DR. HIDEO SATO (Orcid ID : 0000-0003-0942-2923)

Article type : Original Article

Relationship between lip motion detected with a compact 3D camera and swallowing dynamics during bolus flow swallowing in Japanese elderly men

Yushi YAMAMOTO[1], Hideo SATO[2], Hisako KANADA[1], Yoichiro IWASHITA[3], Makiko HASHIGUCHI[1], Youichi YAMASAKI[1]

[1] Department of Pediatric Dentistry, Graduate School of Medical and Dental Sciences, Kagoshima University

[2] Department of Pediatric Dentistry, Kagoshima University Hospital

[3] Department of Dental Education, Graduate School of Medical and Dental Sciences, Kagoshima University

*Corresponding author:

Hideo SATO

8-35-1 Sakuragaoka, Kagoshima City, Kagoshima, 890-8544, Japan

Phone: +81-99-275-6262

Fax: +81-99-275-6268

E-mail: hideo-sato@dent.kagoshima-u.ac.jp

Running title: Lip motion and swallowing dynamics

Keywords: 3D camera, Lip, Mouth, Swallowing, Videofluoroscopy

Acknowledgment

This research was conducted with the support of JSPS Research Fund JP16K11864, JP18K09916, and the support from the 8020 public research project (adopted number 18-4-11) of the fiscal year 2018 of 8020 Foundation.

There is no conflict of interest to declare.

Abstract

Background: Clinical application of a swallowing function evaluation system that is minimally invasive and enables an objective evaluation is necessary.

Objectives: We constructed a system that can synchronize and analyze lip motion with a three-dimensional (3D) camera and swallowing dynamics with videofluoroscopy (VF) and clarified the relationship between lip movement and swallowing dynamics.

Methods: A compact 3D camera was adapted to Microsoft XBox One Kinect Sensor®. We examined Kinect's accuracy and repeatability and analyzed the highest measurement accuracy and repeatability of the distance between anguli oris muscles. The constructed system simultaneously measured, synchronized, and analyzed lip motion by Kinect and swallowing dynamics by VF. Fourteen elderly men without dysphagia were included. Barium turbid solution (5, 10, 15, and 20 mL) was used for swallowing. Measurement parameters were the 3D distance between angulus oris displacement (TDDD), swallowing quantity (SQ), oral transit time (OTT), stage transit duration (STD), pharyngeal transit time (PTT), and total swallowing duration (TSD). Statistical analyses were performed.

Results: The measurement accuracy and reproducibility were optimum within a 10° horizontal imaging angle at a 120-cm measurement distance. SQ, TDDD, OTT, STD, PTT, and TSD showed significant differences, and correlation was found between TDDD and OTT.

Conclusion: SQ affected TDDD, OTT, STD, PTT, and TSD; OTT could be predicted from TDDD. Improvement in the system and analysis method was considered to enable prediction of swallowing dynamics from body surface movement of the head and neck, including realization of measurement conditions with higher precision and reproducibility, and from lip motion.

Background

In recent years, the Japanese aging rate is 27.7%, and Japan has entered the super-aged society.[1] Elderly patients may experience a reduction in the swallowing function due to complications, such as Parkinson's disease, cerebrovascular disease, and sarcopenia. [2-4]

In Japan, pneumonia was the third leading cause of death in 2016.[5] Since a decrease in the swallowing function may cause aspiration pneumonia, leading to fatality, maintenance of the swallowing function and pneumonia prevention are urgent issues for the super-aged society. Furthermore, the Japanese baby-boom generation will be 75 years old in 2025, and it is predicted that the demand for medical care and care for the elderly will increase. The Ministry of Health, Labor and Welfare has been promoting the construction of a comprehensive community care system for the purpose of self-support at home.[6]

It is surmised that users of nursing care services and home services will increase in the future. The need for home care is expected to grow. Patients in need of home care are concerned about the decline in swallowing function due to flailing. Improvement and maintenance of quality of life include maintenance of nutritional status and swallowing function by appropriate assessment and intervention for functional decline and of diet by safe oral intake.[7]

Dysphagia is recognized in a wide age range from childhood to old age, and appropriate evaluation, diagnosis, and intervention at the acquisition, maturation, maintenance, and declining stages of dysphagia function response are extremely important. At present, the gold standard of the swallowing function test includes videofluoroscopy (VF) and videoendoscopy, and it is necessary and indispensable as an evaluation and diagnostic method of the swallowing function. However, the equipment is expensive and difficult to use at home or at remote places, requires specialized technology, and involves an invasive procedure, increasing the risk of exposure. Ayano et al.[8] reported a non-invasive method for classifying the function acquisition process into eight stages by observing and evaluating lip motion. Functional evaluation focusing on lip motion is considered very effective as a screening test for swallowing function. However, the swallowing dynamics of swallowed food cannot be assessed. Therefore, it is essential to clarify the relationship between lip motion and swallowing dynamics and establish it as a new evaluation method. It can be applied to elderly patients in home care, who are increasing in Japan, and can be easily evaluated even by medical workers of various occupations. However, at present, no study has yet reported the relationship between lip movement and swallowing dynamics. Therefore, in this study, we measured the distance between the corners of the mouth in lip motion during swallowing with a small three-dimensional (3D) camera and constructed a system to simultaneously capture, synchronize, and analyze the swallowing dynamics with VF. We also report on the effects of swallowing quantity, distance between angulus oris muscles, and swallowing dynamics.

2. Materials and Methods

2.1 Ethical considerations

This study complies with the ethical principles of the Helsinki Declaration and was approved by the Ethics Committee on Epidemiological Studies, Kagoshima University (authorization number: 170117 (704) Epidemic-Revised 2). The participants provided written informed consent.

2.2 Measurement and analysis System

In this research, the compact 3D camera was adapted to Microsoft XBox One Kinect Sensor® (Kinect). Kinect is a device that three-dimensionally acquires the shape and motion information of an object without using a marker. The software development platform was Microsoft Visual Studio 2015, and the programming language was C#. Kinect for Windows SDK v2.0 is a publicly available development software,[9] in which a software that recognizes facial shapes (Face Tracking) is used.

