ORIGINAL ARTICLE

ORAL REHABILITATION

WILEY

Reliability of the mixing ability test testing masticatory performance in patients with head and neck cancer and healthy controls

Jorine A. Vermaire[1] | Florine M. Weinberg[2] | Cornelis P. J. Raaijmakers[1] | Irma M. Verdonck-de Leeuw[3,4] | Chris H. J. Terhaard[1] | Caroline M. Speksnijder[2,5]

[1]Imaging Division, Department of Radiation Oncology, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

[2]Department of Oral and Maxillofacial Surgery and Special Dental Care, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

[3]Department of Otolaryngology-Head and Neck Surgery and Cancer Center Amsterdam, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

[4]Department of Clinical, Neuro- and Developmental Psychology, Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

[5]Department of Head and Neck Surgical Oncology, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

Correspondence
Caroline M. Department of Head and Neck Surgical Oncology, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands.
Email: C.M.Speksnijder@umcutrecht.nl

Funding information
KWF Kankerbestrijding, Grant/Award Number: 2013-5930

Abstract

Background: Treatment of patients with head and neck cancer can result in disrupted mastication. To measure masticatory performance in people with compromised mastication, the mixing ability test (MAT) was developed.

Objective: In this study, the reliability of the MAT was evaluated in patients with head and neck cancer and healthy controls.

Methods: Thirty-four patients with head and neck cancer and 42 healthy controls performed the MAT twice on the same day. To assess reliability, the intra-class correlation coefficient ($ICC_{2,1}$), standard error of measurement (SEM), smallest detectable change (SDC) and limits of agreement (LoA) were calculated.

Results: A good (ICC = 0.886) and moderate correlation (ICC = 0.525) were found for patients and healthy controls, respectively. Patients had a worse mixing ability (mean = 19.12, SD = 4.56) in comparison with healthy controls (mean = 16.42, SD = 2.04). The SEM was 0.76 in patients and 1.45 in healthy controls, with a SDC of 2.12 and 4.02, respectively. The LoA was −4.46 to 4.42 in patients and −3.65 to 4.59 in healthy controls.

Conclusion: The MAT has a good reliability in patients with head and neck cancer and a moderate reliability in healthy controls.

KEYWORDS
chewing, head and neck cancer, mastication, mixing ability test, oral health, reliability

1 | BACKGROUND

Mastication is a learned automatic complex process involving interaction of hard and soft tissues in order to grind a food bolus prior to swallowing.[1] It involves several nerves, muscles and connective tissue structures, such as the tongue, masseter, temporalis and pterygoid muscles.[2] Many factors can affect efficient mastication, such as maximal biting force, maximal mouth opening, tongue function,

tongue force and number of occlusal units.[3] Loss of teeth, cavities, inadequate restorations, malocclusion and periodontal diseases can negatively affect the chewing function.[4,5]

In patients with head and neck cancer (HNC), mastication may be disrupted due to HNC or cancer treatment, which can result in affected motor and oral functioning. Because of this compromised mastication, tougher foods are more difficult to process because they require a higher muscle force and more chewing cycles.[6] Therefore, some patients switch their diet to softer foods, because the muscle force needed to break down food is too high.[7]

Treatment for HNC may consist of radiation therapy (RT), chemoradiation therapy (CHT), surgery or a combination of these. Early-stage cancers are usually treated with either surgery or radiotherapy, while locally advanced cancers are treated with surgery followed by adjuvant radiation or chemoradiation therapy.[8] Radiation therapy damages all cells receiving a radiation dose, including normal tissue cells surrounding the tumour. This damage to normal tissues can result in acute or long-term damage. Acute effects include pain, mucositis, dermatitis, decreased saliva production or oedema. Long-term damage can consist of dysphagia, fibrosis, oedema, ulcers, vascular toxicity or osteoradionecrosis.[1,9,10] One of the most feared complications is osteoradionecrosis of the jaws. In order to prevent this, teeth can be extracted pre-treatment, or hyperbaric oxygen treatment post-treatment can be prescribed. However, this cannot always be achieved, causing serious deterioration in dental health.[1] Chemotherapy may show additional toxicities, for example by the enhancement of radiation-induced fibrosis of the muscles, oedema or neuropathy.[9]

