Received: 13 August 2020 | Revised: 8 March 2021 | Accepted: 14 April 2021

DOI: 10.1002/hed.26723

ORIGINAL ARTICLE

WILEY

Reliability of the 100 mL water swallow test in patients with head and neck cancer and healthy subjects

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

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

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

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

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

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

Correspondence
Caroline M. Speksnijder, Department of Oral and Maxillofacial Surgery and Special Dental Care, University Medical Center Utrecht, G05.122, P.O. Box 85.500, 3508 GA Utrecht, The Netherlands.
Email: c.m.speksnijder@umcutrecht.nl

Funding information
KWF Kankerbestrijding, Grant/Award Number: NL45051.029.13

Section Editor: Katherine Hutcheson

Abstract

Background: Dysphagia may occur in up to 44% of patients with head and neck cancer (HNC) treated with radiation therapy and up to 84% of patients treated with surgery. To test the extent of dysphagia, the 100 mL water swallow test (WST) was developed. In this study, reliability of the 100 mL WST was determined in patients with HNC and healthy subjects.

Methods: Thirty-three patients and 40 healthy subjects performed the WST twice on the same day. To assess reliability, the intraclass correlation coefficient ($ICC_{2,1}$), standard error of measurement, smallest detectable change, and limits of agreement were calculated.

Results: Good to excellent correlations were found for patients with HNC (number of swallows; ICC = 0.923, duration; ICC = 0.893), and excellent correlations for healthy subjects (number of swallows; ICC = 0.950, duration; ICC = 0.916).

Conclusion: The 100 mL WST has a good to excellent reliability in patients with HNC and healthy subjects.

KEYWORDS

100 mL water swallow test, dysphagia, head and neck cancer, reliability, swallowing

1 | INTRODUCTION

Swallowing is a physiological process formed by oral, pharyngeal, and esophageal phases.[1] It occurs due to neuromuscular actions involving sensitive cranial, motor, and parasympathetic nerves.[2] Its purpose is to transport food from the mouth to the stomach, promoting hydration and nutrition. In order to be successful at this, a number of rapid, coordinated, and accurate events have to occur, such as soft palate elevation, vocal fold closure, pharyngeal muscle contraction, laryngeal elevation and anteriorization, and epiglottis lowering.[3] These mechanisms occur involuntarily after stimulation of sensory receptors, especially located in the oropharyngeal cavity.[2] A lack of onset or delayed onset of these events can be a sign of dysphagia. Dysphagia is a significant toxicity resulting in difficulty in swallowing, caused by abnormalities in structure or function of cartilaginous, bony, muscular, or neural anatomy involved in normal swallowing.[4] Complications such as malnutrition, aspiration, and subsequent pneumonia can occur.[4] Dysphagia can not only lead to a reduction of intake, but to a reduction in peoples' activities and social interactions as well, with corresponding negative changes to quality of life.[5]

Dysphagia may occur in up to 44% of patients with head and neck cancer (HNC) treated with radiation therapy (RT) and up to 84% of patients treated with surgery.[6,7] In addition, up to two out of three HNC patients may present with dysphagia at the time of diagnosis, and silent aspiration is present in 14% to 18% of patients pretreatment.[8] RT-related toxicity may consist of dysphagia caused by the irradiation of swallowing related normal tissues, fibrosis, edema, ulcers, vascular toxicity, and osteoradionecrosis.[9,10] Chemotherapy can add to the effects of RT and cause edema, mucositis, and fibrosis.[4] Surgical resection of the soft palate, floor of mouth, or base of tongue can cause severe swallowing dysfunction as well,[6] compromising lingual mobility, muscle strength, mastication, muscle action, and muscle coordination.[4,8] The most common procedure to evaluate dysphagia, swallowing safety, and efficiency in patients with HNC is based on video-endoscopy, such as fiberoptic endoscopic evaluation of swallowing (FEES).[8,11] However, these procedures are time consuming and require special equipment. Therefore, the 100 mL water swallow test (WST) was developed.[11,12] This test requires minimal equipment, is easily accessed, and provides quantitative measures of swallowing performance. It is therefore used as a standardized test for screening dysphagia.[13] In addition, the WST may be better in reflecting swallowing in everyday life in comparison to FEES, because it allows the patients to self-select the size of each bolus swallowed.[11] In previous research, the WST was performed in neurological patients, where it had high inter-rater reliability, a difference on average of 2.4% between two measurements, when assessing videotaped swallowing movies.[12,14] Besides, the WST has been validated using video fluoroscopy in patients with neurogenic dysphagia, with a sensitivity and specificity up to 85.5% and 91.7%.[15] It showed no significant inter-rater differences or differences between tests over a 48 h period.[12] The WST has proven to be an excellent test to help identify patients at risk for dysphagia and aspiration, and can be used to monitor swallowing performance over time.[11,16] In order to detect changes that may occur in the WST outcomes after treatment, test–retest reliability is an important test criterion, most often measured with an intraclass correlation coefficient (ICC).[17] Besides, to interpret repeated measurement scores, it is important to use the smallest detectable change (SDC) scores to determine whether a change in scores is significant and not a measurement uncertainty. The SDC is crucial for clinicians and researchers to determine the real change in repeated measurements for individual patients.[18] The reliability of the WST has been tested in patients with motor neuron disease, in which a high inter-rater reliability was found.[19] However, to our knowledge, test–retest reliability has not been performed in patients with HNC yet. The purpose of this study was therefore to assess the reliability of the WST in patients with HNC. In order to detect differences in reliability that may occur in a different population, the reliability was tested in healthy subjects as well.

