American Cleft Palate-Craniofacial Association
75th Anniversary

Original Article

The Cleft Palate-Craniofacial Journal
1-10

DOI: 10.1177/1055665617753481
journals.sagepub.com/home/cpc
SAGE

Facial Expressiveness in Infants With and Without Craniofacial Microsomia: Preliminary Findings

Zakia Hammal, PhD[1], Jeffrey F. Cohn, PhD[1,2], Erin R. Wallace, PhD[3], Carrie L. Heike, MD, MS[3,4,5], Craig B. Birgfeld, MD[3,4,5], Harriet Oster, PhD[6], and Matthew L. Speltz, PhD[3,5]

Abstract

Objective: To compare facial expressiveness (FE) of infants with and without craniofacial macrosomia (cases and controls, respectively) and to compare phenotypic variation among cases in relation to FE.

Design: Positive and negative affect was elicited in response to standardized emotion inductions, video recorded, and manually coded from video using the Facial Action Coding System for Infants and Young Children.

Setting: Five craniofacial centers: Children's Hospital of Los Angeles, Children's Hospital of Philadelphia, Seattle Children's Hospital, University of Illinois–Chicago, and University of North Carolina–Chapel Hill.

Participants: Eighty ethnically diverse 12- to 14-month-old infants.

Main Outcome Measures: FE was measured on a frame-by-frame basis as the sum of 9 observed facial action units (AUs) representative of positive and negative affect.

Results: FE differed between conditions intended to elicit positive and negative affect (95% confidence interval = 0.09-0.66, $P = .01$). FE failed to differ between cases and controls (ES = –0.16 to –0.02, $P = .47$ to .92). Among cases, those with and without mandibular hypoplasia showed similar levels of FE (ES = –0.38 to 0.54, $P = .10$ to .66).

Conclusions: FE varied between positive and negative affect, and cases and controls responded similarly. Null findings for case/control differences may be attributable to a lower than anticipated prevalence of nerve palsy among cases, the selection of AUs, or the use of manual coding. In future research, we will reexamine group differences using an automated, computer vision approach that can cover a broader range of facial movements and their dynamics.

Keywords

craniofacial microsomia, facial expressiveness, AUs, infants

Introduction

Craniofacial microsomia (CFM) is a complex congenital condition, typically involving underdevelopment of the mandible and ear. It occurs in approximately 1 in 3500 to 5600 live births (Poswillo, 1988), with higher than expected prevalence among individuals of Hispanic and Native American ancestry (Harris et al., 1996). CFM has been characterized as a spectrum of phenotypic presentations ranging from isolated unilateral microtia to bilateral malformations of the ear, mandible, and facial soft tissue and orbit (Cole et al., 2004); other cranial and extracranial malformations may co-occur (eg, lateral oral clefts and vertebral anomalies). CFM has several functional consequences and correlates that often require treatment, including upper airway obstruction, feeding difficulties, speech and hearing impairment, developmental delays, and facial palsy. The latter area is the focus

[1] Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, USA
[2] Department of Psychology, University of Pittsburgh, Pittsburgh, PA, USA
[3] Seattle Children's Research Institute, Seattle, WA, USA
[4] Seattle Children's Hospital, Seattle, WA, USA
[5] University of Washington School of Medicine, Seattle, WA, USA
[6] NYU School of Professional Studies, New York, NY, USA

Corresponding Author:
Zakia Hammal, PhD, Robotics Institute, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA 15213, USA.
Email: zakia_hammal@yahoo.fr

of this study. Estimates of the incidence of cranial nerve involvement in CFM have varied widely among clinical samples, from approximately 10% to 60% (Barisic et al., 2014; Cline et al., 2014; Cohen et al., 2017), depending on how patients are ascertained and investigators' exact definition of the CFM spectrum.

Although CFM is a congenital condition, little is known about the prevalence and clinical impact of facial palsy in infants and preschool children with CFM. This is due in large part to the challenge of assessing facial nerve function in children of this age. Whereas older children are capable of imitating specific facial expressions that can reveal the effects of different branches of the facial nerve, assessment of infants and younger children largely relies on observations of spontaneous facial expressions in the clinical setting, ideally during moments of heightened affect (eg, laughing or crying). This approach to assessment may underidentify nerve palsy in young children with CFM, delay the planning of relevant treatments (eg, reanimation surgery, feeding/eating, and speech interventions), and hamper our ability to investigate the potential developmental impact of facial palsy.

In the current study, we explored the use of an alternative assessment method using standardized emotion induction tasks to elicit positive and negative affective displays in 12- to 14-month-old infants. The use of emotion induction is well established in other areas of infancy research, primarily studies of early temperament and emotion regulation (Tronick, 1989; Segal et al., 1995; Campos et al., 2004; Cole et al., 2004; Oster, 2005). Our working assumption was that the induction of positive and negative mood states in infants would produce corresponding facial expressions (eg, happy vs frustrated/angry) that would allow for the reliable coding of facial movements relevant to the diagnosis of facial nerve dysfunction. It was also of interest to determine whether the facial expressiveness of children with CFM is discrepant from that of typical children, as such differences might account for previous findings of elevated rates of internalizing behaviors (eg, social inhibition and socialization problems) found in some studies of children and adolescents with CFM (Pertschuk and Whitaker, 1985; Pillemer and Cook, 1989; Padwa et al., 1991; Maris et al., 1999; Snyder et al., 2005).

