The following publication Chan, R. S. M., Wong, W. W. S., Zhou, T. Y., Liu, Y., Tsang, H. T., Luk, C. L., Chan, T. H., Lam, W. Y. H., & Tsoi, J. K. H. (2024). Human-AI collaboration improves adults' oral biomechanical functions: A multi-centre, self-controlled clinical trial. Journal of Dentistry, 150, 105354 is available at https://doi.org/10.1016/j.jdent.2024.105354.

Human-AI collaboration improves adults' oral biomechanical functions: a multi-centre, self-controlled clinical trial

Abstract

Objectives

Maintenance of oral muscle functions is important for survival and communication. Utilizing Artificial Intelligence (AI) as a self-health-management material has shown promise. Here we developed a functional and AI-enabled smartphone *e-Oral* APP that provided real-time feedback features to let humans collaborate with AI, targeting to improve oral biomechanical and speech functions of adults.

Methods

A total of 113 neurologically healthy Hong Kong citizens over 50 years old were recruited in this study from June to October 2022. A set of 12 exercises of the tongue, lips, and jaw were conducted with guidance provided in the self-developed *e-Oral* APP, in which visual-audio demonstrations and immediate feedback on facial movements by an open-source AI-pretrained 68-point Ensemble of Regression Tree (ERT) face landmark detector was given to the users. After 8-week exercises, assessments after 1-week and 4-week post-training were conducted. Primary outcomes were measured on oral functions, including occlusal force, masticatory efficiency, tongue strength, along with dry mouth condition, and oral health-related quality of life (OHQoL), while secondary outcomes were comprised of oral diadochokinesis (DDK) and swallowing ability. The usability of the APP was evaluated by a self-written questionnaire.

Results

After exercising a set of 12 exercises on the tongue, lips, and jaw for 8 weeks, 70 out of 113 participants (average age of 67.70±4.93 y.o., 52 female and 18 male) had a significant improvement in the occlusal force (p<0.001), while masticatory efficiency (p=0.002), tongue pressure (p<0.001) and endurance (p=0.004) were also improved. 80.3% of respondents rated the APP with an overall rating of 3 out of 5 points, and 68.6% would recommend the APP to others.

Conclusions

In conclusion, AI-based APP can be an effective approach to help healthy adults improve their occlusal force, masticatory efficiency, tongue functions, and oral diadochokinesis after 8-week home oral exercises. Furthermore, these improvements can be sustained for at least four weeks.

Clinical Significance

This is the first AI-assisted APP developed for oral muscle training. Our findings demonstrated that a self-administrated human-AI collaboration APP can improve clinically oral muscle biomechanics and functions for healthy adults. AI technology in smartphone gadgets provides a cost-effective, convenient, and reliable means for oral muscles training for adults.

Keywords: Artificial intelligence; oral functions; biomechanical; occlusal force; masticatory efficiency; tongue strength; dry mouth; oral diadochokinesis; swallowing ability.

1. Introduction

Globally, there are approximately 761 million people aged over 65 years. In the next 30 years, the number of older individuals worldwide is expected to more than double, exceeding 1.6 billion people by 2050, with Asia having the largest elderly population [1]. Like other developed regions, Hong Kong will enter a "super-aging" society in 2034. By 2069, the number of citizens aged 50 or over will reach 4.1 million, accounting for about 55% of the total population, with the old-age dependency ratio increasing to 2:1 or close to 1:1 [2]. In the context of an increasing elderly population worldwide, it becomes increasingly important to prolong healthy life expectancy. Of clinical importance, oral health is among of the significant factors contributing to the psychological and social well-being of the elderly [3]. Oral motor skills, including occlusal force, mastication, swallowing, and diadochokinesis, are among one the oral health factors strongly correlated with frailty of the elderly [3]. Oral motor functions deteriorate with age, resulting in oral-related problems like dysphagia [4], decreased masticatory function, and instability in oral diadochokinesis [5], as well as holistic health concerns such as cognitive decline [6], sarcopenia and malnutrition [7]. Therefore, there has been a growing interest in enhancing oral muscle function of the elderly, with positive outcomes reported in recent studies and reviews [8-12].

Initially, oral exercises were conducted to improve poor oral motor skills of the neurologically impaired population due to stroke [13] or Parkinson's Disease [14]. Since the 2000s, researchers have been keen on studying the effects of oral motor exercises on healthy senior adults and found noteworthy outcomes [8, 15]. Since the past decade, oral exercise programs have been introduced to Japanese and Korean societies for the elderly, and are now being advocated in many countries for improving weakened oral function[8].

While most were implemented under the assistance or supervision of health care professionals and taken place at senior centers or clinics [16, 17], the use of self-administered home-based oral muscle exercises, if conducted accurately and in high intensity, will serve as a convenient and cost-effective means for enhancing oral functions. An initial study on improving swallowing tongue pressure in 11 older adults yielded positive outcomes in tongue strength immediate post-training [18]. However, such improvement was not maintained at 12 weeks post-intervention.

Utilizing AI technology in health management has flourished in terms of identification, assessment and treatment of different health issues [19, 20]. An ensemble of regression trees (ERT), one of the machine learning algorithms, has been used to analyze human motion movement by achieving super-real-time performance with high-quality predictions by estimating landmark positions directly from a sparse subset of pixel intensities [21]. ERT has been applied in the medical research such as facial palsy classification [22]. Its application in oral motor training seems to be promising, as it tracks facial movements such as mouth opening immediately and provides concurrent biofeedback to the user while the numeric facial data captured can be used to evaluate the user's performance. Initial attempts to demonstrate the feasibility of utilizing facial feature detection in oral exercises in older adults have been reported [23]. Nevertheless, both platforms ran on personal computers with the assistance of experimenter's smartphone. Although some studies have been conducted on oral muscle exercises, as well as initial attempts to demonstrate that facial feature detection can be utilized to perform oral exercises in older individuals, most of the oral motor training was provided either upon professional guidance or with the use of costly computers and cameras that require a high level of expertise for operation yet there is no guarantee that all elderly persons will reach these resources conveniently and reliably in their communities.

Therefore, to allow adults to conduct oral muscle training with minimal guidance from healthcare professionals, an AI-enabled smartphone APP – *e-Oral* APP – was developed that could feedback provision and performance tracking in oral exercises. This multi-centre and self-controlled clinical trial approach aimed to evaluate the efficacy and user experience of this mobile application in facial movement tracking, which allows people easily equipped and ready-to-exercises anytime they want, in improving oral functions of healthy adults.

2. Methods

2.1 Study Design and Population

The study was opened to all citizens in Hong Kong between June and October 2022. Recruitment leaflets were posted on social media networks and sent to elderly centers in most regional communities. The study was approved by the Institutional Review Board of the University of Hong Kong (IRB UW21-324), and it was preregistered on ClinicalTrials.gov (ID NCT05549648). The trial followed the SPIRIT-AI extension guideline for clinical trial involving artificial intelligence [24].

