Interactive Journal of Medical Research

Impact Factor (2022): 2.0
Volume 8 (2019), Issue 3   ISSN 1929-073X

Contents

Original Papers

Original Paper

Mobile App Use for Insomnia Self-Management: Pilot Findings on Sleep Outcomes in Veterans

Erin D Reilly[1*], PhD; Stephanie A Robinson[2*], PhD; Beth Ann Petrakis[2], MPA; Eric Kuhn[3,4], PhD; Wilfred R Pigeon[5,6], PhD; Renda Soylemez Wiener[2,7], MD, MPH; D Keith McInnes[2,8], ScD, MS; Karen S Quigley[1,2,9], PhD

[1]Center for Social and Community Reintegration Research, Edith Nourse Rogers Memorial VA Hospital, Bedford, MA, United States

[2]Center for Healthcare Outcomes and Implementation Research, Edith Nourse Rogers Memorial VA Hospital, Bedford, MA, United States

[3]National Center for PTSD, Veterans Affairs Palo Alto Health Care System, Palo Alto, CA, United States

[4]Stanford University School of Medicine, Standford, CA, United States

[5]Center of Excellence for Suicide Prevention, Canandaigua VA Medical Center, Canandaigua, NY, United States

[6]University of Rochester Medical Center, Rochester, NY, United States

[7]Boston University School of Medicine, Boston, MA, United States

[8]Boston University School of Public Health, Boston, MA, United States

[9]Northeastern University, Boston, MA, United States

[*]these authors contributed equally

Corresponding Author:
Erin D Reilly, PhD
Center for Social and Community Reintegration Research
Edith Nourse Rogers Memorial VA Hospital
200 Springs Rd
Bldg 9, Room 106
Bedford, MA, 01730
United States
Phone: 1 781 687 4191
Email: Erin.Reilly@va.gov

Abstract

Background: Sleep disturbance is a major health concern among US veterans who have served since 2001 in a combat theater in Iraq or Afghanistan. We report subjective and objective sleep results from a pilot trial assessing self-management–guided use of a mobile app (*CBT-i Coach*, which is based on cognitive behavioral therapy for insomnia) as an intervention for insomnia in military veterans.

Objective: The primary aim of this study was to evaluate changes in subjective and objective sleep outcomes from pre to postintervention.

Methods: Subjective outcomes included the Insomnia Severity Index, the Pittsburgh Sleep Quality Inventory, and sleep-related functional status. A wearable sleep monitor (WatchPAT) measured objective sleep outcomes, including sleep efficiency, percent rapid eye movement (REM) during sleep, sleep time, and sleep apnea. A total of 38 participants were enrolled in the study, with 18 participants being withdrawn per the protocol because of moderate or severe sleep apnea and 9 others who dropped out or withdrew. Thus, 11 participants completed the full 6-week *CBT-i Coach* self-management intervention (ie, completers).

Results: Completer results indicated significant changes in subjective sleep measures, including reduced reports of insomnia (Z=–2.68, P=.007) from pre (mean 16.63, SD 5.55) to postintervention (mean 12.82, SD 3.74), improved sleep quality (Z=–2.37, P=.02) from pre (mean 12.82, SD 4.60) to postintervention (mean 10.73, SD 3.32), and sleep-related functioning (Z=2.675, P=.007) from pre (mean 13.86, SD 3.69) to postintervention (mean 15.379, SD 2.94). Among the objective measures, unexpectedly, objective sleep time significantly decreased from pre to postintervention (χ^2_2=7.8, P=.02). There were no significant changes in percent REM sleep or sleep efficiency.

Conclusions: These findings suggest that the *CBT-i Coach* app can improve subjective sleep and that incorporating objective sleep measures into future, larger clinical trials or clinical practice may yield important information, particularly by detecting previously undetected sleep apnea.

Trial Registration: ClinicalTrials.gov NCT02392000; http://clinicaltrials.gov/ct2/show/NCT02392000

(Interact J Med Res 2019;8(3):e12408) doi:10.2196/12408

KEYWORDS

cognitive behavioral therapy; mobile apps; insomnia; sleep apnea

Introduction

Chronic Insomnia in Veterans

Sleep disturbance, especially chronic insomnia (difficulty falling and staying asleep), is a serious and prevalent problem among US veterans who have served since 2001 in a combat theater in Iraq and/or Afghanistan [1,2-5]. From 2000 to 2010, there was a 7-fold increase in the diagnosis of insomnia across veterans of all ages who were seeking care in the Veterans Health Administration (VHA) [5], with an upward trend in rates of insomnia and sleep apnea in military veterans continuing into this decade [6]. Sleep difficulties rarely occur in isolation from other physical and psychiatric concerns. For instance, insomnia and other self-reported sleep disturbances commonly cooccur with mental health diagnoses, such as posttraumatic stress disorder (PTSD) and depression [1,7-9], and they also cooccur with significant pain symptoms and other functional impairments [10-14]. Among veterans with multiple comorbidities, such as pain, PTSD, and traumatic brain injury, sleep disturbance rates are even higher. In VHA polytrauma clinics, 94% of US veterans who have served since 2001 in Iraq and/or Afghanistan report some sleep disturbance [4,15,16]. Thus, insomnia appears to be a prevalent problem for the most recent cohort of military veterans.

