

OBJECTIVE ASSESSMENT OF SUCCESSFUL SMILE REANIMATION

A Multi-Modal Approach

Joseph Richard Dusseldorp

A thesis submitted in fulfilment of the requirement for the degree of
Doctorate of Philosophy

Faculty of Medicine and Health
The University of Sydney

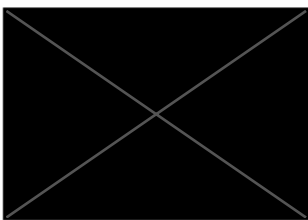
2023

CANDIDATE'S CERTIFICATE

I, Joseph Dusseldorp, hereby declare that the work contained within this thesis is my own and has not been submitted to any other university or institution for any higher degree.

I, Joseph Dusseldorp, hereby declare that I was the principal researcher of all work contained in this thesis, including work published with multiple authors.

I, Joseph Dusseldorp, understand that if I am awarded a higher degree for my thesis entitled **‘OBJECTIVE ASSESSMENT OF SUCCESSFUL SMILE REANIMATION: A Multi-Modal Approach’** being submitted herewith for examination, the thesis will be lodged in the University Library and be will available immediately for use. I agree that the University Librarian (or in the case of the department, the Head of the Department) may supply a photocopy or microform of the thesis to an individual for research or study or to a library.

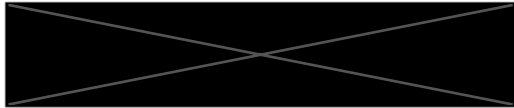


Joseph R. Dusseldorp

Date: 1st September, 2023

SUPERVISOR'S CERTIFICATE

This is to certify that the thesis entitled '**OBJECTIVE ASSESSMENT OF SUCCESSFUL SMILE REANIMATION: A Multi-Modal Approach**' submitted by Joseph Dusseldorp in fulfilment of the requirements for the degree of Doctorate of Philosophy is in a form suitable for examination.



Professor Peter Maitz

Date: 3rd September 2023

ACKNOWLEDGEMENTS

A chance meeting at the world-first Emotion Artificial Intelligence (AI) Summit at the Massachusetts Institute of Technology (MIT) in Boston in 2017 was the genesis of this thesis. I was working as the Facial Reanimation Fellow at the Facial Nerve Centre at the Massachusetts Eye and Ear Infirmary (MEEI) and Harvard Medical School at the time. I was seeking a technological solution that could improve the way I measured my patients' ability to smile after smile reanimation surgery. At the conference I heard a marketing executive talk about how they could use AI to automatically tell which points during a TV advertisement were the most engaging by how much their viewers smiled. I was then fortunate to meet Dr Rana el Kaliouby, the developer of the computer vision and facial recognition AI behind this capability.

Dr el Kaliouby was no ordinary computer science engineer. She was also an entrepreneur, and founder of Affectiva, the world's largest "Emotion AI" company. When I explained to her what I was hoping to achieve for facial palsy patients, she immediately saw the application of her technology and generously gifted us the use of her software throughout our study at no cost. Without her generosity this project would not have been possible.

Translating an AI algorithm into the medical field was not without its challenges and I could not have succeeded without the assistance of the Facial Nerve Centre's entire research team and my co-authors: Tessa Hadlock, Nate Jowett, Diego Guarin, Tijn van Veen, Jacqueline Greene, Olivia Quatela, Suresh Mohan, Matt Naunheim, Matt Miller, Joanna Tavares-Brito, Emily Fortier, Kerry Camp, Julie Arria, Jane Baiungo, and Mara Robinson.

I loved my time working in the disruptive environment that is MEEI and Harvard in general and I hope to be able to make my own contributions to the field of reconstructive plastic surgery in the future that could live up to the great dreams of the MEEI team.

To my supervisors, Peter Maitz, Tessa Hadlock and Pierre Chapuis, my sincere thanks for your guidance and support and for helping me to craft a cohesive narrative around this incessant focus of mine that is *Successful Smile Reanimation*.

DEDICATION

This thesis is dedicated to my loving wife, Mally, who has always supported me despite my long hours away studying, operating or on mission trips, even when those hours away may have been after having dragged you, along with our newborn babies, to live overseas, where you may have not even spoken the language.

I will always remember the sacrifices you have made to make our family as great as it is and anything we are able to achieve together has this greatness as its foundation.

I cannot thank you enough.

PROLOGUE

Facial palsy is a disabling condition that has a profound effect on a patient that is not only physical, but also psychological and social. It is a condition that in many respects originates in the eye of the beholder due to the difficulties that arise when a person living with facial palsy tries to communicate their emotions to others using their facial expressions. After facial palsy this ability is significantly impaired and sufferers are left to mourn the loss of their smile, previously an effortless part of their personality.

Over the last century, clinicians have developed techniques to restore facial balance and movement to facial palsy patients, now collectively known as facial reanimation procedures. However, whilst surgical techniques have progressed rapidly, our ability to quantify the outcomes of such interventions has not kept pace and presently there is no measure that objectively captures all domains that facial reanimation impacts for an affected individual.

An objective assessment of the success of any one facial reanimation procedure should address all domains impacted by the disorder, measuring the physical impact through objective forms of smile measurements and analysis, the psychological effect on the individual through patient reported outcomes measures, and the social effect through layperson assessments.

The aim of this thesis was to develop novel techniques that might enable a multi-modal approach to objectively quantify the success of smile reanimation surgery to restore normal smiles to facial palsy patients. To achieve this aim, this thesis had five objectives. **First**, to introduce the concepts of voluntary and spontaneous smiling and discuss the evolution of

surgical techniques for the reanimation of the smile in patients afflicted by facial paralysis. **Second**, to review the ability of existing measures of facial palsy severity and clinical outcome grading scales to quantify outcomes of smile reanimation procedures. **Third**, to determine the relative importance of voluntary and spontaneous smiling. **Fourth**, to measure facial movements on the healthy and diseased sides of the face and quantify the formation of the nasolabial fold, a key feature of successful smiling. **Fifth**, to automatically quantify the ability of smile reanimation techniques to restore successful spontaneous smiling, the most elusive attribute of a normal smile.

To address the first objective, a review of facial reanimation principles and techniques was performed (**Chapter One**). To address the second objective, a literature review was conducted investigating facial palsy severity grading scales and outcome measures (**Chapter Two**). To address the third objective, a ranked preference methodology known as discrete choice experiment was used to determine the relative preferences of laypersons when choosing between surgical techniques that restore voluntary smiling versus those that also aim to restore spontaneous smiling (**Chapter Three**). To achieve the fourth objective, state of the art facial landmark recognition tools were utilised to quantify the formation of the nasolabial fold during smiling in unilateral facial palsy patients, comparing movement on the affected and healthy sides of the face (**Chapter Four**). To address the final objective, a stepwise process was undertaken to first define a technique to manually time-stamp spontaneous smiles in a videoclip then to use automated facial landmarking software to time-stamp in an automated fashion, and finally to use an emotion AI algorithm to quantify the ability of those detected spontaneous smiles to express joy, a novel measure of successful spontaneous smiling (**Chapter Five**).

The thesis concludes with an **Epilogue** to draw together the linkages from the preceding chapters and suggest areas for further research. In the final **Addendum**, excerpts from a chapter

in Dr Rana el Kaliouby's Autobiography are included to demonstrate the parallel perspective of a non-surgeon as she unexpectedly discovered the capabilities of her emotion AI algorithm, and her amazement that such a tool could impact the life of someone living with facial palsy.

Table of Contents

Candidate's certificate	i
Supervisor's certificate	ii
Acknowledgements	iii
Dedication	v
Prologue	vi
Table of contents	ix
Outline of the thesis	xi
Dissemination of published research	xiv

Chapter One

The Role of Smile Reanimation Surgery in Facial Palsy.....	1
- Preface	2
- Why do we Smile?	3
- Evolution of Surgical Techniques to Restore Smiling in Facial Palsy	7
- List of References	11

Chapter Two

Objective Outcome Assessment in Facial Reanimation	13
- Preface	14
- Authorship Statement	16
- Appendix 1	25
- The Spectrum of Facial Palsy: The MEEI Facial Palsy Photo Video Standard Set	25
- Authorship Statement	26

Chapter Three

Relative Importance Voluntary & Spontaneous Smiling in Facial Palsy...	33
- Preface	34
- Authorship Statement	37
- Appendix 2	47
- Patient Experience Nerve-To-Masseter-Driven Smile Reanimation	47
- Authorship Statement	48
- Appendix 3	56
- Association Regional Facial Dysfunction Facial Palsy–Related QOL	56
- Authorship Statement	57
- Appendix 4	65
- Facial Palsy-Specific Quality Of Life In 920 Patients: Correlation With Clinician-Graded Severity And Predicting Factors	65
- Authorship Statement	66

Chapter Four

Nasolabial Fold Dynamics: Implications for Facial Paralysis and Facial Reanimation Surgery.....72

- Preface 73
- Authorship Statement 75
- Appendix 5 84
- A Machine Learning Approach Automated Facial Measurements In Facial Palsy 84
- Authorship Statement 85
- Appendix 6 88
- Toward An Automatic System For Computer-Aided Assessment In Facial Palsy 88
- Authorship Statement 89

Chapter Five

Objective Measurement of Smile Spontaneity in Facial Reanimation..... 98

- Preface 99
- Authorship Statement 102
- Automated Spontaneity Assessment After Smile Reanimation: A Machine Learning Approach 110
- Authorship Statement 111
- Appendix 7 122
- In The Eye Of The Beholder: Changes In Perceived Emotion Expression After Smile Reanimation 122
- And Invited Discussion and Letters to the Editor of PRS 122
- Authorship Statement 123

Epilogue

Conclusions, Limitations and Suggestions for Future Research.....146

- Conclusions 147
- Limitations of the Thesis 151
- Suggestions for Future Research 152
- Concluding Remarks 153

Addendum 155

OUTLINE OF THE THESIS

Normal smiling is a complex phenomenon comprising both a social, communicative function which is voluntary, and a spontaneous, emotional expression of joy which is innate. Loss of the ability to smile has a profound effect on quality of life. This thesis outlines the series of steps taken to objectively quantify the success of smile reanimation surgery to restore normal smiling in facial paralysis patients who have lost the ability to move one side of their face. Using advanced techniques from the fields of computer science and health economics, this thesis seeks to demonstrate the value of successful smile reanimation surgery and to establish a multi-modal, objective, automated, and quantifiable approach to measure its presence. This thesis contains five chapters, an introductory chapter that summarises the current state of facial reanimation surgery, and four subsequent investigative chapters as detailed below.

Chapter One introduces the concepts of voluntary and spontaneous smiling and discusses the evolution of surgical techniques for the reanimation of the smile in patients afflicted by facial palsy.

Chapter Two provides a literature review of the evolution of facial palsy grading scales, severity assessment measures and facial reanimation outcome measurement techniques. This work has been published in *Facial Palsy, Springer*. This chapter is accompanied by Appendix 1, a published comprehensive database of the full spectrum of facial palsy patients with complete photographic and videographic assessments, an aide to the development of rigorous novel outcome measures. This database was made freely available to all researchers and published in *Laryngoscope*.

Chapter Three investigates the relative importance placed on voluntary and spontaneous smiling. The chapter utilises the methodology of a novel health economic technique, known as the discrete choice experiment, to determine the preferred choice of nerve source to achieve both voluntary and spontaneous smiling. Conducted in 250 laypersons, this paper was published in *Plastic and Reconstructive Surgery*. The three appendices to this chapter focus on patient reported outcomes of quality of life following smile reanimation. First masseteric nerve driven smile reanimation was studied to discern the true amount of successful smile utilisation and patient satisfaction with this technique. Appendix 3 and 4 investigated the relative importance of loss of smile ability to the overall reduction in quality of life experienced by facial palsy patients. These studies were published in the *Journal of Plastic Reconstructive and Aesthetic Surgery*, *JAMA Facial Plastic Surgery* and *Laryngoscope*.

Chapter Four details the use of state-of-the-art techniques for objectively measuring facial movements using facial landmark tracking software. The chapter demonstrates that manually placed facial tracking markers can enable identification of the movement of the nasolabial fold which is a notoriously difficult facial landmark for computer vision software algorithms to recognise given its absence in repose in young people and wide variation in morphology. This paper was published in *Surgery and Orthodontic Craniofacial Research*. The appendices to this chapter outline efforts to use a machine learning approach to enable automated placement of facial landmarking dots, known as Emotrics, automated measurements of facial displacements, and ultimately the development of automated facial palsy severity grading scales, known as Auto eFace. These papers have been published in *Facial Plastic Surgery Aesthetic Medicine* and *JAMA Facial Plastic Surgery*.

Chapter Five comprises two papers that document a stepwise approach to development of a novel objective measure of spontaneous smiling in facial palsy. The first step was to

reliably define and identify when spontaneous smiles occurred in healthy and diseased faces. An objective technique known as time-stamping was defined to manually quantify spontaneous smile occurrences from video recordings of facial expressions of healthy volunteers and facial palsy patients while they watched humorous videoclips. This paper was published in *JAMA Facial Plastic Surgery*. The second step was to combine the capabilities of automated facial landmarking, Emotrics, and automated emotion detection to create a system of automated identification and quantification of successful spontaneous smile reanimation. This paper was published in *Plastic and Reconstructive Surgery*. The appendix to this chapter, and its ensuing correspondence in this literature, is a paper that outlines the development and validation of the use of an emotion detection algorithm to measure the probability of a perceived emotion from the facial expressions of facial palsy patients when attempting to smile voluntarily. This paper and its accompanying Invited Discussion, Letter to the Editor and Response were also published in *Plastic and Reconstructive Surgery*.

The Epilogue summarises the main findings of the thesis, discusses limitations and details implications for clinical practice and future research.

In the final **Addendum**, excerpts from a chapter in Dr Rana el Kaliouby's Autobiography are included to demonstrate the parallel perspective of a computer software engineer who was amazed to discover that a facial recognition artificial intelligence algorithm that she had developed to measure the facial expressions of people watching TV advertisements could impact the life of someone living with facial palsy.

DISSEMINATION OF PUBLISHED RESEARCH

Parts of the work presented in this thesis have been published in the following Texts and Journals:

- | | |
|--|----------------------------|
| • Facial Palsy 2021 | Springer |
| • Plastic and Reconstructive Surgery | Impact Factor 5.169 (2022) |
| • Laryngoscope | Impact Factor 3.325 (2020) |
| • Facial Plastic Surgery and Aesthetic Medicine | Impact Factor 3.185 (2021) |
| • JAMA Facial Plastic Surgery | Impact Factor 4.667 (2021) |
| • Journal of Plastic, Reconstructive & Aesthetic Surgery | Impact Factor 3.215 (2021) |
| • Orthodontics and Craniofacial Research | Impact Factor 2.563 (2021) |

Specific Publications:

Chapter One: The Role of Smile Reanimation Surgery in Facial Palsy

- No published work is included in Chapter One.

Chapter Two: Objective Outcome Assessment in Facial Reanimation

- **Dusseldorp JR**, Meares C, Hadlock TA. A Multi-modal Approach to Outcome Tracking in Facial Palsy. In: Tzou CH.J., Rodríguez-Lorenzo A. (eds) Facial Palsy. 2021. Springer, Cham. https://doi.org/10.1007/978-3-030-50784-8_2

Appendix 1:

Greene JJ, Guarin DL, Tavares J, Fortier E, Robinson M, **Dusseldorp JR**, Quatela O, Jowett N, Hadlock TA. The Spectrum of Facial Palsy: The MEEI Facial Palsy Photo and Video Standard Set. Laryngoscope. 2020 Jan;130(1):32-37

Chapter Three: Relative Importance of Voluntary and Spontaneous Smiling in Facial Palsy

- **Dusseldorp JR**, Naunheim MR, Quatela O, Fortier E, Hadlock TA, Jowett N. Neurotization Preferences in Smile Reanimation: A Discrete Choice Experiment. Plast. Recon. Surg. 2021 Sep;148(3):407e-415e

Appendices 2 – 4:

van Veen MM, **Dusseldorp JR**, Quatela O, Baiungo J, Robinson M, Jowett N, Hadlock TA. Patient Experience in Nerve-to-Masseter-Driven Smile Reanimation. J Plast Reconstr Aesthet Surg. 2019 Aug;72(8):1265-1271
van Veen MM, Tavares-Brito J, van Veen BM, **Dusseldorp JR**, Werker P, Dijkstra PU, Hadlock TA.

Association of Regional Facial Dysfunction with Facial Palsy–Related Quality of Life. *JAMA Facial Plast Surg.* 2019 Jan 1;21(1):32-37

Tavares-Brito J, van Veen MM, **Dusseldorp JR**, Bahmad F, Hadlock TA. Facial Palsy-Specific Quality of Life in 920 Patients: Correlation with Clinician-Graded Severity and Predicting Factors. *Laryngoscope.* 2019 Jan;129(1):100-104

Chapter 4: Utilising Objective Facial Measurements in Facial Palsy

- **Dusseldorp JR**, Faraway J, Razavi L, Hadlock TA, Trotman CA. Nasolabial Fold Dynamics: Implications for Facial Paralysis and Facial Reanimation Surgery. *Orthod. Craniofac. Res.* 2020. 00:1–8.

Appendices 5 – 6:

Guarin DL, Yunusova Y, Taati B, **Dusseldorp JR**, Mohan S, Tavares J, van Veen MM, Fortier E, Hadlock TA, Jowett N. Toward an Automatic System for Computer-Aided Assessment in Facial Palsy. *Facial Plast Surg Aesthet Med.* 2020 Jan;22(1):42-49

Guarin D, **Dusseldorp JR**, Hadlock TA, Jowett N. A Machine Learning Approach for Automated Facial Measurements in Facial Palsy. *JAMA Facial Plast Surg.* 2018. Jul 1;20(4):335-337

Chapter 5: Measurement of Smile Spontaneity in Facial Reanimation

- **Dusseldorp JR**, van Veen MM, Guarin DL, Quatela O, Jowett N, Hadlock TA. Spontaneity Assessment in Dually Innervated Gracilis Smile Reanimation Surgery. *JAMA Facial Plast Surg.* 2019 Dec; 21(6):551-557

- **Dusseldorp JR**, Guarin DL, van Veen MM, Miller, M, Jowett N, Hadlock TA. Automated Spontaneity Assessment after Smile Reanimation: A Machine Learning Approach. *Plast. Recon. Surg.* 2022 June; 149(6):1393-1402

Appendix 7:

Dusseldorp JR, Guarin DL, van Veen MM, Jowett N, Hadlock TA. In the Eye of the Beholder: Changes in Perceived Emotion Expression after Smile Reanimation. *Plast. Recon. Surg.* 2019 Aug; 144(2):457-471 and Invited Discussion, Letter to the Editor and Response

Awards

The candidate has been awarded the following Scholarships during the candidature:

- Herbert and Gloria Kees Scholarship 2021 and 2022 by the Royal Australasian College of Surgeons
- Foundation for Surgery Research Scholarship 2020 by the Royal Australasian College of Surgeons
- Herbert and Gloria Kees Scholarship 2019 by the Royal Australasian College of Surgeons
- Commonwealth Research Training Program Scholarship 2019-2022

CHAPTER ONE

The Role of Smile Reanimation Surgery in Facial Palsy

Preface

Smiling is a fundamental part of our personalities simultaneously enabling effective non-verbal communication and exposing our irrepressible emotions such as joy and laughter. A primary goal of patients living with facial palsy is to restore their ability to express joy by smiling. In the 20th century, surgeons began utilising nerve and muscle transfer procedures to reanimate the paralysed sections of the face through a series of procedures collectively known as *Facial Reanimation*.

The aim of this chapter is to review the pertinent literature to introduce the concepts of voluntary and spontaneous smiling and to provide a summary of the surgical techniques that have been developed to restore animation to the smiles of patients living with facial palsy.

Why do we smile?

“The power of communication between the members of the same tribe by means of language has been of paramount importance in the development of man; and the force of language is much aided by the expressive movements of the face ...

... every true or inherited movement of expression seems to have had some natural and independent origin. But when once acquired, such movements may be voluntarily and consciously employed as a means of communication ...

... certain expressive states of the mind are the direct result of the constitution of the nervous system, independent of will and, to a large extent, of habit. These movements are the result of excessive nerve-force dependent on the connections of the nerve cells ...”

Charles Darwin *The Expression of Emotions in Man and Animals*. 1872.¹⁰

In the study of human emotion, one of the most intensely analyzed questions over the past century has been: “Why do we smile?”. To approach this question, it is useful to divide smiling into two categories: voluntary and spontaneous. A voluntary smile is a highly effective means of non-verbal communication, commonly used as a greeting and to send messages of agreement and empathy.^{1,2} A smile may also be spontaneous, believed to represent the experience of a positive emotion such as happiness or joy, often occurring during amusement or laughter.³ In facial palsy (FP), the absence of voluntary smiling can have profound impacts on one’s ability to engage in non-verbal communication and represents one of the most important problems reported by patients and their families, severely reducing their quality of life.⁴ However, amongst smile reanimation specialists, the recreation of a spontaneous smile (also known as an involuntary, true, or emotional smile) is often considered the most aspirational goal of reconstruction.^{5–7} Whether voluntary or spontaneous, many nuances of

smiling exist. Affective psychologists have tried to classify this variance, using our facial expressions to decode our emotional experience.

Sir Charles Bell, in his *Essays on the Anatomy of Expression in Paintings* (1806), was the first to analyze the expressions of man and relate them to an anatomical understanding of the facial muscles and nerves.⁸ In 1862, Duchenne de Boulogne, a French neurologist and electrophysiologist, first used electrical current and clinical photography to characterize individual facial expressions in living subjects.⁹ In 1872, Charles Darwin published *The Expression of the Emotions in Man and Animals*, in which he asserted that several of the chief facial expressions were universal between cultures and even between species, serving as evidence of a common ancestral linkage between humans and to an extent animals as well.¹⁰ During the twentieth century, many authors studied the specific muscles implicit in individual facial expressions of the basic emotions: anger, disgust, surprise, joy, fear, and sadness.^{3,11} Perhaps the most important contribution was development of a system for quantifying contributions of individual muscle groups to facial expressions, originally developed by Swedish anatomist, Carl-Herman Hjortsjö.¹¹ In 1978, Paul Ekman and Wallace Friesen published the facial action coding system (FACS) using frame-by-frame, expert-trained analysis of videos taken in emotion-generating situations.^{12–14} This system created a common data-driven language enabling the myriad facial expressions to be decoded into discrete muscle action units (AU).^{13,14}

According to Ekman, happiness, or joy, is expressed by activation of bilateral lip corner pull, cheek raise, and lid tighten leading to fine crease-lines in the area lateral to the orbital region colloquially known as crow's feet. The combination of all three of these action units is also commonly known as a Duchenne smile due to his earlier recognition of this pattern

characterizing a true, emotional smile. Disgust is composed of upper lip raise, wrinkled nose, and cheek raise, and contempt is a variant of disgust characterized by a unilateral upper lip raise also known as a smirk.^{13,15} See Table 1. Each AU can be modified by intensity codes ranging from weak to strong usually qualified by the depth of skin creases such as the nasolabial fold, crow's feet, and lower eyelid wrinkles. Psychologists have used this system to ask fundamental questions about the nature of emotion and also sought to quantify methods of suppressing emotions and the facial expressions of lying.^{3,13,15} However, whilst clearly a useful tool, FACS analysis of a video of facial expressions requires a high degree of training, is time-consuming, and has low inter-rater reliability.¹²

AU Number	FACS Name	Muscular Basis	Facial Display Modifier	Contribution to Emotion
6	Cheek Raiser	<i>orbicularis oculi (pars orbitalis)</i>	Crow's feet	Bilateral: Joy Unilateral: Contempt
7	Lid Tightener	<i>orbicularis oculi (pars palpebralis)</i>	Lower eyelid creases	Bilateral: Joy Unilateral: Contempt
9	Nose Wrinkler	<i>levator labii superioris alaeque nasi, nasalis</i>	Bunny lines	Bilateral: Disgust Unilateral: Contempt
10	Upper Lip Raiser	<i>levator labii superioris, caput infraorbitalis</i>	Bunny lines, Show canine	Bilateral: Disgust Unilateral: Contempt
11	Nasolabial Deepener	<i>zygomaticus minor</i>	Nasolabial fold	Bilateral: Joy Unilateral: Contempt
12	Lip Corner Puller	<i>zygomaticus major</i>	Nasolabial fold, Show premolars	Bilateral: Joy Unilateral: Contempt
14	Dimpler	<i>buccinator</i>	Dimple	Bilateral: Joy Unilateral: Contempt

Table 1. Muscle Action Units (AU) relevant to the facial expressions associated with smiling. Excerpt taken from the Facial Action Coding System (FACS) by Ekman and Friesen (1978).¹⁴

Recent advancements in machine learning facial recognition algorithms have enabled analysis of facial expressions to become automated.¹⁶ A commercially-available facial recognition and emotion detection system has been developed from 5 million videos of normal subjects from 75 countries resulting in over 2 billion facial frames (Affectiva, Boston, USA).^{17–}
²¹ This advance has enabled real-time analysis of the phenomenon known as emotionality, the observable behavioral and physiological component of emotion.²² It is important to note that facial expression analysis alone is unable to definitively determine the felt emotional state of an individual, as the system has no way of comprehending the context of the individual. Therefore, the results of any facial expression analysis ought to be considered a prediction of felt emotional state or, in the case of voluntary facial expression, the emotional state that the subject is trying to communicate.

Evolution of Surgical Techniques to Restore Smiling in Facial Palsy

There are many purported goals of facial reanimation including restoration of oral competence, eye closure, brow elevation, lip pucker and smile. However, when patients are asked to prioritise the goals of their own treatment, they usually focus on improvement of eye comfort, speech, swallowing, and return of their normal smile. Facial reanimation procedures can be divided into static and dynamic categories, where the primary goal of static procedures is to improve the symmetry of the face at rest and that of dynamic procedures is to improve the symmetry of the face during animation.

Smile reanimation procedures can be broken down into nerve transfers, regional muscle transfers and neurotised free muscle transfers. The respective advantages and disadvantages of each approach have been well studied and are traditionally measured by the overall success rate of the procedure, the complication rates and the degree of oral commissure excursion created by the procedure (measured in mm).²³ The ability of each procedure to deliver spontaneous smiling is the subject of much debate and until the present study, no objective tool has existed to measure its absence or presence in post-operative patients (see Chapter 5).

Smile reanimation procedures can also be divided into procedures that utilize the ipsilateral or contralateral facial nerve as a source of motor nerve input, versus procedures that utilize non- facial nerve sources, such as nerve to masseter, hypoglossal nerve, deep temporal nerve, and / or spinal accessory nerve.²³ In recent times dual or triple innervation techniques have also been developed to improve reliability of single nerve driven approaches and to improve spontaneity of smiling.⁵⁻⁷

It is widely accepted that the facial nerve is the best source of nerve input for creating spontaneous smiling in unilateral facial palsy. However, this neural input usually must be diverted from the healthy contralateral side of the face and can diminish in potency by the time the regenerated nerve axons reach the affected side.²⁴ Conversely, the nerve to masseter has a reliable pool of axons that can be diverted in very close proximity to the facial nerve branches and give motor input to the denervated facial muscles or free tissue muscle transfer.²³ The nerve to masseter is renowned for yielding high success rates of muscle reinnervation and strong oral commissure excursion measurements but lower rates of spontaneous smile.²⁵ Other neural inputs have also been used with variable degrees of reported success at restoring spontaneity including hypoglossal nerve, spinal accessory nerve and deep temporal nerve. What is less well known is the potential rate of spontaneity that can be generated from non-facial nerve inputs. Various authors attest that cortical adaptation occurs after non-facial nerve smile reanimation and that this is believed to be the cause of spontaneous smiles in the absence of facial nerve input.²⁶ An alternative hypothesis, particularly for nerve to masseter driven smiles, is that a subgroup of patients may co-contract their masseter and facial muscles when smiling *a priori* and thus post-operative spontaneity is a function of this pre-operative co-incidence.²⁷

In a recent systematic review of smile reanimation techniques, it was found that facial nerve driven reanimation did deliver the best chance of smile spontaneity and dual innervation techniques were superior to single non-facial nerve sources.²⁸ They also noted that only a minority of studies defined how they elicited or measured spontaneous smiles and an even smaller proportion used an independent means of assessment.

Facial palsy affects patients of all ages and all backgrounds. When it occurs at a very young age it may potentially be due to congenital birth defect, birth trauma or a developmental anomaly. It can also occur due to a defined cause such as a benign or malignant tumour, trauma, from infection or post-viral reactivation, such as in Bell's Palsy or Ramsay Hunt Syndrome (also known as Herpes Zoster Oticus). These clinical scenarios may present differing degrees of denervation to the facial musculature yet one constant remains that if the intrinsic facial musculature lacks a nerve supply completely for more than 18 months, the neuromuscular junction is very unlikely to be recovered even with the most robust reinnervation procedure.²³ As such, differing approaches are taken by facial reanimation specialists in various timepoints of recovery.

When within the 18 month window of potential reinnervation, direct nerve coaptation, nerve grafting, and nerve transfer procedures are the mainstay of treatments. However, once the facial musculature has been lost, after the 18 month window has closed, regional and free muscle transfers become the procedures of choice to restore animation to the smile. The most performed regional flap procedures in this situation are the temporalis muscle transfer, made popular by Gillies²⁹, McLaughlin³⁰ and most recently Labbé³¹, and the masseter muscle transfer has also been described.³² As expected the motor source for these muscle transfers is via the mandibular division of the trigeminal nerve and activation of the smile is achieved with a voluntary masticatory effort. It is believed that some smile spontaneity can be achieved in this patient population, particularly with extensive physiotherapy and neuromuscular retraining although objective testing is lacking.³³ The most frequently performed procedure to reanimate a long-standing facial palsy is functional free muscle transfer, utilising donor muscle options such as gracilis, pectoralis minor, latissimus dorsi, and serratus anterior. Free tissue transfers

are most often innervated by the nerve to masseter, hypoglossal, cross facial nerve graft (in one or two stages) or using a combination of multiple nerve sources.²³

A recent systematic review of 68 studies reporting the results of 2674 smile reanimation patients, comprising nerve transfer (15%), regional muscle transfer (6%) and free muscle transplantation (79%) procedures, found that facial nerve-driven smile reanimation resulted in the highest proportion of spontaneous smiles (83%).²⁸ This was followed by hypoglossal (45%), masseteric (32%) and spinal accessory nerve-driven smile reanimation (32%). They also found that dual innervation resulted in a higher proportion of spontaneous smiles and postoperative physical therapy resulted in a significantly higher proportion of spontaneous smiles. Importantly, they noted that a minority of studies reported their method of assessing spontaneous smile outcome and even fewer used an independent means of assessment. Given the amount of time both patient and surgeon have invested in generating a positive outcome in smile reanimation procedures there is a strong possibility of reporter bias, and the lack of objective assessment methods raises concern as to the validity of these figures.³⁴

List of References

1. Adolphs R. The Social Brain: Neural Basis of Social Knowledge. *Annu Rev Psychol.* 2009;60(1):693-716. doi:10.1146/annurev.psych.60.110707.163514
2. Frith C. Role of facial expressions in social interactions. *Philos Trans R Soc B Biol Sci.* 2009;364(1535):3453-3458. doi:10.1098/rstb.2009.0142
3. Ekman P, Davidson R. *The Nature of Emotion: Fundamental Questions.* New York: Oxford University Press; 1994.
4. Coulson SE, O'dwyer NJ, Adams RD, Croxson GR. Expression of emotion and quality of life after facial nerve paralysis. *Otol Neurotol.* 2004;25(6):1014-1019. doi:10.1097/00129492-200411000-00026
5. Biglioli F, Colombo V, Tarabbia F, et al. Double innervation in free-flap surgery for long-standing facial paralysis. *J Plast Reconstr Aesthet Surg.* 2012;65(10):1343-1349. doi:10.1016/j.bjps.2012.04.030
6. Watanabe Y, Akizuki T, Ozawa T, Yoshimura K, Agawa K, Ota T. Dual innervation method using one-stage reconstruction with free latissimus dorsi muscle transfer for re-animation of established facial paralysis: simultaneous reinnervation of the ipsilateral masseter motor nerve and the contralateral facial nerve to im. *J Plast Reconstr Aesthet Surg.* 2009;62(12):1589-1597. doi:10.1016/j.bjps.2008.07.025
7. Cardenas-Mejia A, Covarrubias-Ramirez JV, Bello-Margolis A, Rozen S. Double innervated free functional muscle transfer for facial reanimation. *J Plast Surg Hand Surg.* 2015;49(3):183-188. doi:10.3109/2000656X.2014.988218
8. Bell C. *Essays on the Anatomy of Expression in Painting.* London: Longman, Hurst, Rees, and Orme; 1806.
9. Duchenne. *Mécanisme de La Physionomie Humaine.*; 1876.
10. Darwin C. *The Expression of Emotions in Man and Animals.* Vol 2.; 1872.
11. Hjortsjö CH. *Man's Face and Mimic Language.* Studentlitteratur; 1970. <https://books.google.com/books?id=r68ZMwEACAAJ>.
12. Ekman P. Facial expression and emotion. *Am Psychol.* 1993;48(4):384-392. doi:10.1037/0003-066x.48.4.384
13. Ekman P, Rosenberg E. *What the Face Reveals.*; 2005.
14. Ekman P, Friesen W V. *Facial Action Coding System: A Technique for the Measurement of Facial Movement.*; 1978.
15. Ekman P, Friesen W V. *Unmasking the Face: A Guide to Recognizing Emotions from Facial Clues.*; 1975. doi:10.1163/1574-9347_bnp_e804940
16. Pantic M, Rothkrantz LJM. Automatic analysis of facial expressions: The state of the art. *IEEE Trans Pattern Anal Mach Intell.* 2000;22(12):1424-1445. doi:10.1109/34.895976
17. McDuff D, El Kaliouby R, Senechal T, Demirdjian D, Picard R. Automatic measurement of ad preferences from facial responses gathered over the Internet. *Image Vis Comput.* 2014;32(10):630-640. doi:10.1016/j.imavis.2014.01.004
18. El Kaliouby R, Robinson P. Real-time inference of complex mental states from facial expressions and head gestures. In: *Real-Time Vision for Human-Computer Interaction.* ; 2005:181-200. doi:10.1007/0-387-27890-7_11
19. McDuff D, El Kaliouby R, Picard RW. Crowdsourcing facial responses to online videos: Extended abstract. In: *2015 International Conference on Affective Computing and Intelligent Interaction, ACII 2015.* ; 2015:512-518. doi:10.1109/ACII.2015.7344618
20. Baltrušaitis T, McDuff D, Banda N, et al. Real-time inference of mental states from facial expressions and upper body gestures. In: *2011 IEEE International Conference on Automatic Face and Gesture Recognition and Workshops, FG 2011.* ; 2011:909-914.

- doi:10.1109/FG.2011.5771372
21. Rana el Kaliouby, Evan Kodra, Pankaj Jha NM. Does Facial Coding Generalize Across Cultures? A Spotlight on Asia. *Asia Pacific 2014 - Celebr Asian Creat.* 2014;(May):1-18. http://www.affectiva.com/wp-content/uploads/2014/09/Does_Facial_Coding_Generalize_Across_Cultures_ASIA.pdf.
 22. Reber AS, Allen R, Reber ES. Penguin Dictionary of Psychology. *Penguin Dict Psychol.* 2009:891.
 23. Pan DR, Clark NW, Chiang H, Kahmke RR, Phillips BT, Barrett DM. The evolution of facial reanimation techniques. *Am J Otolaryngol.* 2023 May-Jun;44(3):103822. doi: 10.1016/j.amjoto.2023.103822.
 24. Hembd A, et al. Correlation between facial nerve axonal load and age and its relevance to facial reanimation. *Plast Reconstr Surg* 2017;139(6):1459–64.
 25. Klebuc MJA. Facial reanimation using the masseter-to-facial nerve transfer. *Plast Reconstr Surg* 2011;127(5):1909–15
 26. Buendia J, Loayza FR, Luis EO, Celorrio M, Pastor MA, Hontanilla B. Functional and anatomical basis for brain plasticity in facial palsy rehabilitation using the masseteric nerve. *J Plast Reconstr Aesthet Surg.* 2016 Mar;69(3):417-26. doi: 10.1016/j.bjps.2015.10.033.
 27. Lenz Y, Kiefer J, Dietrich F, Stark GB, Eisenhardt SU. Pre-operative masseter muscle EMG activation during smile predicts synchronicity of smile development in facial palsy patients undergoing reanimation with the masseter nerve: A prospective cohort study. *J Plast Reconstr Aesthet Surg.* 2019 Mar;72(3):505-512. doi: 10.1016/j.bjps.2018.11.011.
 28. van Veen MM, Dekker A, Werker P, Dijkstra P. Spontaneous smile after smile reanimation surgery: a systematic review and meta-analysis. Submitted as part of a PhD Thesis, University of Groningen, Groningen, the Netherlands
 29. Gillies H. Experiences with fascia lata grafts in the operative treatment of facial paralysis. *Proc R Soc Med.* 1934;27:1372–1382.
 30. McLaughlin CR. Surgical support in permanent facial paralysis. *Plast Reconstr Surg.* 1953;11:302–314. (1946)
 31. Labbè D, Bussu F, Iodice A. A comprehensive approach to long-standing facial paralysis based on lengthening temporalis myoplasty. *Acta Otorhinolaryngol Ital.* 2012 Jun;32(3):145-53.
 32. Lesavoy MA, Fan KL, Goldberg AG, Dickinson BP, Herrera F. Facial reanimation by staged, split masseter muscle transfer. *Ann Plast Surg.* 2014 Jul;73(1):33-8. doi: 10.1097/SAP.0b013e31826cb267.
 33. Hontanilla B, Cabello A. Spontaneity of smile after facial paralysis rehabilitation when using a non-facial donor nerve. *J Craniomaxillofac Surg.* 2016 Sep;44(9):1305-9. doi: 10.1016/j.jcms.2016.06.031. Epub 2016 Jul 9.
 34. Dusseldorp JR, van Veen MM, Guarin DL, Quatela O, Jowett N, Hadlock TA. Spontaneity Assessment in Dually Innervated Gracilis Smile Reanimation Surgery. *JAMA Facial Plast Surg.* 2019 Dec 1;21(6):551-557. doi: 10.1001/jamafacial.2019.1090.

CHAPTER TWO

Objective Outcome Assessment in Facial Reanimation

Preface

“When you can measure what you are speaking about, and express it in numbers, you know something about it, but when you cannot express it in numbers, your knowledge is of a meager and unsatisfactory kind.”

Lord Kelvin. Electrical Units of Measurement, delivered 3 May 1883

As new surgical techniques have evolved and evidence-based practice has been embraced within the medical specialties, it has become necessary to express the likely post-surgical outcomes in an objective manner. In the field of facial palsy, clinician-graded tools were created to assess the severity of facial palsy at diagnosis. However, these are not sufficiently sensitive to quantify post-operative changes and were never developed to do so. In the late 20th century, a new raft of outcome measures were developed to capture the various domains of post-operative recovery more fully. The aim of this chapter was to detail and critique outcome measures following facial reanimation surgery and propose a multi-modal approach to objective outcome measurement. This chapter was published in *Facial Palsy*. Tzou CH.J., Rodríguez-Lorenzo A. (eds). Springer, Cham., 2021.

Dusseldorp JR, Meares C, Hadlock TA. A Multi-modal Approach to Outcome Tracking in Facial Palsy. In: Tzou CH.J., Rodríguez-Lorenzo A. (eds) *Facial Palsy*. 2021. Springer, Cham. https://doi.org/10.1007/978-3-030-50784-8_2

One of the major challenges to successfully validating outcome measures, particularly automated ones, is to ensure their ability to measure various degrees of difference between patients of all degrees of severity within the spectrum of facial palsy. Appendix 1 details the development steps taken to create a freely-available, comprehensive database of facial

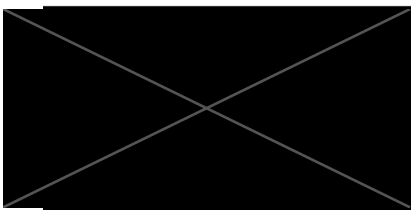
photographs of patients with all possible degrees of facial palsy. This data set is the ideal testing group to validate a facial palsy outcome measure against the full spectrum of disease severity and its open source nature makes rapid validation of novel tools for outcome assessment possible. This paper was evaluated and findings published in *Laryngoscope*, 2020.

Greene JJ, Guarin DL, Tavares J, Fortier E, Robinson M, **Dusseldorp JR**, Quatela O, Jowett N, Hadlock TA. The Spectrum of Facial Palsy: The MEEI Facial Palsy Photo and Video Standard Set. *Laryngoscope*. 2020 Jan;130(1):32-37

Authorship Statement

I, Joseph Dusseldorp, confirm for the ensuing textbook chapter: ‘**A Multi-modal Approach to Outcome Tracking in Facial Palsy**’ that I have made the following contributions to its authorship:

- Conception and design of the research
- Collection and extraction of data
- Analysis and interpretation of the findings
- Drafting and revising of the manuscript and critical appraisal of content



Joseph R. Dusseldorp

Date 14.09. 2023

As Auxiliary Supervisor for the candidature upon which this thesis is based, I can confirm that the above authorship attribution is correct.

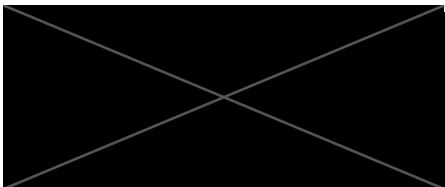


Professor Tessa Hadlock, MD

Date 27.07. 2023

Department of Otolaryngology
Massachusetts Eye and Ear Infirmary
Harvard University, Boston, MA, USA

As Primary Supervisor of the candidate, I concur with the above authorship attestation:



Professor Peter Maitz

Date 14.09.2023

Department of Burns Surgery
Concord Repatriation General Hospital
Sydney University, Sydney, NSW, Australia

A Multi-modal Approach to Outcome Tracking in Facial Palsy

Joseph R. Dusseldorp, Charles Meares,
and Tessa A. Hadlock

Key Points

- Outcome assessment in facial palsy ought to be multi-modal, including patient-reported outcome measures, clinician-graded scoring systems, objective assessment tools, and layperson and spontaneity assessment.
- Patient-reported outcome measures are required to assess the psychosocial impact of facial nerve dysfunction, as experienced by patients.
- Clinician-graded scoring systems and objective assessment tools are needed to rate severity of disease and the effects of interventions. Time and cost constraints are being decreased with automation.
- Layperson assessments have been used to determine impacts of facial palsy on attractiveness, disfigurement, and non-verbal communication.

J. R. Dusseldorp (✉)

Department of Otolaryngology/Head and Neck Surgery, Massachusetts Eye and Ear Infirmary and Harvard Medical School, Boston, MA, USA

Department of Plastic and Reconstructive Surgery, Royal Australasian College of Surgeons and University of Sydney, Sydney, NSW, Australia

C. Meares

Department of Plastic and Reconstructive Surgery, Royal Australasian College of Surgeons and University of Sydney, Sydney, NSW, Australia

T. A. Hadlock

Department of Otolaryngology/Head and Neck Surgery, Massachusetts Eye and Ear Infirmary and Harvard Medical School, Boston, MA, USA
e-mail: Tessa_Hadlock@meei.harvard.edu

- A novel approach using machine learning algorithms to assess the emotionality of smiling has been developed and shown to correlate with layperson assessments. A means of objectively quantifying spontaneity of smile has also been developed.

2.1 Introduction

Outcome tracking in facial palsy is critical in order to improve future outcomes. Currently, outcome measurement tools in facial palsy can be classified into patient-reported outcome measures, clinician-graded scoring systems, objective assessment tools, and emerging systems assessing spontaneity and incorporating layperson assessments. The ideal outcome measurement approach must address the complex dysfunction that occurs in facial palsy, by assessing patient-experienced morbidity, and outcomes of interventions from both the patient and clinician perspective. This approach should also make objective measurements and provide a common language for future research [1]. Clearly a panel of outcome measures are required to capture all of these various domains, and such a panel should be sensitive to change, simple, and low cost to encourage widespread implementation [2].

2.2 Patient-Reported Outcome Measures

Patient-reported outcome measures (PROMs) should be utilized to capture the significant psychological and quality of life impacts associated with facial nerve palsy [3–5]. However, a recent survey of clinicians observed that the majority do not employ a PROM in the assessment of facial palsy patients [6]. The Facial Disability Index (FDI) and the Facial Clinimetric Evaluation (FaCE) scale are facial palsy specific surveys that assess the functioning of facial muscles from the patient's perspective and the effect that dysfunction has on patients' quality of life [7, 8]. Both have been validated; however, neither specifically assess the role of synkinesis in the patient's burden of morbidity [9]. To this end, the Synkinesis Assessment Questionnaire (SAQ) can be used as an adjunct or independent PROM tool for assessment of synkinetic impacts on patients' quality of life.

Introduced in 2001, the Facial Clinimetric Evaluation (FaCE) scale aims to measure both facial impairment and quality of life [8]. It has quickly become the most widely used PROM in facial palsy patients [6, 10]. The tool includes Likert-type questions and visual analogue scales covering facial movement, oral movement, lacrimal control, social function, facial comfort, and eye comfort. Scores for these domains are summed to yield the FaCE scale score. This scale ranges from 0 (worst) to 100 (best). The FaCE scale has demonstrated concurrent and convergent validity with the HBGS and is thought to assess psychometric outcomes more completely than other PROMs [9].

The Facial Disability Index (FDI) consists of ten Likert-type questions addressing physical function and social function [7]. The index includes questions covering normal activities of daily living, such as drinking from a cup. The scores are collated giving a score from 0 (worst) to 100 (best). Developed in 1996, the tool has undergone validation testing [7, 9].

The Synkinesis Assessment Questionnaire (SAQ) was developed in 2007 in order to specifically address the deficiency of established

PROMs to observe synkinesis-specific issues for patients [11]. The SAQ consists of nine items answered with five Likert-type responses. The answers are summated yielding a range from 0 to 100, higher scores representing worse synkinesis. The SAQ has demonstrated good retest reliability and construct validity [11, 12]. SAQ focuses on patient experience with synkinesis and can be combined as an adjunct with a facial palsy specific PROM.

2.3 Clinician Grading Scoring Systems

Historically, various clinician-graded scoring systems have been advocated with no single scale being universally utilized. Such tools are subjective assessment of facial function and have continually been proposed over the past 50 years [13]. Each tool assesses static and dynamic asymmetry. Only some scales contain regional assessment or determine degree of synkinesis.

The most widely used scale for facial nerve severity assessment is House–Brackmann grading scale (HBGS) [14]. After being adopted by the American Academy of Otolaryngology-Head and Neck Surgery in 1985, it remains the standard in the United States. The HBGS assesses overall facial nerve function using a six-point scale, from normal function to total paralysis. It accounts for secondary defects and synkinesis, can measure large changes over time and has good interobserver and intraobserver reliability [14–22]. HBGS has been modified several times and used to measure outcomes in facial nerve reanimation, though this was not its original purpose [20, 23]. There are no regional subdomain scores in HBGS and so it is insensitive to regional improvements, which limits its utility in outcome measurement [24–27]. However, its simplicity and widespread use have ensured it remains a favored communication tool between clinicians.

The Sunnybrook Facial Grading Scale (SFGS) is a regional-based composite score, assessing facial symmetry at rest, symmetry with voluntary facial movements and synkinesis [26]. It was developed in 1996 in order to address limita-

tions of the HBGS. Calculation of the SFGS is a three-step process: quantify the symmetry of the eye, cheek, and mouth at rest; assess five standard facial movements; and evaluate the face for synkinesis during facial expressions. These component scores are weighted and the composite score is produced on a continuous scale (0–100). Compared to HBGS, SFGS has a greater ability to detect change [28] and has high intraobserver and interobserver reliability irrespective of observer experience [19, 24, 29–31]. The scale is time-consuming to employ and retains an element of subjectivity.

The Sydney facial grading system is similar to the Sunnybrook system, involving a regional assessment of facial movement. Unlike the Sunnybrook scale, the Sydney approach breaks down the facial regions based on branches of the facial nerve and corresponding muscle groups [19]. The Sydney system has strong correlation with the Sunnybrook system [19]; however, it is not widely utilized and lacks proper assessment of intraobserver reliability.

Noting the limitations of the HBGS, the revised Facial Nerve Grading Scale 2.0 (FNGS2.0) was introduced in 2006 [32]. This system assesses four regions of the face based on movement (scored 1–6) and accounts for synkinesis (scored 0–3). The summed score is then converted to a HBGS grade. Studies have shown high correlation between the FNGS2.0 and HBGS with consistently high interobserver reliability [13, 21, 27, 32, 33]. The FNGS2.0 has been shown to better measure synkinesis when compared to the HBGS over time [21].

The Yanagihara scale was introduced in 1976 and was one of the earliest widely accepted outcome measure systems and remains in use in Japan [34]. The Yanagihara method is a regional system that assesses ten facial movements which are graded on a four-point scale. The summed score characterizes the spectrum of facial palsy from normal movement to total paralysis. Facial palsy resolution is defined as a score greater than 36 with no associated synkinesis or facial muscle contraction [34, 35]. The Yanagihara scale provides useful regional information and correlates well with patient self-evaluation [36]. However,

the scale is unweighted, potentially resulting in an inability to highlight specific facial movements [18], only captures ocular synkinesis, and has low intraobserver reliability.

The Terzis Facial Grading System (TFGS) observes the cosmetic and functional outcome of facial reanimation with respect to contraction and symmetry [37]. It was developed in 1997 specifically as a grading tool for comparing surgical outcome with pre-operative state. It is easy to implement, offers dynamic movement assessment and evaluates important endpoints: symmetry at rest, and with smiling [38, 39]. Its use has been widespread due to its ease of application [40–45]. Unfortunately, it lacks regional assessment of function, ignores synkinesis and has not been comprehensively validated.

eFACE is a comprehensive application for clinician grading of facial symmetry in unilateral facial palsy [46]. Using a digital interface, the observer scores 15 visual analogue scales with regard to static, dynamic movement and synkinesis. The complete spectrum of possible dysfunction is represented by the sum score of 0–100, with 100 representing normal function. The eFACE application has been demonstrated to be quick, sensitive to change, have high resolution for regional scoring (which have been weighted according to expert consensus [47]), accurately assesses synkinesis and have high inter-, intraobserver and retest reliability [46, 48, 49]. The eFACE application enables a quick assessment of facial palsy outcome in a zonal manner, yet remains a subjective tool.

2.4 Objective Assessment Systems

Objective assessment systems have been developed to evaluate outcomes in a consistent manner [2, 13]. Photography with staged facial movements is the most commonly utilized objective tool by clinicians [6], yet video-based analysis techniques remain poorly integrated into clinical practice as they can be expensive and time-consuming. The increasing utilization of machine learning in medicine has enabled development of

efficient, cheap, and objective measurement systems in facial palsy patients [1, 50, 51].

The Facial Assessment by Computer Evaluation (FACE) application is more commonly referred to as the MEEI Facegram. Developed in 2012 and available from the Sir Charles Bell Society website, the MEEI Facegram can assess any frontal view photograph [2]. It is relatively quick and easy to implement as it does not require markers to be placed. Utilizing the average iris diameter of 11.8 mm, measurements of excursion of the eyebrow, palpebral fissure width, excursion of the commissure, and lip measurements are calculated by the user, scaled and converted into a vector.

The application of machine learning to facial palsy photography assessment is represented by the Emotrics objective tool [50]. Developed in 2018 and available as a freeware application, the Emotrics package is a facial landmark algorithm that automatically calculates facial displacements from standard frontal photographs. Smile excursion and symmetry, brow symmetry, and palpebral width are measured. Additionally, the application can also assess differences between two separate images, thus proving useful tool to automatically compare pre- and post-operative photographs [50].

An innovative videography assessment tool was developed by Frey et al. in 1994 [52]. Utilizing markers attached to facial landmarks and four video cameras in conjunction with a computer, a dynamic 3D video analysis was obtained. The system identified pairs of facial points to compare movement of the brows, eyes, nose, and mouth. Unfortunately, implementing such a system in clinical practice was time-consuming and expensive, preventing its widespread adoption. Subsequent work simplified this method to a single camera, calibration grid and mirror system located behind the patient [53]. Paired with a facial analysis software package, this has resulted in an automatic standardized objective assessment tool that is a sensitive and reliable quantitative facial outcome measurement tool [54].

An alternative 3D videography assessment tool is the FACIAL CLIMA system [55].

Developed in 2008, the system requires reflective markers and three positioned infrared cameras to evaluate facial movements for the patient. The system is capable of calculating distances between landmarks, velocities, and areas for facial palsy analysis [55]. It is easy to use and quick; however, similar to other 3D videography tools, the need for specialized equipment and cost have prevented its widespread adoption.

3D stereophotogrammetry techniques are becoming more commonplace and may ultimately supersede 2D landmark-based assessment [56], though presently these tools are used predominantly in research settings [57]. A major advantage of 2D outcome measurement systems is that imaging is simple to record using standard cameras, large databases currently exist, to which machine learning approaches can be applied [50, 58].

2.5 Emerging Systems

Recent studies have characterized the layperson impression of facial palsy, highlighting the important role of dynamic facial movements, particularly smiling [59–62]. The layperson may be the ideal assessor of a facial palsy outcome as they ultimately decide if patients are able to communicate emotion non-verbally or appear disfigured when attempting to smile [1, 63]. In practice, routine layperson assessments are impractical due to logistical issues and patient confidentiality.

Layperson assessment studies have identified the negative aesthetic consequence of unilateral facial palsy in patients [60, 64]. This is a complex domain for assessment and is potentially best suited to the layperson interpretation [1]. In a layperson comparison of normal subjects and facial palsy patients, significantly lower attractiveness scores were reported for patients with facial paralysis, which was improved by reanimation surgery [60]. In an effort to better assess disfigurement in facial paralysis, over 500 laypersons assessed disfigurement in facial paralysis patients performing standard facial expressions. The layperson's disfigurement rating was signifi-

cantly correlated with the clinician-graded facial function (eFACE). Additionally, it was noted that patients with worse dynamic subscores were considered more disfigured by laypersons [62].

Facial movements play an integral role in non-verbal communication, which is greatly impaired in facial palsy patients [3, 65, 66]. The layperson assessment is sensitive to these differences. Machine learning tools developed on large databases of facial expressions once validated could potentially serve as a proxy for the layperson assessment, circumventing the issues of privacy and practicality involved with utilizing layperson assessment in practice [1, 51].

Spontaneous smiling is an important outcome of smile reanimation and there are a number of different approaches to its assessment. Though various studies have defined spontaneous smile, currently no standardized assessment tool has been adopted in practice [67–71]. Despite the proposed “Tickle test” to elicit the patient to laugh, many authors make their assessment during a conversation with the patient [69, 70, 72, 73]. Such a process is open to confounding and confirmation bias [1, 74]. A standardized means of eliciting a spontaneous smile is through a standardized series of humorous videoclips. The spontaneous smile assay has been validated to elicit at least one spontaneous smile in over 95% of both normal subjects and facial palsy patients [67]. Quantifying the spontaneous smile outcome following reanimation surgery is extremely difficult and also prone to bias and error when performed by clinicians [74]. Another machine learning program has been developed to assess the emotionality of spontaneous smiling from videoclips of facial palsy patients watching humorous videos [51, 74]. Hopefully, objective tools such as these will lead to improvement of spontaneity outcomes for patients.

2.6 Conclusion

No ideal single tool exists that is appropriate for facial palsy outcome tracking. Clinician grading scoring systems remain widely utilized, particularly for ease of communication. PROMs should

be implemented in order to gain an understanding of the impact of interventions on patient-reported quality of life. Objective assessment tools are available and will gain greater adoption in clinical practice as their utility is better understood. Layperson assessments, and automated equivalents, will continue to give valuable insights into interesting subdomains of facial function such as emotionality and spontaneity following reanimation.

In summary, the complete approach to outcome measures in facial palsy (P.A.L.Sy) should include [1]:

- Patient-reported outcome measures
- Automated and clinician-graded scoring systems
- Layperson assessment equivalent
- Spontaneous smile analysis

Disclosure Statement Dr. Hadlock, Dr. Dusseldorp, and Dr. Meares have nothing to disclose. No funding was received for this article.

References

1. Dusseldorp JR, van Veen MM, Mohan S, Hadlock TA. Outcome tracking in facial palsy. *Otolaryngol Clin N Am*. 2018;51(6):1033–50. <https://doi.org/10.1016/j.otc.2018.07.003>.
2. Hadlock TA, Urban LS. Toward a universal, automated facial measurement tool in facial reanimation. *Arch Facial Plast Surg*. 2012;14(4):277–82. <https://doi.org/10.1001/archfacial.2012.111>.
3. Coulson SE, O’dwyer NJ, Adams RD, Croxson GR. Expression of emotion and quality of life after facial nerve paralysis. *Otol Neurotol*. 2004;25(6):1014–9. <https://doi.org/10.1097/00129492-200411000-00026>.
4. Kleiss IJ, Beurskens CHG, Stalmeier PFM, Ingels KJAO, Marres HAM. Quality of life assessment in facial palsy: validation of the Dutch Facial Clinimetric Evaluation Scale. *Eur Arch Otorhinolaryngol*. 2015;272(8):2055–61. <https://doi.org/10.1007/s00405-015-3508-x>.
5. Lindsay RW, Bhama P, Hadlock TA. Quality-of-life improvement after free gracilis muscle transfer for smile restoration in patients with facial paralysis. *JAMA Facial Plast Surg*. 2014;16(6):419. <https://doi.org/10.1001/jamafacial.2014.679>.
6. Fattah AY, Gavilan J, Hadlock TA, et al. Survey of methods of facial palsy documentation in use by members of the Sir Charles Bell Society. *Laryngoscope*. 2014;124(10):2247–51. <https://doi.org/10.1002/lary.24636>.

7. VanSwearingen JM, Brach JS. The Facial Disability Index: reliability and validity of a disability assessment instrument for disorders of the facial neuromuscular system. *Phys Ther.* 1996;76(12):1288–98; discussion 1298–300. <http://www.ncbi.nlm.nih.gov/pubmed/8959998>.
8. Kahn JB, Gliklich RE, Boyev KP, Stewart MG, Metson RB, McKenna MJ. Validation of a patient-graded instrument for facial nerve paralysis: the FaCE scale. *Laryngoscope.* 2001;111(3):387–98. <https://doi.org/10.1097/00005537-200103000-00005>.
9. Ho AL, Scott AM, Klassen AF, Cano SJ, Pusic AL, Van Laeken N. Measuring quality of life and patient satisfaction in facial paralysis patients: a systematic review of patient-reported outcome measures. *Plast Reconstr Surg.* 2012;130(1):91–9. <https://doi.org/10.1097/PRS.0b013e318254b08d>.
10. Luijmes RE, Pouwels S, Beurskens CHG, Kleiss IJ, Siemann I, Ingels KJAO. Quality of life before and after different treatment modalities in peripheral facial palsy: a systematic review. *Laryngoscope.* 2017;127(5):1044–51. <https://doi.org/10.1002/lary.26356>.
11. Mehta RP, WernickRobinson M, Hadlock TA. Validation of the Synkinesis Assessment Questionnaire. *Laryngoscope.* 2007;117(5):923–6. <https://doi.org/10.1097/MLG.0b013e3180412460>.
12. Rhee JS, McMullin BT. Outcome measures in facial plastic surgery: patient-reported and clinical efficacy measures. *Arch Facial Plast Surg.* 2008;10(3):194–207. <https://doi.org/10.1001/archfaci.10.3.194>.
13. Fattah AY, Gurusinge ADR, Gavilan J, et al. Facial nerve grading instruments: systematic review of the literature and suggestion for uniformity. *Plast Reconstr Surg.* 2015;135(2):569–79. <https://doi.org/10.1097/PRS.0000000000000905>.
14. House JW, Brackmann DE. Facial nerve grading system. *Otolaryngol Head Neck Surg.* 1985;93(2):146–7. <https://doi.org/10.1177/019459988509300202>.
15. Burres S, Fisch U. The comparison of facial grading systems. *Arch Otolaryngol Head Neck Surg.* 1986;112(7):755–8.
16. Lewis BI, Adour KK. An analysis of the Adour-Swanson and House-Brackmann grading systems for facial nerve recovery. *Eur Arch Otorhinolaryngol.* 1995;252(5):265–9. <https://doi.org/10.1007/BF00185387>.
17. Saito H. A simple objective evaluation and grading for facial paralysis outcomes. *Acta Otolaryngol.* 2012;132(1):101–5. <https://doi.org/10.3109/00016489.2011.626793>.
18. Berg T, Jonsson L, Engström M. Agreement between the Sunnybrook, House-Brackmann, and Yanagihara facial nerve grading systems in Bell's palsy. *Otol Neurotol.* 2004;25(6):1020–6. <https://doi.org/10.1097/00129492-200411000-00027>.
19. Coulson SE, Croxson GR, Adams RD, O'Dwyer NJ. Reliability of the "Sydney," "Sunnybrook," and "House Brackmann" facial grading systems to assess voluntary movement and synkinesis after facial nerve paralysis. *Otolaryngol Head Neck Surg.* 2005;132(4):543–9. <https://doi.org/10.1016/j.otohns.2005.01.027>.
20. Reitzen SD, Babb JS, Lalwani AK. Significance and reliability of the House-Brackmann grading system for regional facial nerve function. *Otolaryngol Head Neck Surg.* 2009;140(2):154–8. <https://doi.org/10.1016/j.otohns.2008.11.021>.
21. Lee HY, Moon SP, Byun JY, Chung JH, Na SY, Yeo SG. Agreement between the facial nerve grading system 2.0 and the House-Brackmann grading system in patients with Bell palsy. *Clin Exp Otorhinolaryngol.* 2013;6(3):135–9. <https://doi.org/10.3342/ceo.2013.6.3.135>.
22. Smith IM, Murray JAM, Cull RE, Slattery J. A comparison of facial grading systems. *Clin Otolaryngol Allied Sci.* 1992;17(4):303–7. <https://doi.org/10.1111/j.1365-2273.1992.tb01001.x>.
23. Lazarini P, Mitre E, Takatu E, Tidei R. Graphic-visual adaptation of House-Brackmann facial nerve grading for peripheral facial palsy. *Clin Otolaryngol.* 2006;31(3):192–7. <https://doi.org/10.1111/j.1749-4486.2006.01197.x>.
24. Kanerva M, Poussa T, Pitkäranta A. Sunnybrook and House-Brackmann Facial Grading Systems: intrarater repeatability and interrater agreement. *Otolaryngol Head Neck Surg.* 2006;135(6):865–71. <https://doi.org/10.1016/j.otohns.2006.05.748>.
25. Croxson G, May M, Mester SJ. Grading facial nerve function: House-Brackmann versus Burres-Fisch methods. *Am J Otol.* 1990;11(4):240–6.
26. Ross BG, Fradet G, Nedzelski JM. Development of a sensitive clinical facial grading system. *Otolaryngol Head Neck Surg.* 1996;114(3):380–6. [https://doi.org/10.1016/S0194-5998\(96\)70206-1](https://doi.org/10.1016/S0194-5998(96)70206-1).
27. Henstrom DK, Skilbeck CJ, Weinberg J, Knox C, Cheney ML, Hadlock TA. Good correlation between original and modified House Brackmann facial grading systems. *Laryngoscope.* 2011;121(1):47–50. <https://doi.org/10.1002/lary.21163>.
28. Kanerva M, Jonsson L, Berg T, et al. Sunnybrook and House-Brackmann systems in 5397 facial gradings. *Otolaryngol Head Neck Surg.* 2011;144(4):570–4. <https://doi.org/10.1177/0194599810397497>.
29. Ahrens A, Skarada D, Wallace M, Cheung JY, Neely JG. Rapid simultaneous comparison system for subjective grading scales grading scales for facial paralysis. *Am J Otol.* 1999;20(5):667–71.
30. Kayhan FT, Zurakowski D, Rauch SD. Toronto facial grading system: interobserver reliability. *Otolaryngol Head Neck Surg.* 2000;122(2):212–5. [https://doi.org/10.1016/S0194-5998\(00\)70241-5](https://doi.org/10.1016/S0194-5998(00)70241-5).
31. Hu WL, Ross B, Nedzelski J. Reliability of the Sunnybrook Facial Grading System by novice users. *J Otolaryngol.* 2001;30(4):208–11.
32. Vrabec JT, Backous DD, Djalilian HR, et al. Facial Nerve Grading System 2.0. *Otolaryngol Head Neck Surg.* 2009;140(4):445–50. <https://doi.org/10.1016/j.otohns.2008.12.031>.
33. Lee LN, Susarla SM, Hohman MH, Henstrom DK, Cheney ML, Hadlock TA. A comparison of facial nerve

- grading systems. *Ann Plast Surg.* 2013;70(3):313–6. <https://doi.org/10.1097/SAP.0b013e31826acb2c>.
34. Yanagihara N. Grading of facial nerve palsy. In: Fisch U, editor. *Facial nerve surgery, proceedings: third international symposium on facial nerve surgery.* Birmingham, AL: Aesculapius Publishing; 1977. p. 533–5.
 35. Hato N, Fujiwara T, Gyo K, Yanagihara N. Yanagihara facial nerve grading system as a prognostic tool in Bell's palsy. *Otol Neurotol.* 2014;35(9):1669–72. <https://doi.org/10.1097/MAO.0000000000000468>.
 36. Ikeda M, Nakazato H, Hiroshige K, Abiko Y, Sugiura M. 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. *Otol Neurotol.* 2003;24(2):334–8.
 37. Terzis JK, Noah ME. Analysis of 100 cases of free-muscle transplantation for facial paralysis. *Plast Reconstr Surg.* 1997;99(7):1905–21.
 38. Biglioli F, Frigerio A, Rabbiosi D, Brusati R. Single-stage facial reanimation in the surgical treatment of unilateral established facial paralysis. *Plast Reconstr Surg.* 2009;124(1):124–33. <https://doi.org/10.1097/PRS.0b013e3181aa0e2b>.
 39. Brichacek M, Sultan B, Boahene KD, Ishii L, Byrne PJ. Objective outcomes of minimally invasive temporalis tendon transfer for prolonged complete facial paralysis. *Plast Surg (Oakville, Ont).* 2017;25(3):200–10. <https://doi.org/10.1177/2292550317728033>.
 40. Balaji S. Temporalis pull-through vs fascia lata augmentation in facial reanimation for facial paralysis. *Ann Maxillofac Surg.* 2016;6(2):267. <https://doi.org/10.4103/2231-0746.200323>.
 41. Allevi F, Motta G, Colombo V, Biglioli F. Double-bellied latissimus dorsi free flap to correct full dental smile palsy. *BMJ Case Rep.* 2015;2015. <https://doi.org/10.1136/bcr-2015-210436>.
 42. Ahuja R, Chatterjee P, Gupta R, Gupta G, Shrivastava P. A new paradigm in facial reanimation for long-standing palsies? *Indian J Plast Surg.* 2015;48(1):30. <https://doi.org/10.4103/0970-0358.155265>.
 43. Gordin E, Lee TS, Ducic Y, Arnaoutakis D. Facial nerve trauma: evaluation and considerations in management. *Craniomaxillofac Trauma Reconstr.* 2015;8(1):1–13. <https://doi.org/10.1055/s-0034-1372522>.
 44. Leckenby J, Grobbelaar A. Smile restoration for permanent facial paralysis. *Arch Plast Surg.* 2013;40(5):633–8. <https://doi.org/10.5999/aps.2013.40.5.633>.
 45. Labbè D, Bussu F, Iodice A. A comprehensive approach to long-standing facial paralysis based on lengthening temporalis myoplasty. *Acta Otorhinolaryngol Ital.* 2012;32(3):145–53.
 46. Banks CA, Bhamra PK, Park J, Hadlock CR, Hadlock TA. Clinician-graded electronic facial paralysis assessment. *Plast Reconstr Surg.* 2015;136(2):223e–30e. <https://doi.org/10.1097/PRS.0000000000001447>.
 47. Banks CA, Jowett N, Hadlock CR, Hadlock TA. Weighting of facial grading variables to disfigurement in facial palsy. *JAMA Facial Plast Surg.* 2016;18(4):292–8. <https://doi.org/10.1001/jamafacial.2016.0226>.
 48. Banks CA, Jowett N, Hadlock TA. Test-retest reliability and agreement between in-person and video assessment of facial mimetic function using the eFACE facial grading system. *JAMA Facial Plast Surg.* 2017;19(3):206–11. <https://doi.org/10.1001/jamafacial.2016.1620>.
 49. Banks CA, Jowett N, Azizzadeh B, et al. Worldwide testing of the eFACE facial nerve clinician-graded scale. *Plast Reconstr Surg.* 2017;139(2):491e–8e. <https://doi.org/10.1097/PRS.0000000000002954>.
 50. Guarin DL, Dusseldorp J, Hadlock TA, Jowett N. A machine learning approach for automated facial measurements in facial palsy. *JAMA Facial Plast Surg.* 2018;20(4):335–7. <https://doi.org/10.1001/jamafacial.2018.0030>.
 51. Dusseldorp JR, Guarin DL, van Veen MM, Jowett N, Hadlock TA. In the eye of the beholder: changes in perceived emotion expression after smile reanimation. *Plast Reconstr Surg.* 2019;144(2):457–71. <https://doi.org/10.1097/PRS.0000000000005865>.
 52. Frey M, Jenny A, Giovanoli P, Stüssi E. Development of a new documentation system for facial movements as a basis for the international registry for neuromuscular reconstruction in the face. *Plast Reconstr Surg.* 1994;93(7):1334–49. <https://doi.org/10.1097/00006534-199406000-00002>.
 53. Frey M, Giovanoli P, Gerber H, Slameczka M, Stüssi E. Three-dimensional video analysis of facial movements: a new method to assess the quantity and quality of the smile. *Plast Reconstr Surg.* 1999;104(7):2032–9. <https://doi.org/10.1097/00006534-199912000-00013>.
 54. Tzou C-HJ, Pona I, Placheta E, et al. Evolution of the 3-dimensional video system for facial motion analysis: ten years' experiences and recent developments. *Ann Plast Surg.* 2012;69(2):173–85. <https://doi.org/10.1097/SAP.0b013e3182223d96>.
 55. Hontanilla B, Aubá C. Automatic three-dimensional quantitative analysis for evaluation of facial movement. *J Plast Reconstr Aesthetic Surg.* 2008;61(1):18–30. <https://doi.org/10.1016/j.bjps.2007.03.037>.
 56. Verhoeven T, Xi T, Schreurs R, Berge S, Maal T. Quantification of facial asymmetry: a comparative study of landmark-based and surface-based registrations. *J Craniomaxillofac Surg.* 2016;44(9):1131–6. <https://doi.org/10.1016/j.jcms.2016.07.017>.
 57. Codari M, Pucciarelli V, Stangoni F, et al. Facial thirds-based evaluation of facial asymmetry using stereophotogrammetric devices: application to facial palsy subjects. *J Craniomaxillofac Surg.* 2017;45(1):76–81. <https://doi.org/10.1016/j.jcms.2016.11.003>.
 58. McDuff D, El Kaliouby R, Senechal T, Amr M, Cohn JF, Picard R. Affective-MIT facial expression dataset (AM-FED): naturalistic and spontaneous facial expressions collected “in-the-wild”. In: *IEEE Computer Society conference on computer vision and pattern recognition workshops*; 2013. p. 881–8. <https://doi.org/10.1109/CVPRW.2013.130>.