Inclusion criteria included: 1) neurologically healthy Hong Kong citizens over 50 years old, 2) possessed an iPhone (iOS 14 or above) or Android (version 10 or above) smartphone with internet access (either WIFI or mobile data) and was able to use it on their own or with the help of at least one family member, 3) passed the Montreal Cognitive Assessment 5-Minute Protocol (Hong Kong Version) [25], and 4) possessed at least one pair of natural occluding molars. Exclusion criteria included: 1) subjects who had a cerebrovascular accident, head and neck cancers, neurologic diseases, or other severe systemic diseases, 2) receiving radiotherapy and chemotherapy, 3) inability to read Chinese and communicate in Cantonese, and 4) uncontrolled diabetics to avoid interference on masticatory efficiency test results.

Screening was conducted by trained researchers and speech therapists, written informed consent was obtained from the participants and they might withdraw from the study at any time, the reason would be documented if they provided. As a result, six centers were enrolled, 18 out of 131 participants did not meet the inclusion criteria, with a total of 113 participants were recruited.

2.2 Study Procedures

Figure 1 illustrates the study procedure. Following screening, subjects underwent a first examination, followed by a second examination four weeks later. The results of these examinations served as a baseline. In the second assessment, subjects installed and used the *e-Oral* APP under the researchers' guidance, following which they began an 8-week oral muscle training program. Participants were advised to carry out oral muscle training on a daily basis or as frequently as they could arrange. Speech therapists scheduled regular video/phone consultations every two weeks to answer their queries and troubleshoot the problems in APP use, if any. If the participant felt any discomfort during the exercises, they could discontinue the exercises and inform the speech therapists. No other concomitant oral motor training or interventions were implemented. Participants' oral functions were evaluated again in the third post-training examination upon competition of the training period, followed by the final maintenance examination four weeks later. Subjects were given toothpaste as souvenirs after the first, third, and final examinations to promote retention, and a phone stand after installing the APP.

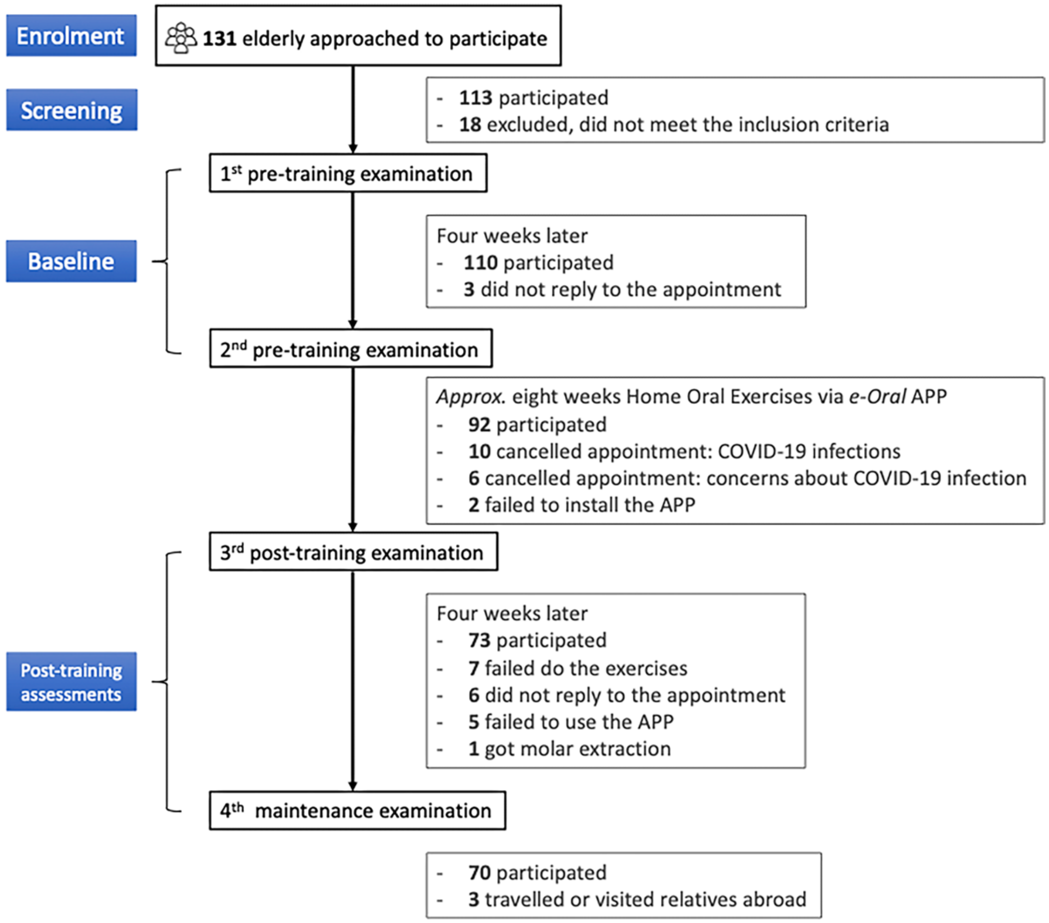

Figure 1. The flow diagram of the clinical trial.

2.3 Home Oral Exercises and Outcomes

Figure 2 depicts the pre-trained cascade regression model used in our study, the *e-Oral* APP interface, and the target of each exercise based on previous studies [15, 26]. A set of 12 oral exercises of the tongue, lips, and jaw were conducted with guidance provided in the APP. The facial landmark detection was achieved via the Ensemble of Regression Trees (ERT) [21], an open-source cascade regression gradient-boosting algorithm that estimates different facial features, including the nose, the eyes, the mouth, and the edge of the face, in 68 points. A human-AI interaction was achieved via the dynamic movements of the 68 detection points in accordance with facial movements of the users, in which the level of expertise required was the capability of using a smartphone. During training, a video demonstrating the respective oromotor exercise was shown in Fig. 2. Then training proceeded in which the participant would face the screen of the mobile phone, with his/her facial features picked up by the front-facing camera, identified by ERT, and shown on the screen as immediate feedback to the user, these landmarks can help users achieved the target motion properly during training. Since the distance between the user and the camera might vary between participants and different occasions of training, the ratio between the anchor and the target would inform us more precisely about the degree of jaw opening performed by the user. Due to the limitations of the ERT, only movements of the lips and jaw were reliably detected. When the application encounters errors, the user would be notified and instructed to suspend use until the issue was resolved.

