J Osaka Dent Univ 2023 (April); **57 (1)**: 71−81.

Reliability and interpretability of mandibular activity measurements using a three-axis accelerometer

Akane Omori[1], *Masako Kishima[2, 3], Kosuke Kashiwagi[4], Kayoko Otsuka[2] and Masataka Itoda[3]

[1]Osaka Dental University Graduate School of Health Sciences, 1-4-4 Makinohonmachi, Hirakata-shi, Osaka 573-1144, Japan, [2]Wakakusa-Tatsuma Rehabilitation Hospital,1580 Tastuma, Daito-shi, Osaka 574-0012, Japan, [3]Faculty of Health Sciences, Osaka Dental University, 1-4-4 Makinohonmachi, Hirakata-shi, Osaka 573-1144, Japan, and [4]Department of Fixed Prosthodontics and Occlusion, Osaka Dental University, 8-1 Kuzuhahanazono-cho, Hirakata-shi, Osaka 573-1121, Japan
**E-mail: kishima-m@cc.osaka-dent.ac.jp*

The purpose of this study was to examine the reliability and interpretability of mandibular activity obtained from two different signal processing methods employing a three-axis accelerometer during mastication of mastication initiation training foods using the generalizability theory. The subjects were 17 healthy dentulous individuals with a mean age of 23.2±1.3 years who had no problems with mastication or swallowing. We calculated two types of mandibular activity, namely the mandibular activity (MA) detected from the central region of the chin and corrected mandibular activity that corrected MA using the activity of the forehead. For each subject, there were two measurement days at 1-week intervals, and three repeated measurements. To examine reliability, a generalizability study (two-facet crossed design) was performed with the subjects being the object of measurement, and the date of measurement and the number of measurements as the two facets. We found that when measuring mandibular activity, the use of mandibular activity corrected for forehead motion (head sway) and the use of the mean measurement over multiple days improved reliability and interpretability. (J Osaka Dent Univ 2023; 57: 71-81)

Key words: Generalizability theory; Reliability; Measurement error; Interpretability; Three-acceleration sensor

INTRODUCTION

Masticatory movements skillfully performed under the control of the cerebrum are being quantified by measuring muscle activity of the masticatory muscles and mandibular movements.[1-6] In particular, it has been reported that the stability of the movement path and movement rhythm of mandibular movement is closely related to masticatory ability,[1] and is used as an index for judging masticatory ability. However, these measuring devices are difficult to use on people with physical or masticatory/swallowing disabilities because they require postural maintenance for a certain period of time and attachment of the device to the oral cavity. On the other hand, given their ease of attachment, there have been reports that used accelerometers to measure the mandibular movement and showed they are useful for evaluating mastication.[7-9] However, they have not yet been widely used. We focused on an accelerometer that can be easily used in clinical settings for the purpose of detecting the movement of the entire mandible during masticatory movements, defined the degree of displacement obtained from these measurements as the mandibular activity, and attempted to apply this to the rehabilitation of masticatory function.[10] Otsuka *et al.*[11] reported that it is necessary to examine conditions for measuring mandibular activity with high reliability and responsiveness in order to apply them to longitudinal measurements in patients with dysphagia due to cerebrovascular disease. By setting the conditions for measuring mandibular activity in patients with cerebrovascular disease, the effects of future rehabilitation on masticatory function can be verified with high accuracy by eliminating the influence of measurement error. This is ultimately expected to lead to implementation of rehabillitation and improvement of appropriate eating, mastication, and swallowing functions. In other words, it is necessary to quantify the measurement error and examine its interpretability in order to judge whether the change in the measurement value is simply within the range of the measurement error or is due to the rehabillitation effect. In the field of rehabilitation medicine, ways to quantify reliability, measurement error, and interpretability have been developed by applying the generalizability theory that estimates the reliability of test scores developed in psychological/educational measurements.[12, 13] In the present study, we investigated the reliability and interpretability of mandibular activity obtained from different signal processing methods using a three-axis accelerometer and generalizability theory.

MATERIALS AND METHODS

Study Population

The subjects were 17 healthy dentulous individuals without abnormalities in mastication or swallowing, specifically 9 males and 8 females with a mean age of 23.2 ± 1.3 years, who were given explanations about the purpose of the experiment in advance and who gave written consent. We considered the sample size for this study design according to the report by Walter *et al.*[14] We set the α error to 0.05, the β error to 0.2, the number of measurements to three, the desired reliability coefficient to 0.95, and the minimum acceptable reliability coefficient to 0.85. We calculated the sample size to be 15 or more. This study was conducted with the approval of the Ethics Committee of Osaka Dental University (Approval No.111211-0) and the Ethics Committee of Wakakusa−Tatsuma Rehabilitation Hospital (Approval No.22040694).

Test Food

Mastication training food (chew and swallow managing food (CSMs]) (Process Lead®; Green Tea Flavor; Otsuka Pharmaceutical Factory, Tokushima, Japan) was used as the test food. Four grams of test food was given to each subject.

Three-axis accelerometry and the signal processing of mandibular movements.