59. Su P, Ishii LE, Joseph A, et al. Societal value of surgery for facial reanimation. *JAMA Facial Plast Surg*. 2016;21287(2):139–46. <https://doi.org/10.1001/jamafacial.2016.1419>.
60. Ishii L, Godoy A, Encarnacion CO, Byrne PJ, Boahene KDO, Ishii M. Not just another face in the crowd: society's perceptions of facial paralysis. *Laryngoscope*. 2012;122(3):533–8. <https://doi.org/10.1002/lary.22481>.
61. Dey JK, Ishii LE, Byrne PJ, Boahene KDO, Ishii M. Seeing is believing: objectively evaluating the impact of facial reanimation surgery on social perception. *Laryngoscope*. 2014;124(11):2489–97. <https://doi.org/10.1002/lary.24801>.
62. Lyford-Pike S, Helwig NE, Sohre NE, Guy SJHT, Lyford-Pike S, Helwig NE, Sohre NE, Guy SJ, Hadlock TA. Predicting perceived disfigurement from facial function in patients with unilateral paralysis. *Plast Reconstr Surg*. 2018;142(5):722e–8e. <https://doi.org/10.1097/PRS.0000000000004851>.
63. Hadlock T. Standard outcome measures in facial paralysis getting on the same page. *JAMA Facial Plast Surg*. 2016;18(2):85–6. <https://doi.org/10.1001/jamafacial.2015.2095>.
64. Dey JK, Ishii M, Boahene KDO, Byrne PJ, Ishii LE. Changing perception: facial reanimation surgery improves attractiveness and decreases negative facial perception. *Laryngoscope*. 2014;124(1):84–90. <https://doi.org/10.1002/lary.24262>.
65. Surakka V, Hietanen JK. Facial and emotional reactions to Duchenne and non-Duchenne smiles. *Int J Psychophysiol*. 1998;29(1):23–33. [https://doi.org/10.1016/S0167-8760\(97\)00088-3](https://doi.org/10.1016/S0167-8760(97)00088-3).
66. Hall JA, Carter JD, Horgan TG. Gender differences in nonverbal communication of emotion. In: *Gender and emotion: social psychological perspectives*; 2000. p. 97–117. <https://doi.org/10.1017/CBO9780511628191.006>.
67. Iacolucci CM, Banks C, Jowett N, et al. Development and validation of a spontaneous smile assay. *JAMA Facial Plast Surg*. 2015;17(3):191–6. <https://doi.org/10.1001/jamafacial.2015.0083>.
68. Buendia J, Loayza FR, Luis EO, Celorrio M, Pastor MA, Hontanilla B. Functional and anatomical basis for brain plasticity in facial palsy rehabilitation using the masseteric nerve. *J Plast Reconstr Aesthetic Surg*. 2016;69(3):417–26. <https://doi.org/10.1016/j.bjps.2015.10.033>.
69. Biglioli F, Bayoudh W, Colombo V, Pedrazzoli M, Rabbiosi D. [Double innervation (facial/masseter) on the gracilis flap, in the middle face reanimation in the management of facial paralysis: a new concept]. *Ann Chir Plast Esthet*. 2013;58(2):89–95. <https://doi.org/10.1016/j.anplas.2012.12.001>.
70. Cardenas-Mejia A, Covarrubias-Ramirez JV, Bello-Margolis A, Rozen S. Double innervated free functional muscle transfer for facial reanimation. *J Plast Surg Hand Surg*. 2015;49(3):183–8. <https://doi.org/10.3109/2000656X.2014.988218>.
71. Watanabe Y, Akizuki T, Ozawa T, Yoshimura K, Agawa K, Ota T. Dual innervation method using one-stage reconstruction with free latissimus dorsi muscle transfer for re-animation of established facial paralysis: simultaneous reinnervation of the ipsilateral masseter motor nerve and the contralateral facial nerve to im. *J Plast Reconstr Aesthetic Surg*. 2009;62(12):1589–97. <https://doi.org/10.1016/j.bjps.2008.07.025>.
72. Chuang DCC, Lu JCY, Anesti K. One-stage procedure using spinal accessory nerve (XI)-innervated free muscle for facial paralysis reconstruction. *Plast Reconstr Surg*. 2013;132(1):117e–29e. <https://doi.org/10.1097/PRS.0b013e318290f8cd>.
73. Sforza C, Frigerio A, Mapelli A, et al. Double-powered free gracilis muscle transfer for smile reanimation: a longitudinal optoelectronic study. *J Plast Reconstr Aesthetic Surg*. 2015;68(7):930–9. <https://doi.org/10.1016/j.bjps.2015.03.029>.
74. Dusseldorp JR, Van Veen MM, Guarin DL, Quatela O, Jowett N, Hadlock TA. Spontaneity assessment in dually innervated gracilis smile reanimation surgery. *JAMA Facial Plast Surg*. 2019;21(6):551–7. <https://doi.org/10.1001/jamafacial.2019.1090>.

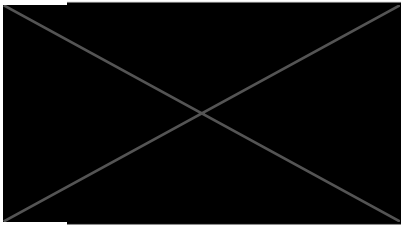
Appendix 1

The Spectrum of Facial Palsy: The MEEI Facial Palsy Photo and Video Standard Set

Authorship Statement

I, Joseph Dusseldorp, confirm for the ensuing journal article: ‘**The Spectrum of Facial Palsy: The MEEI Facial Palsy Photo and Video Standard Set**’ that I have made the following contributions to its authorship:

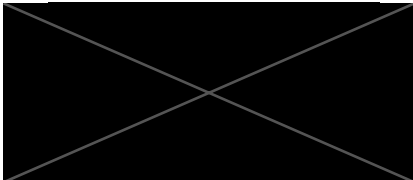
- Conception and design of the research
- Revising of the manuscript and critical appraisal of content



Joseph R. Dusseldorp

Date 14.09. 2023

As Auxiliary Supervisor for the candidature upon which this thesis is based, I can confirm that the above authorship attribution is correct.

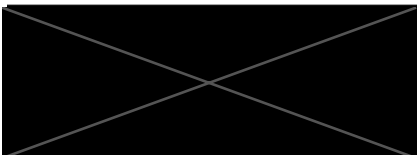


Professor Tessa Hadlock, MD

Date 27.07. 2023

Department of Otolaryngology
Massachusetts Eye and Ear Infirmary
Harvard University, Boston, MA, USA

As Primary Supervisor of the candidate, I concur with the above authorship attestation:



Professor Peter Maitz

Date 14.09.2023

Department of Burns Surgery
Concord Repatriation General Hospital
Sydney University, Sydney, NSW, Australia

The Spectrum of Facial Palsy: The MEEI Facial Palsy Photo and Video Standard Set

Jacqueline J. Greene, MD[†] ; Diego L. Guarin, PhD[†]; Joana Tavares, MD ; Emily Fortier; Mara Robinson; Joseph Dusseldorp, MD; Olivia Quatela; Nate Jowett, MD; Tessa Hadlock, MD

Objectives: Facial palsy causes variable facial disfigurement ranging from subtle asymmetry to crippling deformity. There is no existing standard database to serve as a resource for facial palsy education and research. We present a standardized set of facial photographs and videos representing the entire spectrum of flaccid and nonflaccid (aberrantly regenerated or synkinetic) facial palsy. To demonstrate the utility of the dataset, we describe the relationship between level of facial function and perceived emotion expression as determined by an automated emotion detection, machine learning-based algorithm.

Methods: Photographs and videos of patients with both flaccid and nonflaccid facial palsy were prospectively gathered. The degree of facial palsy was quantified using eFACE, House-Brackmann, and Sunnybrook scales. Perceived emotion during a standard video of facial movements was determined using an automated, machine learning algorithm.

Results: Sixty participants were enrolled and categorized by eFACE score across the range of facial function. Patients with complete flaccid facial palsy (eFACE <60) had a significant loss of perceived joy compared to the nonflaccid and normal groups. Additionally, patients with only moderate flaccid and nonflaccid facial palsy had a significant increase in perceived negative emotion (contempt) when compared to the normal group.

Conclusion: We provide this open-source database to assist in comparing current and future scales of facial function as well as facilitate comprehensive investigation of the entire spectrum of facial palsy. The automated machine learning-based algorithm detected negative emotions at moderate levels of facial palsy and suggested a threshold severity of flaccid facial palsy beyond which joy was not perceived.

Key Words: Facial palsy, facial paralysis, machine learning, joy perception, emotion, facial expression, affectiva, standard set, nonflaccid facial palsy.

Level of Evidence: NA

Laryngoscope, 130:32–37, 2020

INTRODUCTION

Facial nerve function and regeneration outcomes following insult vary widely, depending on myriad clinical factors.^{1–3} Clinical outcomes research in facial palsy is often limited by lack of objective and quantitative evaluation of facial function.⁴ Multiple clinician-reported scales are available, including the eFACE, Sunnybrook Facial Grading system, and modified House-Brackmann grading scale.^{5,6} Facial grading scale usage depends on institutional and individual preferences. A standardized set of photographs and videos of facial function in patients with facial palsy would facilitate comparison and validation of facial grading scales.

Disconnect between internally experienced and communicated emotion remains a major clinical issue for many facial palsy patients. Quantification of nonverbal communication has historically relied on layperson assessment,⁷ which is time-consuming, nonstandardized, and cost-prohibitive in routine clinical practice. New developments in machine learning algorithms to detect facial expressions based on Ekman's Facial Action Coding System (FACS) have led to automated, objective recognition of emotional facial expressions and recently have been applied to patients with facial palsy.⁸

The goal of this study was to present an open-source, prospectively gathered, standardized set of facial photographs and videos representing the entire spectrum of flaccid and nonflaccid (aberrantly regenerated) facial palsy. This open source database will be identified as the *MEEI* [Massachusetts Eye and Ear Infirmary] *Facial Palsy Photo and Video Standard Set* and will be hosted on the Sir Charles Bell Society website. In order to demonstrate the utility of the dataset, we sought to characterize the relationship between facial function and perceived emotional expression using an automated emotion detection, machine learning-based algorithm.

From the Massachusetts Eye & Ear Infirmary, Harvard Medical School (J.J.G., D.L.G., J.T., E.F., M.R., J.D., O.Q., N.J., T.H.), Boston, Massachusetts, U.S.A.

Editor's Note: This Manuscript was accepted for publication on March 18, 2019.

[†]These authors contributed equally to this work.

Presented at the Triological Society Sections Meeting, San Diego, California, U.S.A., January 24–26, 2019.

The authors have no funding, financial relationships, or conflicts of interest to disclose.

Send correspondence to Jacqueline Greene, MD, Massachusetts Eye & Ear Infirmary, Harvard Medical School, 243 Charles St., 9th floor Facial Nerve Center, Boston, MA 02114. E-mail: jacqueline_greene@meei.harvard.edu

DOI: 10.1002/lary.27986

MATERIALS AND METHODS

Institutional review board approval was obtained from the MEEI Human Studies Committee prior to beginning this study. Patients were recruited prospectively from the MEEI Facial Nerve

Center between November 2017 to November 2018. The plan for patients' photographs, videos, and clinical information to be posted as an open source database online was discussed explicitly with each patient and family. Normal intake protocol was followed, including a clinician assessment of facial function (eFACE) as well as a full set of photographs and a video. Exclusion criteria included prior facial reanimation surgery or extensive facial scarring, currently active chemodeneurvation, and bilateral facial palsy. Cases from both flaccid and nonflaccid (aberrantly regenerated or synkinetic) states were included; however, there is no standardized method for distinguishing flaccid from synkinetic facial palsy.

Participants were enrolled sequentially by decile of eFACE score in each state (flaccid or nonflaccid/aberrantly regenerated). Ten patients with normal facial function (eFACE 96–100) were also recruited. Subjects who consented to enroll in the standard set were categorized by their eFACE score into the following: normal (96–100), near-normal (91–95), mild (80–90), moderate (70–79), severe (60–69), and complete flaccid or nonflaccid facial palsy (<60). The degree of palsy was quantified using eFACE, House-Brackmann, and Sunnybrook scales by the first and senior authors (J.J.G., T.H.).

Probability of perceived emotion (0%–100%) was determined using readily available, proprietary software Affdex (Affectiva, Boston, MA; <https://www.affectiva.com/product/emotion-sdk>) that was incorporated using a custom-built graphical user interface created with MatLab version 9.5 (MathWorks, Natick, MA) to analyze the videos taken at each clinic visit, which include the standard facial cues (face at rest; eyebrows raised; gentle eye closure; forceful eye closure; gentle smile; forceful smile; pucker; showing bottom teeth; and repeating “1,2,3,” “happy birthday,” and “hello, it’s a pleasure to meet you”). The probability of joy, disgust, and contempt was compared across groups. Perception of joy, contempt, disgust, sadness, anger, and fear was calculated by taking the area under the curve (AUC) of the perceived emotional probability calculated during the video and normalized by the length of the video (AUC/s).

Data Collection

Demographic data including age, sex, laterality of facial palsy, and clinical presentation and etiology of facial paralysis were recorded. Photography and videography were acquired in a specialized media room with strict adherence to surgical photography guidelines of standard facial movements. Photographs were assembled using GNU image manipulation program (GIMP) version 2.8.22 (Spencer Kimball, Peter Mattis, GIMP Development Team), a free online photography analysis software.

Statistical Analysis

The average AUC of the perceived emotional probability (joy, contempt, fear, disgust, sadness, and anger) normalized by the video time (AUC/s) were compared between the flaccid and nonflaccid facial palsy groups at each category of facial function (separated into deciles by eFACE score). A one-way ANOVA comparing the normal group with the flaccid and nonflaccid facial palsy groups at each category (complete, severe, moderate, mild, and near-normal) was performed using the R Statistical Package version 3.2.3 (R Foundation for Statistical Computing, Bell Laboratory / University of Auckland, New Jersey / New Zealand). Post-hoc comparisons were then performed utilizing the Tukey Honest Significance Difference test. Standard deviations were included for all measured values.

RESULTS

Subjects who consented to enroll in the standard set were categorized by their eFACE score into the following:

normal (96–100), near-normal (91–95), mild (80–90), moderate (70–79), severe (60–69), and complete flaccid or nonflaccid facial palsy (<60). Subjects with any detectable synkinesis were placed in the nonflaccid cohort. The standardized set across the entire spectrum of facial palsy arranged by eFACE score is represented in Figure 1 by a single photograph of full-effort smile. Table 1 provides demographic details for the standard set. The average eFACE, Sunnybrook, and House-Brackmann scores for each category are provided in Figure 2. Each participant in the standard set has eight standard photographs (Fig. 3) as well as a video of the standard facial cues.

Utilizing a machine learning algorithm that assesses perceived emotion (the perceived joy, disgust, and contempt for each member of the standardized set) was calculated and normalized by the length of the video (Fig. 4) and then compared across groups (Fig. 5).

Perceived Joy

The complete flaccid facial palsy group (0.014 ± 0.01 AUC/s) differed significantly from the normal group (11.24 ± 6.08) (Fig. 5A, 5C) ($P = 0.001$) in perception of joy, indicating a possible threshold for expression of joy in this group of facial palsy. The complete flaccid cohort also differed significantly in perceived joy from the complete nonflaccid facial palsy group (13.32 ± 4.01) ($P = 0.001$).

Perceived Contempt

There was a significant difference in perceived contempt between the normal group (0.53 ± 0.88 AUC/s) and the severe and complete nonflaccid (6.97 ± 6.47 , 12.89 ± 13.24 AUC/s, respectively) and severe and complete flaccid facial palsy groups (14.15 ± 7.12 , 14.49 ± 7.71 AUC/s, respectively) (Fig. 5B, 5D) ($P < 0.005$). There was no significant difference between the nonflaccid and flaccid facial palsy groups in perceived contempt, although a visible trend was apparent. Other perceived emotions such as disgust, sadness, anger, or fear were not detected in either the normal or the facial palsy patients, likely due to the fact that the facial movements in the standard video did not include those particular facial action coding units. Absence of these emotions does not indicate inability to express these emotions because these were not elicited during the standard cues video in the facial palsy or normal groups.

DISCUSSION

The goal of this study was to provide a standardized set of patient photographs and videos representing the complete spectrum of facial palsy, from normal facial function to complete flaccid and nonflaccid facial palsy. To our knowledge, this is the first open-source data set for facial palsy. We used the dataset to automatically assess perceived emotion across the spectrum of facial palsy utilizing a newly developed facial emotional recognition algorithm (Affdex) rather than relying on layperson assessment.^{7,9–11} Although we did analyze the dataset utilizing existing facial palsy evaluation systems, including those based on the clinician perspective

Category eFACE

Normal (no facial palsy) 96-100



Flaccid FP Near Normal 91-95



Flaccid FP Mild 80-90



Flaccid FP Moderate 70-79



Flaccid FP Severe 60-69



Flaccid FP Complete <60



Nonflaccid FP Near Normal 91-95



Nonflaccid FP Mild 80-90



Nonflaccid FP Moderate 70-79



Nonflaccid FP Severe 60-69



Nonflaccid FP Complete <60



Fig. 1. Massachusetts Eye and Ear Infirmary Facial Palsy Standard Set. Representative effortful smile photographs of each member of the standard set are shown in the: 1) normal cohort (no facial palsy) with eFACE scores (96–100), 2) flaccid facial palsy, and 3) non-flaccid facial palsy. The complete data set contains a standard set of photographs and video of facial movements and will be hosted on the Sir Charles Bell Society website.

FP = facial palsy. [Color figure can be viewed in the online issue, which is available at www.laryngoscope.com.]

TABLE I.
Demographics of the Standard Set.*

	N	%
Age		
average $\pm$ standard deviation	47 $\pm$ 17	
range	16–92	
Gender		
Male	27	45.0
Female	33	55.0
Side		
Right	26	43.3
Left	25	41.7
Etiology		
Bell palsy	12	20.0
Iatrogenic injury	7	11.7
Ramsay Hunt	5	8.3
Vestibular schwannoma	4	6.7
Benign facial nerve tumor	3	5.0
Congenital facial palsy	3	5.0
Lyme disease	3	5.0
Benign parotid tumor	2	3.3
Parotid malignancy	2	3.3
Neurofibromatosis type II	2	3.3
Trauma	2	3.3
Unknown etiology	2	3.3
Pregnancy-induced Bell palsy	1	1.7
Stroke	1	1.7
Teflon granuloma	1	1.7
CNS lesion	1	1.7

*Of note, 51 patients with facial palsy were included, as well as 9 volunteers with normal facial function, which is why the total number of patients with right and left facial palsy does not add up to 100%.

CNS = central nervous system.

(eFACE, House-Brackmann, Sunnybrook), these grading systems are subjective and fail to capture communicated emotion.⁸ The automated facial emotion detection algorithm in which video segments for smile were isolated and analyzed has been used previously to demonstrate significant increase in perception of joy following smile reanimation surgery (Emotionality Quotient).⁸ The automated facial emotion detection algorithm is derived from facial expressions of 5 million expressions of normal subjects from 75 countries, which were quantified with the Ekman FACS, a well-established method of quantifying facial expressions.^{12–16} This large database was used to train a machine learning algorithm that can automatically detect the probability of expressed emotions in photographs and videos, which can be used as an equivalent to a layperson assessment of facial emotional expression. It must be recognized that all machine learning algorithms are limited by the dataset from which they are built—in this case, people with normal facial function in the Affdex dataset. It is possible that patients with facial palsy are able to convey more joy than the algorithm is able to detect through other means (e.g., head position, body movements, voice tone changes).

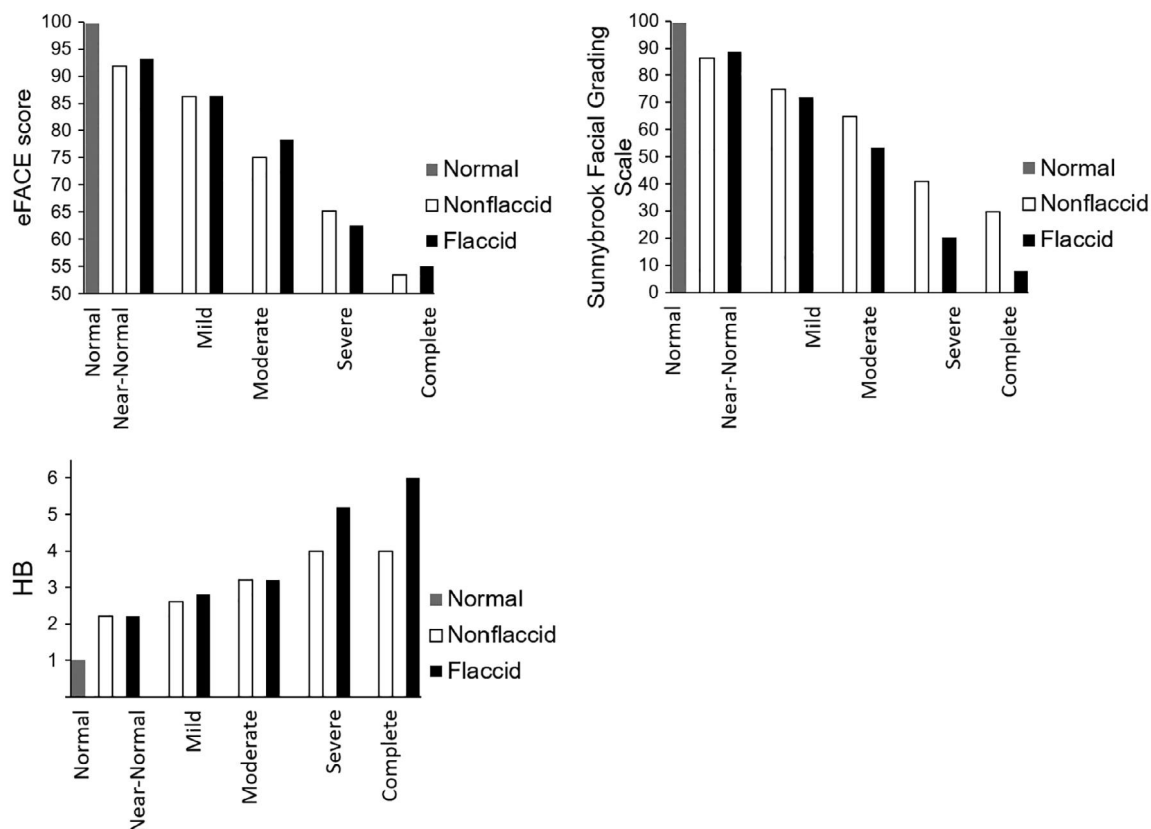


Fig. 2. Facial grading scores. The distribution of Sunnybrook, House-Brackmann, and eFACE scores within each category in the Massachusetts Eye and Ear Infirmary Facial Palsy Standard Set. Of note, the maximum HB score for a patient with synkinesis is 4.²² HB = House-Brackmann.

We found that no perceived joy could be detected in patients whose eFACE score fell below 60, indicating a possible threshold for the ability to express joy in flaccid facial palsy. Interestingly, nonflaccid facial palsy patients were perceived as expressing joy across all eFACE scores, likely related to their partial function and narrowing of the palpebral fissure on the affected side. However, despite the fact that the patients were not instructed to express any negative emotions in the standard facial cues video, negative emotions such as contempt were perceived in flaccid and nonflaccid patients with moderate to complete loss of facial function, potentially elucidating a key factor in the mechanism of disfigurement in facial palsy patients, which can cause psychological morbidity and social isolation.^{17,18}

The creation of this open-source, standardized set of facial photographs and videos representing the entire spectrum of flaccid and aberrantly regenerated facial palsy also allows insights into existing facial grading systems. Clinical outcomes research in facial palsy is limited by institutional and individual preference for different grading schemes and facial analysis, as well as a lack of objective and quantitative evaluation.⁴ Although all clinician-based grading scales are inescapably subjective, categorical scales such as the House-Brackmann and Sunnybrook scales can be faulted for oversimplification, whereas continuous scales such as the eFACE may be overly complex.

The House-Brackmann grading system has been criticized for its lack of zonal facial analysis that both the eFACE and Sunnybrook scales provide.¹⁹ Both the

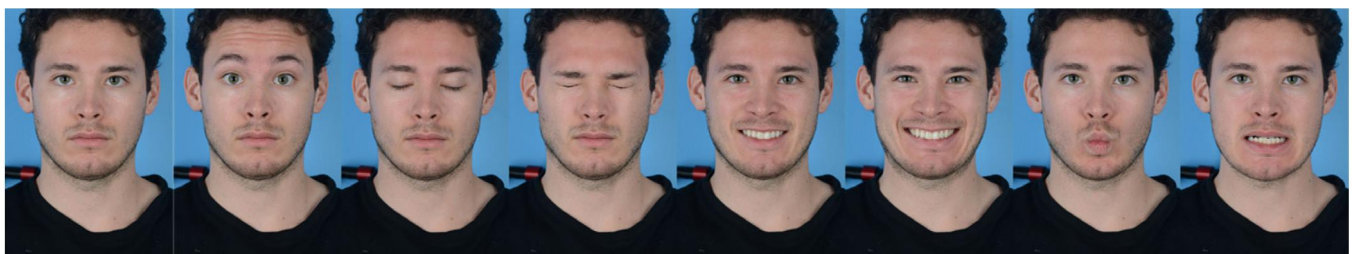


Fig. 3. Complete set of facial movement photographs. Example of standard set of patient photographs as well as video (supplementary data) will be available on the Sir Charles Bell website as an open-source database for facial palsy research and education. From left to right, this standard set of facial movements includes face at rest, eyebrows raised, gentle eye closure, forceful eye closure, gentle smile, forceful smile, pucker, and showing bottom teeth (saying “eee”). [Color figure can be viewed in the online issue, which is available at www.laryngoscope.com.]

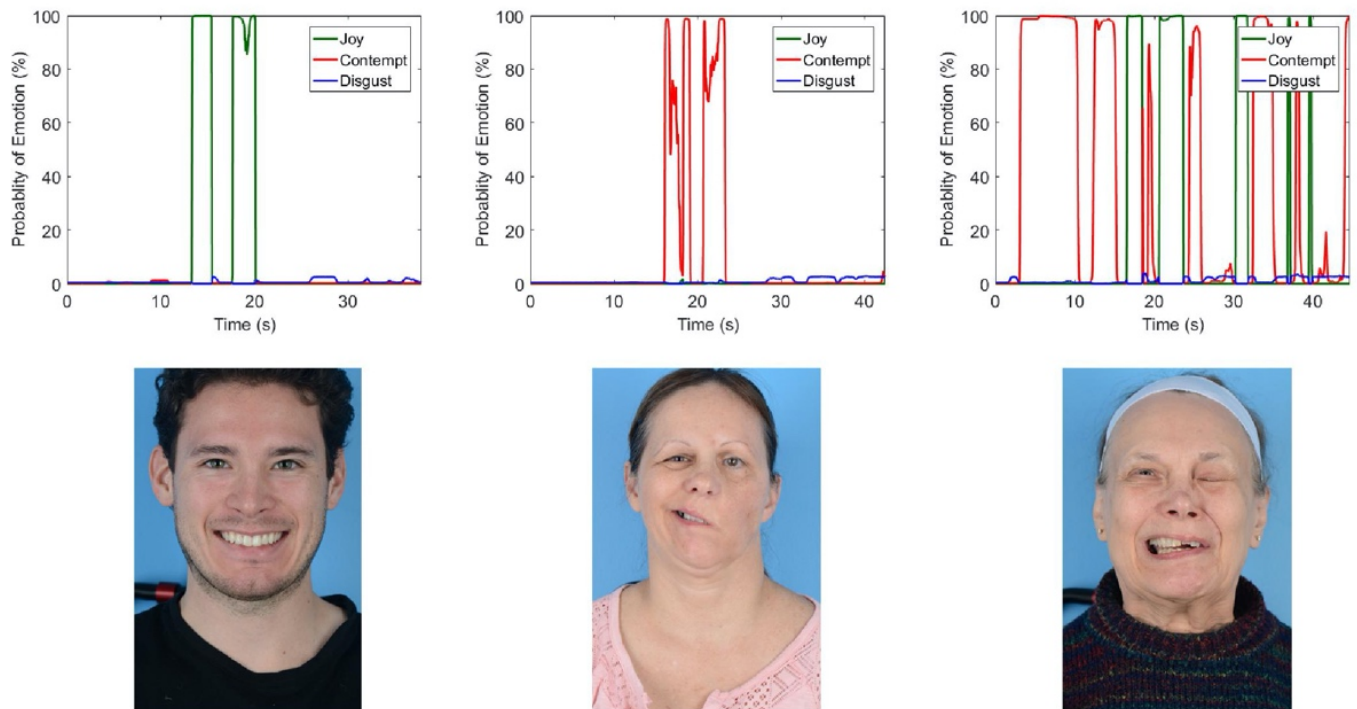


Fig. 4. Quantifying perceived emotion: An automated, machine-learning algorithm calculation of perceived joy, disgust, and contempt was performed during a cued facial expression video of standard set participants with (A) normal facial function, (B) complete flaccid, and (C) complete non-flaccid facial palsy. The video was taken of the same set of facial cues (face at rest; eyebrow raise; gentle eyelid closure; effortful eyelid closure; gentle smile; forceful smile; pucker; showing lower teeth; repeating “happy birthday,” “it’s a pleasure to meet you”). [Color figure can be viewed in the online issue, which is available at www.laryngoscope.com.]

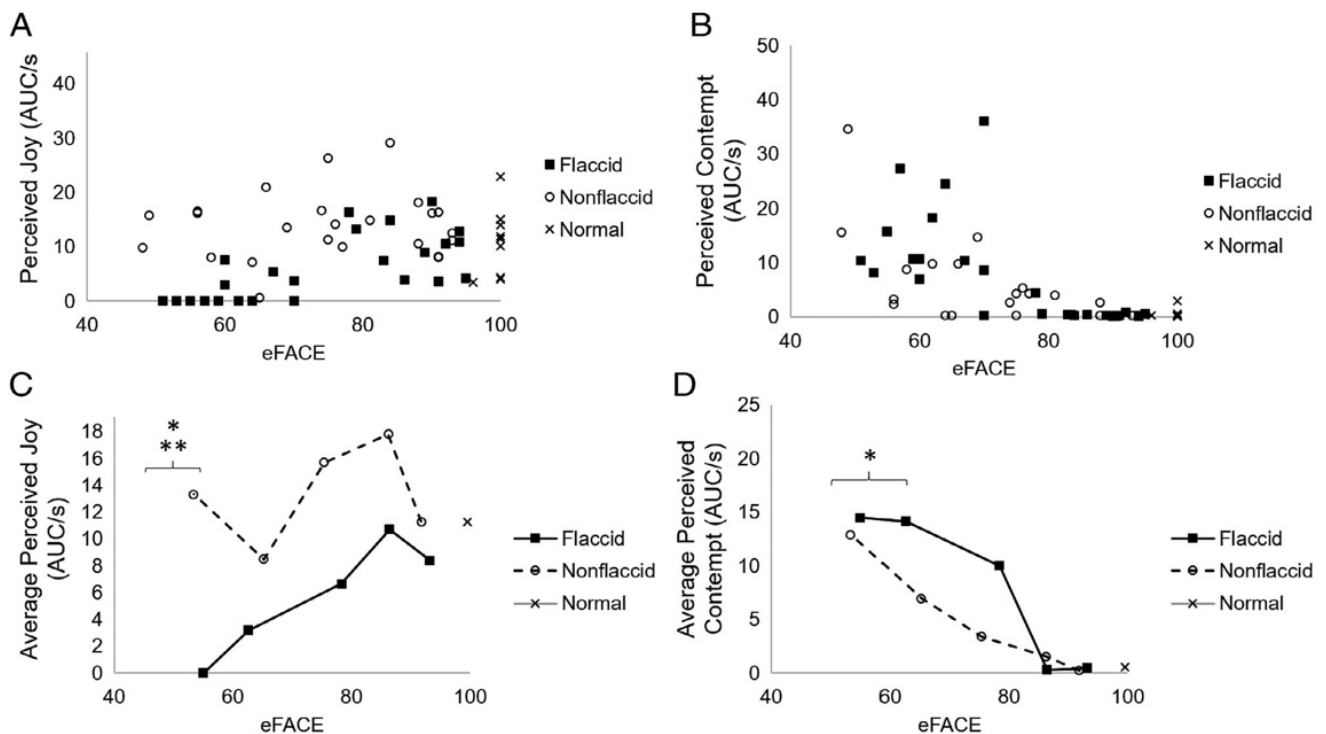


Fig. 5. Perceived emotion across the spectrum of facial palsy. The area under the curve normalized by time of individual participants’ probability of (A) perceived joy and (B) perceived contempt during the standard facial movements video is shown here. Average values in each decile of flaccid and nonflaccid facial palsy are shown for (C) perceived joy and (D) perceived contempt. (*) indicates statistically significant difference between the designated flaccid and nonflaccid facial palsy groups and the normal group (eFACE 96–100), whereas (**) denotes statistically significant difference between the flaccid and nonflaccid groups.

Sunnybrook and HB grading systems refer to “disfiguring synkinesis,” but what differentiates “obvious synkinesis but not disfiguring” from “disfiguring synkinesis” in the Sunnybrook scale or from a HB III (“noticeable but not severe synkinesis”) to a HB IV (“synkinesis ... severe enough to interfere with function”) is unclear. In the HB grading scale, the most severe synkinetic facial palsy patients are rated a score of IV rather than VI,²⁰ which underestimates the severity of the condition. In the most severe nonflaccid facial palsy patients, distorted facial expression is interpreted as a simultaneous expression of both joy and contempt. The Sunnybrook scale prioritizes the “snarl” facial expression rather than assessing lower lip depressor function (saying “eee”), which is associated with smile asymmetry, articulation difficulties, and oral incompetence. Patients with isolated marginal branch weakness or with congenital unilateral lower lip palsy highlight the importance of lower lip depressor function; only the eFACE scale can measure the potential benefit of contralateral lower lip depressor weakening or resection.²¹

This study possessed important limitations. Although all attempts were made to balance gender, age, and diversity in the enrolled subjects, there are limitations based on the available patient population who consented to enroll. The standard set consists of adults who were only given the plan for an open-source database; analysis of this data set may not necessarily apply to pediatric facial palsy. Larger patient cohorts could be obtained in future studies, particularly for the nonflaccid and complete flaccid cases; however, our primary goal was to provide examples of the complete spectrum of facial palsy as a standardized set for facial palsy research and education. Quantifying facial function in terms of perceived emotion allows discernment between subtle asymmetry and crippling deformity and could greatly aid in analyzing the perceived emotion outcome of different surgical interventions for facial palsy.⁸

There is a wide spectrum of facial palsy and regeneration from facial nerve insult. Herein, we present a prospectively gathered, standardized set of facial photographs and videos representing the entire spectrum of flaccid and aberrantly regenerated facial palsy. This tool may facilitate further research in facial palsy across different institutions and educate the medical community and public about the many faces of facial palsy.

CONCLUSION

We present an open-source facial palsy photograph and video standard set that will be posted on the Sir Charles

Bell Society website (www.sircharlesbell.com/standardset). Analysis using a machine learning algorithm for facial emotion recognition suggests a threshold severity of facial palsy below which joy is not perceived during smiling and below which contempt is perceived during smiling.

BIBLIOGRAPHY

- Hadlock TA, Greenfield LJ, Wernick-Robinson M, Cheney ML. Multimodality approach to management of the paralyzed face. *Laryngoscope* 2006;116:1385–1389.
- Hohman MH, Hadlock TA. Etiology, diagnosis, and management of facial palsy: 2000 patients at a facial nerve center. *Laryngoscope* 2014;124:E283–E293.
- Jowett N, Hadlock TA. A contemporary approach to facial reanimation. *Facial Plast Surg Clin North Am* 2015;23:313–334.
- Bhama P, Gliklich RE, Weinberg JS, et al. Optimizing total facial nerve patient management for effective clinical outcomes research. *JAMA Facial Plast Surg* 2014;16:9–14.
- Rozen SM. Discussion: worldwide testing of the eFACE facial nerve clinician-graded scale. *Plast Reconstr Surg* 2017;139:499e–500e.
- Banks CA, Jowett N, Hadlock TA. Test-retest reliability and agreement between in-person and video assessment of facial mimetic function using the eFACE facial grading system. 2017;19:206–211.
- Lyford-Pike S, Helwig NE, Sohre NE, Guy SJ, Hadlock T. Predicting perceived disfigurement from facial function in patients with unilateral paralysis. *Plast Reconstr Surg* 2018;142:722e–728e.
8. Dusseldorp, JR, Guarin, DL, vanVeen, MM, Jowett, N, Hadlock T. In the eye of the beholder: changes in perceived emotion expression after smile reanimation. *Plast Reconstr Surg* 2018.
9. Dey JK, Ishii LE, Byrne PJ, Boahene KDO, Ishii M. Seeing is believing: objectively evaluating the impact of facial reanimation surgery on social perception. *Laryngoscope*. 2014;124:2489–2497.
10. Hadlock T. Standard outcome measures in facial paralysis: getting on the same page. *JAMA Facial Plast Surg* 2016;18:85–86.
11. Helwig NE, Sohre NE, Ruprecht MR, Guy SJ, Lyford-Pike S. Dynamic properties of successful smiles. *PLoS One* 2017;12.
12. McDuff D, El Kaliouby R, Senechal T, Demirdjian D, Picard R. Automatic measurement of ad preferences from facial responses gathered over the internet. *Image Vis Comput* 2014;32:630–640.
13. El Kaliouby R, Robinson P. Real-time inference of complex mental states from facial expressions and head gestures. *Real Time Vis Human Comput Interact*. 2005.
14. McDuff D, El Kaliouby R, Picard RW. Crowdsourcing facial responses to online videos: *IEEE Transactions on Affective Computing*, Vol 3, No. 4, October-December 2012.
15. Baltrusaitis T, McDuff D, Banda N, et al. Real-time inference of mental states from facial expressions and upper body gestures. *IEEE Int Conf Autom Face Gesture Recognit Workshops* 2011:909–914.
16. Ekman P, Oster H. Facial expressions of emotion. *Ann Rev Psychol* 1979;30:527–554.
17. Banks CA, Jowett N, Hadlock CR, Hadlock TA. Weighting of facial grading variables to disfigurement in facial palsy. *JAMA Facial Plast Surg* 2016;18:292–298.
18. Su P, Ishii LE, Joseph A, et al. Societal value of surgery for facial reanimation. *JAMA Facial Plast Surg* 2016;21287:139–146.
19. Fattah AY, Gurusinge ADR, Gavilan J, et al. Facial nerve grading instruments: systematic review of the literature and suggestion for uniformity. *Plast Reconstr Surg* 2015;135:569–579.
20. House JW, Brackmann DE. Facial nerve grading system. *Otolaryngol Head Neck Surg* 1985;93:146–147.
21. Lindsay RW, Edwards C, Smitson C, Cheney ML, Hadlock TA. A systematic algorithm for the management of lower lip asymmetry. *Am J Otolaryngol* 2011;32:1–7.
22. Reitzen SD, Babb JS, Lalwani AK. Significance and reliability of the House-Brackmann grading system for regional facial nerve function. *Otolaryngol Head Neck Surg*. 2009;140:154–158.

CHAPTER THREE

Relative Importance of Voluntary and Spontaneous Smiling in Facial Palsy

Preface

“... To understand, as far as possible, the source or origin of the various expressions which may be hourly seen on the faces of the men around us, not to mention our domesticated animals, ought to possess much interest for us. From these several causes, we may conclude that the philosophy of our subject has well deserved the attention which it has already received from several excellent observers, and that it deserves still further attention, especially from any able physiologist ...”

Charles Darwin *The Expression of Emotions in Man and Animals*. 1872.

Facial reanimation specialists may perceive a successful smile reanimation result differently to patients or layperson observers. For example, if a patient can make a powerful, symmetrical voluntary smile by biting down when the patient visits the clinic the surgeon may think the result is perfectly successful. However, if the patient then leaves the clinic and rarely uses this ability in their daily lives, and still feels self-conscious of their facial difference when laughing in public or eating meals is this truly a successful outcome? Furthermore, if a layperson observer were to see a persistent facial asymmetry while the patient laughs at a joke, what would their impression be of the success of the procedure?

These fundamental questions about the importance of spontaneous versus voluntary smiling in the eyes of patients and laypersons form the basis of investigation in this chapter. The aim of this chapter was to understand what factors would influence a naïve layperson to choose a masseteric nerve-driven smile with lower chance of spontaneity versus a cross facial nerve graft-driven smile with a higher chance of spontaneity. A market research tool known as

“discrete choice experiment” was employed to investigate the key decision-making factors used by 250 laypersons to decide between spontaneous and voluntary smile outcome and to determine the relative importance of each factor. This chapter has been published in *Plastic and Reconstructive Surgery*.

Dusseldorp JR, Naunheim MR, Quatela O, Fortier E, Hadlock TA, Jowett N. Neurotization Preferences in Smile Reanimation: A Discrete Choice Experiment. *Plast. Recon. Surg.* 2021 Sep 1;148(3):407e-415e

In the appendices presented the patient-reported experience of nerve to masseter driven smile reanimation was investigated using a retrospective survey and the effect of facial dysfunction on patient-reported quality of life was also analysed from two distinct perspectives. After masseteric nerve-driven smiling patients reported reliable rates of voluntary smile ability but disappointingly low rates of spontaneous smiling. Poor oral commissure excursion with smiling was the single most important eFACE item predicting lower quality of life and poorer quality of life was predicted by severity of palsy. These additional studies demonstrate the patient reported importance of normal smiling in determining quality of life and the inability of nerve to masseter to achieve this outcome successfully in most cases surveyed in a single institution’s experience.

Appendix 2

van Veen MM, **Dusseldorp JR**, Quatela O, Baiungo J, Robinson M, Jowett N, Hadlock TA. Patient Experience in Nerve-to-Masseter-Driven Smile Reanimation. *J Plast Reconstr Aesthet Surg.* 2019 Aug;72(8):1265-1271

Appendix 3

van Veen MM, Tavares-Brito J, van Veen BM, **Dusseldorp JR**, Werker P, Dijkstra PU, Hadlock TA. Association of Regional Facial Dysfunction with Facial Palsy–Related Quality of Life. JAMA Facial Plast Surg. 2019 Jan 1;21(1):32-37

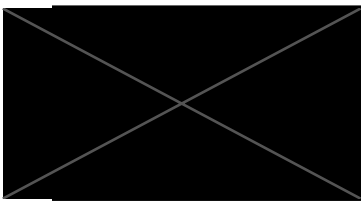
Appendix 4

Tavares-Brito J, van Veen MM, **Dusseldorp JR**, Bahmad F, Hadlock TA. Facial Palsy-Specific Quality of Life in 920 Patients: Correlation with Clinician-Graded Severity and Predicting Factors. Laryngoscope. 2019 Jan;129(1):100-104

Authorship Statement

I, Joseph Dusseldorp, confirm for the ensuing journal article: ‘**Neurotization Preferences in Smile Reanimation: A Discrete Choice Experiment**’ that I have made the following contributions to its authorship:

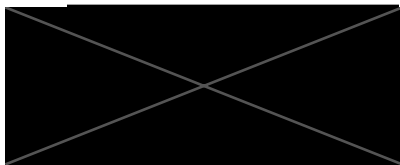
- Conception and design of the research
- Collection and extraction of data
- Drafting of the manuscript



Joseph R. Dusseldorp

Date 14.09. 2023

As Auxiliary Supervisor for the candidature upon which this thesis is based, I can confirm that the above authorship attribution is correct.

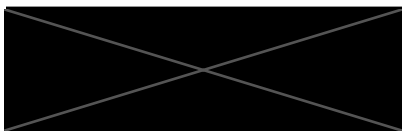


Professor Tessa Hadlock, MD

Date 27.07. 2023

Department of Otolaryngology
Massachusetts Eye and Ear Infirmary
Harvard University, Boston, MA, USA

As Primary Supervisor of the candidate, I concur with the above authorship attestation:



Professor Peter Maitz

Date 14.09.2023

Department of Burns Surgery
Concord Repatriation General Hospital
Sydney University, Sydney, NSW, Australia

Neurotization Preferences in Smile Reanimation: A Discrete Choice Experiment

Joseph R. Dusseldorp,
M.B.B.S., M.S.

Matthew R. Naunheim, M.D.,
M.B.A.

Olivia Quatela, B.S.

Emily Fortier, B.S.

Tessa A. Hadlock, M.D.

Nate Jowett, M.D., Ph.D.

Boston, Mass.; and Sydney, New South
Wales, Australia



Background: Common donor nerve options in smile reanimation include ipsilateral trigeminal motor or contralateral facial nerve branches. Neurotization preference may be influenced by multiple factors, whose relative importance remains poorly understood. In this article, decision-making in smile reanimation is assessed using a stated preference model.

Methods: Qualitative interviews with facial palsy patients identified five relevant attributes for study: smile type (“smile when biting” versus “smile spontaneously” as proxies for trigeminal versus cross-facial neurotization), number of operations, success rates, complication rates, and side effects. Community volunteers ($n = 250$) completed a discrete-choice experiment relevant to free muscle transfer for smile reanimation. Preoperative and postoperative states were demonstrated through video vignettes, together with explanation of surgical risks, consequences, and benefits. Attribute importance was modeled using hierarchical Bayes estimation.

Results: Two hundred forty-one responses met quality controls. Attribute importance ranked as follows: chance of success, 37.3 percent; smile type, 21.4 percent; side effects, 13.9 percent; complication rates, 13.8; and number of operations, 13.6 percent. All attributes significantly correlated with decision making ($p < 0.0001$). An aggregate response model revealed most participants (67.6 percent; standard error, 3.0 percent) preferred smile reanimation by cross-facial (assuming a success rate of 80 percent) as opposed to ipsilateral trigeminal motor branch neurotization. When the success rate for cross-facial neurotization was reduced below 67 percent, trigeminal neurotization was preferred.

Conclusions: Despite a higher risk of failure, most respondents preferred a cross-facial as opposed to trigeminal neurotization strategy for smile reanimation. These findings highlight the complexity of decision-making and need for individualized risk tolerance assessment in the field of facial reanimation. (*Plast. Reconstr. Surg.* 148: 407e, 2021.)

Smile facilitates social interactions by conveying happiness, intelligence, and status.¹⁻⁵ Reanimation of smile is a mainstay of facial palsy treatment. In 1976, Harii et al. described free functional muscle transfer neurotized by ipsilateral trigeminal motor branches for smile reanimation in longstanding facial paralysis.⁶ In 2004, Klebuc and Shenag described direct transfer of an ipsilateral trigeminal motor branch to

distal zygomaticobuccal facial nerve branches for smile reanimation in patients whose native facial musculature remains receptive to reinnervation.^{7,8} Trigeminal neurotization techniques restore voluntary smile with high reliability, yet rarely restore spontaneous expression of joy on the affected side.⁹ In 1971, Smith described neurotization of recently denervated midfacial musculature in hemifacial palsy by routing redundant healthy-side smile axons across the face using a sural nerve interposition graft.¹⁰ In 1980, O'Brien

From the Department of Otolaryngology–Head and Neck Surgery, Massachusetts Eye and Ear and Harvard Medical School; and Department of Plastic and Reconstructive Surgery, Royal Australasian College of Surgeons and University of Sydney.

Received for publication April 9, 2019; accepted December 17, 2020.

The first two authors contributed equally to this work.

Copyright © 2021 by the American Society of Plastic Surgeons

DOI: 10.1097/PRS.00000000000008302

Disclosure: The authors have no financial interest to declare in relation to the content of this article.

Related digital media are available in the full-text version of the article on www.PRSJournal.com.

et al. described neurotization of free transferred muscle for smile reanimation in hemifacial palsy using a cross-facial nerve graft.¹¹ Although cross-facial nerve graft techniques may restore smile spontaneity on the paralyzed side, they carry a higher chance of procedural failure compared to those using ipsilateral trigeminal motor nerve branches.^{9,12–14} Beyond their divergent success rates in achieving meaningful oral commissure excursions and spontaneity, the two techniques differ in that cross-facial neurotization generally requires multiple procedures and sural nerve graft harvest (with resultant sensory deficit in the lower extremity), whereas ipsilateral trigeminal neurotization results in oral commissure movement with prandial activity.

Although decision-making in smile reanimation is complex, surgical preferences have not been well characterized. Do individuals prefer a surgical strategy having potential to restore spontaneous smile if it also carries a higher risk of treatment failure (i.e., failure to restore meaningful commissure excursion)? Do individuals prefer a surgical strategy that avoids sensory deficit in the foot and requires only a single procedure, or would they prefer to avoid smile activation with prandial activity? The answers to such questions may be elucidated through use of survey-based statistical methods known as stated preference techniques.¹⁵ Traditionally used in market research for characterizing consumer valuation of specific attributes of products or services, stated preference techniques may also be used to assess value in health care.^{16,17} Discrete-choice experiments are one type of stated preference technique useful in guiding surgical decision-making. By having participants repeatedly select between surgical strategies when presented with varying and hypothetical versions of their particular attributes, discrete-choice experiments permit quantification of the relative importance of procedure-specific risks, benefits, and consequences.¹⁸ Recently, discrete-choice experiments have been used to understand preferences for breast reconstruction,¹⁹ airway reconstruction,²⁰ bariatric procedures,²¹ Dupuytren's treatment,²² and total knee arthroplasty.²³

Preferences for therapeutic strategies may be assessed among patients, health care providers, or the general population. Community opinion is particularly valuable, as allocation of public funds is guided by societal preferences.²⁴ In this study, we performed a computer-based discrete-choice experiment to elicit neurotization preferences for smile reanimation among a convenience sample of volunteers.

PATIENTS AND METHODS

Institutional review board approval was obtained. Published guidelines in the field of stated preference research were used to guide study design.^{25–27}

Attribute Selection

Exploratory analysis in patients with unilateral facial palsy was undertaken to understand decision-making attributes. Patients had met with the surgical team regarding their diagnosis and prognosis, and none had undergone prior facial reanimation surgery. Open-ended inquiry in the form of semistructured interviews was undertaken among 10 patients. Questions included background knowledge of the disease, treatment expectations, and concerns. Responses were coded using formal thematic content analysis. Several attributes of decision-making surfaced as major themes: likelihood of obtaining a meaningful smile, spontaneity of smile, number of operations, side effects and complications, and expert surgeon recommendation. Less mentioned themes included preference for noninvasive techniques, convenience of treatment, and referring doctor's opinion. The resultant list was reduced to the eight attributes, incorporating patient responses, literature review, and expert opinions. A separate group of 15 additional patients then ranked the five most important attributes from this list for inclusion in the discrete-choice experiment. Surgically relevant levels were defined for each attribute, based on available literature and expert opinions (Table 1).^{9,14,28–32} Expert surgeon recommendation was excluded as an attribute despite ranked importance, as the experiment was designed to analyze patient preference *a priori*. We hypothesized that participants would prefer reanimation of spontaneous smile, greater chance of success, fewer procedures, less risk of complications, and fewer side effects.

Survey Design

Educational Component

A computer-based educational module was designed. Case vignettes of two age-, sex-, and race-matched patients with longstanding hemifacial palsy were used, consisting of photographs and video of preoperative and postoperative voluntary smile, and postoperative spontaneous smile. Spontaneous smile was captured by eliciting laughter using amusing video clips.³³ One patient underwent free gracilis transfer neurotized by the ipsilateral nerve-to-masseter muscle, achieving

Table 1. Attributes and Levels of the Discrete-Choice Experiment

Attribute	Metric	Levels
Smile type	Neurotization preference (ipsilateral trigeminal vs. cross-facial donor nerve)	Smile when biting, smile spontaneously
Chance of success	Likelihood that the procedure is successful	95%, 80%, 70%, 60%, 50%
No. of operations	The number of separate surgical procedures required to restore the smile (each spaced 6–9 mo apart, and requiring a 1- to 4-night stay in the hospital)	1, 2, 3
Chance of complication	Likelihood that an unexpected consequence of surgery occurs, such as wound infection, bleeding, or blood clot	5%, 10%, 20%
Side effects	Expected consequences of surgery	No side effects, numb patch of skin on the outer foot, smile activation when chewing food

favorable voluntary but absent spontaneous smile. [See [Video 1 \(online\)](#), which is an educational video shown to participants of preoperative and postoperative voluntary and spontaneous smile in a patient who underwent smile reanimation by nerve-to-masseter muscle neurotization of a free gracilis flap. Eye blocks were present as shown in all survey images and videos.] The other patient underwent free gracilis transfer neurotized by a cross-facial sural nerve graft, achieving favorable voluntary and spontaneous smile. [See [Video 2 \(online\)](#), which is an educational video shown to participants of preoperative and postoperative voluntary and spontaneous smile in a patient who underwent smile reanimation by cross-facial neurotization of a free gracilis flap. Eye blocks were present as shown in all survey images and videos.] The terms “smile when biting” and “smile

spontaneously” were used as proxies when presenting ipsilateral trigeminal motor and cross-facial neurotization techniques. Typical surgical success rates, number of required procedures, complication rates, and associated side effects for each neurotization strategy were explained, and reiterated in chart format ([Fig. 1](#)) before discrete-choice experiment completion.

Discrete-Choice Experiments

Participants were then asked their neurotization strategy preference through a single non-discrete-choice experiment question. Next, hypothetical treatment scenarios with level variations of each of the five attributes were generated using Lighthouse Studio (Sawtooth Software, Orem, Utah). A partial profile design was used to prevent participant fatigue. Twelve discrete choice sets were

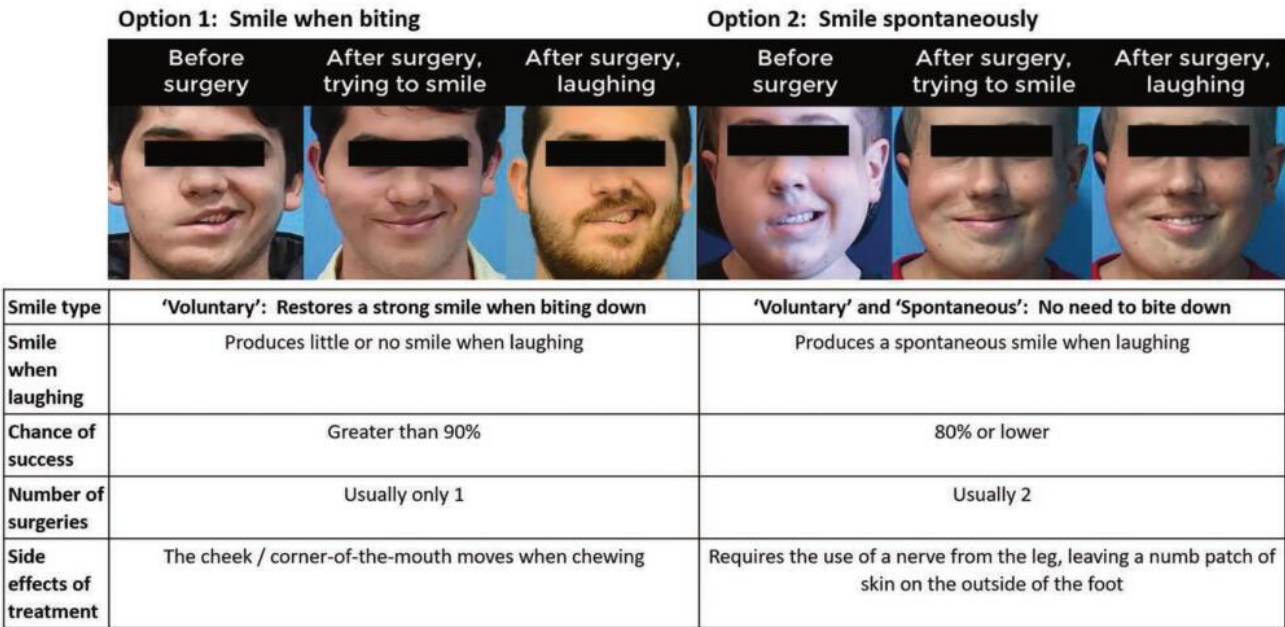


Fig. 1. Summary figure as seen by participants in educational component of survey highlighting de facto attribute levels between ipsilateral trigeminal motor branch (option 1: smile when biting) and cross-facial (option 2: smile spontaneously) neurotization strategies. Eye blocks were present as shown in all survey images and videos.

included, each presenting multiple hypothetical treatments and option to do nothing (Fig. 2). Ten sets presented experimental choice sets with randomized attribute levels for use in utility calculations, whereas two sets were used as holdout tasks (i.e., fixed-choice tasks) to assess reliability. One of the two holdout tasks represented the “base case” scenario, having de facto attribute levels for the two most common neurotization strategies in free muscle transfer for smile reanimation (based on literature review^{9,14,28–32} and expert opinions): use of the ipsilateral nerve-to-masseter muscle (smile when biting, 95 percent chance of success, one procedure, 5 percent chance of complications, smile activation when chewing food) and use of cross-facial nerve graft (smile spontaneously, 80 percent chance of success, two procedures, 5 percent chance of complication, numb patch of skin on the outer foot). The second fixed-choice set was identical to the first, except the chance of success for smile spontaneously was lowered to 60 percent.

Survey Validation

Surveys comprised informed consent, educational and discrete-choice experiment components, and self-reported demographics (i.e., age, sex, race, education level, income, and marriage status). Surveys were optimized for clarity using feedback from three facial nerve clinicians,

10 patients with facial palsy, and 20 individuals without facial palsy by means of open-ended and directed questioning. Excellent comprehension of the optimized survey was demonstrated through monotonicity of responses from the first 50 participants across attributes (i.e., logical ordinal ranking of levels was present for each attribute). (See Appendix, Supplemental Digital Content 1, which shows final optimized survey as used herein. *Blue dotted-line text boxes* were not visible to survey participants, <http://links.lww.com/PRS/E577>.)

Survey Participants

English-speaking, adult volunteers were enrolled by the Harvard University Decision Science Laboratory (Cambridge, Mass.), and completed anonymous, computer-based surveys in a supervised setting. Remuneration was \$15 per participant. There is no consensus for the ideal sample size in discrete-choice experiments; most range between 100 and 300 respondents.²⁶ This study aimed for 200 to 250 responses.

Quality Control and Statistical Analysis

Participants were excluded if the survey was aborted without answering all questions or if they were deemed “speeders,” defined as those who took less than 50 percent of the median

Smile type	Smile when biting	Smile when biting	Smile spontaneously	I would not choose any of these options.
Chance of success	95%	60%	95%	I would prefer to live with facial palsy
Number of surgeries	3	1	3	
Chance of complication	20%	5%	10%	
Side effects	Smile activation when chewing food	Numb patch of skin on the outer foot	No side effects	
	Select	Select	☒	Select

Fig. 2. One of the randomized choice sets seen by participants, with example selection made. Participants must choose one of the four presented options for each choice set.

time to complete the survey.³⁴ Finally, root likelihood (a measure of fit between individual preference scores and choices) was calculated for each participant. Participants having a root likelihood less than 0.25 [calculated as $1/(\text{number of choice tasks})$ —indicating statistically less reliable responses—were also excluded.³⁵

Bayesian hierarchical modeling was used to characterize attribute importance and generate parameter estimates representing part-worth utilities (defined as a unit of desirability associated with each level of each attribute). (See Appendix, Supplemental Digital Content 2, which shows further clarification of the hierarchical Bayes methodology used in this study, <http://links.lww.com/PRS/E578>.) Attribute importances (defined as each attribute's part-worth utility range divided by utility range total) were calculated for each participant and averaged.³⁶ Association of attribute utilities with demographic, medical, and socioeconomic factors was evaluated with one-way analysis of variance. Response consistency on holdout tasks was used to assess validity. Part-worth utilities of holdout questions were included in the survey for reliability testing, but were excluded from final population utility. All statistical analyses were performed with JMP Pro Version 14 (SAS Institute, Inc., Cary, N.C.) and Lighthouse Studio, with alpha set at 0.05.