Treating Chronic Insomnia With Cognitive Behavioral Therapy

Cognitive behavioral therapy for insomnia (CBTI) is a manualized therapy that uses cognitive behavioral techniques, including reconditioning, sleep restriction, sleep hygiene education, and relaxation skills, to help individuals manage and reduce chronic insomnia. There is strong evidence for the efficacy of CBTI [17], including brief interventions, such as 6 to 8 weeks of structured, weekly CBTI sessions [18,19], as well as briefer treatments [20]. Furthermore, CBTI has been successfully deployed in populations with comorbid health conditions [21], such as chronic pain [22-25], PTSD [17-19,26], and depression [27,28]. In addition, evidence from 2 uncontrolled studies of outpatient clinic patients [27] and a large randomized trial of internet-delivered CBTI [29] suggests that CBTI can reduce suicidal ideation, which is of paramount importance given the public health crisis of suicide among post-9/11 veterans [30]. Although the VHA has prioritized training additional clinicians to use CBTI, the demand for behavioral sleep treatments still outstrips supply [23]. In addition, adherence to CBTI can be poor, especially among younger, working-age veterans who face many competing life demands, including health concerns, work, and/or school and family needs [31,32]. Given these challenges, the use of mobile CBTI interventions has been increasing steadily over the last decade. Research suggests that telephone-delivered CBTI is feasible and acceptable [32], with many veterans preferring either individual-based or internet-delivered insomnia treatments rather than in-person, group treatment delivery [31]. This mirrors research in civilian populations as well, where electronically delivered versions of CBTI have been found to improve sleep in college students [33]. Thus, we hypothesized that use of a mobile app to deliver CBTI elements would be well received by veterans with chronic insomnia, given their treatment delivery preferences and high levels of interest in and use of computers and the internet for medical purposes [34]. A publicly available version of a mobile app to deliver CBTI elements, *CBT-i Coach*, was released in 2013 by the Veterans Affairs (VA) National PTSD Center and the Department of Defense National Center for Telehealth and Technology [35-37]. *CBT-i Coach* was based on a CBTI manual created by VA and university sleep researchers [38], and it includes sleep diaries, sleep health education, sleep restriction guidelines, and tools to encourage relaxation. Other features include reminders for bedtime/wake time, stopping caffeine, and scheduled wind down and worry time, with surveys to guide behavioral or environmental changes. For more details about the app, see Kuhn et al 2016 [36].

Utilization of Subjective and Objective Sleep Measurement for Intervention Assessments

Subjective and objective measures of sleep often provide nonredundant data in the context of an intervention, as they often do not correlate highly in patients with insomnia [39], including some individuals with subjectively defined insomnia but objectively normal sleep [40]. Subjective/objective concordance varies as a function of health status, with lower concordance between subjective and objective sleep duration in those with poorer functional health, obesity, depressive symptoms, or lower sleep efficiency [39]. Objective and subjective sleep measures can differentially predict treatment outcomes [41]. As subjectively measured insomnia is related to numerous comorbid health outcomes (eg, chronic pain, posttraumatic stress, and depression), subjective measures are critical for measuring the impact of CBTI intervention success, and these are frequently measured through daily sleep diaries [42], self-report measures, such as the Pittsburgh Sleep Quality Inventory (PSQI) [43], or the Insomnia Severity Index (ISI) [44]). Given the difficulty younger veterans have with committing to the time required for in-person CBTI, further investigation of the impact of the *CBT-i Coach* app on subjective sleep in veterans is warranted. Measures of sleep disturbance, such as a reduced proportion of time spent in rapid eye movement (REM) sleep over a night and the occurrence of sleep apnea, have been proposed as potentially important objective sleep problems in veterans [3,45-48]. Information about REM abnormalities (eg, such as shorter more frequent bouts of REM [49-51]) also may be used as an indication of the need to assess for PTSD symptoms and/or to titrate PTSD treatments. The difficulty in efficiently capturing objective sleep data is a barrier to incorporating objective data into clinical practice. On the one hand, the costs and patient burden of laboratory-based overnight polysomnography are prohibitive. Furthermore, wrist-worn actigraphy devices measure sleep-wake activity, but they do not provide sleep stage information. As objective sleep measures in a self-management trial could have the benefit of improving the tailoring of CBTI interventions and sleep self-management, efforts to gather such data, using patient-friendly, well-validated devices, are worthwhile.