2 | MATERIALS AND METHODS

Patients were included when they had been diagnosed with oral, oropharyngeal, hypopharyngeal, laryngeal, or unknown primary HNC. Patients were included at the University Medical Center Utrecht (UMCU), the Netherlands, and were referred for either RT, chemo radiation therapy, or surgery, with a curative intent, between September 2016 and June 2018. Patients with recurrent or residual disease, cognitive impairments, and patients having trouble understanding and reading the Dutch language were excluded. Healthy subjects could respond to a flyer outside the hospital, and were included when they were 18 years or older. The study protocol for patients with HNC is part of the NET-QUBIC research,[20] and was approved by the Medical Ethics Committee of the Netherlands (NL45051.029.13). A random selection of the total NET-QUBIC research ($N = 154$) was taken, and patients were asked before the start of the measurements if they would want to perform the WST twice. The study protocol for healthy subjects was approved by the Medical Ethics Committee of the UMCU (18/701). General information about age, sex, tumor site, tumor stage, and treatment were collected for patients with HNC, and about age and sex for healthy subjects. Before participating, all subjects received oral and written information about the study, before signing written informed consent.

TABLE 1 Subject characteristics of patients with head and neck cancer and healthy subjects

Characteristics	Patients ($n = 33$)	Healthy subjects ($n = 40$)	p value
Age (median, IQR)	65 (12)	31 (28)	<0.001*[a]
Sex			
Male	30 (91%)	20 (50%)	0.002**[b]
Female	3 (9%)	20 (50%)	
Tumor site			
Oropharynx	15 (46%)	NA	-
Larynx	10 (30%)		
Oral cavity	5 (15%)		
Hypopharynx	2 (6%)		
Unknown primary	1 (3%)		
Tumor stage			
I	7 (21%)	NA	-
II	9 (27%)		
III	2 (6%)		
IV	15 (46%)		
Primary treatment			
RT	17 (52%)	NA	-
CRT	11 (33%)		
Surgery	3 (9%)		
Surgery with PORT	2 (6%)		

Abbreviations: CRT, chemo radiation therapy; IQR, interquartile range, PORT, postoperative radiation therapy; RT, radiation therapy.
[a]Mann–Whitney U test.
[b]Chi-square test.
*$p < 0.001$; **$p < 0.05$;

2.1 | 100 mL WST

During the 100 mL WST, a subject was asked to drink 100 mL of water as quickly as is comfortably possible. The time to swallow this 100 mL (in seconds) and the number of swallows were counted. The researcher counted the number of swallows by touching the larynx, and the subject was asked to count the number of swallows simultaneously, as a control reference. Timing started when the water touched the bottom lip, and stopped when the larynx came to rest after the last swallow.[14] From these measurements, the following parameters could be calculated: the swallowing volume (the amount of mL per swallow), the swallowing capacity (the amount of mL per second) and the swallowing speed (the time per swallow). Swallowing volume was calculated by dividing the number of mL by the number of swallows. Swallowing capacity was calculated by dividing the number of mL by the duration. Swallowing speed was calculated by dividing the duration by the number of swallows. Subjects failed the test when they coughed or choked post swallow, had a wet voice quality post swallow, or were unable to drink the whole 100 mL.[11] When a person was unable to drink the 100 mL, the residual water was measured and noted. Subjects were instructed to perform the WST two times, with an interim period between 15 min and 2 h, with the same rater and testing conditions for all subjects.