A three-axis accelerometer (UB-301 BT; A&D, Tokyo, Japan) was used for accelerometry (Fig. 1). Accelerometers were attached to the subject's forehead and to the center of the mandible with medical tape, and the acceleration from the opening of the mouth when ingesting the test food to the end of the first swallowing reflex was measured with the subjects staying in a 90-degree sitting position. The subjects were instructed to freely chew and swallow after the operator fed them the test food with a spoon. The subject movements were also video-recorded, and the same speech therapist evaluated and confirmed the chewing movements and swallowing reflexes by external observation. This procedure was repeated three times, on two measurement days each time, at 1-week intervals for each subject. The subjects were given ample time to rest in between each trial.

The acceleration signals obtained from each part were subjected to full-wave rectification and smoothening processing by removing the direct-current (DC) component. The amount of activity was determined by calculating the integrated value for each of the waveforms in the lateral, vertical, anterior and posterior directions, and adding up the components in the three directions. Two types of activity were calculated: the amount of activity detected from the central part of the chin (MA), and the corrected mandibular activity (cMA) obtained by subtracting the forehead activity from the MA. Ac-

Fig. 1 Life Recoder® three-axis accelerometer.

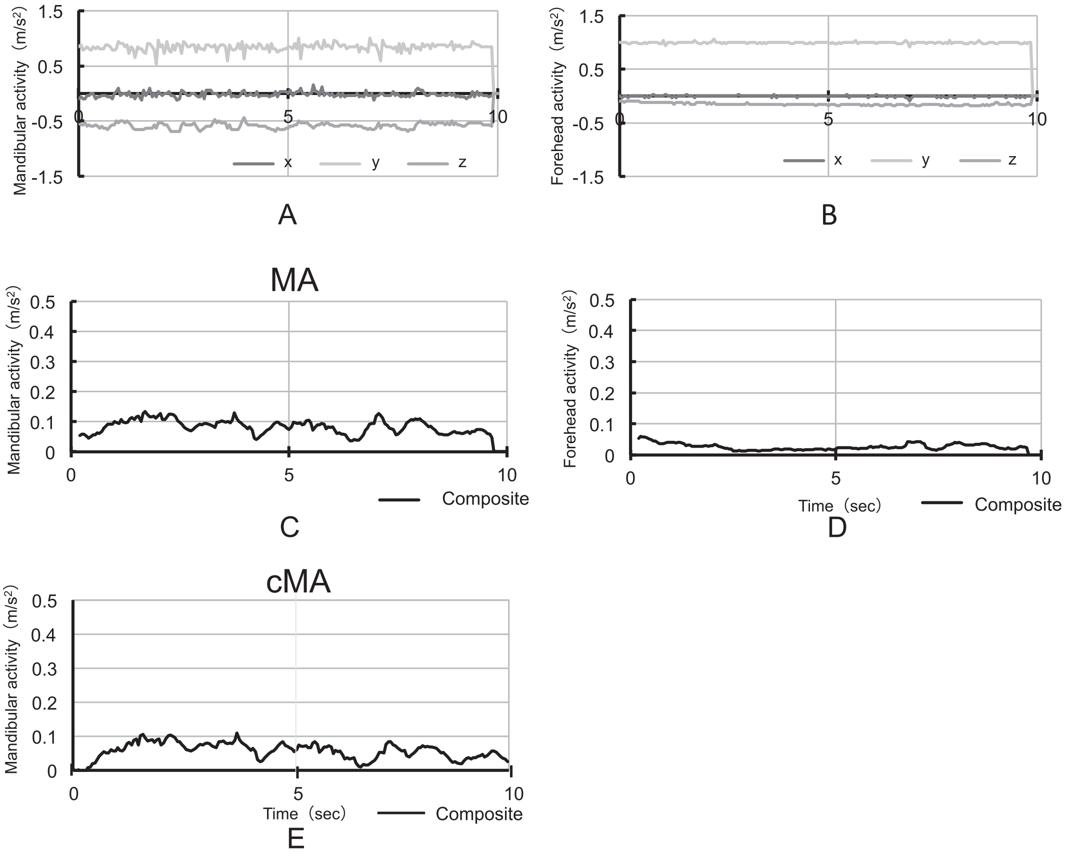

Fig. 2 Processing acceleration in healthy subjects in graph form. (A) Mandibular motion acceleration waveforms in the horizontal (x), vertical (y), and anterior-posterior (z) directions, (B) Forehead motion acceleration waveforms, (C) Mandibular activity (MA), (D) Forehead activity, and (E) Corrected mandibular activity (cMA).

celeration signal processing was performed by a person other than the one taking measurements, while blinding information on the subject and the detection site. Figure 2 shows how acceleration waveform was processed during mastication.

Statistical analysis

Normality of data

The data for each measurement condition were tested for normality using the Shapiro-Wilk test.[15]

Estimation of reliability and interpretability

Generalizability theory[12, 13] was applied. The following two analyses were carried out based on the generalizability theory.