Surgery may require wide resections of one or multiple subsites, including tongue, floor of mouth or lower gingiva.[11] Surgery may be combined with neck dissection or reconstruction of the tumour site by a tissue transfer. Impairments after surgery depend on volume of resection, tumour site and type of reconstruction. Patients may develop defects on soft tissues, bone or dentition, which can lead to functional deficits in mastication, swallowing and speech. For example, tongue resection compromises lingual mobility and strength, and dental and mandibular surgery affects mastication.[1] Although survival rates have improved over time, morbidity remains high.[1] In order to determine the influence of HNC treatment on oral function, it is important to evaluate the masticatory performance in these patients.[12]

In previous research, different tests have been developed to measure mastication, such as comminution methods, sieve and optical scanning methods, gummy jelly as test food, and mixing ability methods.[13] Construct validity was positive in one method measuring mixing ability to test oral function: the mixing ability test (MAT).[12] The MAT was specifically developed for patients with HNC,[7,12,14] consisting of a relatively soft material (wax), to make sure patients with compromised mastication would still be able to perform this test. This test has proven to be sensitive in measuring mastication in adults with dental deficits and children with cerebral palsy.[15,16] However, reliability of this test has not been evaluated in patients with HNC yet.

The aim of this study was therefore to provide insight into the reliability of the MAT, by investigating test-retest reproducibility, standard error of measurement, smallest detectable change and limits of agreement in patients with HNC. In order to make a comparison between patients and healthy subjects, reliability of the MAT was tested in healthy controls as well.

2 | METHODS

Patients were included when they were 18 years or older, were diagnosed with oral, oropharyngeal, hypopharyngeal, laryngeal or unknown primary HNC and were treated with a curative intent at the University Medical Center Utrecht (UMCU), the Netherlands, between September 2016 and June 2018. Patients with recurrent or residual disease, cognitive impairments, and patients having trouble understanding or reading the Dutch language were excluded. Healthy controls were recruited through a poster at the outpatient clinic, between November 2018 and February 2019. Healthy controls were included when they were 18 years or older. The study protocol for patients with HNC was approved by the Medical Ethics Committee of the Netherlands (NL45051.029.13), which is part of the NET-QUBIC research.[17] The study protocol for healthy controls was approved by the Medical Ethics Committee of the UMCU (18/701). Patient data about age, sex, tumour stage,[18] tumour location, treatment, number of teeth (maximal 32 teeth) and number of occlusal units (maximal 16 units)[19] were collected. In addition, data about age, sex, number of teeth and number of occlusal units were collected for healthy controls. All participants signed informed consent before participation.

2.1 | Mixing ability test

The mixing ability test (MAT) consists of two layers of wax, with the colours red and blue (Plasticine modelling wax, non-toxic DIN EN-71, art. nos. crimson 52801 and blue 52809, Stockmar, Kalten Kirchen, Germany).[7,12,14,20] The total thickness is 3 mm, with a diameter of 30 mm. The outcome variable is called the Mixing Ability Index (MAI), and ranges between 5 and 30, where a lower MAI score implies a better mixed tablet and better masticatory performance. A subject was asked to chew on this tablet 20 times in order to mix the two colours. The tablet is then flattened, pressed to a thickness of 2 mm and scanned on both sides using a high-quality scanner (Epson® V750). The scanned images are then processed using Adobe Photoshop CS3 extended (Adobe). The histograms of both sides of the flattened and scanned wax tablet are added to obtain red and blue intensity distributions. The spread of the colour intensities is measured.[12] Subjects were instructed to chew 20 times on two different tablets in order to test reliability. The interim period between the two tests was approximately 2 hours for patients with HNC and 30 minutes for healthy controls, with the same testing conditions for all participants.