RESULTS

Of the 250 participants, 241 responses met quality controls. One participant was excluded for incomplete responses to choice questions; four were excluded as speeders; and four were excluded based on root likelihood criteria. Participant demographics are shown in Table 2. Average survey completion time was 12.2 ± 4.5 minutes.

Choice Model

All tested attributes significantly correlated with decision-making ($p < 0.0001$). In order of decreasing importance on decision-making, attributes ranked as follows: chance of success, 37.3 percent; smile type, 21.4 percent; side effects, 13.9 percent; complication rates, 13.8 percent; and number of operations, 13.6 percent. Holding other factors constant, participants preferred cross-facial neurotization (smile spontaneously) over ipsilateral trigeminal motor neurotization (smile when biting), higher chance of success, fewer procedures, lower risk of complications, and fewer side effects. Figure 3

Table 2. Participant Demographics

Characteristic	Value (%)
Mean age $\pm$ SD, yr	33.0 $\pm$ 13.5
Sex	
Female	120 (49.8)
Male	115 (47.7)
Other	6 (2.5)
Race/ethnicity	
Caucasian	139 (57.7)
Asian/Pacific Islander	30 (12.4)
African American	35 (14.5)
Latino or Hispanic	23 (9.5)
Arab/Middle Eastern	4 (1.7)
Bangladeshi, Pakistani, Indian	14 (5.8)
Native American/Aleut/Aboriginal	5 (2.1)
Other	10 (4.1)
Decline to answer	7 (2.9)
Education level	
Some high school	3 (1.2)
High school diploma or GED	17 (7.1)
Some college	51 (21.1)
Associate's degree	10 (4.1)
Bachelor's degree	72 (29.9)
Some graduate school	26 (10.8)
Graduate or professional degree	58 (24.0)
Professional certification	2 (0.8)
Decline to answer	2 (0.8)
Annual household income	
<\$15,000	41 (17.0)
\$15,001–\$25,000	25 (10.4)
\$25,001–\$35,000	11 (4.6)
\$35,001–\$50,000	30 (12.4)
\$50,001–\$75,000	43 (17.8)
\$75,001–\$100,000	23 (9.5)
>\$100,001	36 (14.9)
Decline to answer	32 (13.3)
Marital status	
Single	145 (60.2)
Committed relationship	44 (18.3)
Married	34 (14.1)
Separated	3 (1.2)
Divorced	9 (3.7)
Decline to answer	6 (2.5)

GED, General Educational Development.

demonstrates the utility graph for each attribute. In addition to being logical, preferences of the sample were monotonic. Analysis of variance against sex, age, marital status, income, and education did not show statistically significant differences between groups.

Using model-derived utilities, the percentage of participants choosing ipsilateral trigeminal motor branch neurotization (mean $\pm$ SEM) (32.4 ± 3.0 percent) was far lower than those favoring cross-facial neurotization (67.6 ± 3.0 percent) at de facto attribute levels. These results were similar to the single, non-discrete-choice experiment question asked at the beginning of the survey ("What is your overall preference for surgery type to restore your smile?"), wherein 76.8 percent of participants favored a spontaneous smile strategy. When the model was expanded to include the option of neither surgery, no significant change in participant preference for the two surgical options

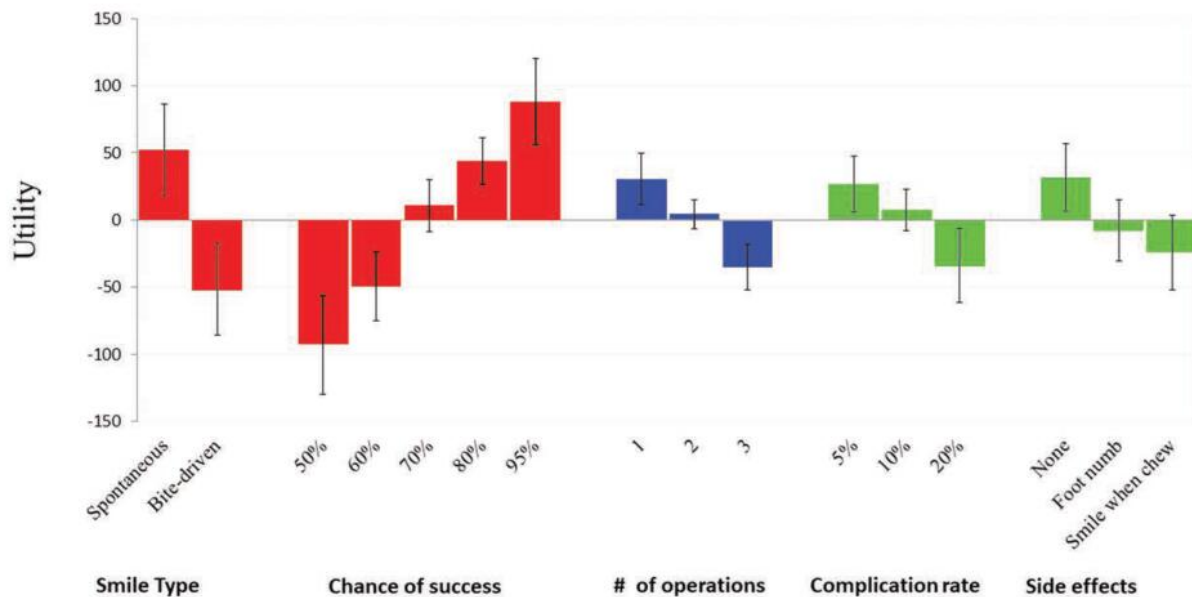


Fig. 3. Utilities of tested attributes. Positive utilities are preferred to negative utilities. Error bars = SD.

was observed (two-stage cross-facial neurotization, 65.6 ± 3.1 percent); single-stage ipsilateral trigeminal motor branch neurotization, 30.3 ± 3.0 percent; and neither, 4.1 ± 1.3 percent). These results were nearly identical to the fixed-choice task held out from analysis for reliability testing, whose respective values were 64.7, 32.8, and 2.5 percent. Notably, the choice model demonstrated a high hit rate for prediction of the two fixed-choice sets, indicating a robust fit. The model predicted the two fixed-choice tasks with 76.3 percent and 75.1 percent accuracy, values that fall within the expected range (75 to 85 percent) of a typical choice model.³⁷

Impact of Failure Rate on Neurotization Preference

Sensitivity analysis was performed on the most important attribute—chance of success of the surgical strategy. When the success rate of the two-stage cross-facial neurotization strategy was reduced to 70 percent, participants continued to favor this strategy (57.7 ± 3.2 percent) over trigeminal motor neurotization (42.3 ± 3.2 percent). However, when the success rate was dropped to 60 percent, cross-facial neurotization was not preferred (35.7 ± 3.1 percent) to trigeminal motor neurotization (64.3 ± 3.1 percent). By further modeling success rate as a linear continuous variable, the threshold failure rate of two-stage cross-facial neurotization above which the majority of participants favored ipsilateral trigeminal motor branch neurotization was estimated to be 33 percent.

Impact of a Theoretical Surgical Strategy on Neurotization Preference

To further demonstrate the power of choice models in therapeutic decision-making in facial reanimation, preference for a hypothetical single-stage cross-facial neurotization strategy was assessed by adjusting model attribute levels (smile spontaneously, single procedure, 5 percent chance of complication, no surgical side effects). Sensitivity analysis was then performed on the chance of procedural success of this unique strategy, whose feasibility has been previously described.^{38–40} When the success rate of a single-stage cross-facial neurotization strategy was set at 80 percent, the model-estimated participants would strongly favor this strategy (85.9 ± 2.2 percent) over single-stage trigeminal motor neurotization (14.1 ± 2.2 percent). When the success rate was dropped to 50 percent, the model-estimated single-stage cross-facial neurotization would no longer be preferred (42.3 ± 3.2 percent) over trigeminal motor neurotization (57.7 ± 3.2 percent). The predicted threshold failure rate of single-stage cross-facial neurotization above which the majority of participants would favor ipsilateral trigeminal motor branch neurotization was 44 percent.

DISCUSSION

Free muscle transfer remains the gold standard for smile reanimation in longstanding facial palsy. Herein, ipsilateral trigeminal motor^{6,41} and cross-facial^{10,11,38–40} neurotization strategies are most commonly used. Alternate neurotization

options include the ipsilateral accessory^{9,41} or hypoglossal nerves,⁴¹ or ipsilateral distal facial nerve branches in select cases.⁴² In this study, we sought to determine societal preference for the two most commonly used neurotization techniques. In developing our survey, we determined that the two most important attributes relevant to patient decision-making in smile reanimation are the chance of success and smile type.

Despite their strong association with decision-making, side effects, complication rates, and number of operations were deemed less important attributes of smile reanimation. Although oral commissure movements with prandial activity occurs in patients with trigeminal motor neurotized smile, few patients report this as problematic.⁴³ Patient surveys indicate sural nerve harvest is well tolerated.⁴⁴ Serious complications following free-muscle transfer for smile reanimation are rare.^{9,12,14} Although two procedures are typically required, single-stage cross-facial neurotization of free transferred muscle for smile reanimation has been reported using gracilis³⁸ and latissimus dorsi muscles.^{39,40} Compared with two-stage cross-facial neurotization, our model predicted a majority of participants would accept an 11 percent higher risk of procedural failure in selecting a single-stage cross-facial neurotization strategy over ipsilateral trigeminal neurotization.

Advantages of smile reanimation in hemifacial palsy are clear.^{30,45,46} In this study, we investigated whether the perceived benefit of rehabilitation of spontaneous—as opposed to exclusively volitional—smile outweighs the higher risk of failure associated with cross-facial neurotization procedures. We found that a majority of participants would be willing to accept procedural failure rates ranging between 33 and 44 percent for the chance to restore spontaneous smile. Although one study noted a high rate of patient-reported smile spontaneity following nerve-to-masseter muscle neurotization of free transferred muscle in children,⁴⁷ other groups have reported complete absence of smile spontaneity with use of trigeminal motor donor nerves.^{9,14} Emerging evidence suggests that a dual-innervation strategy (comprising end-to-end coaptation of the nerve-to-masseter muscle and end-to-side coaptation of a cross-facial nerve graft to the nerve of the transferred muscle) may be effective in restoring spontaneous smile with a lower risk of treatment failure as compared to cross-facial neurotization alone.^{48,49} In our experience, approximately 20 percent of patients with free muscle neurotized by the ipsilateral nerve-to-masseter muscle and 33 percent of those with

dually innervated flaps will demonstrate some affected-side spontaneous smile movements.⁵⁰ The tendency of some individuals to naturally contract their masseter muscles during normal smile production may explain the smile spontaneity seen among a subset of patients following nerve-to-masseter muscle transfer procedures.⁵¹ In this study, we sought to educate participants as to the expected outcomes and consequences of ipsilateral trigeminal motor and cross-facial neurotization strategies as we currently counsel patients preoperatively; namely, as providing little to no smile when laughing as opposed to a smile when laughing, respectively.

This study highlights the complexity of neurotization choice in smile reanimation, and the advantages of stated preference models in guiding surgical decision-making and innovation. Despite ranking the chance of success as the most important attribute, the majority of participants preferred a cross-facial neurotization strategy, notwithstanding its higher chance of failure and potential need for multiple procedures. Our findings suggest that two-stage cross-facial neurotization could be considered first-line therapy in smile reanimation, where the chance of success is deemed at least 67 percent, with ipsilateral trigeminal motor neurotization reserved for salvage.^{52,53} To avoid bias, participants in this study were blinded to the possibility of salvage procedures.

Just as stated preference techniques can be used to predict consumer valuation and inform the design of hypothetical products or services,¹⁵ in this study, we modeled participant preferences for a hypothetical surgical strategy. Our model suggests single-stage cross-facial neurotization of free muscle could be considered first-line therapy for smile reanimation, where the chance of success is at least 56 percent. Of course, surgical decision-making is highly individualized and dependent on patient risk tolerance. Recent work has demonstrated the usefulness of discrete-choice experiments in guiding surgical decision-making and managing expectations among individual patients for total knee arthroscopy.²³ In a similar fashion, the discrete-choice experiment developed in this study could be used to determine individualized thresholds at which patients would prefer an ipsilateral trigeminal over a cross-facial neurotization strategy.

This study has several limitations. Participant demographics were skewed toward younger, well-educated patients, limiting the generalizability of our findings. To minimize survey length, examples of treatment failures were not shown to

participants. Differences in postoperative rehabilitation times were not explicitly discussed. Although postoperative volitional smile outcomes shown to participants were similar for both neurotization techniques (Fig. 1), the cross-facial neurotization example demonstrated favorable dentition-show spontaneous smile, which may have biased responses. This study characterized smile neurotization preferences of participants from the general population. Future work will assess the external validity of the discrete-choice experiment among patients afflicted with facial palsy, whose familiarity and experiences with the disease might impact treatment preferences. This survey was based on neurotization strategies for free muscle transfer in smile reanimation. Although the results in this study might inform decision-making in neurotization strategy for reinnervation of native facial musculature, a tailored discrete-choice experiment would provide more reliable data. Results of stated preference techniques are dependent on educational component quality and are limited by the qualitative nature of attribute and level determination.^{15,25} In line with best practices, we sought to minimize potential bias through extensive patient interviews, literature reviews, expert opinions, and pilot testing of the educational and discrete-choice experiment components of our survey. Regardless, there is a need for validated preference measures in facial reanimation, similar to those developed in other fields.⁵⁴

CONCLUSIONS

Decision-making and outcomes in facial reanimation vary considerably with surgeon experience and reconstructive philosophy. Although the chance of surgical success was rated the most important decision-making factor in smile reanimation, a majority of respondents preferred a cross-facial neurotization strategy, despite its higher failure rate and potential need for multiple procedures. These findings highlight the complexity of decision-making and the need for individualized risk tolerance assessment in the field of facial reanimation.

Nate Jowett, M.D.

Massachusetts Eye and Ear
243 Charles Street
Boston, Mass. 02114

nate_jowett@meei.harvard.edu

Twitter: @NateJowett, @TheFacialNerve

PATIENT CONSENT

Patients provided written consent for the use of their images.

ACKNOWLEDGMENT

The authors thank the Harvard Decision Science Laboratory at Harvard's Kennedy School of Government for the use of their facilities.

REFERENCES

1. Darwin C. *The Expression of the Emotions in Man and Animals*. London: J. Murray; 1872.
2. Duchenne GB. *Mécanisme de la physionomie humaine; ou, Analyse électro-physiologique de l'expression des passions*. Paris: Renouard; 1862.
3. Schmidt KL, Cohn JF. Human facial expressions as adaptations: Evolutionary questions in facial expression research. *Am J Phys Anthropol*. 2001;33(Suppl):3–24.
4. Otta E, Folladore Abrosio F, Hoshino RL. Reading a smiling face: Messages conveyed by various forms of smiling. *Percept Mot Skills* 1996;82:1111–1121.
5. Hecht MA, LaFrance M. License or obligation to smile: The effect of power and sex on amount and type of smiling. *Pers Soc Psychol Bull*. 1998;24:1332–1342.
6. Harii K, Ohmori K, Torii S. Free gracilis muscle transplantation, with microvascular anastomoses for the treatment of facial paralysis: A preliminary report. *Plast Reconstr Surg*. 1976;57:133–143.
7. Klebuc M. Masseter-to-facial nerve transfer: A new technique for facial reanimation. *J Reconstr Microsurg*. 2006;22:A101.
8. Klebuc M, Shenaq SM. Donor nerve selection in facial reanimation surgery. *Semin Plast Surg*. 2004;18:53–60.
9. Chuang DC, Lu JC, Chang TN, Laurence VG. Comparison of functional results after cross-face nerve graft, spinal accessory nerve, and masseter nerve-innervated gracilis for facial paralysis reconstruction: The Chang Gung experience. *Ann Plast Surg*. 2018;81(Suppl 1):S21–S29.
10. Smith JW. A new technique of facial animation. In: *Transactions of the Vth International Congress in Plastic Surgery*; February 22–26, 1971; Melbourne, Australia. Sydney: Butterworth; 1971.
11. O'Brien BM, Franklin JD, Morrison WA. Cross-facial nerve grafts and microvascular free muscle transfer for long established facial palsy. *Br J Plast Surg*. 1980;33:202–215.
12. Bhama PK, Weinberg JS, Lindsay RW, Hohman MH, Cheney ML, Hadlock TA. Objective outcomes analysis following microvascular gracilis transfer for facial reanimation: A review of 10 years' experience. *JAMA Facial Plast Surg*. 2014;16:85–92.
13. Snyder-Warwick AK, Fattah AY, Zive L, Halliday W, Borschel GH, Zuker RM. The degree of facial movement following microvascular muscle transfer in pediatric facial reanimation depends on donor motor nerve axonal density. *Plast Reconstr Surg*. 2015;135:370e–381e.
14. Faria JC, Scopel GP, Busnardo FF, Ferreira MC. Nerve sources for facial reanimation with muscle transplant in patients with unilateral facial palsy: Clinical analysis of 3 techniques. *Ann Plast Surg*. 2007;59:87–91.
15. Bateman JJ, Carson RT, Day B, et al. *Economic Valuation with Stated Preference Techniques: A Manual*. Cheltenham, United Kingdom: Edward Elgar; 2002.
16. Bridges JF. Stated preference methods in health care evaluation: An emerging methodological paradigm in health economics. *Appl Health Econ Health Policy* 2003;2:213–224.
17. Ryan M, Bate A, Eastmond CJ, Ludbrook A. Use of discrete choice experiments to elicit preferences. *Qual Health Care* 2001;10(Suppl 1):i55–i60.

18. Naunheim MR, Wittenberg E, Shrimel MG. Patient preference research in otolaryngology: What do patients want? *JAMA Otolaryngol Head Neck Surg*. 2017;143:971–972.
19. Damen TH, de Bekker-Grob EW, Mureau MA, et al. Patients' preferences for breast reconstruction: A discrete choice experiment. *J Plast Reconstr Aesthet Surg*. 2011;64:75–83.
20. Naunheim MR, Naunheim ML, Rath VK, Franco RA, Shrimel MG, Song PC. Patient preferences in subglottic stenosis treatment: A discrete choice experiment. *Otolaryngol Head Neck Surg*. 2018;158:520–526.
21. Rozier MD, Ghaferi AA, Rose A, Simon NJ, Birkmeyer N, Prosser LA. Patient preferences for bariatric surgery: Findings from a survey using discrete choice experiment methodology. *JAMA Surg*. 2019;154:e184375.
22. Kan HJ, de Bekker-Grob EW, van Marion ES, et al. Patients' preferences for treatment for Dupuytren's disease: A discrete choice experiment. *Plast Reconstr Surg*. 2016;137:165–173.
23. Dowsey MM, Scott A, Nelson EA, et al. Using discrete choice experiments as a decision aid in total knee arthroplasty: Study protocol for a randomised controlled trial. *Trials* 2016;17:416.
24. Gold MR. *Cost-Effectiveness in Health and Medicine*. New York: Oxford University Press; 1996.
25. Lancsar E, Louviere J. Conducting discrete choice experiments to inform healthcare decision making: A user's guide. *Pharmacoeconomics* 2008;26:661–677.
26. Bridges JF, Hauber AB, Marshall D, et al. Conjoint analysis applications in health: A checklist. A report of the ISPOR Good Research Practices for Conjoint Analysis Task Force. *Value Health* 2011;14:403–413.
27. Reed Johnson F, Lancsar E, Marshall D, et al. Constructing experimental designs for discrete-choice experiments: Report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force. *Value Health* 2013;16:3–13.
28. Greene JJ, Tavares J, Mohan S, Jowett N, Hadlock T. Long-term outcomes of free gracilis muscle transfer for smile reanimation in children. *J Pediatr*. 2018;202:279–284.e2.
29. Hadlock TA, Malo JS, Cheney ML, Henstrom DK. Free gracilis transfer for smile in children: The Massachusetts Eye and Ear Infirmary experience in excursion and quality-of-life changes. *Arch Facial Plast Surg*. 2011;13:190–194.
30. Lindsay RW, Bhama P, Hadlock TA. Quality-of-life improvement after free gracilis muscle transfer for smile restoration in patients with facial paralysis. *JAMA Facial Plast Surg*. 2014;16:419–424.
31. Lindsay RW, Bhama P, Weinberg J, Hadlock TA. The success of free gracilis muscle transfer to restore smile in patients with nonflaccid facial paralysis. *Ann Plast Surg*. 2014;73:177–182.
32. Kumar PA, Hassan KM. Cross-face nerve graft with free-muscle transfer for reanimation of the paralyzed face: A comparative study of the single-stage and two-stage procedures. *Plast Reconstr Surg*. 2002;109:451–462; discussion 463–464.
33. Iacolucci CM, Banks C, Jowett N, et al. Development and validation of a spontaneous smile assay. *JAMA Facial Plast Surg*. 2015;17:191–196.
34. Greszki R, Meyer M, Schoen H. The impact of speeding on data quality in nonprobability and freshly recruited probability-based online panels. In Callegaro M, Baker R, Bethlehem J, Göritz AS, Krosnick JA, Lavrakas PJ, eds. *Online Panel Research*. New York: John Wiley & Sons, Ltd.; 2014.
35. Sawtooth Software. *The CBC System for Choice-Based Conjoint Analysis Version 8*. Orem, Utah: Sawtooth Software; 2013.
36. Louviere JJ, Hensher DA, Swait JD. *Stated Choice Methods: Analysis and Application*. New York: Cambridge University Press; 2000.
37. Orme B. Chapter 7: Sample size issues for conjoint analysis. In: *Getting Started with Conjoint Analysis: Strategies for Product Design and Pricing Research Second Edition*. Madison, Wis.: Research Publishers; 2010:57–65.
38. Kumar PA. Cross-face reanimation of the paralyzed face, with a single stage microvascular gracilis transfer without nerve graft: A preliminary report. *Br J Plast Surg*. 1995;48:83–88.
39. Harii K, Asato H, Yoshimura K, Sugawara Y, Nakatsuka T, Ueda K. One-stage transfer of the latissimus dorsi muscle for reanimation of a paralyzed face: A new alternative. *Plast Reconstr Surg*. 1998;102:941–951.
40. Jones BM. Cross-face reanimation of the paralyzed face, with single stage microvascular gracilis transfer without nerve graft. *Br J Plast Surg*. 1995;48:519–520.
41. Zuker RM, Manktelow RT. A smile for the Möbius' syndrome patient. *Ann Plast Surg*. 1989;22:188–194.
42. Gur E, Zuker RM, Zaretski A, et al. Incomplete facial paralysis: The use of the ipsilateral residual facial nerve as a donor nerve for facial reanimation. *Plast Reconstr Surg*. 2018;142:202–214.
43. Rozen S, Harrison B. Involuntary movement during mastication in patients with long-term facial paralysis reanimated with a partial gracilis free neuromuscular flap innervated by the masseteric nerve. *Plast Reconstr Surg*. 2013;132:110e–116e.
44. Hallgren A, Björkman A, Chemnitz A, Dahlin LB. Subjective outcome related to donor site morbidity after sural nerve graft harvesting: A survey in 41 patients. *BMC Surg*. 2013;13:39.
45. Su P, Ishii LE, Joseph A, et al. Societal value of surgery for facial reanimation. *JAMA Facial Plast Surg*. 2017;19:139–146.
46. Faris C, Tessler O, Heiser A, Hadlock T, Jowett N. Evaluation of societal health utility of facial palsy and facial reanimation. *JAMA Facial Plast Surg*. 2018;20:480–487.
47. Manktelow RT, Tomat LR, Zuker RM, Chang M. Smile reconstruction in adults with free muscle transfer innervated by the masseter motor nerve: Effectiveness and cerebral adaptation. *Plast Reconstr Surg*. 2006;118:885–899.
48. Biglioli F, Colombo V, Tarabbia F, et al. Double innervation in free-flap surgery for long-standing facial paralysis. *J Plast Reconstr Aesthet Surg*. 2012;65:1343–1349.
49. Okazaki M, Mutsumi O, Kentaro T, et al. One-stage dual latissimus dorsi muscle flap transfer with a pair of vascular anastomoses and double nerve suturing for long-standing facial paralysis. *J Plast Reconstr Aesthet Surg*. 2015;68:e113–e119.
50. Dusseldorp JR, van Veen MM, Guarin DL, Quatela O, Jowett N, Hadlock TA. Spontaneity assessment in dually innervated gracilis smile reanimation surgery. *JAMA Facial Plast Surg*. 2019;21:551–557.
51. Schaverien M, Moran G, Stewart K, Addison P. Activation of the masseter muscle during normal smile production and the implications for dynamic reanimation surgery for facial paralysis. *J Plast Reconstr Aesthet Surg*. 2011;64:1585–1588.
52. Eisenhardt SU, Eisenhardt NA, Thiele JR, Stark GB, Bannasch H. Salvage procedures after failed facial reanimation surgery using the masseteric nerve as the motor nerve for free functional gracilis muscle transfer. *JAMA Facial Plast Surg*. 2014;16:359–363.
53. Horta R, Silva P, Silva A, et al. Facial reanimation with gracilis muscle transplantation and obturator nerve coaptation to the motor nerve of masseter muscle as a salvage procedure in an unreliable cross-face nerve graft. *Microsurgery* 2011;31:164–166.
54. Spatz ES, Krumholz HM, Moulton BW. Prime time for shared decision making. *JAMA* 2017;317:1309–1310.

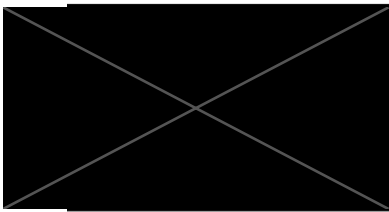
Appendix 2

Patient Experience in Nerve-to-Masseter-Driven Smile Reanimation

Authorship Statement

I, Joseph Dusseldorp, confirm for the ensuing journal article: ‘**Patient Experience in Nerve-to-Masseter-Driven Smile Reanimation**’ that I have made the following contributions to its authorship:

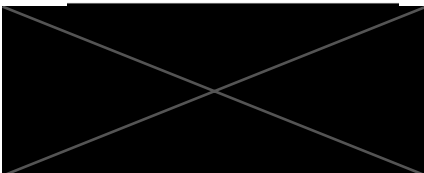
- Conception and design of the research
- Interpretation of the findings
- Revising of the manuscript and critical appraisal of content
- Responding to peer reviewers



Joseph R. Dusseldorp

Date 14.09. 2023

As Auxiliary Supervisor for the candidature upon which this thesis is based, I can confirm that the above authorship attribution is correct.

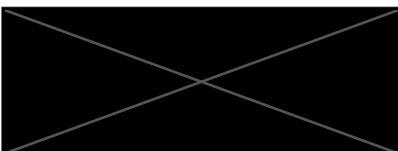


Professor Tessa Hadlock, MD

Date 27.07. 2023

Department of Otolaryngology
Massachusetts Eye and Ear Infirmary
Harvard University, Boston, MA, USA

As Primary Supervisor of the candidate, I concur with the above authorship attestation:



Professor Peter Maitz

Date 14.09.2023

Department of Burns Surgery
Concord Repatriation General Hospital
Sydney University, Sydney, NSW, Australia



Patient experience in nerve-to-masseter-driven smile reanimation

Martinus M. van Veen^{a,b,*}, Joseph R. Dusseldorp^{a,c},
Olivia Quatela^a, Jennifer Baiungo^a, Mara Robinson^a,
Nate Jowett^a, Tessa A. Hadlock^a

^a Department of Otolaryngology/Head and Neck surgery, Facial Nerve Center, Massachusetts Eye and Ear Infirmary and Harvard Medical School, 243 Charles St., Boston, MA 02114, USA

^b Department of Plastic surgery, University Medical Center Groningen and University of Groningen, P.O. Box 30.001, NL-9700 RB Groningen, the Netherlands

^c Department of Plastic and Reconstructive Surgery, Royal Australasian College of Surgeons and University of Sydney, City Road, Camperdown/Darlington NSW 2006, Australia

Received 10 September 2018; accepted 24 March 2019

KEYWORDS

Facial paralysis;
Facial paresis;
Patient-reported
outcome;
Nerve to masseter

Abstract *Introduction:* The nerve-to-masseter is one of the most frequently used neural sources in smile reanimation surgery. Very little information has been reported on patient experience with regard to reanimated smile usage and sequelae following transfer. The aim of this study was to quantify patient perception of nerve-to-masseter use in smile reanimation surgery.

Methods: An online questionnaire was developed based on the clinical expertise of our team, patient interviews, and existing questionnaires of facial palsy-related quality of life and temporomandibular joint dysfunction. All patients treated with nerve-to-masseter-driven smile reanimation surgery, both nerve transfers and muscle transplantations, between 2007 and 2016 with a valid email address were invited to participate.

Results: Of 171 operated patients, 122 with a valid email address were invited to participate. Seventy-one patients responded (63.4% female, mean age 51.1 years) after a median follow-up of 3.8 years. A voluntary smile while biting down at least “most of the time” was reported by 83.1% of patients; 46.5% reported ability to smile on the affected side without bite. A “normal” or “almost normal” spontaneous smile was reported in 23.9% of patients. A total of 18.3% of patients self-reported masseter muscle atrophy, and 1.4–14.1% reported temporomandibular

* Corresponding author at: Department of Plastic Surgery, University Medical Centre Groningen and University of Groningen, P.O. Box 30.001, NL-9700 RB Groningen, the Netherlands.

E-mail address: m.m.van.veen@umcg.nl (M.M. van Veen).

<https://doi.org/10.1016/j.bjps.2019.03.037>

1748-6815/© 2019 British Association of Plastic, Reconstructive and Aesthetic Surgeons. Published by Elsevier Ltd. All rights reserved.

joint dysfunction. Forty-one patients (57.7%) reported prandial movement of the face at least "most of the time," with 9 patients (12.7%) considering this bothersome.

Conclusion: Patients report good voluntary smiling ability following nerve-to-masseter-driven smile reanimation surgery, with low rates of sequelae.

© 2019 British Association of Plastic, Reconstructive and Aesthetic Surgeons. Published by Elsevier Ltd. All rights reserved.

Introduction

Ongoing debate exists regarding the optimal neural source to drive smile reanimation in both the early setting (less than 18-24 months),¹ when facial muscles are receptive to reinnervation, and long-standing palsy, when they are not. Recently, the most commonly used neural source to augment facial function has become the nerve-to-masseter (NTM) because of its high axonal density and favorable anatomical location.²⁻⁵ Other neural sources include the ipsilateral facial nerve stump through cable graft, cross-facial nerve graft (CFNG), deep temporal nerve, hypoglossal nerve, and accessory nerve.

In the acute setting, a masseteric-to-facial (V-to-VII) nerve transfer can be performed, whereby a branch of the NTM is transposed to an ideal distal facial nerve branch that supplies the zygomaticus and orbicularis oculi muscles or to the main trunk of the facial nerve to reanimate the smile and enable some voluntary eye closure.⁵⁻⁸ In longstanding facial palsy, free or regional muscle transfers are needed to reanimate a smile. Cross-facial nerve grafting is considered as the more ideal neural source to reinnervate free muscle transfers, as it provides spontaneity.⁹ However, in recent years, the NTM has been shown to provide more reliable voluntary contraction of the transplanted muscle.^{4,10,11}

While successful reanimation of voluntary smiling is highly reliable after this procedure, opinions vary on the ability of the NTM to produce a spontaneous or emotional smile.¹²⁻¹⁸ Relatively few studies have assessed the patient experience of NTM-driven smile reanimation,¹⁹ although there have been many studies that report low morbidity following its use for smile reanimation.^{7,20-22} Recognized limitations of NTM are that it does not provide resting tone to the face, and complete division of the nerve can lead to atrophy of the masseter muscle and a contour irregularity over the mandibular angle.²²⁻²⁴ Prandial activation is inherent to the use of the NTM in smile reanimation, although a recent study found that this was rarely bothersome to patients.¹⁹ Varying physical therapy regimens have been suggested to optimize training the "bite-smile" to achieve a smile that is balanced with the contralateral healthy side. These exercises could lead to temporomandibular joint (TMJ) pain due to overuse.²⁵ The aim of this study was to understand the patients' perspective of the NTM-driven smile and to investigate the rates of possible accessory complaints.

Methods

Institutional review board approval was obtained before the start of this cross-sectional study. All patients treated with a V-to-VII nerve transfer or gracilis functional free muscle

transfer (FFMT) innervated by the NTM between January 2007 and December 2016 at our center were invited to complete an online questionnaire. Patient charts were reviewed for patient characteristics, and the eFace score for "oral commissure movement with smile."

Questionnaire development and administration

A questionnaire was developed to assess patient experience following NTM-driven smile reanimation surgery based on the clinical expertise of our surgeons and physical therapists. Relevant items from the "Facial Clinimetric Evaluation (FaCE) scale" (facial palsy-related quality of life) and the "Research Diagnostic Criteria for Temporomandibular Disorders" were included in our survey. During a qualitative stage of survey development, four surgeons, two physical therapists, and five patients were interviewed. This was done to ensure direct input from patients into the content of our study questionnaire. Four domains were identified as being of critical importance: smile and symmetry, eating and social behavior, TMJ complaints, and physical therapy exercises. The questionnaire comprised closed-ended questions in both Likert scale and yes/no formats (Appendix - Full questionnaire). The questionnaire was administered through an online survey platform (REDCap, Vanderbilt University, Nashville, USA) to all patients for whom an email address was valid.

Data analysis

Nominal data were presented as frequencies and percentages. Ordinal data were presented as medians and interquartile ranges. Interval data were presented as means and standard deviations or medians and interquartile ranges, depending on whether or not the data were normally distributed. Testing between groups was performed using chi-squared tests and Mann-Whitney U tests. Correlation coefficients were tested by Spearman correlation. Statistical analysis was performed in SPSS version 25 with an alpha of 0.05.

Results

NTM-driven smile reanimation was performed in 171 patients between January 2007 and December 2016 (104 gracilis FFMT and 67 V-to-VII nerve transfers). A valid email address was known for 122 patients, and two patients declined to participate in the study. A total of 71 (58.2%) patients responded to our invitation to participate.

Table 1 Patient characteristics ($n = 71$).

Gender (n (%))	
Male	26 (36.6)
Female	45 (63.4)
Side (n (%))	
Left	33 (46.5)
Right	34 (47.9)
Bilateral	4 (5.6)
Etiology (n (%))	
Acoustic neuroma	20 (28.2)
Bell's palsy	9 (12.7)
Benign head and neck tumor	19 (26.8)
Malignant head and neck tumor	8 (11.3)
Otologic surgery	3 (4.2)
Trauma	4 (5.6)
Other	8 (11.3)
Age at the time of study, years (mean (SD))	
51.1 (15.6)	
Follow-up duration, years (median (IQR))	
3.8 (2.2; 5.0)	
Type surgery (n (%))	
Gracilis by V	42 (59.2)
V-to-VII nerve transfer	29 (40.8)
Additional static midface procedure (n (%))	
29 (40.8)	
Type of facial palsy (n (%))	
Flaccid	47 (66.2)
Non-flaccid	24 (33.8)

Abbreviations: IQR: interquartile range, n : number, SD: standard deviation, V: nerve-to-masseter, VII: facial nerve.

Of the 71 respondents, 42 (59.2%) underwent gracilis FFMT (innervated by the NTM alone) and 29 (40.8%) underwent V-to-VII nerve transfer. Mean (SD) age at the time of study was 51.1 (15.6) years. Median (IQR) follow-up was 3.8 (2.2; 5.0) years. Most patients presented with facial palsy because of an acoustic neuroma (28.2%). In addition to smile reanimation surgery, static suspension with fascia lata was performed in 29 (40.8%) patients. Forty-seven patients (66.2%) had flaccid facial palsy, and the remaining patients had nonflaccid facial palsy ($n = 24$, 33.8%) (Table 1). The patients included in this study did not differ from the rest of the operative cohort in gender ($p = 0.749$) or age ($p = 0.128$). The time between operation and study date was shorter for respondents than for nonrespondents ($p = 0.011$).

A majority of patients ($n = 59$, 83.1%) described being able to smile while biting down at least “most of the time,” while fewer patients ($n = 33$, 46.5%) were able to smile without biting down (Figure 1). Seventeen patients (23.9%) reported a “normal” or “almost normal” spontaneous smile (Figure 2). No statistically significant difference in self-reported ability to smile (voluntary with and without biting down, and spontaneous) was found between patients with flaccid and nonflaccid facial palsy ($p = 0.902$, $p = 0.450$, and $p = 0.233$, respectively). Self-reported smiling abilities were not statistically significantly different for patients with a follow-up shorter than 2 years than for patients with a follow-up longer than 2 years for all three smile types ($p = 0.692$, $p = 0.930$, and $p = 0.668$, respectively). A weak and nonsignificant correlation was found

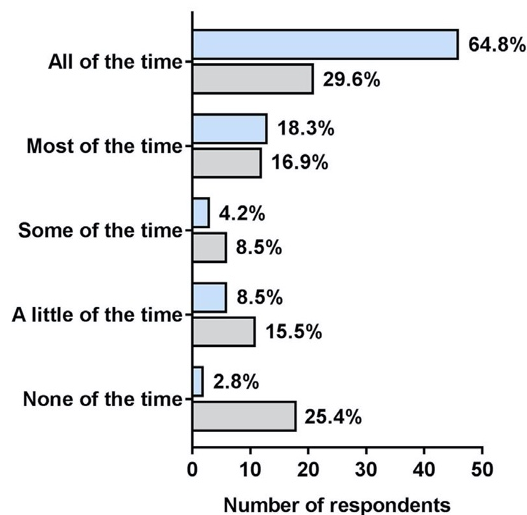


Figure 1 Bar charts of patient self-reported ability to smile voluntary with biting down (blue) and without biting down (gray) (missing: $n = 1$ (1.4%) and $n = 3$ (4.2%), respectively).

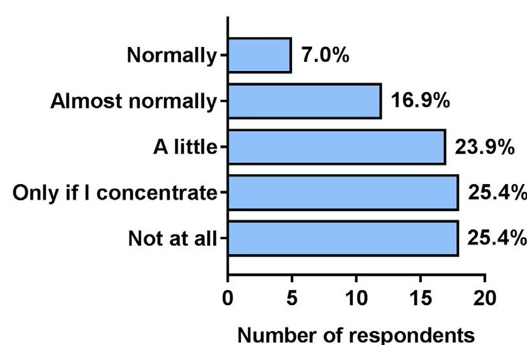


Figure 2 Patient self-reported spontaneity of smile (missing: $n = 1$ (1.4%)).

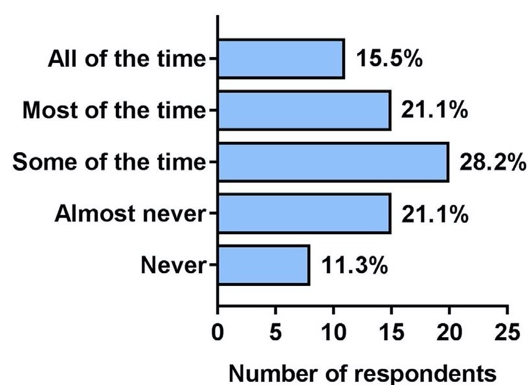


Figure 3 Answers to the question: “How often do you use your bite-smile?” (missing: $n = 2$ (2.8%)).

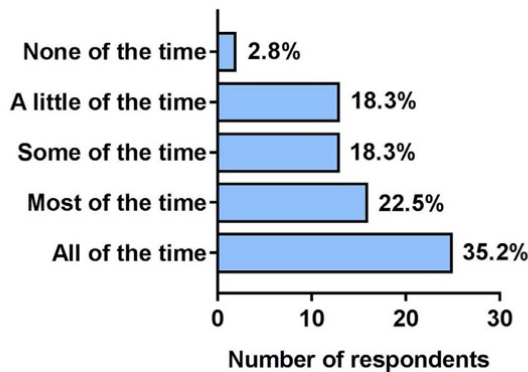
between clinician-observed smiling ability and patient self-reported voluntary smiling ability (Spearman correlation coefficient = 0.175, $p = 0.235$). Twenty-three patients (32.4%) reported “never” or “almost never” using their bite-driven smile (Figure 3), with 14 patients (19.8%) reporting the need to bite down to smile as being “bothersome” or “very bothersome”. In descending order of frequency, patients

Table 2 Answers to the question “Have you noticed any loss of volume or asymmetry over your jaw on your affected side compared to your healthy side?”.

	All patients (<i>n</i> = 71)	Gracilis by V (<i>n</i> = 42)	V-to-VII nerve transfer (<i>n</i> = 29)
Yes (<i>n</i> (%) ^a)	13 (18.3)	6 (14.3)	7 (24.1)
No (<i>n</i> (%) ^a)	38 (53.5)	25 (59.5)	13 (44.8)
I am not sure (<i>n</i> (%) ^a)	18 (25.4)	10 (23.8)	8 (27.6)
Missing (<i>n</i> (%) ^a)	2 (2.8)	1 (2.4)	1 (3.4)

^a Percentages are column percentages.

Abbreviations: *n*: number, V: nerve-to-masseter, VII: facial nerve.

**Figure 4** Answers to the question: “When you eat, does the affected side of your face move?” (missing: *n* = 2 (2.8%)).

reported using their bite smile for “photos” (*n* = 34, 47.9%), “meeting new people” (*n* = 20, 28.2%), “interacting with acquaintances” (*n* = 16, 22.5%), and “interacting with close family or friends” (*n* = 14, 19.7%).

Sixty-five patients (91.6%) reported improved symmetry of the face after surgery, and 63 patients (88.8%) reported improved symmetry of their smile. Forty-three patients (60.5%) scored their face as symmetric, and 30 patients (42.3%) scored their smile as symmetric. No significant difference was seen between patients who were and who were not treated with a concurrent static midface suspension procedure (using a Bonferroni corrected *p*-value for four tests (*p* = 0.013)).

Thirteen patients (18.3%) perceived masseter muscle atrophy or facial contour asymmetry following NTM transfer (Table 2). Although this difference was not statistically significant (*p* = 0.429), a larger proportion of the “V-to-VII nerve transfer” patients reported self-perceived masseter muscle atrophy or asymmetry than the “gracilis by V” patients (24.1% versus 14.3%, respectively) (Table 2). More than half of respondents perceived cheek movement while eating “all of the time” (*n* = 25, 35.2%) or “most of the time” (*n* = 16, 22.5%) (Figure 4), with eight patients (11.3%) reporting movements such as “bothersome”, and one patient (1.4%) reporting “very bothersome.” Patients generally reported few complaints regarding TMJ function (Table 3). Fifteen patients (21.1%) reported preoperative existence of problems with the TMJ or a preoperative diagnosis of TMJ disorder.

Fifty-five of the 71 respondents (77.5%) received targeted facial nerve physical therapy. Thirty-five patients (49.3%) reported adherence to the physical therapy treatment plan. Twenty-seven patients (38.0%) stated that phys-

ical therapy was an important part of their treatment plan, and 36 patients (50.7%) stated that physical therapy helped them to understand how to use their “bite-driven smile”.

A full list of the answers to all questions in the questionnaire is given in the appendix.

Discussion

The NTM is growing in popularity as a first-line neural source to reinnervate smile in facial palsy in our center. The aim of this study was to better understand patients’ perspectives regarding NTM-driven smile reanimation and to quantify potential and reported side effects of its use.

We found that the vast majority of respondents were able to smile voluntarily (83.1%) while biting down. A “normal” or “almost normal” spontaneous smile was self-reported by only 23.9% of patients in our study, which is in line with the 23% of spontaneous smile rates after NTM-driven smile reanimation found by a recent systematic review.⁸ The ability to smile emotionally or spontaneously following NTM-driven smile reanimation has been reported, and several explanations may account for this.⁸ The most commonly attributed rationale is cortical plasticity, whereby new neuronal connections are made in the brain allowing signals from the facial nerve nucleus to cause activation of the motor nucleus of trigeminal nerve.^{16,17} This is believed to occur more frequently in younger patients. Additionally, cortical plasticity is reported to occur more frequently in bilateral facial palsy patients.¹⁶ With only four patients with bilateral palsy in our series, we were not able to test this hypothesis statistically. Other groups describe improvements in spontaneity from intensive neuromuscular retraining.^{18,26} Another more direct explanation is that nearly 40% of normal subjects studied co-contract masseter and zygomaticus major while smiling.¹³ Thus, the presence of spontaneity in some patients may be due to this pre-existing co-contraction of the masseter muscle with smiling, making them ideal candidates for NTM-driven smile reanimation.¹³

While self-reported voluntary smiling ability is generally good following NTM-driven smile reanimation, a relevant proportion of patients (32.4%) report infrequent use of their bite-smile. Simultaneously, patients report less use of the bite smile when interacting with close friends or family than taking photos or meeting new people. This difference may indicate that the bite smile is seen as a good tool for non-verbal communication, which is known to be a critically important for psychosocial functioning after facial palsy.^{27–29}

Masseter muscle atrophy was self-reported in 18.3% of all patients, although some patients (25.4%) reported not being

Table 3 Answers to patient self-perceived complaints regarding the TMJ.

	All of the time	Most of the time	Some of the time	Almost never	Never	Missing
Pain/discomfort ^b (n (%))	1 (1.4)	2 (2.8)	16 (22.5)	9 (12.7)	41 (57.7)	2 (2.8)
Clicking/locking ^c (n (%))	1 (1.4)	2 (2.8)	12 (16.9)	16 (22.5)	39 (54.9)	1 (1.4)
Difficulty chewing ^d (n (%))	4 (5.6)	5 (7.0)	15 (21.1)	23 (32.4)	22 (31.0)	2 (2.8)
Difficulty opening ^e (n (%))	10 (14.1)	17 (23.9)	17 (23.9)	7 (9.9)	15 (21.1)	5 (7.0)

^a Percentages are row percentages.

^b Do you experience pain or discomfort around the joint of the jaw in front of the ear?

^c Do you experience clicking or locking of the lower jaw?

^d Do you have difficulty chewing certain foods?

^e Do you have difficulty opening your mouth wide to eat certain foods (e.g., an apple or hamburger)? Abbreviations: n: number, TMJ: temporomandibular joint.

sure whether masseter muscle atrophy was present. The authors of a case series of seven patients treated with V-to-VII nerve transfer report no postoperative problems with chewing and only one case (14.3%) of facial contour asymmetry due to masseter muscle atrophy.²² An important option in preventing masseter muscle atrophy from happening is preserving the anterior branch of the NTM during donor nerve mobilization, which we have started to do more recently as well.

Only 2.8% of respondents reported never having experienced movement of the oral commissure with prandial activity. Postoperative prandial movement of the cheek was reported by 97.2% of all patients, ranging from “a little of the time” to “all of the time.” A small proportion of patients reported this involuntary movement to be “bothersome” (11.3%) or “very bothersome” (1.4%). In a study solely focused on examining involuntary movement during mastication in NTM-driven gracilis FFMT patients, the authors acquired postoperative videos of 16 patients eating and evaluated the pattern of mastication, observer-perceived cheek movement, and social hindrance according to the patient.¹⁹ The authors observed prandial movement in all 16 patients, but only 33% of the patients self-report being aware of prandial movement. Similar to our study, prandial movement was not considered bothersome by the majority of patients.

TMJ complaints are relatively common in the general population, with prevalence rates of 6-12% being reported.³⁰ In our study, “pain or discomfort” and “clicking or locking” of the TMJ were reported by a small proportion of patients. Difficulty chewing and difficulty opening the mouth, in particular, were reported more often (Table 3), but 21.1% of patients also reported a pre-existing diagnosis of TMJ disorders or complaints. Given the high rate of preoperative TMJ complaints, it seems unlikely that NTM-driven smile reanimation techniques cause complaints of the TMJ. This is further supported by a case series of 11 patients after V-to-VII nerve transfer.²⁵ The authors evaluated the rehabilitation of the facial muscles after the procedure with clinical examination of the masseter muscle and TMJ, finding no symptoms or signs of TMJ dysfunction after 18 months. However, this is a clinician report, and symptoms such as patient self-perceived pain of the TMJ were not evaluated.

Physical therapy is known to be of benefit in the rehabilitation of NTM-driven smile reanimation.^{5,25,26} Physical therapy is important in not only helping patients initially using their bite-driven smile but also teaching how to generate

a smile without active bite movement and how to avoid cheek movement when eating. Patients generally share this view, with relatively high proportions of patients stating that physical therapy is important, helped them to understand how to activate their smile, and a good self-reported adherence to the physical therapy treatment plan. It is our opinion that optimal smile reanimation requires a comprehensive team effort including the surgeon and physical therapist.

This study has several limitations. All responses were retrospective and thus subject to recall and responder bias, although a response rate of approximately 60% is sufficient to make some conclusions from the data presented. There were also very few differences found between the responder and nonresponder groups, thereby limiting the degree of responder bias. All rates presented were self-reported by patients through a questionnaire, and the use of closed-ended questions may have limited the answer options for some patients. Although the present study shows that rates of accessory complaints among NTM-driven smile reanimation patients are low, qualitative studies specifically addressing these individual areas could be useful in the future. Additionally, we chose to utilize an online platform for administering this questionnaire, ruling out patients without an email address, which may represent a source of selection bias.

This study was not designed to predict which patients would be more likely to have a successful outcome after smile reanimation driven by the NTM versus other innervation strategies. Other studies such as discrete choice experiments are needed to answer this important question. Similarly, patients who underwent cross-facial nerve grafting or dual-innervated reanimations were not surveyed, and hence, it is not reasonable to make comparisons between techniques using these data. Future studies should attempt to objectively quantify rates of spontaneous smiling after various innervation strategies in addition to including patient-reported data.

Conclusion

Patients report good voluntary smiling ability following NTM-driven smile reanimation surgery, with low rates of sequelae. Patients report generally low rates of spontaneous smiling, and one-third of patients report rarely using their bite-driven smile. The present study supports the acclaimed

low morbidity of NTM-driven smile reanimation surgery. Further prospective studies are required to analyze the effect of dual-innervation strategies and to be able to objectively quantify spontaneous smiling.

Conflict of interest

The authors have no personal or financial relationships to disclose.

Funding

None.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.bjps.2019.03.037.

References

- Jowett N, Hadlock TA. Free gracilis transfer and static facial suspension for midfacial reanimation in long-standing flaccid facial palsy. *Otolaryngol Clin North Am* 2018;51(6):1129-39 doi:S0030-6665(18)30131-2 [pii].
- Borschel GH, Kawamura DH, Kasukurthi R, Hunter DA, Zuker RM, Woo AS. The motor nerve to the masseter muscle: an anatomic and histomorphometric study to facilitate its use in facial reanimation. *J Plast Reconstr Aesth Surg* 2011;65(3):363-6. doi:10.1016/j.bjps.2011.09.026.
- Bhama PK, Weinberg JS, Lindsay RW, Hohman MH, Cheney ML, Hadlock TA. Objective outcomes analysis following microvascular gracilis transfer for facial reanimation: a review of 10 years' experience. *JAMA Fac Plast Surg* 2014;16(2):85-92. doi:10.1001/jamafacial.2013.2463.
- Snyder-Warwick AK, Fattah AY, Zive L, Halliday W, Borschel GH, Zuker RM. The degree of facial movement following microvascular muscle transfer in pediatric facial reanimation depends on donor motor nerve axonal density. *Plast Reconstr Surg* 2015;135(2):81e [doi]. doi:10.1097/PRS.0000000000000860.
- García R, Hadlock T, Klebuc M, Simpson R, Zenn M, Marcus J. Contemporary solutions for the treatment of facial nerve paralysis. *Plast Reconstr Surg* 2015;135(6):1046e. doi:10.1097/PRS.0000000000001273.
- Coombs CJ, Ek EW, Wu T, Cleland H, Leung MK. Masseteric-facial nerve coaptation - an alternative technique for facial nerve reinnervation. *J Plast Reconstr Aesthet Surg* 2009;62(12):1580-8 [doi]. doi:10.1016/j.bjps.2008.05.046.
- Klebuc MJA. Facial reanimation using the masseter-to-facial nerve transfer. *Plast Reconstr Surg* 2011;127(5):1909-15. doi:10.1097/PRS.0b013e31820e9138.
- Murphy AW, Clinkscales WB, Oyer SL. Masseteric nerve transfer for facial nerve paralysis: a systematic review and meta-analysis. *JAMA Fac Plast Surg* 2018. doi:10.1001/jamafacial.2017.1780.
- O'Brien BM, Franklin JD, Morrison WA. Cross-facial nerve grafts and microvascular free muscle transfer for long established facial palsy. *Br J Plast Surg* 1980;33(2):202-15 doi:0007-1226(80)90013-2 [pii].
- Zuker RM, Goldberg CS, Manktelow RT. Facial animation in children with mobius syndrome after segmental gracilis muscle transplant. *Plast Reconstr Surg* 2000;106(1):8 discussion 9.
- Bae YC, Zuker RM, Manktelow RT, Wade S. A comparison of commissure excursion following gracilis muscle transplantation for facial paralysis using a cross-face nerve graft versus the motor nerve to the masseter nerve. *Plast Reconstr Surg* 2006;117(7):2407-13 [doi]. doi:10.1097/01.prs.0000218798.95027.21.
- Buendia J, Loayza FR, Luis EO, Celorrio M, Pastor MA, Hontanilla B. Functional and anatomical basis for brain plasticity in facial palsy rehabilitation using the masseteric nerve. *J Plast Reconstr Aesth Surg* 2015;69(3):417-26. doi:10.1016/j.bjps.2015.10.033.
- Schaverien M, Moran G, Stewart K, Addison P. Activation of the masseter muscle during normal smile production and the implications for dynamic reanimation surgery for facial paralysis. *J Plast Reconstr Aesthet Surg* 2011;64(12):1585-8 [doi]. doi:10.1016/j.bjps.2011.07.012.
- Hontanilla B, Cabello A. Spontaneity of smile after facial paralysis rehabilitation when using a non-facial donor nerve. *J Cranio Maxillo Fac Surg* 2016;44(9):1305-9. doi:10.1016/j.jcms.2016.06.031.
- Hontanilla B, Marre D. Differences between sexes in dissociation and spontaneity of smile in facial paralysis reanimation with the masseteric nerve. *Head Neck* 2014;36(8):1176-80. doi:10.1002/hed.23432.
- Manktelow RT, Tomat LR, Zuker RM, Chang M. Smile reconstruction in adults with free muscle transfer innervated by the masseter motor nerve: effectiveness and cerebral adaptation. *Plast Reconstr Surg* 2006;118(4):885-99. doi:10.1097/01.prs.0000232195.20293.bd.
- Garmi R, Labbé D, Coskun O, Compère J, Bénateau H. Lengthening temporalis myoplasty and brain plasticity: a functional magnetic resonance imaging study. *Annales de Chirurgie Plastique et Esthétique* 2013;58(4):271. <http://www.ncbi.nlm.nih.gov/pubmed/23623628>.
- Labbe D, Huault M. Lengthening temporalis myoplasty and lip reanimation. *Plast Reconstr Surg* 2000;105(4):97 discussion 1298.
- Rozen S, Harrison B. Involuntary movement during mastication in patients with long-term facial paralysis reanimated with a partial gracilis free neuromuscular flap innervated by the masseteric nerve. *Plast Reconstr Surg* 2013;132(1):116e. doi:10.1097/PRS.0b013e318290f644.
- Yoshioka N, Tominaga S. Masseteric nerve transfer for short-term facial paralysis following skull base surgery. *J Plast Reconstr Aesth Surg* 2015;68(6):764-70. doi:10.1016/j.bjps.2015.02.031.
- Biglioli F, Colombo V, Rabbiosi D, et al. Masseteric-facial nerve neurotization: results of a case series. *J Neurosurg* 2017;126(1):312-18. doi:10.3171/2015.12.JNS14601.
- Biglioli F, Frigerio A, Colombo V, et al. Masseteric-facial nerve anastomosis for early facial reanimation. *J Cranio Maxillo Fac Surg* 2011;40(2):149-55. doi:10.1016/j.jcms.2011.03.005.
- Chen G, Wang W, Wang W, Ding W, Yang X. Symmetry restoration at rest after masseter-to-facial nerve transfer: is it as efficient as smile reanimation. *Plast Reconstr Surg* 2017;140(4):793-801 [doi]. doi:10.1097/PRS.0000000000003698.
- Rozen SM. Discussion: symmetry restoration at rest after masseter-to-facial nerve transfer: is it as efficient as smile reanimation? *Plast Reconstr Surg* 2017;140(4):802-4 [doi]. doi:10.1097/PRS.0000000000003736.
- Pavesi C, Cecini M, Lozza A, et al. Rehabilitation and functional recovery after masseteric-facial nerve anastomosis. *Eur J Phys Rehabil Med* 2016;52(3):379-88 doi:R33Y999N00A150026 [pii].

26. Tzou CH, Chuang DC, Chen HY. Cortical adaptation staging system: a new and simple staging for result evaluation of functioning free-muscle transplantation for facial reanimation. *Ann Plast Surg* 2014;**73**(1):50-3.
27. Coulson SE, O'dwyer NJ, Adams RD, Croxson GR. Expression of emotion and quality of life after facial nerve paralysis. *Otol Neurotol* 2004;**25**(6):1014-19 doi:00129492-200411000-00026 [pii].
28. VanSwearingen JM, Cohn JF, Bajaj-Luthra A. Specific impairment of smiling increases the severity of depressive symptoms in patients with facial neuromuscular disorders. *Aesthetic Plast Surg* 1999;**23**(6):416-23 APS1273 [pii].
29. van Veen MM, Tavares-Brito J, van Veen BM, et al. Association of regional facial dysfunction with facial palsy-related quality of life. *JAMA Fac Plast Surg* 2018. doi:10.1001/jamafacial.2018.0804.
30. Liu F, Steinkeler A. Epidemiology, diagnosis, and treatment of temporomandibular disorders. *Dent Clin North Am* 2013;**57**(3):465-79. doi:10.1016/j.cden.2013.04.006.

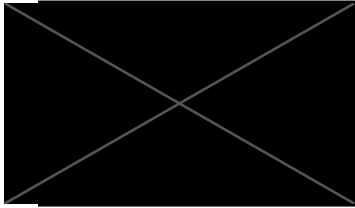
Appendix 3

Association of Regional Facial Dysfunction with Facial Palsy–Related Quality of Life

Authorship Statement

I, Joseph Dusseldorp, confirm for the ensuing journal article: ‘**Association of Regional Facial Dysfunction with Facial Palsy–Related Quality of Life**’ that I have made the following contributions to its authorship:

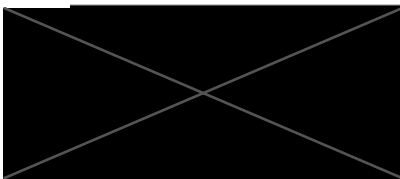
- Analysis and interpretation of the findings
- Revising of the manuscript and critical appraisal of content



Joseph R. Dusseldorp

Date 14.09. 2023

As Auxiliary Supervisor for the candidature upon which this thesis is based, I can confirm that the above authorship attribution is correct.

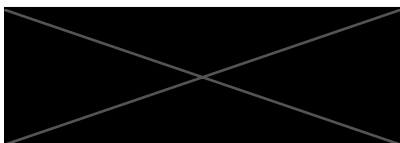


Professor Tessa Hadlock, MD

Date 27.07. 2023

Department of Otolaryngology
Massachusetts Eye and Ear Infirmary
Harvard University, Boston, MA, USA

As Primary Supervisor of the candidate, I concur with the above authorship attestation:



Professor Peter Maitz

Date 14.09.2023

Department of Burns Surgery
Concord Repatriation General Hospital
Sydney University, Sydney, NSW, Australia

Association of Regional Facial Dysfunction With Facial Palsy–Related Quality of Life

Martinus M. van Veen, MD; Joana Tavares-Brito, MD; Britt M. van Veen, BSc; Joseph R. Dusseldorp, MBBS, MS; Paul M. N. Werker, MD, PhD; Pieter U. Dijkstra, PT, PhD; Tessa A. Hadlock, MD

 Supplemental content

IMPORTANCE Study of the association of regional facial dysfunction with quality of life will lead to a better understanding of quality of life in facial palsy.

OBJECTIVE To determine the association of regional facial dysfunction with facial palsy–related quality of life.

DESIGN, SETTING, AND PARTICIPANTS This retrospective cohort analysis included patients with flaccid and nonflaccid (synkinetic) facial palsy treated at a tertiary care facial nerve center; the flaccid facial palsy group included 529 patients, and the nonflaccid facial palsy group included 391 patients. Data were included from all patients with facial palsy who had an eFACE score and Facial Clinimetric Evaluation (FaCE) scale total score acquired at the same time from February 1, 2014, through October 31, 2017. Linear regression analysis was performed to calculate the amount of variance in quality of life explained by the severity of facial palsy (eFACE). A relative weight analysis was performed for the contribution of each individual eFACE item in estimating quality of life.

MAIN OUTCOMES AND MEASURES Facial palsy severity was measured using all 15 individual eFACE items (rated on a scale of 0 to 200, where 0 represents complete flaccidity, 100 represents a balanced aesthetic appearance, and 200 represents the worst imaginable hypertonia of a patient with synkinesis, with a transformation used for values from 101 to 200), and facial palsy–related quality of life was measured using the FaCE scale total score (range, 0 [worst] to 100 [best]).

RESULTS Data of 920 individual patients (59.5% female; mean [SD] age, 48.6 [16.6] years) were available. The eFACE composite score accounted for 21.2% of the quality-of-life variance in the flaccid group and 13.9% in the nonflaccid group. With the use of all 15 individual eFACE items, these proportions increased to 29.7% and 16.8%, respectively. In both groups, oral commissure movement with smile was found to be the most important contributing item (relative weight, 0.108 [95% CI, 0.075–0.148] for the flaccid group and 0.025 [95% CI, 0.005–0.052] for the nonflaccid group). Items related to the function of periocular muscles were found to be of low importance.

CONCLUSIONS AND RELEVANCE The present study suggests that the function of individual facial regions is not equally important for estimating facial palsy–related quality of life. The ability to smile is of greatest importance among patients with flaccid and nonflaccid facial palsy. The true importance of periocular function in the estimation of quality of life should be studied further in future research.

LEVEL OF EVIDENCE NA.

JAMA Facial Plast Surg. doi:10.1001/jamafacial.2018.0804
Published online August 2, 2018.

Author Affiliations: Author affiliations are listed at the end of this article.

Corresponding Author: Martinus M. van Veen, MD, Department of Plastic Surgery, University Medical Center Groningen, University of Groningen, PO Box 30.001, NL-9700 RB Groningen, the Netherlands (m.m.van.veen@umcg.nl).

Facial palsy may range in severity from near-normal facial function to crippling facial disfigurement. Facial palsy severity has been recognized to influence patient-perceived quality of life to varying degrees, with some studies describing an association¹⁻³ and others describing no such association.^{4,5} Measurement scales and study populations differ greatly between studies, and most studies that did not find an association between facial palsy severity and quality of life used the House-Brackmann Scale.⁶ Studies that have found an association between severity of facial palsy and quality of life have found the strength of the association to be weak¹⁻³ and often used the composite score of the Sunnybrook Facial Grading System,⁷ thereby calculating the mean scores for disability of different facial zones. Most well-established facial grading systems apply a regional assessment of facial function. For example, the recently developed eFACE facial grading system assesses 15 items, including 4 static, 7 dynamic, and 4 synkinesis items, enabling generation of a composite score and static, dynamic, and synkinesis subscores.⁸ Although the composite score is an effective approach for presenting a gross impression of facial function, valuable information about regional facial function is lost in composite scores. Individual regional facial grading items may each have a different influence on perceived quality of life that is not represented in composite scores. The aim of this study was to examine the influence of individual facial grading items on the statistical estimation of facial palsy-related quality of life by means of relative weight analysis and thereby achieve a better understanding of the association of regional facial dysfunction with quality of life.

Methods

Patients and Study Overview

We selected a retrospective cohort of all patients with facial palsy treated at Massachusetts Eye and Ear Infirmary, Boston, from February 1, 2014, through October 31, 2017. Criteria for inclusion in this study were the presence of an eFACE assessment and a Facial Clinimetric Evaluation (FaCE) scale score obtained at the same time. Patients were excluded if clinical notes were missing, the FaCE scale was incomplete, or they had bilateral facial palsy because the eFACE system is primarily intended for unilateral facial palsy. This study was approved by the institutional review board of the Massachusetts Eye and Ear Infirmary, and all patients provided written informed consent.