Generalizability study

First, the presence of systematic errors in the measurement day and number of measurements were analyzed by two-factor within-subjects analysis of variance.[16, 17] Next, as a generalizability study, a two-facet crossed design[12, 13] was carried out, in which the object of measurement was the subject, the date of measurement, and the number of measurements were 2-facet. We calculated the estimated variance components for each facet, and calculated the percentage of the total variance. Figure 3 shows the Venn diagram of the two-facet crossed design. The entire Venn diagram represents all measurement variations. In this study, measurement error was defined as measurement variation caused by factors other than the subject.[12, 13]

Decision study

From the estimated variance components obtained

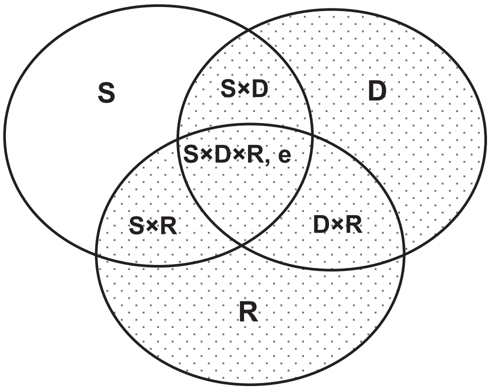

Fig. 3 Venn diagram for a two-facet, crossed design. Dotted areas are the source of measurement error. (S: Subject, D: Day, R: Repetition, e: Residual error).

Table 1 Normality of mandibular activity (MA) and corrected mandibular activity (cMA)

	Days	Repetitions	Mean±1 SD	Shapiro-Wilks test W (p)
MA	1	1	0.863±0.32	0.952 (0.486)
	1	2	0.766±0.30	0.924 (0.171)
	1	3	0.864±0.26	0.977 (0.929)
	2	1	0.778±0.30	0.925 (0.176)
	2	2	0.822±0.27	0.941 (0.336)
	2	3	0.795±0.27	0.972 (0.853)
cMA	1	1	0.375±0.91	0.948 (0.419)
	1	2	0.396±0.23	0.923 (0.167)
	1	3	0.412±0.22	0.955 (0.542)
	2	1	0.387±0.19	0.913 (0.113)
	2	2	0.376±0.18	0.897 (0.060)
	2	3	0.381±0.17	0.949 (0.435)

Table 2 Analysis of variance table for 6 measurements of mandibular activity (MA)

Source	Sum of squares	df	Mean squares	F-value	p-value
Sub	5.828	16	0.364		
Day (D)	0.028	1	0.028	0.360	0.557
Error	1.243	16	0.078		
Repetition (R)	0.023	2	0.012	0.528	0.595
Error	0.706	32	0.022		
D×R	0.102	2	0.051	2.476	0.100
Error	0.660	32	0.021		

df: Degrees of freedom

Table 3 Analysis of variance table for 6 measurements of corrected mandibular activity (cMA)

Source	Sum of squares	df	Mean squares	F-value	p-value
Sub	3.517	16	0.220		
Day (D)	0.004	1	0.004	0.296	0.594
Error	0.230	16	0.014		
Repetition (R)	0.004	2	0.002	0.752	0.480
Error	0.095	32	0.003		
D×R	0.009	2	0.004	1.349	0.274
Error	0.103	32	0.003		

df: Degrees of freedom

in the generalizability study, we calculated the index of dependability (ID) as relative reliability,[12, 13] the standard error of measurement (SEM)[18] as absolute reliability, and minimal detectable change at the statistical significance level of 5% (95% CI of minimal detectable change: MDC$_{95}$), as an index of interpretability.[19] IBM SPSS Statistics Ver. 28 (IBM, Armonk, NY, USA) and GENOVA (The American College Testing Program, Iowa City, Iowa, USA)[20] were used as statistical analysis software. Statistical analysis was performed by a different analyst than the two who were in charge of acceleration measurements and signal processing, with the information on the subject and the detection site blinded.

RESULTS

Normality of data

The mean cMA was lower than MA. The mean for MA for each measurement condition ranged from 0.77 to 0.86 m/s^2, and cMA ranged from 0.37 to 0.41 m/s^2. Normality was confirmed under all measurement conditions (Fig. 4 and Table 1).

Estimation of reliability and interpretability

Generalizability study

In both MA and cMA, the analysis of variance did not show any significant differences in factors such as the measurement day, the number of measurements, or the interaction (Tables 2 and 3). Subject variation accounted for 53.24% of the total measurement variation in MA (Table 4). When the measurement error was 100%, the ratio of each facet was 0% for the measurement day and number of measurements, and 45.7% for the interaction be-

