Supportive Care in Cancer (2019) 27:1673–1677
https://doi.org/10.1007/s00520-018-4412-x

ORIGINAL ARTICLE

Inter-rater reliability of the Oral Assessment Guide for oral cancer patients between nurses and dental hygienists: the difficulties in objectively assessing oral health

Takayuki Aoki[1] · Miyuki Kudo[2] · Maho Endo[3] · Yuka Nakayama[2] · Akiyuki Amano[3] · Mariko Naito[4] · Yoshihide Ota[1]

Received: 27 September 2017 / Accepted: 8 August 2018 / Published online: 15 August 2018

Abstract

Purpose Perioperative oral care is effective for the prevention and reduction of complications following surgery. However, oral cancer patients' oral health is often poor. During hospitalization, oral cancer patients frequently receive oral care from ward nurses as well as professional oral care from dental hygienists. Maintenance of good oral hygiene in these patients ideally requires cooperation between nurses and dental hygienists. Consequently, communication tools used to share information about the status of patients' oral health are needed. One such tool is the Oral Assessment Guide (OAG). However, the inter-rater reliability of the OAG has not been consistently reported; therefore, we examined this between nurses and dental hygienists.

Methods Participants comprised 76 patients hospitalized for oral cancer treatment. After surgery, a nurse and a dental hygienist performed oral assessments using the OAG. Comparative statistical analyses were conducted to examine differences and consistencies in the scores of nurses and dental hygienists.

Results Although almost all patients' oral health status was poor, none were given the worst score in the mucous membrane or gingiva categories. Further, the tongue, saliva, mucous membrane, gingiva, and teeth/denture categories had low κ coefficients, indicating poor nurse–dental hygienist inter-rater reliability. In contrast, the κ coefficients and agreement rates for voice and swallowing were high. Dental hygienists' scores were significantly higher for the tongue, gingiva, and teeth/denture categories than were nurses' scores.

Conclusions This study showed low nurse–dental hygienist inter-rater reliability for the OAG and highlighted the difficulties in objectively assessing patients' symptoms and oral health conditions. Therefore, rather than only relying on an objective assessment of symptoms by a clinician, assessments should also include patients' subjective reporting of symptoms. OAG will likely be used until a new assessment tool is developed. Objective assessment training and/or user manual development may be required to improve the reliability of OAG. The present training of one lesson a year is insufficient, and efforts should be made to improve this to get more reliable outcomes.

Keywords Nurses · Dental hygienists · Oral assessment guide · Oral cancer · Inter-rater reliability

✉ Takayuki Aoki
taoki123jp@ybb.ne.jp

[1] Department of Oral and Maxillofacial Surgery, Tokai University School of Medicine, 143 Shimokasuya, Isehara, Kanagawa 259-1193, Japan

[2] Department of Clinical Technology, Medical Technology Support Division, Tokai University Hospital, 143 Shimokasuya, Isehara, Kanagawa 259-1193, Japan

[3] Division of Nursing, Tokai University Hospital, 143 Shimokasuya, Isehara, Kanagawa 259-1193, Japan

[4] Department of Maxillofacial Functional Development, Graduate School of Biomedical and Health Sciences, Hiroshima University, 1-2-3 Kasumi, Minami-ku, Hiroshima 734-8553, Japan

Introduction

Much evidence supports the value of oral care in managing health [1–4]. Regular performance of oral care to maintain good oral hygiene is known to prevent the occurrence of respiratory infections in the elderly [5]. Performing oral care in nursing homes has been shown to be effective in the prevention of pneumonia and is also cost effective [6]. Oral care is also believed to contribute to the prevention and reduction of complications in compromised hosts, not only in the elderly, but also in cancer patients who have had invasive surgery [1–4]. The provision of regular oral care among cancer patients receiving radiation and chemotherapy treatments is also

effective in managing the severity of oral mucositis, a common adverse event [3, 4]. Consequent inclusion of regular oral care as part of cancer supportive care for these patients is effective, and, in Japan, perioperative oral care is now covered by national health insurance. The principal objectives of providing oral care are to mitigate pain and suffering, improve treatment completion rates and outcomes, and reduce treatment duration and medical costs [1–6].

Among malignancies, oral cancer is one of the worst clinical conditions affecting oral health. According to the guidelines of the National Comprehensive Cancer Network, the standard treatment for oral cancer is surgery [7], which is often accompanied by various functional impairments. Specifically, impaired mastication and impaired swallowing increase the risk of pulmonary aspiration [2–5]. Additionally, poor oral hygiene can negatively impact wound healing or cause wound and respiratory infections [8]. Further, bacterial translocation is known to already occur in oral cancer patients before surgery [9], suggesting that poor oral hygiene is not only a local problem, but that it may have a key impact on the entire body. Therefore, it is important to provide oral cancer patients with perioperative oral care.

