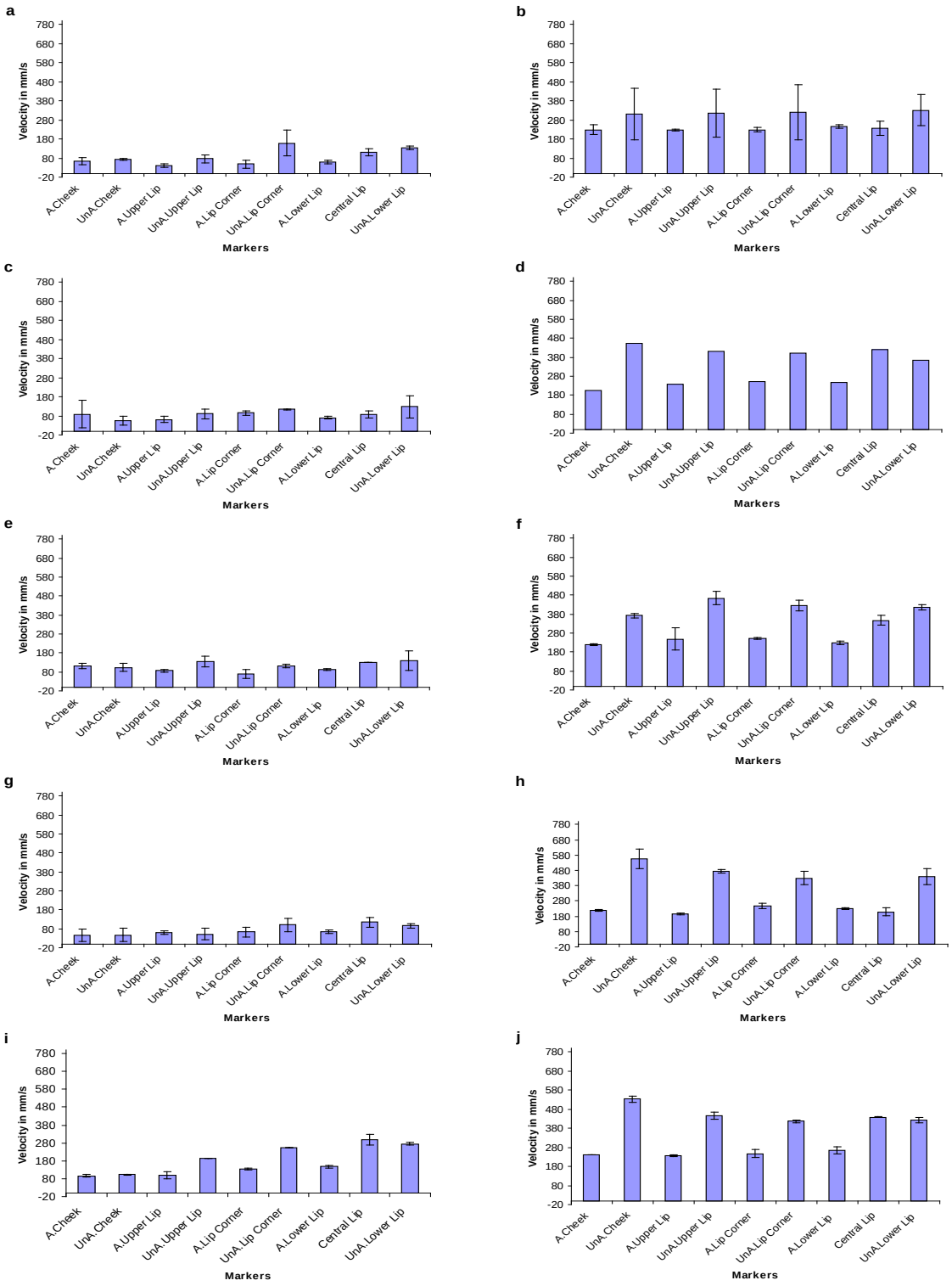

Figure 34. Error bar graphs of maximal velocity (in mm/s) traveled by all markers for S5 for (a) Smile in 2D, (b) Smile in 3D, (c) Pucker in 2D, (d) Pucker in 3D, (e) Wi in 2D, (f) Wi in 3D, (g) Ni in 2D, (h) Ni in 3D, (i) Jack and Jill in 2D, and (j) Jack and Jill in 3D. *Note.* A = Affected side of face, UnA = Unaffected side of face.