We used the well-studied Facial Action Coding System for Infants and Young Children (Baby FACS; Oster, 2003) to observe the video-recorded facial expressions of infants with CFM ("cases") and demographically similar infants without a craniofacial anomaly ("controls"). The study addressed 2 primary questions: (1) During emotion induction, do cases and controls evidence discernible differences in facial expressiveness that potentially reveal facial nerve impairment? (2) Among cases, is facial expressiveness related to phenotypic differences in facial structures (eg, microtia only vs microtia and mandibular hypoplasia)? Although infants with mandibular hypoplasia might be expected to more likely demonstrate limitations in facial movement indicative of facial nerve palsy, previous studies of older children and adults with CFM have not observed consistent associations with different patterns of facial malformations (Cline et al., 2014). Secondary analyses involving all participants examined the potential moderating influence of infant sex, ethnicity, and type of emotion induction (positive vs negative affect) on observed facial movements. Ethnicity was of particular interest given the higher than expected prevalence of CFM among individuals of Hispanic and Native American descent. Finally, in addition to overall expressiveness, we examined group differences by specific regions of the face (eg, eyebrow vs lip movements).

Methods

Participants

This study is part of an observational, longitudinal, multicenter project called Craniofacial Microsomia: Longitudinal Outcomes in Children Pre-Kindergarten (CLOCK), which is tracking the neurodevelopmental, speech, and hearing outcomes and phenotypic features of infants and young children with and without CFM ("cases" and "controls," respectively). Participants have been enrolled since 2012 from one of 5 craniofacial centers: Children's Hospital of Los Angeles, Children's Hospital of Philadelphia, Seattle Children's Hospital, University of Illinois, Chicago, and the University of North Carolina–Chapel Hill. Enrollment is still under way and will continue until approximately 195 participants are enrolled (110 cases and 85 frequency-matched controls).

Participants are assessed at 12, 24, and 36 months of age on a variety of developmental and observational measures. Here we report on findings for the first 80 participants (44 cases and 36 controls) who completed the 12-month assessment and whose facial responses to emotion induction were coded both manually using Baby FACS and automatically using a computer vision-based approach called automated face analysis (AFA; Cohn and De la Torre, 2015). The current paper reports on the results of the manual coding for these 80 participants (all other participants in this research will only be coded with AFA, which will be the focus of a future article when the full sample is ascertained and coded).

This research was approved by the institutional review boards at all participating centers. All parents gave informed consent for their infant to participate in the study. Informed consent was obtained for all images that appear below.

Cases. Cases were recruited from each site's hospital-based craniofacial centers, hospital-based centers seeing infants or young children with CFM (eg, hearing screening programs, ENT programs), and research study websites (eg, clinical trials.gov). To be eligible, cases had to (1) have at least 1 of the CFM inclusion criteria developed by the Facial Asymmetry Collaborative for Interdisciplinary Analysis and Learning (FACIAL) network (see Table 1); (2) be diagnosed by a regional craniofacial team; (3) be between the ages of 12 and 24 months (or corrected age, born between 34 and 36 weeks' gestation); and (4) have a legal guardian who is able to provide informed written consent, be willing to comply with all study procedures, and be available for the duration of the study. Exclusion criteria for cases included (1) diagnosis of a known syndrome (eg, Townes-Brocks, Treacher Collins, branchio-

Table 1. Cases Inclusion Criteria.

One or more of the following:
1. Microtia
2. Anotia
3. Facial asymmetry *and* preauricular tag(s)
4. Facial asymmetry *and* facial tag(s)
5. Facial asymmetry *and* epibulbar dermoid
6. Facial asymmetry *and* macrostomia
7. Preauricular tag *and* epibulbar dermoid
8. Preauricular tag *and* macrostomia
9. Facial tag *and* epibulbar dermoid
10. Macrostomia *and* epibulbar dermoid

oto-renal or Nager syndromes); (2) presence of an abnormal karyotype or major medical or neurologic conditions (eg, cancer and cerebral palsy); (3) premature birth (less than 34 weeks' gestation); (4) any circumstance that would preclude the family's ability to participate fully in the research; (5) a sibling already participating in the CLOCK study; and (6) infant's consenting parent unable to speak English or Spanish.

Control group participants. We identified eligible participants with demographic characteristics that met our frequency-matching criteria for the case cohort with respect to infant age and sex, family socioeconomic status (SES), and language spoken in the home (English or Spanish). Exclusion criteria for controls included (1) meeting 1 or more of the exclusionary criteria for cases and (2) diagnosis or history of any disorder, condition, or injury that would affect facial features (eg, craniofacial malformation or deformation; facial surgery or trauma).

Emotion Induction

Infants' facial expressiveness was observed in response to 2 standardized emotion inductions, one intended to elicit positive affect (eg, smile, surprise, interest, and amusement) and the other, negative affect (eg, frustration, anger, and distress). For each task, infants were seated in a highchair in front of a table with an experimenter and their mother seated on the other side of the table. The experimenter sat to the mother's left, out of camera view and closer to the table. In the *positive emotion task* (PosET), soap bubbles were blown toward the child and the center of the table, just below camera view. In the *negative emotion task* (NegET), the examiner first demonstrated and allowed the infant to play with an attractive toy car, followed by the toy's removal and placement for 30 seconds under a clear plastic bin just out of the infant's reach; this procedure followed a protocol described by Goldsmith and Rothbart (1999). These procedures (ie, blowing bubbles or toy removal) were repeated 1 to 3 times, depending on the infant's response. The NegET was terminated if the infant became too upset or mother became uncomfortable with the procedure. Both tasks were recorded using a Sony DXC190 compact camera at 60 frames per second (seeFigure 1A). Participants' face orientation to the cameras was approximately 15° from frontal.