Since its development in early 2014, the eFACE facial grading system has been our standard measure of clinician-graded facial function. The eFACE consists of 4 static, 7 dynamic, and 4 synkinesis items.⁸ Ratings of each item are used to calculate static, dynamic, and synkinesis subscores and an eFACE total score (range, 0 [worst] to 100 [best]). Nine eFACE items are rated on a scale of 0 (worst) to 100 (best). Four items regarding static characteristics and 2 items regarding dynamic characteristics relate to the nasolabial fold (NLF) and are rated on a scale of 0 to 200, where 0 represents complete flaccidity, 100 represents a balanced

Key Points

Question Which individual facial grading item contributes most to the estimation of facial palsy-related quality of life?

Findings In this cohort study of 920 patients with facial palsy, there were 529 with absent synkinesis (flaccid group) and 391 with synkinesis (nonflaccid group). Oral commissure movement with smile was the eFACE item that contributed most in the estimation of facial palsy-related quality of life in groups with facial palsy (relative weight, 0.108 for the flaccid group and 0.025 for the nonflaccid group).

Meaning The present study demonstrates that the function of individual facial regions is not equally important for the estimation of facial palsy-related quality of life; the ability to smile is of greatest importance in patients with flaccid and nonflaccid facial palsy.

aesthetic appearance, and 200 represents the worst imaginable hypertonia of a patient with synkinesis. When calculating the eFACE total score or subscores, a transformation is used for values from 101 to 200. These hyperkinetic values are transformed to their absolute distance from 200 ($200 - \text{item score}$) to be comparable to the flaccid values. The same transformation into standardized eFACE scores was used for all analyses.

We routinely use the patient-reported FaCE scale¹ as a measure of facial palsy-related quality of life at our institution. The FaCE scale consists of 15 items with 5-point Likert scale-type answers that are transformed to a FaCE total score (range, 0 [worst] to 100 [best]).

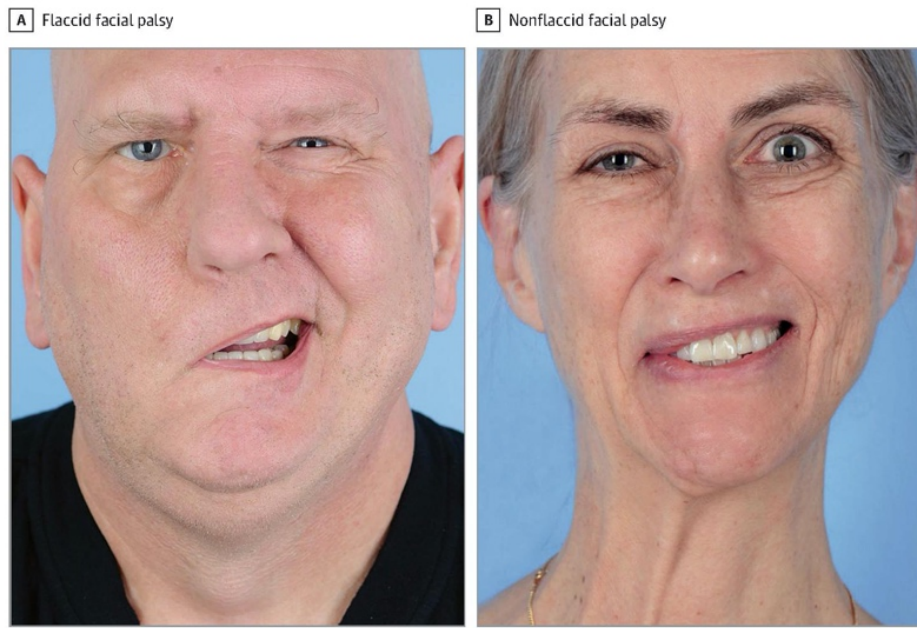
The eFACE and FaCE scale assessments are routinely performed during the initial consultation and thereafter, by indication. If more than 1 pair of FaCE and eFACE scores was present, the assessments of the senior author (T.A.H.) were used in this study for consistency. In addition, we searched medical records for general patient characteristics.

We divided the cohort between a group in whom synkinesis was absent (flaccid group) and a group in whom synkinesis was present (nonflaccid group), as determined by the eFACE synkinesis score, for all analyses. Patients with a synkinesis score of 100 were included in the flaccid group, and patients with a synkinesis score ranging from 0 to 99 were included in the nonflaccid group. Representative photographs of patients from each group are presented in the Figure.

Statistical Analysis

Descriptive statistics are given using numbers and frequencies, medians and interquartile ranges, and means and SDs when appropriate. We performed linear regression analyses to study the size of the association between the eFACE and FaCE scale total scores for both groups. The same analyses were also performed for the association of individual eFACE items with the FaCE scale total score. In linear regression, the model correlation coefficient (R^2) represents the proportion of variance in the output variable (FaCE scale total score) that can be explained by the input variable (the eFACE total score or the 15 individual eFACE item subscores).

Figure. Representative Photographs of Patients With Right-Sided Facial Palsy



When determining the relative weight (defined as the proportionate contribution to the correlation coefficient) of individual eFACE items, multicollinearity may occur. Multicollinearity is a phenomenon whereby an estimator variable can be derived from another estimator variable, which might influence the regression coefficients of the estimator variables in a multiple regression model. The technique of relative weight analysis described by Johnson⁹ overcomes this problem and was used. In this technique, noncorrelating estimator variables are derived from the original correlating variables and are used to calculate the proportionate contribution of each individual original input variable.⁹ We determined the significance of relative weights based on the 95% CI.¹⁰ Relative weight estimation was performed using R software (version 3.4; R Foundation for Statistical Computing), and all other analyses were performed using SPSS software (version 25.0; IBM Corporation).

Results

The eFACE scores of 1304 unique patients were acquired during the study period. For 975 patients, a FaCE scale score was available from the same time. Fifty-five patients were excluded owing to missing data regarding clinical notes and facial photographs because some FaCE scale questions were not filled out or because they had bilateral facial palsy, leaving data from 920 patients (70.6%) available for this study (373 men [40.5%] and 547 women [59.5%]; mean [SD] age, 48.6 [16.6] years). The flaccid group included 529 patients; the nonflaccid group, 391. In the flaccid group, 282 patients (53.3%) were women, and the mean (SD) age was 49.5 (17.4) years. In the nonflaccid group, 265 patients (67.8%) were women, and the mean (SD) age was 47.3 (15.5) years. Median

(interquartile range) eFACE total scores, 15 individual eFACE subscores, and FaCE scale total scores for both groups are presented in Table 1.

Two linear regression analyses were performed with the eFACE total score as the input variable and FaCE scale total score as output variable. For the flaccid group, the model correlation was 0.212 ($P < .001$) (eFigure 1 in the Supplement). For the nonflaccid group, the model correlation was 0.139 ($P < .001$) (eFigure 2 in the Supplement). In a multiple regression model, with all 15 eFACE items as input variables instead of the eFACE total score and the FaCE scale total score as the output variables, the model R^2 value increased with 0.085 to 0.297 ($P < .001$) for the flaccid group. In the nonflaccid group, the model R^2 value increased with 0.029 to 0.168 ($P < .001$).

Results of the relative weight analysis showed that, for both groups, oral commissure movement with smile was the most important item in the estimation of the FaCE scale total score (relative weight, 0.108 [95% CI, 0.075-0.148] for the flaccid group and 0.025 [95% CI, 0.005-0.052] for the nonflaccid group). All 7 eFACE dynamic items (brow elevation [relative weight, 0.038; 95% CI, 0.021-0.062], NLF depth with smile [relative weight, 0.038; 95% CI, 0.023-0.057], NLF orientation with smile [relative weight, 0.037; 95% CI, 0.024-0.053], lower lip movement with /ee/ pronunciation [relative weight, 0.028; 95% CI, 0.011-0.054], gentle eye closure [relative weight, 0.010; 95% CI, 0.004-0.020], full eye closure [relative weight, 0.006; 95% CI, 0.001-0.012], and oral commissure with smile [relative weight, 0.108; 95% CI, 0.075-0.148]) and 2 static eFACE items (NLF depth at rest [relative weight, 0.015; 95% CI, 0.006-0.030] and oral commissure at rest [relative weight, 0.008; 95% CI, 0.002-0.018]) were significant in the estimation of the FaCE scale total score in the flaccid group.

Table 1. Spread of Individual eFACE Items in the Flaccid and Nonflaccid Groups

Item ^a	Facial Palsy Group, Median (IQR) Score	
	Flaccid (n = 529)	Nonflaccid (n = 391)
eFACE total score	51 (36-61)	72 (65-83)
eFACE individual items		
Resting brow ^b	85 (69-100)	90 (80-100)
Resting PFW ^b	80 (61-91)	87 (75-98)
NLF depth at rest ^b	77 (51-91)	87 (78-100)
OC at rest ^b	80 (57-94)	90 (82-100)
Brow elevation	4 (0-62)	39 (8-89)
Gentle eye closure	81 (58-97)	100 (91-100)
Full eye closure	94 (75-100)	100 (100-100)
NLF depth with smile ^b	38 (0-87)	88 (75-100)
OC movement with smile	7 (0-61)	32 (10-76)
NLF orientation with smile ^b	37 (0-82)	78 (56-90)
Lower lip movement with /ee/ pronunciation	54 (28-86)	67 (38-89)
Ocular synkinesis	100 (100-100)	67 (30-93)
Midfacial synkinesis	100 (100-100)	78 (56-92)
Mentalis dimpling	100 (100-100)	74 (49-93)
Platysmal synkinesis	100 (100-100)	79 (52-94)
FACE scale total score	45 (31-59)	67 (58-81)

Abbreviations: FaCE, Facial Clinimetric Evaluation; IQR, interquartile range; NLF, nasolabial fold; OC, oral commissure; PFW, palpebral fissure width.

^a Unless otherwise indicated, transformed scores range from 0 (worst) to 100 (best).

^b Scores range from 0 to 200, where 0 represents complete flaccidity, 100 represents a balanced aesthetic appearance, and 200 represents the worst imaginable hypertonia of a patient with synkinesis.

In the nonflaccid group, the eFACE items oral commissure with smile (relative weight, 0.025; 95% CI, 0.005-0.052), platysmal synkinesis (relative weight, 0.024; 95% CI, 0.002-0.055), and lower lip movement with /ee/ pronunciation (relative weight, 0.023; 95% CI, 0.003-0.056) were significant in the estimation of the FaCE scale total score (Table 2 and eFigure 3 in the Supplement).

Discussion

The aim of this study was to better understand the association of individual components of regional facial function with estimating quality of life in patients with facial palsy. We found that facial palsy-related quality of life is better estimated by regional facial function items than a composite score of facial function. In addition, quality of life can be estimated better in patients with flaccid compared with nonflaccid facial palsy.

Overall facial palsy severity is associated with facial palsy-related quality of life more closely in patients who do not have synkinesis (the flaccid group) compared with patients in whom synkinesis is present (nonflaccid group), with an explained variance of 21.2% and 13.9%, respectively. Estimating quality of life using all 15 eFACE items increased the model correlation to 29.7% in the flaccid group but only 16.8% in the nonflaccid group. This observation demonstrates that use of the regional

Table 2. Relative Weights and 95% CIs in the Flaccid and Nonflaccid Groups

Item	Relative Weight (95% CI)	
	Flaccid Group (n = 529)	Synkinesis Group (n = 391)
eFACE static variables		
Resting brow	0.006 (−0.001 to 0.023)	0.004 (−0.010 to 0.027)
Resting PFW	0.004 (−0.001 to 0.013)	0.006 (−0.009 to 0.026)
OC at rest	0.008 (0.002 to 0.018) ^a	0.002 (−0.022 to 0.006)
NLF depth at rest	0.015 (0.006 to 0.030) ^a	0.010 (−0.007 to 0.030)
eFACE dynamic variables		
Brow elevation	0.038 (0.021 to 0.062) ^a	0.017 (−0.001 to 0.040)
Gentle eye closure	0.010 (0.004 to 0.020) ^a	0.002 (−0.022 to 0.006)
Full eye closure	0.006 (0.001 to 0.012) ^a	0.003 (−0.012 to 0.013)
NLF depth with smile	0.038 (0.023 to 0.057) ^a	0.008 (−0.008 to 0.032)
OC movement with smile	0.108 (0.075 to 0.148) ^a	0.025 (0.005 to 0.052) ^a
NLF orientation with smile	0.037 (0.024 to 0.053) ^a	0.014 (−0.002 to 0.038)
Lower lip movement with /ee/	0.028 (0.011 to 0.054) ^a	0.023 (0.003 to 0.056) ^a
eFACE synkinesis variables		
Ocular synkinesis	NA	0.011 (−0.005 to 0.034)
Midface synkinesis	NA	0.017 (−0.003 to 0.045)
Mentalis dimpling	NA	0.003 (−0.021 to 0.007)
Platysmal synkinesis	NA	0.024 (0.002 to 0.055) ^a

Abbreviations: NA, not available; NLF, nasolabial fold; OC, oral commissure; PFW, palpebral fissure width.

^a Indicates items contribute significantly based on the 95% CI.

facial function items instead of composite scores helps to better estimate facial palsy-related quality of life.

Although the proportion of variance explained may seem low (ie, 13.9%-29.7%), these levels are in line with findings in other studies of disease severity and disease-specific quality of life. Compared with other variables estimating disease-specific quality of life in facial palsy, our R^2 values are relatively large. In a study investigating age, sex, and etiology and duration of palsy,³ the total proportion of variance was found to be 3.8% compared with 13.9% to 29.7% explained solely by facial palsy severity in the present study. Other factors influencing quality of life in facial palsy are yet to be unraveled.

The difference in explained variance between the flaccid and nonflaccid groups suggests that expert-rated severity of synkinesis correlates poorly with patient-reported severity of synkinesis. A future study of this association, with the use of the synkinesis-specific Synkinesis Assessment Questionnaire,¹¹ for example, on an individual regional item level would provide additional insights. In addition, the stronger correlation between eFACE scores and FaCE scale total score in the flaccid group could partly

be owing to the lack of specific synkinesis-related questions included in the FaCE scale.¹ Development of a synkinesis-specific extension to the FaCE scale may improve understanding of quality of life in a population with nonflaccid facial palsy.

The single most important contributing eFACE item in the flaccid group was found to be oral commissure movement with smile. The third and fourth most important contributing factors were NLF depth with smile and NLF orientation with smile. Impairment in these items may be owing to lack of smile but also to midfacial flaccidity, which results in problems with speaking, eating, and drinking. Together, these 3 items account for two-thirds of the total R^2 value, meaning that midfacial flaccidity is the single most important estimator of facial palsy-related quality of life. The relative importance of oral commissure movement with smile compared with oral commissure at rest also emphasizes the importance of a dynamic facial reconstruction for patients with flaccid facial palsy (Table 2 and eFigure 3 in the Supplement).

In the nonflaccid group, a more balanced contribution of each individual item to facial palsy-related quality of life was seen, although only 3 items contributed significantly. Although the most important contributor again was oral commissure movement with smile, the contribution of the items related to the NLF in the smile and midfacial synkinesis were much less important and not statistically significant. Patients with synkinesis usually have a good resting tone of the face and generally have fewer issues with speaking, eating, or drinking. Therefore, the inability to produce a balanced and sufficient smile likely represents the cause of quality-of-life impairment in this patient group.¹² Only 1 synkinesis item (platysmal synkinesis) was found to have a significant effect on facial palsy-related quality of life. The relatively low importance of the other 3 synkinesis items suggests a poor correlation between expert-rated and patient-experienced severity of synkinesis.

Of interest, lower lip movement with /ee/ pronunciation contributes significantly in both groups. Thus, asymmetry of the lower lip has a relatively high influence on quality of life in our sample. A low score on this item indicates weakness of the lower lip in both groups of patients because hypertonicity of the lower lip is very uncommon. Even patients with synkinesis are more likely to have persistent lower lip weakness after facial nerve insult and aberrant regeneration. Asymmetry of the lower lip can be effectively treated with botulinum toxin injections or depressor labii inferioris myectomy to the contralateral healthy lower lip.¹³⁻¹⁵ These interventions improve the esthetic appearance of the smile but do not restore movement. Several surgical techniques for dynamic restoration of the lower lip have been published,¹⁶ but only 1 nonrandomized study of 18 patients¹⁷ compared the effects of contralateral botulinum toxin injections with the effects of ipsilateral digastric muscle transfer. In that study,¹⁷ recipients of digastric muscle transfer were more likely to be dissatisfied with their outcome. Furthermore, pretreatment and post-treatment quality-of-life measurements are not available

in the literature. Future prospective and randomized studies using validated preoperative and postoperative quality-of-life outcome measures are needed to understand the benefits of static vs dynamic treatment of lower lip asymmetry.

Items related to the function of the periocular muscles (gentle eye closure, full eye closure, and ocular synkinesis) were found to be of low importance in the flaccid and the nonflaccid groups. The effect of periocular function on quality of life has been investigated previously.¹⁸ However, the relative importance of individual regional items that make up the composite score has not. Clinically, we expected to find these items to be of greater importance because impairment in eye closure can cause great discomfort. We considered that this low weighting of periocular items may be owing to prior periocular surgery and so cross-referenced all treatment records and a consecutive sample of 50 medical records and found only a 3% to 6% rate of prior surgery. This finding leaves room for further investigation because we do not fully understand why these items are less important. Alternatively, the FaCE scale only incorporates 3 items related to ocular problems, all aimed at problems in patients with flaccid facial palsy.¹ Thus, the use of the FaCE scale could have partly masked the true extent of the eye-related disability, especially in those with synkinesis.

In a recent study,¹⁹ the relative contribution of all eFACE items was compared with expert-rated disfigurement. Of the 5 significantly contributing items that were found, 3 were static (NLF depth at rest, oral commissure at rest, and palpebral fissure at rest). This finding may seem contrary to the finding of the present study that dynamic variables are of much greater importance in estimating quality of life. We hypothesize that dynamic variables are of greater importance in estimating quality of life because of the influence that dynamic asymmetries may have on interpersonal communication and the expression of emotion. A direct comparison is limited because of the difference between patient-reported quality of life vs expert-rated disfigurement constructs. Experts may be focused on different items of the eFACE, such as symmetry at rest, because of the clinical problems they know might arise from certain regional dysfunctions. In another study,²⁰ layperson perception of global and regional facial palsy was studied. Of the regional facial palsies, zygomatic and/or buccal facial palsy generally scored worst on domains such as perceived normality, distress, intelligence, and trustworthiness as rated by a layperson. These negative impressions can cause problems in social interaction for patients, which is known to negatively influence quality of life.²¹ The importance of the midface variables may be related to this zone's importance in interacting with other people.

Limitations

A limitation of our study is that we used clinician grading to determine severity of facial palsy. Although the eFACE instrument has been shown to be a reliable and valid method of assessing facial palsy severity,²²⁻²⁴ the use of

scores by different clinicians can introduce an element of error. However, all eFACE assessments in our study sample were performed by clinicians who are experienced in the whole spectrum of facial palsy thereby limiting the degree of error. Data of 920 patients (70.6%) were included in this study, presenting a potential source of selection bias. The patients excluded did not differ from the included patients in characteristics such as sex and age. The FaCE scale total scores of the 55 patients excluded because of missing data were relatively high compared with the scores of the included patients. These high quality-of-life scores suggests that the present cohort might represent a slightly lower quality of life than other cohorts of patients with facial palsy because of the tertiary care facial nerve center setting. However, the excluded patients had relatively high eFACE scores as well, suggesting that the associations found can be generalized.

Another limitation of our study is the constitution of the cohort studied. Our tertiary care center delivers care to the full spectrum of patients with facial palsy. This spectrum often includes patients who were not helped elsewhere or who referred themselves to our center. These patients may present

with problems regarding their facial function when findings of the clinical examination of the face are nearly normal. We hypothesize that quality of life in this patient subgroup may be disproportionately low. We included these patients because we set out to study the full range of facial palsy and the effect of different regional impairments. For that reason, patients with minor problems in only 1 region should be included.

Conclusions

Our data demonstrate the importance and value of regional assessment of facial dysfunction in facial palsy, instead of the use of composite scores. The study of regional facial function items provides a better understanding of facial palsy-specific quality of life. The relative importance of regional facial function items can point a direction toward future clinical and research aims. The present study shows a discrepancy in the relative importance of regional facial function items between facial palsy-related quality of life and expert-rated disfigurement.

ARTICLE INFORMATION

Accepted for Publication: May 19, 2018.

Published Online: August 2, 2018.
doi:10.1001/jamafacial.2018.0804

Author Affiliations: Facial Nerve Center, Department of Otolaryngology, Massachusetts Eye and Ear Infirmary, Harvard Medical School, Boston (M. M. van Veen, Tavares-Brito, Dusseldorp, Hadlock); Department of Plastic Surgery, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands (M. M. van Veen, Werker); Department of Economics, Econometrics and Finance, University of Groningen, Groningen, the Netherlands (B. M. van Veen); Department of Plastic and Reconstructive Surgery, Royal Australasian College of Surgeons, University of Sydney, Sydney, Australia (Dusseldorp); Department of Rehabilitation Medicine, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands (Dijkstra); Department of Oral and Maxillofacial Surgery, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands (Dijkstra).

Author Contributions: Dr M. M. van Veen had full access to all the data in the study and takes full responsibility for the integrity of the data and the accuracy of the data analysis.

Concept and design: M. M. van Veen, Tavares-Brito, Dusseldorp, Werker, Dijkstra, Hadlock.

Acquisition, analysis, or interpretation of data: M. M. van Veen, Tavares-Brito, B. M. van Veen, Dusseldorp, Dijkstra, Hadlock.

Drafting of the manuscript: M. M. van Veen, Dusseldorp, Dijkstra, Hadlock.

Critical revision of the manuscript for important intellectual content: All authors.

Statistical analysis: M. M. van Veen, B. M. van Veen.
Administrative, technical, or material support: M. M. van Veen, Tavares-Brito, Hadlock.

Supervision: Dusseldorp, Werker, Dijkstra.

Conflict of Interest Disclosures: None reported.

Additional Contributions: We thank the patients for granting permission to publish this information.

REFERENCES

- Kahn JB, Glicklich RE, Boyev KP, Stewart MG, Metson RB, McKenna MJ. Validation of a patient-graded instrument for facial nerve paralysis: the FaCE scale. *Laryngoscope*. 2001;111(3):387-398. doi:10.1097/00005537-200103000-00005
- Ng JH, Ngo RY. The use of the Facial Clinimetric Evaluation Scale as a patient-based grading system in Bell's palsy. *Laryngoscope*. 2013;123(5):1256-1260. doi:10.1002/lary.23790
- Kleiss IJ, Hohman MH, Susarla SM, Marres HA, Hadlock TA. Health-related quality of life in 794 patients with a peripheral facial palsy using the FaCE Scale: a retrospective cohort study. *Clin Otolaryngol*. 2015;40(6):651-656. doi:10.1111/coa.12434
- Lassaletta L, Alfonso C, Del Rio L, Roda JM, Gavilan J. Impact of facial dysfunction on quality of life after vestibular schwannoma surgery. *Ann Otol Rhinol Laryngol*. 2006;115(9):694-698. doi:10.1177/00034894061500908
- Lee J, Fung K, Lownie SP, Parnes LS. Assessing impairment and disability of facial paralysis in patients with vestibular schwannoma. *Arch Otolaryngol Head Neck Surg*. 2007;133(1):56-60. doi:10.1001/archotol.133.1.56
- House JW, Brackmann DE. Facial nerve grading system. *Otolaryngol Head Neck Surg*. 1985;93(2):146-147. doi:10.1177/019459988509300202
- Ross BG, Fradet G, Nedzelski JM. Development of a sensitive clinical facial grading system. *Otolaryngol Head Neck Surg*. 1996;114(3):380-386. doi:10.1016/S0194-5998(96)70206-1
- Banks CA, Bhama PK, Park J, Hadlock CR, Hadlock TA. Clinician-graded electronic facial paralysis assessment: the eFACE. *Plast Reconstr Surg*. 2015;136(2):223e-230e. doi:10.1097/PRS.0000000000001447
- Johnson JW. A heuristic method for estimating the relative weight of predictor variables in multiple regression. *Multivariate Behav Res*. 2000;35(1):1-19. doi:10.1207/S15327906MBR3501_1
- Tonidandel S, Lebreton JM, Johnson JW. Determining the statistical significance of relative weights. *Psychol Methods*. 2009;14(4):387-399. doi:10.1037/a0017735
- Mehta RP, WernickRobinson M, Hadlock TA. Validation of the Synkinesis Assessment Questionnaire. *Laryngoscope*. 2007;117(5):923-926. doi:10.1097/MLG.0b013e3180412460
- Coulson SE, O'dwyer NJ, Adams RD, Croxson GR. Expression of emotion and quality of life after facial nerve paralysis. *Otol Neurotol*. 2004;25(6):1014-1019. doi:10.1097/00129492-200411000-00026
- Chen CK, Tang YB. Myectomy and botulinum toxin for paralysis of the marginal mandibular branch of the facial nerve: a series of 76 cases. *Plast Reconstr Surg*. 2007;120(7):1859-1864. doi:10.1097/01.prs.0000287136.22709.77
- Hussain G, Manktelow RT, Tomat LR. Depressor labii inferioris resection: an effective treatment for marginal mandibular nerve paralysis. *Br J Plast Surg*. 2004;57(6):502-510. doi:10.1016/j.bjps.2004.04.003
- Tulley P, Webb A, Chana JS, et al. Paralysis of the marginal mandibular branch of the facial nerve: treatment options. *Br J Plast Surg*. 2000;53(5):378-385. doi:10.1054/bjps.2000.3318
- Terzis JK, Kalantarian B. Microsurgical strategies in 74 patients for restoration of dynamic depressor muscle mechanism: a neglected target in facial reanimation. *Plast Reconstr Surg*. 2000;105(6):1917-1931. doi:10.1097/00006534-200005000-00001
- Butler DP, Leckenby JL, Miranda BH, Grobbelaar AO. Botulinum toxin therapy versus anterior belly of digastric transfer in the management of marginal mandibular branch of the facial nerve palsy:

a patient satisfaction survey. *Arch Plast Surg*. 2015; 42(6):735-740. doi:10.5999/aps.2015.42.6.735

18. Henstrom DK, Lindsay RW, Cheney ML, Hadlock TA. Surgical treatment of the periocular complex and improvement of quality of life in patients with facial paralysis. *Arch Facial Plast Surg*. 2011;13(2):125-128. doi:10.1001/archfacial.2011.9

19. Banks CA, Jowett N, Hadlock CR, Hadlock TA. Weighting of facial grading variables to disfigurement in facial palsy. *JAMA Facial Plast Surg*. 2016;18(4):292-298. doi:10.1001/jamafacial.2016.0226

20. Li MK, Niles N, Gore S, Ebrahimi A, McGuinness J, Clark JR. Social perception of morbidity in facial nerve paralysis. *Head Neck*. 2016;38(8):1158-1163. doi:10.1002/hed.24299

21. Macgregor FC. Facial disfigurement: problems and management of social interaction and implications for mental health. *Aesthetic Plast Surg*. 1990;14(4):249-257. doi:10.1007/BF01578358

22. Banks CA, Jowett N, Hadlock TA. Test-retest reliability and agreement between in-person and video assessment of facial mimetic function using the eFACE facial grading system. *JAMA Facial Plast Surg*. 2017;19(3):206-211. doi:10.1001/jamafacial.2016.1620

23. Chong LSH, Eviston TJ, Low TH, Hasmat S, Coulson SE, Clark JR. Validation of the clinician-graded Electronic Facial Paralysis Assessment. *Plast Reconstr Surg*. 2017;140(1):159-167. doi:10.1097/PRS.0000000000003447

24. Gaudin RA, Robinson M, Banks CA, Baiungo J, Jowett N, Hadlock TA. Emerging vs time-tested methods of facial grading among patients with facial paralysis. *JAMA Facial Plast Surg*. 2016;18(4):251-257. doi:10.1001/jamafacial.2016.0025

Appendix 4

Facial Palsy-Specific Quality of Life in 920 Patients: Correlation with Clinician-Graded Severity and Predicting Factors

Authorship Statement

I, Joseph Dusseldorp, confirm for the ensuing journal article: **‘Facial Palsy-Specific Quality of Life in 920 Patients: Correlation with Clinician-Graded Severity and Predicting Factors’** that I have made the following contributions to its authorship:

- Analysis and interpretation of the findings
- Revising of the manuscript and critical appraisal of content



Joseph R. Dusseldorp

Date 14.09. 2023

As Auxiliary Supervisor for the candidature upon which this thesis is based, I can confirm that the above authorship attribution is correct.



Professor Tessa Hadlock, MD

Date 27.07. 2023

Department of Otolaryngology
Massachusetts Eye and Ear Infirmary
Harvard University, Boston, MA, USA

As Primary Supervisor of the candidate, I concur with the above authorship attestation:



Professor Peter Maitz

Date 14.09.2023

Department of Burns Surgery
Concord Repatriation General Hospital
Sydney University, Sydney, NSW, Australia



Facial Palsy-Specific Quality of Life in 920 Patients: Correlation With Clinician-Graded Severity and Predicting Factors

Joana Tavares-Brito, MD; Martinus M. van Veen, MD; Joseph R. Dusseldorp, MBBS, MS, FRACS;
Fayez Bahmad Jr MD, PhD; Tessa A. Hadlock, MD

Objectives: To investigate the correlation between facial palsy severity and quality of life in a broad cohort of facial palsy patients and to elucidate factors that influence this relationship.

Study Design: Retrospective study.

Methods: Records of patients presenting with a clinician-graded facial function (eFACE) and facial palsy-specific quality-of-life patient-reported outcome measure (FaCE) scale from the same moment were reviewed. Multiple linear regression was performed to study the effect of various variables on FaCE total score.

Results: A total of 920 of 1,304 patients were included, 59.9% female with a mean (standard deviation) age of 48.6 (16.7) years and a median (interquartile range) palsy duration of 9.6 [2.2; 42.2] months. A multiple linear regression model predicting FaCE total score was established, finding 10 significant variables: eFACE; viral, malignant, and congenital etiologies; overweight status; anxiety; chronic pain; previous treatment; radiotherapy; and duration of palsy ($R^2 = 0.261$, $P < 0.001$). Gender, age, laterality, surgical etiology, depression, and timing of evaluation (at initial intake or at follow up) were not found to predict FaCE total scores.

Conclusion: A correlation between facial palsy severity and quality of life was found in a large cohort of patients comprising various etiologies. Additionally, novel factors that predict quality of life in facial palsy were revealed. This information may help specialists to predict which facial palsy patients are at higher risk of a poorer quality of life, regardless of severity.

Key Words: Facial paralysis, quality of life, Facial Clinimetric Evaluation Scale, predictive factors.

Level of Evidence: 4.

Laryngoscope, 00:1–5, 2018

INTRODUCTION

Facial palsy (FP) yields myriad deficits affecting form and function. It can be a devastating and disfiguring condition, leading to psychological difficulties and restrictions of facial expressions, and may have a dramatic impact on interpersonal relationships.¹ The impact of FP on quality of life (QoL) is well known,^{2–19} although sometimes no correlation has been found between severity of FP and QoL.^{3–6,8,20} When it has been found, the correlation ranges from 0.13 to 0.66,^{2,10,12,14,15,17,21,22} meaning that other factors may be important contributors to QoL in patients with FP. Some factors, including gender, age, laterality, and duration of disease, have already been

studied. However, the results are conflicting, and it is unclear which factors influence QoL in FP patients, and in what way.^{2–4,6–9,11–14,16–18} Increasing knowledge about factors influencing QoL in FP is clinically relevant and will help the multidisciplinary team guide patient care.

The aim of this study was to investigate the correlation between severity of FP and QoL, and if found, determine the strength of this relationship in a broad cohort of patients. Additionally, we aim to analyze possible predictors of QoL after FP.

MATERIALS AND METHODS

Prior to beginning this study, approval was obtained from the Massachusetts Eye and Ear Infirmary Institutional Review Board. Since eFACE²³ was implemented in our Center (February 2014), patients had their degree of facial palsy evaluated using it. In our retrospective study, every patient who underwent assessment between February 2014 and October 2017 and who had both a QoL questionnaire and an eFACE score from the same time point were included. Patients with no FP, bilateral disease, or missing data were excluded.

Data Collection

The eFACE, a reliable and valid method of documenting facial function,^{23–26} and the Facial Clinimetric Evaluation (FaCE) scale,¹⁰ a widely adopted patient-reported outcome measure of facial palsy-related QoL, were used. The FaCE total score

From the Facial Nerve Center, Department of Otolaryngology, Massachusetts Eye and Ear Infirmary and Harvard Medical School, Boston, Massachusetts, U.S.A.(T.-B., M.M.V.-V., J.R.D., T.A.H.); Healthy Science School (J.T.-B., F.B.), University of Brasilia, Brasilia, Brazil; Department of Plastic Surgery (M.M.V.-V.), University Medical Center Groningen and University of Groningen, Groningen, the Netherlands; Department of Plastic and Reconstructive Surgery (J.R.D.), Royal Australasian College of Surgeons and University of Sydney, Sydney, Australia

Editor's Note: This Manuscript was accepted for publication on July 9, 2018.

The authors have no funding, financial relationships, or conflicts of interest to disclose.

Additional supporting information may be found in the online version of this article.

Send correspondence to Joana Tavares-Brito, Universidade de Brasilia, Faculdade de Ciencias da Saude, Campus Darcy Ribeiro, Asa Norte, Brazil 70.910-900. E-mail: joana_tavares@hotmail.com

DOI: 10.1002/lary.27481

was used as our outcome measure because we set out to study the effect of FP severity on overall FP-related QoL.

Patient charts were reviewed for variables, including gender, side of palsy, etiology, age, duration of palsy, previous treatment for FP, previous radiotherapy, and whether assessment was at initial intake or at follow-up. Medical history and medication lists were analyzed to determine if there was evidence of depression, anxiety, and/or chronic pain. Patients were classified as “not overweight,” “unsure,” or “clearly overweight” independently by two observers based on photographs of the head and neck. Scores of both observers were averaged to obtain a 5-point scale of body mass index (BMI). Our weight impression was validated in a subset of patients who did have BMI recorded in their medical chart ($n = 325$). For further analysis, only the patients that both observers considered “clearly overweight” were classified as “overweight”; all others were classified as “not overweight.” Etiologies were divided into standard groups aligned with previous studies^{13,27} and considering prognosis and gravity of the underlying condition. A dichotomous variable was created for each etiology, and patients were classified as having or not either: congenital, malignant, postsurgical, or viral etiologies. Treatments were recorded as medication, surgery, chemodeneration, or physical therapy. A dichotomous variable was created for previous treatment grouping patients who received at least one of the mentioned treatments and those that had no previous treatment.

Statistical Analysis

A simple linear regression was performed to study the relation between the eFace score and FaCE scale total score. A multiple linear regression model was used to generate a model of all predictors of QoL as measured by the FaCE total score. First, independent t tests and Pearson correlation coefficients (coef.), for nominal and continuous data, respectively, were used to determine which factors might be of influence on the FaCE total score. A typical P value less than 0.2 was used as a cutoff. Secondly, the multiple regression model was used to determine which of these variables predicts FaCE scale total score. All previously selected variables were entered one by one. Variables with P value higher than 0.1 were excluded. All statistical analyses were performed using SPSS version 24 (IBM Corp., Armonk, NY).

RESULTS

Of 3,032 eFACE assessments for 1,304 different patients between February 2014 and October 2017, 975 patients had a corresponding FaCE scale from the same time point. We excluded 55 patients with no clinically apparent FP, bilateral FP, or missing data; the other 920 were included in our analysis. Patients excluded (29%) had their demographic characteristics compared to the patients included, and they were not different with regard to age, gender, duration of palsy, and etiology. A slight majority of included patients were female (59.5%); mean age was 48.6 years (standard deviation [SD] 16.7); and median duration of FP was 9.6 months (interquartile range [IQR] 2.2; 42.2). Affected sides were evenly distributed. The most common etiology was Bell palsy (40.8%), followed by acoustic neuroma (10.2%), Varicella zoster (9.8%), and head and neck cancer (8.0%), as determined by the treating physician. Other etiologies included trauma, benign tumors, iatrogenic injuries, central nervous system lesion, Lyme disease,

congenital, and others. Six hundred ninety-six patients (75.7%) had received previous treatment at the time of evaluation. Data were obtained at the initial intake in 847 patients (95.1%). The degree of facial function measured by eFACE was 71.3 on average (SD 14.6), and the mean FaCE scale total score was 47.6 (SD 20.1) (Table I).

Weight impression assessment was done by two independent investigators in three categories. A quadratic weighted kappa for interrater agreement was 0.66 (thus good). The weight impression scale was plotted against BMI values for a subset of 325 patients for whom

TABLE I.
Demographic Description of 920 Patients

Gender (n (%))	
Male	373 (40.5)
Female	547 (59.5)
Age, years (mean (SD))	48.6 (16.7)
Duration of palsy, months (median (IQR))	9.6 (2.2;42.2)
Side (n (%))	
Right	474(51.5)
Left	446(48.5)
Etiology (n (%))	
Bell palsy	375 (40.8)
Pregnancy-associated	35
Recurrent	20
Acoustic neuroma	94 (10.2)
Varicella zoster	90 (9.8)
Head and neck cancer	74 (8.0)
Postresection	57
Trauma	52 (5.7)
Soft tissue trauma	23
Temporal bone fracture	23
Benign tumor	49 (5.3)
Facial nerve tumor	26
Iatrogenic injury	40 (4.3)
CNS lesion	37 (4.0)
Lyme	34 (3.7)
Congenital	17 (1.8)
Otologic disease	9 (1.0)
Stroke	9 (1.0)
Other	8 (0.9)
Unclear	32 (3.5)
Previous treatment	
Medication	517 (56.2)
Surgery	160 (17.4)
Chemodeneration	76 (8.3)
Physical therapy	162 (17.6)
Timing of evaluation	
Intake	847 (92.1)
Follow-up	73 (7.9)
eFACE (mean (SD))	71.3(14.6)
FaCE scale total score (mean (SD))	47.6 (20.1)

eFACE = clinician-graded facial function scale; FaCE = Facial Clinical Metric Evaluation; IQR = interquartile range; SD = standard deviation.

a BMI was available. A concurrent increase in BMI with the weight impression scale was seen and found to be significant (Spearman rho 0.677, $P < .001$) (Appendix I).

Bivariate analyses were used to select possible predictors of the FaCE total score. The eFACE score was positively correlated to QoL ($R = 0.434$, $P < 0.001$) (Fig. 1). Correlations between age and duration of palsy were -0.088 ($P = 0.007$) and 0.097 ($P = 0.003$), respectively, indicating that higher age was associated with a lower QoL and that a longer duration was associated with a higher QoL. The dichotomous variables “viral etiology,” “congenital etiology,” “malignant etiology,” “overweight status,” “anxiety,” “chronic pain,” “previous treatment,” “radiotherapy,” and “timing of evaluation” were found to be possible predictors of FaCE total score (Table II).

Possible FaCE total score predictors of univariate analyses were used to establish multiple linear regression models. Two variables with P values > 0.1 were excluded from the model in the following order: age ($P = 0.920$) and timing of evaluation ($P = 0.460$). In the final model, the severity of FP measured by eFACE was associated with FaCE total score (coef. = 0.60, $P < 0.001$). Viral etiology was associated with lower QoL (coef. = -3.57 , $P = 0.008$), whereas malignant and congenital etiologies were associated with higher QoL (coef. = 4.18 and 10.23; $P = 0.096$ and 0.027, respectively). Overweight status, anxiety, and chronic pain were predictors of lower QoL (coef. = -7.61 ; -7.32 and -3.07 ; $P = < 0.001$, < 0.001 , and 0.098, respectively). Longer duration FP was associated with a higher QoL (coef. = 0.01, $P = 0.030$). Having been previously treated or having undergone radiation therapy were associated with lower QoL (coef. -2.65 and -7.64 ; $P = 0.073$ and 0.011, respectively). The explained variance in FaCE total score of the multiple regression model was 26.1% ($R^2 = 0.261$) (Table III). Gender, age, laterality, surgical etiology, depression, and timing of evaluation were found not to be predictors of FaCE total scores.

DISCUSSION

This study demonstrates a relationship between the degree of FP and QoL in a large cohort of FP patients with a wide variety of etiologies. Among the 10 variables

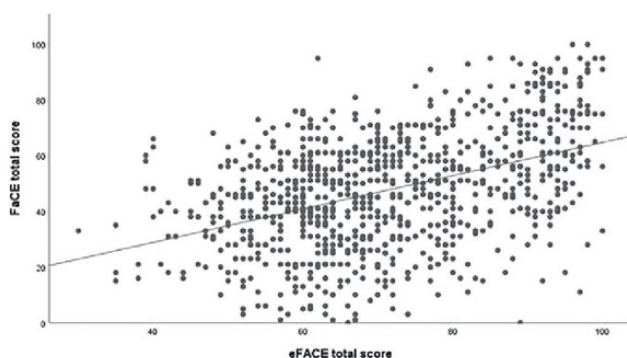


Fig. 1. Scatterplot of eFACE scores and FaCE scale total scores. Line represents the regression line ($R^2 = 0.189$, $P < 0.001$). eFACE = clinician-graded facial function scale; FaCE = Facial Clinimetric Evaluation.

Laryngoscope

TABLE II.
Initial Dichotomous Variables Analysis.

	FaCE total score (mean (SD))	P Value
Gender		.594
Male	48.0(20.1)	
Female	47.3(20.1)	
Side		.711
Left	47.8(19.6)	
Right	47.3(20.6)	
Viral etiology		.002
Yes	45.6(20.4)	
No	49.6(19.6)	
Congenital etiology		< .001
Yes	65.4(10.9)	
No	47.2(20.1)	
Malignant etiology		.138
Yes	44.4(19.9)	
No	47.9(20.1)	
Postsurgical etiology		.553
Yes	46.9(19.7)	
No	47.8(20.3)	
Overweight status		< .001
Yes	38.8(19.5)	
No	48.7(19.9)	
Depression		.232
Yes	45.9(20.1)	
No	48.0(20.1)	
Anxiety		.001
Yes	41.0 (21.0)	
No	48.3(19.9)	
Chronic pain		.004
Yes	42.4(19.1)	
No	48.3(20.2)	
Previous treatment		.001
Yes	46.3 (19.5)	
No	51.6 (21.5)	
Radiotherapy		.002
Yes	39.1 (20.4)	
No	48.1 (20.0)	
Timing of evaluation		.002
Intake	47.0 (20.3)	
Follow-up	53.6 (16.7)	

Possible predictors of FaCE total score are presented in bold.
FaCE = Facial Clinimetric Evaluation; SD = standard deviation.

found to predict FaCE total score, of all 16 variables studied, severity of FP was the most important factor, explaining 18.9% out of the total 26.1% of the FaCE score variance predicted by eFACE. Similar to previous studies,^{2,12,17,21,22} the correlation found in our data between severity of FP and patient-reported outcome (PROM) was moderate. A systematic review examining the impact of disease-related impairments on health-related QoL found that the association between impairment and QoL in patients with various disorders usually

Tavares-Brito et al.: Quality of Life in Facial Palsy

TABLE III.
Results of Stepwise Multiple Regression Model Showing Predictive Factors of FaCE Scale Score

Predictor	Coefficient	P Value
eFACE	0.60	< .001
Etiology viral	-3.57	.008
Etiology malignant	4.18	.096
Etiology congenital	10.23	.027
Overweight status	-7.61	< .001
Anxiety	-7.32	< .001
Chronic pain	-3.07	.098
Previous treatment	-2.65	.073
Radiotherapy	-7.64	.011
Duration of FP	0.01	.045

eFACE = clinician-graded facial function scale; FaCE = Facial Clinical Evaluation; FP = Facial Palsy.

Appendix I. Weight impression and mean BMI scores for each category. Weight classification and BMI were found to be significantly correlated (Spearman rho 0.677, $p < .001$). BMI = Body Mass Index.

has a correlation coefficient of less than 0.50.²⁸ They concluded that QoL scores reflect the patient's perception of the consequences of disease and depend on numerous additional, usually psychosocial, factors other than the disease itself.

There is relatively scant published information about FP disability as it relates to etiology. Most authors who studied it found that etiology was not a predicting factor of PROMs.^{2,16,17} One previous study compared the morbidity of facial nerve dysfunction arising from surgical intervention for VS ($n = 53$) with those resulting from Bell palsy ($n = 22$). Investigators found that patients with facial nerve dysfunction arising from VS surgery experienced less morbidity than those with FP caused by Bell palsy.¹³ However, we found that viral etiology predicted lower QoL, whereas surgical etiology was found not to be a predictor of QoL. Furthermore, in our data, malignant and congenital etiologies predicted higher QoL. Those interesting findings are novel in the FP scenario and invite further investigations.

Numerous studies in various populations have shown that obesity adversely affects QoL.²⁹⁻³¹ In this study, patients with FP were analyzed for overweight status and its impact on a specific health-related QoL. Overweight status was a predicting factor of lower QoL, suggesting that FP patients who are obese are more likely to experience impaired QoL than patients with normal weight. This hypothesis was generated based on our clinical experience and requires prospective formal analysis to be confirmed.

Longer duration of palsy was associated with a higher QoL. This may suggest that patients learn to adapt to their facial impairment over time, leading to improved QoL, as occurs in other diseases.^{32,33} Although this association may seem obvious, the impact of duration on FaCE total scores was small, with a coefficient of 0.01. Other authors who studied the effect of length of time after palsy on QoL found that duration was not a significant factor.^{4,5,12,13}

Our study represents the largest series describing QoL in patients with FP from a wide variety of etiologies. This analysis reinforces the relevance of the degree of paralysis and its correlation to health-related QoL and provides insight into new factors influencing it. The fact that viral etiology and malignant etiology were predictive for QoL is novel. Congenital etiology and overweight status have already been investigated in patients with other diseases,³⁴⁻³⁶ but their impact on QoL in FP is herewith studied for the first time.

Retrospective data collection is the main limitation of the present study. The information contained in charts was not acquired for research purpose and thus may lack precision. We used medication lists to supplement the medical charts on evidence for anxiety, depression, and chronic pain. The presence of overweight status was defined by an observer rating of patients' photographs. Both methods give an approximation of the variable studied but are inherently inaccurate. Our findings based on these assumptions should be interpreted with caution and help to generate hypothesis to study formally rather than direct treatment.

CONCLUSION

The degree of FP is the main predicting factor of health-related QoL. Duration of palsy, malignant and congenital etiologies (associated with higher QoL), as well as viral etiology, overweight status, anxiety, chronic pain, previous treatment, and radiotherapy (associated with lower QoL) also predict the FaCE total score to varying degrees. Understanding how these factors may predict QoL outcomes in patients with FP can assist physicians to minimize the impact of FP by optimizing assessment and clinical decision-making processes. Early psychological counseling may be a potential strategy targeted at those at risk for poorer QoL outcomes, decreasing the morbidity of FP for those who need it the most.

BIBLIOGRAPHY

1. Ishii LE, Godoy A, Encarnacion CO, Byrne PJ, Boahene KD, Ishii M. What faces reveal: impaired affect display in facial paralysis. *Laryngoscope* 2011;121:1138-1143.
2. Kleiss LJ, Hohman MH, Susarla SM, Marres HA, Hadlock TA. Health-related quality of life in 794 patients with a peripheral facial palsy using the FaCE scale: A retrospective cohort study. *Clin Otolaryngol* 2015;40:651-656.
3. Baumann I, Polligkeit J, Blumenstock G, Mauz PS, Zalaman IM, Maassen MM. Quality of life after unilateral acoustic neuroma surgery via middle cranial fossa approach. *Acta Otolaryngol* 2005;125:585-591.
4. Cross T, Sheard CE, Garrud P, Nikolopoulos TP, O'Donoghue GM. Impact of facial paralysis on patients with acoustic neuroma. *Laryngoscope* 2000;110:1539-1542.
5. Lassaletta L, Alfonso C, Del Rio L, Roda JM, Gavilan J. Impact of facial dysfunction on quality of life after vestibular schwannoma surgery. *Ann Otol Rhinol Laryngol* 2006;115:694-698.
6. Lee J, Fung K, Lowrie SP, Parnes LS. Assessing impairment and disability of facial paralysis in patients with vestibular schwannoma. *Arch Otolaryngol Head Neck Surg* 2007;133:56-60.
7. Nikolopoulos TP, Johnson I, O'Donoghue GM. Quality of life after acoustic neuroma surgery. *Laryngoscope*. 1998;108:1382-1385.
8. Fu L, Bundy C, Sadiq SA. Psychological distress in people with disfigurement from facial palsy. *Eye (Lond)* 2011;25:1322-1326.
9. da Cruz MJ, Moffat DA, Hardy DG. Postoperative quality of life in vestibular schwannoma patients measured by the SF36 health questionnaire. *Laryngoscope* 2000;110:151-155.

10. Kahn JB, Gliklich RE, Boyev KP, Stewart MG, Metson RB, McKenna MJ. Validation of a patient-graded instrument for facial nerve paralysis: The FaCE scale. *Laryngoscope* 2001;111:387–398.
11. Ryzenman JM, Pensak ML, Tew JM. Facial paralysis and surgical rehabilitation: A quality of life analysis in a cohort of 1,595 patients after acoustic neuroma surgery. *Otol Neurotol* 2005;26:516–521; discussion 521.
12. Nicoucar K, Momjian S, Vader JP, De Tribolet N. Surgery for large vestibular schwannomas: How patients and surgeons perceive quality of life. *J Neurosurg* 2006;105:205–212.
13. Saito DM, Cheung SW. A comparison of facial nerve disability between patients with Bell's palsy and vestibular schwannoma. *J Clin Neurosci* 2010;17:1122–1125.
14. Ng JH, Ngo RY. The use of the facial clinimetric evaluation scale as a patient-based grading system in Bell's palsy. *Laryngoscope* 2013;123:1256–1260.
15. Diaz-Aristizabal U, Valdes-Vilches M, Fernandez-Ferreras TR, Calero-Munoz E, Bienzobas-Allue E, Moracen-Naranjo T. Correlations between impairment, psychological distress, disability, and quality of life in peripheral facial palsy. *Neurologia* 2017. doi: 10.1016/j.nrl.2017.03.004.
16. Nellis JC, Ishii M, Byrne PJ, Boahene KDO, Dey JK, Ishii LE. Association among facial paralysis, depression, and quality of life in facial plastic surgery patients. *JAMA Facial Plast Surg* 2017;19:190–196.
17. Volk GF, Granitzka T, Kreysa H, Klingner CM, Guntinas-Lichius O. Initial severity of motor and non-motor disabilities in patients with facial palsy: An assessment using patient-reported outcome measures. *Eur Arch Otorhinolaryngol*. 2017;274:45–52.
18. Starmer HM, Ward BK, Best SR, et al. Patient-perceived long-term communication and swallow function following cerebellopontine angle surgery. *Laryngoscope* 2014;124:476–480.
19. Volk GF, Granitzka T, Kreysa H, Klingner CM, Guntinas-Lichius O. Non-motor disabilities in patients with facial palsy measured by patient-reported outcome measures. *Laryngoscope* 2016;126:1516–1523.
20. Guntinas-Lichius O, Straesser A, Streppel M. Quality of life after facial nerve repair. *Laryngoscope* 2007;117:421–426.
21. Volk GF, Steigerwald F, Vitek P, Finkensieper M, Kreysa H, Guntinas-Lichius O. Facial disability index and facial clinimetric evaluation scale: Validation of the German versions. *Laryngorhinootologie* 2015;94:163–168.
22. Gonzalez-Cardero E, Infante-Cossio P, Cayuela A, Acosta-Feria M, Gutierrez-Perez JL. Facial disability index (FDI): adaptation to Spanish, reliability and validity. *Med Oral Patol Oral Cir Bucal*. 2012;17:1006.
23. Banks CA, Bhama PK, Park J, Hadlock CR, Hadlock TA. Clinician-graded electronic facial paralysis assessment: the eFACE. *Plast Reconstr Surg* 2015;136:230e.
24. Gaudin RA, Robinson M, Banks CA, Baiungo J, Jowett N, Hadlock TA. Emerging vs time-tested methods of facial grading among patients with facial paralysis. *JAMA Facial Plast Surg* 2016;18:251–257.
25. Banks CA, Jowett N, Hadlock TA. Test-retest reliability and agreement between in-person and video assessment of facial mimetic function using the eFACE facial grading system. *JAMA Facial Plast Surg* 2017;19:206–211.
26. Banks CA, Jowett N, Azizzadeh B, et al. Worldwide testing of the eFACE facial nerve clinician-graded scale. *Plast Reconstr Surg* 2017;139:498e.
27. van den Elzen, Marijke EP, Versnel SL, Hovius SER, Passchier J, Duivenvoorden HJ, Mathijssen IMJ. Adults with congenital or acquired facial disfigurement: impact of appearance on social functioning. *J Cranio-maxillofac Surg* 2012;40:777–782.
28. Weisscher N, de Haan RJ, Vermeulen M. The impact of disease-related impairments on disability and health-related quality of life: a systematic review. *BMC Med Res Methodol* 2007;7:24.
29. Jia H, Lubetkin EI. The impact of obesity on health-related quality-of-life in the general adult US population. *J Public Health (Oxf)*. 2005;27:156–164.
30. Xu Y, Zhou Z, Li Y, et al. Exploring the nonlinear relationship between body mass index and health-related quality of life among adults: a cross-sectional study in Shaanxi Province, China. *Health Qual Life Outcomes* 2015;13:9.
31. Wee HL, Wu Y, Thumboo J, Lee J, Tai ES. Association of body mass index with short-form 36 physical and mental component summary scores in a multiethnic Asian population. *Int J Obes (Lond)* 2010;34:1034–1043.
32. Horgan O, MacLachlan M. Psychosocial adjustment to lower-limb amputation: a review. *Disabil Rehabil*. 2005;26:837–850.
33. Sojevic Timotijevic Z, Majcan P, Trajkovic G, et al. The impact of changes in psoriasis area and severity index by body region on quality of life in patients with psoriasis. *Acta Dermatovenereol Croat* 2017;25:215–222.
34. Garcia-Poma A, Segami MI, Mora CS, et al. Obesity is independently associated with impaired quality of life in patients with rheumatoid arthritis. *Clin Rheumatol* 2007;26:1831–1835.
35. Bogart KR, Tickle-Degnen L, Ambady N. Compensatory expressive behavior for facial paralysis: adaptation to congenital or acquired disability. *Rehabil Psychol* 2012;57:43–51.
36. Robinson E, Rumsey N, Partridge J. An evaluation of the impact of social interaction skills training for facially disfigured people. *Br J Plast Surg* 1996;49:281–289.

CHAPTER FOUR

Nasolabial Fold Dynamics: Implications for Facial Paralysis and Facial Reanimation Surgery

Preface

“... By the drawing backwards and upwards of the corners of the mouth, through the contraction of the great zygomatic muscles, and by the raising of the upper lip, the cheeks are drawn upwards. Wrinkles are thus formed under the eyes, and, with old people, at their outer ends; and these are highly characteristic of laughter or smiling ...”

Charles Darwin *The Expression of Emotions in Man and Animals*. 1872.

Surgical recreation of smiling requires a precise understanding of what movements the smile muscles usually create and how neural and muscular interventions can and cannot restore those movements. The key muscles responsible for oral commissure excursion are the zygomaticus major and minor muscles which arise from the zygomatic bone and insert into a condensation of connective tissue in the cheek just adjacent to the corner of the mouth, known as the modiolus. There are also muscular fibres that insert into the ala of the nose and the dermis between the ala and the modiolus. With smiling, these additional fibres form a crease line known as the nasolabial fold. This fold is absent at rest except in those of advanced age and only usually visible with strong smiling, such as with laughter.

The nasolabial fold holds a paramount importance to the aesthetic appearance of normal smiling as evidenced by its inclusion in multiple items of the eFACE assessment tool yet remains an elusive characteristic to measure in an objective way because it is absent at rest and varies considerably in its morphology between patients. The aim of this chapter was to use state of the art facial landmark recognition tools to quantify nasolabial fold during smiling and thus define nasolabial fold dynamics in unilateral facial palsy by comparing the affected side to the healthy side. The chapter was published in *Orthodontic and Craniofacial Research*.

Dusseldorp JR, Faraway J, Razavi L, Hadlock TA, Trotman CA. Nasolabial Fold Dynamics: Implications for Facial Paralysis and Facial Reanimation Surgery. *Orthod. Craniofac. Res.* 2020. 00:1–8.

In the appendices presented, an algorithm for automated identification of facial landmarks in facial palsy patients was developed and used to quantify facial displacements both manually, and subsequently in an automated fashion, leading to an automated facial symmetry assessment tool in facial palsy, known as Emotrics. This tool is now routinely being employed in the field of facial reanimation and the paper has received 126 citations since publication.

Appendix 5:

Guarin D, **Dusseldorp JR**, Hadlock TA, Jowett N. A Machine Learning Approach for Automated Facial Measurements in Facial Palsy. *JAMA Facial Plast Surg.* 2018. Jul 1;20(4):335-337

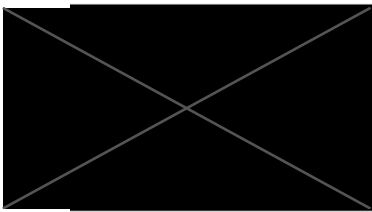
Appendix 6:

Guarin DL, Yunusova Y, Taati B, **Dusseldorp JR**, Mohan S, Tavares J, van Veen MM, Fortier E, Hadlock TA, Jowett N. Toward an Automatic System for Computer-Aided Assessment in Facial Palsy. *Facial Plast Surg Aesthet Med.* 2020 Jan;22(1):42-49

Authorship Statement

I, Joseph Dusseldorp, confirm for the ensuing journal article: '**Nasolabial Fold Dynamics: Implications for Facial Paralysis and Facial Reanimation Surgery**' that I have made the following contributions to its authorship:

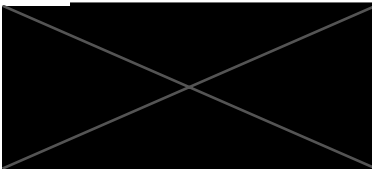
- Conception and design of the research
- Analysis and interpretation of the findings
- Drafting and revising of the manuscript and critical appraisal of content
- Responding to multiple stages of peer review



Joseph R. Dusseldorp

Date 14.09. 2023

As Auxiliary Supervisor for the candidature upon which this thesis is based, I can confirm that the above authorship attribution is correct.

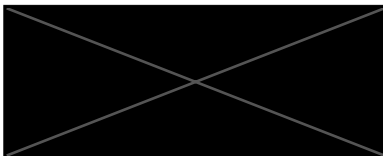


Professor Tessa Hadlock, MD

Date 27.07. 2023

Department of Otolaryngology
Massachusetts Eye and Ear Infirmary
Harvard University, Boston, MA, USA

As Primary Supervisor of the candidate, I concur with the above authorship attestation:



Professor Peter Maitz

Date 14.09.2023

Department of Burns Surgery
Concord Repatriation General Hospital
Sydney University, Sydney, NSW, Australia

Nasolabial fold dynamics: Implications for facial paralysis and facial reanimation surgery

Joseph R. Dusseldorp¹ | Julian Faraway² | Lily Razavi³ | Tessa A. Hadlock⁴ |
Carroll Ann Trotman⁵ 

¹Plastic and Reconstructive Surgery,
Strathfield Private Hospital Medical Centre,
3-5 Everton Road, Strathfield, NSW,
Australia

²Department of Mathematical Sciences,
University of Bath, Bath, UK

³DMD Candidate D'21, Tufts University
School of Dental Medicine, Boston, MA,
USA

⁴Facial Nerve Center, Massachusetts Eye
and Ear Infirmary, Boston, MA, USA

⁵Department of Orthodontics, Tufts
University School of Dental Medicine,
Boston, MA, USA

Correspondence

Carroll Ann Trotman, Department of
Orthodontics, Tufts University School of
Dental Medicine, 1 Kneeland Street, Boston,
MA 02111.
Email: carroll_ann.trotman@tufts.edu

Funding information

This article was funded by NIDCR Grant
DE025295 awarded to Dr Carroll Ann
Trotman as Principal Investigator. None of
the authors have a financial interest in any of
the products, devices or drugs mentioned in
this manuscript.

Abstract

Objectives: In patients with facial paralysis, facial reanimation surgery may be needed to normalize facial soft tissue function/movements. Critical for this normalization is the dynamics of the nasolabial folds (NLFs). The objective of this prospective, observational study was to determine the 3D morphologic dynamics of the NLFs in patients with unilateral facial palsy and normal subjects.

Settings and Sample Population: 3D facial soft tissue movement data collected from adults with unilateral, facial paralysis (Bell's Palsy, $n = 36$); and (2) an age- and sex-frequency matched control group ($n = 68$).

Materials and methods: Movement data were collected during repeated animations from participants using a video-based motion capture system. Movement in terms of displacement and asymmetry of the NLFs, nasal and circumoral regions were analyzed in the lateral, vertical and depth planes; as well as movement of the commissure and NLFs relative to the lower lip midline. Two-sample t tests were used to test for significant group differences.

Results: Patients NLFs had less mean displacement, greater mean asymmetry and uncoordinated movements compared with the controls. For both groups during smiling, the NLF and commissure landmarks had approximately similar magnitudes of displacement (control range = 11-14mm; patient range = 7-10mm).

Conclusion: NLF dynamics during smiling were as significant as oral commissure excursion. Thus, an immobile NLF is an unnatural feature of facial animations. Surgical treatments that address impaired NFL movements must be considered to create a more natural surgical outcome especially during smiling.

KEYWORDS

Bell's palsy, facial paralysis, nasolabial fold, smile reanimation

Presented at: This work has been presented at AADR/IADR meeting in Vancouver, BC, Canada.

© 2020 John Wiley & Sons A/S. Published by John Wiley & Sons Ltd

1 | INTRODUCTION

Facial palsy is a disabling and disfiguring condition that results in impaired facial soft tissue function/movement. The impact is profound. Afflicted patients consistently report negative psychosocial sequelae and functioning.¹ Many patients experience permanent disfigurement either flaccid weakness or non-flaccid hypertonicity and may require reanimation surgery to improve facial function and expression. Critical to successful surgical outcomes is a thorough understanding of the soft tissue muscle anatomy during facial expression movements. For example, an obvious and important factor to surgically recreate for a smile movement is oral commissure excursion, the degree to which the corner of the mouth moves from rest. In normal subjects, this movement ranges from 7 to 22 mm.^{1,2} Equally important for normal esthetics is the *dynamics* of the nasolabial fold (NLF), the natural crease between the nose and mouth (Figure 1A). The orientation, depth and direction of the NLF movement during smiling are a major predictor of paralysis severity and patients' quality of life (QOL).³ Anatomically, the NLFs are facial crease lines or rhytids made visible by shadow formation at the junctions of the upper lip, cheek and nasal alar on either side of the face.⁴⁻⁶ With age, they become more pronounced. An important aspect of reanimation surgery is normalization of facial form and dynamics, of which the NLFs are a critical component; however, NLF dynamics remain poorly understood. Recently, we developed a suite of objective three-dimensional (3D) measures to dynamically characterize and map facial soft tissue movements for diagnosis and assessment of treatment outcomes in patients with facial paralysis.^{7,8} In this study, the objective was to use these measures to determine the 3D morphologic dynamics of the NLFs during different facial expressions in normal subjects and patients with unilateral facial palsy. Emphasis was placed on understanding the NLF differences between patients and controls when the patients first presented for treatment.

2 | MATERIALS AND METHODS

The study participants were recruited as part of a prospective study (NIDCR Grant # DE025295) to determine changes in facial soft tissue movements in adults with unilateral facial paralysis and to compare the patients' facial movements with control participants.^{7,8} Consecutive participants were recruited between June 2016 and March 2018 on first presenting for treatment at the Facial Nerve Center at Massachusetts Eye and Ear Infirmary (MEEI) and the controls were patients and

volunteers from Tufts University School of Dental Medicine (TUSDM) who were invited to participate either by personal contact or by responding to a posted flyer/advertisement. Eligible participants were screened by telephone based on selection criteria (Table S1). Those who agreed to participate attended TUSDM Facial Animation laboratory where they were consented and data collection was completed. Based on previously published power calculations,⁸ the final study sample consisted of two groups: patients ($n = 36$; mean age = 43.3 years ± 12.9) with acute, unilateral and flaccid facial paralysis (Bell's Palsy); and an age- and sex-frequency matched 'normal' control group ($n = 68$; mean age = 42.7 years ± 14.5). The patient group consisted of 18 males and 18 females. Five patients self-identified as Hispanic, one as Black, two as Asian and 28 as Caucasian. The control group consisted of 34 males and 34 females. Four of the control participants self-identified as Hispanic, five as Black, six as Asian and 53 as Caucasian. Study consent and HIPAA documents were approved by the Tufts Health Sciences IRB. All participants had facial movement data collected.

2.1 | Data collection

The methods and analyses for the collection of the facial movement data specific for patients with facial paralysis were described in previous publications,^{7,8} and a brief summary is provided here. A video-based motion tracking system (Motion Analysis™, Motion Analysis Corporation; Figure 1B and C) was used to record the movement during eleven facial animations of 64 retro-reflective markers placed on specific facial soft tissue landmarks (Figure 1B—facial landmarks in gray). The animations were gentle eye closure (gec), natural smile (nsm) and the following maximum movements—brow raise (br), tight eye closure (tec), 'ee' sound (ee), 'oo' sound (oo), smile (msm), grimace (gr), lip purse (lp), cheek puff (cp) and mouth opening (mo). Each animation was repeated ten times by the subjects. The raw data were tracked off-line, and the tracked data for each landmark consisted of a time series of 3-D vectors defined by x, y and z where x, y and z represented the position in space at 60 frames per second (60 Hz) for 4 seconds.