Face Tracking can obtain 3D positional information of facial feature points and was used to recognize the 3D position of the corners on both sides of the object.

A proprietary program was used to record the time change of the distance between the corners of the mouth. This program was used in this study. The constructed system simultaneously acquired, synchronized, and analyzed body surface information using Kinect and internal dynamic imaging by fluoroscope (TOSHIBA SXT-6600A) (Fig. 1). Other equipment used was a personal computer (PC) for measurement and analysis (G-Tune, Little Gear i310) and a high-definition multimedia interface (HDMI) recorder (AREA SD-2WAYCUP) digital timer (Apple iPad). Kinect and VF can record at 30 fps. Video-audio synchronization software ELAN (Max-Planck-Institut für Psycholinguistik) (ELAN) was used for synchronization and analysis.

2.3. Measurement accuracy and repeatability of Kinect

In this study, the created program needed to verify the measurement accuracy of the distance between the corners of the mouth. Wheat et al.[10] obtained the morphological information of a mannequin with Kinect and verified the accuracy and reproducibility. In this study, a commercial mannequin (Nanasai LUX 3V) was used. Kinect was made to recognize angulus oris muscles of both sides on the mannequin and examined the accuracy and repeatability of measuring the distance between the angulus oris muscles at the time of rotational movement.

2.3.1 Measuring method

The mannequin was placed in front of Kinect and adjusted such that the line connecting the inferior margin of the ala nasi and the inferior margin of the tragus was parallel to the floor. The mannequin was placed at a height of 100 cm above the floor where the entire upper body would be reflected by the Kinect camera. Kinect can recognize an object from a distance of 80 cm, and the closer it is, the higher the accuracy. The distances between Kinect and the mannequin were 80 cm, 100 cm, 120 cm, and 140 cm. When taking a clinical measurement, imaging was performed from the front, but movement of the measurement object was considered. The mannequin was placed on a turntable. The front of Kinect was set to 0°, and it was rotated 10°, 15°, and 20° to the left and right. Kinect records at 30 fps, and measurements can be performed for 10 seconds to obtain a sufficient amount of data. Measurements were made three times under each condition. The representative value was the average value of each.

2.3.2 Analysis

The measurement distance (MD) was the distance between the angulus oris measured by Kinect. The actual measurement distance (AMD) was the distance measured with digital calipers (Mitutoyo, SuperCaliper) on both sides of the mannequin. The difference value (Diff-V) between MD and AMD was determined. MD was calculated as an average (AVE), standard deviation (SD), and coefficient of variation (CV), and Diff-V was calculated as an AVE. The most appropriate measurement conditions for the Kinect measurement were examined from the values obtained.

2.4. Simultaneous filming of Kinect and VF

2.4.1 Participants

The participants were 14 Japanese elderly men (mean age 75.8 ± 8.4 years).

The average body mass index was 23.3 ± 2.8 kg/m^2.

Participants have observed clinical findings suspected of aspiration such as mussels during meals. They had no history of dysphagia, neurological disorders, pulmonary disease, head and neck structural damage, psychiatric disturbance, or current drug or alcohol abuse.

In addition, they conducted the modified water swallow test[11], however found no unusual findings.

No abnormalities were found in the screening test, however there was a possibility of occult aspiration, and participants who were considered to require detailed examination by videofluoroscorpy were selected as participants.

2.4.2 Samples

The samples used were 40 W/V% barium sulfate suspension. Swallowing quantity (SQ) was: 5, 10, 15, and 20 mL solutions.

2.4.3 Filming conditions

VF involved exposure to radiation; thus, the measurer was required to ensure a reduction in the radiation dose. Everyone involved in the measurement was obliged to wear a protector.

The measurer measured at a position separated by a radius of 100 cm or more from the X-ray tube bulb. Tube voltage of VF was 64.0 kV, and tube current was set to 1.5 mA. Measurements were performed under exposure to the minimum irradiation necessary. Kinect was placed 100 cm above the floor. The distance to the participant was 120 cm. Kinect was taken from the front and VF was taken from the side.

The fluoroscope was positioned such that the oral cavity and the pharynx entered the radiation field. The participant sat on a chair with both feet on the floor. The position of the headrest was adjusted such that the inferior margin of the ala nasi and the inferior margin of the tragus were parallel to the floor. The measurer instructed the participant not to move the head as much as possible. Electronic sound of the digital timer was recorded by the microphone of VF and Kinect camera. The digital timer was started immediately after the irradiation. The samples were taken in at the first electronic sound and held in the oral cavity. The sample bolus was swallowed with the second electronic sound in a single swallow. The measurement was terminated after the sample showed confirmed passage through the esophagus. Four samples were randomly swallowed, and each was repeated three times. The safety of the swallowing function was evaluated using the Penetration-Aspiration Scale (PAS). PAS assesses the severity of laryngeal invasion and aspiration in eight grades. After completion of bolus swallowing, the presence or absence of aspiration was evaluated by PAS.[12]

2.4.4 Synchronization and analysis

The data obtained by the measurement were the video and electronic sound of swallowing dynamics by VF, the video and electronic sound of Kinect recorded on the HDMI recorder, and distance between the angulus oris measured by Face Tracking. In Face Tracking, measurement time and distance between corners were recorded using spreadsheet software (Excel). The cell number corresponding to Excel was displayed at 30 fps in the screen. According to the electronic sound recorded in VF and Kinect, the film of VF and the data on the distance between the corners were synchronized by ELAN.

Daniels et al.[13] classified the bolus swallowing dynamics of a sample into oral transit time (OTT), stage transit duration (STD), pharyngeal transit time (PTT), and total swallowing duration (TSD) by VF and evaluated the swallowing time. Furthermore, they also compared instructional swallowing with free swallowing, but in this study, we only used instructional swallowing. The 3D distance between angulus oris displacement (TDDD) was defined as the difference between the maximum distance between angulus oris muscles and minimum distance between angulus oris muscles in TSD. TDDD was obtained from the recorded data, and the relationship between TDDD and swallowing dynamics was examined.