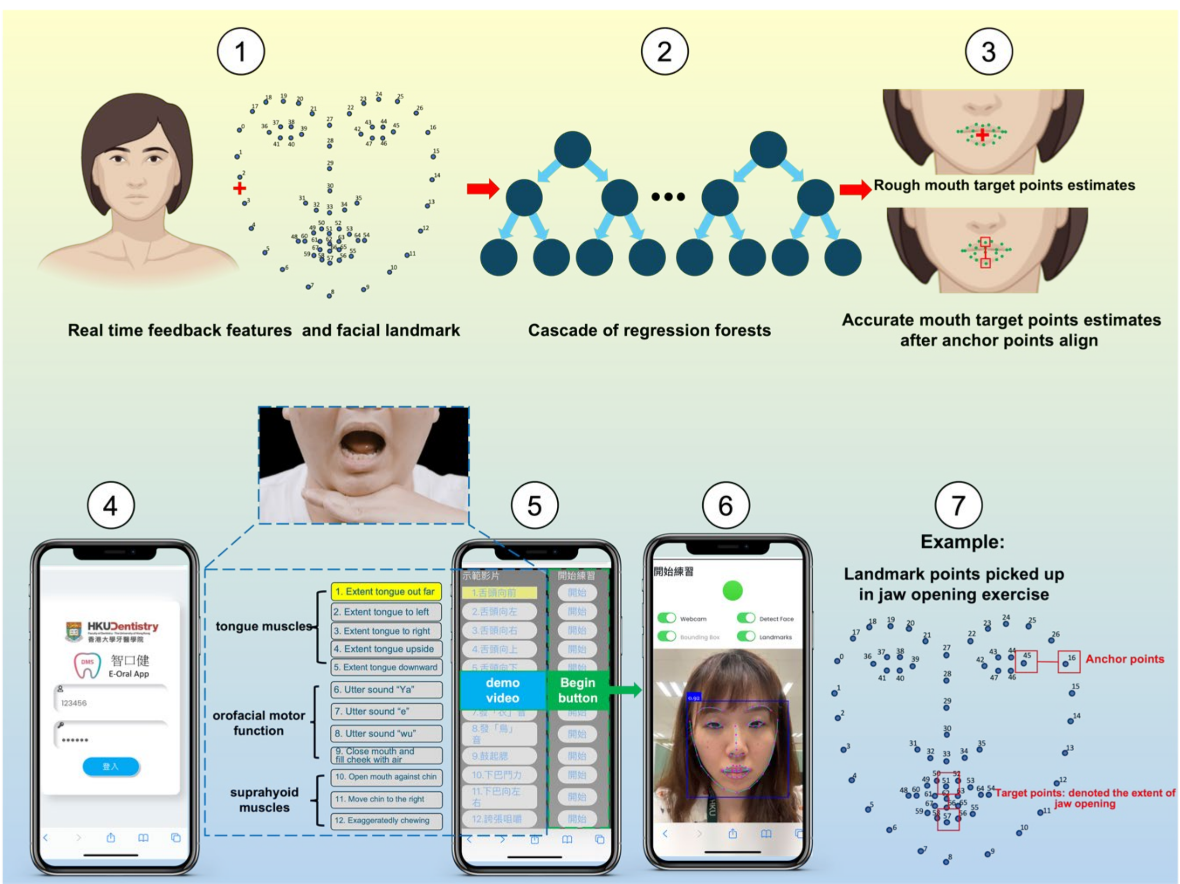

Figure 2. **Pre-trained cascade regression model illustration and the *e-Oral* APP interface.** ①-③. In the pre-trained cascade regression model used in our study, real-time features and facial landmarks were used to regress the mouth target points estimates. ④-⑥ *e-Oral* APP interface display, whereas ④ The APP login interface; ⑤ The main interface, the left column illustrates the demonstration video clicking buttons, and the right column leads to the exercise page; ⑥ When face detection was enabled, the four buttons above the live image window would turn from red to green; ⑦Visualization of the task output: red circled points are examples of anchor points and target points. Displacement of landmark points more than 3.7mm could be tracked reliably.

2.4 Primary Outcomes

Oral function examination, including occlusal force, masticatory efficiency, tongue pressure and endurance, together with dry mouth condition and oral health-related quality of life, served as primary outcomes. The occlusal force was measured by occlusal pressure measurement film [27] (Dental Prescale II, GC Corp), following scanning through a dedicated scanner (GT-X830, Epson) with analysis software (Bite force analyzer, GC Corp). A glucose sensor (GS-II, GC Corp) was employed to quantify the masticatory efficiency by examining the filtrate resulting from chewing a 2 g of gummy jelly for 20 seconds. Their maximum tongue pressure and endurance were assessed by a tongue pressure sensor [28] (TPM-02, JMS). Tongue endurance is defined as the time it takes to reach 60% or more of maximum tongue strength. Three trials were performed while the average value was taken to determine tongue function. Participants using partially removable dentures should rest them in place during examination. The Chinese-validated version of the Oral Health Impact Profile-14 (OHIP-14) questionnaire [29] was applied to assess oral health-related quality of life [30]. Higher scores represent poorer oral condition-related quality of life. A Chinese (Cantonese) version of the dry mouth questionnaire [31] consisting of 8-item xerostomia questions with a maximum total score of 32 points was used for subjective assessment, with a Likert format rated on a 4 -point scale (1= not at all, 4= very much). Participants graded each aspect, with a higher score indicating impaired salivary function.

2.5 Secondary Outcomes

A Chinese version of Eat-10, a self-reported outcome instrument on swallowing function [32], consisted of 10 questions arranged on a 5-point Likert scale (0 not at all, 4 very much). Physical symptoms of chewing and swallowing disorders were assessed via the Dysphagia Risk Assessment of Community-dwelling elderly persons (DRACE) questionnaire [33]. DRACE, which was forward and backward translated by our research team, included 12 variables involving chewing, swallowing, and coughing problem, etc., based on a 3-level scale from 0 (never occurs) to 2 (occurs frequently). DDK of the lips and tongue, a speech-related measure, was examined via alternating motion rate of /pa/, /ta/, and /ka/ and sequential motion rate (i.e., /pataka/), during a sustained production of the target syllables for 10 s [34].

2.6 Evaluation of APP Usability and Likability

Following the completion of training, participants were given a questionnaire consisting of 19 multiple-choice questions and one open-ended question (**Supplementary A.1**) in four aspects, including APP usage adherence, APP design, exercise engagement, and satisfaction with APP usage. Participants were encouraged to give any other suggestions regarding improvements to APP in the open-ended question.

2.7 Statistical Analysis

The planned sample size was calculated based on a previous study [15], using

G*Power (version 3.1). A power of 90% and a significance level of 5% (two-sided) were adopted. With an effect size of 0.5 and a drop-out rate of 25%, a total of 106 participants were required.

SPSS (v.28, IBM) was used for the data analysis, with statistical significance set at α=0.05. Primary and secondary outcomes were described using median, interquartile range (IQR), mean, and standard deviation (SD). Skewness between -0.5 and 0.5, or a skewness ratio of <2, was considered symmetric; otherwise, the data distribution was considered asymmetrical. Data with a negative skewness of <0.5 would undergo a log transformation before further analysis was conducted.