2.2 | Data analysis

For each repeated movement of an animation, the average 'at-rest' facial landmark configuration of each participant was computed by extracting the first 'at-rest' frames and computing the Procrustes mean.

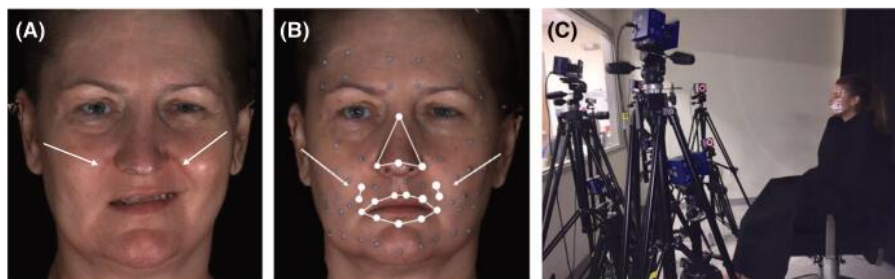


FIGURE 1 A-C, Arrows point to the nasolabial fold areas (A and B), the facial landmarks (B) and the video-based motion tracking system (C)

Then, the mean at-rest facial landmark configuration for the control participants was computed. This mean control face was symmetrized and rotated into a level and upright position and served as a standardized template upon which the at-rest faces of each participant were rotated onto using Procrustes conformation which also controls for facial size. Then, for each animation, each frame of the movement was rotated on to the corresponding at-rest frame for that participant, and the frame at which the greatest distance from rest occurred was selected. The latter was translated so that the landmark on the bridge of the nose coincided with the rest frame. The nose bridge landmark was considered a fixed point during the movements. Further analysis focused on two landmarks on each side of face that localized the right and left NLFs and landmarks on the nose and around the mouth (Figure 1B—facial landmarks highlighted in white).

Three different comparisons of displacement were made of the right and left NLF landmarks of the patients and controls. For these calculations, the data were reflected so that the left side of the patients' face was always the affected (paralyzed) side. Also, given that there were ten replicates for each participant/animation combination, this was reduced to a single measure by taking the median.

1. **NLF Displacement.** For the control participants, the mean displacement of the left and right pairs of NLF landmarks was computed; and for the patients, the mean displacement of the NLF landmark pairs for the affected or paralyzed left side of the face and the contralateral right side of the face was computed.
2. **NLF Asymmetry.** For each control and patient participant, the mean absolute difference in displacement between the two NLF landmarks on the left and the corresponding two landmarks on the right sides of the face was calculated.
3. **Lateral, Vertical and Depth (Antero-Posterior) NLF Displacement.** For the right and left landmark pairs located at the upper and lower ends of the NLFs, respectively, the distance (mm) travelled from rest to maximum displacement was calculated for each animation. For the patients and controls, further analyses were based on separate plots of the mean displacement in the lateral, vertical and depth (antero-posterior) facial planes of space for upper and lower NLF landmark pairs.

In addition, to determine the displacement of the NLFs relative to the mouth corners or commissures, the distances at rest and at the maximum of the smile animation were calculated from the landmark on the midpoint of the lower lip to the right and left commissure landmarks, the right and left upper NLF landmarks, and the right and left lower NLF landmarks. The midpoint of the lower lip landmark was considered relatively stable during smiling.

2.3 | Statistical analysis

A 2-sample *t* test was used to test for significant differences in displacement and asymmetry between the patients and control participants.

3 | RESULTS

The results for mean displacement and asymmetry in the patient and control groups (Tables 1 and 2) show that, as might be expected, there was less displacement and greater asymmetry of the NLFs for the patients especially for the maximum smile (msm), natural smile (nsm) and grimace (gr) animations.

3.1 | Lateral Movement of the NLF

Figure 2 is a plot of the lateral displacement of the control subjects' right and left landmark pairs at the upper end of the NLF for each animation. The horizontal axis shows the displacements of the landmark on the left side of the face where values to the left of the origin ('0') indicate displacement to the left of the face (displacement outwards). The vertical axis shows the displacements of the landmark on the right side where values above the origin indicate movement to the right side of the face (displacement outwards). Points close to and on the origin would indicate little or no displacement as seen for brow raise (br) and gentle eye closure (gec). In general, the plots show a strong negative correlation which is expected since normal displacement would be symmetric (outwards on either side of the face). For the smile animations (nsm & msm) and 'eeee' sound (ee), both landmarks are in the upper left quadrant because they both move outwards while the landmarks for lip purse (lp) animation and 'oooo' sound (oo) are in the lower right quadrant because they move inwards. Figure 3 shows the same plot for the patients. As can be seen, the strength of the correlation is much reduced.

3.2 | Vertical movement of the NLF

The vertical displacement of upper NLF landmark pairs for the control subjects is shown in Figure S1. The correlations are strong and positive because both landmark pairs move in the same direction—either upwards or downwards. For the patients (Figure S2), similar plots show a weaker correlation but stronger overall than the plots for the lateral displacement.

3.3 | Depth (Antero-Posterior) movement of the NLF

The depth displacement for the control subjects upper NLF landmark pairs is shown in Figure S3. Again, the correlations are strongly positive. For some animations, the landmark pairs move mostly forward such as for lip purse (lp) and cheek puff (cp) and 'oo' sound (oo). For other animations, the landmark pairs move mostly backwards, for example, both smiles (msm & nsm) and the 'ee' sound (ee). The plots for the patients (Figure S4) show a weak correlation. Similar lateral, vertical, and depth plots for the landmark pairs at the lower

Animations	FP Mean (mm)	NS Mean (mm)	t Test statistic	DF	LCI	UCI	P value
br	2.92	5.10	-5.53	66.24	-2.97	-1.39	.000***
cp	14.54	18.08	-2.79	62.37	-6.07	-1.01	.007**
ee	7.27	17.43	-8.28	70.27	-12.62	-7.72	.000***
gec	1.44	1.61	-0.67	40.71	-0.70	0.35	.509n
gr	11.02	22.09	-8.11	60.65	-13.80	-8.34	.000***
lp	8.79	15.54	-7.32	67.77	-8.59	-4.91	.000***
mo	17.03	15.40	1.38	50.68	-0.74	3.99	.174n
msm	9.31	22.32	-9.94	51.39	-15.63	-10.38	.000***
nsm	10.77	22.97	-8.78	41.99	-15.01	-9.39	.000***
oo	7.32	13.21	-5.97	71.07	-7.85	-3.92	.000***
tec	8.64	10.25	-1.00	49.29	-4.83	1.62	.322n

Note: The displacement for the control group was calculated as the mean of the left and right landmark pair maximum displacement from rest; and for the patients as the affected side maximum displacement from rest. t Tests for significant differences between the patient and control group means. DF is degrees of freedom. LCI and UCI are the 95% confidence intervals for the difference.

Abbreviations: br, brow raise; ee, eeee sound; gec, gentle eye closure; gr, maximum grimace; lp, maximum lip purse; mcp, maximum cheek puff; mo, maximum mouth opening; msm, maximum smile; nsm, natural smile; oo, oooh sound; tec, tight eye closure.

Significance levels: n = non-significant;

* $P < .05$,

** $P < .01$,

*** $P < .001$.

TABLE 1 Mean displacement (mm) of the NLF landmark pairs for the patient (FP) and control (NS) groups for each animation

Animations	FP Mean	NS Mean	t Test statistic	DF	LCI	UCI	P value
br	0.94	0.56	2.82	44.54	0.11	0.65	.007**
cp	2.83	1.24	3.44	32.91	0.65	2.54	.002**
ee	6.83	1.61	5.61	30.01	3.32	7.11	.000***
gec	0.49	0.42	0.65	60.99	-0.15	0.30	.516n
gr	8.52	1.91	6.56	28.16	4.55	8.67	.000***
lp	4.17	1.50	4.89	33.76	1.56	3.78	.000***
mo	2.52	1.50	2.04	34.63	0.00	2.04	.049*
msm	10.04	1.98	7.77	29.50	5.94	10.18	.000***
nsm	8.74	1.80	6.56	28.42	4.78	9.11	.000***
oo	3.65	1.46	3.93	31.72	1.05	3.32	.000***
tec	2.45	1.76	1.37	36.73	-0.33	1.70	.179n

Note: The asymmetry for the patient and control groups was calculated as the mean absolute difference in displacement from rest between the landmark pairs on the left and right sides of the face. t tests for significant differences between the patient and control group means. DF is degrees of freedom. LCI and UCI are the 95% confidence intervals for the difference

Abbreviations: br, brow raise; ee, eeee sound; gec, gentle eye closure; gr, maximum grimace; lp, maximum lip purse; mcp, maximum cheek puff; mo, maximum mouth opening; msm, maximum smile; nsm, natural smile; oo, oooh sound; tec, tight eye closure.

Significance levels: n = non-significant;

* $P < .05$;

** $P < .01$,

*** $P < .001$.

TABLE 2 Mean asymmetry of the NLF landmark pairs for the patient (FP) and control (NS) groups for each animation

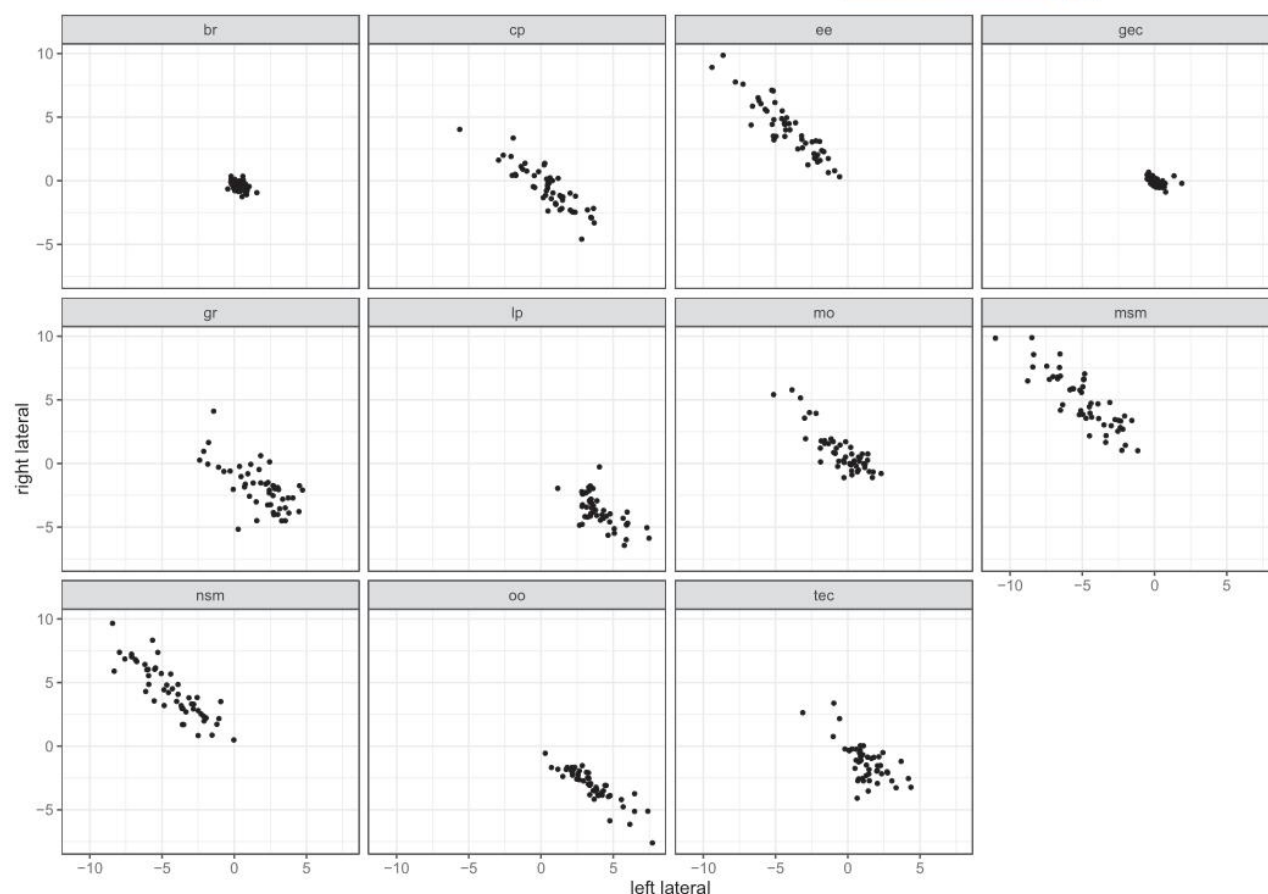


FIGURE 2 Lateral displacement of the control subjects right and left landmark pairs at the upper end of the NLFs for each animation

end of the NLFs demonstrated similar correlations and these plots are available upon request.

Figure 4 is a composite set of schematic plots for each animation comparing the mean positions of the NLFs at rest and at maximum displacement for the patient and control groups. As stated previously, the left side of the face was set as the affected or paralyzed side. The mean landmark positions for the control face at rest are in black, and lines connect the landmarks on the nose, mouth and NLF regions. The upper and lower NLF landmarks at maximum displacement are connected by a red line for the control group and by a blue line for the patient group. The plots show that the greatest differences in maximum displacement of the NLFs between the patients and controls for the affected (left) side are seen for the smile (msm and nsm) and grimace (gr) animations and 'ee' sound.

Tables S2 and S3 give the mean displacement from rest to the maximum of the smile movements for the upper NLF landmarks, lower NLF landmarks and commissure landmarks relative to the middle of the lower lip for the patients and controls, respectively. Because the smile animation is the focus of facial reanimation surgery, the results for the smile are reported here. The results for the other animations are available on request. Overall, for the smile, the mean displacement from rest for the commissure landmarks ranged from 11.3 to 14.0 millimeters for the control group

and 7.2-10.2 millimeters for the patient group. For the patients, the landmarks on both the affected (paralyzed) and presumed unaffected sides of the face demonstrated less mean displacement than the corresponding landmarks of the controls. Nonetheless, for both the patients and controls, the upper and lower NLF landmarks had approximately similar magnitudes of displacement as the commissure landmarks (~ 1cm for the controls and variable for the patients). Also, the 'at rest' distances from the middle of the lower lip to the upper NLF landmarks ($P < .05$) and from the middle of the lower lip to the commissure landmarks ($P < .01$) were less for the patients. Finally, the distributions and normality of the data can be seen in the plots in the Supplementary Material. P-values were reported so the reader can choose their own alpha level. Any age, sex and ethnicity effects that exist were not possible to detect without a much larger study. The analysis in this study controls for size and initial shape so that these effects would be relatively small.

4 | DISCUSSION

This study is the first to provide an in-depth analysis of the NLF dynamics and associated asymmetry in patients with facial paralysis and

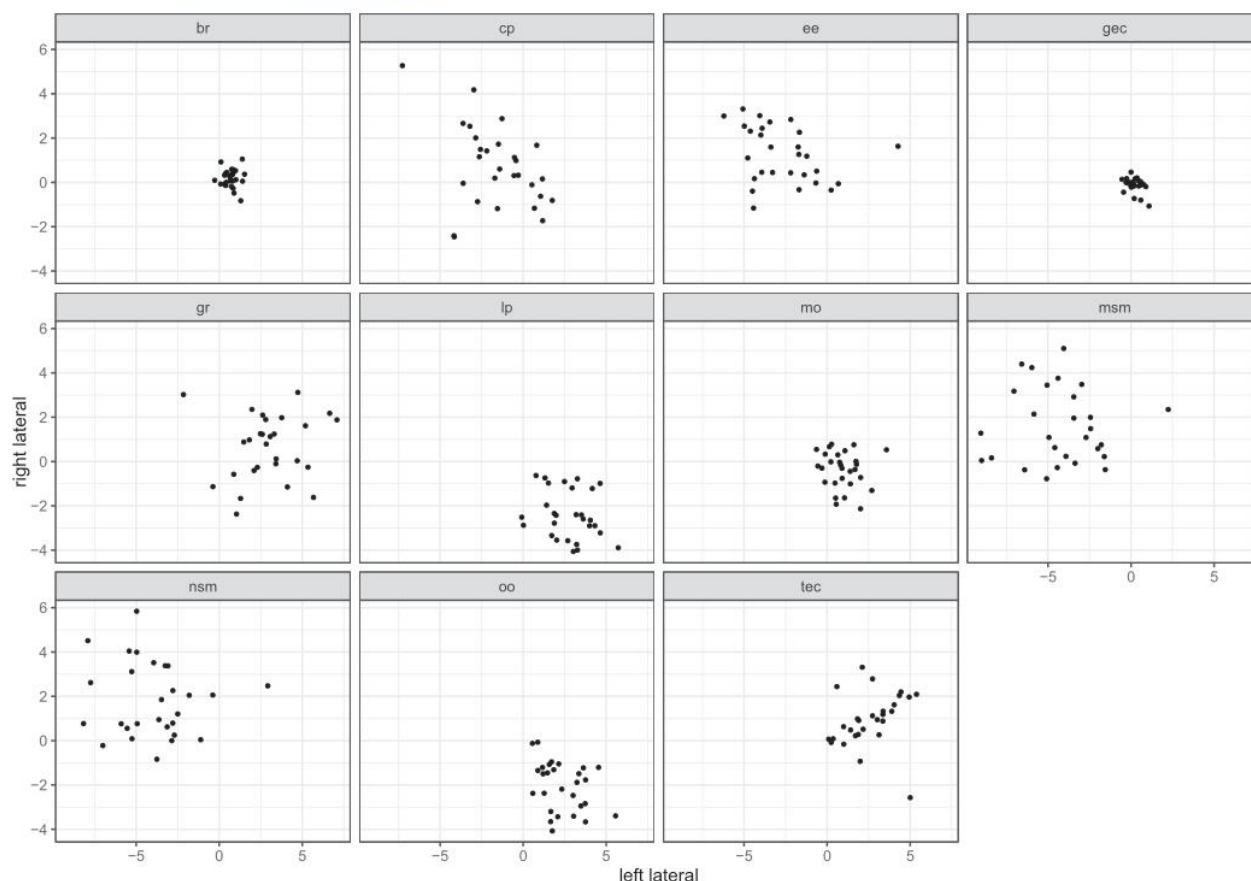


FIGURE 3 Lateral displacement of the patients' right and left landmark pairs at the upper end of the NLFs for each animation

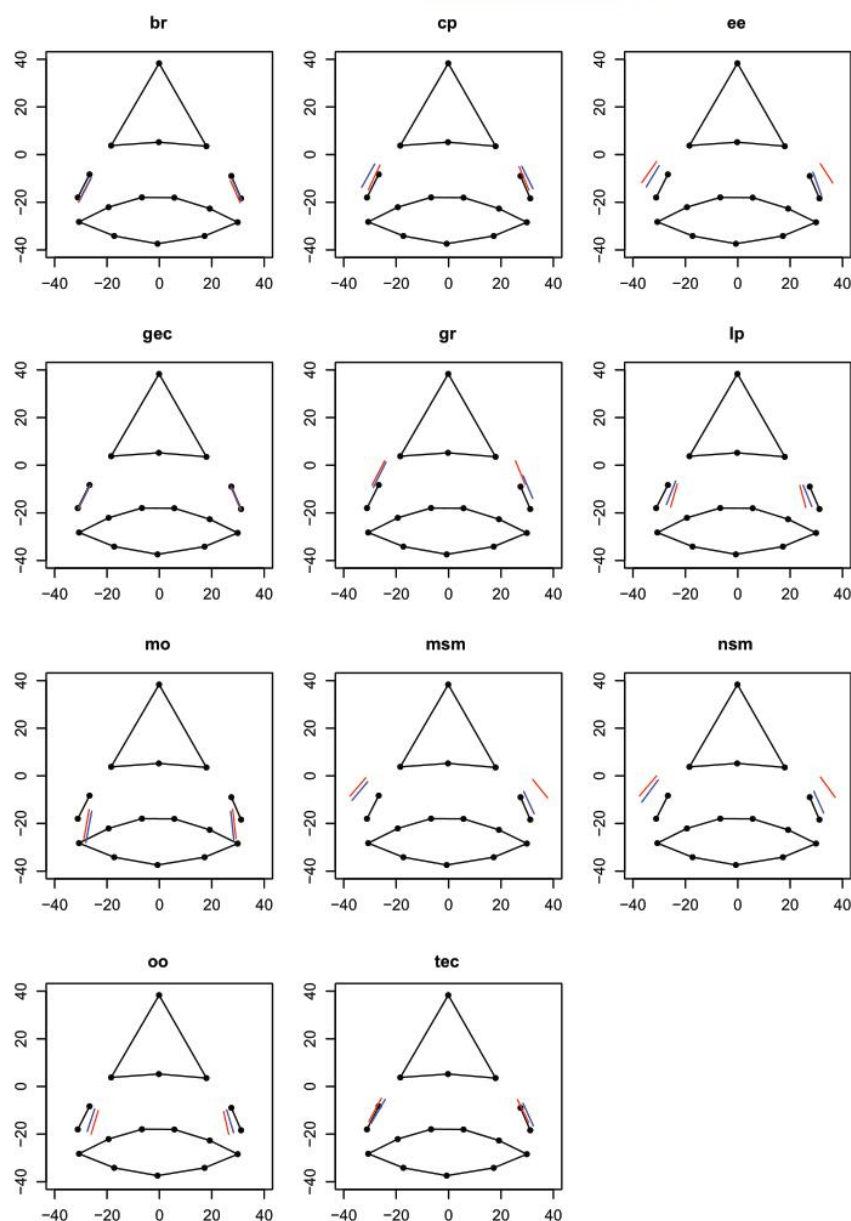
control subjects. Previous methods of quantification involved the use of 2D measures and analyses^{9,10} and focused on the asymmetry of the brow, upper and lower eyelids and oral commissures—much more limited in scope compared with the analyses presented here. Moreover, the issue of validity of 2D analyses is a concern.^{7,8} In our study, an objective and highly detailed 3D facial movement analysis was conducted that provided insights into both normal NLF animation dynamics and the dynamics of the folds in patients with unilateral facial paralysis. Such an analysis is important to understand the relationships of normal movement of the NLF in all three planes of space—transversely, vertically and antero-posteriorly (depth), but as demonstrated, the analysis also can be used to diagnose impaired movements in patients as well as assess outcomes of reconstructive surgeries.

Currently, there is no universal agreement on a measurement system or outcome measure for facial palsy.¹¹ A more recent review of the literature highlighted the need for multiple validated outcome measures in order to capture all domains of successful surgical reanimation and included Patient-Reported Outcome Measures, Automated and Clinician-Graded Facial Palsy Grading Systems, Layperson Assessment Equivalent, and Spontaneous Smile Analysis.¹¹ In the present study, the focus was on measuring facial soft tissue movement with an emphasis on the NLFs. One popular measure for this purpose is the eFACE which is a scale used by clinicians to grade

the severity of the facial paralysis using specific software.^{12,13} The clinician subjectively rates or scores facial regions according to the degree of facial symmetry at rest and at the maximum of the smile movement only. Experts who rated patients' disfigurement with this scale deemed that the NLF morphology and symmetry were two very influential parameters when assessing paralysis.¹⁴

In fact, static NFL suspension is one of the most common ancillary surgical procedures performed in treating facial palsy because the cheek position at rest, as delineated by the NLFs, is of critical importance (Figure S5).¹⁵ Yet restoration of NLF dynamics is not generally considered important by reanimation specialists (Figure S6).¹⁵ From a dynamic perspective, we found herein that there was considerable movement of the NLFs during smiling and other animated movements. During normal smiling, the mean 3D movement of the NLFs was of similar magnitude to that of the oral commissures, around the order of one centimeter. For the patients, although the NFL movement also was of similar magnitude as the commissures, the overall mean movement was reduced on both sides of the face when compared with the controls. In a previous study,⁸ we found that the affected side of the face may tether the movement of the unaffected side. Additionally, symmetry of the *normal* NFL movements was highly coordinated in all three planes of space between the right and left sides of the face. However, in affected patients

FIGURE 4 Schematic plots for each animation showing the mean landmark positions for the control face (black lines) at rest and a comparison of the NLF landmarks at maximum displacement (controls in red, patients in blue)



the movement was limited in magnitude and highly uncoordinated, and the movement in the lateral and antero-posterior directions was much more uncoordinated than the vertical movements. These findings suggest that an immobile NLF is an unnatural feature of the smile movement/animation and that surgical treatment strategies that address impaired NFL movements during animations must be considered to create a more natural surgical outcome. It is likely that the level of NFL dynamism will need to be tailored to an individual patient's dynamic anatomy and many surgeons may already incorporate strategies to restore dynamic movement to the NLF. Increased awareness of the importance of restoring this dynamic feature of smiling is needed. The 3D outcome measures described in this and other studies afford an approach to personalize facial soft tissue movement outcomes in patients.

5 | CONCLUSION

This study is the first to provide an in-depth, 3D analysis of the NLF dynamics in patients with facial paralysis and control subjects. This analysis can be used to diagnose impaired facial soft tissue movements in patients and assess outcomes of reconstructive surgeries. From a dynamic perspective based on the analysis, we found considerable movement of the NLFs during smiling and other animated movements. During normal smiling, movement of the NLFs was highly coordinated, symmetrical, and of similar magnitude to that of the oral commissures—around the order of one centimeter. Affected patients had highly uncoordinated and asymmetrical movements that were limited in magnitude suggesting that an immobile NLF is an unnatural feature of the smile movement/

animation, and surgical treatment strategies tailored to a patients' specific needs that address impaired NFL movements during animations must be considered to create a more natural surgical outcome.

ACKNOWLEDGEMENTS

The authors would like to thank Cassandra O'Connell for her assistance.

ORCID

Carroll Ann Trotman  <https://orcid.org/0000-0003-0778-7494>

REFERENCES

1. Snyder-Warwick AK, Fattah AY, Zive L, Halliday W, Borschel GH, Zuker RM. The degree of facial movement following microvascular muscle transfer in pediatric facial reanimation depends on donor motor nerve axonal density. *Plast Reconstr Surg*. 2015;135(2):370e-381e.
2. Manktelow RT, Zuker RM, Tomat LR. Facial paralysis measurement with a handheld ruler. *Plast Reconstr Surg*. 2008;121(2):435-442.
3. Van Veen MM, Tavares-Brito J, Van Veen BM, et al. Association of regional facial dysfunction with facial palsy-related quality of life. *JAMA Facial Plast Surg*. 2019;21(1):32-37.
4. Barton FE Jr, Gyimesi IM. Anatomy of the nasolabial fold. *Plast Reconstr Surg*. 1997;100(5):1276-1280.
5. Yousif NJ, Gosain A, Matloub HS, et al. The nasolabial fold: an anatomic and histologic reappraisal. *Plast Reconstr Surg*. 1994;93(1):60-69.
6. Millard DR Jr, Yuan RT, Devine JW Jr. A challenge to the undefeated nasolabial folds. *Plast Reconstr Surg*. 1987;80(1):37-46.
7. Trotman CA, Faraway J, Hadlock T, Banks C, Jowett N, Jung HJ. Facial soft-tissue mobility: baseline dynamics of patients with unilateral facial paralysis. *Plast Reconstr Surg Glob Open*. 2018;6(10):e1955.
8. Trotman CA, Faraway J, Hadlock TA. Facial mobility and recovery in patients with unilateral facial paralysis. *Orthod Craniofac Res*. 2020;23(1):82-91.
9. Hadlock TA, Urban LS. Toward a universal, automated facial measurement tool in facial reanimation. *Arch Facial Plast Surg*. 2012;14(4):277-282.
10. Guarin DL, Dusseldorp JR, Hadlock TA, Jowett N. A machine learning approach for Automated Facial Measurements in Facial Palsy. *JAMA Fac Plast Surg*. 2018;20(4):335-337.
11. Dusseldorp JR, van Veen MM, Mohan S, Hadlock TA. Outcome tracking in facial palsy. *Otolaryngol Clin North Am*. 2018;51(6):1033-1050.
12. Eviston TJ, Chong LSH, Clark JR. Validation of the clinician-graded electronic facial paralysis assessment. *Plastic and Reconstr Surg*. 2018;141(4):615e-616e.
13. Banks CA, Bhama PK, Park J, Hadlock CR, Hadlock TA. Clinician-graded electronic facial paralysis assessment. *Plast Reconstr Surg*. 2015;136(2):223e-230e.
14. Banks CA, Jowett N, Hadlock CR, Hadlock TA. Weighting of facial grading variables to disfigurement in facial palsy. *JAMA Facial Plast Surg*. 2016;18(4):292-298.
15. Dusseldorp JR, Guarin DL, van Veen MM, Jowett N, Hadlock T. In the eye of the beholder: changes in perceived emotion expression after smile reanimation. *Plast. Reconstr Surg*. 2019;144(2):457-471.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Dusseldorp JR, Faraway J, Razavi L, Hadlock TA, Trotman CA. Nasolabial fold dynamics: Implications for facial paralysis and facial reanimation surgery. *Orthod. Craniofac. Res.* 2020;00:1-8. <https://doi.org/10.1111/ocr.12400>

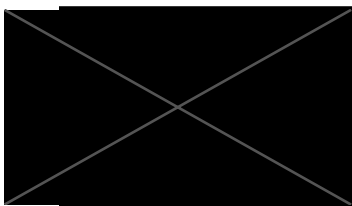
Appendix 5

A Machine Learning Approach for Automated Facial Measurements in Facial Palsy

Authorship Statement

I, Joseph Dusseldorp, confirm for the ensuing journal article: ‘**A Machine Learning Approach for Automated Facial Measurements in Facial Palsy**’ that I have made the following contributions to its authorship:

- Conception and design of the research
- Analysis and interpretation of the findings
- Drafting and revising of the manuscript and critical appraisal of content
- Responding to multiple stages of peer review



Joseph R. Dusseldorp

Date 14.09. 2023

As Auxiliary Supervisor for the candidature upon which this thesis is based, I can confirm that the above authorship attribution is correct.

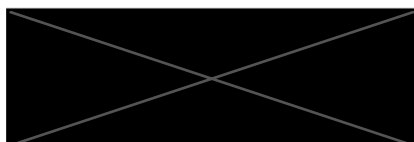


Professor Tessa Hadlock, MD

Date 27.07. 2023

Department of Otolaryngology
Massachusetts Eye and Ear Infirmary
Harvard University, Boston, MA, USA

As Primary Supervisor of the candidate, I concur with the above authorship attestation:



Professor Peter Maitz

Date 14.09.2023

Department of Burns Surgery
Concord Repatriation General Hospital
Sydney University, Sydney, NSW, Australia

Letters

OBSERVATION

A Machine Learning Approach for Automated Facial Measurements in Facial Palsy

An ongoing problem in the management of facial neuromotor disorders is the lack of a universal outcome measurement system.¹ Objective systems have been developed to quantify facial symmetry.¹⁻³ However, these systems are time consuming, require expensive and complicated hardware, and can be difficult to implement in clinical practice. Recently, machine learning techniques have been developed that enable automatic localization of facial landmarks using large data sets of facial photographs.⁴⁻⁶ We have leveraged these technological advances to develop Emotrics, a simple, high-throughput software platform that enables automatic facial landmark localization and computation of facial measurements.

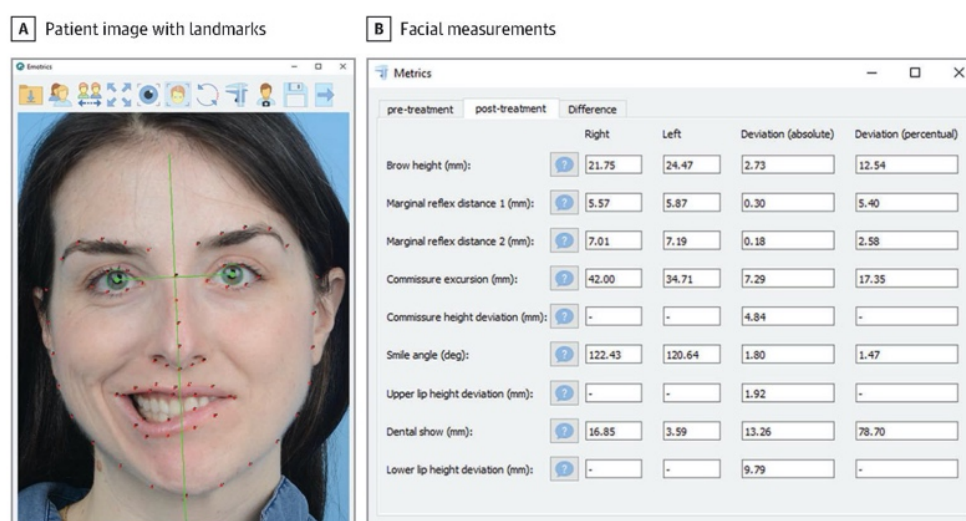
Emotrics is designed for use with frontal-view clinical photographs, automatically placing facial landmark dots on an uploaded image. Emotrics automatically generates multiple facial measurements by scaling iris diameter to pixel width in each image. This measurement technique uses a mean human population iris diameter of 11.77 mm; this value is comparable to that used in the Massachusetts Eye and Ear Infirmary's FACE-gram software (Sir Charles Bell Society). However, Emotrics has 2 important advantages over FACE-gram. Emotrics rapidly computes multiple relevant facial measure-

ments simultaneously, with full analysis of one image taking less than 5 seconds on average. Emotrics can also analyze the differences between 2 photographs, which allows automated calculation of smile excursion.

Functionalities | The user can manually reposition any facial landmark dot requiring refinement (red dots). Facial landmark dots should outline the superior border of the brow, the free margin of the upper and lower eyelids, the nasal midline, the nasal base, the mucosal edge and vermillion-cutaneous junction of the upper and lower lips, and the lower two-thirds of the face. The locations of the eye centers (green dots) and iris borders (green circles) may also be finely adjusted, which provides additional flexibility and accuracy. The facial midline may be displayed to ensure that the user is satisfied with its position (Figure 1); this line is computed as a line perpendicular to the interpupillary plane at the midpupillary point.

Emotrics automatically computes a literature-established set of facial measurements relevant to facial palsy (Figure 1). Brow symmetry, palpebral fissure width, smile excursion, and smile symmetry outputs are automatically generated. During future development of Emotrics, outputs will appear in an even more user-friendly fashion. Emotrics can also measure differences between any 2 images manually selected by the user (Figure 2), although ultimately this process will be automated. By comparing images obtained before and after intervention, users can rapidly obtain measures of treat-

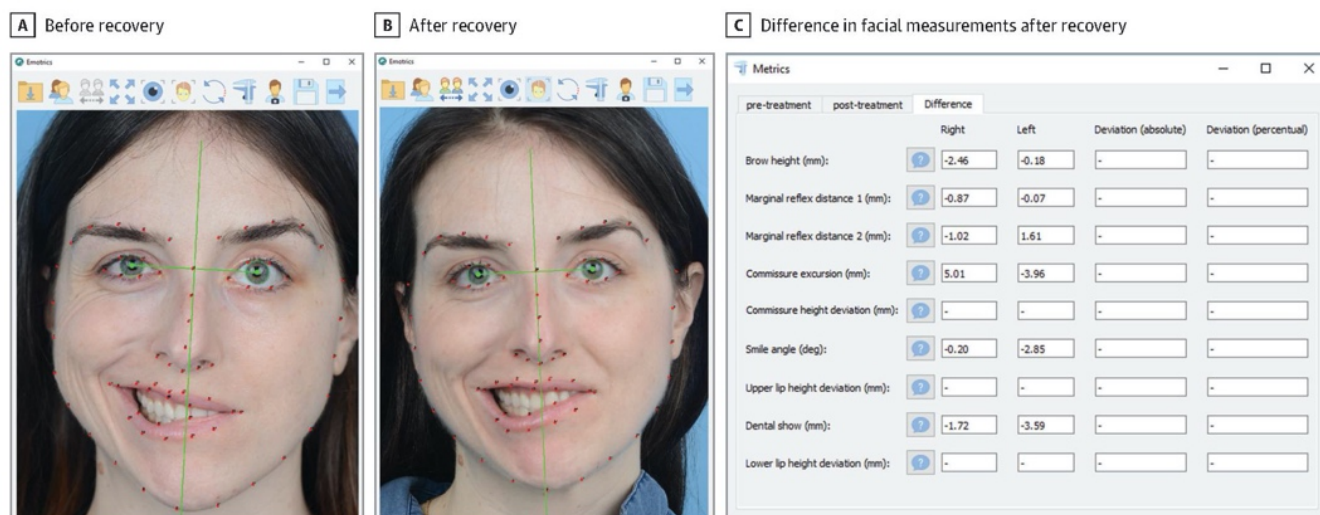
Figure 1. Graphical User Interface of Emotrics and Facial Measurements Computed From the Facial Landmarks



A, Icon bar displays the different functionalities: load image, load patient, compare 2 images, change image, fit image to window, resize iris diameter, find facial midline, toggle landmarks, show facial measurements, take screenshot, save results, and exit. Each function is fully described in the video tutorial. The window of Emotrics displays the facial photograph with the 68 facial landmarks

(red dots) and bilateral iris positions (green circle). The face midline (vertical green line) can be easily estimated and displayed as a reference. B, Pixel width is automatically normalized using a mean iris diameter of 11.77 mm to produce millimetric values. A complete description of each facial measurement is available within Emotrics by clicking on the help icon.

Figure 2. Case Example of a Patient With Aberrant Recovery From Left-Sided Bell Palsy



This example shows photographs taken before (A) and after (B) recovery. The facial midline is shown in both patient images, computed as a line perpendicular to the interpupillary plane at the midpupillary point. The difference in facial

measurements between photographs obtained before and after recovery (for both sides of the face) is shown in panel C.

ment effects. Similarly, comparing the facial position during maximum smile effort with the resting facial position enables automated measurement of smile excursion.

Discussion | Emotrics was created using a database of normal faces, which renders the program prone to localization errors in cases of gross facial asymmetry. User verification of landmarks is recommended for photographs of patients with facial palsy. An additional limitation is that the accuracy of the measurements is dependent on image quality. Image resolution of at least 1 megapixel is ideal; at least 1 iris should be visible, and there should be no head yaw or tilt (head roll is acceptable). If the eyes are closed, the user can import pupil positions from a previous photograph in the same clinical series, which would enable measurements in these images. Where there is strabismus or orbital dystopia, the user may manually select an ideal pupil position that enables facial measurements.

An objective means of characterizing facial displacements is essential to the management of facial palsy. Emotrics harnesses recent advances in machine learning to automatically compute facial displacements from standard photographs. This tool may facilitate communication of disease severity and outcomes within the field of facial reanimation by means of shared, objective, data-driven language. Emotrics is freely available and can be downloaded, along with detailed tutorial videos, from the Sir Charles Bell Society website (<http://www.sircharlesbell.org/>).

Diego L. Guarin, PhD
Joseph Dusseldorp, MD
Tessa A. Hadlock, MD
Nate Jowett, MD, FRCSC

Author Affiliations: Department of Otolaryngology, Massachusetts Eye and Ear Infirmary, Harvard Medical School, Cambridge (Guarin, Dusseldorp, Hadlock, Jowett); University of Sydney, Sydney, Australia (Dusseldorp).

Corresponding Author: Diego L. Guarin, PhD, Massachusetts Eye and Ear Infirmary, Harvard Medical School, 243 Charles St, Boston, MA 02114 (diego_guarin@meei.harvard.edu).

Published Online: March 15, 2018. doi:10.1001/jamafacial.2018.0030

Conflict of Interest Disclosures: None reported.

Additional Contributions: We thank the patient for granting permission to publish this information.

- Hadlock TA, Urban LS. Toward a universal, automated facial measurement tool in facial reanimation. *Arch Facial Plast Surg*. 2012;14(4):277-282.
- Bray D, Henstrom DK, Cheney ML, Hadlock TA. Assessing outcomes in facial reanimation: evaluation and validation of the SMILE system for measuring lip excursion during smiling. *Arch Facial Plast Surg*. 2010;12(5):352-354.
- Coulson SE, Croxson GR, Gilleard WL. Three-dimensional quantification of "still" points during normal facial movement. *Ann Otol Rhinol Laryngol*. 1999; 108(3):265-268.
- Vahid K, Sullivan J. One millisecond face alignment with an ensemble of regression trees. In: *27th IEEE Conference on Computer Vision and Pattern Recognition*. Piscataway, NJ: IEEE; 2014:1867-1874.
- King DE. Dlib-ml: a machine learning toolkit. *J Mach Learn Res*. 2009;10:1755-1758.
- Sagonas C, Georgios T, Stefanos Z, Maja P. A semi-automatic methodology for facial landmark annotation. In: *26th IEEE Conference on Computer Vision and Pattern Recognition*. Piscataway, NJ: IEEE; 2013:896-903.

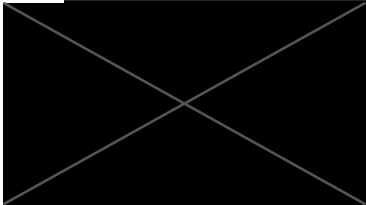
Appendix 6

Toward an Automatic System for Computer-Aided Assessment in Facial Palsy

Authorship Statement

I, Joseph Dusseldorp, confirm for the ensuing journal article: ‘**Toward an Automatic System for Computer-Aided Assessment in Facial Palsy**’ that I have made the following contributions to its authorship:

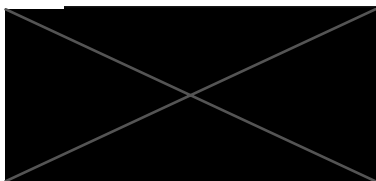
- Interpretation of the findings
- Drafting and revising of the manuscript



Joseph R. Dusseldorp

Date 14.09. 2023

As Auxiliary Supervisor for the candidature upon which this thesis is based, I can confirm that the above authorship attribution is correct.



Professor Tessa Hadlock, MD

Date 27.07. 2023

Department of Otolaryngology
Massachusetts Eye and Ear Infirmary
Harvard University, Boston, MA, USA

As Primary Supervisor of the candidate, I concur with the above authorship attestation:



Professor Peter Maitz

Date 14.09.2023

Department of Burns Surgery
Concord Repatriation General Hospital
Sydney University, Sydney, NSW, Australia

ORIGINAL INVESTIGATION

Toward an Automatic System for Computer-Aided Assessment in Facial Palsy

Diego L. Guarin, PhD,^{1,2,*} Yana Yunusova, PhD, CCC-SLP,^{1,3,4} Babak Taati, PhD,^{1,5,6}
Joseph R. Dusseldorp, MBBS, MS, FRACS,⁷ Suresh Mohan, MD,² Joana Tavares, MD,⁸
Martinus M. van Veen, MD,² Emily Fortier,² Tessa A. Hadlock, MD,² and Nate Jowett, MD²

Abstract

Importance: Quantitative assessment of facial function is challenging, and subjective grading scales such as House–Brackmann, Sunnybrook, and eFACE have well-recognized limitations. Machine learning (ML) approaches to facial landmark localization carry great clinical potential as they enable high-throughput automated quantification of relevant facial metrics from photographs and videos. However, the translation from research settings to clinical application still requires important improvements.

Objective: To develop a novel ML algorithm for fast and accurate localization of facial landmarks in photographs of facial palsy patients and utilize this technology as part of an automated computer-aided diagnosis system.

Design, Setting, and Participants: Portrait photographs of 8 expressions obtained from 200 facial palsy patients and 10 healthy participants were manually annotated by localizing 68 facial landmarks in each photograph and by 3 trained clinicians using a custom graphical user interface. A novel ML model for automated facial landmark localization was trained using this disease-specific database. Algorithm accuracy was compared with manual markings and the output of a model trained using a larger database consisting only of healthy subjects.

Main Outcomes and Measurements: Root mean square error normalized by the interocular distance (NRMSE) of facial landmark localization between prediction of ML algorithm and manually localized landmarks.

Results: Publicly available algorithms for facial landmark localization provide poor localization accuracy when applied to photographs of patients compared with photographs of healthy controls (NRMSE, 8.56 ± 2.16 vs. 7.09 ± 2.34 , $p \ll 0.01$). We found significant improvement in facial landmark localization accuracy for the facial palsy patient population when using a model trained with a relatively small number photographs (1440) of patients compared with a model trained using several thousand more images of healthy faces (NRMSE, 6.03 ± 2.43 vs. 8.56 ± 2.16 , $p \ll 0.01$).

Conclusions and Relevance: Retraining a computer vision facial landmark detection model with fewer than 1600 annotated images of patients significantly improved landmark detection performance in frontal view photographs of this population. The new annotated database and facial landmark localization model represent the first steps toward an automatic system for computer-aided assessment in facial palsy.

Level of Evidence: 4

¹KITE | Toronto Rehabilitation Institute—University Health Network, Toronto, Canada.

²Department of Otolaryngology/Head and Neck Surgery, Massachusetts Eye and Ear Infirmary and Harvard Medical School, Boston, Massachusetts.

³Department of Speech Language Pathology, University of Toronto, Toronto, Canada.

⁴Hurvitz Brain Sciences Program, Sunnybrook Research Institute, Toronto, Canada.

⁵Department of Computer Science, University of Toronto, Toronto, Canada.

⁶Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, Canada.

⁷Department of Plastic and Reconstructive Surgery, Royal Australasian College of Surgeons and University of Sydney, Sydney, Australia.

⁸Faculty of Health Sciences, Brasília University, Brasília, Brazil.

*Address correspondence to: Diego L. Guarin, PhD, Toronto Rehabilitation Institute—University Health Network, 550 University Avenue, Toronto ON M5G 2A2, Canada, E-mail: diego.guarin@uhn.ca

KEY POINTS

Question: Can recent developments in ML and computer vision be used to develop an objective and automatic system for computer-aided assessment in facial palsy?

Findings: In this research article, we found that by using a relatively small number of manually annotated photographs for a patient-specific database, it is possible to obtain significant improvement in the accuracy of facial measurements provided by a popular ML algorithm.

Meaning: The results presented in this article represent the first steps toward the development of an automatic system for computer-aided assessment in facial palsy.

Introduction

Management of facial neuromuscular disorders is hampered by lack of a universal and objective grading system to characterize disease severity, recovery, and response to therapeutic interventions.¹ Quantifying static facial features and displacements occurring with facial expressions is a promising technique for standardizing assessment in facial palsy, whose reported U.S. incidence exceeds 150,000 cases per year.² Several methods exist for measuring facial features and movements. Caliper assessments offer high accuracy yet are tedious and must be performed in person. Computer-based techniques to quantify facial displacements are now widely employed.^{3–7} Early approaches comprised manual identification of facial landmarks on digital images within specialized software, from which relevant distances and angles could be readily calculated. Although such techniques enabled retrospective assessment of facial function, manual tagging of digital images is resource intensive, error prone, and infeasible for dynamic tracking of facial movements from video. To automate measurement of facial displacements, physical markers placed at specific facial landmarks have been employed, and their location was tracked using customized software and hardware.⁶ Marker-based tracking is limited as manual marker localization is time consuming, subjective, and requires specialized recording conditions.

Machine learning (ML)-based computer vision algorithms enable rapid and fully automated tracking of facial displacements from digital images and videos recorded under typical conditions with consumer-grade cameras. Such facial landmark detection algorithms are usually trained using databases of manually annotated facial photographs. Once trained, these ML algorithms can predict the position of facial landmarks in a new photograph without human intervention, with high accuracy.^{8–13} ML algorithms for facial landmark localization are increasingly being used to study facial palsy,^{7,14–17}

Parkinson disease,¹⁸ stroke,¹⁹ amyotrophic lateral sclerosis,²⁰ and dementia.²¹ Owing to their training using predominantly normal subjects, current ML models for facial landmark recognition may be biased against patients, and demonstrate inadequate accuracy when presented with faces of patients with neuromuscular disease impacting facial movements and expression.²¹ Herein, we hypothesize that training a ML model for facial landmark localization with facial photographs from a disease-specific clinical database will demonstrate improved tracking accuracy when presented with faces of patients with the condition, in comparison with the model trained using a much larger database of normal subjects.

To evaluate our hypothesis, we introduce the first database of annotated clinical photographs of patients with unilateral facial palsy, employ it to evaluate the bias against patients of a popular facial landmark localization algorithm, and train a ML model for automated facial landmark localization in this population. We further demonstrate the utility of an open-access and user-friendly software, Emotrics, for extracting clinically relevant measurements from photographs and videos in automated manner using the ML model trained herein.

Methods**Automatic facial landmark localization**

We employed a popular approach for automatic facial landmark localization in facial photographs known as cascade of regression trees.^{11,22–25} Specifically, we employed the algorithm proposed by Kazemi and Sullivan,¹¹ which provides accurate facial landmark localization results under multitude of pose, illumination, and expression (PIE) conditions²⁶ and can process medical images in just a few hundred milliseconds without the use of graphical processing units or other specialized hardware.¹⁰ Implementation of this algorithm for facial landmark localization is readily available in open source ML libraries such as OpenCV²⁷ and Dlib.²⁸ These implementations were trained using the 300-W data set from the Intelligent Behavior Understanding Group (iBUG).¹² This data set is a concatenation of several freely available databases (LFPW,²⁹ HELEN,³⁰ AFW,³¹ and iBUG³²) and comprises 11,500 in-the-wild photographs of >4000 healthy subjects. Model training has been performed by manually annotating a set of facial landmarks in each photograph using the 68-point Carnegie Mellon University multiple pose, illumination, and expression (multi-PIE) database approach.^{33,34} Manually annotated landmarks outlined the superior border of the brow, the free margin of the upper and lower eyelids, the nasal midline, the nasal base, the mucosal edge and vermilion-cutaneous junction of the upper and lower lips, and the lower two-thirds of the face.

MEEI database

We obtained institutional review board approval to access the Massachusetts Eye and Ear (MEE) Facial Nerve Data Repository, a digital collection of facial photographs, and video clips of patients with unilateral facial palsy at the MEE Facial Nerve Center. The patients had consented to use of the facial photographs for research purposes. A clinical photographer captured high-resolution photographs (1080×720 pixels) of a standardized series of facial expressions used to evaluate facial mimetic function.³⁵ Figure 1 exemplifies the type of photographs taken from each patient; all images were taken from the frontal view using a digital camera with optimal lighting.

In addition to the patient's photographs, we obtained high-resolution photographs of healthy controls from the MEEI Facial Palsy Photo and Video Standard Set¹⁴ performing the same standardized facial expressions as the patients.

Photograph annotation

Three trained clinicians independently annotated the 68 multi-PIE landmark points for each photograph in the clinical database using a custom graphical user interface¹⁰ (Emotrics Software, Mass Eye and Ear, Boston, MA).

Images were uploaded into Emotrics, which provided an initial estimation of the 68 landmark points. Landmark positions for each photograph were verified and manually repositioned as necessary by each marker. Landmark positions of three manual annotators were averaged for each photograph to define ground-truth locations.

Evaluating model bias and training a new model for landmark localization

Marked photographs were clustered by patient and randomly divided into three nonoverlapping groups, all images of the same patients were only in one group. The groups were model training ($N=180$ subjects; equating to 90% of the database or 1440 photographs), validation ($N=10$ subjects; equating to 5% of the database or 80 photographs), and test ($N=10$ subjects; equating to 5% of the database or 80 photographs). Healthy control photographs ($N=10$ subjects, equating to 80 photographs) were used for model testing only. Computer operations were performed in the Python programming language (version 3.6.7) on a Lenovo Thinkpad personal computer (T470, Intel Core i7-7600U processor running at 2.8 GHz with 32 GB of RAM) that did not include specialized hardware for training and evaluation of ML models (graphics processing units or tensor processing units).



Fig. 1. Example of the standardized set of facial expressions among facial palsy patients in the annotated database. These facial expressions enable global and zonal evaluation of facial function from still photographs.

Evaluation of model bias against patients

Evaluation of model bias against patients was performed using publicly available implementation of the ensemble of regression trees algorithm for facial landmark localization proposed by Kazemi and Sullivan.¹¹ Available algorithm was used to estimate the position of 68 facial landmarks in the test group of our database and the photographs of healthy controls; predicted landmark positions were compared with the ground-truth locations provided by the manual annotation procedure. Errors in landmark localization were quantified using the root-mean-square error (RMSE) between ground-truth and algorithm-predicted landmark positions normalized by the interocular distance (NRMSE).³⁸

Training a specialized model for facial landmark localization in patients

Next, photographs from the training group were employed to retrain the ensemble of regression trees model. Photographs from the validation group were used to determine the model parameters that provided the minimum landmark localization error using cross-validation. Standard model parameters such as number of estimators, tree deep, minimum number of samples per leaf, and number of features were selected using a grid search; a total of 400 permutations were assessed. The parameter set yielding the lowest NRMSE for the validation data set was chosen for model testing and comparison.

The retrained model was compared against the implementation trained with the 300-W database by computing the NRMSE between the ground-truth and model-predicted landmark positions yielded by both models for the test group (patients and controls) of our database.

Statistical analysis

Differences between the groups (i.e., healthy vs. patients) were sought using the Kruskal–Wallis one-way analysis of variance; differences between models (i.e., publicly available vs. retrained using patients' photographs) were sought using the Wilcoxon signed rank test. Statistical significance was considered at $p < 0.01$.

Results

MEE database

All patients who visited the Facial Nerve Center at MEE between October 2017 and October 2018 and provided written consent to use their photographs for research purposes were considered candidates for the database. Patients who demonstrated gross facial deformities or whose facial features were not clearly visible (due to obstruction by clothing or hair) were excluded. Media from 200 adult and pediatric patients with varying facial palsy

severity and etiology, totaling 1600 photographs, were collected, manually annotated, and employed for algorithm training and testing; Table 1 presents relevant demographic and clinical information.

Model bias against patients

Figure 2A shows a box and whiskers plot representing the NRMSE yielded by the standard implementation of the ensemble of regression trees model trained with the 300-W database, when applied to photographs of healthy controls and patients with facial palsy. Our results demonstrated that this implementation of the algorithm yielded significantly worst facial landmark localization accuracy, as quantified by the NRMSE, for patients than for healthy subjects (7.09 ± 2.34 vs. 8.56 ± 2.16 , $p \ll 0.01$).

Table 1. Patients' demographics

Demographics	N
Age	
Range (years)	7–89
Mean $\pm$ SD (years)	48.9 $\pm$ 17.1
Gender	
Female	135
Male	65
Race	
White	160
Hispanic	15
Asian	14
Black	11
Etiology	
Infectious	
Bell's palsy	65
Ramsay Hunt syndrome	14
Lyme disease	12
Pregnancy-associated Bell's palsy	7
Zoster sine herpete	6
Meningitis	1
Congenital	
Congenital facial palsy	5
Otolgic	
Cholesteatoma	2
Neoplastic	
Vestibular schwannoma	21
Parotid neoplasm	14
Facial nerve schwannoma	8
Brainstem neoplasm	4
CNS metastasis	1
Basal cell carcinoma	1
Trigeminal nerve neoplasm	1
Trauma	
Trauma	6
Iatrogenic	
Iatrogenic	7
Neurologic	
Hemifacial spasm	1
Multiple sclerosis	1
Vascular	
Brainstem stroke	1
Venous malformation	1
Cavernous brainstem hemangioma	1

CNS, central nervous system.

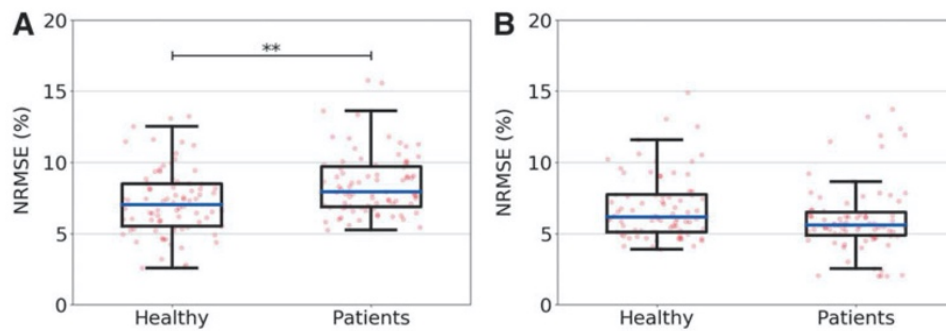


Fig. 2. NRMSE of facial landmark localization **(A)**. Model trained with photographs of healthy subjects applied to photographs of healthy subjects and patients. **(B)** Model trained with photographs of patients suffering from facial palsy applied to photographs of healthy subjects and patients. NRMSE, normalized root-mean-square error.

Specialized model accuracy

Figure 2B shows a box and whiskers plot representing the NRMSE yielded by the ensemble of regression trees model retrained with the MEEI database, when applied to photographs of healthy controls and patients suffering from facial palsy. Our analysis demonstrated that there was no significant difference in the facial landmark localization accuracy between patients and healthy subjects (6.87 ± 2.28 vs. 6.03 ± 2.43 , $p = 0.03$).

There was no significant difference in the landmark localization accuracy provided by the models trained with

the 300-W and MEEI databases when applied to healthy subjects (7.09 ± 2.34 vs. 6.87 ± 2.28 , $p = 0.162$). In contrast, the retrained model provided significantly improved accuracy when applied to patients (8.56 ± 2.16 vs. 6.03 ± 2.43 , $p \ll 0.01$). Figure 3 illustrates the improved accuracy of the model trained with the MEEI database over the model trained with the 300-W database. Figure 3A and C shows obvious localization errors for points defining the facial contour and lips yielded by the 300-W model output; Figure 3B and D shows that fewer and smaller errors are present for the output of the MEEI model.



Fig. 3. Comparison of automatic facial landmark localizations by models trained using the 300-W **(A, C)** and MEEI **(B, D)** databases among two patients with facial palsy in the validation group.

Discussion

Objective quantification of disease severity facilitates improved understanding of disease progression, recovery, and treatment response, and patient–clinician and interclinician communication. Facial palsy severity is typically assessed using subjective clinician facial grading systems including the House–Brackmann,³⁶ Sunnybrook,³⁷ and eFACE³⁸ scales. Such approaches are limited by high inter-rater variability, and require considerable training for proper use and interpretation.^{39,40} Although more objective methods for quantifying facial displacements have been described, no single tool has achieved widespread use.

In this study we demonstrated that ML approaches can provide objective, automatic, and accurate facial measurements in photographs of patients suffering from facial palsy, so that these methods have the potential of disrupting the current clinical practice for diagnosis and assessment of the condition. However, our results demonstrated that publicly available models, trained with databases of healthy subjects, provide significantly worst landmark localization accuracy when applied to photographs of patients. We also demonstrated that by

retraining the facial landmark localization model using a small number of photographs from a disease-specific clinical database, it is possible to significantly improve the facial landmark localization accuracy in this patient population. The standardized recording conditions (pose, illumination, expression, and background) of photographs

in the clinical database likely explain the observed high accuracy of the model, despite use of a relatively small training data set. These results supported our hypothesis that a novel model for facial landmark localization trained using disease-specific photographs would demonstrate improved tracking accuracy when presented

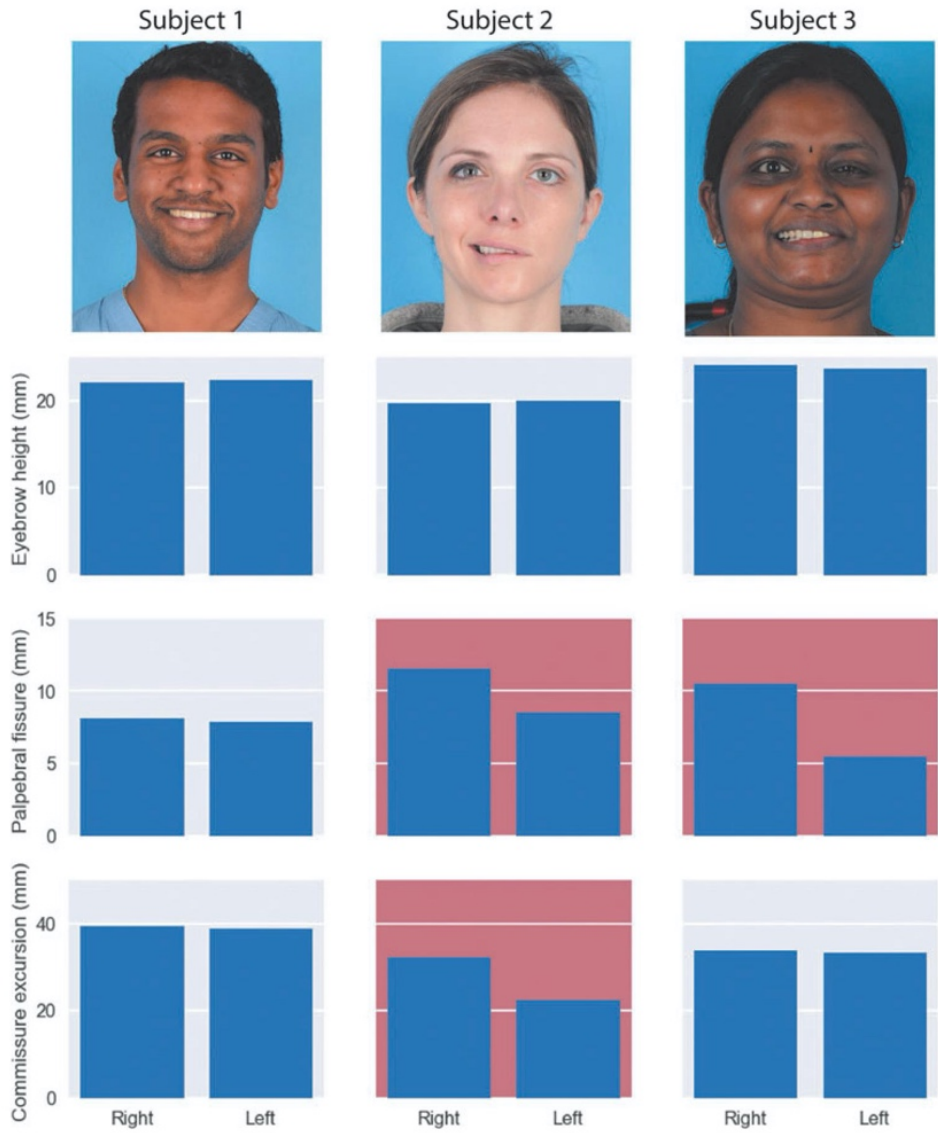


Fig. 4. Examples of automated facial metrics calculated using Emotrics software. Measurements were computed using estimated landmark position provided by the ML model trained using photographs of facial palsy patients. Red background indicates facial measurements that are markedly different among facial sides. Subject 1 is normal, Subject 2 suffers from flaccid facial palsy, and Subject 3 suffers from postparalytic facial palsy. Presented measures include eyebrow height (vertical distance from the midpupillary point to the superior border of the brow), palpebral fissure height (vertical distance between central portions of upper and lower lid margins), and commissure excursion (distance from the facial midline at the lower lip vermillion junction to the oral commissure). ML, machine learning.

with faces of patients with the condition, in comparison with the model trained using a much larger database of normal subjects.

Finally, we found no significant difference in the landmark localization accuracy yielded by the models trained with the 300-W and MEEI databases when applied to photographs of healthy subjects. This is an unexpected but welcomed result, as it indicates that the new model can be used to track the recovery of a patient across a continuum from highly impaired to normal.

Clinical application: computer-aided assessment in facial palsy

The facial landmark localization model trained with the MEEI database was packaged into the latest version of a previously characterized open-source customized software platform for automatic estimation of clinically relevant facial metrics—Emotrics.¹⁰ This tool uses landmark positions provided by a designated 68-point facial landmark localization ML model to estimate facial metrics relevant to the field of facial palsy in high-throughput manner.^{10,14–17} Figure 4 illustrates various facial metrics computed automatically using Emotrics from clinical photographs of three subjects during full-effort smile. Photographs were taken from the MEEI Facial Palsy Photo and Video Standard Set¹⁴ and were not used to train or validate the models presented here. The first subject demonstrates normal facial function and near-symmetric facial metrics. The second subject developed left-sided flaccid facial paralysis following Bell's palsy. The third subject developed left-sided postparalytic facial palsy and synkinesis following Bell's palsy. Marked differences in oral commissure position and palpebral fissure heights between sides and subjects are readily quantified.

Limitations

There are several limitations with the facial landmark localization model described herein. The database comprises patient photographs from a single center and reflects its demographics. The database includes more female ($N=135$) than male ($N=65$) patients, and its racial demographic is mostly white ($N=160$), with small representation of minorities, including Hispanic ($N=15$), Asian ($N=14$), and black ($N=11$). In addition, the database comprises more middle-aged adults (age group = (24, 64] years, $N=142$) than younger adults (age group = (18, 25] years, $N=12$), older adults (age group = 64+ years, $N=41$), and children (age group = (0, 18] years, $N=5$). Other sources of model bias such as the presence of facial hair were not assessed. Nonsymmetric distribution of patient demographics might lead to prediction

error bias; for example, the model might demonstrate higher performance among adult middle-aged white women as they comprised the largest cohort of the training data set. The model is further limited in that recording conditions (PIE) were specific to our clinical center. Although patient pose (requiring frontal view of face with neutral roll, tilt, and yaw) and expression may be readily standardized across clinical centers, illumination conditions are more challenging to standardize and their impact on model accuracy has yet to be assessed. Further work will seek to expand the training data set to include patient photographs from multiple clinical centers to improve model accuracy across a wider range of patient demographics and disease severities. The facial landmark localization model was applied only to patient photographs of eight fixed facial expressions. Future study will seek to assess the performance of this model for dynamic tracking of facial landmarks during expression from videos of patients with unilateral facial palsy.