2.4.5 Quantitative analysis

The reproducibility of TDDD, OTT, STD, PTT, and TSD was verified when 5, 10, 15, and 20 mL samples were swallowed three times by each participant.

Repeated measures one-way analysis of variance (ANOVA) was performed with the swallowing variables as explanatory variables and TDDD, OTT, STD, PTT, and TSD as dependent variables. When no significant difference was observed, it was judged that the reproducibility in each participant was recognized, and the average value of three times was taken as the representative value of each parameter. Furthermore, repeated measures ANOVA used SQ as the extraneous variable; TDDD, OTT, STD, PTT, and TSD values as dependent variables; and multiple comparison tests (Tukey's test) were performed when significant differences were observed. Pearson's product-moment correlation coefficient was calculated and examined for these variables. Prior to these corrections, $\alpha = 0.05$ was considered for all analyses.

3. Results

3.1 Measurement accuracy and repeatability of Kinect

The measurement results are shown in Table 1. The measurement distance of 120 cm and the rotation angle of 10 ° showed the highest precision. The measured values were AVE = 47.707 mm, SD = ± 0.516 mm, and CV = 0.011. Diff-V was AVE = 0.467 mm. This result demonstrates an accuracy of 0.467 mm and a repeatability of 0.516 mm.

3.2 Simultaneous measurement of Kinect and VF

Video-audio synchronization software ELAN enabled the synchronization and analysis of swallowing dynamics based on the data obtained by simultaneous measurement. In this experiment, the total bolus swallowing for all participants was 168 times. PAS was 92.9% for score 1 and 7.1% for score 2. The score 2 swallow was 41.7% for 15 mL and 58.3% for 20 mL. Therefore, it was considered that there was no aspiration at all swallowing volumes. The reproducibility during swallowing was verified by repeated measures one-way ANOVA. The verification results are shown in Table 2. No significant difference was found in the TDDD, OTT, STD, PTT, and TSD of each participant. Therefore, it was judged that reproducibility at the time of swallowing existed. The representative value was the average value for each participant.

The distance between the angulus oris muscles measured by Kinect was shorter than the distance between the angulus oris muscles at rest at the time of intraoral retention of the sample (after the first electronic sound and immediately before the second electronic sound). After the swallowing instruction (the second electronic sound), the lips were stretched, and the maximum and minimum values of displacement were recorded between TSD. After completion of swallowing, a displacement of about 1 mm between the angulus oris muscles was observed, and finally the distance between the angulus oris muscles at rest was approached. SQ and representative values of TDDD, OTT, STD, PTT, and TSD were used for multivariate analysis of variance (Table 3).

There were significant differences between SQ and TDDD, OTT, TSD ($P < 0.001$, $P < 0.001$, $P = 0.005$). With the increase in SQ, TDDD, OTT, and TSD tended to increase, and STD and PTT tended to decrease. Parameters with significant differences were tested for multiple comparisons (Fig. 2). SQ and TDDD were significantly different between 5 mL and 15 mL, 5 mL and 20 mL, and 10 mL and 20 mL ($P = 0.013$, $P < 0.0001$, $P = 0.006$). SQ and OTT were significantly different between 5 mL and 15 mL and 5 mL and 20 mL ($P = 0.001$, $P < 0.001$). SQ and TSD were significantly different between 5 mL and 20 mL ($P = 0.004$).

However, there was no correlation between the PAS score and SQ, TDDD, OTT, STD, PTT, and TSD. SQ and STD, SP and PTT were not significantly different. The correlation for bivariate analysis was determined for SQ, TDDD, OTT, STD, PTT, and TSD. The correlation is shown in Table 4.

4. Discussion

4.1 Measurement accuracy and repeatability of Kinect

In this study, the measurement accuracy and reproducibility of Kinect were verified using a mannequin. As a result of verification, when the distance to the object was 120 cm and the rotation angle was within 10°, the accuracy of the distance of the mouth angle was 0.467 mm, and the reproducibility was 0.516 mm.

Body surface information (e.g. lip, buccal, neck) was mainly measured by video analysis and high-precision motion capture.[14,15] The high performance and miniaturization of equipment and the sophistication of programs have made it possible to obtain measurements with inexpensive high-performance 3D cameras. The Kinect adopted for this study was a relatively new device that has already been applied as a rehabilitation tool in the nursing and medical fields.[16,17] Kinect measurement achieved high reproducibility and accuracy by satisfying certain conditions and can be used as a measuring instrument.[18,19]

However, there is no research that has used Kinect to recognize the lips, measure the distance between the angulus oris, and examine the accuracy and reproducibility. Therefore, it was necessary to examine the measurement accuracy by the measurement system constructed for this study. The mannequin was judged as a suitable object of measurement, that is, a rigid body that Kinect could recognize.

Face Tracking recognized the participant, predicted the positional relationship between the upper body and the face from the information of the participant obtained by the Kinect camera and the pre-installed human body information, and extracted and measured feature points on the face .[9]

However, it took several seconds of motion to recognize the upper body shape of the human body, and around ten seconds on average of facial motion operation to recognize the face shape. The distance between the angulus oris of the mannequin was measured by Kinect, and the effect on the measurement results was verified by rotational movement of the mannequin. Accuracy and repeatability of Kinect involved distance and rotation angle. Diff-V was considered to be caused by the fact that the rotation angle becomes larger, the mouth angle on the rotation side deviates from the detection recognition range of the face shape of the 3D camera, and measurement of the distance between the angulus oris becomes impossible. Therefore, it was thought that high accuracy and repeatability could be obtained by minimizing and measuring the rotation angle.

Face Tracking was more accurate in detecting face position information as the Kinect camera projected the entire upper body of the participant. The distance to the participant was 80 cm or more. Measurements with high accuracy and repeatability were achieved by shortening the distance from the participant as much as possible.