Masticatory efficiency and tongue strength were normally distributed while occlusion force, tongue endurance, DDK, dry mouth questionnaire, OHIP-14, DRACE, and EAT-10 were not normally distributed. To examine the stability of the performance collected in the first and second examinations, which served as a baseline for comparison after training, participants' scores obtained in these two examinations were compared using Paired samples-t test and Wilcoxon test before averaging them to determine baseline value. For those questionnaire outcomes designed as non-negative integers, the outcomes at baseline value would be defined after rounding before subsequent statistical analyses were conducted.

To estimate the training effects on the outcomes over time, Generalized Estimating Equations (GEE) model with linear distribution for scale outcomes and negative binomial distribution for non-negative integer outcomes were performed using unstructured correlation for repeated measures within the same participants. Fixed factors (assessment time point, gender, state of employment and marital status) and covariates were included (the number of years of education and age) as intercept variables, while primary/secondary measures as depended on variables were fitted in the model, factorial interaction effects were further explored. Details on statistical methods and model specifications are shown in **Figure A.1**.

2.8 Interrater Reliability

5% of the total data were randomly selected for interrater reliability at the end of the study. The DDK results were analyzed by two different speech therapists, while data entry reliability for the rest of the outcomes was evaluated by two researchers. Intraclass Correlation Coefficients (ICC) using a one-way random approach were performed.

2.9 Data Management

Personal information and datasets were stored on a locked cabinet and a password-protected cloud, which is only accessible by the researchers. At the same time, the backend database of the APP was only available to the programmer and principal investigator (PI). The results of individuals were informed after the final examination, while overall findings were announced via social media and publications.

Adverse event was defined as a medical condition resulting in participants consulting a physician due to circumstances possibly related to the intervention. The occurrences of any adverse events or protocol modifications were recorded and reviewed in interim reviews held on a monthly basis by the research team. The interim results were available to team members.

3. Results

3.1 Participants

In total, 113 participants consented and enrolled, 92 completed the pre-training inspections, and 90 (90/92, 97.8%) successfully installed and used the APP. Lastly, 70 people completed all four inspections by June 2023 and were included in the analytical sample **(Figure 1)**. Of those, the mean (SD) age of participants was 67.70 (4.93) years old, while the number of years of education was 10.21 (2.615); the mean (SD) score of MOCA was 26.67 (2.84); 52 (74.3%) of them were female and 18 (25.7%) were male; 65 (92.9%) was retired and 5 (7.1%) was still employed while 55 (78.6%) was married and 15 (21.4%) was single or divorced. No adverse events were reported among the interventions.

3.2 Primary and Secondary Outcomes

All the primary outcome variables measured in the two baseline examinations were not statistically different ($p>0.05$); hence, baseline performance was determined by averaging the two evaluations. In the secondary outcome, participants performed similarly in the four questionnaires and the alternating motion of the lips and tongue (i.e., /pa/, /ta/, and /ka/) in the two baseline evaluations. However, the performance on /pataka/ of the diadochokinetic test differed statistically, which might subsequently affect the evaluation of training effect. Therefore, the data on /pataka/ was excluded from further investigations. The performance of the averaged baseline, post-training, and maintenance examinations is shown in **Table 1**.

After the 8-week oral exercises, the GEE model reported a significant training effect ($p<0.001$) on the occlusal force, masticatory efficiency, tongue strength, tongue endurance, and DDK across three assessments (**Table A.1**), indicating improvements in the above functions. Such improvements were evident post-training and in the maintenance phase. The dry mouth questionnaire found a negative training effect at post-training ($p=0.005$), which meant heavier dryness after training, while other questionnaire items exhibited no influence. Among different characteristics of the participants, gender and age were found to interact with one of the primary outcomes in the post-training assessment. In masticatory efficiency, women made significant smaller improvements compared to men. When age increased, the participants gained less in chewing ability ($p<0.001$) after training. Interestingly, a significant three-way interaction existed among timepoint, gender and age: older women achieved greater gains in masticatory efficiency than that of men ($p<0.001$) in the post-training examination. Another interaction effect with timepoints concerned the employment status of the participants; also, it was found that in immediate post-training, retirees experienced a greater decrease in mouth dryness than employed individuals ($p<0.001$).

Apart from the timepoint effect, there was a significant age-related difference in masticatory efficiency and scores of the dry mouth questionnaire, with older individuals chewing sugar less efficiently ($p=0.043$), in addition to having a drier mouth behavior ($p=0.043$). Gender effect was observed, with females demonstrating lower masticatory efficiency than males ($p=0.034$) and less maximum tongue strength ($p<0.001$). GEE also revealed that retired participants scored lower on the dry mouth questionnaire than those still employed ($p<0.001$); Participants receiving more years of education had a higher DDK rate of /pa/, /ta/, and /ka/ [DDK-pa ($p=0.005$); DDK-ta ($p=0.007$); DDK-ka ($p=0.017$)] **(Table A.2)**. **Figure 3** summarizes the pairwise comparison result based on the GEE model.

Table 1. Descriptive outcome characteristics of baseline, post-training, and maintenance phases.

Note. Baseline value was derived by averaging two pre-training inspections after no statistical differences were obtained. n=70 represents the total number of participants who completed all four assessments and were included for analysis; three missing data points in the DDK test were not considered, and only valid value was used.

n=70	Baseline	Post-training	Maintenance
Occlusal force (Newton)			
Mean	648.30	870.88	842.79
SD	391.48	603.98	506.04
Median	596.73	805.20	807.60
IQR	338.84-851.34	437.23-1102.77	435.48-1097.88
Masticatory efficiency (ml/dL)			
Mean	215.87	238.87	252.66
SD	75.72	74.84	78.74
Median	208.25	226.50	250.00
IQR	157.00-282.13	184.25-288.75	211.25-298.25
Tongue pressure (KPa)			
Mean	28.25	31.73	33.30
SD	6.66	6.64	6.43
Median	27.51	31.65	32.92
IQR	23.16-33.51	26.56-36.80	28.28-38.92
Tongue endurance (s)			
Mean	7.02	8.48	8.74
SD	5.73	4.96	5.16
Median	5.38	7.85	7.50
IQR	3.36-9.20	4.52-12.85	4.61-11.42
Dry mouth (points)			
Mean	16.51	17.63	17.11
SD	3.77	4.51	5.00
Median	16.00	17.00	16.00
IQR	13.38-19.50	14.00-22.00	13.00-21.00
OHIP-14 (points)			
Mean	11.54	12.61	11.67
SD	9.02	9.94	9.59
Median	10.75	14.00	10.50
IQR	4.38-17.00	3.00-20.00	3.00-18.25
DRACE (points)			
Mean	5.84	5.96	6.14
SD	4.07	4.02	5.14