During hospitalization, oral cancer patients frequently receive oral care from ward nurses as well as professional oral care from dental hygienists. Maintenance of good oral hygiene in these patients, ideally, requires cooperation between nurses and dental hygienists. Consequently, communication tools used to share information about the status of patients' oral health are needed. One such tool is the Oral Assessment Guide (OAG) [10]. The OAG is an oral health assessment tool developed by Eilers, an oncology clinical nurse specialist at the University of Nebraska Medical Center. Eilers and colleagues showed that the OAG had high nurse-to-nurse inter-rater reliability, suggesting its usefulness, and why it is currently used in many countries [10–14].

Given the recent attention to the importance of interdisciplinary collaboration in healthcare and oral care, an assessment tool that can be used between disciplines with high reliability and validity is needed. Therefore, in this study, we explored the OAG's inter-rater reliability between nurses and dental hygienists to identify any limitations in performing objective oral assessments.

Participants and methods

Participants

Participants were oral cancer patients admitted to the Department of Oral and Maxillofacial Surgery at Tokai University Hospital from March 2016 through June 2017. All patients underwent surgery according to the National Comprehensive Cancer Network's guidelines [7]. All patients provided written, informed consent before participating in this study.

Oral assessment tool

For the oral assessment tool, we used the Japanese version of the OAG [15]. The OAG consists of eight assessment categories: voice, swallowing, lips, tongue, saliva, mucous membrane, gingiva, and teeth/dentures. Each category is expressed in three stages. The OAG score is the sum of the scores for each category; each category is scored from 1 to 3. The best possible score of 8 points is indicative of good oral health; the worst possible score of 24 points is indicative of poor oral health.

Procedure

This study was a prospective, comparative study conducted at the Department of Oral and Maxillofacial Surgery, Tokai University School of Medicine. Our department has long provided perioperative oral care to oral cancer patients. Because oral health status deteriorates after surgery, dental hygienists provide patients with professional oral health care treatments and oral hygiene instructions before surgery. Concurrently, they rate patients' oral health status using the OAG. The nurses who participated in this study are used to observing the oral cavity because they work in an oral surgery ward. In addition, at the beginning of the fiscal year, nurses and dental hygienists attend a lecture showing them how to use and score the OAG. Based on that information, ward nurses provide oral care and complete OAG assessments several times a day. The patients are often unable to perform oral care by themselves for a while after surgery or when the patients' overall conditions are bad. At that time, oral health care is performed more frequently by nurses and dental hygienists.

The data of this study were taken at the time of the first post-surgery oral care. Just before the dental hygienist performed treatment, both a nurse and a dental hygienist completed OAG assessments concurrently. Nurses and dental hygienists assessed the same oral areas separate from the wounded area. They were instructed not to share their scores with each other. Then, their scores were statistically compared. The OAG was scored by 26 nurses and 7 dental hygienists.

Statistical analyses

Nurses' and dental hygienists' OAG total scores and category scores were compared using the Wilcoxon signed-rank tests and, in addition, we calculated the agreement rates and Kappa (κ) coefficients for the scores for each category. Statistical analyses were performed using IBM SPSS Statistics, version 24 (Chicago, IL). $P < 0.05$ was considered statistically significant.

Results

Seventy-six patients agreed to participate in the study. The patients were aged between 23 and 98 years (median age = 70 years, mean age ± standard error= 66.64 ± 1.76 years, 44 men and 32 women). Concerning cancer sites, there were 26 cases involving the tongue, 21 cases involving the lower jaw, 11 involving the upper jaw, 9 involving the buccal mucosa, 5 involving the floor of the mouth, 3 involving the palate, and 1 involving the lower lip. All data in this study were collected within 3 days following surgery, and there were no patients with wound infection.

Comparisons of OAG scores between the nurses and the dental hygienists are shown in Table 1. All agreement rates and κ coefficients for the eight categories are shown in Table 2. Concerning total OAG, dental hygienists' scores were significantly higher than the nurses' scores. Concerning the categorical scores, no significant differences were found between nurses' and dental hygienists' scores for voice, swallowing, lips, saliva, and mucous membrane. The data distributions for both nurses' and dental hygienists' scores on the mucous membrane and gingiva were skewed because the highest score for both categories was 2 points. For tongue, gingiva, and teeth/dentures, dental hygienists' scores were significantly higher than the nurses' scores.