2.2 | Statistical analyses

Test-retest reproducibility of the MAT was tested by a two-way random, single measurement, absolute agreement, intra-class correlation coefficient ($ICC_{2,1}$), calculated as $\frac{MS_R - MS_E}{MS_R + (k-1)MS_E + \frac{k}{n}(MS_C - MS_E)}$, in which MS_R = mean square of rows; MS_E = mean square for error; MS_C = mean square for columns; k = number of measurements; and n = number of subjects. Cut-off points for the ICC were chosen as poor (<0.5), moderate (0.5-0.75), good (0.75-0.90) and excellent (>0.90).[21,22] Standard error of measurement (SEM) was calculated as $SEM = SD * \sqrt{(1 - ICC)}$.[23] For the SD, the standard deviation of the difference between the two MATs was used. Standard error of measurement per cent change was calculated as

$SEM\% = (SEM/\overline{X}) \times 100$, in which $\overline{X}$ = the mean of all measurements of test and retest. Smallest detectable change (SDC) was calculated as $SDC = 1.96 \times \sqrt{2} \times SEM$.[24,25] The SDC per cent change was

calculated as $SDC\% = (SDC/\overline{X}) \times 100$, in which $\overline{X}$ = the mean of all measurements of test and retest. In order to check for systematic bias, variability and agreement, Bland-Altman plots were constructed by plotting the test-retest difference versus the mean value of the test and retest. Agreement between test and retest was summarised using the mean difference and SD of the difference, and the 95% limits of agreement (LoA) were calculated as $LoA = Mean \pm 1.96 \times SD$.[26]

A power analysis was conducted, in which an ICC of at least 0.7 was expected. A P_1 value of .9 was chosen; therefore, the sample size had to be at least 18.4.[27] In addition, a comparable study in children with cerebral palsy showed a sample size of 25-30 patients[15]; therefore, it was chosen to include at least 30 subjects.

Data were tested for normality using a Shapiro-Wilk test. Because data were not normally distributed, a Wilcoxon signed-rank test was conducted to examine differences between test and retest for both patients with HNC and healthy controls, and a Kruskal-Wallis test was conducted to examine differences in MAT scores according to age and sex. A Mann-Whitney U test was run to test for differences between patients and healthy controls regarding age, number of teeth, and number of occlusal units, and a chi-squared test was run to test for differences regarding sex. All analyses were performed using Statistical Package for the Social Sciences (SPSS) version 25. A P-value below .05 was considered statistically significant.

3 | RESULTS

Thirty-four patients with HNC and 42 healthy controls performed the MAT twice within a time interval of 2 hours. In the patient group, eleven patients performed the test before HNC treatment, six patients 3 months after treatment, five patients 6 months after treatment, five patients 12 months after treatment and seven patients 24 months after treatment. No missing data were reported. In Table 1, subject characteristics are depicted for both patients with HNC and healthy controls.

As seen in Table 2, the ICC for patients with HNC (0.886; 95% CI = 0.784-0.942) showed a good correlation between the test and retest, and a moderate correlation for healthy controls (ICC = 0.525; 95% CI = 0.272-0.712). The SEM was 0.76 (4.0%) in patients and 1.45 (9.0%) in healthy controls, with an SDC of 2.12 (11.1%) and 4.02 (24.8%), respectively. The SEM values indicate that there was an expected random variation in all MAT scores of 0.76 points (4.0%) for patients with HNC and 1.45 points (9.0%) for healthy controls.[21] The SDC values indicate that the difference between two tests needs to be at least 2.12 points (11.1%) for patients with HNC and 4.02 points (24.8%) for healthy controls to be considered a true change in masticatory performance which is not caused by a measurement uncertainty. The Bland-Altman plots (Figures 1 and 2) show that 95% of the data lie between the LoA, with a consistent variability, indicating no systematic variation in performance between two measurements.