Measures

Observational measures. We used the manual Baby FACS (Oster, 2003) to code emotion-related facial action units (AUs). AUs correspond to discrete, minimally distinguishable actions of the facial muscles (Ekman et al., 2002). For example, among independent muscle actions in the brow region, there are codes for elevation of the inner and outer corners of the brows (AUs 1 and 2, respectively; see Table 2) and narrowing or "knitting" of the inner brow corners (AU 3; see Table 2). Because action units are exhaustive and mutually exclusive, nearly all possible facial expressions can be precisely and unambiguously identified in relation to combinations and sequences of its constituent actions. Baby FACS validity has been demonstrated by its cross-cultural invariance and ability to distinguish responses to different emotion elicitors (Rosenstein and Oster; 1988; Camras et al., 2003; Oster, 2003, 2005, Bolzani Dinehart et al., 2005; Mattson et al., 2013). Importantly, Oster et al. have shown that the facial expressions of infants with craniofacial anomalies can be reliably coded with Baby FACS (Oster, 2003).

Three Baby FACS–certified coders (blind to case/control status) manually coded 9 preselected AUs on a frame-by-frame basis for both tasks (see Table 2). The goal was to sample a range of AUs from across the upper, middle, and lower face that are central to the communication of positive and negative affect. The selection of the specific 9 AUs (see Table 2) was informed by prior research (Matias and Cohn, 1993; Camras et al., 2003; Oster, 2003, 2005; Messinger et al., 2012). Smiles are indexed by AU 12 (lip corner puller) and cry faces by AU 20 (lip stretcher). AU 6 (cheek raiser) differentiates felt smiles from social smiles and is an intensifier of positive and negative affect. AU 1+2 (brow raiser) is a key component of surprise. AU 3 and AU 4 figure in interest, concentration, and also negative affect. AU 9 (nose wrinkler) signals disgust and distress. AU 28 (lip suck) was selected as one of several candidate lip movements that are common in infants.

Coders continuously coded on a frame-by-frame basis the first 45 seconds of the PosET. For the NegET, coders continuously coded on a frame-by-frame basis the first 15 seconds following each toy removal (45 seconds total). For each frame, AUs were coded for presence/absence by one of 3 coders (Ekman et al., 2002; Oster, 2003). To assess intercoder agreement, 2 or more of the coders independently coded on a frame-by-frame basis 15 seconds of randomly selected segments from the PosET and NegET tasks for 68 infants (30 cases and 38 controls). Agreement between coders was quantified using free-margin kappa (Brennan and Prediger, 1981), which corrects for chance agreement. Intercoder agreement was good to high for all AUs for both cases and controls (see Table 2).

Using the results of the Baby FACS coding, facial expressiveness was operationalized as the continuous sum of all manually observed AUs on a frame-by-frame basis during the 45 seconds coded for each.

Phenotypic classification. We classified the participant's phenotype based on the integration of standardized ratings of facial

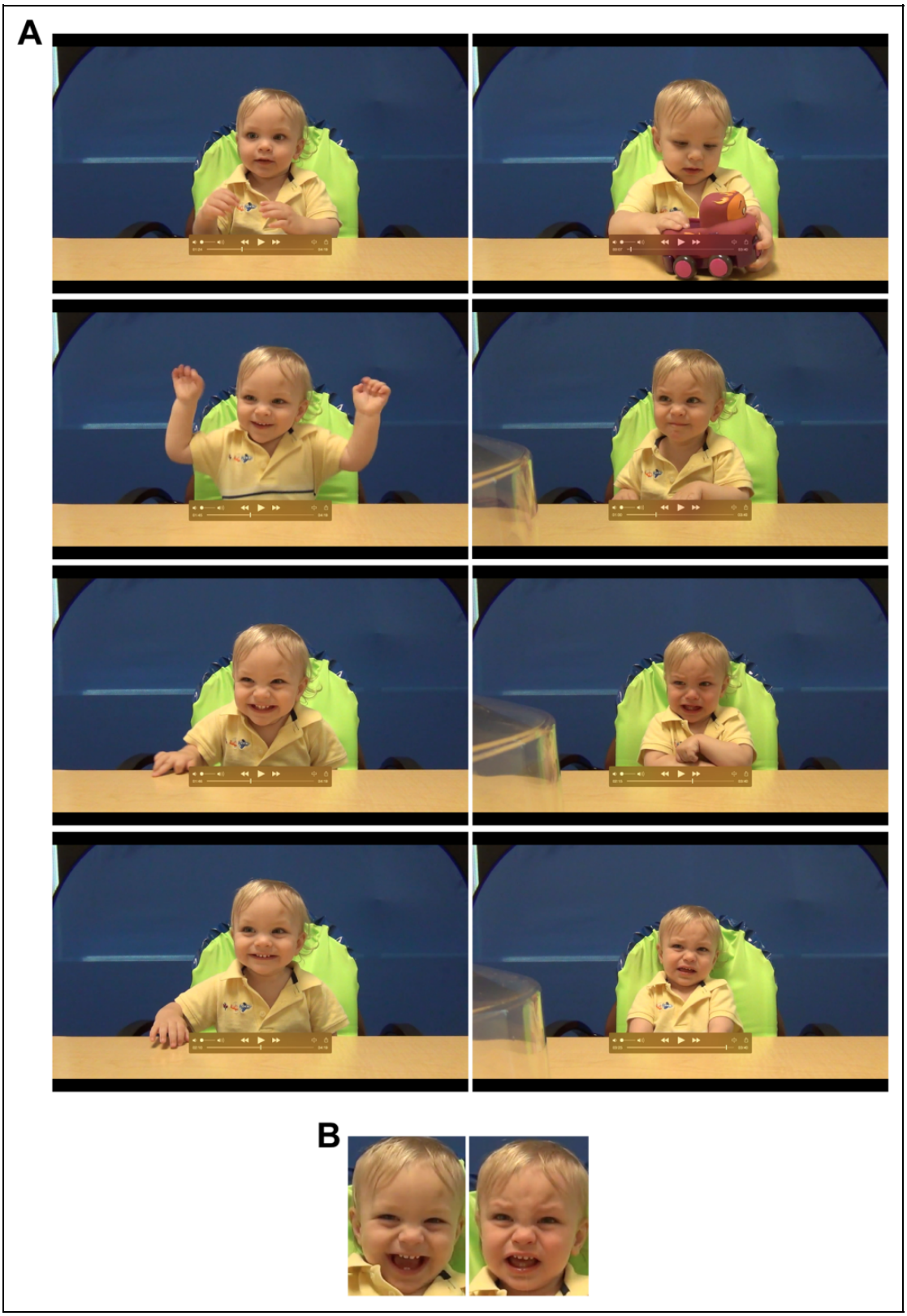
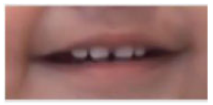
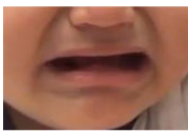
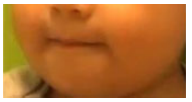