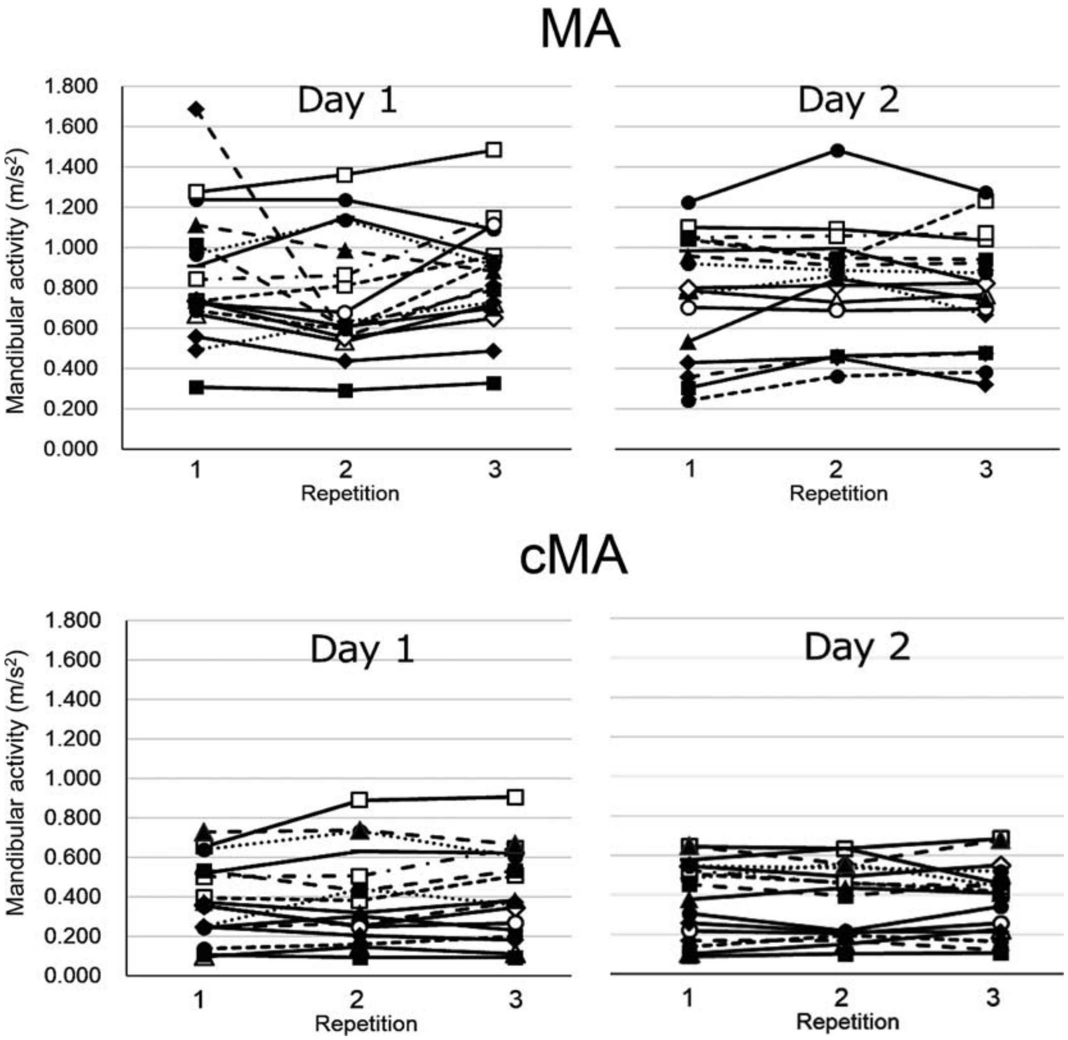

Fig. 4 MA and cMA measurements over 2 days (n＝17).

tween the subject and the measurement day (Fig. 5). Subject variation accounted for 83.07% of total measurement variation in cMA (Table 5). The ratio of each facet in the measurement error was 0% for the measurement day and number of measurements, and 53.16% for the interaction between the subject and the measurement day (Fig. 6).

Decision study

Index of dependability (ID)

Under the current measurement conditions (two measurement days and three measurements), the ID was 0.78 for MA and 0.93 for cMA. Overall, the ID for cMA was higher than that of MA. The ID for both MA and cMA increased as the number of

Table 4 Results of generalizability study for mandibular activity (MA)

Facet	n	Estimated variance components	Total variance (%)
Subject (S)	17	4.778×10^{-2}	53.235
Day (D)	2	-1.578×10^{-3}	0.000*
Repetition (R)	3	-1.190×10^{-3}	0.000*
S×D	34	1.891×10^{-2}	21.075
S×R	51	5.722×10^{-4}	0.638
D×R	6	1.779×10^{-3}	1.982
S×D×R, e	102	2.070×10^{-2}	23.070

e: Residual error, *Negative estimates were replaced by zero.

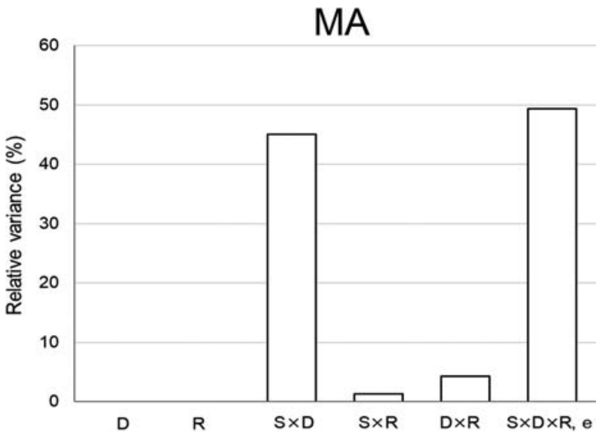

Fig. 5 Relative variance in measurement error. Estimates of variance components are normalized to error variance (100%) (D: Day, R: Repetition, S: Subject, e: Residual error).