When the distance was 120 cm, Kinect was at the shortest distance from which it could project the entire upper body of the mannequin. On the other hand, at 80 cm and 100 cm, Kinect could not project the entire upper body. At 140 cm, it was possible for Kinect to project the upper body; however, a decrease in measurement accuracy was recognized because of the large distance. From the results shown in Table 3, SD in MV indicates repeatability, and AVE in Diff-V indicates accuracy. It became clear that measurement of the distance between the angulus oris with the smallest error with high reproducibility was possible by setting the measurement distance to 120 cm and the rotation angle within ± 10°. Therefore, it was necessary to consider that the measurement had an accuracy of 0.467 mm and a repeatability of 0.516 mm under optimal conditions. For the measurement under other conditions (distance and angle), it was necessary to consider the repeatability and accuracy shown in Table 3.

4.2 Simultaneous measurement of Kinect and VF

There are few studies that measured the specific dose of VF. Wright et al.[20] reported an average dose of about 0.4 millisieverts. Kim et al.[21] reported an average dose of 1.23 millisieverts (corresponding to about half of upper gastrointestinal series). In this study, the setting of measurement conditions and protection against exposure enabled as much control of exposure dose as possible.

Previous reports have included measurements of the head and neck during swallowing using conventional high-precision motion capture and the swallowing dynamics with VF.

It was clear that there was interest in the motion of the body surface and the oral/pharynx during swallowing.[22,23] The relationship between head and neck motion and swallowing dynamics has been assessed as a method of confirming the end of swallowing by a swallowing sound or a sign made by the participant. However, it was impossible to visually confirm that the sample had actually passed through the pharynx.

The synchronization system constructed enabled simultaneous measure lip motion and swallowing dynamics by enhancing and downsizing the 3D camera. To the best of our knowledge, there is no report that has performed synchronization and analysis in a similar way, and our analysis is considered a novel approach. The elements of the synchronization and analysis system were facial surface information obtained from Kinect, swallowing dynamics by VF, filming of Kinect camera, and electronic sound. VF and Kinect acquired data at 30 fps. ELAN enabled data synchronization by superimposing a waveform on electronic sound extracted from both sides, and matching the time recorded by Kinect and Face Tracking data.

In this study, swallowing dynamics were classified into four types (i.e. OTT, STD, PTT, and TSD). Several studies have reported using this classification method.[13,24] There are no studies that analyzed 10 mL, 15 mL, and 20 mL of swallowing dynamics. However, Daniels et al.[13] considered swallowing dynamics with directed cued swallowing and none cued for 5 mL sample for participant. Compared to that in this study, no outlier deviation was observed for a 5-mL swallowing dynamics. Therefore, the analysis method of VF was considered appropriate.

The reproducibility of sample swallowing was verified by swallowing the prepared samples three times each. The order of swallowing the samples was randomly chosen.

There were no significant differences in each article regarding TDDD and swallowing dynamics. Reproducibility was confirmed, and it was considered that the verification method was appropriate. Therefore, the average value of each article as a representative value was considered for SQ, TDDD, and swallowing dynamics.

The average value of TDDD, OTT, and TSD increased with an increase in SQ. TDDD showed lip motion during bolus swallowing, OTT indicated oral phase swallowing dynamics, and was a voluntary movement. TSD was a mixture of voluntary movement and involuntary movement from the oral phase to the pharyngeal phase. SQ, TDDD, OTT, and TSD showed significant differences, and a positive correlation was observed between two variables. The increase in SQ was considered to contribute to the movement of an oral cavity marginal muscle and to pass the bolus to the pharynx. Dynamic oral cavity marginal muscle movements were involved in prolonging the swallowing time in the oral phase.

Therefore, the increase in TDDD and the extension of OTT and TSD were recognized. During the measurement, changes in the distance between the angulus oris muscles due to lip motion were observed. The maximum and minimum values of the displacement amount of lip expansion and contraction after swallowing instruction were recorded, and the bolus was sent to the pharynx from the video of VF. It was clear that the lip motion involved in swallowing was measured. Although a displacement of about 1 mm was recognized after the end of swallowing, it was not considered as the lip motion involved in swallowing because it was less than the value recognizable by Kinect. On the other hand, the average value of STD and PTT decreased with the increase of SQ. STD showed the transition from the oral phase to the pharyngeal phase, and PTT indicated the pharyngeal phase swallowing dynamics, and was involuntary movement. SQ, STD, and PTT had a negative correlation between the two variables.

Hasegawa et al.[25] reported that pharyngeal pressure increased with an increase in SQ. An increase in SQ was considered to increase pharyngeal swallowing pressure and reduce pharyngeal transit time. Therefore, it was considered that STD and PTT were shortened. The reason for the negative correlation between TDDD and PTT is that the large lip motion increases the force of bolus delivery to the pharynx and increases the swallowing pressure. It was clarified that SQ and TDDD were involved in the oral phase, and SQ was involved in all bolus timing. TDDD and SQ suggested that it was possible to predict the swallowing time of the oral phase. Furthermore, it was suggested that SQ was an effective parameter for estimation of bolus swallowing time.

4.3 Limitations

The participants were required to be facing the front, and when it was difficult to follow the instructions due to body movement, accurate lip movement could not be recorded. Although using multiple Kinects led to relaxation of measurement conditions, operation became complicated and the system could not be considered a highly versatile system.

The use of higher-performance 3D cameras and the creation of improved programs were pressing issues for clinical applications. In this study, only the expansion and contraction of the distance between the corners of the angulus oris muscles were recorded by Kinect. The program was improved to obtain 3D positional information of the mouth and to enable more detailed measurement of lip motion. In addition, it is necessary to conduct measurements on individuals of different sex and age and collect the data.