	Median	5.00	6.00	6.00
	IQR	2.50-8.50	3.00-9.00	2.00-8.25
Secondary outcomes	**EAT-10 (points)**			
	Mean	4.67	4.80	5.01
	SD	5.91	6.01	5.74
	Median	2.50	2.00	3.00
	IQR	0.50-6.00	0.00-7.25	1.00-8.00
	DKK-/pa/ (rep/s)			
	Mean	5.29	5.60	6.07
	SD	1.43	1.08	1.00
	Median	5.46	6.21	6.08
	IQR	3.96-6.48	5.41-6.69	5.51-6.87
	DKK-/ta/ (rep/s)			
	Mean	4.73	5.24	5.45
	SD	1.48	1.20	1.22
	Median	4.91	5.29	5.52
	IQR	3.39-6.04	4.17-6.21	4.75-6.30
	DDK-/ka/ (rep/s)			
	Mean	4.56	5.31	5.25
	SD	1.42	1.09	1.16
	Median	4.67	5.52	5.36
	IQR	3.41-5.74	4.53-6.07	4.55-6.16

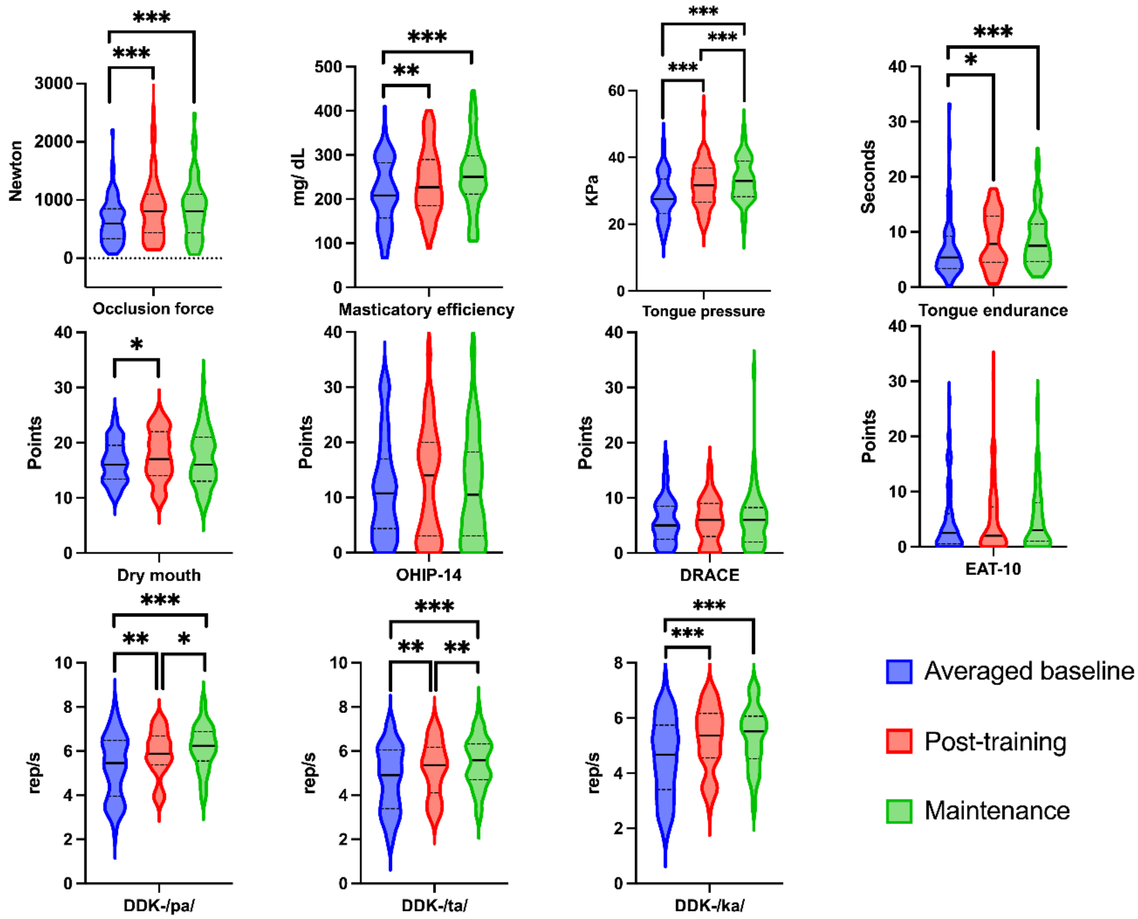

Figure 3. Comparisons before and after oral exercise training using *e-Oral* APP. The violin plots display overall distribution with upper adjacent and lower adjacent values of each outcome. The solid line refers to the median while the upper and lower dashed lines refer to the third quartile and first quartile respectively. *$p < 0.05$, **$p < 0.01$, *** $p < 0.001$.

3.3 APP Feasibility and Acceptability Questionnaire

In total, 92 (97.87%, 92/94) participants successfully installed the *e-Oral* APP and 73 of them enrolled in the third assessment, 71 completed the questionnaire (**Figure 1**). Six (8.6%) participants used the APP on a daily basis, while 38(54.3%) users used the APP at least three times a week (1.4% missing). When using the APP, participants experienced issues with mouth movement detection, and video demonstration or login problems. Participants also made some other suggestions in the open response section, including that the exercise time might take too long, which produced fatigue of the mouth, as well as that the APP ran slow. 80.3% of respondents gave the APP an overall rating of 3 points or higher, and 68.6% (1.4% missing) would recommend the APP to others. **Figure 4** presents the results of the rest of the rating items in the questionnaire.