Figure. 1. (A) Examples of PosET (right) and NegET (left), (B) example of combination of AUs. Right AUs 6+12 (smile), left AUs 4+20 (cry).

features based on photographs and data taken from a medical history interview and medical charts (Heike et al., 2016). The standardized photographic protocol consisted of 4 views (frontal views with neutral and smiling expressions, right and left lateral views) of the face (adapted from Heike et al., 2011) and a classification method described by Birgfeld et al. (2016),

Table 2. Intercoder Agreement for Baby FACS Action Units.

Action Unit (AU)		Free-Margin Kappa	
		Cases (n = 30)	Controls (n = 38)
1 and 2	Inner corner of eyebrow raised	0.62	0.65
	Outer corner of eyebrow raised	0.71	0.73
3	Inner corners of the brows drawn together	0.72	0.74
			
4	Inner brows lowered	0.89	0.87
			
6	Cheeks raised	0.78	0.82
			
9	Upper lip raised, superior part of nasolabial furrow deepened, nose wrinkled	0.92	0.90
			
12	Lip corners pulled up and orthogonally	0.86	0.85
			
20	Lip corners pulled laterally	0.77	0.72
			
28	Lips sucked	0.93	0.86
			
	$\bar{x}$	0.80	0.79

which used a modified version of the Orbital, Ear, Mandible, Nerve, Soft tissue (OMENS) pictorial rating scale (Cousley, 1993; Horgan et al., 1995; Gougoutas et al., 2007). In previous research, ratings by physicians of photos using this method correlated highly with physical examination for most features and demonstrated high inter-rater reliability (kappa coefficients > 0.7 for each of the OMENS features) (Heike et al., 2016).

In the current study, 2 of the investigators, a craniofacial pediatrician (C.H.) and a geneticist, were blind to case/control status and rated all photographs. All discrepancies were reviewed by the raters to develop consensus. For each feature on each study participant, data from all 3 sources (ie, consensus OMENS ratings based on photographs, medical history interview, and medical chart abstraction) were reviewed to establish the phenotype. Using the phenotype data, 3 subgroups among the case cohort were identified: (1) microtia *only* (in the absence of other CFM-related features such as mandibular hypoplasia and epibulbar dermoids; n = 12); (2) microtia and mandibular hypoplasia; n = 27), and (3) other combinations of CFM-associated malformations (2 or more were required; n = 6). In the latter subgroup, nearly all (5 of 6 cases) had microtia (in the absence of mandibular hypoplasia) plus preauricular or facial tags; 1 additional case had these features plus an epibulbar dermoid.

Statistical Analyses

Dependent variables. Facial expressiveness was the primary outcome and was operationalized using Baby FACS coding results during both the PosET and NegET. Facial expressiveness was calculated as the total number of manually coded AUs. So that any minor differences in the amount of coded video for individual participants would not influence the measures, expressiveness was normalized by the duration of the coded segments.

Analyses. To confirm that the induction tasks had elicited the desired emotional states, we compared the proportion of positive affect (AUs 6+12, smiles, Figure 1B) and negative affect (AUs 4+20, cry-face, Figure 1B) shown by infants in the PosET and NegET, respectively.

General estimating equations (GEEs) were used to examine differences in facial expressiveness between PosET and NegET and corresponding 95% confidence intervals (CIs), using an independent correlation matrix. Linear regression with robust standard error estimates was used to examine differences in AUs and facial expressiveness between cases and controls, as well as differences across phenotype, with controls serving as the referent category. Corresponding 95% CIs were calculated using robust standard error estimates. Wald tests were used to evaluate evidence for phenotypic group differences. To facilitate the interpretation of coefficients from the linear regression models, we estimated standardized effect sizes (ESs) using a modification of Cohen *d*, calculated as the estimated mean difference divided by the root mean square error for the model (Cohen, 1988). In secondary analyses, we used linear regression with robust standard error estimates to examine differences in facial expressiveness by sex (males vs females) and ethnicity (Hispanic/Latino vs non-Hispanic or Latino) and corresponding 95% CIs.