2.2 | Statistical analyses

Test–retest reproducibility of the WST outcomes was tested by a two-way random, single measurement, absolute agreement, $ICC_{2,1}$ calculated as $\frac{\mathbf{MS_R - MS_E}}{\mathbf{MS_R + (k-1)MS_E + \frac{k}{n}(MS_C - MS_E)}}$, 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. Cutoff 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] The standard error of measurement (SEM) was calculated as $SD \times \sqrt{(1 - ICC)}$.[23] For the SD, the standard deviation of the difference between the two WSTs was used. The SEM percent change (SEM%) was calculated as $(SEM/\bar{X}) \times 100$, in which $\bar{X}$ is the mean of all

measurements of test and retest. The SDC was calculated as $1.96 \times \sqrt{2} \times \text{SEM}$.[24,25] The SDC percent change (SDC %) was calculated as $(\text{SDC}/\bar{X}) \times 100$, in which $\bar{X}$ is 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. The agreement between the two tests was summarized using the mean difference and SD of the difference, and the 95% limits of agreement (LoA) were calculated as Mean $\pm$ 1.96 $\times$ SD.[26]

A power analysis was conducted, with an expected ICC of at least 0.7. A p_1 value of 0.9 was chosen, therefore the sample size had to be at least 18.4.[27] Data were tested for normality using a Shapiro–Wilk test. Because data were not normally distributed, a Wilcoxon signed ranks test was conducted to examine differences between the test outcomes of the WST for both patients with HNC and healthy subjects. A paired samples *t*-test was conducted to examine differences between the number of swallows reported by the researcher in comparison to the number of swallows reported by the patient or healthy subject. A Kruskal–Wallis test was conducted to examine differences in WST outcomes according to sex and age.

All analyses were performed using Statistical Package for the Social Sciences (SPSS) version 25 (Chicago, Illinois). A *p*-value below 0.05 was considered statistically significant.

3 | RESULTS

Thirty-three patients with HNC and 40 healthy subjects performed the WST twice on the same day. In Table 1, subject characteristics are depicted for patients with HNC and healthy subjects. The median age for patients with HNC was 65 years (91% male), and 31 years (50% male) for healthy subjects. All subjects were able to drink the 100 mL of water, and no missing data were reported. For the patient group, 10 patients performed the test before treatment, 5 patients 3 months after treatment, 5 patients 6 months after treatment, 5 patients 12 months after treatment, and 8 patients 24 months after treatment. No significant differences ($p = 1.00$) were reported between number of swallows reported by the researcher (mean = 4.25, SD = 2.41) in comparison to the subject (mean = 4.25, SD = 2.43). Significant differences were found for age and sex between patients with HNC and healthy subjects ($p < 0.001$ and $p = 0.002$, respectively). Mean and median scores are depicted in Table 2 for patients with HNC and healthy subjects, and for test and retest.

TABLE 2 WST characteristics of patients with head and neck cancer and healthy subjects. Differences between test and retest outcome for both patients with HNC and healthy subjects are depicted

	Patients ($n = 33$)					Healthy subjects ($n = 40$)				
	Test		Retest			Test		Retest		
	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)	*p* value[a]	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)	*p* value[a]
Number of swallows	4.94 (2.69)	4.00 (3.50)	4.61 (2.68)	4.00 (2.00)	0.069	3.68 (2.02)	3.00 (2.75)	3.70 (2.11)	3.00 (2.75)	0.808
Duration (s)	12.73 (17.57)	7.00 (7.50)	10.76 (12.18)	6.00 (6.50)	0.080	5.69 (5.19)	3.90 (3.91)	5.28 (4.01)	4.11 (3.38)	0.089
Swallowing volume (mL)	25.78 (12.14)	25.00 (17.86)	29.08 (17.66)	25.00 (13.33)	0.046**	36.96 (24.09)	33.33 (28.75)	36.21 (22.09)	33.33 (28.75)	0.875
Swallowing capacity (mL/s)	15.44 (10.53)	14.29 (11.88)	16.09 (9.41)	16.67 (13.41)	0.336	26.29 (14.03)	25.71 (19.88)	26.28 (13.24)	24.33 (17.63)	0.371
Swallowing speed (s)	2.37 (2.35)	1.75 (1.02)	2.28 (1.89)	1.75 (0.96)	0.509	1.49 (0.62)	1.35 (0.58)	1.42 (0.43)	1.34 (0.55)	0.446

Abbreviations: HNC, head and neck cancer; IQR, interquartile range.