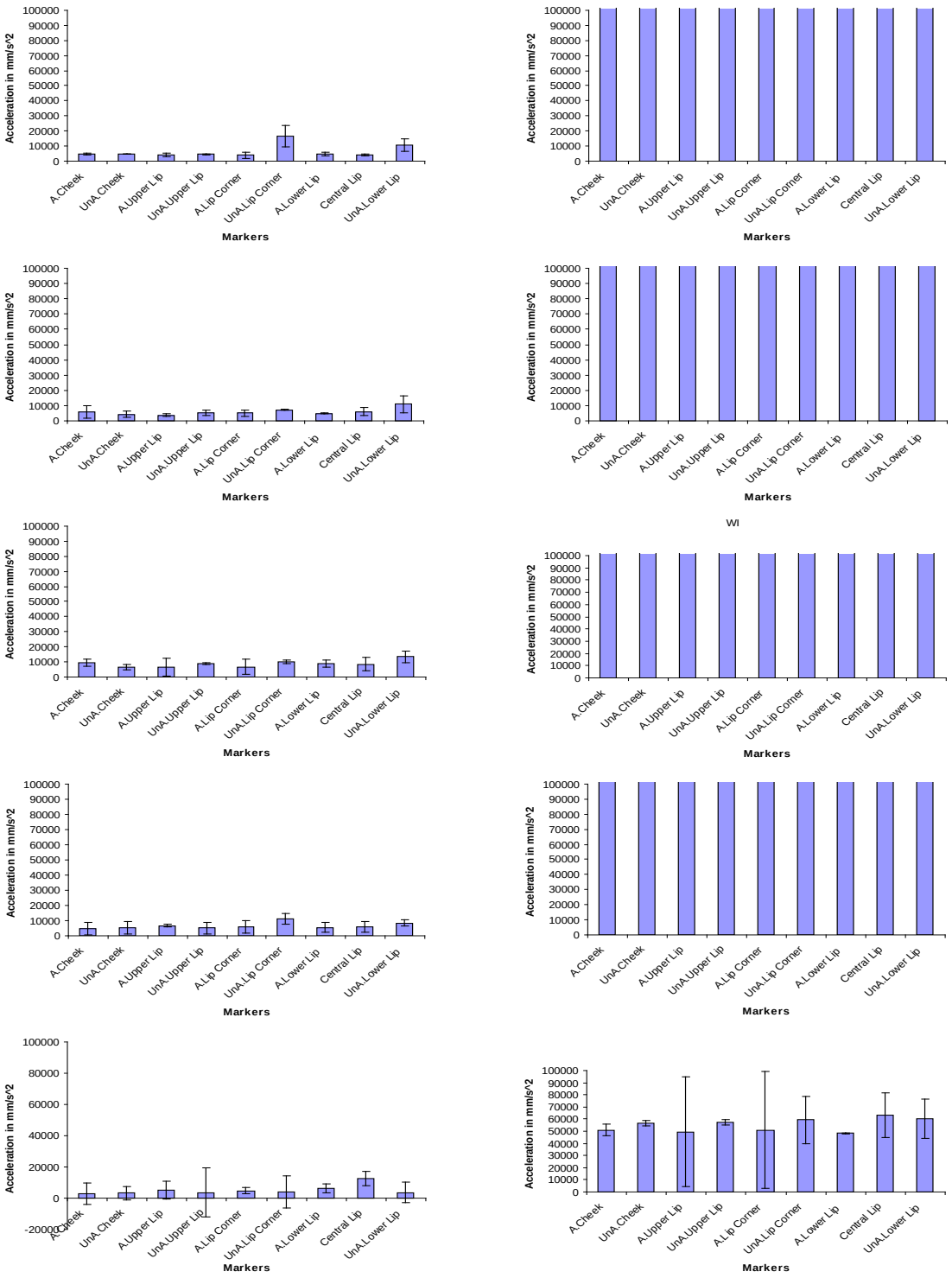

Figure 35. Error bar graphs of maximal acceleration (in mm/s^2) traveled by all markers for S5 for (a) Smile in 2D, (b) Smile in 3D, (c) Pucker in 2D, (d) Pucker in 3D, (e) Wi in 2D, (f) Wi in 3D, (g) Ni in 2D, (h) Ni in 3D, (i) Jack and Jill in 2D, and (j) Jack and Jill in 3D. *Note*. A = Affected side of face, UnA = Unaffected side of face.