Because of the exploratory nature of this research, *P* values were not adjusted for multiple comparisons and they did not serve as the sole basis for estimating the strength of findings. Instead, we assessed the magnitude of observed effect sizes, their precision, and the consistency of these estimates across multiple measures and the 2 emotion induction tasks (Rothman, 2014).

Table 3. Baseline Characteristics of Children With and Without CFM.

	Cases		Controls	
Characteristic	n	(%)	n	(%)
Total	44	(100.0)	36	(100.0)
Sex				
Male	26	(59.1)	19	(52.8)
Female	18	(40.9)	17	(47.2)
Age, mo				
Mean (SD)	13.1	(0.6)	12.9	(0.5)
<13	21	(47.7)	22	(61.1)
13-14	22	(50.0)	14	(38.9)
>14	1	(2.3)	0	(0.0)
SES				
Mean (SD)	34.7	(12.7)	39.2	(14.9)
I	5	(11.4)	7	(19.4)
II	7	(15.9)	8	(22.2)
III	17	(38.6)	9	(25.0)
IV	10	(22.7)	8	(22.2)
V	5	(11.4)	3	(8.3)
Hispanic				
No	19	(43.2)	27	(75.0)
Yes	25	(56.8)	8	(22.2)
Race				
White	33	(75.0)	27	(75.0)
Black/African American	1	(2.3)	1	(2.8)
Asian	5	(11.4)	0	(0.0)
American Indian/Alaska Native	1	(2.3)	0	(0.0)
Native Hawaiian/ Other PI	0	(0.0)	0	(0.0)
Other race	0	(0.0)	0	(0.0)
Multiracial	3	(6.8)	7	(19.4)
Testing language (based on PDP)				
100% English	30	(68.2)	32	(88.9)
100% Spanish	6	(13.6)	0	(0.0)
Combined English and Spanish	8	(18.2)	4	(11.1)
Phenotype				
Microtia only	11	(25.0)	0	(0)
Microtia + Mandibular hypoplasia	27	(61.4)	0	(0)
Other anomaly(s)	6	(13.6)	0	(0.0)
No discernible anomaly	0	(0.0)	36	(100.0)
Recruitment site				
CHLA	16	(36.4)	3	(8.3)
CHOP	0	(0.0)	0	(0.0)
SCH	18	(40.9)	30	(83.3)
UNC	8	(18.2)	2	(5.6)
UIC	2	(4.5)	1	(2.8)

Abbreviations: CHLA, Children's Hospital of Los Angeles; CHOP, Children's Hospital of Philadelphia; PI, Pacific Islander; SCH, Seattle Children's Hospital; SES, socioeconomic status; UIC, University of Illinois, Chicago; UNC, University of North Carolina–Chapel Hill.

Results

The demographic characteristics of cases and controls are shown in Table 3. Mean age at the time of the assessment was 13.0 months (SD = 0.6). Compared to controls, cases were more likely to be male and of Hispanic ethnicity. Among cases, the most common phenotypic presentation was microtia plus mandibular hypoplasia (61%), followed by microtia alone (25%).

Differences Between Emotion Induction Tasks

As expected, infants were more positive in the PosET and more negative in the NegET. The ratio of smiles (AUs 6+12; see Figure 1B, right) to cries (AUs 4+20; see Figure 1B, left) was higher in the PosET compared to the NegET ($t = 4.54$, $df = 79$, $P < .01$). Similarly, the ratio of cries (AUs 4+20) to smiles (AUs 6+12) was higher in the NegET than during the PosET ($t = 4.54$, $df = 79$, $P < .01$).

Differences in Facial Expressiveness by Emotion Task

Facial expressiveness using the sum of AUs detected at the frame-by-frame basis was higher in the NegET than during the PosET (Table 4). The mean number of AUs per frame during the NegET was an estimated 0.37 points higher (95% CI 0.09-0.66, $P = .01$) than during the PosET.

Case-Control Differences

There was little evidence for differences in individual AUs between cases and controls (Table 5). Standardized differences ranged from –0.19 to 0.39 for the PosET (P values ranged from .07 to .94) and between –0.02 and 0.25 for the NegET (P values ranged from .24 to .91). Likewise, there was little evidence for group differences in total facial expressiveness scores (sum of AUs) for either the PosET or NegET (ES = –0.16 to –0.02; P values ranged from .47 to .92) (Table 4).

Analyses by Phenotype

Cases with microtia and mandibular hypoplasia and other CFM-associated features had lower levels of facial expressiveness than controls, as measured by AUs, but the magnitude of the differences was generally small, they were imprecise, and all estimates included the null (ES = –0.38 to –0.16; P values ranged from .33 to .66). Estimates for cases with microtia only relative to controls ranged from 0.10 to 0.54 (P values ranged from .25 to .81).

Secondary Analyses

Males and females did not differ in facial expressiveness (ES = –0.18 to 0.13; P values ranged from .41 to .57) (Table 4). There was only scant evidence for differences in facial expressiveness by ethnicity (ES = –0.08 to 0.37; P values ranged from .14 to .71) (Table 4).

Discussion

There are several advantages to the early identification of nerve palsy in infants and young children with CFM. These include the possibility of assessing the potential clinical impact of facial palsy on developing toddlers and the benefits and feasibility of reanimation surgery, which can be performed in the preschool years (Petersson et al., 2014). Information about facial nerve functioning could be used to develop interventions

Table 4. Estimated Mean Difference in Expressiveness by Task, Case Status, Sex, and Ethnicity.