[a]Wilcoxon signed ranks test.

**p < 0.05.

TABLE 3 Reliability of the WST for the parameters number of swallows, duration, swallowing volume, swallowing capacity, and swallowing speed. Both patients with head and neck cancer and healthy subjects are depicted

	Diff. test–retest mean (SD)	95% LoA	$ICC_{2,1}$ (95% CI)	SEM	SEM%	SDC	SDC%
Patients ($n = 33$)							
Number of swallows	0.33 (1.02)	2.33 to −1.67	0.923 (0.846 to 0.962)	0.28	5.9%	0.79	16.5%
Duration (s)	1.97 (6.84)	15.38 to −11.44	0.893 (0.793 to 0.946)	2.24	19.1%	6.21	52.8%
Swallowing volume (mL)	−3.30 (10.02)	16.34 to −22.94	0.768 (0.577 to 0.879)	4.83	17.6%	13.38	48.8%
Swallowing capacity (mL/s)	−0.65 (4.74)	8.64 to −9.94	0.888 (0.787 to 0.943)	1.59	10.1%	4.40	27.9%
Swallowing speed (s)	0.09 (0.77)	1.61 to −1.43	0.935 (0.873 to 0.967)	0.20	8.5%	0.55	23.6%
Healthy subjects ($n = 40$)							
Number of swallows	−0.03 (0.66)	1.27 to −1.32	0.950 (0.908 to 0.973)	0.15	4.0%	0.41	11.1%
Duration (s)	0.40 (1.89)	4.11 to −3.29	0.916 (0.847 to 0.954)	0.55	10.0%	1.52	27.6%
Swallowing volume (mL)	0.75 (8.83)	18.06 to −16.56	0.928 (0.869 to 0.961)	2.37	6.5%	6.57	18.0%
Swallowing capacity (mL/s)	0.01 (6.75)	13.23 to −13.21	0.880 (0.785 to 0.935)	2.34	8.9%	6.48	24.6%
Swallowing speed (s)	0.07 (0.39)	0.84 to −0.69	0.733 (0.551 to 0.849)	0.20	13.8%	0.56	38.2%

Abbreviations: CI, confidence interval; ICC, intraclass correlation coefficient; LoA, limits of agreement; SEM, standard error of measurement; SDC, smallest detectable change.

All swallowing parameters (number of swallows, duration, swallowing volume, swallowing capacity, and swallowing speed) showed good to excellent test–retest correlations for patients with HNC (ICC > 0.75), and moderate to excellent correlations for healthy subjects (ICC > 0.70) (Table 3). The SEM values indicated that there is an expected random variation in the different parameters of 5.9% to 19.1% for patients with HNC, and of 4.0% to 13.8% for healthy subjects. The SDC values indicated that the difference between two tests needs to be higher than this SDC value to be considered a true change in swallowing, which is not caused by a measurement uncertainty. Therefore, the difference for the different parameters needs to be higher than 16.5% to 52.8% for patients with HNC, and 11.1% to 38.2% for healthy subjects. The Bland–Altman plots (Figures 1 and 2) showed that 95% of the data lie between the LoA, indicating no systematic variation in performance between two measurements.

The Wilcoxon signed ranks test showed no significant differences between test and retest for all swallowing parameters except swallowing volume in patients with HNC. The Kruskal–Wallis test showed a significant effect for age for all parameters (number of swallows, $p < 0.001$, duration, $p < 0.001$, swallowing volume, $p = 0.001$, swallowing capacity, $p < 0.001$, swallowing speed, $p = 0.005$). Number of swallows and duration increase with age, and swallowing volume, swallowing capacity, and swallowing speed decrease with increasing age. In addition, a significant effect was found for sex for number of swallows ($p = 0.033$) and swallowing volume ($p = 0.044$). Women need a higher number of swallows and have a lower swallowing volume in comparison to men. Duration ($p = 0.257$), swallowing capacity ($p = 0.257$), and swallowing speed ($p = 0.373$) did not show an affect for sex.