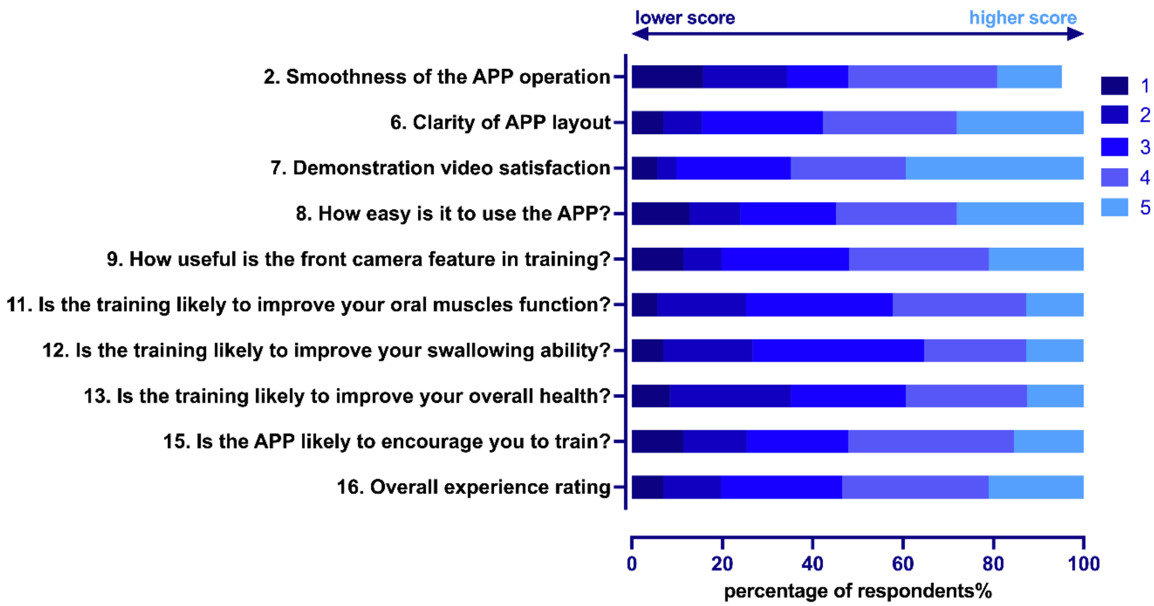

Figure 4. The distribution of responses to the rating questions about *e-Oral* APP. The serial numbers in the figure correspond to the serial numbers in **Supplementary Questionnaire A.1.**

3.4 Interrater Reliability

Table 2 displays results on interrater reliability among expert raters, in which data entry reliability and DDK raters met excellent agreement (Cronbach's Alpha>0.9, ICC>0.9, $p<0.001$) in all tests.

Table 2. Interrater reliability in various examinations.

Note. Cronbach's α= Cronbach's Alpha. ICC=Intraclass Correlation Coefficient. 95% CI=95% Confidence Interval. Single measures results were reported; Intraclass

Correlation>0.74 and Cronbach's Alpha>0.8 was accepted for good agreement.

	Cronbach's α	ICC	95% CI	*p*-value
Occlusal force	0.999	0.999	0.997-1.000	<0.001
Masticatory efficiency	0.997	0.994	0.982-0.998	<0.001
Tongue pressure	0.998	0.996	0.988-0.999	<0.001
Tongue endurance	0.998	0.996	0.989-0.999	<0.001
Dry mouth	0.993	0.987	0.965-0.996	<0.001
OHIP-14	0.999	0.999	0.997-1.000	<0.001
DRACE	0.999	0.998	0.996-0.999	<0.001
EAT-10	0.995	0.991	0.975-0.997	<0.001
DKK-/pa/	1.000	0.999	0.998-1.000	<0.001
DKK-/ta/	1.000	0.999	0.997-1.000	<0.001
DDK-/ka/	1.000	0.999	0.998-1.000	<0.001

4. Discussion

This clinical trial was among the first that investigated the efficacy of AI-enable and self-administered smartphone applications in improving oral muscle functions for healthy adults. The results suggested that oral motor training offered by *e-Oral* APP, which involved range-of-motion, endurance (isometric), and some isotonic exercises of lips, tongue, and jaw, were effective in enhancing occlusal force, masticatory efficiency, tongue functions, and oral diadochokinesis. The effects could be observed immediate post-training and maintained for at least four weeks ($p>0.05$).

Age-related declines in oral muscle functions lead to oral frailty, resulting in reduced general well-being and quality of life [3, 7]. Therefore, the 12 exercises (**Figure 2.5**) we selected aimed at training the tongue, orofacial motor functions and suprahyoid muscles. These exercises were able to activate muscles associated with oral function: the orofacial myofunctional muscles, which control facial expressions [35], lip movements [36], and articulatory oral motor skills [37], as well as the suprahyoid, which participates in jaw-opening [26]. Whilst these training movements can be captured by the front camera of the APP, the collected data can provide users with biofeedback immediately. Previous research has shown that muscle strength is related to neural factors and muscle hypertrophy [17, 38] or muscle thickness [39] in elderly, which could be the mechanisms responsible for the increase in muscle strength observed in this study.

Although many studies have been conducted on oral muscle exercises, as well as initial attempts to demonstrate that facial feature detection can be utilized to perform oral exercises in older individuals, most of the oral motor training was provided either upon professional guidance or with the use of costly computers and cameras that require a high level of expertise for operation yet there is no guarantee that all elderly persons will reach these resources conveniently and reliably in their communities. The positive outcomes in oral functions obtained from our study provides initial evidence on the application of AI technology in smartphone gadgets, which provides a cost-effective, convenient, and reliable means for oral motor training for the elderly.

Based on the feedback from the users' questionnaire, we attributed the success of *e-Oral* APP to its user-friendly interface, the human-AI interaction which provides immediate biofeedback during training, high accessibility in their own mobile phones, and regular follow-up video/voice calls by the research team to maintain subjects' motives in training. Being one of the leading facial alignment models with advantages in both speed and accuracy, its error rates range from 3.8% to 6.4%, as validated in different databases [40]. The error was mainly caused by dim lighting and misaligned faces, which prevented 68 points captured. To ensure that users can operate the APP correctly and reduce the error rate, various measures have been implemented, such as teaching users how to use the APP and following up every fortnight. Participation was rewarded with a phone stand, which enabled users to position the camera directly on their faces more easily and stably. Thus, there was no plan for the identification and analysis of performance errors. Nevertheless, some users reported slow facial tracking responses and long exercise times during use. Possible reasons might include insufficient software/hardware resources of their smartphones or compatibility issues with different operating systems.

Our findings are generally consistent with previous reports on home-based oral exercises, in which improvements in mastication [8] and tongue pressure [18] have been found. Participants made improvements regardless of gender and age in most of the outcomes, indicating that the *e-Oral* APP was generally effective for the population. One unexpected finding concerned the drier mouth symptoms after training, which might be related to prolonged mouth opening during exercises. We attributed the absence of improvements in oral health-related quality of life to the lack of attention paid to teeth and periodontal health in the *e-Oral* APP. The non-noticeable improvement in swallowing function, as reflected in DRACE and Eat-10, might be explained by a ceiling effect, as all subjects were on a normal diet.

There are also some limitations to this study. Firstly, since tongue movements were sometimes limited to the mouth, our algorithm could not accurately capture tongue dynamics and therefore unable to provide users with accurate tongue guidance. Nonetheless, improvements in tongue functions were still evident. Secondly, the outbreak of COVID-19 in Hong Kong during the study period made follow-up visits challenging, resulting in a dropout rate of 38%, exceeding the expected 25%. As such, future research should focus on improving algorithm capabilities for capturing tongue details, as well as enhancing visual and verbal reinforcement to improve user interfaces. Moreover, we will explore its potential for supporting physically frail populations, such as patients with neurological impairment, sarcopenia, or dyspepsia.

5. Conclusions

In conclusion, the AI-enabled APP can be an effective approach to help healthy adults improve their occlusal force, masticatory efficiency, tongue functions, and oral diadochokinesis after 8-week home oral exercises, and these improvements can be sustained for at least four weeks.