In this study, SQ, TDDD, OTT, STD, PTT, and TSD were used as indices that can be extracted from Kinect and VF. The relevance of these factors was derived by their comparison and by correlation of the mean values. However, swallowing is a complex movement, consisting of coordinated movements of nerves and muscles in the head and neck. In past reports, other factors involved in swallowing (e.g. muscle activity, lip pressure, tongue pressure, and viscosity) were considered, and additional measurement and analysis methods were required.[26,27] In this study, instructions of swallowing a single bolus were given. Ertekin et al.[28] reported that the maximum single bolus swallowing for adults was 20 mL. In this study, participants swallowed once for all samples. However, practically, the patient often swallows freely or continuously, and it is necessary to examine the relationship between lip motion and swallowing dynamics at that time. Furthermore, as the sample was only a liquid, measurement with samples of various forms, such as paste and solid, should be performed.

At the current stage, we only measured the displacement of the distance between the angulus oris when swallowing. In the future, we aim to develop a new measurement system that will allow us to acquire 3D positional information on both sides of the angulus oris muscles in more detail for analysis. It is also necessary to clarify the relationship between temporal changes in lip motion and swallowing dynamics. Body surface information (e.g. head, neck, and face), including that on the lips, can be used as an index for swallowing function evaluation, and further study is required. To clarify the relationship between body surface information and swallowing dynamics, measurements in various participants (e.g. age, sex difference, disease, and presence or absence of dysphagia) will be performed with synchronized analysis.

4.4 Clinical application

As a clinical application of this study, it is conceivable to compare the body surface movement of participants having no problem with swallowing function with that of participants having an abnormality by applying the screening test. If it is possible to predict the swallowing dynamics by measuring the body surface movement using Kinect, the screening test of swallowing function can be performed in home care patients only by photographing with Kinect. By detecting an operational or a functional abnormality, it is possible to determine the need for an early intervention or a close examination and to perform a close examination at a medical institution. It is possible to use this information to determine the appropriate intervention time and to prevent unnecessary exposure and invasion. In the future, we aim to build a system that can be evaluated for function.

Conclusion

SQ affected TDDD, OTT, STD, PTT, and TSD; OTT could be predicted from TDDD. Improvement in the system and analysis method was considered to enable prediction of swallowing dynamics from body surface movement of the head and neck, including realization of measurement conditions with higher precision and reproducibility, and from lip motion. In the future, it is necessary to develop a program that enables detailed analysis of the head and neck to clarify the relationship between body surface information and swallowing dynamics.

References

1. Cabinet Office. Annual Report on the Ageing Society: 2018. https://www8.cao.go.jp/kourei/english/annualreport/2018/pdf/c1-1.pdf. Accessed April 22, 2019

2. Suttrup I, Warnecke T. Dysphagia in Parkinson's Disease. Dysphagia 2016;31:24-32.

3. Azzolino D, Damanti S, Bertagnoli L, Lucchi T, Cesari M. Sarcopenia and swallowing disorders in older people. Aging Clin Exp Res 2019;31:799-805.

4. Watanabe S, Oh-Shige H, Oh-Iwa I, Miyachi H, Shimozato K, Nagao T. Reconsideration of three screening tests for dysphagia in patients with cerebrovascular disease performed by non-expert examiners. Odontology 2019. doi: 10.1007/s10266-019-00431-9. [Epub ahead of print]

5. Ministry of Health, Labour and Welfare. Annual Health, Labour and Welfare Report 2017. https://www.mhlw.go.jp/english/wp/wp-hw11/dl/01e.pdf. Accessed April 22, 2019

6. Ministry of Health, Labour and Welfare. Annual Health, Labour and Welfare Report 2017. https://www.mhlw.go.jp/english/wp/wp-hw11/dl/10e.pdf. Accessed April 22, 2019

7. Yoneyama T , Yoshida M, Ohrui T, Mukaiyama H, Okamoto H, Hoshiba K, Ihara S, Yanagisawa S, Ariumi S, Morita T, Mizuno Y, Ohsawa T, Akagawa Y, Hashimoto K, Sasaki H; Oral Care Working Group. Oral care reduces pneumonia in older patients in nursing homes. J Am Geriatr Soc 2002;50:430-433.

8. Ayano R, Tamura F, Ohtsuka Y, Mukai Y. The development of normal feeding and swallowing: Showa University study of the feeding function. Int J Orofacial Myology 2000;26:24-32.

9. Microsoft, Kinect for Windows, Developer resources, Human Interface Guidelines (PDF). https://developer.microsoft.com/en-us/windows/kinect. Accessed April 22, 2019

10. Wheat JS, Choppin S, Goyal A. Development and assessment of a Microsoft Kinect based system for imaging in the breast with three dimensions. Med Eng Phys 2014;36:732-738.

11. Tohara H, Saitoh E, Mays KA, Kuhlemeier K, Palmer JB. Three tests for predicting aspiration

without videofluorography. Dysphagia 2003;18:126-134.

12. Rosenbek JC, Robbins JA, Roecker EB, Coyle JL, Wood JL. A penetration-aspiration scale. Dysphagia 1996;11:93-98.

13. Daniels SK, Schroeder MF, DeGeorge PC, Corey DM, Rosenbek JC. Effects of verbal cue on bolus flow during swallowing. Am J Speech Lang Pathol 2007;16:140-147.

14. Hong WH, Chen HC, Yang FP, Wu CY, Chen CL, Wong AM. Speech-associated labiomandibular movement in Mandarin-speaking children with quadriplegic cerebral palsy: a kinematic study. Res Dev Disabil 2011;32:2595-2601.

15. Mishima K, Umeda H, Nakano A, Shiraishi R, Hori S, Ueyama Y. Three-dimensional intra-rater and inter-rater reliability during a posed smile using a video-based motion analyzing system. J Craniomaxillofac Surg 2014;42:428-431.

16. Nuic D, Vinti M, Karachi C, Foulon P, Van Hamme A, Welter ML. The feasibility and positive effects of a customized videogame rehabilitation programme for freezing of gait and falls in Parkinson's disease patients: a pilot study. J Neuroeng Rehabil 2018;15:31.

17. Leightley D, Yap MH. Digital Analysis of Sit-to-Stand in Masters Athletes, Healthy Old People, and Young Adults Using a Depth Sensor. Healthcare 2018;6:E21.