4 | DISCUSSION

The aim of this study was to determine the reliability of the WST for patients with HNC and healthy subjects. The results showed moderate to excellent reliability for all measures (ICC > 0.70). The SEM values for patients with HNC were 0.28 (number of swallows), 2.24 (duration [s]), 4.83 (swallowing volume [mL]), 1.59 (swallowing capacity [mL/s]), and 0.20 (swallowing speed [s]), which are small considering the range of outcome possibilities. The SDC values were 0.79 (number of swallows), 6.21 (duration), 13.38 (swallowing volume), 4.40 (swallowing capacity), and 0.55 (swallowing speed), indicating that the outcomes of the WST have to change with at least these values before the observed change over time can be considered a true change in swallowing function and not potentially the result of a measurement uncertainty. The Bland–Altman plots show that 95% of the measures lie between the upper and lower LoA with a consistent variability.

In all measures except swallowing speed, the SEM% and SDC% values were lower in healthy subjects in comparison to patients with HNC, indicating the importance of calculating these values for a specific population.

In previous research, no significant differences were found in swallowing speed between the first and fourth

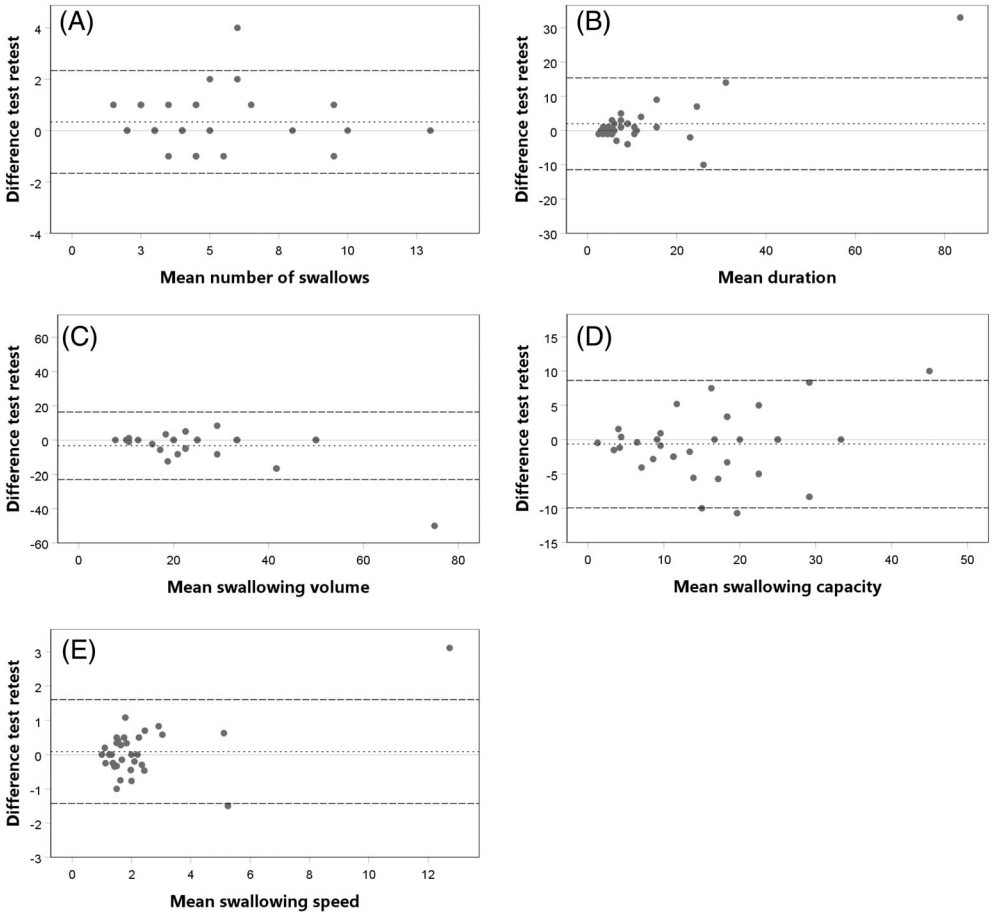

FIGURE 1 Bland–Altman plots for patients with head and neck cancer for the number of swallows (A), duration (B), swallowing volume (C), swallowing capacity (D), and swallowing speed (E). The dashed line represents the mean difference between test and retest and the striped lines represent the 95% limits of agreement

test over a 48-h period.[12] This is in correspondence to the results found in this research, were there is a high reliability between the first and second test, over a 2-h period. Swallowing speed is correlated with age, as found in this research.[12,14,19] However, previous research is inconclusive about the correlation between swallowing speed and sex: although most research found a correlation,[12,14,19] this was not always the case (including this research).[13] With increasing age, speed decreases while time per swallow increases, and speed is most often lower in women in comparison to men. In addition, volume per swallow and swallowing capacity are greater in men.[19] This is in correspondence to the results found in this research, where swallowing volume is correlated to sex. In addition, in this research, a significant effect for age on number of swallows, duration, and swallowing capacity was found, and an effect for sex on number of swallows.