This Study

This pilot study investigated the effects of a mobile insomnia self-management intervention. Self-management of insomnia was facilitated with the *CBT-i Coach* mobile app, including suggested activities in the app over a 6-week intervention period. Our primary aim was to assess the effectiveness of the *CBT-i Coach* app. We hypothesized there would be pre to postintervention improvements in subjective sleep outcomes (self-reported insomnia severity, sleep quality, and functional sleep) and objective sleep-related variables (total sleep time, sleep efficiency, and percent time in REM sleep). We also explored the relationship between demographic factors and self-reported mental and physical health outcomes, and we assessed the prevalence of positive sleep apnea screens in our sample.

Methods

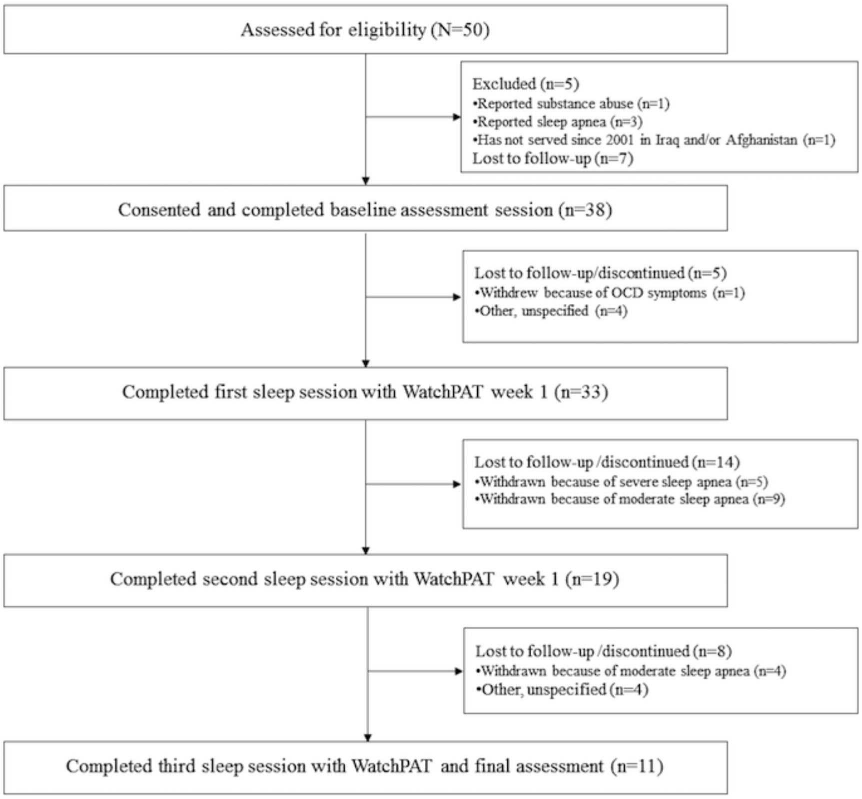
Overview

This study was approved by the Institutional Review Board of the Edith Nourse Rogers Memorial veterans Hospital, Bedford, Massachusetts. We report subjective and objective sleep and functional health outcomes from a 6-week open trial pilot intervention of self-management–based use of the *CBT-i Coach* app. Owing to space restrictions, other measures of feasibility and usability testing of the features of this study will be reported elsewhere.

Participant Recruitment

Veterans were recruited via flyers, presentations, outreach to local community organizations, referrals from VA behavioral and sleep health providers, and recruitment letters. Recruitment letters were sent to a list of potentially eligible veterans that was generated by VA Informatics and Computing Infrastructure, followed by a telephone call no sooner than 2 weeks later to veterans who did not reply to the letter. Interested veterans were screened for study eligibility by phone. To be eligible, participants must have served since 2001 in a combat theater in Iraq and/or Afghanistan, reported current insomnia lasting at least 1 month, as defined by an ISI score greater than 10 [32,52], and impaired daytime functioning (as measured by endorsing *Much* or *Very Much* on ISI Item 7 about how much sleep problems interfere with daily functioning). Participants were excluded if they demonstrated moderate-to-severe cognitive impairment (defined by scores on the Telephone Mini Mental State Exam [53]), self-reported sleep apnea, periodic leg movements, or circadian rhythm disorder (delayed or advanced sleep phase). Figure 1 shows participant flow through the study.

Figure 1. Consolidated Standards of Reporting Trials diagram of participant inclusion and attrition. OCD: obsessive-compulsive disorder.