18. Huber ME, Seitz AL, Leeser M, Sternad D. Validity and reliability of Linect skeleton for measuring shoulder joint angles: a feasibility study. Physiotherapy 2015;101:389-393.

19. Bonnechère B, Jansen B, Salvia P, Bouzahouene H, Sholukha V, Cornelis J, Rooze M, Van Sint Jan S. Determination of the precision and accuracy of morphological measurements using the Kinect™ sensor: comparison with standard stereophotogrammetry. Ergonomics 2014;57:622-631.

20. Wright RE, Boyd CS, Workman A. Radiation doses to patients during pharyngeal videofluoroscopy. Dysphagia. 1998;13:113-115.

21. Kim HM, Choi KH, Kim TW. Patients' radiation dose during videofluoroscopic swallowing

studies according to underlying characteristics. Dysphagia 2013;28:153-158.

22. Nip ISB, Wilson EM, Kearney L. Spatial characteristics of jaw movements during chewing in children with cerebral palsy: a pilot study. Dysphagia 2018;33:33-40.

23. Perry BJ, Martino R, Yunusova Y, Plowman EK, Green JR. Lingual and jaw kinematic abnormalities precede speech and swallowing impairments in ALS. Dysphagia 2018;33:840-847.

24. Waito AA, Steele CM, Peladeau-Pigeon M, Genge A, Argov Z. A Preliminary videofluoroscopic investigation of swallowing physiology and function in individuals with oculopharyngeal muscular dystrophy (OPMD). Dysphagia 2018;33:789-802.

25. Hasegawa M, Kurose M, Okamoto K, Yamada Y, Tsujimura T, Inoue M, Sato T, Narumi T, Fujii N, Yamamura K. Differential response pattern of oropharyngeal pressure by bolus and dry swallows. Dysphagia 2018;33:83-90.

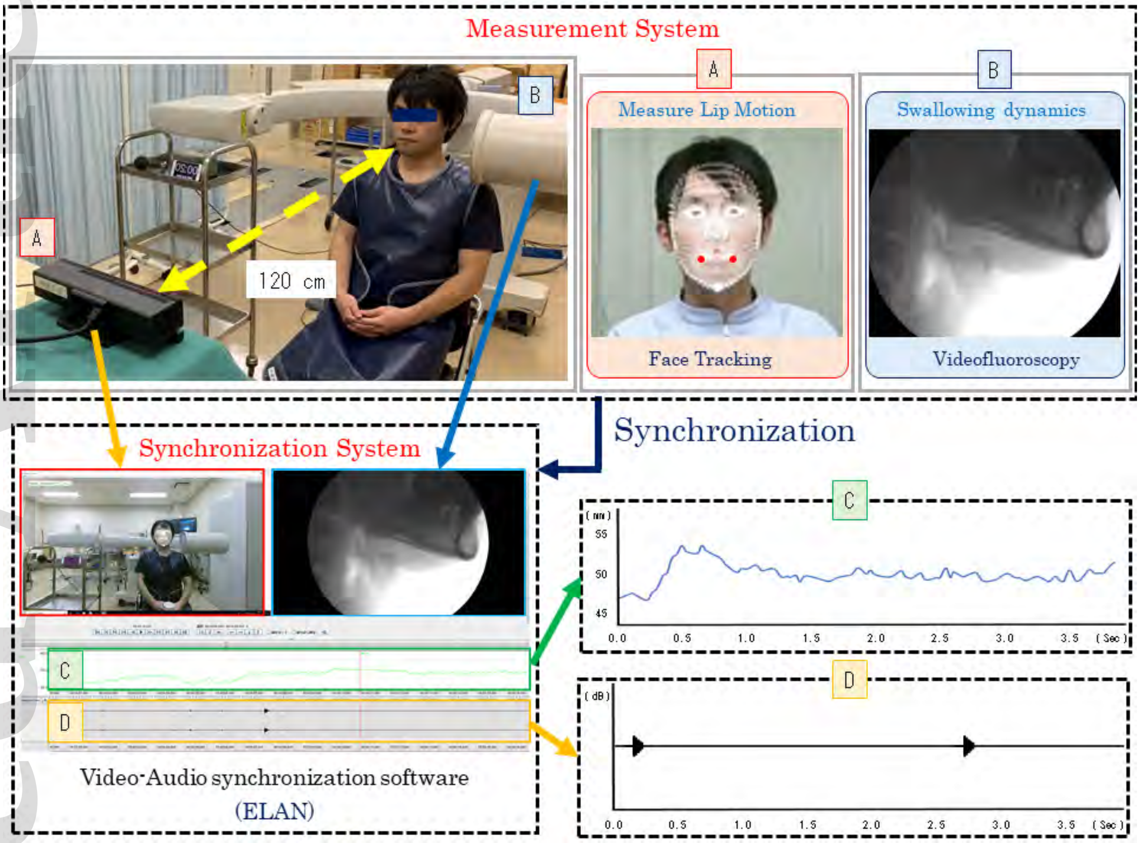
26. Rosen SP, Jones CA, McCulloch TM. Pharyngeal swallowing pressures in the base-of-tongue and hypopharynx regions identified with three-dimensional manometry. Laryngoscope 2017;127:1989-1995.

27. Okuno K, Tachimura T, Sakai T. Influences of swallowing volume and viscosity on regulation of levator veli palatini muscle activity during swallowing. J Oral Rehabil 2013;40:657-663.

28. Ertekin C, Aydoğdu I, Yűceyar N. Piecemeal deglutition and dysphagia limit in normal subjects and in patients with swallowing disorders. J Neurol Neurosurg Psychiatry 1996;61:491-496.

Figure Legends:

Fig. 1: Measurement and Synchronization System



A: Information of three-dimensional location and video, audio.