	PosET		NegET		NegET vs PosET Task			
Measure	Mean	SD	Mean	SD	Mean Difference	95% CI		*P* Value
Expressiveness	1.24	0.82	1.61	1.02	0.37	0.09, 0.66		0.01
Case Status								
	Controls		Cases		Cases vs Controls			
Task	Mean	SD	Mean	SD	Mean Difference	95% CI	ES	*P* Value
Expressiveness								
PosET	1.25	0.75	1.24	0.89	–0.02	–0.38, 0.34	–0.02	.92
NegET	1.7	0.94	1.54	1.08	–0.16	–0.61, 0.28	–0.16	.47
Sex of Child								
	Male		Female		Females vs Males			
Task	Mean	SD	Mean	SD	Mean Difference	95% CI	ES	*P* Value
Expressiveness								
PosET	1.2	0.74	1.31	0.92	0.11	–0.27, 0.48	0.13	.57
NegET	1.69	1.15	1.51	0.82	–0.18	–0.61, 0.25	–0.18	.41
Ethnicity								
	Non-Hispanic		Hispanic/ Latino		Hispanic/Latino vs Non-Hispanic			
Task	Mean	SD	Mean	SD	Mean Difference	95% CI	ES	*P* Value
Expressiveness								
PosET	1.15	0.68	1.45	0.98	0.3	–0.09, 0.7	0.37	0.14
NegET	1.64	1.13	1.56	0.86	–0.09	–0.53, 0.36	–0.08	0.71

Table 5. Mean Differences in Individual AUs Between Cases and Controls.

	Task							
	PosET				NegET			
AU	Mean Difference	95% CI	ES	*P* Value	Mean Difference	95% CI	ES	*P* Value
1	0	–0.11, 0.1	–0.02	0.94	–0.04	–0.16, 0.07	-0.16	0.46
2	0.02	–0.08, 0.12	0.1	0.65	–0.01	–0.12, 0.1	–0.02	0.91
3	0	–0.11, 0.12	0.02	0.94	–0.01	–0.15, 0.13	–0.03	0.88
4	0.02	–0.06, 0.1	0.1	0.62	0.04	–0.09, 0.17	0.14	0.52
6	0.07	–0.05, 0.2	0.24	0.27	0.01	–0.11, 0.14	0.04	0.84
9	0.06	–0.02, 0.13	0.29	0.16	0.04	–0.05, 0.13	0.17	0.41
12	0.06	–0.07, 0.2	0.2	0.37	0.04	–0.05, 0.13	0.18	0.4
20	0.09	–0.01, 0.19	0.39	0.07	0.07	–0.04, 0.18	0.25	0.24
28	–0.04	–0.13, 0.05	–0.19	0.38	0.01	–0.08, 0.11	0.07	0.75

Abbreviations: AUs, action units; ES, effect size.

for feeding, speech, and nonverbal communication and provide anticipatory guidance for parents who struggle to "read" the facial expressions and related affective communications of infants or toddlers with limited facial movement. In an effort to facilitate earlier identification of facial nerve function in individuals with CFM, we explored the use of standardized emotion induction procedures—commonly used in other areas of infancy research—to elicit affectively charged facial expressions in infants with CFM and demographically similar infants without craniofacial anomalies. The primary aims were to examine case-control group differences in manually coded, anatomically based facial movements, and among cases, to determine whether facial movement would vary across phenotypic subgroups. To our knowledge, this is the first study of CFM to code children's faces in real time from video recordings, rather than using ratings of static images.

We observed little difference in facial expressiveness between cases and controls. Facial expressiveness was similar in both groups across multiple indicators of expression, including total expressiveness scores for positive and negative emotion tasks and in specific regions of the face as measured by individual AUs. Nor did we find statistically meaningful differences in facial expression among subgroups of cases distinguished by facial phenotype (eg, microtia with and without mandibular hypoplasia).

Several factors may have accounted for these null findings. First, there are some components of facial movement important to the assessment of nerve palsy that are difficult to elicit with typical emotion induction procedures (eg, lower lip suppression, which usually requires baring of the lower teeth). In this study, we targeted movements related to basic emotions (eg, the cheek raising and lip-corner pull observed in displays of positive affect).

Second, this study relied upon manually coded observations of facial movement, and small and subtle, but important, movements may have been missed by the coders. As noted earlier, we will be conducting computer vision-based AFA (Cohn and De la Torre, 2015) for all participants in the full sample. AFA may be more sensitive to subtle movements and better able to capture the dynamics of facial movement. Using AFA, recent findings indicate that the dynamics of head and facial movement reliably measure the automatic assessment of depression severity in adults (Dibeklioglu et al., 2015, 2017). AFA measures of head and facial dynamics in children with CFM may prove to be an important indicator of facial nerve function.

A third factor accounting for our findings may be the way in which infants' behavior was time sampled. Each of the emotion induction tasks lasted about 5 minutes, and we coded the first 45 seconds of each task. There is evidence that some children with CFM are more socially inhibited than typical children (Pillemer and Cook, 1989; Padwa et al., 1991; Dufton et al., 2011), and it is possible that cases were slower to warm up than the control group participants. In future studies, we plan to investigate this by sampling behavior throughout the entire duration of emotion induction (eg, first, middle, and final 45 seconds), something that is far more feasible with AFA than manual coding.