Test and retest showed no significant differences for both patients ($Z = -0.206$, $P = .837$) and healthy controls ($Z = -1.406$, $P = .160$). Age and sex were significantly different between patients and healthy controls ($P < .001$ and $P = .001$, respectively). A significant effect of age on MAT outcome in both patients and healthy controls was observed (test: $\chi^2(6) = 19.812$; $P = .003$, retest: $\chi^2(6) = 16.127$; $P = .013$), in which a higher age leads to a higher MAT score and therefore a lower mixing ability. Sex (test: $\chi^2(1) = 0.054$; $P = .815$, retest: $\chi^2(1) = 0.611$; $P = .434$) did not show an effect.

4 | CONCLUSIONS

The aim of this study was to determine the reliability of the MAT for patients with HNC and healthy controls. The results showed a good reliability in patients with HNC and a moderate reliability in healthy controls. In addition, healthy controls showed a higher SEM and SDC in comparison with patients with HNC, indicating a greater difference between test and retest.

In patients with HNC, an ICC of 0.886 was found, indicating a good reliability. In comparison, previous research tested the reproducibility in children with cerebral palsy and healthy children,[15] in which a moderate ICC of 0.69 was found. The higher ICC in patients with HNC in comparison with children indicates that this MAT is more suitable to use in (older) patients with HNC. In comparison, a moderate correlation between test and retest was found for healthy controls (ICC = 0.525), indicating that this MAT is less suitable for healthy subjects.

Healthy controls displayed a better retest result in comparison with the test result, indicating a learning effect (Table 2). This effect was not visible in patients with HNC. In previous research, no learning effect or apparent optimisation of jaw muscle activity was induced by a 1-hour training task.[28] Therefore, it is unlikely that the second MAT shows a better result caused by a learning effect

TABLE 1 Characteristics of patients with head and neck cancer and healthy controls

Characteristics	Patients(n = 34)	Healthy controls (n = 42)	P-value
Age (median, IQR)	64 (9)	31 (27)	<.001[*,a]
Sex			
Male	29 (85%)	20 (48%)	.001[*,b]
Female	5 (15%)	22 (52%)	
Number of teeth (median, IQR)	30 (16)	30 (4)	.968[a]
Number of occlusal units (median, IQR)	12 (16)	12 (4)	.641[a]
Tumour site			
Oropharynx	16 (47%)	NA	—
Larynx	10 (29%)		
Oral cavity	5 (15%)		
Hypopharynx	2 (6%)		
Unknown primary tumour	1 (3%)		
Tumour stage			
I	7 (20.5%)	NA	—
II	7 (20.5%)		
III	5 (15%)		
IV	15 (44%)		
Primary treatment			
RT	15 (44%)	NA	—
CRT	13 (38%)		
Surgery	4 (12%)		
Surgery with PORT	2 (6%)		

Abbreviations: CRT, chemoradiation therapy; IQR, interquartile range; PORT, post-operative radiation therapy; RT, radiation therapy.

[a]Mann-Whitney U test.

[b]Chi-squared test.

[*]P ≤ .001.

in masticatory performance after just 20 chewing strokes and with at least half an hour time difference. The variability between test and retest can be influenced by natural individual variability, unfamiliarity with the wax tablet, or adjustment to the taste and structure.[4,15] In addition, healthy controls have no problems regarding masticatory performance and oral functioning. They need less monitoring and regulating of their movements in comparison with patients, because their movements occur implicitly,[29] which may lead to more variation in chewing outcome. Patients are more aware of their chewing ability, due to for example pain or reduced oral sensibility, and therefore perform their movements more consciously (explicitly).[30] Healthy controls can show more variation in their chewing pattern, whereas patients already reached a ceiling effect.