Table 5 Results of generalizability study for corrected mandibular activity (cMA)

Facet	n	Estimated variance components	Total variance (%)
Subject (S)	17	3.428×10^{-2}	83.070
Day (D)	2	-2.200×10^{-4}	0.000*
Repetition (R)	3	-5.461×10^{-5}	0.000*
S×D	34	3.714×10^{-3}	8.999
S×R	51	-1.156×10^{-4}	0.000*
D×R	6	6.576×10^{-5}	0.159
S×D×R, e	102	3.207×10^{-3}	7.771

e: Residual error, *Negative estimates were replaced by zero.

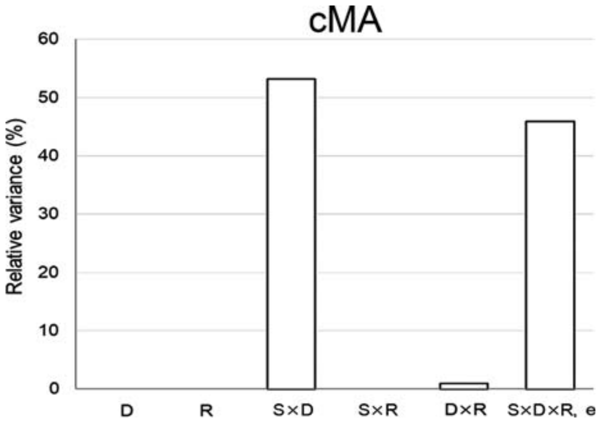

Fig. 6 Relative variance in measurement error. Estimates of variance components are normalized to error variance (100%) (D: Day, R: Repetition, S: Subject, e: Residual error).

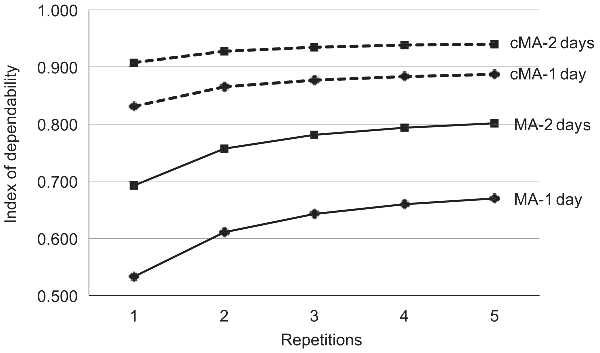

Fig. 7 Index of dependability for MA and cMA.

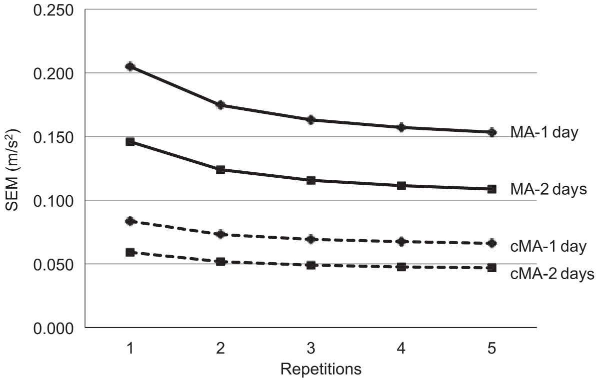

Fig. 8 Standard error of measurement for MA and cMA.

measurement days and the number of measurements increased. Increasing the number of measurement days increased the ID more than increasing the number of measurements per day (Fig. 7).

Standard error of measurement (SEM)

The SEM under these measurement conditions was 0.12 m/s^2 for MA and 0.05 m/s^2 for cMA. Overall, SEM of cMA was lower than that of MA. The SEM for both MA and cMA decreased as the number of measurement days and number of measurements increased. Increasing the number of measurement days decreased the SEM more than increasing the number of measurements per day (Fig. 8).

Interpretability (MDC$_{95}$)

The MDC$_{95}$ under these measurement conditions was 0.32 m/s^2 for MA and 0.14 m/s^2 for cMA. Overall, the MDC$_{95}$ for cMA was lower than that for MA. The MDC$_{95}$ for both MA and cMA decreased as the

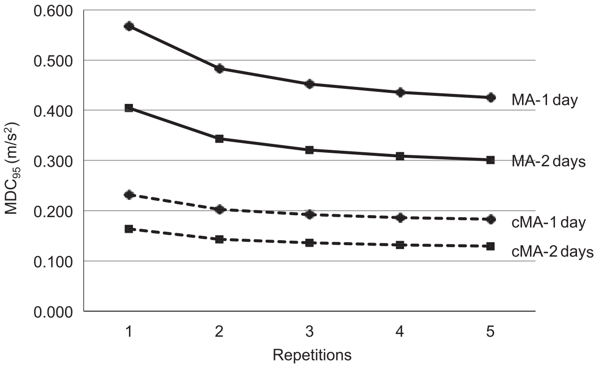

Fig. 9 95% CI of minimal detectable change for MA and cMA.

number of measurement days and the number of measurements increased. Increasing the number of measurement days decreased the MDC_{95} more than increasing the number of measurements per day (Fig. 9).