4.1 | Strengths and limitations

This study followed the COSMIN checklist (Consensus-based Standards for the selection of health Measurement INstruments) to ensure methodological and statistical quality, and to reduce bias.[28] A large research population was used to test the reliability, and data were collected by the same author (J. A. Vermaire). However, only three female patients with HNC were tested, making it possible to have missed sex effects in this population. Therefore, the results found on sex differences between men and women should be tested again in a larger population. Because there were significant differences in age and sex between patients with HNC and healthy subjects, these groups are not comparable. Therefore, results should be interpreted separately and can only be applied to subjects with the same sex and age distribution.

Although both patients and healthy subjects performed different tests and filled in questionnaires between test and retest, it is possible that the time between test and retest of approximately 15 min (healthy subjects) to 2 h (patients) has caused recall bias, because previous research used a 48-h time frame.[12] However, no response shift was found between the second and first test; the second test did not always show an improvement compared to the first test, which otherwise would have been visible in the Bland Altman plots in Figure 1. Therefore, it is believed that this possible bias is negligible. No inter-rater reliabilities were tested, because this was believed to be too time consuming for patients with HNC.

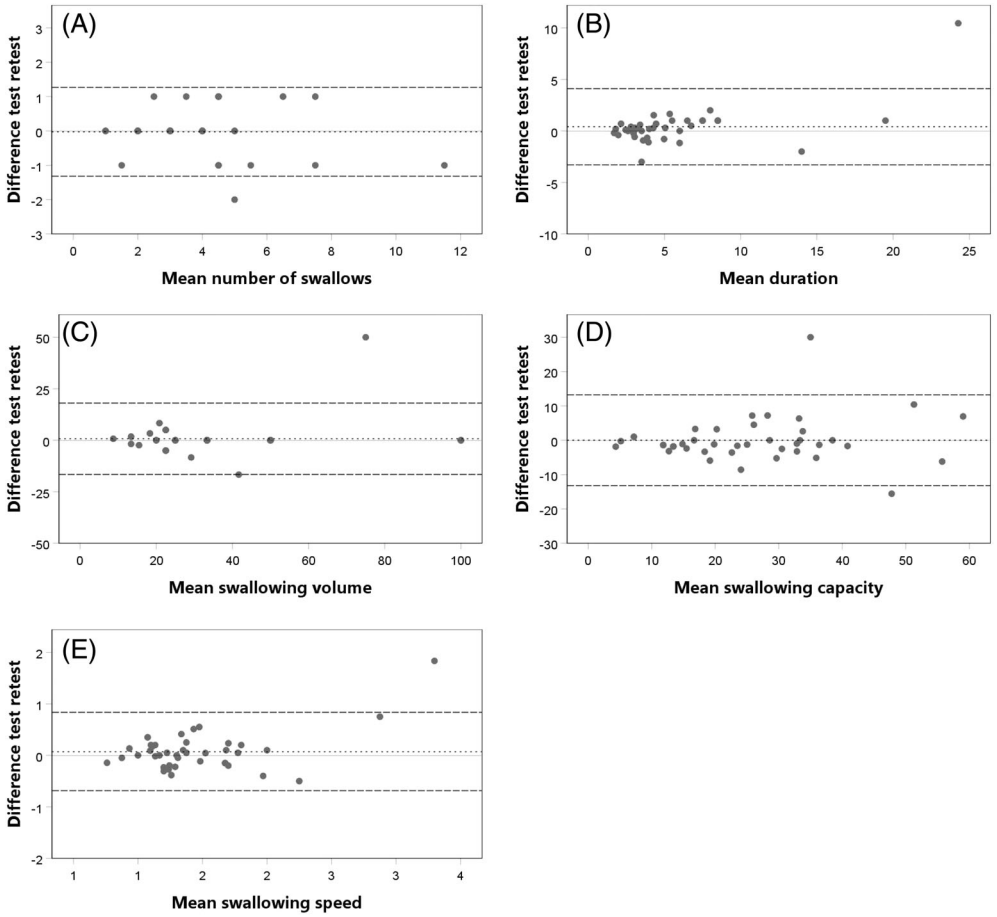

FIGURE 2 Bland–Altman plots for healthy subjects for the number of swallows (A), duration (B), swallowing volume (C), swallowing capacity (D), and swallowing speed (E). The dashed line represents the mean difference between test and retest and the striped lines represent the 95% limits of agreement