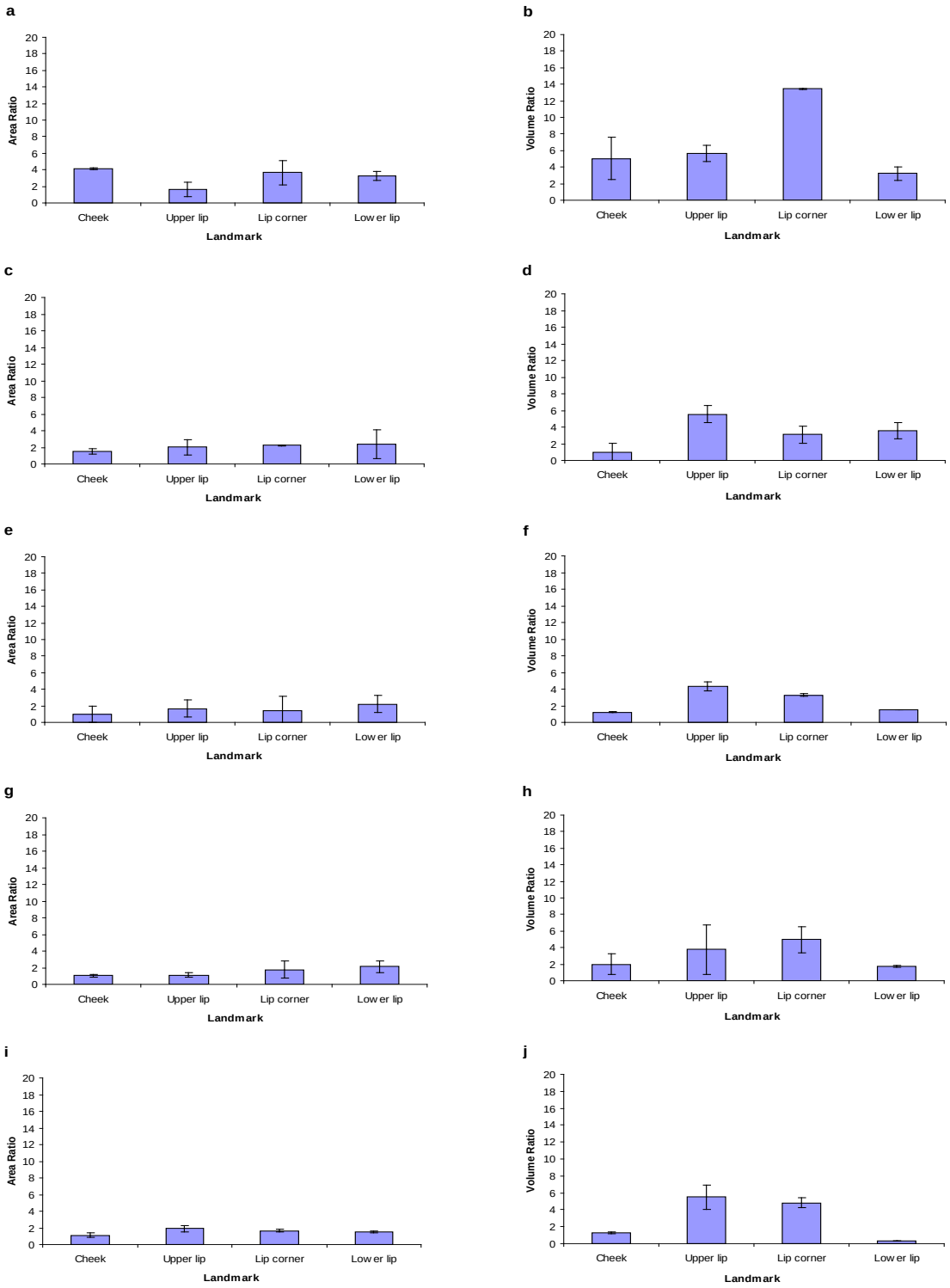

Figure 36. Error bar graphs of are and volume ratios for all markers for S5 for (a) Smile in 2D, (b) Smile in 3D, (c) Pucker in 2D, (d) Pucker in 3D, (e) Wi in 2D, (f) Wi in 3D, (g) Ni in 2D, (h) Ni in 3D, (i) Jack and Jill in 2D, and (j) Jack and Jill in 3D. *Note.* A = Affected side of face, UnA = Unaffected side of face.