Significant differences were found between patients with HNC and healthy controls for age and sex; significantly more people were male in the patient group, with a higher age in comparison with the healthy control group. Age had a significant effect on MAT outcome. Previous research showed that age has a negative influence on mastication, because total body muscle mass and muscle mechanical performance decrease, indicating that elderly persons need more

time and more chewing strokes before food can be safely swallowed.[31] In addition, younger people may automatically chew food without additional effort for monitoring or regulating their movement.[29] Older people have a more distinct experience in chewing, where they monitor their oral status continually. This can also be caused by poorer oral conditions[29] such as fewer teeth or occlusal units. However, no significant differences were found between patients with HNC and healthy controls for number of teeth or number of occlusal units, indicating that these oral conditions were similar.
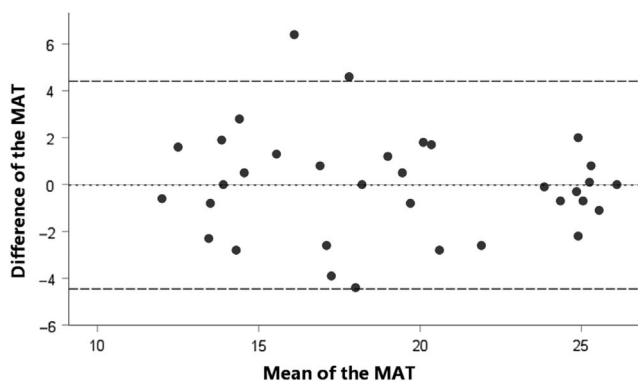
4.1 | Strengths and limitations

The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) checklist was followed to reduce bias and ensure methodological and statistical quality.[32] Reliability was tested in a large research population. Data were collected by the same authors (JAV and FMW), and the MAT score was calculated by the same observer (CMS). However, no inter-rater reliabilities were tested, because this was believed to be too time consuming for patients with HNC. In addition, significant differences

TABLE 2 Reliability of the mixing ability test for patients with head and neck cancer and healthy controls

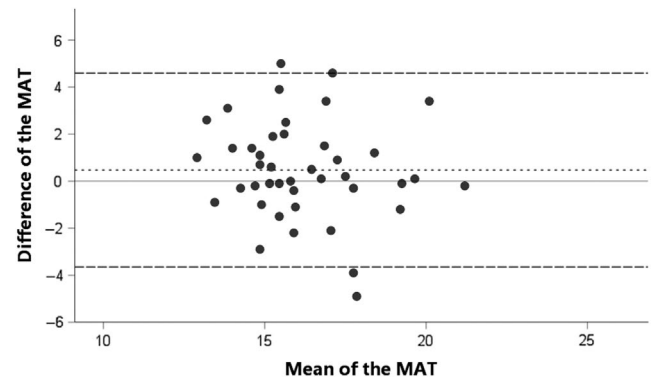|	Patients (n = 34)	Healthy Controls (n = 42)
Test mean (SD)	19.12 (4.56)	16.42 (2.04)
Test median (IQR)	19.30 (8.68)	15.80 (2.60)
Retest mean (SD)	19.14 (4.80)	15.95 (2.30)
Retest median (IQR)	19.20 (9.58)	15.65 (3.13)
Difference test-retest, mean (SD)	−0.02 (2.26)	0.47 (2.10)
$ICC_{2,1}$	0.886	0.525
95% CI	0.784-0.942	0.272-0.712
SEM	0.76	1.45
SEM%	4.0%	9.0%
SDC	2.12	4.02
SDC%	11.1%	24.8%
95% LoA	−4.46 to 4.42	−3.65 to 4.59

Abbreviations: CI, confidence interval; ICC, intra-class correlation coefficient; IQR, interquartile range; LoA, limits of agreement; SDC, smallest detectable change; SEM, standard error of measurement.

FIGURE 2 Bland-Altman plot for the difference between test and retest of the mixing ability test (MAT) for healthy controls. The dashed line represents the mean difference between test and retest, and the striped lines represent the 95% limits of agreement

FIGURE 1 Bland-Altman plot for the difference between test and retest of the mixing ability test (MAT) for patients with head and neck cancer. The dashed line represents the mean difference between test and retest, and the striped lines represent the 95% limits of agreement

random variation of a MAT score at any given time. The SDC and SDC% values can be used to describe minimal changes needed over time in order to be clinically significant. When values between different measurements are larger than the SDC, these changes are not caused by measurement uncertainty and are actual changes over time.[22] The SEM% and SDC% values generate a fair comparison between different measures and indicate that only small changes are needed to indicate a real change in mixing ability over time for patients with HNC. In healthy controls, bigger changes are needed to determine whether performance has truly changed over time.