Discussion

The OAG consists of eight categories and each is allocated 3 points. A total of the category scores provides a total OAG score, which indicates patients' overall oral health status.

In this study, the data distributions for the mucous membrane and gingiva scores were skewed. Although almost all patients had poor oral health, none of them were given the worst score for these two categories. The low κ coefficients for the lips, tongue, saliva, mucous membranes, gingiva, and teeth/denture categories indicated that nurse–dental hygienist inter-rater reliability was poor. On the other hand, the κ coefficients and agreement rates for the voice and swallowing categories suggested that the validity of those scores was high for oral cancer patients. In particular, when aspiration occurs, the patient almost chokes due to the reflex action. For assessment of swallowing, expedient objective methods have been established, such as a repetitive saliva swallowing test and modified water swallowing test. Therefore, swallowing may be evaluated in the same way by anyone.

Further, the dental hygienists assigned significantly worse scores than the nurses for the tongue, gingiva, and teeth/denture categories. Perhaps, this was because dental hygienists were more familiar with performing oral cavity assessments than nurses. However, the problem is not whether the dental hygienists or the nurses were more correct in their assessments, but that the inter-rater reliability was poor, and assessments differed from person-to-person.

Unfortunately, it appears that the OAG was developed without effectively considering reliability [10]. To reduce disparity between raters using the OAG, a scholarly discussion is needed to reach a consensus beforehand [14]. However, it is difficult amid the routines of a medical care setting to hold frequent interdisciplinary OAG consensus meetings. Further, categorical differences in inter-rater reliability may not only be due to the scale definition problem, but insufficient clarity in the descriptions themselves. Because the phrasing of the descriptions may affect the raters' psychometrical factors, reliability testing is essential.

Physical evaluations of patients by clinicians tends to be subjective. Difficulty developing assessment scales for physical findings with high inter-rater reliability may not be limited to the OAG. It has been suggested that the objective assessment of patient symptoms and health conditions through clinician's inspection is—in and of itself—difficult [16]. For example, when two doctors examined the same cancer patients and assessed the severity of their symptoms using the Common Terminology Criteria for Adverse Events (CTCAE), a comparison of their assessments showed there were often differences, resulting in low inter-rater reliability [16].

Consequently, many recent studies have suggested the importance of patient-reported outcomes (PRO) [16–19]. These studies, not only in clinical research, but in clinical settings,

Table 1 Comparison of OAG scores between NS and DH

	Voice	Swallow	Lips	Tongue	Saliva	Mucous membrane	Gingiva	Teeth or denture	OAG (total)
NS (mean ± SE)	1.84 ± 0.099	1.89 ± 0.105	1.36 ± 0.061	1.43 ± 0.057	1.45 ± 0.069	1.16 ± 0.042	1.11 ± 0.035	1.24 ± 0.056	11.47 ± 0.353
DH (mean ± SE)	1.88 ± 0.099	1.88 ± 0.099	1.36 ± 0.061	1.79 ± 0.047	1.39 ± 0.062	1.09 ± 0.033	1.24 ± 0.049	1.51 ± 0.071	12.14 ± 0.302
P value	0.439	0.854	1.000	0.000 *	0.553	0.166	0.025 *	0.002 *	0.009 *

OAG Oral assessment guide, *DH* Dental hygienists, *NS* Nurses, *SE* Standard error

Wilcoxon signed-rank test: $*P < 0.05$ was considered statistically significant

Table 2 Agreement rate and κ coefficient

(a) The voice category

agreement rate = 84.21%, κ coefficient = 0.755

Voice		NS			Total
DH	Score	1	2	3	
	1	28	5	0	33
	2	6	13	0	19
	3	1	0	23	24
Total		35	18	23	76

(b) The swallow category

agreement rate = 80.26%, κ coefficient = 0.691

Swallow		NS			Total
DH	Score	1	2	3	
	1	30	3	0	33
	2	4	9	6	19
	3	2	0	22	24
Total		36	12	28	76

(c) The lips category

agreement rate = 67.11%, κ coefficient = 0.281

Lips		NS			Total
DH	Score	1	2	3	
	1	39	12	0	51
	2	11	11	1	23
	3	1	0	1	2
Total		51	23	2	76

(d) The tongue category

agreement rate = 51.32%, κ coefficient = 0.095

Tongue		NS			Total
DH	Score	1	2	3	
	1	29	17	2	48
	2	16	9	1	26
	3	1	0	1	2
Total		46	26	4	76

(e) The saliva category

agreement rate = 51.32%, κ coefficient = 0.025

Saliva		NS			Total
DH	Score	1	2	3	
	1	29	17	2	48
	2	16	9	1	26
	3	1	0	1	2
Total		46	26	4	76