Availability

Emotrics and the two ML models for facial landmark localization discussed here are freely available online on GitHub (www.github.com/dguari1) and the Sir Charles Bell Society website (www.sircharlesbell.com). Emotrics software is open-access and open-source; the Python-based code can be modified as desired. Request from research institutes to share the data will be reviewed on case-by-case basis by the MEEI Institutional Review Board. Owing to patient privacy concerns, the annotated database cannot be shared online.

Conclusions

We introduced the first manually annotated database of standardized PIE photographs among patients with unilateral facial palsy. Using this data set, we demonstrated that a ML model for automatic facial landmark localization in this patient population outperforms a model trained using a much larger data set of healthy subjects. We demonstrated the clinical utility of this approach in the quantification of facial palsy disease severity from database photographs and characterized an open-access software tool that facilitates rapid calculation of relevant facial metrics in this patient population.

Author Disclosure Statement

No competing financial interests exist.

Funding Information

D.L.G. was supported by the Fonds de Recherche du Québec – Nature et technologies, grant number 208637.

References

- Hadlock TA, Urban LS. Toward a universal, automated facial measurement tool in facial reanimation. *Arch Facial Plast Surg*. 2012;14(4):277–282.
- Bleicher JN, Hamiel S, Gengler JS, Antimario J. A survey of facial paralysis: etiology and incidence. *Ear Nose Throat J*. 1996;75(6):355–358.
- Bray D, Henstrom DK, Cheney ML, Hadlock TA. Assessing outcomes in facial reanimation: evaluation and validation of the SMILE system for measuring lip excursion during smiling. *Arch Facial Plast Surg*. 2010;12(5):352–354.
- Coulson SE, Croxson GR, Gilleard WL. Three-dimensional quantification of “still” points during normal facial movement. *Ann Otol Rhinol Laryngol*. 1999;108(3):265–268.
- Frey M, Michaelidou M, Tzou CH, et al. Three-dimensional video analysis of the paralyzed face reanimated by cross-face nerve grafting and free gracilis muscle transplantation: quantification of the functional outcome. *Plast Reconstr Surg*. 2008;122(6):1709–1722.
- Gerós A, Horta R, Aguiar P. Facegram—objective quantitative analysis in facial reconstructive surgery. *J Biomed Inform*. 2016;61:1–9.
- Guo Z, Dan G, Xiang J, et al. An unobtrusive computerized assessment framework for unilateral peripheral facial paralysis. *IEEE J Biomed Health Inform*. 2018;22(3):835–841.
- Cootes TF, Edwards GJ, Taylor CJ. Active appearance models. *IEEE Trans Pattern Anal Mach Intell*. 2001;23(6):681–685.
- Fan H, Zhou E. Approaching human level facial landmark localization by deep learning. *Image Vision Comput*. 2016;47:27–35.
- Guarin DL, Dusseldorp JR, Hadlock A, Jowett N. Automated facial measurements in facial paralysis: a machine learning approach. *JAMA Facial Plast Surg*. 2017;20(4):335–337.
- Kazemi V, Sullivan J. One millisecond face alignment with an ensemble of regression trees. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. New Jersey: IEEE Publishing; 2014.
- Sagonas C, Tzimiropoulos G, Zafeiriou S, Pantic M. 300 faces in-the-wild challenge: the first facial landmark localization challenge. In *Proceedings of the IEEE International Conference on Computer Vision Workshops*. New Jersey: IEEE Publishing; 2013.
- Zhang Z, Luo P, Loy CC, Tang X. Facial landmark detection by deep multi-task learning. In *European Conference on Computer Vision*. New York: Springer; 2014.
- Greene JJ, Tavares J, Guarin DL, et al. The spectrum of facial palsy: the MEEI facial palsy photo & video standard set [published online April 25, 2019]. *Laryngoscope*. doi: 10.1002/lary.27986.
- Greene JJ, Tavares J, Guarin DL, Jowett N, Hadlock T. Surgical refinement following free gracilis transfer for smile reanimation. *Ann Plast Surg*. 2018;81(3):329–334.
- Greene JJ, Tavares J, Mohan S, Jowett N, Hadlock T. Long-term outcomes of free gracilis muscle transfer for smile reanimation in children. *J Pediatr*. 2018;202:279–284. e2.
- Jowett N, Hadlock TA. Facial palsy: diagnostic and therapeutic management. *Otolaryngol Clin North Am*. 2018;51(6):xvii–xviii.
- Li MH, Mestre TA, Fox SH, Taati B. Vision-based assessment of parkinsonism and levodopa-induced dyskinesia with pose estimation. *J Neuroeng Rehabil*. 2018;15(1):97.
- Lanz C, Olgay BS, Denzler J, Gross H-M. Facial landmark localization and feature extraction for therapeutic face exercise classification. In *International Conference on Computer Vision, Imaging and Computer Graphics*. New York: Springer; 2013.
- Bandini A, Green JR, Taati B, Orlandi S, Zinman L, Yunusova Y. Automatic detection of amyotrophic lateral sclerosis (ALS) from video-based analysis of facial movements: speech and non-speech tasks. In *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*. New Jersey: IEEE Publishing; 2018.
- Taati B, Zhao S, Ashraf AB, et al. Algorithmic bias in clinical populations—evaluating and improving facial analysis technology in older adults with dementia. *IEEE Access*. 2019;7:25527–25534.
- Dollár P, Welinder P, Perona P. Cascaded pose regression. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. New Jersey: IEEE Publishing; 2010.
- Johnston B, de Chazal P. A review of image-based automatic facial landmark identification techniques. *EURASIP J Image Video Process*. 2018;2018(1):86.
- Sun Y, Wang X, Tang X. Deep convolutional network cascade for facial point detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. New Jersey: IEEE Publishing; 2013.
- Tzimiropoulos G. Project-out cascaded regression with an application to face alignment. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. New Jersey: IEEE Publishing; 2015.
- Bulat A, Tzimiropoulos G. How far are we from solving the 2d & 3d face alignment problem?(and a dataset of 230,000 3d facial landmarks). In *Proceedings of the IEEE International Conference on Computer Vision*. New Jersey: IEEE Publishing; 2017.
- Baksheev A, Erumihov V, Korniyakov K, Pulli K. Realtime computer vision with Opencv. Queue. 2012;10(4):40 pages.
- King DE. Dlib-ml: a machine learning toolkit. *J Mach Learn Res*. 2009;10(Jul):1755–1758.
- Belhumeur PN, Jacobs DW, Kriegman DJ, Kumar N. Localizing parts of faces using a consensus of exemplars. *IEEE Trans Pattern Anal Mach Intell*. 2013;35(12):2930–2940.
- Le V, Brandt J, Lin Z, Bourdev L, Huang TS. Interactive facial feature localization. In *European Conference on Computer Vision*. New York: Springer; 2012.
- Ramanan D, Zhu X. Face detection, pose estimation, and landmark localization in the wild. In *Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Citeseer. New Jersey: IEEE Publishing; 2012.
- Sagonas C, Tzimiropoulos G, Zafeiriou S, Pantic M. A semi-automatic methodology for facial landmark annotation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. New Jersey: IEEE Publishing; 2013.
- Gross R, Matthews I, Cohn J, Kanade T, Baker S. Multi-pie. *Image Vision Comput*. 2010;28(5):807–813.
- Sim T, Baker S, Bsat M. The CMU pose, illumination, and expression (PIE) database. In *Proceedings of Fifth IEEE International Conference on Automatic Face Gesture Recognition*. New Jersey: IEEE Publishing; 2002.
- Santosa KB, Fattah A, Gavilán J, Hadlock TA, Snyder-Warwick AK. Photographic standards for patients with facial palsy and recommendations by members of the Sir Charles Bell Society. *JAMA Facial Plast Surg*. 2017;19(4):275–281.
- House W. Facial nerve grading system. *Otolaryngol Head Neck Surg*. 1985;93:184–193.
- Ross BG, Fradet G, Nedzelski JM. Development of a sensitive clinical facial grading system. *Otolaryngol Head Neck Surg*. 1996;114(3):380–386.
- Banks CA, Bhama PK, Park J, Hadlock CR, Hadlock TA. Clinician-graded electronic facial paralysis assessment: the eFACE. *Plast Reconstructive Surg*. 2015;136(2):223e–230e.
- Banks CA, Jowett N, Azizzadeh B, et al. Worldwide testing of the eFACE facial nerve clinician-graded scale. *Plast Reconstructive Surg*. 2017;139(2):491e–498e.
- Reitzen SD, Babb JS, Lalwani AK. Significance and reliability of the House-Brackmann grading system for regional facial nerve function. *Otolaryngol Head Neck Surg*. 2009;140(2):154–158.

CHAPTER FIVE

Objective Measurement of Smile Spontaneity in Facial Reanimation

Preface

“It has often struck me as a curious fact that so many shades of expression are instantly recognized without any conscious process of analysis on our part ...

... We have also seen that expression in itself, or the language of the emotions, as it has sometimes been called, is certainly of importance for the welfare of mankind.”

Charles Darwin *The Expression of Emotions in Man and Animals*. 1872.

There has never been a tool that could quantify the spontaneous smiling ability of a reanimated smile. Yet clinicians and patients continue to make complex decisions regarding their chosen surgical treatment approach based on their likelihood of developing this elusive capability. The preceding chapters have demonstrated the importance of smile spontaneity and the capabilities of advanced facial recognitions to recognise subtle changes in naso-labial fold morphology. The aim of this chapter was to describe a stepwise approach to developing an objective measure of smile spontaneity.

The first step was to reliably define and identify when spontaneous smiles occurred. A technique, known as time-stamping, was developed to manually quantify spontaneous smile occurrences from video recordings of facial expressions of healthy volunteers and facial palsy patients while they watched humorous videoclips. This paper was published in *JAMA Facial Plastic Surgery*.

Dusseldorp JR, van Veen MM, Guarin DL, Quatela O, Jowett N, Hadlock TA. Spontaneity Assessment in Dually Innervated Gracilis Smile Reanimation Surgery. *JAMA Facial Plast Surg*. 2019 Dec 1;21(6):551-557

The second step was to combine the capabilities of automated facial landmarking (Emotrics – Appendix 5) and automated emotion detection algorithms to create a system of automated identification and quantification of successful spontaneous smile reanimation. This paper was published in *Plastic and Reconstructive Surgery*.

Dusseldorp JR, Guarin DL, van Veen MM, Miller, M, Jowett N, Hadlock TA. Automated Spontaneity Assessment after Smile Reanimation: A Machine Learning Approach. *Plast. Recon. Surg*. 2022 June; 149(6):1393-1402

The presented appendix in this chapter outlines the development and validation of the use of an emotion detection algorithm to measure the probability of a perceived emotion from the facial expressions of facial palsy patients when attempting to smile voluntarily. This paper was controversial, as it represented the first time an Artificial Intelligence algorithm had been utilised in the field of facial reanimation surgery, and so it was accompanied by an Invited Discussion at the time of publication and subsequent correspondence was published in Letters to the Editor of PRS. This correspondence and the Response are included at the end of the appendix having been published in *Plastic and Reconstructive Surgery*.

Appendix 7:

Dusseldorp JR, Guarin DL, van Veen MM, Jowett N, Hadlock TA. In the Eye of the Beholder: Changes in Perceived Emotion Expression after Smile Reanimation. *Plast. Recon. Surg*. 2019 Aug; 144(2):457-471

For context, I have included the Invited Discussion written by David Chuang:

Chuang D. In the Eye of the Beholder: Changes in Perceived Emotion Expression after Smile Reanimation. *Plast. Recon. Surg.* 2019 Aug; 144(2):472-473

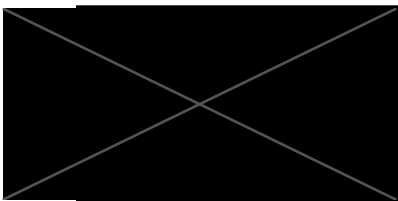
And Letter to the Editor with Response:

Response to Gulbitti, et. al. **Dusseldorp JR**, Hadlock TA. Using artificial intelligence to measure facial expression following facial reanimation surgery. *Plast. Recon. Surg.* 2022 Jan;149(2):343e-345e

Authorship Statement

I, Joseph Dusseldorp, confirm for the ensuing journal article: '**Spontaneity Assessment in Dually Innervated Gracilis Smile Reanimation Surgery**' that I have made the following contributions to its authorship:

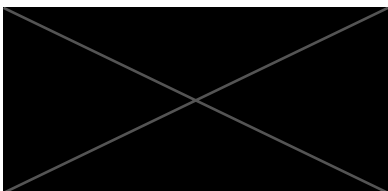
- Conception and design of the research
- Collection and extraction of data
- Analysis and interpretation of the findings
- Drafting and revising of the manuscript and critical appraisal of content
- Responding to multiple stages of peer review



Joseph R. Dusseldorp

Date 14.09. 2023

As Auxiliary Supervisor for the candidature upon which this thesis is based, I can confirm that the above authorship attribution is correct.

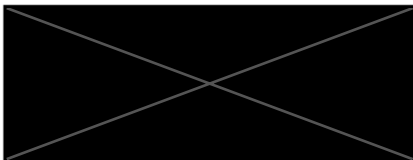


Professor Tessa Hadlock, MD

Date 27.07. 2023

Department of Otolaryngology
Massachusetts Eye and Ear Infirmary
Harvard University, Boston, MA, USA

As Primary Supervisor of the candidate, I concur with the above authorship attestation:



Professor Peter Maitz

Date 14.09.2023

Department of Burns Surgery
Concord Repatriation General Hospital
Sydney University, Sydney, NSW, Australia

Spontaneity Assessment in Dually Innervated Gracilis Smile Reanimation Surgery

Joseph R. Dusseldorp, FRACS; Martinus M. van Veen, MD; Diego L. Guarin, PhD; Olivia Quatela, BS; Nate Jowett, MD; Tessa A. Hadlock, MD



IMPORTANCE Surgeons have sought to optimize outcomes of smile reanimation surgery by combining inputs from nerve-to-masseter and cross-face nerve grafts. An objective assessment tool could help surgeons evaluate outcomes to determine the optimal neural sources for smile reanimation.

OBJECTIVE To evaluate the use of a novel video time-stamping method and standard outcome measurement tools to assess outcomes of facial reanimation surgery using various innervation strategies.

DESIGN, SETTING, AND PARTICIPANTS Cohort study assessing the outcomes of dually innervated gracilis free muscle transfers vs single-source innervated gracilis transfer performed at a tertiary care facial nerve center between 2007 and 2017 using a novel, video time-stamping spontaneity assessment method. The statistical analyses were performed in 2018.

INTERVENTIONS Dually innervated gracilis free muscle transfers or single-source innervated gracilis transfer.

MAIN OUTCOMES AND MEASURES Spontaneous smiling was assessed by clinicians and quantified using blinded time-stamped video recordings of smiling elicited while viewing humorous video clips.

RESULTS This retrospective cohort study included 25 patients (12 men and 13 women; median [range] age, 38.4 [29.3-46.0] years) treated with dually innervated gracilis free functional muscle graft for unilateral facial palsy between 2007 and 2017. Smile spontaneity assessment was performed in 17 patients and was compared with assessment performed in 24 patients treated with single-source innervated gracilis transfer (ie, nerve-to-masseter-driven or cross-face nerve graft-driven gracilis [$n = 13$]) (demographic data not available for NTM and CFNG cohorts). The use of time-stamped video assessment revealed that spontaneous synchronous oral commissure movement in a median percentage of smiles was 33% in patients with dually innervated gracilis (interquartile range [IQR], 0%-71%), 20% of smiles in patients with nerve-to-masseter-driven gracilis (IQR, 0%-50%), and 75% of smiles in patients with cross-face nerve graft-driven gracilis (IQR, 0%-100%). Clinicians graded smile spontaneity in dually innervated cases as absent in 40% ($n = 6$ of 15), trace in 33% ($n = 5$ of 15) and present in 27% ($n = 4$ of 15). No association was demonstrated between clinician-reported spontaneity and objectively measured synchronicity.

CONCLUSIONS AND RELEVANCE Dually innervated gracilis free muscle transfers may improve smile spontaneity compared with masseteric nerve-driven transfers but not to the level of cross-face nerve graft-driven gracilis transfers. Quantifying spontaneity is notoriously difficult, and most authors rely on clinical assessment. Our results suggest that clinicians may rate presence of spontaneity higher than objective measures, highlighting the importance of standardized assessment techniques.

LEVEL OF EVIDENCE 4.

Author Affiliations: Department of Otolaryngology-Head and Neck Surgery, Massachusetts Eye and Ear Infirmary and Harvard Medical School, Boston (Dusseldorp, van Veen, Guarin, Quatela, Jowett, Hadlock); Department of Plastic and Reconstructive Surgery, Royal Australasian College of Surgeons and University of Sydney, Sydney, Australia (Dusseldorp); Department of Plastic Surgery, University Medical Center Groningen, and University of Groningen, Groningen, the Netherlands (van Veen).

Corresponding Author: Joseph R. Dusseldorp, FRACS, Massachusetts Eye and Ear Infirmary, 243 Charles St, Boston, MA 02114 (drjoedusseldorp@icloud.com).

JAMA Facial Plast Surg. 2019;21(6):551-557. doi:10.1001/jamafacial.2019.1090
Published online October 31, 2019.

Two neural sources are widely used to innervate the gracilis free functional muscle transfer (FFMT) in facial reanimation: the ipsilateral nerve to masseter (NTM) or branches of the contralateral facial nerve via a cross-face nerve graft (CFNG). Both sources have distinct advantages and disadvantages. The NTM provides a robust source of motor axons and reliable postoperative oral commissure excursion with voluntary smiling.¹⁻⁴ However, conscious muscle activation is required to produce a smile, facial movements are present with prandial activity, and low rates of spontaneity have been reported.^{5,6} Cross-face nerve graft results in spontaneous movement but can lead to less oral commissure excursion and a higher chance of failure to achieve a meaningful smile on the surgical side.^{1,3,7}

In 2009, Watanabe et al⁸ suggested that combining voluntary motor input from neurotization of the masseter and gracilis with spontaneous movement from the CFNG may improve outcomes in smile reanimation. Several other groups have since performed a wide variety of dual innervation techniques.⁹⁻¹³ Good to excellent results have been reported, although various outcome measures have been used and very few studies have systematically assessed smile spontaneity.

The aim of this study was to develop an objective assessment tool to quantify spontaneous smiling and use this tool to analyze facial reanimation outcomes from various innervation strategies.

Methods

Institutional review board approval from the Massachusetts Eye and Ear Infirmary and patient written informed consent were obtained for this retrospective cohort study. All patients who underwent dually innervated gracilis FFMT for unilateral facial palsy at our institution between 2007 and 2017 were identified ($n = 25$). In addition, 24 patients treated with single-source (NTM- or CFNG-driven) innervated gracilis from the same time period were analyzed. Demographic data were recorded, including age, sex, the cause of facial palsy, follow-up duration at time of spontaneity assessment, and preoperative facial nerve status (flaccid vs nonflaccid). Facial nerve status was defined as long-standing flaccid facial palsy (FFP) if there was no sign of activity in any facial muscle on the affected side of the face for at least 18 months prior to reanimation surgery. The remaining patients were classified as having nonflaccid facial palsy (NFFP).¹⁴ Intraoperative neural coaptation patterns were classified as either interfascicular split, Y-shaped neurotomy, or end-to-side neurotomy (Figure 1).

Standardized outcome measures were collected using a patient-reported facial palsy-specific quality-of-life questionnaire (Facial Clinimetric Evaluation [FaCE] scale),¹⁵ photographs and videos of the face in repose and in 7 standardized facial expressions,¹⁶ clinician-graded electronic facial palsy severity assessment (using the eFACE scale),¹⁷ and objective facial landmark detection and measurement with a machine-learning model graphical user interface (Emotrics).¹⁸ Spontaneous smile was elicited using the Spontaneous Smile

Key Points

Question Can a video time-stamping method be used to measure spontaneity of smile after dually innervated gracilis smile reanimation surgery?

Findings In this cohort study of 25 patients treated with dually innervated gracilis free muscle transfers vs 24 patients treated with masseteric nerve-driven transfers ($n = 11$) or cross-face nerve graft-driven gracilis ($n = 13$), a spontaneous smile was present in a median of 33% of smiles, which was more than was present in masseteric nerve-driven transfers (20%) but less than in cross-face nerve graft-driven smile reanimation surgery (75%).

Meaning Dual innervation may improve synchronicity compared with masseteric nerve transfer but not to the level of cross-face nerve graft-driven gracilis free muscle transfer.

Assay (SSA),¹⁹ a validated technique of obtaining video recordings of elicited spontaneous smiles from patients viewing humorous video clips.

Two techniques were used to quantify spontaneous smiling. First, a blinded assessor analyzed each SSA video clip in which the upper face and the affected side of the lower face were blacked out. The assessor watched the video and time-stamped the filmstrip over the time period that a healthy side commissure movement was clearly elicited (Figure 2). The blackout was then reversed so that only the affected half of the lower face was visible, and oral commissure movements were time-stamped for a second time. The time-stamped videos were compared in a binary fashion (spontaneous smile present vs absent) and a mean of all smile moments was calculated, yielding an overall percentage of smiles within a SSA video that had synchronous oral commissure movements (Video 1 and Video 2). This method is similar to a recently published study that used a slow-motion video analysis method; however, blinding of the 2 sides of the face was not performed in that study.²⁰ Second, the medical record was reviewed to see how the clinician reported the presence or absence of spontaneous smile activity. Clinician-graded spontaneity was categorized as either absent, trace, or present. Correlation between synchronicity and clinician-reported spontaneity was assessed. Synchronicity was compared between dually innervated NTM-driven and CFNG-driven gracilis FFMTs.

Statistical Analysis

Descriptive statistics were expressed as numbers and frequencies and medians and interquartile ranges (IQRs). Differences between preoperative and postoperative data were tested using Wilcoxon signed rank tests. Correlation between synchronicity and clinician-reported spontaneity was tested using Spearman rank correlation. A Kruskal-Wallis test was used to analyze differences in oral commissure excursion on the affected side of the face for 3 neurotomy techniques used in dual innervation. Mann-Whitney U tests were used to test differences in synchronicity and follow-up time between dually innervated and NTM-driven or CFNG-driven gracilis FFMT. An alpha of .05 was selected, and all statistical tests were performed using SPSS version 25 (IBM).

Figure 1. Neuroorrhaphy Techniques

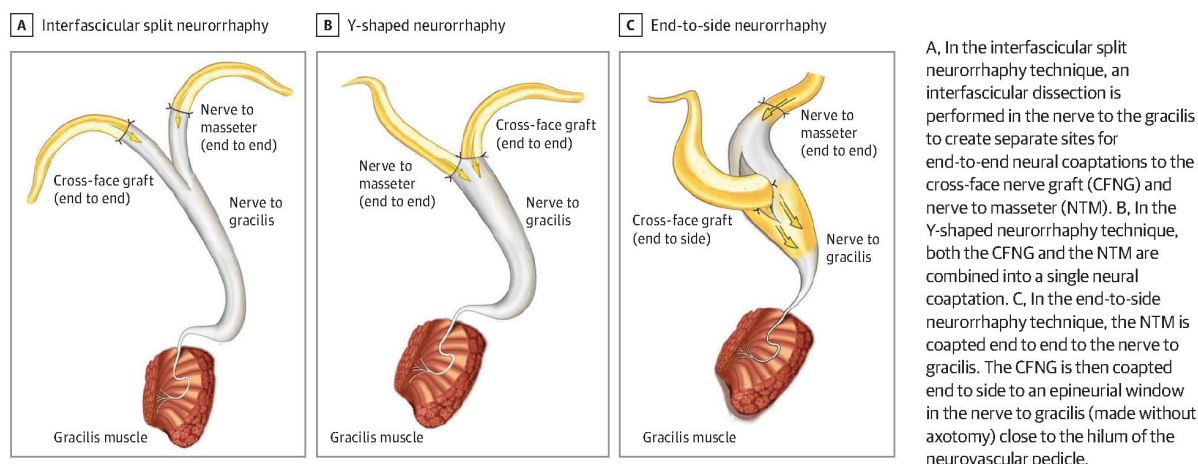
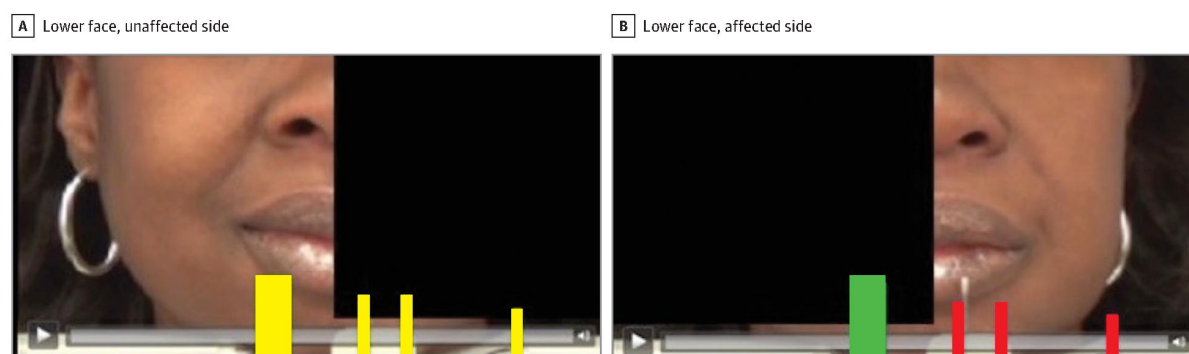


Figure 2. Manual Time-Stamping Technique



A, A blinded assessor analyzed the spontaneous smile assay (SSA) video clip in which the upper face and the affected side of the lower face were blacked out. The assessor time-stamped the filmstrip over the time period that a healthy side commissure movement was clearly elicited (yellow bars). B, The blackout was then reversed so that only the affected half of the lower face was visible, and oral commissure movements were time-stamped for a second time.

Time-stamps were labeled (movement present, green bars; movement absent, red bars) and a mean was calculated, yielding an overall percentage of smiles within the SSA video clip that demonstrated synchronous oral commissure movements. Synchronicity in this case was 25% (1 of 4 smiles). See Video 1 and Video 2 for further demonstration of this technique.

Results

Demographics

Dually innervated gracilis FFMT for unilateral facial palsy was performed in 37 patients between 2007 and 2017 (29 patients between 2016 and 2017, and 8 between 2007 and 2015). Eight patients had insufficient follow-up of less than 6 months at time of spontaneity assessment and 4 patients were lost to follow-up; thus, 25 patients (12 men and 13 women; median [range] age, 38.4 [29.3-46.0] years) in the dual innervation group and 24 patients (demographic data not available for NTM and CFNG cohorts) in the single-source innervation group were included in the final analysis. Median postoperative follow-up was 0.8 (IQR, 0.6-1.4) years (Table 1). Most patients (15 [60%] of 25) presented with facial palsy after tumor or trauma. Seventeen (68%) of 25 patients had FFP and 8 (32%) had NFFP. An interfascicular split neuroorrhaphy technique was used to

perform dual innervation in 15 (60%) patients, Y-shaped neuroorrhaphy was used in 6 (24%), and the end-to-side technique was used in 4 (16%) (Figure 1).

Outcome Measures

Preoperative and postoperative FaCE scale scores were available for 15 (60%) of 25 patients, and patient-reported quality-of-life scores rose from a preoperative median of 38 points (IQR, 35-53) to a postoperative median of 51 points (IQR, 38-70) ($P = .004$). Significant increases were found in static, dynamic, and total eFACE scores (Table 2). Excursion of the affected corner of the mouth with smile improved significantly from a preoperative median (IQR) of -1.2 mm (-2.6 to 2.4) to a postoperative median of 5.8 mm (IQR, 2.4 - 6.4) ($P < .001$). Additionally, symmetry of the mouth in repose and with smile improved significantly (Table 2). No difference in postoperative oral commissure excursion was found between the 3 different dual innervation neuroorrhaphy techniques.

Spontaneity Assessment

The SSA was performed in 19 (76%) of 25 patients treated with a dually innervated technique.¹⁹ Of these, 17 (89%) of 19 patients generated a spontaneous smile during the video that could be analyzed. The median (IQR) percentage of synchronous oral commissure movements in dually innervated patients was 33.0% (0%-71%). Median (IQR) synchronicity in FFP was 16.5% (0%-64%) and NFFP was 60.0% (0%-100%), although this difference was not statistically significant ($P = .33$, Mann-Whitney U test). The SSA was performed in dual innervation after a median (IQR) follow-up of 0.8 (0.7-1.1) years. Synchronous movement was not correlated with the duration of follow-up.

Table 1. Characteristics, Follow-up Time, and Innervation Techniques in Patients Treated With Dual Innervation

Characteristic	No. of Patients n = 25
Sex	
Female	13
Male	12
Median age (range), y	38.4 (29.3-46.0)
Affected side	
Left	14
Right	11
Etiology	
Acoustic neuroma resection	6
Iatrogenic injury/trauma	5
Benign tumor resection	4
Viral paralysis	4
Congenital	1
Other	5
Duration of paralysis, median (IQR), y	4.1 (2.4-12.9)
Flaccid facial palsy	17
Nonflaccid facial palsy	8
Follow-up time, median (IQR), y	0.8 (0.6-1.4)
Innervation technique	
Interfascicular split	15
Y-shaped neuroorrhaphy	6
End-to-side	4

Abbreviation: IQR, interquartile range.

Clinician-assessed spontaneity was documented in 15 (88%) of 17 medical records of patients who were noted to have smiled during the SSA. Of these, clinician-graded spontaneity was rated as absent in 40.0% ($n = 6$ of 15), trace in 33.3% ($n = 5$ of 15), and present in 26.7% ($n = 4$ of 15). No significant correlation between synchronicity and clinician-graded spontaneity was noted (correlation coefficient, 0.347; $P = .20$).

Comparison of CFNG, Dual Innervation, and NTM

Synchronous oral commissure excursion elicited with humorous video clips occurred variably in patients treated with CFNG-driven, dually innervated, and NTM-driven gracilis FFMTs. The use of time-stamped video assessment revealed that spontaneous synchronous oral commissure movement occurred in a median percentage of 33% in patients with dually innervated gracilis (interquartile range [IQR], 0%-71%; $n = 17$ patients), 20% of smiles in patients with NTM-driven gracilis (IQR, 0%-50%; $n = 11$ patients), and 75% of smiles in patients with CFNG-driven gracilis (IQR, 0%-100%; $n = 13$ patients) (Figure 3). No statistical differences among groups were detected. Synchronous movement analysis was performed for CFNG-driven, dually innervated, and NTM-driven gracilis FFMTs at a median (IQR) follow-up of 1.3 (1.0-2.0) years, 0.8 (0.7-1.0) years, and 1.4 (0.4-2.0) years, respectively. There was a significantly shorter follow-up time for dually innervated cases than CFNG-driven gracilis FFMTs. Synchronous movement was more dichotomous for CFNG-driven gracilis FFMTs than dually innervated and NTM-driven gracilis FFMTs, yielding high levels of synchronicity when any movement was present, and no synchronicity in cases of complete failure of reanimation. A wider spread of synchronicity was present in dually innervated or NTM-driven cases (Figure 3).

Discussion

Several studies using various dual innervation techniques have been published since the concept was first described by Watanabe et al⁸ in 2009, and some excellent results have been reported.⁹⁻¹³ The most commonly reported outcome

Table 2. Patient-Reported Outcome Measure Improvements, Clinician-Graded Improvements, and Objective Smile Measurements Following Dual Innervation

Outcome Measure	Median (IQR)		P Value
	Preoperative	Postoperative	
FaCE scale total scores ^a	38 (35 to 53)	51 (38 to 70)	.01
eFace scores ^b			
Total	60 (57 to 66)	80 (73.3 to 83.8)	<.001
Static	74 (60 to 82.5)	88 (84.3 to 92)	<.001
Dynamic	35 (27 to 47.5)	65.5 (58.3 to 74)	<.001
Smile measurements, ^c mm			
Excursion healthy side	11.4 (8.6 to 12.8)	9.1 (5.6 to 11.7)	.05
Excursion affected side	-1.2 (-2.6 to 2.4)	5.8 (2.4 to 6.4)	<.001
Symmetry repose	5.1 (-0.3 to 10.6)	-0.6 (-4.3 to 2.7)	.002
Symmetry smile	15.7 (10.7 to 21.7)	3.2 (-0.1 to 6.7)	<.001

Abbreviations: FaCE, Facial Clinimetric Evaluation scale; IQR, interquartile range; mm, millimeters.

^a Patient-reported outcome measure.

^b Clinician-graded improvement.

^c Emotrics facial recognition software.

assessment tool is the Terzis and Noah scale,²⁰ which consists of a clinician-graded assessment of aesthetic and functional outcome that ranges from poor to excellent. Spontaneity of smile is not included in this grading system. In numerous case series, spontaneous smile was reported to be present in nearly all cases following dual innervation, as assessed by the treating physicians.⁹⁻¹³ In 2015, Sforza et al¹³ systematically analyzed dually innervated gracilis FFMT using a humorous video and motion tracking technology and found spontaneous smiling to be present in 9 (70%) of 13 cases. No data were presented on the duration of palsy prior to surgery or the chance of spontaneous recovery. Dual innervation was performed in a single stage, and it was unclear whether any other nerve transfer procedure (ie, direct CFNG) was simultaneously performed.

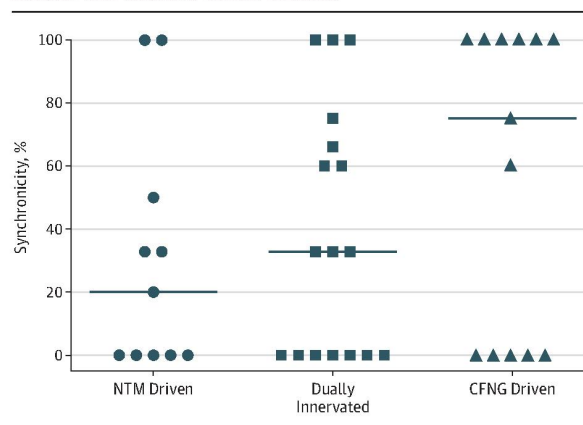
Initially, we preferred to perform 2-stage CFNG, routinely assessing for the presence of a Tinel sign at the tip of the CFNG before performing the second-stage gracilis FFMT. We performed our first case of dually innervated gracilis FFMT for smile reanimation in 2007. At that time, our indication for using dual neural sources was to optimize axonal input when questionable axonal supply was suspected within the CFNG (eg, owing to weak evoked stimulation in the donor nerve branch, poor Tinel sign, absence of a neuroma at the tip of the CFNG, or if CFNG had been used elsewhere). In order to combine 2 neural inputs, we elected to perform an interfascicular split of the anterior branch of the obturator nerve to gracilis so that 2 end-to-end neuroorrhaphies could be performed with the laterally positioned NTM and the short CFNG located medially in the upper lip (Figure 1A).

Later in our experience, we offered dual innervation to older patients as a primary innervation strategy. Our goal was to augment the reliability of a NTM-driven gracilis FFMT smile with an element of the spontaneity generated from the use of CFNG. We wanted to improve on the spontaneity generated using the interfascicular split technique and decided to use a longer CFNG in the first stage, extending the graft across the face to 1 cm in front of the tragus, rather than banking it in the upper lip, as is our preference when performing gracilis transfer by using CFNG alone. This long CFNG was then coapted with the NTM in a single, Y-shaped neural coaptation to the nerve to gracilis (Figure 1B).

Our surgical technique has recently transitioned to performing a true end-to-side neuroorrhaphy (also known as “supercharged” or “reverse end to side”) instead of a Y-shaped neuroorrhaphy based on outcomes experienced by others (F. Biglioli, MD, personal communication, August 20, 2017). During this approach, the NTM is coapted end-to-end to the nerve to gracilis. The CFNG is then coapted end-to-side to an epineurial window in the nerve to gracilis—made without axotomy—at the hilum of the neurovascular pedicle (Figure 1C). We recommend that all patients receive postoperative physical therapy irrespective of the type of smile reanimation technique used.

We sought to assess these dual innervation techniques by using our routine panel of outcome assessment tools, including patient-reported outcome measures, clinician-graded scoring systems, and objective measurements (such

Figure 3. Synchronicity in Nerve-to-Masseter (NTM)-Driven, Dually Innervated, and Cross-Face Nerve Graft (CFNG)-Driven Gracilis Free Functional Muscle Transfers



Horizontal lines represent the median values; geometric symbols represent individual patients.

as oral commissure excursion) (Table 2). No significant differences were found between innervation techniques in terms of eFACE score improvements or oral commissure excursion.

Moreover, because the real advantage of dual innervation over NTM-driven gracilis is the presence of smile spontaneity, we sought to develop a blinded, objective measurement system to quantify the presence of synchronous smile. Although time-consuming, manual time-stamping surpassed our ability to detect spontaneity from a clinical interaction with patients. Our own clinician gradings showed no correlation with manual time-stamping in this group of dually innervated patients. This finding highlights the difficulty of quantifying spontaneous smile using a clinician-grading system alone, and it implies that there may be inherent bias in clinician assessments. Therefore, objective measurement methods of spontaneity are essential.

Synchronous oral commissure movements and spontaneous, emotional smiling are not synonymous. Synchronous oral commissure movements are a measure of how well the affected side smile is linked to the normal side via a CFNG. Spontaneous, emotional smile is much more complex, and although it is well recognized by humans, it can be difficult to quantify for study purposes. Some groups have used layperson assessment to quantify the ability of a smile to convey emotion or communicate nonverbally. For the purposes of this study, we were interested in the additional use of the CFNG and have used synchronous oral commissure movements as a proxy for spontaneous smiling. One problem with this definition is that oral commissure movements on the affected side may also occur with blink on the healthy side, when a suboptimal donor facial nerve branch has been used (Video 2). Although this synchronous movement implies the presence of successful innervation through the CFNG, it does not necessarily give information as to the ability of the smile in conveying emotion or non-verbal information. Other computer-learning techniques

quantifying the emotionality of smiling are currently being developed for this purpose.²¹

We found that synchronous movements were present in a median 33% of spontaneous smiles in our dual innervated cases, with a wide range (0% to 100%), indicating variable degrees of innervation by the CFNG or varying degrees of NTM-driven spontaneity. Interestingly, we found synchronous movement to be present in median 75% (IQR, 0%-100%) of the CFNG-driven cases. We believe that this may be owing to lower axon counts crossing the face and a weaker muscle response, especially when compared with NTM-driven gracilis muscles. Synchronous oral commissure movements were present in median 20% of smiles in NTM-driven gracilis (IQR, 0%-50%).

Existing literature suggests that dual innervation is useful for achieving both voluntary bite-driven smile and spontaneous cross-face-driven smile.⁹⁻¹³ Our experience is less convincing; we have performed many recent cases using a true end-to-side neurotaphy technique but these cases have not yet had sufficient follow-up to be analyzed in this study. Overall, dual innervation performed at least as well as a typical gracilis by NTM alone. Cross-face nerve graft-driven gracilis transfer continues to be the most reliable technique for recreating a spontaneous smile, but may produce less oral commissure excursion than NTM-driven gracilis and has an overall higher failure rate.^{3,4,22}

The scientific basis of dual innervation is not well understood. Since Viterbo et al²³ reintroduced the concept of end-to-side neurotaphy in 1992²⁴ there has been significant interest in optimizing reinnervation of a target muscle using a reverse end-to-side²⁵ (also known as “supercharged” end-to-side²⁶) technique. The clinical benefits of early reinnervation of the hand intrinsic muscles by supercharging end-to-side the deep motor branch of the ulnar nerve with a distal anterior interosseus nerve transfer have been reported.^{26,27} However, it is less clear whether this clinical model in fact constitutes dual innervation because regeneration of the ulnar nerve from the elbow to the hand is unreliable, and the contributions of individual nerve sources are difficult to determine clinically.²⁶ An interesting biological phenomenon to understand is what happens when 2 fronts of regenerating axons come into contact within a regenerating nerve, or when 1 axon arrives at the neuromuscular junction and there is another axon present. Does ephaptic conduction occur between regenerated axons from different origins²⁸ or does one population of axons gain dominance by a process of synaptic elimination, as has been demonstrated in developing neuromuscular junctions in vivo by Turney and Lichtman²⁹ and others?³⁰ These questions remain unanswered.

In the largest previously reported case series of dually innervated gracilis FFMT, CFNG and gracilis FFMT were performed in a single-stage procedure using a true end-to-side innervation technique.¹³ The authors assert that preserving the gracilis motor end plates by the NTM extends the window of muscle receptivity for the time it takes for the CFNG to regenerate across the face. Whether it is possible for axons to arrive at the neuromuscular junction several months after another

front of regenerating axons and either displace existing axons or create new neuromuscular junctions has not yet been investigated in animal models. Further animal studies and rigorous analysis of clinical data may improve our understanding of this biological process.

A significant proportion of healthy patients are known to co-contract the masseter muscle with smiling.³⁰ Lenz et al²⁰ found that patients with preoperative co-contraction of the zygomaticus and masseter muscles who underwent NTM-driven smile reanimation often developed a synchronous smile postoperatively. Theoretically, this finding could mean that some of the spontaneous movement in dual innervation comes from NTM input rather than CFNG input. Conversely, we recently studied patient-reported smile perception in NTM-driven smile reanimation surgery and found that self-reported spontaneous smile rates and NTM-driven smile use were low (<50%).³¹

Limitations

The present study was designed to demonstrate the utility of an objective approach to quantify spontaneous smiling; however, we did not have sufficient power to be able to make definitive conclusions on ideal neural sources in facial reanimation. Other limitations of this study include its retrospective nature, predominantly short-term follow-up (0.8 years for dual innervation vs 1.4 years for CFNG), variation in the indication for dual innervation, various neurotaphy techniques used, and the inclusion of patients with FFP and NFFP. It is possible that the combination of patients with FFP and NFFP makes data interpretation more difficult because the remaining ipsilateral facial nerve function may be mistaken for CFNG-driven movement. However, in the current era, early intervention is becoming more common, and patients with long-standing, unoperated, FFP are rare. Objective assessment tools are required to be able to determine subtle differences in these complex patients. In the future, surface electromyography measurements and local anesthetic injections in the region of the CFNG may also prove useful to assess the presence or absence of a neural connection between the CFNG and the gracilis muscle, although we have not yet used these techniques.

Conclusions

We developed a blinded, objective assessment of synchronous smile and found that synchronous oral commissure movement was present in approximately one-third of the patient cohort in the present study of dually innervated facial reanimations. Additional studies of dually innervated cases with long-term follow-up and larger sample sizes are required to verify any additional utility of the dual innervation technique over the NTM-driven gracilis technique. Time-stamping is a simple and effective method of quantifying synchronicity of oral commissure movement and may help to determine the effectiveness of dual innervation at creating spontaneity, the most elusive attribute of the reanimated smile.

ARTICLE INFORMATION

Accepted for Publication: September 7, 2019.

Conflict of Interest Disclosures: Dr Jowett reported a patent to 16/069341 pending. No other disclosures were reported.

Published Online: October 31, 2019.
doi:10.1001/jamafacial.2019.1090

Author Contributions: Dr Dusseldorp had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Dusseldorp, van Veen, Guarin, Hadlock.

Acquisition, analysis, or interpretation of data: Dusseldorp, van Veen, Quatela, Jowett, Hadlock.

Drafting of the manuscript: Dusseldorp, van Veen, Quatela, Hadlock.

Critical revision of the manuscript for important intellectual content: Dusseldorp, Guarin, Jowett, Hadlock.

Statistical analysis: van Veen.

Obtained funding: Hadlock.

Administrative, technical, or material support:

Dusseldorp, Guarin, Quatela, Hadlock.

Study supervision: Dusseldorp, Jowett, Hadlock.

Meeting Presentation: Elements of this article were presented at the International Facial Nerve Symposium (IFNS); August 3 to 6, 2017, Los Angeles, California.

Additional Contributions: We thank the patients for granting permission to publish this information.

REFERENCE LIST

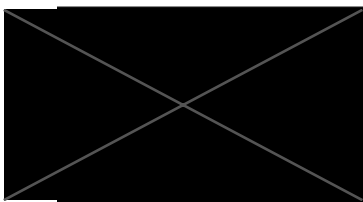
- Bae YC, Zuker RM, Manktelow RT, Wade S. A comparison of commissure excursion following gracilis muscle transplantation for facial paralysis using a cross-face nerve graft versus the motor nerve to the masseter nerve. *Plast Reconstr Surg*. 2006;117(7):2407-2413. doi:10.1097/01.prs.0000218798.95027.21
- Thibaut A, Chatelle C, Ziegler E, Bruno MA, Laureys S, Gosseries O. Spasticity after stroke: physiology, assessment and treatment. *Brain Inj*. 2013;27(10):1093-1105. doi:10.3109/02699052.2013.804202
- Hontanilla B, Marre D, Cabello A. Facial reanimation with gracilis muscle transfer neurotized to cross-facial nerve graft versus masseteric nerve: a comparative study using the FACIAL CLIMA evaluating system. *Plast Reconstr Surg*. 2013;131(6):1241-1252. doi:10.1097/PRS.0b013e31828bd4da
- Snyder-Warwick AK, Fattah AY, Zive L, Halliday W, Borschel GH, Zuker RM. The degree of facial movement following microvascular muscle transfer in pediatric facial reanimation depends on donor motor nerve axonal density. *Plast Reconstr Surg*. 2015;135(2):370e-381e. doi:10.1097/PRS.0000000000000860
- Hontanilla B, Cabello A. Spontaneity of smile after facial paralysis rehabilitation when using a non-facial donor nerve. *J Craniomaxillofac Surg*. 2016;44(9):1305-1309. doi:10.1016/j.jcms.2016.06.031
- Rozen S, Harrison B. Involuntary movement during mastication in patients with long-term facial paralysis reanimated with a partial gracilis free neuromuscular flap innervated by the masseteric nerve. *Plast Reconstr Surg*. 2013;132(1):110e-116e. doi:10.1097/PRS.0b013e318290f644
- Bhama PK, Weinberg JS, Lindsay RW, Hohman MH, Cheney ML, Hadlock TA. Objective outcomes analysis following microvascular gracilis transfer for facial reanimation: a review of 10 years' experience. *JAMA Facial Plast Surg*. 2014;16(2):85-92. doi:10.1001/jamafacial.2013.2463
- Watanabe Y, Akizuki T, Ozawa T, Yoshimura K, Agawa K, Ota T. Dual innervation method using one-stage reconstruction with free latissimus dorsi muscle transfer for re-animation of established facial paralysis: simultaneous reinnervation of the ipsilateral masseter motor nerve and the contralateral facial nerve to improve the quality of smile and emotional facial expressions. *J Plast Reconstr Aesthet Surg*. 2009;62(12):1589-1597. doi:10.1016/j.bjps.2008.07.025
- Okazaki M, Kentaro T, Noriko U, et al. One-stage dual latissimus dorsi muscle flap transfer with a pair of vascular anastomoses and double nerve suturing for long-standing facial paralysis. *J Plast Reconstr Aesthet Surg*. 2015;68(6):e113-e119. doi:10.1016/j.bjps.2015.02.013
- Biglioli F, Bayouth W, Colombo V, Pedrazzoli M, Rabbiosi D. Double innervation (facial/masseter) on the gracilis flap, in the middle face reanimation in the management of facial paralysis: a new concept [in French]. *Ann Chir Plast Esthet*. 2013;58(2):89-95. doi:10.1016/j.anplas.2012.12.001
- Cardenas-Mejia A, Covarrubias-Ramirez JV, Bello-Margolis A, Rozen S. Double innervated free functional muscle transfer for facial reanimation. *J Plast Surg Hand Surg*. 2015;49(3):183-188. doi:10.3109/2000656X.2014.988218
- Biglioli F, Colombo V, Tarabba F, et al. Double innervation in free-flap surgery for long-standing facial paralysis. *J Plast Reconstr Aesthet Surg*. 2012;65(10):1343-1349. doi:10.1016/j.bjps.2012.04.030
- Sforza C, Frigerio A, Mapelli A, et al. Double-powered free gracilis muscle transfer for smile reanimation: a longitudinal optoelectronic study. *J Plast Reconstr Aesthet Surg*. 2015;68(7):930-939. doi:10.1016/j.bjps.2015.03.029
- Lindsay RW, Bhama P, Weinberg J, Hadlock TA. The success of free gracilis muscle transfer to restore smile in patients with nonflaccid facial paralysis. *Ann Plast Surg*. 2014;73(2):177-182. doi:10.1097/SAP.0b013e3182a0df04
- Kahn JB, Gliklich RE, Boyev KP, Stewart MG, Metson RB, McKenna MJ. Validation of a patient-graded instrument for facial nerve paralysis: the FaCE scale. *Laryngoscope*. 2001;111(3):387-398. doi:10.1097/00005537-200103000-00005
- Santosa KB, Fattah A, Gavilán J, Hadlock TA, Snyder-Warwick AK. Photographic standards for patients with facial palsy and recommendations by members of the Sir Charles Bell Society. *JAMA Facial Plast Surg*. 2017;19(4):275-281. doi:10.1001/jamafacial.2016.1883
- Banks CA, Bhama PK, Park J, Hadlock CR, Hadlock TA. Clinician-graded electronic facial paralysis assessment: the eFACE. *Plast Reconstr Surg*. 2015;136(2):223e-230e. doi:10.1097/PRS.0000000000001447
- Guarin DL, Dusseldorp J, Hadlock TA, Jowett N. A machine learning approach for automated facial measurements in facial palsy. *JAMA Facial Plast Surg*. 2018;20(4):335-337. doi:10.1001/jamafacial.2018.0030
- Iacolucci CM, Banks C, Jowett N, et al. Development and validation of a spontaneous smile assay. *JAMA Facial Plast Surg*. 2015;17(3):191-196. doi:10.1001/jamafacial.2015.0083
- Lenz Y, Kiefer J, Dietrich F, Stark GB, Eisenhardt SU. Pre-operative masseter muscle EMG activation during smile predicts synchronicity of smile development in facial palsy patients undergoing reanimation with the masseter nerve: a prospective cohort study. *J Plast Reconstr Aesthet Surg*. 2019;72(3):505-512. doi:10.1016/j.bjps.2018.11.011
- Dusseldorp JR, Guarin DL, van Veen MM, Jowett N, Hadlock TA. Emotion expression following smile reanimation surgery. Presented at: European Association of Plastic Surgeons (EURAPS) 29th Annual Meeting; May 17-19, 2018; Madrid, Spain.
- Viterbo F, Trindade JC, Hoshino K, Mazzoni Neto A. Latero-terminal neurotomy without removal of the epineurial sheath: experimental study in rats. *Rev Paul Med*. 1992;110(6):267-275.
- Viterbo F, Trindade JC, Hoshino K, Mazzoni Neto A. End-to-side neurotomy with removal of the epineurial sheath: an experimental study in rats. *Plast Reconstr Surg*. 1994;94(7):1038-1047. doi:10.1097/00006534-199412000-00019
- Isaacs J, Allen D, Chen LE, Nunley J II. Reverse end-to-side neurotization. *J Reconstr Microsurg*. 2005;21(1):43-48. doi:10.1055/s-2005-862780
- Barbour J, Yee A, Kahn LC, Mackinnon SE. Supercharged end-to-side anterior interosseous to ulnar motor nerve transfer for intrinsic musculature reinnervation. *J Hand Surg Am*. 2012;37(10):2150-2159. doi:10.1016/j.jhsa.2012.07.022
- Davidge KM, Yee A, Moore AM, Mackinnon SE. The supercharge end-to-side anterior interosseous-to-ulnar motor nerve transfer for restoring intrinsic function: clinical experience. *Plast Reconstr Surg*. 2015;136(3):344e-352e. doi:10.1097/PRS.00000000000001514
- Lichtman JW, Livet J, Sanes JR. A technicolour approach to the connectome. *Nat Rev Neurosci*. 2008;9(6):417-422. doi:10.1038/nrn2391
- Terzis JK, Noah ME. Analysis of 100 cases of free-muscle transplantation for facial paralysis. *Plast Reconstr Surg*. 1997;99(7):1905-1921. doi:10.1097/00006534-199706000-00016
- Turney SG, Lichtman JW. Reversing the outcome of synapse elimination at developing neuromuscular junctions in vivo: evidence for synaptic competition and its mechanism. *PLoS Biol*. 2012;10(6):e1001352. doi:10.1371/journal.pbio.1001352
- Schaverien M, Moran G, Stewart K, Addison P. Activation of the masseter muscle during normal smile production and the implications for dynamic reanimation surgery for facial paralysis. *J Plast Reconstr Aesthet Surg*. 2011;64(12):1585-1588. doi:10.1016/j.bjps.2011.07.012
- van Veen MM, Dusseldorp JR, Quatela O, et al. Patient experience in nerve-to-masseter-driven smile reanimation. *J Plast Reconstr Aesthet Surg*. 2019;72(8):1265-1271. doi:10.1016/j.bjps.2019.03.037

**Automated Spontaneity Assessment after
Smile Reanimation:
A Machine Learning Approach**

Authorship Statement

I, Joseph Dusseldorp, confirm for the ensuing journal article: ‘**Automated Spontaneity Assessment after Smile Reanimation: A Machine Learning Approach**’ that I have made the following contributions to its authorship:

- Conception and design of the research
- Collection and extraction of data
- Analysis and interpretation of the findings
- Drafting and revising of the manuscript and critical appraisal of content
- Responding to multiple stages of peer review



Joseph R. Dusseldorp

Date 14.09. 2023

As Auxiliary Supervisor for the candidature upon which this thesis is based, I can confirm that the above authorship attribution is correct.

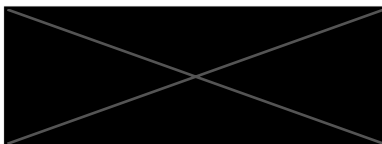


Professor Tessa Hadlock, MD

Date 27.07. 2023

Department of Otolaryngology
Massachusetts Eye and Ear Infirmary
Harvard University, Boston, MA, USA

As Primary Supervisor of the candidate, I concur with the above authorship attestation:



Professor Peter Maitz

Date 14.09.2023

Department of Burns Surgery
Concord Repatriation General Hospital
Sydney University, Sydney, NSW, Australia

Automated Spontaneity Assessment after Smile Reanimation: A Machine Learning Approach

Joseph R. Dusseldorp,
F.R.A.C.S.

Diego L. Guarin, Ph.D.

Martinus M. van Veen, M.D.

Matt Miller, M.D.

Nate Jowett, M.D.

Tessa A. Hadlock, M.D.

Boston, Mass.; Sydney, New South Wales,
Australia; and Groningen,
The Netherlands



Background: Recreation of a spontaneous, emotional smile remains a paramount goal of smile reanimation surgery. However, optimal techniques to reliably restore spontaneity remain unknown. Dual automated machine-learning tools were used to develop an objective tool to analyze spontaneous smiling. The feasibility of this tool was tested in a sample of functional free muscle transfers.

Methods: Validated humorous videos were used to elicit spontaneous smiles. Automated facial landmark recognition (Emotrics) and emotion detection software (Affdex) were used to analyze video clips of spontaneous smiling in nine normal subjects and 39 facial reanimation cases. Emotionality quotient was used to quantify the ability of spontaneous smiles to express joy.

Results: The software could analyze spontaneous smiling in all subjects. Spontaneous smiles of normal subjects exhibited median 100 percent joy and 0 percent negative emotion (emotional quotient score, +100/0). Spontaneous smiles of facial palsy patients after smile reanimation, using cross-facial nerve graft, masseteric nerve, and dual innervation, yielded median emotional quotient scores of +82/0, 0/-48, and +10/-24 respectively (joy, $p = 0.006$; negative emotion, $p = 0.034$).

Conclusions: Computer vision software can objectively quantify spontaneous smiling outcomes. Of the retrospective sample of cases reviewed in this study, cross-facial nerve graft-innervated gracilis functional free muscle transfer achieved a greater degree of emotionality during spontaneous smiling than masseteric or dually innervated transfer. Quantification of spontaneous smiling from standard video clips could facilitate future, blinded, multicenter trials with sufficient long-term follow-up to definitively establish the rates of spontaneity from a range of reanimation procedures. (*Plast. Reconstr. Surg.* 149: 1393, 2022.)

CLINICAL QUESTION/LEVEL OF EVIDENCE: Diagnostic, IV.

Restoration of a normal smile is a critical priority to facial palsy patients and facial reanimation specialists alike.^{1,2} Although voluntary smiling can be reliably reconstructed using a variety of

methods, reanimation and quantification of spontaneous or emotional smiling remains challenging. A recent systematic review and meta-analysis of spontaneity outcomes found that cross-face nerve graft-innervated facial reanimation leads to the highest rates of spontaneous smiling, followed by nonfacial nerve-innervated solutions such as hypoglossal, masseteric, and accessory nerves. However, two-thirds of studies did not report their method of

From the Department of Otolaryngology/Head and Neck Surgery, Massachusetts Eye and Ear Infirmary and Harvard Medical School; Department of Plastic and Reconstructive Surgery, Royal Australasian College of Surgeons and University of Sydney; and Department of Plastic Surgery, University Medical Center Groningen and University of Groningen.

Received for publication April 9, 2020; accepted August 4, 2021.

Presented in part at the 29th Annual Meeting of the European Association of Plastic Surgery Meeting, in Madrid, Spain, May 17 through 19, 2018; and the 97th Annual Scientific Congress of the Surgical Research Society of the Royal Australasian College of Surgeons, in Sydney, New South Wales, Australia, May 7 through 11, 2018.

Copyright © 2022 by the American Society of Plastic Surgeons
DOI: 10.1097/PRS.00000000000009167

Disclosure: None of the authors has a financial interest in any of the products, devices, or drugs mentioned in this article. Dr. Guarin received funding from "Fonds Quebecoise de la recherche sur la nature et technologies" for his postdoctoral work.

Related digital media are available in the full-text version of the article on www.PRSJournal.com.

assessing spontaneity, and of those that did, clinical observation and patient self-reporting were the most common forms of assessment. A prior study found no correlation between clinician-rated spontaneity and objectively measured synchronous oral commissure movement.³ Furthermore, clinician- and patient-rated spontaneity are subject to positive confirmation bias, as both parties have invested considerable time and effort in achieving the desired outcome.³ An automated and objective system for quantification of smile spontaneity would be a welcome tool in the field of facial reanimation, enabling comparison between various innervation techniques across various centers.

Computer vision facial recognition software can track facial landmarks and quantify the ability to express joy in the smiles of facial palsy patients.^{4,5} Machine learning algorithms can also be readily converted into applications deployable on smartphones. As telehealth becomes more commonplace in the era of coronavirus disease of 2019 and with the advent of value-based health care, we anticipate that digital solutions that can objectively quantify outcomes will become more important than ever. The aim of this study was to develop an automated system to quantify smile spontaneity in unilateral facial palsy and to test its feasibility using a consecutive sample of gracilis functional free muscle transfer patients.

PATIENTS AND METHODS

This study was approved by the institutional review board of the Massachusetts Eye and Ear Infirmary. Written consent was obtained from all patients before their inclusion in this study.

Elicitation of Spontaneous Smile Events

All patients were filmed from the frontal aspect while watching a validated series of humorous video clips: the Spontaneous Smile Assay.⁶ Two manual steps were required before automated analysis: the clinician designated which side was the healthy side, and then screened each video clip to ensure there was at least one significant oral commissure movement on the healthy side. If the Spontaneous Smile Assay did not elicit an emotional response on the healthy side, these video clips were excluded from analysis. Video clips of the facial expressions were then automatically analyzed by computer software.

Automated Time-Stamping of Spontaneous Smile Events

Healthy side oral commissure movement was used to automatically define, or time-stamp, the

portions of the video clip of interest for analysis, referred to as spontaneous smile events.³ To define the optimal duration of a spontaneous smile event, we analyzed voluntary and spontaneous smiling across a range of normal ($n = 9$) and facial palsy subjects ($n = 39$), and tested various parameters to determine what percentage of healthy side oral commissure excursion and duration of movement was required to capture at least one spontaneous smile event and not capture other facial movements such as lip puckering, lip biting, or cheek twitching. Clinically significant spontaneous smile events were noted to last longer than 1 second and contain a healthy side oral commissure displacement of more than 80 percent of normalized maximal displacement. Each facial frame within this period was then selected for analysis by dual machine learning algorithms: (1) facial landmarking and (2) emotion detection software.

Automatic Facial Landmarking

The freely available Emotrics application⁴ was used to automatically track facial landmarks in each frame of the video clip, enabling a plot of healthy and affected side oral commissure deviations to be generated over time. Postprocessing of data was performed with MATLAB (MathWorks, Natick, Mass.). Using the established threshold of 80 percent of maximal excursion within that video clip on the healthy side, the frame numbers of onset and offset of all spontaneous smile events occurring during the video clip were automatically generated. To validate this automated time-stamping approach, manual time-stamping was performed for all video clips using previously described methods,³ and differences between time-stamps and number of spontaneous smile events were recorded.

Emotionality Assessment

Using previously reported methods,^{3,5} Affectiva (Affectiva, Boston, Mass.) was used to analyze time-stamped spontaneous smile events, and mean spontaneous emotionality from the entire video clip was calculated. Emotionality was determined using a descriptive means of reporting relative probabilities, the emotionality quotient,⁵ defined as the percentage probability of perceived joy (range, 0 to 100 percent) over the percentage probability of perceived negative emotions (the higher of contempt or disgust, ranging from -100 to 0 percent) during smiling, as predicted by computer vision artificial intelligence software. Mean percentage probabilities of the algorithm

perceiving joy, contempt, and disgust were calculated for all spontaneous smile events during the entire video clip. The number of spontaneous smile events was also recorded.

Ranges of Emotionality Quotient Scores in Normal, Preoperative, and Postoperative Spontaneous Smiles

To test the utility of this machine learning approach, we tested the software on a series of normal control video clips and a consecutive selection of unilateral facial palsy patients who had Spontaneous Smile Assay video clips at least 6 months after surgery available in our database (Figs. 1 and 2). [See Figure, Supplemental Digital Content 1, which shows (left) a single frame of a video clip of a normal subject filmed while smiling spontaneously in response to humorous content. (Right) Outputs of two machine learning algorithms analyzing the video clip. (Above, right, and center, right) Output of Emotrics, an automated facial landmark tracking algorithm. Oral commissure excursion was tracked and normalized on the y axis for both the right and left sides of the face. The threshold of 0.8 is shown by the horizontal black line. Blue shaded areas represent oral commissure excursion above threshold; each of these

frames is automatically time-stamped for further analysis. (Below, right) Probability of joy expression for each frame of the video clip using the Affdex Emotion artificial intelligence algorithm (only every third frame of the video clip was sampled by this algorithm, thus the differing frame reference numbers). The large gray box depicts the playhead of the filmstrip indicating which frame is currently being analyzed from the video clip. The full video is available to view in Video 1, <http://links.lww.com/PRS/F65>. See Video 1 (online), which shows a normal subject smiling spontaneously in response to humorous content. Emotionality quotient score was +100/0, as expected. Graphic outputs of two machine learning algorithms analyzing the video clip are demonstrated. The two upper panels represent the output of Emotrics, an automated facial landmark tracking algorithm. Oral commissure excursion was tracked and normalized on the y axis for both the right and left sides of the face. The threshold of 0.8 is shown by the horizontal black line. Blue shaded areas represent oral commissure excursion above threshold; each of these frames is automatically time-stamped for further analysis. The lower panel represents the probability of joy expression for each frame of the video clip using the Affdex Emotion artificial intelligence



Fig. 1. (Left) Preoperative image of a flaccid right-sided facial palsy patient attempting to smile voluntarily. (Center) Postoperative image demonstrating the effect of facial reanimation with gracilis free functional muscle transfer innervated by the nerve to masseter and nasolabial fold suspension with fascia lata. The patient is smiling voluntarily. (Right) Frame of a video clip of a facial palsy patient filmed while smiling spontaneously in response to humorous content. Full video available to view in Video 2. (Left and center reproduced with permission from Dusseldorp JR, Guarin DL, van Veen MM, Jowett N, Hadlock T. In the eye of the beholder: Changes in perceived emotion expression after smile reanimation. *Plast Reconstr Surg.* 2019;144:457–471.)