DISCUSSION

Mastication and swallowing movements receive sensory input from the periphery nerves and are triggered by motor output with the central pattern generator (CPG). In cerebrovascular disease, the necessary sensory input and motor output in each oral organ and the brain function that integrates them are likely to be impaired. Although some patients with dysphagia due to cerebrovascular disease have not only swallowing problems, but also mastication disorders due to decreased or impaired oral function, there are no reports to clarify this. In addition, since masticatory ability affects the risk of aspiration, suffocation, malnutrition, and selection of food,[21] there is an urgent need to objectively evaluate masticatory movements in dysphagic patients and clarify the aspects of mastication. For smooth swallowing, we believe that rehabillitation of masticatory function is very important. The ultimate goal of this study was to build a mandibular motion measurement system for evaluation of rehabillitation effectiveness from mastication to swallowing function using a three-axis accelerometer that can be easily used at rehabillitation centers, and has a low burden on the patient.

In general, measurements form the basis of medical research and clinical practice.[19, 23, 24] Quality evaluation of measurement instruments is very important. The three main characteristics of high validity, reliability and responsiveness, are required.[22, 23] Reliability is defined as the degree to which there is no measurement error, validity is the degree to which the measurement instrument actually measures the construct it is intended to measure, and responsiveness is defined as the ability of the measurement instrument to detect changes in the longitudinal measurement of the construct of the object of measurement.[19, 22, 23] A prerequisite for obtaining high validity and responsive measurements is to first establish higher reliability.[19] In the past, although interpretability was treated as "internal responsiveness" as a concept included in responsiveness, Mokkink *et al.*[22, 23] re-defined interpretability as "The degree to which a quantitative score or change in score of a measurement instrument can be assigned a qualitative meaning, that is, a clinical or commonly understood meaning." It shows what the score of the measurement instrument means. Mokkink *et al.*[22, 23] noted that often, not enough attention is paid to the interpretability of measurement scores in clinical practice. Correct interpretation of scores is considered a prerequisite for the appropriate use of clinical measurement instruments.

We investigated reliability and interpretability of mandibular activity measurements using a three-axis accelerometer during mastication of starter training food under different measurement conditions.

Study population

Since the score obtained from the measurement instrument is affected by the population of the object of measurement and the measurement conditions, the quality evaluation of the measurement instrument currently focuses not on the "measurement instrument itself" but on the "score generated by the measurement instrument."[22, 23] The ultimate goal of this study was to measure mandibular activity in patients with dysphagia. In order to understand the basic characteristics of the measurement instrument, we chose healthy dentulous individuals as subjects. According to Nunnally and Bernstein,[24] a reliability coefficient of 0.70 is often considered acceptable when the measurement method is used for intergroup measurement (cross-sectional measurement). However, a coefficient of 0.95 is needed for measurement of individual patients (longitudinal measurement). In this study, using the calculation formula of Walter *et al.,*[14] we set the α error to 0.05, the β error to 0.2, the number of measurements to three, the desirable ID to 0.95, and the minimum acceptable ID to 0.85, We determined that a sample size of 15 or more was needed.

Test Food

In this study, we used Process Lead® (Otsuka Pharmaceutical), which is commercially available as food for chewing and swallowing training. This food corresponds to Code 3 or 4 of the 2013 Dysphagia-Adjusted Diet Classification of the Japanese Society of Dysphagia Rehabilitation. It has physical properties that allow smooth movements leading to chewing, food mass formation and swallowing. In dysphagia rehabilitation, this food is considered appropriate for mastication and swallowing, which starts with ingestion training with jelly and paste foods, and then shifts to a form that requires mastication.[25] After evaluating how often food enters the larynx in patients with dysphagia, this food was deemed safe as the paste food. The ultimate goal of this study was to measure the mandibular movement of dysphagia patients with mastication and bolus formation disorders. Based on previous research,[26] we decided to use 4 g of the test food for this study.

Three-axis accelerometry and signal processing of mandibular movements

The accelerometer we used was a wristband-type activity meter that has a length 21 mm, a width of 39 mm, a height of 15.5 mm, and a weight of approximately 22 g. It can be attached to the wrist and can measure human movement in three axial directions. In this study, we removed the wristband from the main unit of the accelerometer. By using a sampling rate of 20 Hz, and by transferring the data to the PC via the cradle, this device can evaluate the daily behavior and the amount of exercise of the subject based on their exercise intensity and frequency. Mastication is a rhythmic movement performed by coordinated activities of masticatory muscles, head and neck muscles, tongue muscles, and other oral soft tissues.[27] Recent research on measuring instruments has shown that it is now possible to measure small movements of the body, allowing verification of small movements of the maxilla and head that can be coordinated with functional movements of the mandible.[28, 29] Based on reports that coordinated movement of the head accompanies mandibular movement during mastication,[30, 31] in this study, we attached accelerometers to the subject's forehead and the center of the chin.

Statistical analysis

The generalizability theory applied in this study was developed by Cronbach *et al.*[12] and has been used to evaluate the reliability of tests in the field of psychological and educational measurements. This theory separates and estimates variation (errors) related to various conditions in measurement (called facets in generalizability theory) and analyzes reliability. The analysis is divided into two parts. First, reliability experiments are performed as a generalizability study, and the variance components for each set facet are calculated. In the generalizability study, the ratio (influence) of each facet to the error is calculated. A decision study simulates the reliability when the measurement conditions are changed.[12, 13] In this generalizability study, we set the measurement day and the number of measurements as facets in mandibular activity measurements using different signal processing methods, and analyzed their influence on reliability. In addition, in this study, the presence or absence of systematic errors was examined by repeated measures analysis of variance before the generalizability study. Weir[32] has recommended that the reliability test should first involve an analysis of variance and evaluation of the F-ratio of the factors in the measurement conditions to check for the presence of systematic errors.