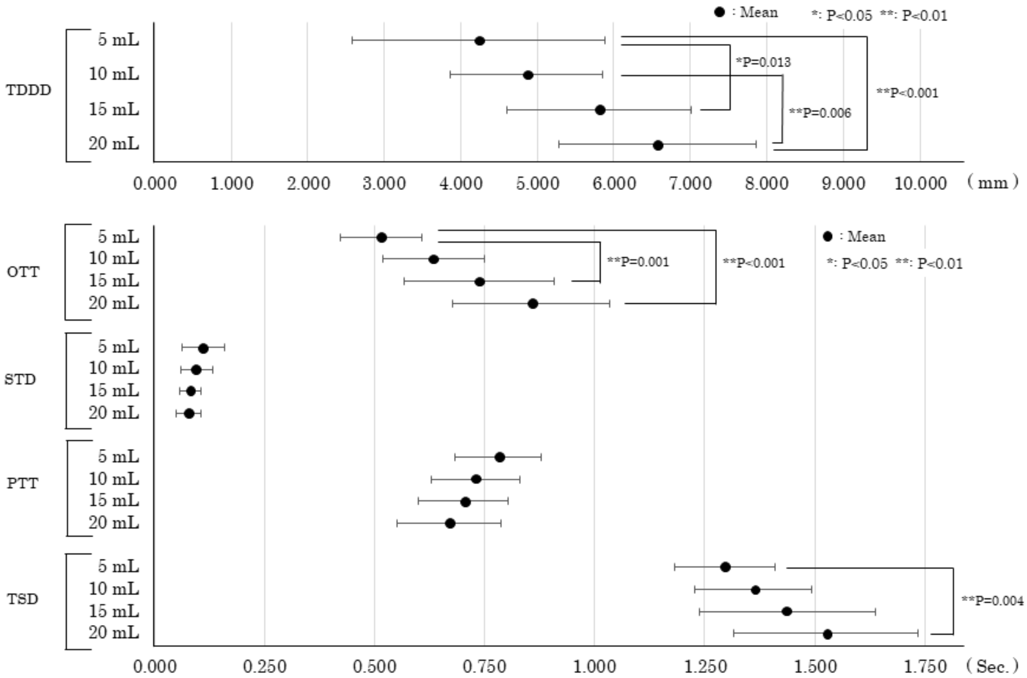
B: Recording fluoroscopy and audio with Videofluoroscopy (VF).

C: Change over time the distance between the angulus oris by Face Tracking.

D: Waveform of acquired sound.

Measurement of lip motion with Kinect and Video filming of swallowing dynamics with VF were simultaneously performed, and synchronization and analysis were performed with the video audio synchronization software ELAN.

Fig. 2: Multiple Comparison Tests of SQ, TDDD and Swallowing dynamics



Three-dimensional Distance between Angulus Oris Displacement (TDDD), Oral Transit Time (OTT), Stage Transit Duration (STD), Pharyngeal Transit Time (PTT), Total Swallowing Duration (TSD), Swallowing Quantity (SQ).

TDDD: Significantly different between 5 mL and 15 mL, 20 mL (P = 0.013, P <0.0001) and 10 mL and 20 mL (P = 0.006).

OTT: Significantly different between 5 mL and 15 mL, 20 mL (P = 0.001, P <0.001).

STD: No significant difference

PTT: No significant difference

TSD: Significantly different between 5 mL and 20 mL (P = 0.004).

Table 1: Verification result of reproducibility and accuracy by Kinect measurement

actual measurement distance (AMD), measurement distance (MD), difference value (Diff-V)

Distance (cm)	Rotation Angle (°)	AMD (mm)	MD			Diff-V
			AVE. (mm)	*S.D.* (mm)	CV	AVE. (mm)
80	10	47.24	48.11	0.57	0.012	0.91
80	15	47.24	48.42	0.64	0.013	1.22
80	20	47.24	48.29	0.70	0.014	1.09
100	10	47.24	47.12	0.58	0.012	0.58
100	15	47.24	47.84	0.65	0.014	0.71
100	20	47.24	47.75	0.73	0.015	0.73
120	10	47.24	47.71	0.52	0.011	0.47
120	15	47.24	47.30	0.73	0.015	0.59
120	20	47.24	47.68	0.81	0.017	0.71
140	10	47.24	46.40	0.63	0.014	0.90
140	15	47.24	46.73	0.82	0.017	0.79
140	20	47.24	46.97	1.09	0.023	0.93

Kinect performs measurements 30 times per second.

Since the measurement for 10 seconds was performed for each item, the statistic is 300.

The conditions which showed the highest precision were a measurement distance of 120 cm and a rotation angle of 10 °. Therefore, it was necessary to consider that measurement under optimal conditions has an accuracy of 0.467 mm and a repeatability of 0.516 mm.

Table 2: Evaluation of the safety of swallowing function by PAS

Subject	5 mL			10 mL			15 mL			20 mL		
	1st	2nd	3rd	1st	2nd	3rd	1st	2nd	3rd	1st	2nd	3rd
No.1	1	1	1	1	1	1	1	1	1	1	1	1
No.2	1	1	1	1	1	1	1	1	1	1	1	1
No.3	1	1	1	1	1	1	2	2	2	2	2	2
No.4	1	1	1	1	1	1	1	1	1	1	1	1
No.5	1	1	1	1	1	1	1	2	2	1	2	2
No.6	1	1	1	1	1	1	1	1	1	1	1	1
No.7	1	1	1	1	1	1	1	1	1	2	2	1
No.8	1	1	1	1	1	1	1	1	1	1	1	1
No.9	1	1	1	1	1	1	1	1	1	1	1	1
No.10	1	1	1	1	1	1	1	1	1	1	1	1
No.11	1	1	1	1	1	1	1	1	1	1	1	1
No.12	1	1	1	1	1	1	1	1	1	1	1	1
No.13	1	1	1	1	1	1	1	1	1	1	1	1
No.14	1	1	1	1	1	1	1	1	1	1	1	1

1: Did not invade the larynx. 2: There was laryngeal invasion; however, it was discharged without reaching the glottis. 3: There was a laryngeal invasion; however, it did not reach the glottal and was not excreted. 4: There was a laryngeal invasion reaching the glottis, however it was excreted. 5: Laryngeal penetration reaching the glottis, not excreted. 6: Bolus entered under the glottis (aspiration) and was excreted from the larynx or subglottis. 7: Bolus entered under the glottis, and coughing did not drain the airway 8: Bolus entered the subglottic area and no action to discharge was seen. The total bolus swallowing for all participants was 168 times. PAS was 92.9% for score 1 and 7.1% for score 2. The score 2 swallow was 41.7% for 15 mL and 58.3% for 20 mL. Therefore, it was considered that there was no aspiration at all swallowing volumes.