Fig. 2. (Left) Preoperative image of a flaccid right-sided facial palsy patient attempting to smile voluntarily. (Center) Postoperative image demonstrating the effect of facial reanimation with gracilis free functional muscle transfer innervated by the nerve to masseter and nasolabial fold suspension with fascia lata. The patient is smiling voluntarily. (Right) Frame depicting a spontaneous smile in response to humorous content. Full video available to view in Video 2. (Left and center reproduced with permission from Dusseldorp JR, Guarin DL, van Veen MM, Jowett N, Hadlock T. In the eye of the beholder: Changes in perceived emotion expression after smile reanimation. *Plast Reconstr Surg.* 2019;144:457–471.)

algorithm (only every third frame of the video clip was sampled by this algorithm, thus the differing frame reference numbers). The *large gray box* depicts the playhead of the filmstrip indicating which frame is currently being analyzed from the video clip. See **Figure, Supplemental Digital Content 2**, which shows outputs of two machine learning algorithms analyzing the video clip in **Fig. 1** and Video 2. (Above and center) Output of Emotrics, an automated facial landmark tracking algorithm. Oral commissure excursion was tracked and normalized on the y axis for both the right and left sides of the face. The threshold of 0.8 is shown by the *horizontal black line*. *Blue shaded areas* represent oral commissure excursion above threshold for the left contralateral healthy side of the face; each of these frames is automatically time-stamped for further analysis. (Below) Probability of joy expression for each frame of the video clip using the Affdex Emotion artificial intelligence algorithm (only every third frame of the video clip was sampled by this algorithm, thus the differing frame reference numbers). The *large gray box* depicts the playhead of the filmstrip indicating which frame is currently being analyzed from the video clip, <http://links.lww.com/PRS/F66>. See **Figure, Supplemental Digital Content 3**, which

shows outputs of two machine learning algorithms analyzing the spontaneous smile video clip in **Fig. 2** and Video 2. (Above and center) Output of Emotrics, an automated facial landmark tracking algorithm. Oral commissure excursion was tracked and normalized on the y axis for both the right and left sides of the face. The threshold of 0.8 is shown by the *horizontal black line*. *Blue shaded areas* represent oral commissure excursion above threshold for the left contralateral healthy side of the face; each of these frames is automatically time-stamped for further analysis. (Below) Probability of joy expression for each frame of the video clip using the Affdex Emotion artificial intelligence algorithm (only every third frame of the video clip was sampled by this algorithm, thus the differing frame reference numbers). The *large gray box* depicts the playhead of the filmstrip indicating which frame is currently being analyzed from the video clip, <http://links.lww.com/PRS/F67>. See **Video 2 (online)**, which shows video clips of two flaccid facial palsy patients, who have undergone smile reanimation with gracilis functional free muscle transfer innervated by the masseteric nerve, smiling spontaneously in response to humorous content. Both patients demonstrate excellent voluntary smile ability; however, with

spontaneous smiling, there is limited ability to coordinate oral commissure movement. In patient 1, there was a very low probability of joy expression during spontaneous smile events. Emotionality quotient score for patient 1 was 0/–50.7. Patient 2 was more successful at combining oral commissure movements and registered a moderate probability of joy expression. Emotionality quotient score for patient 2 was +60.5/–3.3. Graphic outputs of two machine learning algorithms analyzing the video clip are demonstrated. The two *upper panels* represent the output of Emotrics, an automated facial landmark tracking algorithm. Oral commissure excursion was tracked and normalized on the y axis for both the right and left sides of the face. The threshold of 0.8 is shown by the *horizontal black line*. *Blue shaded areas* represent oral commissure excursion above threshold; each of these frames is automatically time-stamped for further analysis. The *lower panel* represents the probability of joy expression for each frame of the video clip using the Affdex Emotion artificial intelligence algorithm (only every third frame of the video clip was sampled by this algorithm, thus the differing frame reference numbers). The *large gray box* depicts the playhead of the filmstrip indicating which frame is currently being analyzed from the video clip. **See Video 3 (online)**, which shows video clips of two facial palsy patients, who have undergone smile reanimation with gracilis free functional muscle transfer innervated by a cross-facial nerve graft, smiling spontaneously in response to humorous content. Patient 1 has left facial palsy with a mixed picture of upper eyelid synkinesis and midface weakness treated with gracilis free functional muscle transfer. She exhibits joy during the early part of the video clip while smiling spontaneously but also exhibits negative emotion in the form of disgust. Her emotionality quotient score was +99.8/–96.9. Patient 2 has right flaccid facial palsy treated with gracilis free functional muscle transfer. Her resting symmetry is good, but with smiling, there is very little movement of her gracilis. The subtle movement that is present on her right side combined with her strong movement on her healthy left side gives her a slight probability of joy and stronger probability of contempt expression with an emotionality quotient score from this video clip of +2.8/–38.8. Graphic outputs of two machine learning algorithms analyzing the video clip are demonstrated. The two *upper panels* represent the output of Emotrics, an automated facial landmark tracking algorithm. Oral commissure excursion was tracked on the y axis for both the

right and left sides of the face. The *lower panel* represents the probability of joy expression for each frame of the video clip using the Affdex Emotion artificial intelligence algorithm (only every third frame of the video clip was sampled by this algorithm, thus the differing frame reference numbers). The *large gray box* depicts the playhead of the filmstrip indicating which frame is currently being analyzed from the video clip. **See Video 4 (online)**, which shows video clips of two flaccid facial palsy patients, who have undergone smile reanimation with gracilis free functional muscle transfer innervated by the masseteric nerve and cross-face nerve graft in dually innervated fashion, smiling spontaneously in response to humorous content. Patient 1 has left-sided flaccid facial palsy treated with gracilis free functional muscle transfer. Only one of four spontaneous smile events on the healthy right side corresponded with a movement of the left oral commissure, and this movement correlated with joy. This patient was assessed by the algorithm as demonstrating a low probability of joy and strong probability of contempt expression, with an emotionality quotient score from this video clip of 0.8/–98.5. Patient 2 has flaccid right facial palsy treated with gracilis free functional muscle transfer dually innervated. His voluntary smile ability was excellent; however, there was no recordable joy when smiling spontaneously. This patient was assessed by the algorithm as having a moderate probability of contempt expression with spontaneous smiling and an emotionality quotient score from this video clip of 0/–41.4. In this case, it could be suggested that the contribution of the cross-facial nerve graft has been minimal, and thus objective assessment of his spontaneous smile ability demonstrates almost no spontaneous movement. Graphic outputs of two machine learning algorithms analyzing the video clip are demonstrated. The two *upper panels* represent the output of Emotrics, an automated facial landmark tracking algorithm. Oral commissure excursion was tracked for both the right and left sides of the face. The *lower panel* represents the probability of joy expression for each frame of the video clip using the Affdex Emotion artificial intelligence algorithm (only every third frame of the video clip was sampled by this algorithm, thus the differing frame reference numbers). The *large gray box* depicts the playhead of the filmstrip indicating which frame is currently being analyzed from the video clip.] Bilateral facial palsy patients were excluded. Postoperative patients were analyzed according to their innervation pattern: cross-face nerve graft, masseteric nerve or dually

innervated (using our standard two stage technique).³ Informed consent procedures were used to select the innervation strategy for each patient. In general terms, we are more likely to recommend cross-face nerve graft and dually innervated reanimation in younger patients and the masseteric nerve alone in older patients (Table 1). Duration of follow-up at the time of spontaneity analysis was also recorded.

Statistical Analysis

Descriptive data were presented as numbers and frequencies, medians and interquartile ranges, and means and standard deviations. Nonparametric tests were used for statistical comparison of perceived emotion between preoperative and postoperative data (Wilcoxon signed ranks test) and gracilis innervation patterns (Kruskal-Wallis test and Mann Whitney *U* tests) because of nonnormality of the data. All statistical analyses were performed using IBM SPSS Version 25.0 (IBM Corp., Armonk, N.Y.), with a significance level of 0.05.

RESULTS

A consecutive sample of 39 unilateral facial palsy patients were analyzed who had undergone cross-face nerve graft–driven ($n = 17$), masseteric nerve–driven ($n = 9$), or dually innervated ($n = 13$) gracilis free functional muscle transfer. Mean follow-up was 1.0 year (interquartile range, 0.7 to 1.4 years). Twenty-four patients (61.5 percent) were

female, and the mean $\pm$ SD age was 36.9 ± 18.1 years. Vestibular schwannoma was the most common cause of facial palsy (Table 1). Nine healthy subjects (staff from our institution) were included as controls: three men and six women, with a median age of 25 years (interquartile range, 23 to 30 years). At least one clinically meaningful healthy oral commissure movement was elicited in 97 percent of cases.

Range of Emotional Quotient Scores in the Normal and Flaccid Preoperative States

In spontaneous smiles of normal subjects, median 100 percent joy was perceived and almost no negative expression, as anticipated. The median emotional quotient score was 100/0 (interquartile range, 2/5). In spontaneous smiles of preoperative flaccid facial palsy patients, median 0 percent joy was perceived, and there was a median 94 percent probability of negative emotion perception. Median emotional quotient score in flaccid facial palsy was 0/–94 (interquartile range, 0/9) (Table 2). (See Table, Supplemental Digital Content 4, which shows computer vision perceived emotionality scores for normal controls and for each patient according to innervation method. The case examples demonstrated in each video are annotated, <http://links.lww.com/PRS/F68>.)

Emotional Quotient Sensitivity to Change in Spontaneous Smiling

In postoperative spontaneous smiles of unilateral facial palsy patients ($n = 39$), median

Table 1. Demographic Data, Cause, and Innervation Pattern of 39 Facial Palsy Patients Who Were Analyzed Postoperatively Having Undergone Gracilis Free Functional Muscle Transfer Using Various Innervation Sources

	All Patients (%)	Gracilis by CFNG (%)	Dual Innervation (%)	Gracilis by Masseteric Nerve (%)	<i>p</i>
No.	39	17	13	9	
Mean age $\pm$ SD, yr	36.9 ± 18.1	24.2 ± 14.8	40.6 ± 10.0	55.3 ± 14.7	$<0.001^*$
Sex					$<0.001^\dagger$
Male	15 (38.5)	3 (17.6)	6 (46.2)	6 (66.7)	
Female	24 (61.5)	14 (82.4)	7 (53.8)	3 (33.3)	
Follow-up, yr					0.136 [‡]
Median	1.0	1.1	0.8	1.3	
IQR	0.7–1.4	0.9–1.7	0.7–1.1	0.7–2.0	
Cause					$<0.001^\dagger$
Vestibular schwannoma	8 (20.5)	—	6 (46.2)	2 (22.2)	
Trauma	6 (15.4)	5 (29.4)	1 (7.7)	—	
CNS lesion	5 (12.8)	4 (23.5)	1 (7.7)	—	
Congenital	4 (10.3)	4 (23.5)	—	—	
Bell palsy	3 (7.7)	—	3 (23.1)	—	
Benign tumor	3 (7.7)	1 (5.9)	—	2 (22.2)	
Iatrogenic	3 (7.7)	1 (5.9)	1 (7.7)	1 (11.1)	
Head and neck cancer	2 (5.1)	—	1 (7.7)	1 (11.1)	
Other	5 (12.8)	2 (11.8)	—	3 (33.3)	

CFNG, cross-facial nerve graft; IQR, interquartile range; CNS, central nervous system.

*Result of one-way analysis of variance test.

†Result of χ^2 test.

‡Result of Kruskal-Wallis test.



Table 2. Perceived Probabilities of Joy and Negative Emotion Expression during Spontaneous Smiling in Normal Subjects and Flaccid Facial Palsy Patients Preoperatively and after Smile Reanimation with Gracilis Functional Free Muscle Transfer Innervated by Either Cross-Face Nerve Graft, Masseteric Nerve, or Dually Innervated*

	% Joy	% Negative Emotion
Normal	100 (2)	0 (5)
Cross-face innervated gracilis	82 (77)	0 (28)
Dually innervated gracilis	10 (98)	24 (56)
Masseteric innervated gracilis	0 (3)	48 (72)
Preoperatively (flaccid)	0 (0)	94 (9)

IQR, interquartile range.

*Data are presented as median (IQR). Raw data of emotionality quotient outcomes are available in Table, **Supplemental Digital Content 4**, <http://links.lww.com/PRS/F68>.

emotional quotient score was +36/−3 (interquartile range, 90/41), compared to median preoperative emotional quotient scores of 0/−94 (interquartile range, 0/9). After smile reanimation surgery, perceived probability of joy expression during spontaneous smiling increased by a mean of 43 percent and negative emotion perception decreased by a mean of 66 percent.

Emotional Quotient Scores by Gracilis Innervation Pattern

Analyzing postoperative patients by innervation pattern yielded median emotional quotient scores for cross-face nerve graft of +82/0 (interquartile range, +77/+28), dual innervation of +10/−24 (interquartile range, +98/+56), and masseteric innervation of 0/−48 (interquartile range, +3/+72) respectively (Table 2) (raw data are available in Table, **Supplemental Digital Content 4**, <http://links.lww.com/PRS/F68>). There were statistically significant differences between these groups [joy, $p = 0.006$ (Kruskal-Wallis test, chi-square value, 10.3); negative emotion, $p = 0.034$ (Kruskal-Wallis test, chi-square value, 6.8)]. In a post hoc analysis, the main difference was found to be between the cross-face nerve graft- and masseteric nerve-driven gracilis free functional muscle transfer [joy, $p = 0.001$ (Mann-Whitney U test, 17.0); negative emotion, $p = 0.008$ (Mann-Whitney U test, 29.0)]. There was a trend toward increased joy expression and decreased negative emotion expression in dually innervated spontaneous smiles compared with those innervated by the masseteric nerve; however, this was not statistically significant [joy, $p = 0.117$ (Mann-Whitney U test, 36.0); negative emotion, $p = 0.262$ (Mann-Whitney U test, 41.0)]. Median follow-up durations in

cross-face nerve graft, dual, and nerve-to-masseter were 1.1 years (interquartile range, 0.9 to 1.7 years), 0.8 year (interquartile range, 0.7 to 1.1 years), and 1.3 years (interquartile range, 0.7 to 2.0 years), respectively.

Manual versus Automatic Time-Stamping

During spontaneous smile assessments, frequency and intensity of spontaneous smile events was highly variable in both normal controls and facial palsy patients. The mean number of spontaneous smile events defined by the automated software in normal and facial palsy patients was five (interquartile range, two/seven), compared with manual time-stamping, which had a mean of three (interquartile range, two/five). At least one large healthy side spontaneous smile event was time-stamped using both the automated and manual techniques in 100 percent of cases, and algorithm-derived time-stamps were found to be more accurate when analyzing discordance between automated and manual time-stamps (data not presented).

DISCUSSION

The goal of this study was to develop and test the feasibility of an objective system to quantify spontaneous smiling. To achieve this, we analyzed spontaneous smiles of normal subjects and unilateral facial palsy patients, before and after reanimation. While watching recordings of spontaneous smiles, we considered which moments we needed to analyze, and how to not measure other facial movements (such as when patients talk, purse or bite their lips, or have cheek twitches). The key to identifying only spontaneous smile events was to automatically identify only those frames with large (>80 percent of maximum) and sustained (>1 second) healthy side oral commissure movements.³

We found that spontaneous smile video clips could be automatically time-stamped and objectively quantified based on the success of the smile to express joy using two freely available machine learning approaches: automatic facial landmarking, known as Emotrics⁴; and emotion detection computer vision software, known as Affdex.⁵ Oral commissure movement away from the midline on the healthy side of the face was tracked throughout the video clip. The largest spontaneous displacement within the video clip was normalized to 100 percent, and we only reviewed movements that were 80 percent or higher when compared to this benchmark. We found this was

essential, as the variability between individuals was so great. Frames of interest were automatically time-stamped, and emotion detection software was used to score the ability of each spontaneous smile to express emotion.

Emotion detection software perceived 100 percent joy in spontaneous smiles of normal controls and 0 percent joy in spontaneous smiles of preoperative flaccid unilateral facial palsy patients. The probability of perceiving joy during a spontaneous smile event was increased by, on average, 43 percent after smile reanimation, whereas the probability of perceiving a negative emotion was reduced by 73 percent, on average (Table 2). These findings were statistically significant ($p < 0.001$). To test the utility of the machine learning approach to quantify differences in spontaneous smiling, we performed an analysis of free functional muscle transfer outcomes based on innervation technique using a consecutive series of free functional muscle transfer results with video clips available in our database. It is important to note that this study was not adequately powered, or designed with sufficient long-term follow-up, to definitively determine the optimal innervation technique in free functional muscle transfer facial reanimation. These findings are preliminary and require further testing. Nevertheless, our findings were consistent with those of a recent meta-analysis of spontaneity outcomes, demonstrating that cross-facial nerve grafting achieved the greatest degree of improvement in emotionality during spontaneous smiling, whereas non-facial nerve-innervated reanimation achieved significantly less (Table 2). In adults, single-stage masseteric-innervated facial reanimation has been consistently demonstrated to produce a good voluntary smile,^{5,7} and in the literature, rates of approximately 30 percent success have been quoted for spontaneity outcomes following masseter-driven reanimation. Long-term pediatric studies quantifying spontaneous smiling outcomes after masseteric nerve-driven reanimation are required to determine likelihood of spontaneity in this patient group.

Long-term follow-up can be very difficult to obtain in facial reanimation, particularly when patients are satisfied with their early results. However, we need these data to ensure we give future patients the most accurate expectations. It is this exact problem that led us to develop this automated machine learning tool, and we now aim to convert this machine learning algorithm into a freely available smartphone application. Having the ability to prospectively assess long-term spontaneous smile outcomes from multiple

centers will lead to an understanding of the optimal techniques to restore spontaneous smiling. In the era of telemedicine and value-based health care, it will become increasingly important to objectively quantify the effectiveness of our interventions according to a specific set of patient-centered outcomes.

A precise means to quantify smile spontaneity has heretofore remained elusive. To quantify spontaneous smiling, clinicians have typically relied on their own clinical observations during the course of a consultation, in conversation, while telling a joke, or using tests such as “the tickle test.”⁸ The accuracy of the clinician’s ability to simultaneously elicit and observe spontaneous smiling is questionable, and no correlation has been found between clinician assessments of spontaneity and objectively measured synchronicity of oral commissure movement in smile reanimation patients.^{3,9} Patient-reported spontaneity may also be unreliable, as both patient and clinician are motivated to report a positive spontaneity outcome given the amount of time and effort vested in the process.¹⁰ Moreover, very high rates of spontaneity are described in the literature from dually innervated (approximately 83 percent) and masseteric techniques (approximately 30 percent), yet these high rates have not been replicated when looking objectively at our experience using similar techniques.^{3,10}

We have previously used synchronous oral commissure movement as a proxy for spontaneous smiling.³ However, this definition lacks the resolution to stratify imperfect results and fails to grade symmetry of oral commissure movements. From a surgical perspective, quantifying degrees of static and dynamic asymmetry in excursion distance and vector, and phase asymmetry can be useful.^{3,4} However, facial symmetry is not, in itself, a patient-centered goal of reconstruction.⁵ The overriding goal from the patient’s perspective is to appear normal in the eyes of a casual observer.^{5,11,12} At present, the best tool to measure this outcome is layperson assessment.^{11–14} However, layperson assessments are cumbersome and limited by patient privacy concerns.^{5,11,15} We have previously demonstrated that layperson perceptions of facial palsy correlate with scores generated by emotion detection software.⁵ For this reason, we have used emotionality to objectively determine the ability of the spontaneous smile to convey joy. We believe that an automated spontaneity assessment tool could be a useful adjunct to existing outcome measures in facial reanimation, as it tells us the functional ability of an emotional smile to express

joy, and how closely the reanimated smile relates to the normal population.

Limitations

This study has several important limitations. We used a validated series of video clips (known as the Spontaneous Smile Assay⁶) to elicit spontaneous smiles. However, once patients viewed the Spontaneous Smile Assay video preoperatively, it became familiar and, in some cases, smiling was less spontaneous postoperatively. Despite this, very few subjects failed to generate at least one clinically relevant spontaneous smile event, lending support to the use of this technique to elicit spontaneity. In the future, it would be preferable to have access to a range of validated video clips to avoid familiarity with the stimulus material.

Another limitation of this study was the assumption that a large, sustained, healthy side oral commissure movement while watching a humorous video clip equated to a spontaneous smile event. This assumption was made after qualitative assessment of multiple normal subjects and facial palsy patients. We observed that the most common pattern of smile elicitation as a result of the Spontaneous Smile Assay was for healthy side oral commissure excursion to slowly increase in amplitude until a full spontaneous smile occurred, coinciding with audible laughter, often followed by talking, lip biting or puckering, and cheek twitches, which were common throughout the video recordings. In these scenarios, manual and automatic time-stamping were similar, and both were able to readily discern large and sustained healthy side spontaneous oral commissure movements. However, if there was only one very large smile and a series of smaller ones, the algorithm time-stamped only the large event and did not time-stamp smaller smiles, and if there was no large smile event, the algorithm would time-stamp several smaller smile events. In these scenarios, there was some discordance between the number of smile events identified using manual and automated time-stamping. However, when these discordant cases were reviewed, the automated approach was found to be more accurate at determining frames that coincided with clinically relevant healthy side oral commissure movement and often audible laughter.

There are several concerns with using an artificially intelligent emotion detection algorithm as an outcome measure in facial palsy.⁵ One criticism is that the algorithm analyzes the whole face and thus may be influenced by activity on the healthy side, in particular, the periocular region.¹⁶

A counterargument is that laypersons also analyze the whole face, and that the “healthy” side is also frequently affected in unilateral facial palsy.¹⁷ What is important is not only the degree of facial symmetry but rather whether or not patients’ facial expressions of the whole face fit into the realm of normal, in the eyes of a casual observer.¹³ The machine learning algorithm used to assess emotion in this study was trained using millions of normal facial expressions gathered from over 75 countries, representing a comprehensive normative data set that is ideally suited for comparison with our smile reanimation outcomes.

CONCLUSIONS

It is critical that patients understand the true likelihood of developing a spontaneous smile after various types of smile reanimation. Previously, comparison of smile spontaneity rates achieved through various surgical approaches has been largely subjective. By leveraging the power of machine learning, we have created an objective tool for automatic quantification of smile spontaneity. This tool may enable long-term, multicenter studies to characterize optimal techniques to restore spontaneous, emotional smiling.

Joseph R. Dusseldorp, F.R.A.C.S.
Massachusetts Eye and Ear Infirmary
243 Charles Street
Boston, Mass. 02114
drjoedusseldorp@icloud.com
Facebook: [drjoedusseldorp](#)
Instagram: [@drjoedusseldorp](#)

PATIENT CONSENT

Patients provide written informed consent for the use of their images.

REFERENCES

1. Coulson SE, O'dwyer NJ, Adams RD, Croxson GR. Expression of emotion and quality of life after facial nerve paralysis. *Otol Neurotol*. 2004;25:1014–1019.
2. Ho AL, Scott AM, Klassen AF, Cano SJ, Pusic AL, Van Laeken N. Measuring quality of life and patient satisfaction in facial paralysis patients: A systematic review of patient-reported outcome measures. *Plast Reconstr Surg*. 2012;130:91–99.
3. Dusseldorp JR, van Veen MM, Guarin DL, Quatela O, Jowett N, Hadlock TA. Spontaneity assessment in dually innervated gracilis smile reanimation surgery. *JAMA Facial Plast Surg*. 2019;21:551–557.
4. Guarin DL, Dusseldorp J, Hadlock TA, Jowett N. A machine learning approach for automated facial measurements in facial palsy. *JAMA Facial Plast Surg*. 2018;20:335–337.
5. Dusseldorp JR, Guarin DL, van Veen MM, Jowett N, Hadlock TA. In the eye of the beholder: Changes in perceived

- emotion expression after smile reanimation. *Plast Reconstr Surg*. 2019;144:457–471.
6. Iacolucci CM, Banks C, Jowett N, et al. Development and validation of a spontaneous smile assay. *JAMA Facial Plast Surg*. 2015;17:191–196.
 7. Snyder-Warwick AK, Fattah AY, Zive L, Halliday W, Borschel GH, Zuker RM. The degree of facial movement following microvascular muscle transfer in pediatric facial reanimation depends on donor motor nerve axonal density. *Plast Reconstr Surg*. 2015;135:370e–381e.
 8. Chuang DCC, Lu JCY, Anesti K. One-stage procedure using spinal accessory nerve (XI)-innervated free muscle for facial paralysis reconstruction. *Plast Reconstr Surg*. 2013;132:117e–129e.
 9. Dusseldorp JR, van Veen MM, Mohan S, Hadlock TA. Outcome tracking in facial palsy. *Otolaryngol Clin North Am*. 2018;51:1033–1050.
 10. van Veen MM, Dusseldorp JR, Quatela O, et al. Patient experience in nerve-to-masseter-driven smile reanimation. *J Plast Reconstr Aesthet Surg*. 2019;72:1265–1271.
 11. Lyford-Pike S, Helwig NE, Sohre NE, Guy SJ, Hadlock TA. Predicting perceived disfigurement from facial function in patients with unilateral paralysis. *Plast Reconstr Surg*. 2018;142:722e–728e.
 12. Ishii L, Godoy A, Encarnacion CO, Byrne PJ, Boahene KD, Ishii M. Not just another face in the crowd: Society's perceptions of facial paralysis. *Laryngoscope* 2012;122:533–538.
 13. Dey JK, Ishii M, Boahene KDO, Byrne PJ, Ishii LE. Changing perception: Facial reanimation surgery improves attractiveness and decreases negative facial perception. *Laryngoscope* 2014;124:84–90.
 14. Ishii L, Dey J, Boahene KD, Byrne PJ, Ishii M. The social distraction of facial paralysis: Objective measurement of social attention using eye-tracking. *Laryngoscope* 2016;126:334–339.
 15. Chuang DC. Discussion: In the eye of the beholder: Changes in perceived emotion expression after smile reanimation. *Plast Reconstr Surg*. 2019;144:472–473.
 16. Gulbitti HA, van der Lei B. In the eye of the beholder: Changes in perceived emotion expression after smile reanimation. (Letter) *Plast Reconstr Surg*. 2020;145:1109e–1110e.
 17. Dusseldorp JR, Hadlock TA. Reply: In the eye of the beholder: Changes in perceived emotion expression after smile reanimation. *Plast Reconstr Surg*. 2020;145:1110e–1112e.

Appendix 7

**In the Eye of the Beholder: Changes in Perceived
Emotion Expression after Smile Reanimation**

And Invited Discussion and Letters to the Editor

Authorship Statement

I, Joseph Dusseldorp, confirm for the ensuing journal article: **‘In the Eye of the Beholder: Changes in Perceived Emotion Expression after Smile Reanimation’** and **‘Reply to In the Eye of the Beholder: Changes in Perceived Emotion Expression after Smile Reanimationx’** that I have made the following contributions to its authorship:

- Conception and design of the research
- Collection and extraction of data
- Analysis and interpretation of the findings
- Drafting and revising of the manuscript and critical appraisal of content
- Responding to multiple stages of peer review



Joseph R. Dusseldorp

Date 14.09. 2023

As Auxiliary Supervisor for the candidature upon which this thesis is based, I can confirm that the above authorship attribution is correct.



Professor Tessa Hadlock, MD

Date 27.07. 2023

Department of Otolaryngology
Massachusetts Eye and Ear Infirmary
Harvard University, Boston, MA, USA

As Primary Supervisor of the candidate, I concur with the above authorship attestation:



Professor Peter Maitz

Date 14.09.2023

Department of Burns Surgery
Concord Repatriation General Hospital
Sydney University, Sydney, NSW, Australia

In the Eye of the Beholder: Changes in Perceived Emotion Expression after Smile Reanimation

Joseph R. Dusseldorp,
F.R.A.C.S.

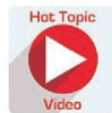
Diego L. Guarin, Ph.D.

Martinus M. van Veen, M.D.

Nate Jowett, M.D.

Tessa A. Hadlock, M.D.

Boston, Mass.; Sydney, New South
Wales, Australia; and Groningen,
The Netherlands



Background: Tools to quantify layperson assessments of facial palsy are lacking. In this study, artificial intelligence was applied to develop a proxy for layperson assessments, and compare sensitivity to existing outcome measures.

Methods: Artificially intelligent emotion detection software was used to develop the emotionality quotient. The emotionality quotient was defined as the percentage probability of perceived joy over the percentage probability of perceived negative emotions during smiling, as predicted by the software. The emotionality quotient was used to analyze the emotionality of voluntary smiles of normal subjects and unilateral facial palsy patients before and after smile reanimation. The emotionality quotient was compared to oral commissure excursion and layperson assessments of facial palsy patients.

Results: In voluntary smiles of 10 normal subjects, 100 percent joy and no negative emotion was detected (interquartile ranges, 0/1). Median preoperative emotionality quotient of 30 facial palsy patients was 15/–60 (interquartile range, 73/62). Postoperatively, median emotionality quotient was 84/0 (interquartile range, 28/5). In 134 smile reanimation patients, no correlation was found between postoperative oral commissure excursion and emotionality quotient score. However, in 61 preoperative patients, a moderate correlation was found between layperson-assessed disfigurement and negative emotion perception (correlation coefficient, 0.516; $p < 0.001$).

Conclusions: Computer vision artificial intelligence software detected less joy and more negative emotion in smiles of facial palsy patients compared with normal subjects. Following smile reanimation, significantly more joy and less negative emotion were detected. The emotionality quotient was correlated with layperson assessments. The simplicity, sensitivity, and objectivity of the emotionality quotient render it an attractive tool to serve as a potential proxy for layperson assessment, an ideal outcome measure in facial palsy. (*Plast. Reconstr. Surg.* 144: 457, 2019.)

CLINICAL QUESTION/LEVEL OF EVIDENCE: Diagnostic, IV.

Smiling is a highly effective means of nonverbal communication,^{1,2} and inability to smile has a profoundly negative impact on quality of life, being the dominant problem reported by facial palsy patients and their families.^{3,4} For

Disclosure: None of the authors has a financial interest in any of the products or devices mentioned in this manuscript. Dr. Guarin received funding from “Fonds Quebecoise de la recherche sur la nature et technologies” for his postdoctoral work.

From the Department of Otolaryngology/Head and Neck Surgery, Massachusetts Eye and Ear Infirmary and Harvard Medical School; the Department of Plastic and Reconstructive Surgery, Royal Australasian College of Surgeons and University of Sydney; and the Department of Plastic Surgery, University Medical Center Groningen and University of Groningen.

Received for publication April 2, 2018; accepted January 7, 2019.

Presented at the 29th Annual Meeting of the European Association of Plastic Surgeons, in Madrid, Spain, May 17 through 19, 2018.

Copyright © 2019 by the American Society of Plastic Surgeons

DOI: 10.1097/PRS.0000000000005865

Supplemental digital content is available for this article. Direct URL citations appear in the text; simply type the URL address into any Web browser to access this content. Clickable links to the material are provided in the HTML text of this article on the *Journal's* website (www.PRSJournal.com).

A “Hot Topic Video” by Editor-in-Chief Rod J. Rohrich, M.D., accompanies this article. Go to PRSJournal.com and click on “Plastic Surgery Hot Topics” in the “Digital Media” tab to watch.

facial reanimation surgeons, the most challenging reconstructive goal is for the patient to appear normal in the eye of the naive observer. Despite the multiple outcome measures available in facial reanimation surgery, none has gained widespread acceptance, and it is impractical to perform layperson assessments routinely in clinical practice. An extremely welcome tool would be an assay that predicts the layperson interpretation of facial expressions in facial palsy.

A method of quantifying normal facial expressions, the Facial Action Coding System, was developed by analyzing the contributions from discrete facial muscle groups (known as action units).^{5,6} Joy is composed of bilateral lip corner pull, cheek raise, and lid tighten; disgust is upper lip raise, wrinkled nose, and cheek raise; and contempt is a variant of disgust, also known as a smirk.^{7,8} [See **Appendix, Supplemental Digital Content 1**, which shows muscle action units (AU) relevant to the facial expressions associated with smiling. *FACS*, Facial Action Coding System. (Excerpt taken from Ekman P, Friesen WV. *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Palo, Alto, Calif: Consulting Psychologists Press; 1978), <http://links.lww.com/PRS/D582>.] Although clearly a useful tool, Facial Action Coding System analysis requires a high degree of training, is time-consuming, and has only moderate interrater reliability.⁹

Recent advances in artificial intelligence have enabled automatic detection of facial movements, and machine learning has trained models to predict perceived expression of standardized emotion categories.¹⁰ Freely available, facial recognition and emotion detection software has been developed from 5 million videos of normal subjects from 75 countries, resulting in over 2 billion facial frames in which Facial Action Coding System action units were expertly recognized and labeled.^{11–15} These data have enabled automatic analysis of the phenomenon known as emotionality, the *observable* behavioral and physiologic component of emotion.¹⁶ One must recognize that facial movement analysis alone, without contextual information, is unable to determine the true or *felt* emotional state of an individual.¹⁷ Therefore, computer vision software may only predict *perceived* emotion or, in the case of voluntary facial movements, emotion that is being *communicated*.

The aim of this study was to leverage recent advances in artificial intelligence software to develop the emotionality quotient, a novel, objective, automated emotion-detection tool in facial palsy. The emotionality quotient was used to characterize the *perceived* emotionality of smiling in

normal subjects and in facial palsy patients before and after smile reanimation with gracilis functional free muscle transfer. Only voluntary smile frames were analyzed where, by definition, the subject was trying to communicate joy. Spontaneous smiling was not assessed. A further aim was to validate the emotionality quotient by comparison with oral commissure excursion and layperson assessments of facial disfigurement.

PATIENTS AND METHODS

This study was approved by the Massachusetts Eye and Ear Infirmary Institutional review board. Written consent was obtained from all patients, before their inclusion in this study.

Freely available, proprietary software (Affectiva; Affectiva, Boston, Mass.) was used to analyze photographs and videos of smiles, generating percentage probabilities of perceived joy, contempt, and disgust (Fig. 1). Postprocessing of data was performed with MATLAB (MathWorks, Natick, Mass.). A descriptive means of reporting relative probabilities, known as the emotionality quotient, was defined as the percentage probability of perceived joy (range, 0 to 100 percent) over the percentage probability of perceived negative emotions (contempt or disgust; range, –100 to 0 percent) during smiling, as predicted by computer vision software. The higher probability between contempt and disgust was selected, as similar action units evoke these expressions. The change in emotionality quotient after an intervention, known as emotionality quotient delta, was defined as the sum of the percentage change in both joy and negative emotion perception, comparing postoperative to preoperative scores (Fig. 2). The maximum possible improvement in emotionality quotient delta is 200 percent.

Range of Emotionality Quotient Scores in Normal, Flaccid Facial Palsy, and Nonflaccid Facial Palsy Patients

To quantify the range of perceived emotion in unilateral facial palsy patients, we analyzed maximal voluntary smile efforts in flaccid and nonflaccid facial palsy patients, and used normal subjects as controls (Fig. 1). Bilateral facial palsy patients were excluded. Flaccid facial palsy was defined by a complete absence of facial movement on the affected side of the face for at least 18 months. Nonflaccid facial palsy was defined as all other cases of facial palsy where there had been some improvement in facial animation on the affected side, with or without synkinesis.

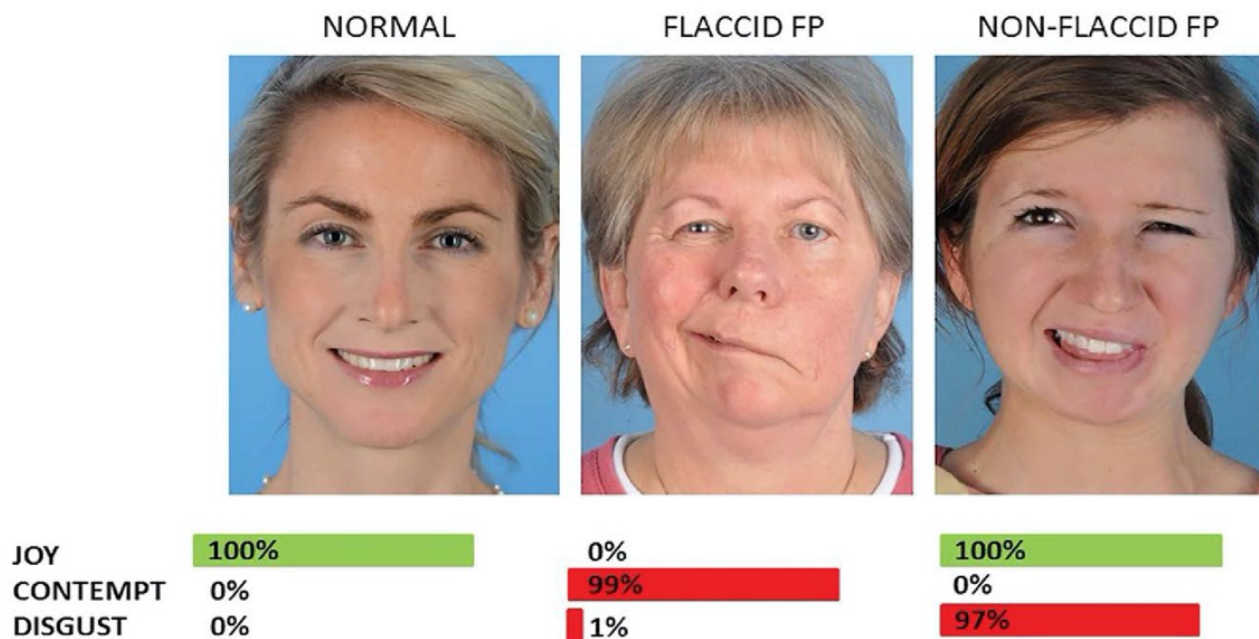


Fig. 1. Maximal smile efforts in (left) a normal subject (high levels of joy and no negative emotion are perceived), (center) a patient with flaccid facial palsy (no joy and high levels of contempt are perceived), and (right) a patient with nonflaccid facial palsy (high levels of joy and disgust are perceived). Both flaccid and nonflaccid patients are affected on their left sides. FP, facial palsy.

Emotionality Quotient Sensitivity to Change and Reliability

Paired preoperative and postoperative smiles of randomly selected gracilis functional free muscle transfer smile reanimation cases were analyzed to determine the reliability and sensitivity to change of emotionality quotient (Table 1). Selected cases were screened to ensure they included equal numbers of flaccid and nonflaccid facial palsy, and cases driven by cross-face nerve graft, nerve-to-masseter, and combined innervation techniques. Reliability was assessed by analyzing the same photographs on multiple occasions.

Comparison of Oral Commissure Excursion to the Emotionality Quotient

To determine the relationship between the emotionality quotient and oral commissure excursion, we reviewed all gracilis functional free muscle transfer smile reanimation cases performed at our institution between 2006 and 2016. Cases were excluded from analysis if preoperative or postoperative images were unavailable or if the algorithm failed to detect the face (e.g., because of a cropped border of the face). Cause, preoperative facial nerve status (i.e., flaccid or nonflaccid), and innervation pattern of the gracilis were recorded (Table 2). Commissure excursion was measured from postoperative voluntary smile and rest photographs using an automated photographic

measurement tool, known as Emotrics.^{18,19} The Emotrics application automatically calculated pixel distance from the point where the facial midline crossed the lower lip to the oral commissure on the affected side. This value was scaled to iris diameter to give an output in millimeters.^{18,19} Negative values were recorded if the affected side oral commissure moved toward the facial midline during smiling.

Comparison of Layperson-Assessed Disfigurement to the Emotionality Quotient

To establish the relationship between the emotionality quotient and layperson assessment of facial palsy, we compared emotionality quotient scores to a validated method of quantifying layperson assessment of disfigurement in unilateral facial palsy. Thanks to generous collaborators, we were able to analyze exactly the same voluntary smile videos that had been shown to 593 laypersons at the 2016 Minnesota State Fair.²⁰ Sixty-one videos of preoperative, unilateral facial palsy patients were selected from a larger population of 500 patient videos provided by the Massachusetts Eye and Ear Infirmary to encompass the complete spectrum of eFACE total scores (range, 40 to 98) and patient ages (range, 18 to 84 years).²⁰ Laypersons ranked disfigurement of facial palsy patients during smiling on a Likert scale from 0 to 10 (with 0 being no disfigurement and 10 being maximal

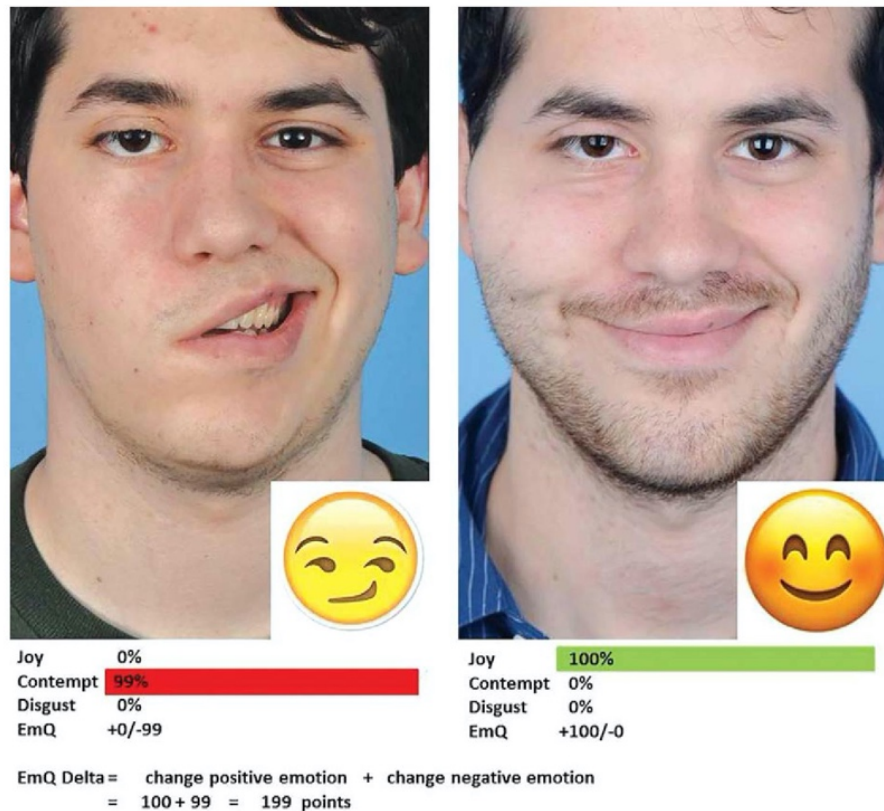


Fig. 2. Maximal smile efforts. (Left) Preoperative flaccid facial palsy and (right) postoperative outcome of smile reanimation with gracilis functional free tissue muscle transfer. Preoperatively, 0 percent perceived joy and 99 percent perceived contempt, and so the emotionality quotient was +0/−99. Postoperatively, 100 percent perceived joy and 0 percent perceived contempt, and so the emotionality quotient was reversed to +100/−0. The emotionality quotient delta is the sum of the improvements in probabilities of both joy and negative emotion perception: emotionality quotient delta = 100-point improvement (joy) + 99-point reduction (negative emotion) = 199 points. This patient had longstanding right unilateral flaccid facial palsy after vestibular schwannoma resection corrected by gracilis functional free muscle transfer dually innervated by the masseteric nerve and a long cross-face nerve graft. Postoperative photograph was taken after 1.6 years. *EmQ*, emotionality quotient; *Preop*, preoperative; *Postop*, postoperative.

disfigurement). The Likert scale was normalized to range from 0 to 100, so that a mean value of layperson assessments could be calculated for each patient. Finally, we compared mean layperson-assessed disfigurement with the probability of perceived joy and perceived negative emotion for each patient.

Statistical Analysis

Descriptive data were presented as numbers, frequencies, medians, interquartile ranges (the difference between the 25th and 75th percentiles), and means and standard deviations. Nonparametric tests were used for statistical comparison of perceived emotion between preoperative and

postoperative data (Wilcoxon signed ranks test) and flaccid and nonflaccid patients (Mann-Whitney *U* test) attributable to nonnormality of the data. Correlations were tested using the Spearman rank correlation. An alpha of 0.05 was selected, and all statistical tests were performed using IBM SPSS Version 25 (IBM Corp., Armonk, N.Y.).

RESULTS

Range of Emotionality Quotient Scores in Normal, Flaccid Facial Palsy, and Nonflaccid Facial Palsy Patients

In maximal effort smiles of normal subjects ($n = 10$), 100 percent joy was perceived and almost

Table 1. Demographic Data, Cause, and Innervation Pattern of Two Representative Groups of Facial Palsy

	FFP	NFFP
No.	15	15
Mean age $\pm$ SD, yr	41.5 $\pm$ 17.4	37.1 $\pm$ 14.7
Sex		
Male	6	3
Female	9	12
Cause		
Vestibular schwannoma	4	4
Temporal bone fracture	—	3
Trauma	—	3
Brainstem hemangioma	2	—
Iatrogenic injury	1	1
Medulloblastoma	1	1
Other	7	3
Innervation source		
Nerve to masseter	5	5
Cross-face nerve graft	5	5
Dually innervated	5	5

FFP, flaccid facial palsy; NFFP, nonflaccid facial palsy.

Table 2. Demographic Data, Cause, and Innervation Pattern of All Patients Analyzed Preoperatively and Postoperatively Who Underwent Gracilis Free Functional Muscle Transfer

Characteristic	Value (%)
No. of gracilis patients analyzed	134
Mean age $\pm$ SD, yr	41.3 $\pm$ 18.1
Sex	
Male	48
Female	86
Cause	
Vestibular schwannoma	25 (18.7)
Central nervous system lesion	25 (18.7)
Benign tumor	15 (11.2)
Congenital	14 (10.4)
Bell palsy	13 (9.7)
Head and neck cancer	13 (9.7)
Trauma	13 (9.7)
Other	16 (11.9)
Innervation source	
Nerve to masseter	69
Cross-face nerve graft	55
Dually innervated	10

no negative expression was detected, as anticipated. Median emotionality quotient was +100/0 (interquartile range, 0/1) (Fig. 1). In maximal effort smiles of flaccid facial palsy patients ($n = 15$), a median 0 percent joy was perceived and there was a median 75 percent probability of negative emotion perception. Median emotionality quotient score in flaccid facial palsy was 0/−75 (interquartile range, 0/47). In maximal effort smiles of nonflaccid facial palsy patients ($n = 15$), there was a median 63 percent probability that joy was perceived and a 16 percent probability of negative emotion perception. Median emotionality quotient score in nonflaccid facial palsy was 63/−16 (interquartile range, 62/57). This result

indicates that, preoperatively, the software was 63 percent more likely to perceive joy in smiles of nonflaccid facial palsy patients than in flaccid facial palsy patients ($p < 0.001$).

Emotionality Quotient Sensitivity to Change and Reliability

In postoperative maximal effort smiles of flaccid facial palsy patients ($n = 15$), the median emotionality quotient score was +84/0 (interquartile range, 26/5), compared to preoperative score of 0/−75 (interquartile range, 0/47). After smile reanimation surgery, the perceived probability of joy expression increased by 84 percent (interquartile range, 92 percent; $p = 0.002$) and negative emotion perception diminished by 73 percent (interquartile range, 51 percent; $p < 0.001$). The median emotionality quotient delta of gracilis functional free muscle transfer smile reanimation in flaccid facial palsy was 144 percent (interquartile range, 122 percent) (Figs. 3 and 4). (See Video, Supplemental Digital Content 2, which shows preoperative and postoperative maximal smile efforts in a flaccid facial palsy patient with graphic representations of automated output of the emotion detection software, <http://links.lww.com/PRS/D583>. See Video, Supplemental Digital Content 3, which shows preoperative and postoperative maximal smile efforts in a flaccid facial palsy patient with graphic representations of automated output of the emotion detection software, <http://links.lww.com/PRS/D584>.) In postoperative maximal effort smiles of nonflaccid facial palsy patients ($n = 15$), the median emotionality quotient score was +85/0 (interquartile range, 33/6), compared with a preoperative score of +63/16 (interquartile range, 62/57). Perceived joy was not significantly increased in the nonflaccid facial palsy group (median, 2 percent; interquartile range, 36 percent; $p = 0.194$). However, a small yet significant reduction in perceived negative emotion was demonstrated (9 percent; interquartile range, 54 percent; $p = 0.004$). The median emotionality quotient delta of gracilis functional free muscle transfer smile reanimation in nonflaccid facial palsy patients was 35 percent (interquartile range, 69 percent). [See Figure, Supplemental Digital Content 4, which shows (above) preoperative and postoperative maximal smile efforts in a nonflaccid facial palsy patient. (Below) Graphic representations of automated output of the emotion-detection software from a video of the following sequence: rest, smile, rest, smile, rest. (Below, left) Preoperatively, no joy was detected and a

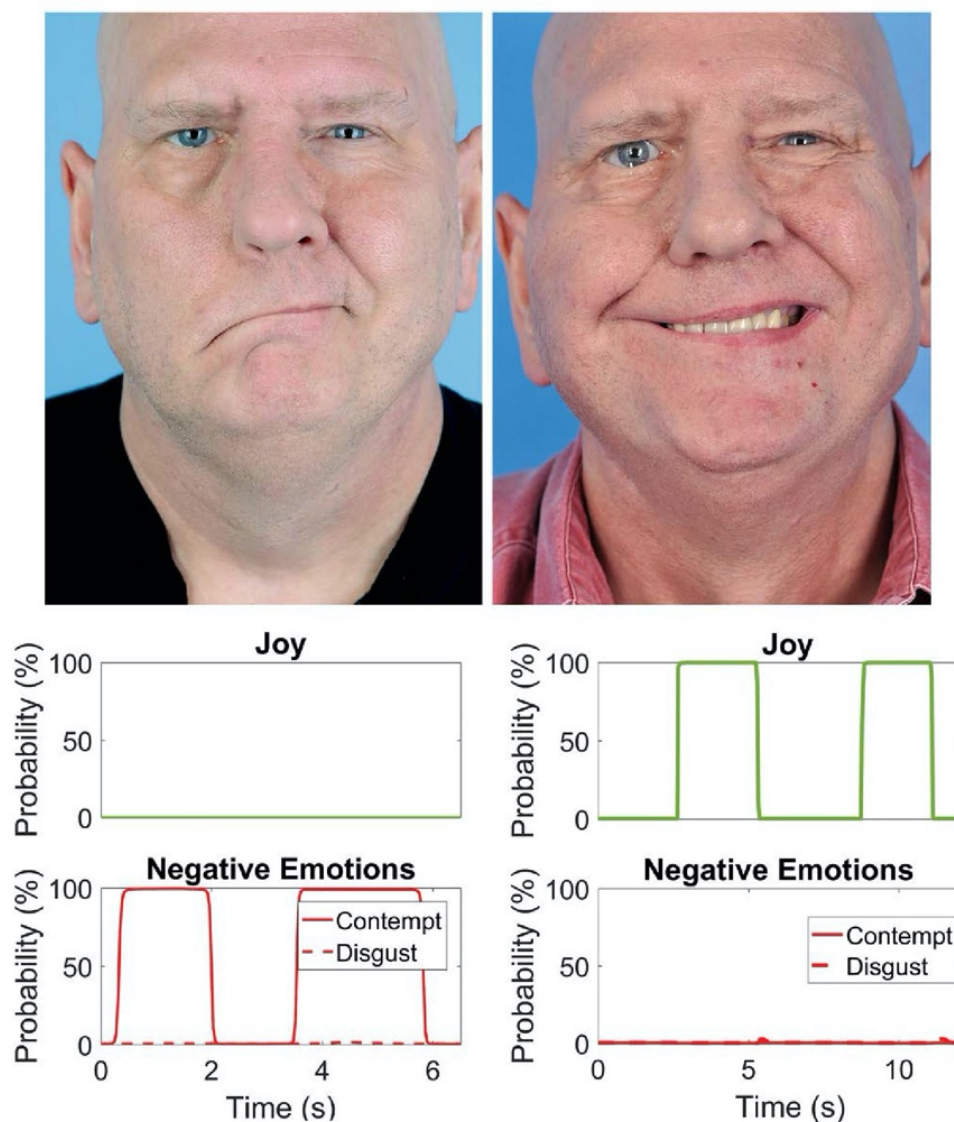


Fig. 3. (Above) Preoperative and postoperative maximal smile efforts in a flaccid facial palsy patient. (Below) Graphic representations of automated output of the emotion-detection software from a video of the following sequence: rest, smile, rest, smile, rest. (Below, left) Preoperatively, no joy was detected and considerable contempt was detected during smile efforts. (Below, right) Postoperatively, this was corrected. This patient had longstanding right unilateral flaccid facial palsy after resection of a cholesteatoma corrected by gracilis functional free muscle transfer innervated by the masseteric nerve. He also underwent static facial suspension with two ribbons of fascia lata. The postoperative photograph was taken after 1.4 years.

mean of 32 percent contempt was detected during smile efforts. (Below, right) Postoperatively, 92 percent joy was detected and no negative emotion was detected. This patient had a temporal bone fracture and developed right facial nerve palsy with subsequent panfacial synkinesis. She underwent two-stage gracilis functional free muscle transfer innervated by a cross-face nerve graft 3 years after her injury. She had no static suspensions. The postoperative photograph was

taken after 0.6 years, <http://links.lww.com/PRS/D587>. See Video, Supplemental Digital Content 5, which shows preoperative and postoperative maximal smile efforts in a nonflaccid facial palsy patient with graphic representations of automated output of the emotion-detection software, <http://links.lww.com/PRS/D589>.] Every time the algorithm was presented with the same photograph, it produced the same numerical output. The emotionality quotient was highly reliable.

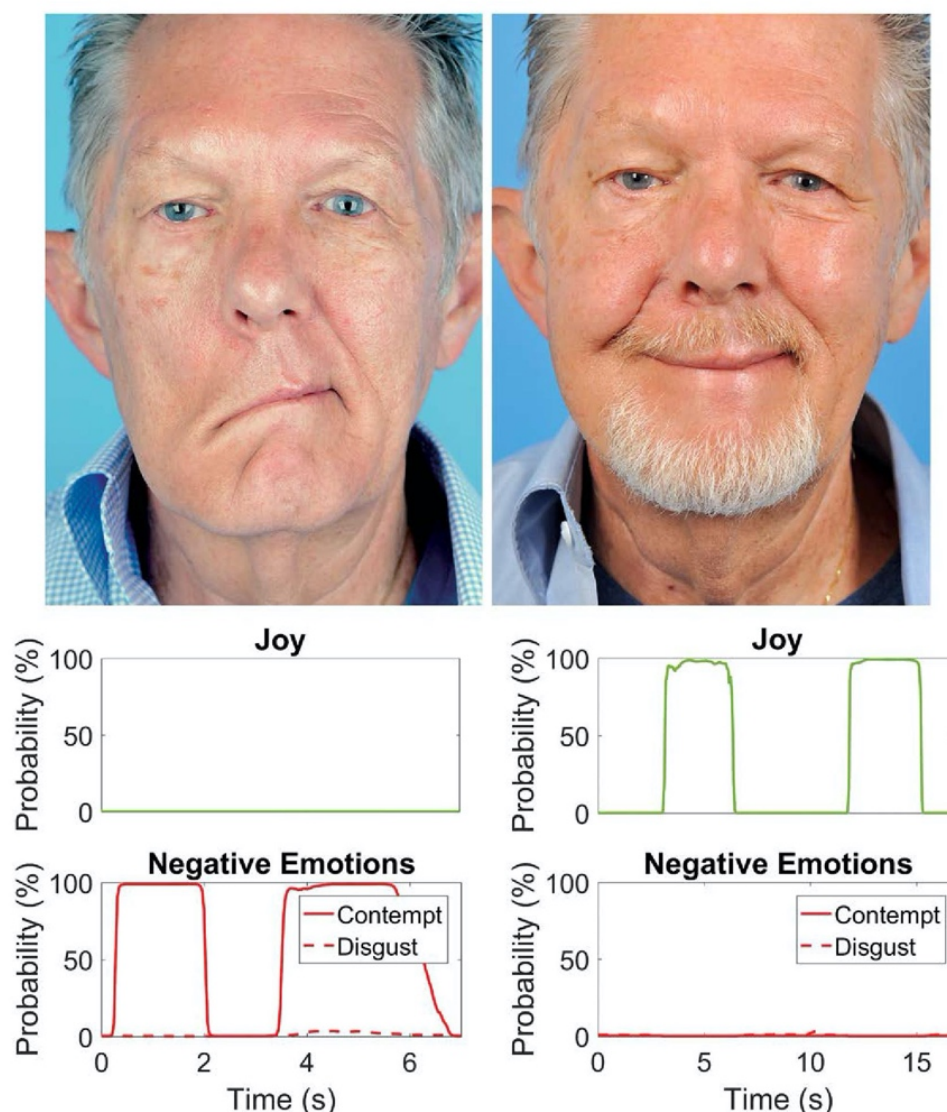
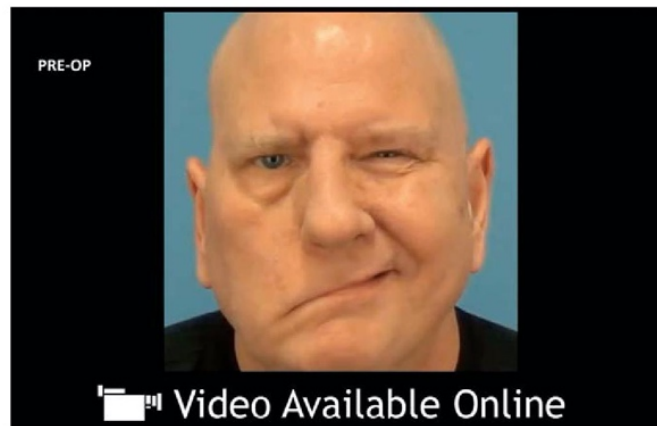


Fig. 4. (Above) Preoperative and postoperative maximal smile efforts in a flaccid facial palsy patient. (Below) Graphic representations of automated output of the emotion-detection software from a video of the following sequence: rest, smile, rest, smile, rest. (Below, left) Preoperatively, no joy was detected and considerable contempt was detected during smile efforts. (Below, right) Postoperatively, this was corrected. This patient had longstanding right unilateral flaccid facial palsy after superficial parotidectomy for an inconclusive parotid mass corrected by gracilis functional free muscle transfer innervated by the masseteric nerve. He also underwent two ribbon static facial suspension with two ribbons of fascia lata. Postoperative photograph was taken after 0.8 year.

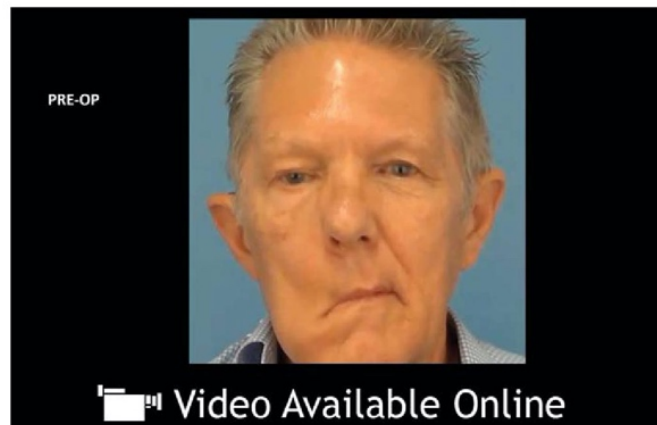
Comparison of Oral Commissure Excursion to the Emotionality Quotient

Of 223 gracilis functional free muscle transfer patients reviewed, 134 met criteria for inclusion. No correlation was found between commissure excursion and the probability of joy perception ($p = 0.433$) or negative emotion perception ($p = 0.527$) (Fig. 5). No correlation was found even when analyzing flaccid and nonflaccid facial palsy

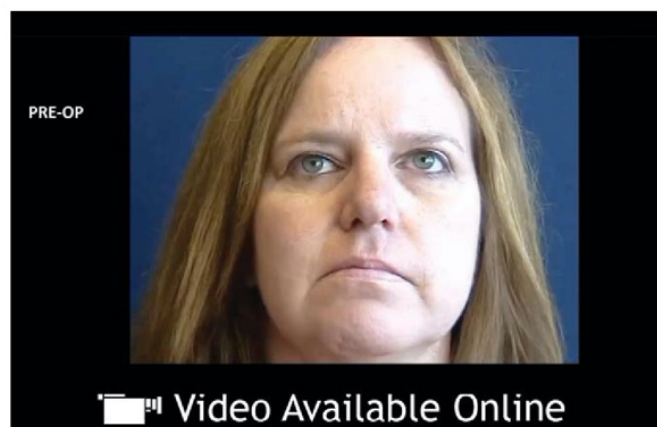
separately (data not shown). Several patients with less than 3-mm commissure excursion were perceived to express considerable levels of joy. [See **Figure, Supplemental Digital Content 6**, which shows rest and voluntary smile photographs of a patient postoperatively with discordant perceived joy expression and oral commissure excursion. (Above) This patient generated –3.3 mm of oral commissure excursion during voluntary smiling



Video 1. Supplemental Digital Content 2 shows preoperative and postoperative maximal smile efforts in a flaccid facial palsy patient with graphic representations of automated output of the emotion detection software, <http://links.lww.com/PRS/D583>.



Video 2. Supplemental Digital Content 3 shows preoperative and postoperative maximal smile efforts in a flaccid facial palsy patient with graphic representations of automated output of the emotion detection software, <http://links.lww.com/PRS/D584>.



Video 3. Supplemental Digital Content 5 shows preoperative and postoperative maximal smile efforts in a nonflaccid facial palsy patient with graphic representations of automated output of the emotion-detection software, <http://links.lww.com/PRS/D589>.



Video 4. Supplemental Digital Content 7 shows two postoperative patients with discordant perceived joy expression and oral commissure excursion. Video refers to Video 1, Supplemental Digital Content 2, and Video 2, Supplemental Digital Content 3, <http://links.lww.com/PRS/D593>.

(i.e., oral commissure moves toward the midline during smiling), but the probability of perceived joy from voluntary smiling was 62 percent. This patient had a right vestibular schwannoma and insidious loss of right facial function. He underwent facial reanimation with gracilis functional free muscle transfer innervated by the masseteric nerve. He also had two ribbons of fascia lata to correct his nasolabial fold position and nasal valve position. The postoperative photograph was taken after 0.6 year. (Below) Graphic representations of automated output of the emotion-detection software from a video of the following sequence: rest, smile, rest, smile, rest, <http://links.lww.com/PRS/D592>. See **Video, Supplemental Digital Content 7**, which shows two postoperative patients with discordant perceived joy expression and oral commissure excursion. Video refers to **Video 1, Supplemental Digital Content 2**, and **Video 2, Supplemental Digital Content 3**, <http://links.lww.com/PRS/D593>.] Furthermore, three of 11 patients (27 percent) with good commissure excursion (between 10 and 15 mm) demonstrated minimal probability of joy perception. [See **Figure, Supplemental Digital Content 8**, which shows rest and voluntary smile photographs of a patient postoperatively with discordant perceived joy expression and oral commissure excursion. (Above) This patient generated 11-mm oral commissure excursion during voluntary smiling, but the probability of perceived joy was 0 percent. This patient had facial nerve schwannoma that was resected, leading to right flaccid facial palsy. He underwent facial reanimation with gracilis functional free muscle transfer innervated by the

masseteric nerve and a cross-face nerve graft in a dual-innervation pattern. He also had nasal valve correction with a band of fascia lata. The postoperative photograph was taken after 0.8 year. (Below) Graphic representations of automated output of the emotion detection software from a video of the following sequence: rest, smile, rest, smile, rest, <http://links.lww.com/PRS/D595>.]

Comparison of Layperson Assessment of Disfigurement to the Emotionality Quotient

In videos of smile efforts of a spectrum of facial palsy patients ($n = 61$), mean layperson-assessed disfigurement scores showed a positive correlation with perceived negative emotion (correlation coefficient, 0.516; $p < 0.001$) and a negative correlation with perceived joy (correlation coefficient, -0.452 ; $p < 0.001$) (Fig. 6).