In the decision study, we sought indices of two types of reliability (relative and absolute reliability) and interpretability when the measurement day and number of measurements were changed. The ID of relative reliability means the proportion of the subject's variation to the total measurement variation. It is calculated as a value from 0 to 1. The closer the ID is to 1, the higher the reliability.[12, 13] Absolute reliability is synonymous with measurement error, and the SEM is one of its representative indices. It expresses one standard deviation of the measurement error that changes each time the measurement is performed, and quantifies the accuracy of the individual score in the measurement.[18] MDC_{95} is an indicator of interpretability, and is the minimum statistically significant difference between two independently obtained measurements at a statistical significance level of 5%. If the difference between the two measurements is large compared to MDC_{95}, it indicates a change that exceeds the measurement error (true rehabilitation effect).[19] In this study, SEM and MDC_{95} are expressed in units of acceleration (m/s^2).

Reliability and interpretability

For both MA and cMA, the analysis of variance did not show significant differences for such factors as measurement day and the number of measurements, or interaction. There was no systematic error in the mandibular activity measurements performed three times over two days in this study. This indicates that there was no apparent increase or decrease in measured values due to fatigue of the subject during the three trials each day, and that the measured values on the first and second days did not show a systematic upward or downward trend. We found that the measurement variation of mandibular activity measurement by the accelerometer was due to random error. In the generalizability study, although the subject-dependent variation accounted for 50% of the total measurement variation in MA, it increased to 80% in cMA. In general, the definition of reliability is the proportion of the variance between subjects (signal) and the total variance of the data (signal + noise).[33, 34] It was found that cMA has a greater ability to discriminate the subject's measured values than MA. In MA and cMA, the proportion of each facet in the measurement error is 0% for the measurement day and the number of measurements, and about 50% for the interaction between the subject and the measurement day. As such there was no significant difference between the two. This indicates that there are subjects whose measured values are stable on both days and subjects whose measured values fluctuate on different days.

In the decision study, ID, which is relative reliability, was higher for cMA than MA in general. In addition, SEM and MDC_{95}, which show absolute reliability, were lower for cMA than MA in general. Because cMA is a measurement that excludes the effects of head and neck movement from the measurement of the acceleration detected in the chin, this method provided highly reliable (low measurement error) mandibular activity measurements. In general, it is recommended to use the mean value of repeated measurements to improve the reliability of measurement (reducing measurement error).[19, 23, 32] This method is effective in reducing random errors, and can be expected to improve reliability (reducing measurement errors), especially when it is possible to repeatedly measure the largest cause of variation. For both MA and cMA, increasing the number of measurement days increased ID more than just increasing the number of measurements. In terms of ID, when using the cMA of this study for comparison between groups, even the measured value obtained from one measurement on one measurement day exceeded 0.70, indicating its utility. However, when considering its use for longitudinal measurements, the mean value obtained from 5 measurements (a total of 10 measurements over 2 days) did not reach an ID of 0.95. In addition, the ratio of MDC_{95} to the mean measurement under the same conditions decreased to only 33.41%. This is probably due to the fact that the analysis interval in this study was set from the opening of the mouth when ingesting the test food to the end of the first swallowing reflex. In the future, in order to apply cMA to individual longitudinal measurements, it is necessary to examine the measurement conditions by splitting up the analysis interval.

Based on the above, we found that reliability and interpretability were improved by using mandibular activity corrected for head sway and by using the mean result from multiple measurement days in the mandibular activity measurements.

Conflicts of interest

There are no conflicts of interest for the authors of this article.

Acknowledgements

The authors are grateful to the study participants for volunteering for this study. We would like To express my appreciation Associate Professor Takashi Yamanaka, department of Oral Health Sciences, Faculty of Health Science, Osaka Dental University and Lecturer Yuto Konishi, department of Radiological Technology, Faculty of Health Science, Butsuryo College of Osaka, for their assistance in this study.