In conclusion, the MAT has a good reliability in patients with head and neck cancer and a moderate reliability in healthy controls.

CONFLICT OF INTEREST

The authors report no conflicts of interest.

were found between patients and healthy controls for age and sex, causing the groups to be non-comparable.

4.2 | Future research

The results of the test-retest reliability can be used in future research to provide insight into differences over time and differences between different treatment modalities for patients with HNC. Because the ICC showed a good reproducibility, we expect the outcomes of the MAT to be of good reliability for future research. The SEM and SEM% values can be used as an indication for the expected

AUTHORS' CONTRIBUTION

Jorine A. Vermaire involved in conceptualisation, methodology, form analysis, investigation, resources, data curation, writing original draft and visualisation. Florine M. Weinberg involved in investigation, resources, data curation, and writing—review and editing. Cornelis PJ Raaijmakers involved in supervision, visualisation, and writing—review and editing. Irma M. Verdonck-de Leeuw involved in investigation, methodology, and writing—review and editing. Chris HJ Terhaard involved in supervision, writing—review and editing, and data curation. Caroline M. Speksnijder involved in conceptualisation, methodology, form analysis, investigation, resources, visualisation, supervision, and writing—review and editing.

ORCID

Irma M. Verdonck-de Leeuw iD https://orcid.
org/0000-0002-4507-4607

Caroline M. Speksnijder iD https://orcid.
org/0000-0003-0540-3741

REFERENCES

1. Pace-Balzan A, Shaw RJ, Butterworth C. Oral rehabilitation following treatment for oral cancer. *Periodontology*. 2000;57(1):102-117.
2. Teguh DN, Levendag PC, Voet P, et al. Trismus in patients with oropharyngeal cancer: relationship with dose in structures of mastication apparatus. *Head Neck*. 2008;30(5):622-630.
3. van der Bilt A, Engelen L, Pereira LJ, van der Glas HW, Abbink JH. Oral physiology and mastication. *Physiol Behav*. 2006;89(1):22-27.
4. Pereira LJ, van der Bilt A. The influence of oral processing, food perception and social aspects on food consumption: a review. *J Oral Rehabil*. 2016;43(8):630-648.
5. Pedroni-Pereira A, Marquezin MCS, Araujo DS, Pereira LJ, Bommarito S, Castelo PM. Lack of agreement between objective and subjective measures in the evaluation of masticatory function: a preliminary study. *Physiol Behav*. 2018;184:220-225.
6. van der Bilt A, Abbink JH. The influence of food consistency on chewing rate and muscular work. *Arch Oral Biol*. 2017;83:105-110.
7. de Groot RJ, Rosenberg AJWP, van der Bilt A, Aalto D, Merkx MAW, Speksnijder CM. The association between a mixing ability test and patient reported chewing ability in patients treated for oral malignancies. *J Oral Rehabil*. 2019;46(2):140-150.
8. Mendenhall WM, Dagan R, Bryant CM, Fernandes RP. Radiation oncology for head and neck cancer: current standards and future changes. *Oral Maxillofac Surg Clin North Am*. 2019;31(1):31-38.
9. Nemeth D, Zaleczna L, Huremovic A, et al. Importance of chewing, saliva, and swallowing function in patients with advanced oral cancer undergoing preoperative chemoradiotherapy: a prospective study of quality of life. *Int J Oral Maxillofac Surg*. 2017;46(10):1229-1236.
10. Owosho AA, Pedreira Ramalho LM, Rosenberg HI, et al. Objective assessment of trismus in oral and oropharyngeal cancer patients treated with intensity-modulated radiation therapy (IMRT). *J Craniomaxillofac Surg*. 2016;44(9):1408-1413.