DISCUSSION

Despite decades of effort, many measures used to assess facial palsy rehabilitation outcomes remain prone to bias and error.²¹ Moreover, measures not developed for the purpose of quantifying success after smile reanimation surgery remain insensitive to outcomes of these procedures.²² Presently, many authors report success after smile reanimation surgery according to clinician-graded scales of success.^{23–26} The Terzis and Noah scale, a five-point Likert scale ranging from poor to excellent, is unreliable, insensitive, and prone to confirmation bias.²² It is also commonplace to quantify successful outcomes using oral commissure excursion, as measured by

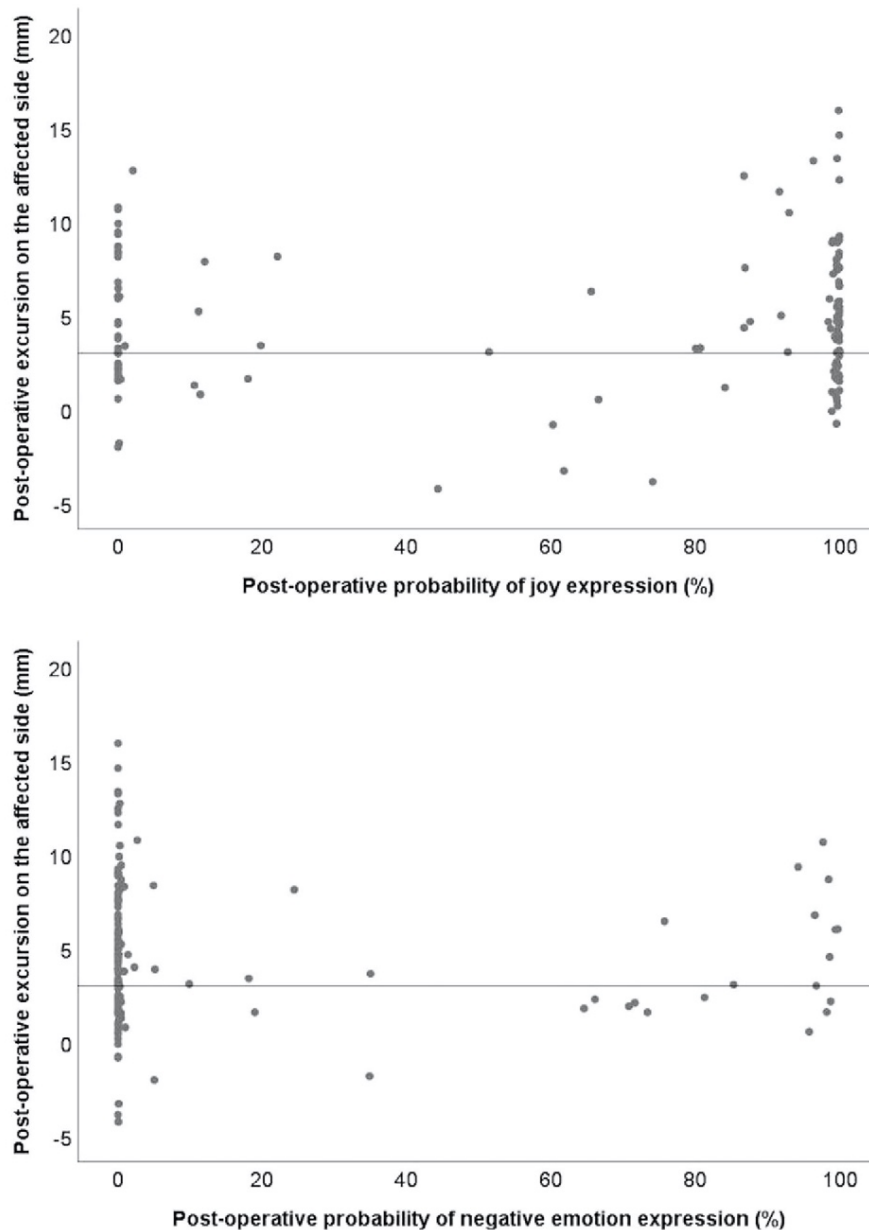


Fig. 5. Scatter plots of postoperative excursion and probability of perceived joy (*above*) and perceived negative emotion (*below*). Horizontal line placed at 3-mm excursion to represent commonly used threshold for defining failure in smile reanimation. Negative values represent a movement of the affected side oral commissure toward the facial midline during a voluntary smile effort, as is commonly seen in flaccid facial palsy preoperatively. No correlation was found between excursion and the probability of perceiving joy (correlation coefficient, 0.068; $p = 0.433$) or negative emotion (correlation coefficient, -0.055 ; $p = 0.527$).

software (e.g., MEEI Facegram and Emotrics) or manually.^{18,22,27} Such measurement systems generate an absolute measure of commissure excursion (usually between 3 and 15 mm) but do not take into consideration the overall aesthetic outcome of the face, or whether the voluntary smile is able to communicate joy. Layperson assessment may

be the ideal method of assessing these additional considerations.^{20,28–30} However, it is not practical to include layperson assessments in routine clinical care and thus a proxy for layperson assessments would be extremely useful.

In this study, artificial intelligence software was used to detect expressions of emotion in the

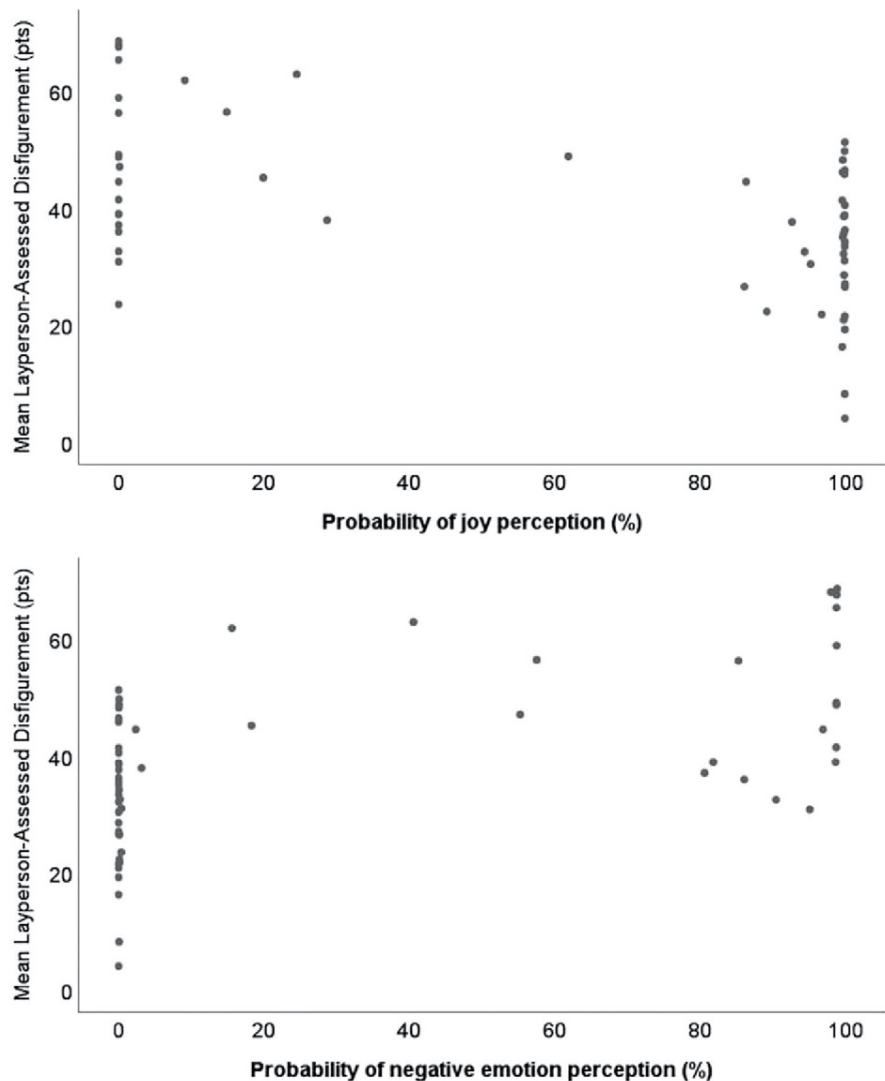


Fig. 6. Scatter plots of mean layperson-assessed disfigurement and probability of perceived joy (*above*) and perceived negative emotion (*below*). Mean layperson-assessed disfigurement score showed a significant positive correlation with perceived negative emotion (correlation coefficient, 0.516; $p < 0.001$) and a significant negative correlation with perceived joy (correlation coefficient, -0.452 ; $p < 0.001$).

voluntary smiles of facial palsy patients, to quantify the deleterious effect of facial palsy on the ability to communicate nonverbally, and to quantify the sensitivity to change of the emotionality quotient by comparing preoperative and postoperative smiles. To determine the role of the emotionality quotient within the current paradigm of facial palsy outcome assessment, we compared emotionality quotient scores to commissure excursion in postoperative patients. We also compared emotionality quotient scores to layperson data, generated from the same videos previously viewed by 593 laypersons.

We found that the software perceived 100 percent joy in smiles of normal controls and 0 percent

joy in flaccid facial palsy. The emotionality quotient was sensitive to change and highly reliable. In flaccid facial palsy, the perception of joy was increased by a median 84 percent after smile reanimation. Both flaccid and nonflaccid facial palsy also demonstrated a high level of perceived negative emotion preoperatively that was reversed to minimal levels after surgery. These findings were highly statistically significant ($p < 0.001$). In comparing the emotionality quotient to oral commissure excursion, we did not find a correlation with either joy or negative emotion perception, regardless of whether patients were flaccid or nonflaccid preoperatively. These unexpected findings

warrant further investigation to identify the key factors that influence communication of nonverbal information in facial palsy. Finally, we found a correlation between layperson-assessed disfigurement and both perceived negative emotion, which was correlated with high facial disfigurement (correlation coefficient, 0.516; $p < 0.001$), and perceived joy, which was negatively correlated (correlation coefficient, -0.452 ; $p < 0.001$). A major strength of this analysis was that the same video clips were viewed by a large number of laypersons and our computer vision software, and that the patients had been selected, a priori, to represent the full spectrum of facial palsy, raising the possibility of emotionality quotient functioning as a proxy for layperson assessment.²⁰

A facial recognition and emotion detection tool trained using machine learning from millions of encounters with normal subjects could represent a standard to which facial reanimation results are compared. It is known that laypersons perceive the facial expressions of facial palsy patients negatively when compared to normal subjects, and significant improvements in layperson assessment of facial disfigurement have been shown following smile reanimation.^{29,31,32} The fact that an artificial intelligence algorithm detects a high probability of a negative emotion does not mean the patient in fact appears contemptuous or disgusting, but rather that there are features of these expressions picked up by the algorithm. That these features correspond to cases where laypersons consider there is a high degree of disfigurement tells us that a high negative emotion score from the emotionality quotient may be a proxy for disfigurement. The emotionality quotient is an objective and automatic outcome measure that could be used as an automated layperson equivalent, assessing facial palsy patients without having to show clinical images of patients to the public.

Our results have quantified the significant nonverbal communication deficits suffered by facial palsy patients and the dramatic improvements afforded by smile reanimation. Attempting to smile with flaccid facial palsy produces a unilateral cheek raise, formation of a nasolabial fold, and upper lip raise (Fig. 2, *left*). These action units are hallmark features of contempt (see **Appendix, Supplemental Digital Content 1**, <http://links.lww.com/PRS/D582>). In contrast to flaccid facial palsy patients, those with nonflaccid facial palsy may also form a nasolabial fold on the affected side (Fig. 1, *right*), allowing for perception of higher levels of joy preoperatively, in addition to increased levels of negative emotion. In addition,

the synkinetic oral commissure is often immobile, leading to what appears like a version of disgust when bilateral lip raise, cheek raise, and nose wrinkle are present without lateral lip corner pull (Fig. 1, *right*).

Interestingly, we did not find a correlation between oral commissure excursion and either joy or negative emotion perception, suggesting that success of smile reanimation to convey meaning may be based on several factors other than oral commissure excursion (Fig. 5). Although formal analysis is required to identify these other factors, we hypothesize that contralateral side movements, preexisting facial function or synkinesis, vector of pull of the gracilis muscle, static procedures, and patient age may all play a role in the discordance between oral commissure excursion and perceived emotionality (see **Figure, Supplemental Digital Content 8**, <http://links.lww.com/PRS/D595>). In contrast to excursion, layperson-assessed disfigurement was moderately yet significantly correlated with perceived emotion (Fig. 6). It is likely that this moderate level of correlation is attributable to the vastly different methodologies used to obtain the results in each group (i.e., mean layperson-assessed disfigurement versus computerized calculation of perceived emotion probability). We expect that a more highly correlated proxy would be generated from comparing emotionality quotient to layperson assessments of perceived emotion in facial palsy. Dedicated studies comparing these outcome measures are required to validate the emotionality quotient as a layperson proxy. It is intriguing that the output of an artificial intelligence-generated facial analysis tool is correlated with layperson assessments of facial disfigurement, yet both perceived joy and negative emotion were correlated and highly statistically significant. It is possible that the emotionality quotient may have the capacity to serve as an automated layperson equivalent, and emotionality quotient delta (the change in the emotionality quotient after an intervention) could supplant or complement commissure excursion as a quantifiable outcome measure of success or failure in smile reanimation. Further studies comparing the emotionality quotient to clinician grading and layperson assessments of outcomes are required to determine what level of emotionality quotient delta is clinically significant.

The present study only applied artificial intelligence emotion detection software to the analysis of voluntary smile efforts. We anticipate many more potential uses for this software as our understanding of emotionality in facial palsy increases.

For example, the real-time assessment capability of the emotionality quotient could be used to give additional feedback during physical therapy rehabilitation. It is also foreseeable that emotionality assessments will be useful preoperatively. For instance, in synkinetic patients, knowledge of preoperatively perceived joy and negative emotion would help to set treatment goals and manage expectations regarding potential improvements.

Limitations and Confounders

As computer vision analyzes the whole face, it may be confounded by multiple factors, including healthy side facial movements, synkinesis, and the effects of other static and dynamic surgical procedures. Affdex is dependent on the quality of photography and may fail to detect a face if the lateral border of the cheek has been cropped, for example, which was a problem in several of our cases. Other factors such as age, facial hair, gross facial abnormality, soft-tissue thickness, scarring, radiotherapy change, or bony abnormality may impair the ability of the software to detect the face or subtle facial movements. Spontaneous facial nerve recovery may also confound results, as we have shown that preoperative emotionality quotient of flaccid facial palsy is much worse than that of nonflaccid facial palsy patients. Thus, any improvement in emotionality quotient delta is likely to be more significant if a preoperative photograph is taken in the flaccid state compared with the nonflaccid state. We chose to focus on flaccid facial palsy for our analysis because of the homogeneity of the patient population and the isolated role of the gracilis muscle in producing smile on the affected side. Heterogeneity of the nonflaccid facial palsy population makes any gross statistical comparisons with flaccid facial palsy invalid. Further studies should be used to understand the emotionality penalty paid by incomplete facial palsy patients and varying degrees of synkinesis.

The artificial intelligence software used was not trained with images of facial palsy patients and thus may be inaccurate for determining the probability of stereotypical expressions of emotion in facial palsy patients in the preoperative state. However, an ideal of facial reanimation specialists is for our patients to appear normal in the eyes of naive observers, and so software validated from millions of recordings of normal facial expressions may well be an appropriate outcome measurement tool to use, particularly in the postoperative state. Interestingly, we did find a moderate correlation between perceived emotionality and layperson assessment of facial palsy in the

preoperative state. A major advantage of using layperson assessments to quantify outcomes is that they are the recipients of nonverbal communication cues expressed by facial palsy patients. However, many laypersons may also be unfamiliar with the appearance of facial palsy. Machine learning models generate predictions of the emotional state being communicated by facial palsy patients based on an enormous database of facial expressions of normal subjects. As such, the emotionality quotient may perhaps be used as an automated layperson equivalent. Future work should investigate whether computer vision-assessed probability of perceived emotion among patients with unilateral facial palsy corresponds to other previously studied domains of layperson assessments, such as attractiveness, approachability, and negative facial perception, and novel domains, such as layperson-assessed emotionality.^{31,33}

Although it would have been very interesting to include, the present study did not analyze spontaneous smiles. To clearly understand the efficacy of the emotionality quotient at detecting the communication of joy, we analyzed voluntary smile frames only, as quantifying spontaneous smiling remains subjective. Likewise, this study did not seek to make inferences about the felt emotional state, because of the difficulties of inferring true emotional state from facial expressions alone.^{34,35}

CONCLUSIONS

Artificial intelligence computer vision software perceived a lower probability of joy and an increased probability of negative emotion expression in smiles of facial palsy patients compared with normal subjects. After smile reanimation, significantly more joy and less negative emotion were detected. The emotionality quotient was reliable, highly sensitive to change after smile reanimation, and correlated with layperson assessments of facial disfigurement for both joy and negative emotion perception. Previously, commissure excursion has been used as an outcome measure of the technical success of gracilis functional free muscle transfer. However, we found that excursion was poorly correlated with perceived joy and negative emotion expression, indicating that there is more to facial expression of joy than simply oral commissure excursion. This facial recognition and emotion detection software has quantified the ability of smile reanimation surgery to alleviate the deficits in nonverbal communication suffered by facial palsy patients. The efficiency, sensitivity, and objectivity of the emotionality quotient render it

an extremely attractive tool that may serve as a proxy for layperson assessment, an ideal outcome measure in facial palsy.

Joseph R. Dusseldorp, F.R.A.C.S.
Massachusetts Eye and Ear Infirmary
243 Charles Street
Boston, Mass. 02114
drjoedusseldorp@icloud.com
Facebook: joedusseldorp
Instagram: @drjoedusseldorp
Twitter: @drjoedusseldorp

PATIENT CONSENT

Patients provided written consent for the use of their images.

ACKNOWLEDGMENTS

The authors would like to acknowledge Lisa Feldman Barrett for advising as to the correct terminology of emotion used in this article; Sofia Lyford-Pike, Nathaniel Helwig, and Stephen J. Guy for sharing the layperson assessment data; and Emily Fortier for assistance with recording of the videos.

REFERENCES

- Adolphs R. The social brain: Neural basis of social knowledge. *Annu Rev Psychol*. 2009;60:693–716.
- Frith C. Role of facial expressions in social interactions. *Philos Trans R Soc Lond B Biol Sci*. 2009;364:3453–3458.
- Coulson SE, O'dwyer NJ, Adams RD, Croxson GR. Expression of emotion and quality of life after facial nerve paralysis. *Otol Neurotol*. 2004;25:1014–1019.
- Ho AL, Scott AM, Klassen AF, Cano SJ, Pusic AL, Van Laeken N. Measuring quality of life and patient satisfaction in facial paralysis patients: A systematic review of patient-reported outcome measures. *Plast Reconstr Surg*. 2012;130:91–99.
- Hjortsjö CH. *Man's Face and Mimic Language*. Lund: Studentlitteratur; 1970. Available at: <https://books.google.com/books?id=r68ZMwEACAAJ>. Accessed March 1, 2018.
- Ekman P, Friesen WV. *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Palo, Alto, Calif: Consulting Psychologists Press; 1978.
- Ekman P, Friesen WV. *Unmasking the Face: A Guide to Recognizing Emotions from Facial Clues*. Available at: doi:10.1163/1574-9347_bnp_e804940. Accessed March 21, 2018.
- Ekman P, Rosenberg E. *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS)*. New York: Oxford University Press; 2005.
- Ekman P. Facial expression and emotion. *Am Psychol*. 1993;48:384–392.
- Pantic M, Rothkrantz LJM. Automatic analysis of facial expressions: The state of the art. *IEEE Trans Pattern Anal Mach Intell*. 2000;22:1424–1445.
- McDuff D, El Kaliouby R, Senechal T, Demirdjian D, Picard R. Automatic measurement of ad preferences from facial responses gathered over the Internet. *Image Vis Comput*. 2014;32:630–640.
- El Kaliouby R, Robinson P. Real-time inference of complex mental states from facial expressions and head gestures. In: Kisačanić B, Pavlović V, Huang TS, eds. *Real-Time Vision for Human-Computer Interaction*. Boston, Mass: Springer; 2005:181–200.
- McDuff D, El Kaliouby R, Picard RW. Crowdsourcing facial responses to online videos: Extended abstract. In: *Proceedings of the 2015 International Conference on Affective Computing and Intelligent Interaction*. New York: IEEE Xplore; 2015:512–518.
- Baltrušaitis T, McDuff D, Banda N, et al. Real-time inference of mental states from facial expressions and upper body gestures. Available at: <https://www.cl.cam.ac.uk/~pr10/publications/fera11.pdf>. Accessed March 24, 2018.
- Kaliouby R, Kodra E, Jha P, Mediratta N. Does facial coding generalize across cultures? A spotlight on Asia. Available at: https://www.affective.com/wp-content/uploads/2017/03/Does_Facial_Coding_Generalize_Across_Cultures_ASIA.pdf. Accessed March 24, 2018.
- Reber AS, Allen R, Reber ES. *Penguin Dictionary of Psychology*. New York: Penguin; 2009.
- Barrett LF. *How Emotions Are Made: The Secret Life of the Brain*. New York: Houghton-Mifflin-Harcourt; 2017.
- Guarin DL, Dusseldorp J, Hadlock TA, Jowett N. A machine learning approach for automated facial measurements in facial palsy. *JAMA Facial Plast Surg*. 2018;20:335–337.
- Bhama PK, Weinberg JS, Lindsay RW, Hohman MH, Cheney ML, Hadlock TA. Objective outcomes analysis following microvascular gracilis transfer for facial reanimation: A review of 10 years' experience. *JAMA Facial Plast Surg*. 2014;16:85–92.
- Lyford-Pike S, Helwig NE, Sohre NE, Guy SJ, Hadlock TA. Predicting perceived disfigurement from facial function in patients with unilateral paralysis. *Plast Reconstr Surg*. 2018;142:722e–728e.
- Manktelow RT, Zuker RM, Tomat LR. Facial paralysis measurement with a handheld ruler. *Plast Reconstr Surg*. 2008;121:435–442.
- Fattah AY, Gurusinghe ADR, Gavilan J, et al. Facial nerve grading instruments: Systematic review of the literature and suggestion for uniformity. *Plast Reconstr Surg*. 2015;135:569–579.
- Terzis JK, Noah ME. Analysis of 100 cases of free-muscle transplantation for facial paralysis. *Plast Reconstr Surg*. 1997;99:1905–1921.
- Biglioli F, Colombo V, Tarabbia F, et al. Double innervation in free-flap surgery for long-standing facial paralysis. *J Plast Reconstr Aesthet Surg*. 2012;65:1343–1349.
- Cardenas-Mejia A, Covarrubias-Ramirez JV, Bello-Margolis A, Rozen S. Double innervated free functional muscle transfer for facial reanimation. *J Plast Surg Hand Surg*. 2015;49:183–188.
- Biglioli F, Bayoudh W, Colombo V, Pedrazzoli M, Rabbiosi D. Double innervation (facial/masseter) on the gracilis flap, in the middle face reanimation in the management of facial paralysis: A new concept (in French). *Ann Chir Plast Esthet*. 2013;58:89–95.
- Hadlock TA, Urban LS. Toward a universal, automated facial measurement tool in facial reanimation. *Arch Facial Plast Surg*. 2012;14:277–282.
- Hadlock T. Standard outcome measures in facial paralysis: Getting on the same page. *JAMA Facial Plast Surg*. 2016;18:85–86.
- Dey JK, Ishii LE, Byrne PJ, Boahene KD, Ishii M. Seeing is believing: Objectively evaluating the impact of facial reanimation surgery on social perception. *Laryngoscope*. 2014;124:2489–2497.
- Helwig NE, Sohre NE, Ruprecht MR, Guy SJ, Lyford-Pike S. Dynamic properties of successful smiles. *PLoS One*. 2017;12:e0179708.



31. Dey JK, Ishii M, Boahene KD, Byrne PJ, Ishii LE. Changing perception: Facial reanimation surgery improves attractiveness and decreases negative facial perception. *Laryngoscope* 2014;124:84–90.
32. Ishii L, Godoy A, Encarnacion CO, Byrne PJ, Boahene KD, Ishii M. Not just another face in the crowd: Society's perceptions of facial paralysis. *Laryngoscope* 2012;122:533–538.
33. Ishii L, Dey J, Boahene KD, Byrne PJ, Ishii M. The social distraction of facial paralysis: Objective measurement of social attention using eye-tracking. *Laryngoscope* 2016;126:334–339.
34. Barrett LF. Solving the emotion paradox: Categorization and the experience of emotion. *Pers Soc Psychol Rev* 2006;10:20–46.
35. Barrett LF, Mesquita B, Ochsner KN, Gross JJ. The experience of emotion. *Annu Rev Psychol* 2007;58:373–403.

To capture the complete picture of the impact of the above article, I have reproduced the Invited Discussion that accompanied publication of the above article in PRS in 2019:

Chuang D. In the Eye of the Beholder: Changes in Perceived Emotion Expression after Smile Reanimation. *Plast. Recon. Surg.* 2019 Aug; 144(2):472-473

DISCUSSION

Discussion: In the Eye of the Beholder: Changes in Perceived Emotion Expression after Smile Reanimation

David Chwei-Chin Chuang, M.D.
Taipei-Linkou, Taiwan

The authors presented an interesting but debatable study.¹ In the Introduction section, they mentioned that the purpose of the study was to discover a tool that can quantify layperson assessments of facial palsy. They found that artificial intelligence computer software could be applied and become a proxy for layperson assessments, and concluded it could be an attractive tool to replace layperson assessment as an ideal outcome measurement in facial palsy. I am reluctant to accept their explanations, and am not convinced.

I agree with the authors that using layperson assessments to perform outcome study of facial palsy patients after facial reanimation surgery (the main operation performed was functioning free muscle transplantation) is impractical. However, I do not agree that of the multiple outcome measures available today, none has gained widespread clinical application. Although many evaluation system have been developed for outcome evaluation,²⁻⁵ they might not achieve consensus. However, they are truly in the eye of the naive "human" beholder. Lack of consensus in a debated issue for outcome assessment is seen quite often in science, isn't it? There are many debates but few conclusions, such as outcome assessments of obstetric brachial plexus palsy, adult brachial plexus injury, and facial paralysis. These developed evaluation systems are straightforward. Once you see the patient, you can immediately give the score. It is, of course, more subjective. This is why they often need more than one (usually from two to five) observers to evaluate the results to avoid its subjectivity.

The Facial Action Coding System is a method of quantifying normal facial expressions. Because of normal people, it can analyze, for example, joy, disgust, contempt, or smirk expressions. As the authors described, it requires a high level of

training for observers, it is time-consuming, and yields only moderate interrater reliability. In the layperson assessment of facial palsy, the authors analyzed 593 layperson assessments of the smile videos (2016 Minnesota State Fair) and 61 videos of preoperative unilateral facial palsy patients (from 500 patients videos in their own hospital). Results showed that perceived negative emotion had positive correlation with layperson assessment disfigurement, but perceived joy (positive emotion) was negatively correlated. This is a strange conclusion because (1) all videos were of facial palsy patients with disfigurement, (2) all were preoperative videos, and (3) all analyses were performed by (fixed) just a few observers. The described results are not surprising. However, it cannot be concluded that the perceived emotion detected by artificial intelligence would be as effective as layperson assessment in the postoperative outcome evaluation. In addition, using layperson assessment of disfigurement for comparison is also incorrect. In layperson assessment, the layperson often overestimates the negative emotion perception preoperatively, and underestimates the results of smile after functional free muscle transfer. Layperson assessment of disfigurement is good for normal people and preoperative facial palsy patients, but not good for postoperative patients. They cannot detect the patient's emotionality postoperatively. This is why the authors found no correlation between oral commissure excursion and emotionality quotient, because both (emotionality quotient and layperson assessment) are incorrect. Our investigation (over 400 functional free muscle transfers for facial reanimation) shows that it is a quite positive relationship: better (Hallock) lip excursion score, and better perceived joy results.

In the emotionality quotient or emotionality quotient delta, the authors used positive emotion (joy) and negative emotion (contempt or disgust) as their evaluation system. Negative emotion does not exist in facial palsy patients preoperatively

From the Department of Plastic Surgery, Chang Gung Memorial Hospital.

Received for publication August 21, 2018; accepted January 7, 2019.

Copyright © 2019 by the American Society of Plastic Surgeons

DOI: 10.1097/PRS.00000000000005872

Disclosure: The author has no financial interest to declare in relation to the content of this Discussion.

because the muscles are paralyzed. Postoperatively, after facial reanimation surgery, the negative emotion of contempt or disgust still does not exist because functional free muscle transfer is only for joy reconstruction (the “smirk” is actually more rare than facial expression for contempt or disgust; smirk is a type of smile but incomplete).

Facial reanimation surgery has many reconstructions, including static and dynamic methods, local muscle, and distal muscle transfer. In this study, the authors used distal muscle transfer or functional free muscle transfer for the majority of cases. However, functional free muscle transfer does not pretend to restore complete facial movement, merely movement of the oral commissure in one or two vectors. It is for joy, not for negative emotion. The emotionality quotient, or emotionality quotient delta, included joy and negative emotions. I might believe that both are only benefits for preoperative judgment, but not for postoperative judgment after functional free muscle transfer reconstruction.

“Perceive,” from *Webster’s Dictionary*, means to “become aware of through the senses.” To the human being, perceived sensation means the senses are inherent, naive, and born with. Perceived positive and negative emotion here is by an artificial intelligence computer. However, its score system is made by a human being and not inherent. In the title “In the Eye of the Beholder,” the beholder is not easily understood by non-English-speaking people. Why not just say, “In the eye of the artificial intelligence machine vision”? This would make it more straightforward and easily understood.

The authors overemphasized the benefits of artificial intelligence: free, easy, quick, and simple. This is much better than any layperson assessment study, which is difficult, slow, and time-consuming. Negative emotions do not exist in facial palsy patients preoperatively. It is also a difficulty seen in patients after functional free muscle transfer

facial reanimation surgery (the only benefit is pulling up the upper lip or oral commissure for smile).

In summary, the article is full of uncertain words and strange explanations. Artificial intelligence software detection is also full of “probability” and “prediction.” Artificial intelligence also cannot analyze spontaneous smiles, which is very important for result evaluation. In the current situation, its applicability is very low. I could not accept it as an ideal outcome measurement in facial palsy. However, this description of new application of artificial intelligence may stimulate others to consider it as another option or to continuously improve the software to evaluate facial palsy patients.

David Chwei-Chin Chuang, M.D.
Department of Plastic Surgery
Chang Gung Memorial Hospital
199 Tung Hwa North Road
Taipei-Linkou 105, Taiwan
dccchuang@gmail.com

REFERENCES

1. Dusseldorp JR, Guarin DL, van Veem MM, Jowett M, Hadlock TA. In the eye of the beholder: Changes in perceived emotion expression after smile reanimation. *Plast Reconstr Surg*. 2019;144:457–471.
2. Coulson SE, Croxson GR, Adams RD, O’Dwyer NJ. Reliability of the “Sydney,” “Sunnybrook,” and “House Brackmann” facial grading systems to assess voluntary movement and synkinesis after facial nerve paralysis. *Otolaryngol Head Neck Surg*. 2005;132:543–549.
3. Terzis JK, Noah ME. Analysis of 100 cases of free-muscle transplantation for facial paralysis. *Plast Reconstr Surg*. 1997;99:1905–1921.
4. Chuang DCC, Lu JCY, Anesti K. One-stage procedure using spinal accessory nerve (XI)-innervated free muscle for facial paralysis reconstruction. *Plast Reconstr Surg*. 2013;132:117e–129e.
5. Lin JT, Lu JC, Chang TN, Chuang DC. Simultaneous reconstruction of the lower lip with gracilis functioning free muscle transplantation for facial reanimation: Comparison of different techniques. *Plast Reconstr Surg*. 2018;142:1307–1317.

I have also included a Letter to the Editor that was received along with my Response:

In the Eye of the Beholder: Changes in Perceived Emotion Expression after Smile Reanimation

Sir:

With great interest, we have read the article by Dusseldorp et al., “In the Eye of the Beholder: Changes in Perceived Emotion Expression after Smile Reanimation.”¹ The authors used computer vision artificial intelligence software to detect less joy and more negative emotion in smiles of facial palsy patients compared with normal subjects. After smile reanimation, perception of joy increased and negative emotion decreased.

We completely support their statement that smile reanimation improves the perception of negative emotions and more joy. However, it is also of great importance to realize that the central area of the face, especially around the eyes, has an essential impact on the perception of beauty and perceived emotion; we

1109e

s. Unauthorized reproduction of this article is prohibited.

also know from aesthetic surgery that rejuvenation around the eyes has a great impact on our perception of “improved” beauty; therefore, one can state that “beauty is around the eyes (and mouth) of the beheld.”² In former days, psychologists demonstrated by eye tracking studies how laypersons view other faces: mainly, the centropalpebral area (around the eyes, nose, and lips).³ Tonnard et al. demonstrated that the appearance of the eyes and lips (next to smooth skin) are universal aspects of beautiful faces.⁴

Therefore, it is quite logical that asymmetry around the mouth and/or eyes in facial palsy patients immediately will be noticed. Improvement of the area of the lips, as demonstrated in the study by Dusseldorp et al., thus results in improved perception of emotion. However, we should not forget to improve also the periorbital area in facial palsy patients. In an earlier study, we showed that Hering’s law is also applicable to the frontal facial branch in facial palsy patients: the eyebrow position of the unaffected side will increase in height, with also an increased amount of frontal lines as compared to the affected side, another aspect that results in asymmetry of the periorbital area.⁵ Knoll et al. already showed that such an eyebrow position has a significant effect on the perception of perceived emotions.⁶ Treating the forehead of the unaffected side in facial palsy patients with botulinum toxin can be used to decrease the overelevation of the eyebrow and to reduce the amount of frontal lines. In addition, thread lifts or direct eyebrow lifts can be used to heighten the lower eyebrow position of the affected side. By means of these treatment modalities, one can more completely restore the asymmetry of the centropalpebral area of facial palsy patients, thus achieving a better postoperative result.

DOI: 10.1097/PRS.00000000000006854

Haydar Aslan Gülbitti, M.D.

Department of Plastic Surgery
University of Groningen
University Medical Centre Groningen
Groningen, The Netherlands

Berend van der Lei, M.D., Ph.D.

Department of Plastic Surgery
University of Groningen
University Medical Centre Groningen
Groningen, The Netherlands
Bey Bergman Clinics
Heerenveen, Zwolle, and Hilversum, The Netherlands

Correspondence to Prof. Dr. Berend van der Lei
University Medical Center Groningen
P.O. Box 30.001
9700 RB Groningen, The Netherlands
h.a.gulbitti@gmail.com
Instagram: @haydaraslan

DISCLOSURE

The authors have no financial interest to declare in relation to the content of this communication, and there has been no financial support for this work.

1110e

REFERENCES

1. Dusseldorp JR, Guarin DL, van Veen MM, Jowett N, Hadlock TA. In the eye of the beholder: Changes in perceived emotion expression after smile reanimation. *Plast Reconstr Surg*. 2019;144:457–471.
2. Gülbitti HA, Bouman TK, Marten TJ, van der Lei B. The orbital oval balance principle: A morphometric clinical analysis. *Plast Reconstr Surg*. 2018;142:451e–461e.
3. Janik SW, Wellens AR, Goldberg ML, Dell’Osso LF. Eyes as the center of focus in the visual examination of human faces. *Percept Mot Skills* 1978;47:857–858.
4. Tonnard PL, Verpaele AM, Bensimon RH. *Centropalpebral Rejuvenation*. New York: Thieme Verlag Medical Publishers; 2018.
5. Gülbitti HA, van der Lei B. Hering’s law of the frontal facial branch. *Plast Reconstr Surg*. 2018;142:991e–992e.
6. Knoll BI, Attkiss KJ, Persing JA. The influence of forehead, brow, and periorbital aesthetics on perceived expression in the youthful face. *Plast Reconstr Surg*. 2008;121:1793–1802.

Response to Gulbitti, et. al. **Dusseldorp JR**, Hadlock TA. Using artificial intelligence to measure facial expression following facial reanimation surgery. *Plast. Recon. Surg.* 2022 Jan;149(2):343e-345e

Reply: In the Eye of the Beholder: Changes in Perceived Emotion Expression after Smile Reanimation

Sir:

We thank our colleagues for their considered response to our article.¹ Eye-tracking studies demonstrate that observers focus mostly on the eyes and the central triangle of the face, yet the layperson's attention when viewing paralyzed hemifaces is distracted toward the mouth.² Layperson assessments of facial palsy patients are considered gold standard, and evaluation systems using "two to five" observers are required to avoid subjectivity.³ Either observers are naive, and thus are laypersons, or they are familiar with the procedure, and thus potentially biased. Human-based measurement systems will always be prone to subjectivity and bias.

Success in our study was defined as a facial expression that a machine learning algorithm perceived as joyful. However, the data also demonstrated there was negative emotion detected in preoperative smiles of facial palsy patients, and this disappeared after successful smile reanimation.¹ It is pleasing to us that the authors have recognized the importance of smile reanimation to reduce the expression of negative emotion. However, others have asserted that "negative emotion does not exist in facial palsy patients preoperatively because the muscles are paralyzed."³ This thinking pattern underscores a failure to realize that the naive observer looks on a person's whole face. When a unilateral facial palsy patient smiles, the healthy side of their face contracts, giving an overall negative impression to an observer (*Fig. 1, center*). Various authors have labeled this phenomenon differently: reduced attractiveness, negative facial perception, disfigurement, negative emotion,¹ smirk,³ or contempt.¹ The data are clear that there is an overall negative perception of the facial appearance of someone who smiles with a paralyzed hemiface, and smile reanimation improves this. This finding is real. We can now quantify it; thus, we

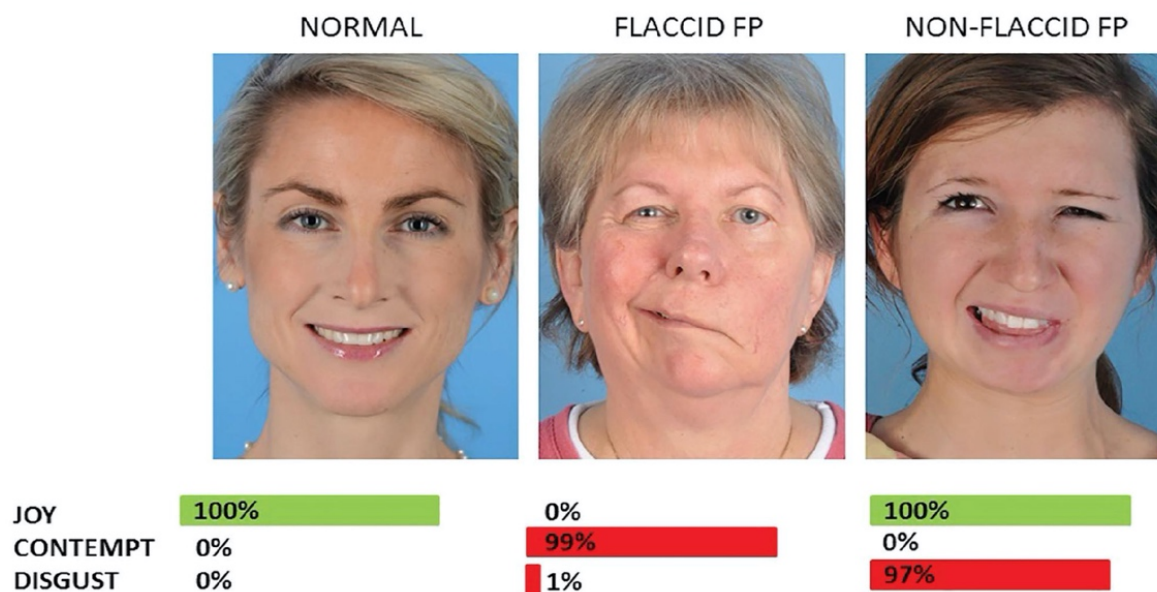


Fig. 1. Maximal smile efforts in (left) a normal subject (high levels of joy and no negative emotion are perceived), (center) one with flaccid facial palsy (no joy and high levels of contempt are perceived), and (right) one with nonflaccid facial palsy (high levels of joy and disgust are perceived). Both flaccid and nonflaccid patients are affected on their left sides. *FP*, facial palsy. (From Dusseldorp JR, Guarin DL, van Veen MM, Jowett N, Hadlock TA. In the eye of the beholder: Changes in perceived emotion expression after smile reanimation. *Plast Reconstr Surg.* 2019;144:457–471)

have represented it as the denominator of the emotionality quotient.

It has been said that if a mental task can be executed in 1 second, it can be automated. There are many advantages to automation, including precision, objectivity, reliability, and standardization. For these reasons, we developed an algorithm that could automatically detect a successful facial expression of joy in a facial palsy patient. Importantly, like a naive observer, a machine learning algorithm does not just look at the mouth to make this assessment—it takes into account the whole face when predicting an emotion expression.

It is true that part of our facial reanimation success may be attributable to the static periocular procedures that we routinely use (including upper eyelid weighting, lateral tarsal strip, lateral tarsorrhaphy, and brow lifting). Using machine learning approaches, we have previously quantified the contributions of periocular deformity to emotion expression across the full spectrum of facial palsy⁴ and also quantified the success of static periocular interventions in restoring normal periocular aesthetics.

As we have consistently published, we consider that a multimodal approach is required for outcome assessment in facial reanimation, including patient-reported outcome measures, automated or clinician-rated scales, layperson assessments (or equivalent), and a spontaneity assessment.⁵ In our view, there will never be a single ideal outcome measure for facial palsy that adequately satisfies all of these domains.

DOI: 10.1097/PRS.00000000000006855

Joseph R. Dusseldorp, F.R.A.C.S.

Department of Otolaryngology/Head and Neck Surgery
Massachusetts Eye and Ear Infirmary and
Harvard Medical School
Boston, Mass.

Department of Plastic and Reconstructive Surgery
Royal Australasian College of Surgeons and
University of Sydney
Sydney, New South Wales, Australia

Department of Plastic and Reconstructive Surgery
Concord Clinical School
Concord, New South Wales, Australia

Tessa A. Hadlock, M.D.

Department of Otolaryngology/Head and Neck Surgery
Massachusetts Eye and Ear Infirmary and
Harvard Medical School
Boston, Mass.

Correspondence to Dr. Dusseldorp
Department of Plastic and Reconstructive Surgery
Concord Clinical School
Hospital Road
Concord, New South Wales 2139, Australia
drjoedusseldorp@icloud.com
Facebook: [joedusseldorp](#)
Instagram: [@drjoedusseldorp](#)
Twitter: [@drjoedusseldorp](#)

PATIENT CONSENT

The patients provided written consent for the use of their images.

1111e

DISCLOSURE

Dr. Dusseldorp has been supported by a Foundation for Surgery Research Scholarship from The Royal Australasian College of Surgeons. The authors have no financial interest to declare in relation to the content of this communication.

REFERENCES

1. Dusseldorp JR, Guarin DL, van Veen MM, Jowett N, Hadlock TA. In the eye of the beholder: Changes in perceived emotion expression after smile reanimation. *Plast Reconstr Surg*. 2019;144:457–471.
2. Dey JK, Ishii LE, Byrne PJ, Boahene KD, Ishii M. Seeing is believing: Objectively evaluating the impact of facial reanimation surgery on social perception. *Laryngoscope* 2014;124:2489–2497.
3. Chuang DC. Discussion: In the eye of the beholder: Changes in perceived emotion expression after smile reanimation. *Plast Reconstr Surg*. 2019;144:472–473.
4. Greene JJ, Guarin DL, Tavares J, et al. The spectrum of facial palsy: The MEEI facial palsy photo and video standard set. *Laryngoscope* 2020;130:32–37.
5. Dusseldorp JR, van Veen MM, Mohan S, Hadlock TA. Outcome tracking in facial palsy. *Otolaryngol Clin North Am*. 2018;51:1033–1050.

1112e

EPILOGUE

Conclusions, Limitations and Suggestions for Future Research

Conclusions

*“Every time you smile at someone, it is an action of love,
a gift to that person, a beautiful thing...”*

We shall never know all the good that a simple smile can do.”

Mother Teresa “The Best Gift is Love: Meditations” Servant Books, 1993

The aim of this thesis was to develop tools that could enable a multi-modal approach to assessing successful smile reanimation outcome in unilateral facial palsy patients. **Chapter One** introduced the concepts of voluntary and spontaneous smiling and the various procedures that are used to reanimate faces of patients living with facial palsy. The key outcome domains that ought to be assessed for someone undergoing surgery to correct their facial palsy were defined as the physical, psychological, and social domains. **Chapter Two** comprised a comprehensive review of existing outcome assessment tools in use in the field of facial reanimation and identified deficiencies in assessing these domains objectively, particularly regarding smile spontaneity. **Chapter Three** demonstrated the importance of spontaneous smiling relative to voluntary smiling and identified the impact of loss of smile ability on overall quality of life using a patient-reported outcome measure: the FaCE scale. It was apparent that existing PROMs such as the FaCE scale adequately captured relevant quality of life data and thus routine implementation of such PROMs to assess the psychological domain of facial reanimation outcomes is warranted.

In the physical domain, it was demonstrated that existing assessment tools were insufficient and so new tools were developed to objectively quantify specific facial movements. The first step in this process was developing the Emotrics app, a machine learning algorithm that was able to predict facial landmark positions from frontal photographs of facial palsy

patients. By training this computational model further it was possible to generate automated assessments of numerous eFACE variables and test their validity against a standardised set of the full spectrum of facial palsy photographs. A major obstacle was to find a way to quantify objectively the nasolabial fold motion in smiles of healthy volunteers and in unilateral facial palsy patients. In **Chapter Four**, state-of-the-art facial landmarking and 3D vector analysis demonstrated that the degree and vector of nasolabial fold excursion is similar to oral commissure excursion, albeit in a 3-dimensional plane that is not well visualised from a 2D frontal facial photograph. As the nasolabial fold is not a facial landmark that has been routinely trained into any existing facial recognition algorithm future work ought to concern itself with manual markings and training of an advanced facial palsy specific model to enable real time quantification of this elusive facial feature.

The focus of the final chapter of the thesis, **Chapter Five**, was to validate tools to objectively quantify the ability of reanimated smiles to communicate joy to a naïve observer, thus addressing the social domain of outcomes assessment. The first step in this process was to develop a standardised method for eliciting and quantifying spontaneous smiles. The validated Spontaneous Smile Assay, a series of humorous videoclips, was used to elicit spontaneous smiles. Next, video recordings were made of the facial expressions of normal subjects, as well as pre-op and post-op reanimation patients while watching these clips. To quantify spontaneous smiling, a blinded method was described for detecting healthy side oral commissure deviations and it found that clinically relevant spontaneous smile events occurred when the oral commissure deviation was more than 80% of normalized maximal displacement and lasted longer than one second. The method was termed “Manual time-stamping” and enabled an objective assessment of *smile synchronicity* which is a measure of how well synchronised the affected side oral commissure is to the healthy side. However, what this approach lacked was

the ability to tell if the resulting smile would have been viewed as joyful in the eye of a naïve observer.

A novel machine learning algorithm, known as Affdex, provided by an Emotion AI company Affectiva, was employed to quantify the success of a smile to demonstrate joy to a naïve observer. The algorithm had been developed from 5 million videos of normal subjects from 75 countries, resulting in over 2 billion facial frames in which Facial Action Coding System action units were expertly recognized and labelled. It was clearly apparent how well the algorithm was able to detect the faces of facial palsy patients despite there being no known cases of facial palsy included in the training set of the algorithm. As the videos of pre-operative flaccid unilateral facial palsy patients were processed it became clear that not only was joy not being expressed but that contempt and disgust were being perceived with high probability. To measure joy and negative emotion expression in spontaneous smiles the next problem was to be able to identify when in a recorded videoclip the patient was smiling spontaneously. This required performing the analysis in real-time, using dual machine learning algorithms. First, Emotrics was used to measure oral commissure displacements via automated facial landmarking. The Emotrics model was trained further specifically to recognise facial palsy patients and became sufficiently reliable to measure oral commissure displacements in real-time and determine during which exact frames Spontaneous Smile Events were occurring during the recording. This method was termed “Automated time-stamping”. Second, selected time-stamped frames were analysed by the Affdex algorithm and probabilities of joy and negative emotion perception were measured during each of the Spontaneous Smile Events, which had been identified by Emotrics from clinically significant healthy side oral commissure deviations.

The significance of this development was immediately apparent. The algorithm was able to differentiate normal smiling from pre-operative flaccidity with 100% probability of joy perception in the normal state and 0% in the flaccid facial palsy state. What was most interesting was, for the first time, to be able to quantify smile spontaneity outcomes following various reconstructive facial reanimation procedures. In a subset of consecutive patients, reviewed retrospectively, there was most improvement in spontaneous joy expression in cross-face nerve graft innervated gracilis smile reanimation (82%, 77 IQR) and this was superior to dual innervation (10%, 98 IQR) and masseteric nerve-driven gracilis (0%, 3 IQR) which exhibited significantly lower probabilities of joy while smiling spontaneously.

Table 2. Perceived Probabilities of Joy and Negative Emotion Expression during Spontaneous Smiling in Normal Subjects and Flaccid Facial Palsy Patients Preoperatively and after Smile Reanimation with Gracilis Functional Free Muscle Transfer Innervated by Either Cross-Face Nerve Graft, Masseteric Nerve, or Dually Innervated*

	% Joy	% Negative Emotion
Normal	100 (2)	0 (5)
Cross-face innervated gracilis	82 (77)	0 (28)
Dually innervated gracilis	10 (98)	24 (56)
Masseteric innervated gracilis	0 (3)	48 (72)
Preoperatively (flaccid)	0 (0)	94 (9)

IQR, interquartile range.

*Data are presented as median (IQR). Raw data of emotionality quotient outcomes are available in Table, **Supplemental Digital Content 4**, <http://links.lww.com/PRS/F68>.

Reproduced from: Dusseldorp JR, Guarin DL, van Veen MM, Miller, M, Jowett N, Hadlock TA. Automated Spontaneity Assessment after Smile Reanimation: A Machine Learning Approach. *Plast. Recon. Surg.* 2022 June; 149(6):1393-1402

Another interesting finding was the reduction in negative emotion expression following smile facial reanimation procedures. This was a feature of both flaccid and non-flaccid (synkinetic) facial palsy and may constitute an equally important outcome of facial reanimation surgery that can now be routinely measured, particularly because increased joy expression does not seem to sufficiently explain all the improvement that can occur in layperson-assessed disfigurement after reanimation.

The ability of the algorithm to process large volumes of data quickly and to identify outliers in a dataset was also useful. For instance, when comparing the results of Affdex to layperson assessments of the same images assessing for disfigurement, there was a moderate correlation between Affdex and layperson-detected disfigurement. Higher probability of joy expression correlated with lower disfigurement scores and higher probability of negative emotion correlated with higher disfigurement scores. This finding suggested future studies formally comparing AI to layperson assessment ought to be considered and raised the possibility of AI becoming a layperson-equivalent in the future avoiding the need for complex logistical arrangements showing confidential post-operative photographs and videos of people's faces before and after smile reanimation procedures to naïve observers.

Limitations of the Thesis

This thesis has several important limitations. The first was the lack of an agreed definition of spontaneous smiling in the field of facial reanimation. From this thesis, the empirical definition of a spontaneous smile event in unilateral facial palsy is a large, sustained, healthy side oral commissure movement triggered while viewing a humorous stimulus. This definition was based on numerous qualitative assessments of normal subjects and facial palsy patients. In these scenarios, manual and automatic time-stamping were similar, but there was

some discordance. Interestingly, when the discordant cases were reviewed, the automated approach was found to be more accurate at determining frames that coincided with clinically relevant healthy side oral commissure movement and often audible laughter. This definition will hopefully help to inform future studies of spontaneity outcomes. Another limitation was the use of the Spontaneous Smile Assay to elicit spontaneous smiles as, once patients had viewed the video, it became familiar, and it is reasonable to expect smiling was less spontaneous on a second viewing. In the future, having access to a range of validated videoclips to avoid familiarity with the stimulus material would be preferable.

There are potential issues with using an artificially intelligent emotion detection algorithm as an outcome measure in facial palsy, particularly when the algorithm has not been trained on facial palsy patients. One criticism is that the algorithm analyses the whole face and thus may be influenced by activity on the healthy side. A counterargument is that laypersons also analyse the whole face, and what is important is if they perceive that the face fits into the realm of “normal”.

Suggestions for Future Research

Further work is required to understand the correlation between emotionality assessment and layperson perceived disfigurement. Studies should compare layperson predicted joy and negative emotion expression to the algorithm to enable a more direct comparison. Similarly future research could try to understand the perception of negative emotion in the flaccid state and its resolution in the reanimated state.

In the current era of evidence-based medicine and value-based healthcare it is critical that patients understand the true likelihood of developing a spontaneous smile after various types of smile reanimation procedures. Previously, reporting of smile spontaneity rates

achieved through various reanimation techniques has been subjective and prone to confirmation bias. By leveraging the power of machine learning, an automated and objective tool for smile spontaneity assessment now exists. This tool should ideally facilitate future long-term, multi-centre outcome studies to delineate the most likely techniques to restore spontaneous, emotional smiling. It is possible that, after further validation efforts, an automated tool employing such an AI algorithm could become a layperson-equivalent obviating the need for cumbersome layperson assessments. Furthermore, this tool could be used to track recovery following facial nerve insult or reanimation procedure. It could also be deployed on a personal device to generate data from a patient's home environment avoiding the need for long distance travel to facial nerve centres for long term follow-up assessments.

Concluding Remarks

This thesis has applied state-of-the-art technology, in particular AI, to outcomes assessment in facial palsy. Furthermore, lessons were learned that are relevant to the field of medicine in general when considering implementation of AI algorithms into clinical practice. Clinicians ought to be aware of the limitations and biases of machine learning algorithms especially regarding the training datasets. Such biases stem from the data that are used to train the algorithm and the precision and accuracy of its training. By increasing the size of the training dataset relative to the overall database the error rate can be reduced but the manual process of training larger and larger datasets can be extremely time-consuming manual work. The training dataset must also represent the full spectrum of the subject matter in question or else the algorithm will be biased to see recognise only those features presented in the training dataset. For instance, if only Caucasian faces are included in a small dataset, a computer vision algorithm will fail to recognise the face of someone with darker skin tones or ethnically diverse facial features.

Regarding the method of training, machine learning models can either be trained in a supervised way (as in this thesis) or in an unsupervised way. There is an inevitable conflict between the use of unsupervised machine learning algorithms, such as neural networks, and evidence-based medicine, given such unsupervised models have not been prescriptively trained using datasets labelled by experts. It is perhaps still possible that we can find uses for such models in low risk, hypothesis generating applications, such as new drug discovery, so long as those utilising the outputs of the model know their limitations and that the findings of an unsupervised AI algorithm cannot be logically understood.

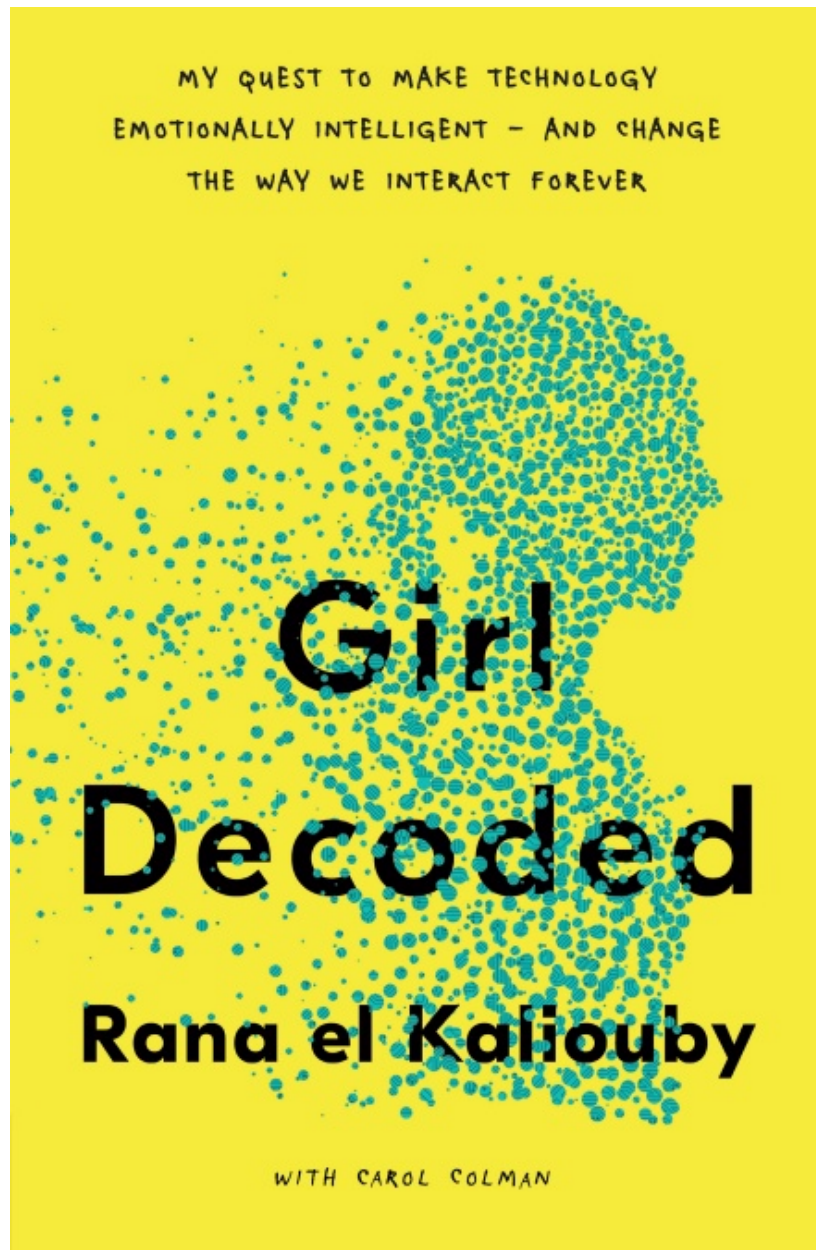
Undertaking this thesis has taught me that the field of computational science has much to offer clinicians and what is lacking is not technological capabilities but rather the willingness to go through the arduous process of clinical translation, validating novel technologies and ensuring they are safely deployed within the various disciplines of medicine. I now firmly believe that we all ought to be looking for ways to incorporate computerisation into our everyday practice to help us to gain valuable insights more quickly, to recognise patterns in common clinical scenarios and to objectively assess outcomes of our interventions, so that we may better serve the needs and desires of our patients.

ADDENDUM

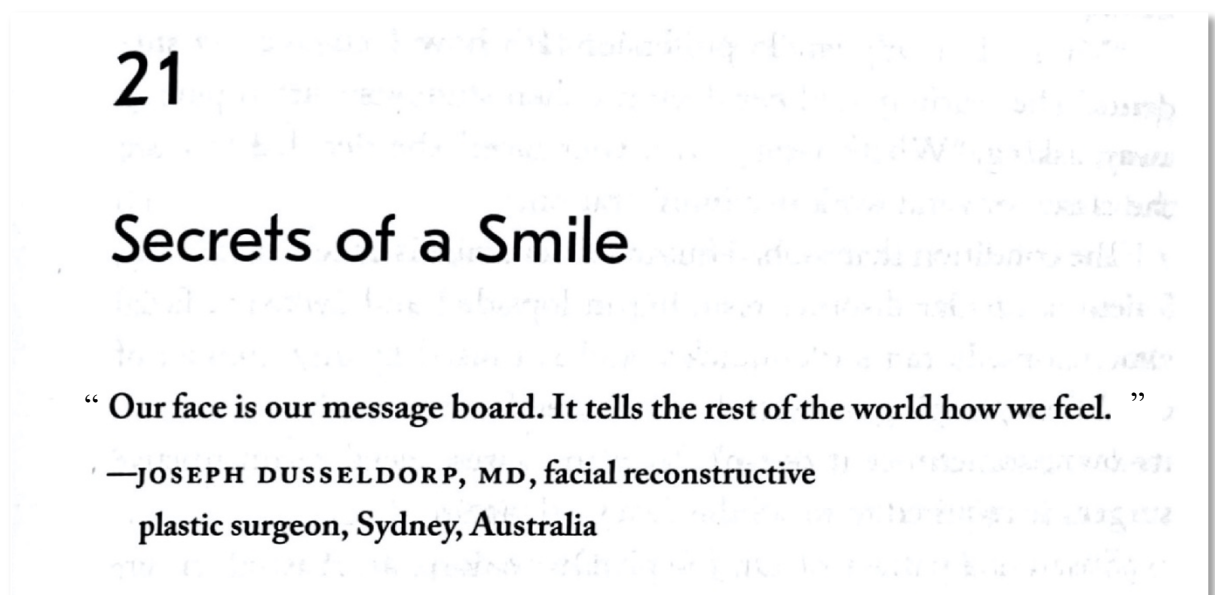
A chance meeting at an Emotion AI Summit in Boston set me on a path to incorporate machine learning into my practice as a reconstructive plastic surgeon. This experience has taught me that all medical fields could benefit from the incorporation of computerisation into our everyday practice.

Dr Rana el Kaliouby was the computer scientist, turned start-up entrepreneur, who gifted us the use of her game-changing technology to test if her AI software could score patients' smile results after facial reanimation surgery. It turns out it could, and the new approach led to insights that had not previously been considered. The algorithm also made it possible to analyse data at scale and make connections to other large datasets in real time opening fascinating opportunities to develop insights and providing the platform for exciting new models of care.

It appears that while the impact of Dr el Kaliouby's technology on me was profound, so too was the impact of our collaboration on her. Dr el Kaliouby reflected on the moment we met and on our collaboration in Chapter 21 of her autobiography "Girl Decoded".¹ In the following pages I have included excerpts from her book chapter titled "Secrets of a Smile" that highlight how using AI to solve a human problem can be very impactful even in the eyes of someone who had never met another person who had experienced facial palsy.



1. Rana el Kaliouby, Carol Colman. Girl Decoded. Penguin Publishing, 2021, ISBN 9781984824783



My smile is my superpower, and so is yours. When I was a foreign PhD student at Cambridge, in England, my smile helped me break down walls, build bridges, and forge strong emotional ties with others.

Smiles are so essential for human interaction that, even before we are born, we practice our smiles in our mother's womb. Yes, smiles are critical for human-to-human communication.

So, imagine losing your ability to smile. That's what happened to Susan (not her real name), an elementary school teacher in Sydney, Australia, who woke up one morning with a, well, *broken* face. While one side of her face still functioned, the other half was frozen. So, her smile was a strange, droopy version of its former self, a distorted

half-smile that didn't express on the outside what she was feeling inside.

"My smile is my whole profession. It's how I engage my students," she tearfully told her doctor. When students started pulling away, asking, "What's wrong with your face?" she decided to leave the classroom and work in administration.

Susan is a patient of Dr. Joseph Dusseldorp, an Australian surgeon and clinical fellow at the Department of Facial Reconstructive Plastic Surgery at Massachusetts Eye and Ear, part of Harvard Medical School. Although the condition interfered with her ability to eat, speak, breathe through her nose, and close her eyes, what she cared about most was getting her smile back.

I met Joe at Affectiva's first Emotion AI Summit, which we organized in 2017, a meeting of experts from around the world focused on the uses of Emotion AI, including in healthcare. (Since then, we've held a summit every year.) Dusseldorp was doing a fellowship at Harvard Medical School and was curious about how Emotion AI technology might apply to his work. The topic quickly turned to smiles.

Dusseldorp told me, whether patients are adults or children, when they seek him out for reconstructive surgery, their reasoning is "I just want my smile back."

to the smile. Even the slightest asymmetry can take you from a perfect smile to a smirk. The goal is to re-create a convincing smile that passes what Dusseldorp calls "the man on the street test." In other words, does it look authentic or would a passerby do a double take and wonder, *Hey, what's wrong with that face?*

"Suppose even though the smile muscle starts to work, and the corner of the mouth is moving, the smile does not tell me that the person is happy," Dusseldorp says. "Maybe it moves too far laterally or too far vertically, but there actually could be all sorts of things about it that are wrong, and the smile actually communicates the wrong message."

I knew from my own research that there are dozens of different types of smiles and that each conveys a different meaning—and not all of them express joy. Even the tiniest variation around the mouth can change the message. But what bothered Dusseldorp most was the lack of an *objective* tool to monitor a repaired smile over time, like those available to other medical specialties. After heart or kidney surgery, a surgeon can use any number of tests to check on a patient's progress. When it comes to facial surgery, though, there is no standardized test to assess function; surgeons have to rely on their own two eyes to track the patient through recovery.

Postsurgical care for smile reanimation involves follow-up visits during which the surgeon typically tries to elicit “spontaneous” smiles from the patient by telling jokes or showing funny videos. Dusseldorp was dissatisfied with the unscientific nature of the postsurgical assessment—the entire process seemed hit or miss—yet he couldn’t figure out an alternative. The answer came to him during a presentation at the Emotion AI Summit, an annual meeting sponsored by Affectiva that includes representatives from a wide range of disciplines to explore potential applications for Emotion AI technology. Graham Page, now head of Affectiva’s Media Analytics Business Unit, but then with Kantar Millward Brown (KMB, formerly Millward Brown), the advertising and branding company that had partnered with Affectiva, described how KMB uses our software to track and measure the facial expressions of consumers watching video commercials. Although on the surface, it had nothing to do with medicine, when Dusseldorp saw Page’s demo, he immediately saw a link between KMB’s work and his patients. *That’s exactly what we reconstructive plastic surgeons are trying to do*, he thought. *Except we don’t have these awesome tools!* He realized that the technology could be used to develop the objective measure of surgical outcomes he had been looking for. So he asked us if he could access the software, and we said yes.

Dusseldorp finally had a tool that not only could measure and rate smiles, but would also eliminate much of the guesswork involved in evaluating postoperative patients.

One of the first things Dusseldorp did with our Emotion AI software was to investigate how the smiles of people with facial palsy were interpreted by others. In one study, he ran videos of the smiles of presurgery patients through our software. He discovered that their smiles were sending others the wrong message. Although the patients were trying to express joy, an asymmetric smile, for most of us,

is associated with contempt. To add to this distortion, people with a form of partially recovered facial palsy, known as synkinesis, may wrinkle their noses involuntarily when they try to smile, which is interpreted as disgust.

After smile reanimation, Affectiva's Emotion AI detected significantly more joy and less negative emotion in the patients' new smiles, a vast improvement. Using our software, Dusseldorp can now rate the type and intensity of a smile postsurgery, just as we do in market research when we test advertisements. The beauty of this approach is that, in the future, when using an app that Dusseldorp is developing with our software developer kit (SDK), postoperative patients will be able to monitor their own smiles in real time, in their own homes, eliminating the need for office visits.

I still find the whole thing rather mind-boggling: Who would have thought that a talk by an advertising executive would plant a novel idea in the mind of a plastic surgeon whose professional work was about helping people smile again? It was precisely the kind of cross-pollination of ideas I had hoped the summit would produce—and why I believe it's so important for people from different disciplines to have an opportunity to meet and share ideas.

I still find the whole thing rather mind-boggling: Who would have thought that a talk by an advertising executive would plant a novel idea in the mind of a plastic surgeon whose professional work was about helping people smile again? It was precisely the kind of cross-pollination of ideas I had hoped the summit would produce—and why I believe it's so important for people from different disciplines to have an opportunity to meet and share ideas.