CRediT authorship contribution statement

Rosalind Sin Man Chan: Writing – original draft, Validation, Methodology, Investigation, Formal analysis, Data curation. **Winsy Wing Sze Wong:** Writing – review & editing, Supervision, Software, Project administration, Conceptualization. **Tian Yu Zhou:** Investigation, Formal analysis, Data curation. **Ying Liu:** Investigation, Formal analysis, Data curation. **Hiu Tung Tsang:** Validation, Investigation, Formal analysis. **Chun Lam Luk:** Investigation, Formal analysis. **Tsz Hei Chan:** Investigation, Formal analysis. **Walter Yu Hang Lam:** Supervision, Conceptualization. **James Kit Hon Tsoi:** Project administration, Methodology, Funding acquisition, Data curation.

Declaration of competing interests

All authors declare no competing interests.

Code and data availability statement

The code of *e-Oral* APP, deidentified participant data, study protocol, statistical analysis methods, and informed consent form are available on reasonable request from the principal investigator (jkhtsoi@hku.hk) who will coordinate with the study senior centers and hospital to ensure that any data sharing complies with the General Data Protection Regulations, the Hong Kong Personal Data (Privacy) Ordinance, and other legal requirements.

Acknowledgments

We would like to thank all our study participants and staff at trial sites who dedicated their time to support our research. This study was supported by the Nakao Foundation. The funding source was not involved in the study design, collection, analysis, and interpretation of data, writing of the report, or decision to submit the manuscripts.

References

[1] 2023 Theme: Fulfilling the Promises of the Universal Declaration of Human Rights for Older Persons: Across Generations. https://www.un.org/en/observances/older-persons-day.

[2] Hong Kong Population Projections 2020-2069, 2020.

[3] V. Dibello, R. Zupo, R. Sardone, M. Lozupone, F. Castellana, A. Dibello, A. Daniele, G. De Pergola, I. Bortone, L. Lampignano, G. Giannelli, F. Panza, Oral frailty and its determinants in older age: a systematic review, Lancet Healthy Longev 2(8) (2021) e507-e520 https://doi.org/10.1016/S2666-7568(21)00143-4

[4] K. Yoshimi, K. Hara, H. Tohara, A. Nakane, K. Nakagawa, K. Yamaguchi, Y. Kurosawa, S. Yoshida, C. Ariya, S. Minakuchi, Relationship between swallowing muscles and trunk muscle mass in healthy elderly individuals: A cross-sectional study, Arch. Gerontol. Geriatr. 79 (2018) 21-26 https://doi.org/10.1016/j.archger.2018.07.018

[5] T. Kikutani, F. Tamura, K. Nishiwaki, M. Kodama, M. Suda, T. Fukui, N. Takahashi, M. Yoshida, Y. Akagawa, M. Kimura, Oral motor function and masticatory performance in the community-dwelling elderly, Odontology 97(1) (2009) 38-42 https://doi.org/10.1007/s10266-008-0094-z

[6] M. Nagatani, T. Tanaka, B.K. Son, J. Kawamura, J. Tagomori, H. Hirano, M. Shirobe, K. Iijima, Oral frailty as a risk factor for mild cognitive impairment in community-dwelling older adults: Kashiwa study, Exp. Gerontol. 172 (2023) 112075 https://doi.org/10.1016/j.exger.2022.112075

[7] V. Dibello, F. Lobbezoo, M. Lozupone, R. Sardone, A. Ballini, G. Berardino, A. Mollica, H.J. Coelho-Junior, G. De Pergola, R. Stallone, A. Dibello, A. Daniele, M. Petruzzi, F. Santarcangelo, V. Solfrizzi, D. Manfredini, F. Panza, Oral frailty indicators to target major adverse health-related outcomes in older age: a systematic review, Geroscience 45(2) (2023)

663-706 https://doi.org/10.1007/s11357-022-00663-8

[8] H.J. Kim, J.Y. Lee, E.S. Lee, H.J. Jung, H.J. Ahn, B.I. Kim, Improvements in oral functions of elderly after simple oral exercise, Clin. Interv. Aging Volume 14 (2019) 915-924 https://doi.org/10.2147/cia.s205236

[9] C. Namiki, K. Hara, H. Tohara, K. Kobayashi, A. Chantaramanee, K. Nakagawa, T. Saitou, K. Yamaguchi, K. Yoshimi, A. Nakane, S. Minakuchi, Tongue-pressure resistance training improves tongue and suprahyoid muscle functions simultaneously, Clin. Interv. Aging 14 (2019) 601-608 https://doi.org/10.2147/CIA.S194808

[10] J.O. Silva, L.D. Giglio, L.V.V. Trawitzki, Effects of tongue strengthening exercises in healthy adults and elderly: an integrative literature review, CoDAS 35(3) (2023) https://doi.org/10.1590/2317-1782/20232021213en

[11] C.J. Lin, Y.S. Lee, C.F. Hsu, S.J. Liu, J.Y. Li, Y.L. Ho, H.H. Chen, Effects of tongue strengthening exercises on tongue muscle strength: a systematic review and meta-analysis of randomized controlled trials, Sci. Rep. 12(1) (2022) 10438 https://doi.org/10.1038/s41598-022-14335-2

[12] S. Nurfatul Jannah, S. Syahrul, K. Kadar, The effectiveness of tongue strengthening exercise in increasing tongue strength among older people with dysphagia: A systematic review, Health Sciences Review 4 (2022) https://doi.org/10.1016/j.hsr.2022.100047

[13] R. Marzouqah, A. Huynh, J.L. Chen, M.I. Boulos, Y. Yunusova, The role of oral and pharyngeal motor exercises in post-stroke recovery: A scoping review, Clin. Rehabil. 37(5) (2023) 620-635 https://doi.org/10.1177/02692155221141395

[14] S. Baram, M. Karlsborg, M. Bakke, Improvement of oral function and hygiene in Parkinson's disease: A randomised controlled clinical trial, J. Oral Rehabil. 47(3) (2020) 370-376 https://doi.org/10.1111/joor.12924

[15] H. Ibayashi, Y. Fujino, T.-M. Pham, S. Matsuda, Intervention study of exercise program for oral function in healthy elderly people, The Tohoku Journal of Experimental Medicine 215(3) (2008) 237-245

[16] E.P. Cho, S.J. Hwang, J.B. Clovis, T.Y. Lee, D.I. Paik, Y.-S. Hwang, Enhancing the quality of life in elderly women through a programme to improve the condition of salivary hypofunction, Gerodontology 29(2) (2012) e972-e980 https://doi.org/10.1111/j.1741-2358.2011.00594.x

[17] J. Robbins, R.E. Gangnon, S.M. Theis, S.A. Kays, A.L. Hewitt, J.A. Hind, The effects of lingual exercise on swallowing in older adults, J. Am. Geriatr. Soc. 53(9) (2005) 1483-1489