Finally, our null findings may be the result of a lower-than-expected percentage of children with nerve palsy in our sample of cases. We do not have an optimal independent measure or clinical diagnosis of nerve palsy because, as already noted, our study relied on phenotypic classification based on photographic images, parent interview, and medical chart abstraction to document the presence of nerve palsy. In future research, we plan to assess our cases when they are old enough to imitate targeted facial expressions (in response to images and/or examiner modeling) that can be photographed or video recorded and rated for extent of nerve palsy, using the OMENS pictorial rating system or similar approach (PAT-CFM). Doing so would allow us to re-examine differences in response to emotion induction between control group infants and a subgroup of infants with CFM who were subsequently diagnosed as having nerve palsy.

We observed modest differences in facial expression between Hispanic and non-Hispanic children, with slightly greater facial expression observed in Hispanic infants for the positive emotion task. Given the small magnitude of this difference and the exploratory nature of the question, these results can only be considered as hypothesis generating. However, they merit further investigation, as differences in facial expression have been observed across other racial and ethnic groups (Camras et al., 1998), although, to our knowledge, Hispanic samples have not been included in that research. This may be an important area of study given the elevated rate of CFM in Hispanic/Latino infants, and ethnicity is therefore potentially confounded with case status. Follow-up studies in larger populations of Hispanic infants with CFM are needed.

Among the study's limitations, two are notable. First, as already mentioned, the true prevalence of facial nerve impairment in our sample is yet unknown. Second, this was a preliminary study of a subsample of a larger cohort that is still being recruited (the subsample is distinguished from the remainder of the cohort in that it was coded both manually and with AFA). The relatively small size of the subsample limited our ability to adjust for differences in the case-control groups, primarily in ethnicity, and among cases there was reduced statistical power for the analyses of facial expression by phenotype. These issues will be resolved when the full sample is recruited.

Conclusion

In summary, specific facial expressions and expressiveness strongly differed between conditions intended to elicit positive and negative affect. However, we observed little difference in facial expressiveness between cases and controls, and among cases, between those with and without mandibular hypoplasia. These null findings may be attributable to several factors, operating separately or in combination, including the difficulty of eliciting with emotion induction the entire range of facial movements relevant to the identification of nerve palsy; limited sampling of infants' optimal affective displays, which may have underidentified cases' limitations in facial movement due to social inhibition; and/or a lower than anticipated prevalence of nerve palsy in our sample of cases. The latter possibility can only be confirmed with an assessment of our case sample at an older age. Finally, human coders may be limited in the extent to which they can detect the often subtle, brief indicators of nerve palsy. As a next step, we plan to investigate this possibility by using AFA, which can cover a broader range of facial movements and, because of the efficiency of machine learning, sample longer sequences of behavior.

Despite these preliminary, null findings, we remain enthusiastic about the use of standardized observational procedures for early detection of nerve palsy in infants and young children with CFM, including emotion induction. Such procedures can potentially lead to earlier identification of facial nerve dysfunction, inform the development of early interventions for infants and parents, and serve as objective measures of pre-post functioning in preschool age children who undergo reanimation surgery.

Acknowledgments

We would like to thank our colleagues at participating craniofacial centers who helped recruit and assess our families, including Drs Kathy Kapp-Simon (University of Illinois–Chicago), Amelia Drake (University of North Carolina–Chapel Hill), Alexis Johns (Children's Hospital of Los Angeles), and Leanne Magee (Children's Hospital of Philadelphia). We would also like to thank the families who so generously volunteered their time to participate in this research.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported by the Center for Clinical and Translational Research at Seattle Children's Research Institute (grant UL1 TR000423) and the National Institute of Dental and Craniofacial Research (grant R01 DE 022438).

ORCID iD

Carrie L. Heike, MD http://orcid.org/0000-0003-2178-7730

References

Barisic I, Odak L, Loane M, Garne E, Wellesley D, Calzolari E, Dolk H, Addor MC, Arriola L, Bergman J, et al. Prevalence, prenatal diagnosis and clinical features of oculo-auriculo-vertebral spectrum: a registry-based study in Europe. *Eur J Hum Genet*. 2014; 22:1026-1033.

Birgfeld CB, Heike CL, Saltzman BS, Leroux BG, Evans KN, Luquetti DV. Reliable classification of facial phenotypic variation in craniofacial microsomia: a comparison of physical exam and photographs. *Head Face Med*. 2016;12:14.

Bolzani-Dinehart L, Messinger DS, Acosta S, Cassel T, Ambadar Z, Cohn J. Adult perceptions of positive and negative infant emotional expressions. *Infancy*. 2005;8:279-303.

Brennan RL, Prediger DJ. Coefficient kappa: some issues, misuses, and alternatives. *Educ Psychol Meas*. 1981;41:687-699.

Camras LA, Oster H, Campos JJ, Bakeman R. Emotional facial expressions in European-American, Japanese, and Chinese children. *Ann N Y Acad Sci*. 2003;1000:135-151.

Camras LA, Oster H, Campos J, Campos R, Ujiie T, Miyake K, Wang L, Meng Z. Production of emotional facial expressions in European American, Japanese, and Chinese infants. *Dev Psychol*. 1998;34: 616-628.

Cline JM, Hicks KE, Patel KG. Characterization of facial paresis in hemifacial microsomia. *Otolaryngol Head Neck Surg*. 2014;150: 188-193.

Campos JJ, Frankel CB, Camras L. On the nature of emotion regulation. *Child Dev*. 2004;75:377-394.

Cohn JF, De la Torre F. Automated face analysis for affective computing. In: Calvo RA, D'Mello SK, Gratch J, Kappas A, eds. *Handbook of Affective Computin*g. New York, NY: Oxford University Press; 2015:131-150.

Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum; 1988:400.