REFERENCES

1. Shiga H, Kobayashi Y. An objective evaluation of masticatory function by analysis of masticatory movements. *J Prosthodont Res* 1990; **34**: 1112-1126.
2. Kumagai H, Hirano Y, Minakuchi S, Sekita T, Suzuki T, Kobayashi K, Nagao M. The movement of the chin during mastication. *Jpn Prosthodont Soc* 1995; **39**: 1132-1141. (Japanese)
3. Soboļeva U, Lauriņa L, Slaidiņa A. The masticatory system−−an overview. *Stomatologija* 2005; **7**: 77-80.
4. Hirai T, Matai N, Anzai T, Nagao M, Hayashi T. Masticatory rhythm analyzing method and clinical application. *J Med Dent Sci* 1985; **32**: 57-65.
5. Kobayashi Y, Shiga H, Yokoyama M, Arakawa I, Nakajima K. Differences in masticatory function of subjects with different closing path. *J Prosthodont Res* 2009; **53**: 142-145.
6. Iwanami Y, Shiga H, Kobayashi Y. Mandibular movement and masticatory efficiency when compared between the two most representative patterns of masticatory path. *Ann Jpn Prosthodont Soc* 2000; **44**: 147-155.
7. Grant AA. Some aspects of mandibular movement: acceleration and horizontal distortion. *Ann Acad Med Singap* 1986; **15**: 305-310.
8. Yamabe Y, Fujii H. An experimental study on body motion during mandibular movement with accelerometers. *J Jpn Soc Stomiatognat Funct* 1997; **3**: 121-129.
9. Sekita T, Hasegawa S, Hayakawa I. Fundamental study on the development of a measuring device using an accelerometer. *J Stomatol Soc Jpn* 2002; **69**: 251-257.
10. Otsuka K, Iwashita A, Konishi Y, Tanaka M. Quantification of mandibular movement during mastication using a 3-axis acceleration sensor. *J Yamato Univ* 2020; **6**: 3-7. (Japanese)
11. Otsuka K, Konishi Y, Kishima M, Imai A, Tanaka J, Itoda M, Kashiwagi K. Measurement of mandibular activity during mastication using a three-axis accelerometer in patients with dysphagia after a stroke. *J Osaka Dent Univ* 2022; **56**: 87-98.
12. Cronbach LJ, Gleser GC, Nanda H, Rajaratnam N. The dependability of behavioral measurements: theory of generalizability for scores and profiles. New York: Johon Wilely & Sons, 1972: 1-33, 161-188.
13. Brennan, RL. Generalizability theory. New York: Springer-Verlag New York Inc, 2001: 1-20, 53-139.
14. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. *Stat Med* 1998; **17**: 101-110.
15. Shapiro SS, Wilk MB. An analysis of variance test for normality (complete samples). *Biometrika* 1965; **52**: 591-611.
16. Kirk RE. Experimental design: procedures for the behavioral sciences. 4th ed. Los Angeles. Sage Publications, 2013.
17. Braver SL, MacKinnon DP, Page M. Levine's guide to SPSS for analysis of variance. Psychology Press, 2003.
18. Harvill LM. Standard error of measurement. *Educ Meas* 1991; **10**: 33-41.
19. Portney LG. Foundations of clinical research: applications to evidence-based practice, 4th ed. Philadelphia: F. A. Davis, 2020 :115-126, 486-508.
20. Crick JE, Brennan RL. Manual for GENOVA: a generalized analysis of variance system, Iowa: American College Testing Program. 1983.
21. Kikutani T. Determination for modified food for the elderly patients with motor masticatory disorders. *Jpn Prosthodont Soc* 2016; **8** :126-131. (Japanese)
22. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, de Vet HC. International consensus on taxonomy, terminology, and definitions of measurement properties for health−related patient− reported outcomes: results of the COSMIN study. *J Clinical Epidemiology* 2010; **63**: 737-45.
23. de Vet HC, Terwee CB, Mokkink LB, Knol DL. Measurement in medicine: a practical guide. Cambridge: Cambridge University Press, 2010.
24. Nunnally JC, Bernstein IH. Psychometric theory. 3rd ed. New York: McGraw-Hill, 1994.
25. Nakagawa K, Matsuo K, Shibata S, Inamoto Y, Ito Y, Abe K, Ishibashi N, Fujii W, Saitoh E. Efficacy of a novel training food based on the process model of feeding for mastication and swallowing −a preliminary study in elderly individuals living at a residential facility−. *Jpn J Comprehensive Rehabili Science* 2014 ; **5**: 72-78. (Japanese)
26. Tohara T, Kikutani T, Yajima Y, Igarashi K, Tanaka Y, Tamura F. An attempt to evaluate masticatory performance using commercially sold masticatory training food. *Jpn J Dysphagia Rehabil* 2017; **21**: 3-10. (Japanese)
27. Lund JP, Enomoto S. The generation of mastication by the mammalian central nervous system. In: Cohen AH, Rossignol S, Grillner S, eds. Neural control of rhythmic movements in vertebrates. New York: Wiley Interscience, 1988: 41-72.
28. Matsuyama T. The head movement cooperated with mandibular tapping movement. *J Jpn Prosthodont Soc* 1996; **40**: 535-543. (Japanese)
29. Eriksson PO, Häggman-Hebrikson B, Nordh E, Zafar H. Coordinated mandibular and head-neck movements during rhythmic jaw activities in man. *J Dent Res* 2000; **79**: 1378-1384.
30. Häggman-Henrikson B, Eriksson PO. Head movements during chewing: relation to size and texture of bolus. *J Dent Res* 2004; **83**: 864-868.
31. Shinya A, Sato T, Hisanaga R, Miho O, Nomoto S. Time

course analysis of influence of food hardness on head posture and pitching of head during masticatory movement. *Bull Tokyo Dent Coll* 2013; **54**: 73-80.

32. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res* 2005; **19**: 231-40.
33. Keating J, Thomas M. Unreliable inferences from reliable measurements. *Aust J Physiother* 1998; **44**: 5-10.
34. Stratford P. Reliability : consistency or differentiating between subjects? [Letter]. *Phys Ther* 1989; **69**: 299-300.