In this study, all patients passed the WST and thus showed no signs of dysphagia or aspiration. This contradicts previous results, which show that up to 84% of patients suffer from dysphagia post-treatment,[6] and that the WST has a good sensitivity for the detection of aspiration.[13,16] One explanation could be that the WST missed latent or silent aspiration in patients.[11,13] In addition, a random selection was made of different patients with HNC before treatment up to 2 years after treatment. It may be possible to have missed patients with severe dysphagia, because dysphagia is mainly seen 3 and 6 months after treatment,[11] and in patients with pharyngeal cancer.[14] However, the swallowing speed calculated from the WST provides an effective tool for screening for FEES referral,[15] in which dysphagia can be further evaluated.[8]

4.2 | Future research

The 100 mL WST has been validated using video fluoroscopy in patients with neurogenic dysphagia with a sensitivity and specificity up to 85.5% and 91.7%.[15] In patients with motor neuron disease, the WST had a high inter-rater reliability, with bigger differences between subjects due to the effects of age and sex.[19] The high sensitivity, specificity, and inter-rater reliability indicate that the WST is an excellent test to use when measuring swallowing performance. These finding are equally important as the reliability testing performed in this research, and should be taken into account as well when reporting outcome measures on swallowing performance.

The results of the test–retest reliability can be used in future research to provide insight in differences over time and differences between different treatment modalities for patients with HNC. Swallowing volume was significantly different between test and retest for patients with HNC. In addition, swallowing speed in healthy subjects had a moderate reliability while all other ICCs show a good to excellent reliability, and duration had a relatively high SEM% and SDC% value. We therefore recommend to especially use the parameter number of swallows in future research, instead of the derivatives swallowing volume, swallowing capacity, and swallowing speed. The SEM and SEM% values can be used to indicate the expected random variation in WST outcomes at any given time point before and after treatment for HNC. The SDC and SDC% can be used to describe minimal changes needed between measurements over time in order to be clinically significant.[22]

In conclusion, this study displays a good to excellent reliability of the WST for the parameters number of swallows, duration, swallowing volume, swallowing capacity, and swallowing speed for both patients with HNC, and moderate to excellent reliability for healthy subjects. We recommend to especially use the parameter number of swallows in future research, because this parameter showed and excellent reliability and displayed the smallest SEM% and SDC%. Based on the results found in this study, we expect the results of the WST to be of good reliability, and therefore reliable conclusions can be made in future research using the WST.

ACKNOWLEDGMENTS

We thank all patients and relatives for participating in this research. This study was carried out using the research infrastructure within the NET-QUBIC project (NETherlands QUality of life and BIomedical Cohort studies in Head and Neck Cancer) sponsored by the Dutch Cancer Society/Alpe d'HuZes.

DATA AVAILABILITY STATEMENT

The collection and integration of large amounts of personal, biological, genetic and diagnostic information precludes open access to the NET-QUBIC research data. In the section Data and sample dissemination (www.kubusproject.nl) is described how the data are made available for the research community.

ORCID

Irma M. Verdonck-de Leeuw https://orcid.org/0000-0002-4507-4607
Caroline M. Speksnijder https://orcid.org/0000-0003-0540-3741