(f) The mucous membrane category

agreement rate = 82.89%, κ coefficient = 0.226

Mucous Membrane		NS			Total
DH	Score	1	2	3	
	1	60	9	0	69
	2	4	3	0	7
	3	0	0	0	0
Total		64	12	0	76

(g) The gingiva category

agreement rate = 73.68%, κ coefficient = 0.100

Gingiva		NS			Total
DH	Score	1	2	3	
	1	53	5	0	58
	2	15	3	0	18
	3	0	0	0	0
Total		68	8	0	76

(h) The teeth or denture category

agreement rate = 82.89%, κ coefficient = 0.226

Teeth or Denture		NS			Total
DH	Score	1	2	3	
	1	34	7	1	42
	2	23	6	0	29
	3	3	1	1	5
Total		60	14	2	76

DH dental hygienists; *NS* nurses

The numbers in red are the same ratings given by dental hygienists and nurses

demonstrated that patients' subjective reporting of symptoms should also be included in an assessment, rather than relying solely on clinicians' objective assessment of patients' symptoms. In addition, some studies have addressed adverse events among patients receiving chemotherapy where CTCAE assessments were conducted by both patients and clinicians and the values were compared [18, 19]. These studies found that while clinicians' assessments of patients' symptoms and health conditions were positively correlated with patients' own reporting of symptoms, the assessment values differed, and they were often underestimated. Therefore, the PRO-CTCAE measurement system has been developed as a tool for measuring symptoms through patient self-assessment by integrating PRO elements into the existing CTCAE [20].

It is entirely possible that the objective assessment of oral health will continue to suffer from low reliability. That is, unless, in the future, a new type of assessment tool is developed for oral health assessment based on the idea that PROs need to be reflected in the assessments.

OAG will likely be used until a new assessment tool is developed. Objective assessment training and/or user manual development may be required to improve the reliability of OAG. The present training of one lesson a year is insufficient, and efforts should be made to improve this to get more reliable outcomes.

Compliance with ethical standards

This study was approved by the Institutional Review Board for Clinical Research at Tokai University School of Medicine (16R-163). This clinical research was conducted in accordance with the Declaration of Helsinki as well as the Ethical Guidelines for Clinical Research by the Japanese Ministry of Health, Labour and Welfare. Written informed consent was obtained from each patient after an explanation of the study protocol. To protect personal information, the obtained data was managed by "anonymity with linking capability."

Competing interests The authors declare that they have no conflicts of interest.