11. Ohkoshi A, Ogawa T, Nakanome A, et al. Predictors of chewing and swallowing disorders after surgery for locally advanced oral cancer with free flap reconstruction: a prospective, observational study. *Surg Oncol*. 2018;27(3):490-494.
12. Speksnijder CM, Abbink JH, van der Glas HW, Janssen NG, van der Bilt A. Mixing ability test compared with a comminution test in persons with normal and compromised masticatory performance. *Eur J Oral Sci*. 2009;117(5):580-586.
13. Elgestad Stjernfeldt P, Sjogren P, Wardh I, Bostrom AM. Systematic review of measurement properties of methods for objectively assessing masticatory performance. *Clin Exp Dent Res*. 2018;5(1):76-104.
14. Speksnijder CM, van der Bilt A, Abbink JH, Merkx MA, Koole R. Mastication in patients treated for malignancies in tongue and/or floor of mouth: a 1-year prospective study. *Head Neck*. 2011;33(7):1013-1020.
15. Remijn L, Vermaire JA, Nijhuis-van de Sanden MWG, Groen BE, Speksnijder CM. Validity and reliability of the mixing ability test as masticatory performance outcome in children with spastic cerebral palsy and children with typical development: a pilot study. *J Oral Rehabil*. 2018;45(10):790-797.
16. Weinberg FM, Vermaire JA, Forouzanfar T, Rosenberg A, Speksnijder CM. Reproducibility and construct validity of the Utrecht Mixing Ability Test to obtain masticatory performance outcome in patients with condylar mandibular fractures. *J Oral Rehabil*. 2019;47(4):460-466.
17. Verdonck-de Leeuw IM, Jansen F, Brakenhoff RH, et al. Advancing interdisciplinary research in head and neck cancer through a multicenter longitudinal prospective cohort study: the NETherlands QUality of life and BIomedical Cohort (NET-QUBIC) data warehouse and biobank. *BMC Cancer*. 2019;19(1):765.
18. Paleri V, Mehanna H, Wight RG. TNM classification of malignant tumours 7th edition: what's new for head and neck? *Clin Otolaryngol*. 2010;35(4):270-272.
19. Gotfredsen K, Walls AW. What dentition assures oral function? *Clin Oral Implants Res*. 2007;3(3):34-45.
20. van der Bilt A, Speksnijder CM, de Liz PR, Abbink JH. Digital image processing versus visual assessment of chewed two-colour wax in mixing ability tests. *J Oral Rehabil*. 2012;39(1):11-17.
21. Portney L, Watkins M. *Foundations of Clinical Research*. Upper Saddle River, NJ: Application to Practice; 2015.
22. de Vet HCW, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine; a practical guide*, 1st edn. New York, NY: Cambridge University Press; 2011.
23. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;86(2):420-428.
24. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res*. 2005;19(1):231-240.
25. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60(1):34-42.
26. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;327(8476):307-310.
27. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. *Stat Med*. 1998;17(1):101-110.
28. Kumar A, Svensson KG, Baad-Hansen L, Trulsson M, Isidor F, Svensson P. Optimization of jaw muscle activity and fine motor control during repeated biting tasks. *Arch Oral Biol*. 2014;59(12):1342-1351.
29. Lin CS, Wu CY, Wu SY, Lin HH, Cheng DH, Lo WL. Age-related difference in functional brain connectivity of mastication. *Front Aging Neurosci*. 2017;9:82.
30. Kal E, Prosee R, Winters M, van der Kamp J. Does implicit motor learning lead to greater automatization of motor skills compared to explicit motor learning? A systematic review. *PLoS One*. 2018;13(9):e0203591.
31. Peyron MA, Woda A, Bourdiol P, Hennequin M. Age-related changes in mastication. *J Oral Rehabil*. 2017;44(4):299-312.
32. Mokkink LB, de Vet HCW, Prinsen CAC, et al. COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. *Quality Life Res*. 2018;27(5):1171-1179.