REFERENCES

1. Costa MMB. Neural control of swallowing. *Arq Gastroenterol*. 2018;55Suppl 1(Suppl 1):61-75.
2. Zancan M, Luchesi KF, Mituuti CT, Furkim AM. Onset locations of the pharyngeal phase of swallowing: meta-analysis. *CoDAS*. 2017;29(2):e20160067. https://doi.org/10.1590/2317-1782/20172016067.
3. Almeida RCA, Haguette RCB, de Andrade ISN. Swallowing with and without verbal commands: videofluroscopic findings. *Rev Soc Bras Fonoaudiol*. 2011;16(3):291-297.
4. Riffat F, Gunaratne DA, Palme CE. Swallowing assessment and management pre and post head and neck cancer treatment. *Curr Opin Otolaryngol Head Neck Surg*. 2015;23(6):440-447.
5. Perry A, Lee SH, Cotton S, Kennedy C. Therapeutic exercises for affecting post-treatment swallowing in people treated for advanced-stage head and neck cancers. *Cochrane Db Syst Rev*. 2016;8:CD011112. https://doi.org/10.1002/14651858.CD011112.pub2.
6. Kreeft AM, van der Molen L, Hilgers FJ, Balm AJ. Speech and swallowing after surgical treatment of advanced oral and oropharyngeal carcinoma: a systematic review of the literature. *Eur Arch Otorhinol*. 2009;266(11):1687-1698.
7. van der Veen J, Nuyts S. Can intensity-modulated-radiotherapy reduce toxicity in head and neck squamous cell carcinoma? *Cancer*. 2017;9(12):135.
8. Russi EG, Corvo R, Merlotti A, et al. Swallowing dysfunction in head and neck cancer patients treated by radiotherapy: review and recommendations of the supportive task group of the Italian Association of Radiation Oncology. *Cancer Treat Rev*. 2012;38(8):1033-1049.
9. Hutcheson KA, Lewin JS, Barringer DA, et al. Late dysphagia after radiotherapy-based treatment of head and neck cancer. *Cancer*. 2012;118(23):5793-5799.
10. Alterio D, Gerardi MA, Cella L, et al. Radiation-induced acute dysphagia: prospective observational study on 42 head and neck cancer patients. *Strahlenther Onkol*. 2017;193(11):971-981.
11. Patterson JM, Hildreth A, McColl E, Carding PN, Hamilton D, Wilson JA. The clinical application of the 100 mL water swallow test in head and neck cancer. *Oral Oncol*. 2011;47(3):180-184.
12. Nathadwarawala KM, Nicklin J, Wiles CM. A timed test of swallowing capacity for neurological patients. *J Neurol Neurosurg Psychiatry*. 1992;55(9):822-825.
13. Osawa A, Maeshima S, Tanahashi N. Water-swallowing test: screening for aspiration in stroke patients. *Cerebrovasc Dis*. 2013;35(3):276-281.
14. Patterson JM, McColl E, Carding PN, Kelly C, Wilson JA. Swallowing performance in patients with head and neck cancer: a simple clinical test. *Oral Oncol*. 2009;45(10):904-907.
15. Wu MC, Chang YC, Wang TG, Lin LC. Evaluating swallowing dysfunction using a 100-ml water swallowing test. *Dysphagia*. 2004;19(1):43-47.
16. Patterson JM, McColl E, Carding PN, Hildreth AJ, Kelly C, Wilson JA. Swallowing in the first year after chemoradiotherapy for head and neck cancer: clinician- and patient-reported outcomes. *Head Neck*. 2014;36(3):352-358.
17. Brennan P, Silman A. Statistical methods for assessing observer variability in clinical measures. *BMJ*. 1992;304:1491-1494.
18. Kovacs FM, Abraira V, Royuela A, et al. Minimum detectable and minimal clinically important changes for pain in patients with nonspecific neck pain. *BMC Musculoskelet Disord*. 2008;9:43.
19. Hughes TAT, Wiles CM. Clinical measurement of swallowing in health and in neurogenic dysphagia. *Qjm-Mon J Assoc Phys*. 1996;89(2):109-116.
20. 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.
21. Portney L, Watkins M. *Foundations of Clinical Research: Application to Practice*. Upper Saddle River, NJ: Pearson/Prentice Hall; 2015.
22. de Vet HCW, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine; A Practical Guide*. New York: Cambridge University Press; 2011.
23. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;86(2):420-428.
24. 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.

25. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res*. 2005;19(1):231-240.
26. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1(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. Mokkink LB, de Vet HCW, Prinsen CAC, et al. COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. *Qual Life Res*. 2018;27(5):1171-1179.

How to cite this article: Vermaire JA, Terhaard CHJ, Verdonck-de Leeuw IM, Raaijmakers CPJ, Speksnijder CM. Reliability of the 100 mL water swallow test in patients with head and neck cancer and healthy subjects. *Head & Neck*. 2021;43:2468–2476. https://doi.org/10.1002/hed.26723