Cohen N, Cohen E, Gaiero A, Zecca S, Fichera G, Baldi F, Giordanetto JF, Mercier JM, Cohen A. Maxillofacial features and systemic malformations in expanded spectrum Hemifacial Microsomia. *Am J Med Genet A*. 2017;173:1208-1218.

Cole PM, Martin SE, Dennis TA. Emotion regulation as a scientific construct: methodological challenges and directions for child development research. *Child Dev*. 2004:75:317-333.

Cousley RR. A comparison of two classification systems for hemifacial microsomia. *Br J Oral Maxillofac Surg*. 1993;31:78-82.

Dibeklioglu H, Hammal Z, Cohn JF. Multimodal measurement of depression severity in the context of clinical interviews. *IEEE J Biomed Health Inform*. 2017;doi:10.1109/JBHI.2017.2676878. [Epub ahead of print].

Dibeklioglu H, Hammal Z, Yang Y, Cohn JF. Multimodal detection of depression in clinical interviews. Paper presented at: Proceedings of the ACM International Conference on Multimodal Interaction (ICMI), Seattle, WA; 2015.

Dufton LM, Speltz ML, Kelly JP, Leroux B, Collett BR, Werler MM. Psychosocial outcomes in children with hemifacial microsomia. *J Pediatr Psychol*. 2011;36:794-805.

Ekman P, Friesen WV, Hager JC. *Facial Action Coding System [e-book]*. Salt Lake City, UT: Research Nexus; 2002.

Goldsmith HH, Rothbart MK. *The Laboratory Temperament Assessment Battery*. Eugene, OR: University of Oregon; 1999.

Gougoutas AJ, Singh DJ, Low DW, Bartlett SP. Hemifacial microsomia: clinical features and pictographic representations of the OMENS classification system. *Plast Reconstr Surg*. 2007;120: 112e-120e.

Harris J, Kallen B, Robert E. The epidemiology of anotia and microtia. *J Med Genet*. 1996;33:809-813.

Heike CL, Stueckle LP, Stuhaug ET, Pimenta LA, Drake AF, Vivaldi D, Sie KC, Birgfeld CB. Photographic protocol for image acquisition in craniofacial microsomia. *Head Face Med*. 2011;7:25.

Heike CL, Wallace E, Speltz ML, Siebold B, Werler MM, Hing AV, Birgfeld CB, Collett BR, Leroux BG, Luquetti DV. Characterizing facial features in individuals with craniofacial microsomia: a systematic approach for clinical research. *Birth Defects Res A Clin Mol Teratol*. 2016;106:915-926.

Horgan JE, Padwa BL, LaBrie RA, Mulliken JB. OMENS-Plus: analysis of craniofacial and extracraniofacial anomalies in hemifacial microsomia. *Cleft Palate Craniofac J*. 1995;32:405-412.

Maris CL, Endriga MC, Omnell ML, Speltz ML. Psychosocial adjustment in twin pairs with and without hemifacial microsomia. *Cleft Palate Craniofac J*. 1999;36:43-50.

Matias R, Cohn JF. Are MAX-specified infant facial expressions during face-to-face interaction consistent with differential emotions theory? *Dev Psychol*. 1993;29:524-531.

Mattson WI, Cohn JF, Mahoor MH, Gangi DN, Messinger DS. Darwin's Duchenne: eye constriction during infant PosET and distress. *PLoS One*. 2013;8:e80161.

Messinger DS, Mattson WI, Mahoor MH, Cohn JF. The eyes have it: Making positive expressions more positive and negative expressions more negative. *Emotion*. 2012;12-430.

Oster H. Emotion in the infant's face: insights from the study of infants with facial anomalies. *Ann N Y Acad Sci*. 2003;1000:197-204.

Oster H. The repertoire of infant facial expressions: an ontogenetic perspective. In: Nadel J, Muir D, eds. *Emotional Development: Recent Research Advance*s. New York, NY: Oxford University Press; 2005:261-292.

Padwa BL, Evans CA, Pillemer FC. Psychosocial adjustment in children with hemifacial microsomia and other craniofacial deformities. *Cleft Palate Craniofac J*. 1991;28:354-359.

Pertschuk MJ, Whitaker LA. Psychosocial adjustment and craniofacial malformations in childhood. *Plast Reconstr Surg*. 1985;75:177-184.

Petersson RS, Sampson DE, Sidman JD. Dynamic facial reanimation with orthodromic temporalis tendon transfer in children. *JAMA Facial Plast Surg*. 2014;16:432-436.

Pillemer FG, Cook KV. The psychosocial adjustment of pediatric craniofacial patients after surgery. *Cleft Palate J*. 1989;26: 201-207.

Poswillo D. The aetiology and pathogenesis of craniofacial deformity. *Development*. 1988;103:207-212.

Rosenstein D, Oster H. Differential facial responses to four basic tastes in newborns. *Child Dev*. 1988;59:1555-1568.

Rothman KJ. Six persistent research misconceptions. *J Gen Intern Med*. 2014;29:1060-1064.

Segal L, Oster H, Cohen M, Caspi B, Meyers M, Brown D. Smiling and fussing in seven-month-old preterm and full-term black infants in the still-face situation. *Child Dev*. 1995;66: 1829-1843.

Snyder HT, Bilboul MJ, Pope AW. Psychosocial adjustment in adolescents with craniofacial anomalies: a comparison of parents and self-reports. *Cleft Palate Craniofac J*. 2005;42:548-555.

Tronick EZ. Emotions and emotional communication in infants. *Am Psychol*. 1989;44:112-119.