			Repeated Measures One-way ANOVA							Repeated Measures MANOVA		
Dependent variable	SQ	Statistic	Mean. (mm)			S.D. (mm)			P-Value	Mean. (mm)	S.D. (mm)	P-Value
			1st	2nd	3rd	1st	2nd	3rd				
TDDD	5 mL	14	4.23	4.39	4.09	2.00	2.05	1.21	0.663	4.23	1.65	–
	10 mL	14	4.52	5.00	5.06	1.14	1.41	1.30	0.299	4.86	1.00	–
	15 mL	14	5.60	5.98	5.84	1.38	1.86	1.25	0.682	5.81	1.20	–
	20 mL	14	6.64	7.23	6.52	1.92	2.13	1.82	0.370	6.57	1.29	–
	ALL	56	–	–	–	–	–	–	–	5.34	1.56	0.000**

Dependent variable	SQ	Statistic	Mean. (Sec.)			S.D. (Sec.)			P-Value	Mean. (Sec.)	S.D. (Sec.)	P-Value
			1st	2nd	3rd	1st	2nd	3rd				
OTT	5 mL	14	0.537	0.543	0.467	0.144	0.128	0.061	0.072	0.515	0.092	–

	10 mL	14	0.661	0.612	0.613	0.201	0.110	0.111	0.438	0.634	0.115	–
	15 mL	14	0.714	0.739	0.757	0.197	0.150	0.187	0.249	0.737	0.171	–
	20 mL	14	0.839	0.862	0.866	0.219	0.168	0.177	0.646	0.856	0.178	–
	ALL	56	–	–	–	–	–	–	–	0.685	0.189	0.000**
STD	5 mL	14	0.112	0.119	0.102	0.068	0.054	0.048	0.461	0.111	0.048	–
	10 mL	14	0.093	0.100	0.091	0.052	0.054	0.024	0.663	0.096	0.037	–
	15 mL	14	0.076	0.079	0.084	0.023	0.040	0.028	0.552	0.081	0.024	–
	20 mL	14	0.069	0.079	0.076	0.027	0.031	0.042	0.333	0.076	0.028	–
	ALL	56	–	–	–	–	–	–	–	0.910	0.037	0.050
PTT	5 mL	14	0.798	0.775	0.767	0.119	0.103	0.101	0.325	0.780	0.098	–
	10 mL	14	0.712	0.749	0.727	0.121	0.095	0.124	0.342	0.730	0.100	–
	15 mL	14	0.706	0.704	0.694	0.101	0.107	0.123	0.873	0.701	0.102	–
	20 mL	14	0.658	0.671	0.680	0.140	0.119	0.149	0.766	0.669	0.119	–
	ALL	56	–	–	–	–	–	–	–	0.720	0.110	0.051
TSD	5 mL	14	1.334	1.318	1.233	0.173	0.168	0.092	0.069	1.296	0.114	–
	10 mL	14	1.373	1.378	1.333	0.210	0.143	0.119	0.540	1.361	0.133	–
	15 mL	14	1.420	1.443	1.451	0.239	0.184	0.202	0.503	1.438	0.201	–
	20 mL	14	1.497	1.533	1.545	0.252	0.204	0.235	0.549	1.526	0.209	–
	ALL	56	–	–	–	–	–	–	–	1.405	0.186	0.005**

Table 3: Reproducibility during swallowing was verified by repeated measures one-way ANOVA and Repeated Measures MANOVA of SQ, TDDD, and Swallowing dynamics

*: $P<0.05$ **: $P<0.01$

No significant difference was found after measuring three times for each participant. Therefore, it was judged that reproducibility at the time of swallowing existed, and the representative value was taken as the average value of each participant.

There were significant differences between SQ and TDDD, OTT, TSD ($P < 0.001$, $P < 0.001$, $P = 0.005$). With the increase in SQ, TDDD, OTT and TSD tended to increase, and STD and PTT tended to decrease. Parameters with significant differences were tested for multiple comparisons.

Three-dimensional distance between angulus oris displacement (TDDD), oral transit time (OTT), stage transit duration (STD), pharyngeal transit time (PTT), total swallowing duration (TSD), swallowing quantity (SQ).

Table 4: Pearson's product-moment correlation coefficient of SQ, TDDD, and swallowing dynamics

		SQ	TDDD	OTT	STD	PTT	TSD
SQ	Pearson's correlation coefficient		0.577	0.673	-0.363	-0.368	0.465
	P-Value		0.000**	0.000**	0.006**	0.005**	0.000**
	N		56	56	56	56	56
TDDD	Pearson's correlation coefficient			0.383	-0.016	-0.336	0.191
	P-Value			0.004**	0.906	0.011*	0.158
	N			56	56	56	56
OTT	Pearson's correlation coefficient				-0.119	-0.290	0.832
	P-Value				0.382	0.030*	0.000**
	N				56	56	56
STD	Pearson's correlation coefficient					0.529	0.178
	P-Value					0.000**	0.189
	N					56	56
PTT	Pearson's correlation coefficient						0.286
	P-Value						0.033*
	N						56

TSD	Pearson's correlation coefficient	
	P-Value	
	N	

SQ was correlated with all paragraph. TDDD correlated with OTT and PTT, OTT with PTT and TSD, STD with PTT, and PTT with TSD.

Three-dimensional distance between angulus oris displacement (TDDD), oral transit time (OTT), stage transit duration (STD), pharyngeal transit time (PTT), total swallowing duration (TSD), swallowing quantity (SQ).