References

1. Mori H, Hirasawa H, Oda S, Shiga H, Matsuda K, Nakamura M (2006) Oral care reduces incidence of ventilator-associated pneumonia in ICU populations. Intensive Care Med 32:230–236. https://doi.org/10.1007/s00134-005-0014-4
2. Yoneyama T, Yoshida M, Matsui T, Sasaki H (1999) Oral care and pneumonia. Oral care working group. Lancet 354:515. https://doi.org/10.1016/s0140-6736(05)75550-1
3. McGuire DB, Fulton JS, Park J, Brown CG, Correa ME, Eilers J, Elad S, Gibson F, Oberle-Edwards LK, Bowen J, Lalla RV (2013) Systematic review of basic oral care for the management of oral mucositis in cancer patients. Support Care Cancer 21:3165–3177. https://doi.org/10.1007/s00520-013-1942-0
4. Saito H, Watanabe Y, Sato K, Ikawa H, Yoshida Y, Katakura A, Takayama S, Sato M (2014) Effects of professional oral health care on reducing the risk of chemotherapy-induced oral mucositis. Support Care Cancer 22:2935–2940. https://doi.org/10.1007/s00520-014-2282-4
5. Yoneyama T, Yoshida M, Ohrui T, Mukaiyama H, Okamoto H, Hoshiba K, Ihara S, Yanagisawa S, Ariumi S, Morita T, Mizuno Y, Ohsawa T, Akagawa Y, Hashimoto K, Sasaki H (1996) Oral hygiene reduces respiratory infections in elderly bed-bound nursing home patients. Arch Gerontol Geriatr 22:11–19. https://doi.org/10.1046/j.1532-5415.2002.50106.x
6. Schwendicke F, Stolpe M, Müller F (2017) Professional oral healthcare for preventing nursing home-acquired pneumonia: a cost-effectiveness and value of information analysis. J Clin Periodontol 13:1236–1244. https://doi.org/10.1111/jcpe.12775
7. Pfister DG, Spencer S, Brizel DM, Burtness B, Busse PM, Caudell JJ, Cmelak AJ, Colevas AD, Dunphy F, Eisele DW, Foote RL, Gilbert J, Gillison ML, Haddad RI, Haughey BH, Hicks WL Jr, Hitchcock YJ, Jimeno A, Kies MS, Lydiatt WM, Maghami E, McCaffrey T, Mell LK, Mittal BB, Pinto HA, Ridge JA, Rodriguez CP, Samant S, Shah JP, Weber RS, Wolf GT, Worden F, Yom SS, McMillian N, Hughes M (2015) Head and neck cancers, version 1.2015. J Natl Compr Cancer Netw 13:847–855. https://doi.org/10.6004/jnccn.2015.0102
8. Shigeishi H, Ohta K, Fujimoto S, Nakagawa T, Mizuta K, Ono S, Shimasue H, Ninomiya Y, Higashikawa K, Tada M, Ishida F, Okui G, Okumura T, Fukui A, Kubozono K, Yamamoto K, Ishida Y, Seino S, Hashikata M, Sasaki K, Naruse T, Rahman MZ, Uetsuki R, Nimiya A, Takamoto M, Dainobu K, Tokikazu T, Nishi H, Sugiyama M, Takechi M (2016) Preoperative oral health care reduces postoperative inflammation and complications in oral cancer patients. Exp Ther Med 12:1922–1928. https://doi.org/10.3892/etm.2016.3532
9. Sakamoto H, Sasaki J, Nord CE (1999) Association between bacterial colonization on the tumor, bacterial translocation to the cervical lymph nodes and subsequent postoperative infection in patients with oral cancer. Clin Microbiol Infect 5:612–616. https://doi.org/10.1111/j.1469-0691.1999.tb00417.x
10. Eilers J, Berger AM, Peterson MC (1988) Development, testing, and application of the oral assessment guide. Oncol Nurs Forum 15:325–330
11. Knoos MM, Ostman M (2010) Oral assessment guide—test of reliability and validity for patients receiving radiotherapy to the head and neck region. Eur J Cancer Care 19:53–60. https://doi.org/10.1111/j.1365-2354.2008.00958.x
12. Andersson P, Persson L, Hallberg IR, Renvert S (1999) Testing an oral assessment guide during chemotherapy treatment in a Swedish care setting: a pilot study. J Clin Nurs 8:150–158. https://doi.org/10.1046/j.1365-2702.1999.00237.x
13. Andersson P, Hallberg IR, Renvert S (2002) Inter-rater reliability of an oral assessment guide for elderly patients residing in a rehabilitation ward. Spec Care Dentist 22:181–186. https://doi.org/10.1111/j.1754-4505.2002.tb00268.x
14. Paulsson G, Wårdh I, Andersson P, Ohrn K (2008) Comparison of oral health assessments between nursing staff and patients on medical wards. Eur J Cancer Care 17:49–55. https://doi.org/10.1111/j.1365-2354.2007.00802.x
15. Shiraishi A, Yoshimura Y, Wakabayashi H, Tsuji Y (2017) Poor oral status is associated with rehabilitation outcome in older people. Geriatr Gerontol Int 17:598–604. https://doi.org/10.1111/ggi.12763
16. Basch E (2014) New frontiers in patient-reported outcomes: adverse event reporting, comparative effectiveness, and quality assessment. Ann Rev Med 65:307–317. https://doi.org/10.1146/annurev-med-010713-141500
17. U.S. Department of Health and Human Services FDA Center for Drug Evaluation and Research; U.S. Department of Health and Human Services FDA Center for Biologics Evaluation and Research; U.S. Department of Health and Human Services FDA Center for Devices and Radiological Health (2006) Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. Health Qual Life Outcomes 4:79
18. Basch E (2010) The missing voice of patients in drug-safety reporting. N Engl J Med 362:865–869. https://doi.org/10.1056/nejmp0911494
19. Basch E, Iasonos A, McDonough T, Barz A, Culkin A, Kris MG, Scher HI, Schrag D (2006) Patient versus clinician symptom reporting using the National Cancer Institute Common Terminology Criteria for Adverse Events: results of a questionnaire-based study. Lancet Oncol 7:903–909. https://doi.org/10.1016/s1470-2045(06)70910-x
20. Basch EM, Reeve BB, Mitchell SA, Clauser SB, Minasian L, Sit L, Chilukuri R, Baumgartner P, Rogak L, Blauel E, Abernethy AP, Bruner D (2011) Electronic toxicity monitoring and patient-reported outcomes. Cancer J 17:231–234. https://doi.org/10.1097/ppo.0b013e31822c28b3