[18] H. Kim, N.B. Cho, J. Kim, K.M. Kim, M. Kang, Y. Choi, M. Kim, H. You, S.I. Nam, S. Shin, Implementation of a Home-Based mHealth App Intervention Program With Human Mediation for Swallowing Tongue Pressure Strengthening Exercises in Older Adults: Longitudinal Observational Study, JMIR mHealth and uHealth 8(10) (2020) e22080 https://doi.org/10.2196/22080

[19] K.H. Yu, A.L. Beam, I.S. Kohane, Artificial intelligence in healthcare, Nature Biomedical Engineering 2(10) (2018) 719-731 https://doi.org/10.1038/s41551-018-0305-z

[20] M. Ducret, C.M. Mörch, T. Karteva, J. Fisher, F. Schwendicke, Artificial intelligence for sustainable oral healthcare, J. Dent. 127 (2022) 104344 https://doi.org/https://doi.org/10.1016/j.jdent.2022.104344

[21] V. Kazemi, J. Sullivan, One millisecond face alignment with an ensemble of regression trees, 2014 IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 1867-1874.

[22] J. Barbosa, W.-K. Seo, J. Kang, paraFaceTest: an ensemble of regression tree-based facial features extraction for efficient facial paralysis classification, BMC Med. Imaging 19(1) (2019) 30 https://doi.org/10.1186/s12880-019-0330-8

[23] T. Ando, A. Masaki, Q. Liu, T. Ooka, S. Sakurai, K. Hirota, T. Nojima, Squachu: A training game to improve oral function via a non-contact tongue-mouth-motion detection system, ACM.

[24] S. Cruz Rivera, X. Liu, A.W. Chan, A.K. Denniston, M.J. Calvert, A.I. Spirit, C.-A.W. Group, Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension, Lancet Digit Health 2(10) (2020) e549-e560 https://doi.org/10.1016/S2589-7500(20)30219-3

[25] A. Wong, D. Nyenhuis, S.E. Black, L.S.N. Law, E.S.K. Lo, P.W.L. Kwan, L. Au, A.Y.Y. Chan, L.K.S. Wong, Z. Nasreddine, V. Mok, Montreal Cognitive Assessment 5-minute protocol is a brief, valid, reliable, and feasible cognitive screen for telephone administration, Stroke 46(4) (2015) 1059-1064 https://doi.org/10.1161/STROKEAHA.114.007253

[26] J.C. Oh, Effects of Resistive Jaw-Opening Exercise with Elastic Resistance Bands on Suprahyoid Muscle Activation and Tongue Strength in the Elderly: A Pilot Study, Folia Phoniatrica Et Logopaedica : Official Organ of the International Association of Logopedics and Phoniatrics (IALP) 73(5) (2021) 376-383 https://doi.org/10.1159/000509441

[27] Y.F. Huang, C.M. Wang, W.Y. Shieh, Y.F. Liao, H.H. Hong, C.T. Chang, The correlation between two occlusal analyzers for the measurement of bite force, BMC Oral Health 22(1) (2022) 472 https://doi.org/10.1186/s12903-022-02484-9

[28] M. Yoshikawa, T. Fukuoka, T. Mori, A. Hiraoka, C. Higa, A. Kuroki, C. Takeda, M. Maruyama, M. Yoshida, K. Tsuga, Comparison of the Iowa Oral Performance Instrument and JMS tongue pressure measurement device, J Dent Sci 16(1) (2021) 214-219

https://doi.org/10.1016/j.jds.2020.06.005

[29] M.C.M. Wong, E.C.M. Lo, A.S. McMillan, Validation of a Chinese version of the Oral Health Impact Profile (OHIP), Community Dent. Oral Epidemiol. 30(6) (2002) 423-430

[30] Y. Feng, J.J. Lu, Z.Y. Ouyang, L.X. Xue, T. Li, Y. Chen, Z.R. Gao, S.H. Zhang, J. Zhao, Y.Q. Zhao, Q. Ye, J. Hu, Y.Z. Feng, Y. Guo, The Chinese version of the Oral Health Impact Profile-14 (OHIP-14) questionnaire among college students: factor structure and measurement invariance across genders, BMC Oral Health 22(1) (2022) 405 https://doi.org/10.1186/s12903-022-02441-6

[31] K. LEUNG Chiu Man, Effect of cevimeline on oral health and quality of life in Sjögren's syndrome patients, Thesis (Ph. D.)--University of Hong Kong, 2007.

[32] D. Pu, T. Murry, C.M. Wong May, M.L. Yiu Edwin, M.K. Chan Karen, Indicators of Dysphagia in Aged Care Facilities, J. Speech. Lang. Hear. Res. 60(9) (2017) 2416-2426 https://doi.org/10.1044/2017_JSLHR-S-17-0028

[33] H. Miura, M. Kariyasu, K. Yamasaki, Y. Arai, Evaluation of chewing and swallowing disorders among frail community-dwelling elderly individuals, J. Oral Rehabil. 34(6) (2007) 422-7 https://doi.org/10.1111/j.1365-2842.2007.01741.x

[34] J.E. Pierce, S. Cotton, A. Perry, Alternating and sequential motion rates in older adults, Int. J. Lang. Commun. Disord. 48(3) (2013) 257-64 https://doi.org/10.1111/1460-6984.12001

[35] A.F.M. Borges, K.V.M. Taveira, J.Y.M. Eduardo, R.V.A. Cavalcanti, Orofacial and cervical myofunctional intervention programmes for older adults: A scoping review, Gerodontology (2023) https://doi.org/10.1111/ger.12719

[36] H. Byeon, Effect of orofacial myofunctional exercise on the improvement of dysphagia patients' orofacial muscle strength and diadochokinetic rate, Journal of Physical

Therapy Science 28(9) (2016) 2611-2614

[37] M. Shirobe, Y. Watanabe, T. Tanaka, H. Hirano, T. Kikutani, K. Nakajo, T. Sato, J. Furuya, S. Minakuchi, K. Iijima, Effect of an Oral Frailty Measures Program on Community-Dwelling Elderly People: A Cluster-Randomized Controlled Trial, Gerontology 68(4) (2022) 377-386 https://doi.org/10.1159/000516968

[38] T. Moritani, Neuromuscular adaptations during the acquisition of muscle strength, power and motor tasks, J. Biomech. 26 Suppl 1 (1993)

[39] J.S. Park, N.K. Hwang, H.H. Kim, J.B. Choi, M.Y. Chang, Y.J. Jung, Effects of lingual strength training on oropharyngeal muscles in South Korean adults, J. Oral Rehabil. 46(11) (2019) 1036-1041 https://doi.org/10.1111/joor.12835

[40] Y. Wu, Q. Ji, Facial Landmark Detection: A Literature Survey, International Journal of Computer Vision 127(2) (2019) 115-142 https://doi.org/10.1007/s11263-018-1097-z