Participant Demographics

The mean age of the sample was 44.29 years, and the mean body mass index (BMI) was 29.0 kg/m^2 (ie, in the overweight range). Participants (32 men, 6 women) had an average baseline ISI score of 15.42 (ie, moderate insomnia severity). Participants could identify with multiple racial categories, with 89% (34/38) participants identifying as white/Caucasian and 13% (5/38) participants identifying as Hispanic. A total of 37% (14/38) participants reported a combined average household income of US $50,000 to US $99,000 and another 37% (14/38) participants reported an average income over US $100,000. 94% (36/38) of participants reported that their health in general was good, very good, or excellent (94%). In addition, 55% (21/38) participants of the sample were married or living with a partner. The sample was well educated, with 61% (23/38) participants having an Associate’s or higher degree. See Table 1 for detailed demographic information.

Table 1. Descriptive statistics by enrollment status (N=38).

Characteristic	Withdrawn because of apnea (n=18)	Self-withdrew (n=9)	Completers (n=11)	Total (N=38)
Age (years), mean (SD)	42.22 (11.58)	44.11 (10.06)	47.82 (10.52)	44.29 (10.92)
BMI[a], mean (SD)	30.68 (4.49)	27.77 (4.11)	27.22 (3.02)	29.00 (4.10)
Baseline ISI[b], mean (SD)	15.33 (4.19)	14.11 (4.68)	16.63 (5.55)	15.42 (4.69)
Race, n (%)				
White	15 (83)	9 (100)	10 (91)	34 (90)
Black/African American	2 (11)	0	0	2 (5)
Puerto Rican	2 (6)	0	1 (9)	2 (5)
Filipino	0	1 (11)	1 (9)	2 (5)
American Indian	0	0	1 (9)	1 (3)
Other	2 (6)	0	1 (9)	2 (5)
Ethnicity, n (%)				
Hispanic/Latino	3 (17)	1 (13)	1 (9)	5 (13)
Not Hispanic/Latino	15 (83)	8 (89)	10 (91)	33 (86)
Income US $, n (%)				
Less than 11,999	0	1 (11)	1 (9)	2 (5)
12,000-24,999	3 (17)	0	0	4 (11)
25,000-49,999	2 (11)	0	2 (18)	4 (11)
50,000-99,000	6 (33)	4 (44)	4 (36)	14 (37)
100,000+	7 (39)	3 (33)	4 (36)	14 (37)

[a] BMI: body mass index.

[b] ISI: Insomnia Severity Index.

Description of the Intervention: Cognitive Behavioral Therapy-i Coach, Self-Management Guidance, and WatchPAT

The intervention comprised 6-week use of the *CBT-i Coach*, with added self-management support and a set of supplemental app-delivered worksheets to address behavioral concerns or prompt behavioral changes. All participants were loaned an Apple iPod Touch, on which the app was installed (iOS ver. 2.0) for the duration of the study.

Cognitive Behavioral Therapy-i Coach

CBT-i Coach offers sleep psychoeducation, tools for tracking sleep (eg, enables and can prompt for completion of daily sleep diaries and the ISI), and provides sleep hygiene recommendations, including cultivating a conducive sleep environment, engaging in regular exercise, and maintaining a healthy diet. Relaxation tools include multiple guided imagery audio clips, tips for winding down, breathing tools, and an audio-guided progressive muscle relaxation. A behavioral plan can also be reviewed and updated in *CBT-i Coach*, including setting reminders for when to go to sleep and get out of bed, complete sleep diaries and take ISI assessments, engage in scheduled worry time, and stop caffeine intake for the day. *CBT-i Coach* allows users to see graphs of their sleep diary data and ISI scores. We also created supplemental worksheets on the basis of the elements of the *Quiet Your Mind andGet to Sleep* manual [54]. These sleep worksheets were embedded in a separate app that participants could also access from the iPod Touch. Supplemental worksheets included Wakeful Activities, Coping Self-Statements, Constructive Worry, and a Relaxation Log.

Self-Management Guide

A self-management guide, in the form of a document accessible on the iPod Touch, along with a paper copy, provided week-by-week suggestions for using elements of the app and the worksheets. For each week of the 6-week intervention, the guide suggested what materials to read in the app, which features of the app to use, such as completing a daily sleep diary each morning, and which worksheets to complete.

WatchPAT Sleep Monitor

Objective sleep was recorded with a WatchPAT (model WP200U) sleep monitor (Itamar Medical Inc). The WatchPAT sleep monitor is a Food and Drug Administration–approved device that assesses objective sleep parameters, including a screen for obstructive sleep apnea. The WatchPAT is worn like a simple wristwatch with a plethysmographic-based finger-mounted probe and a small sensor on the chest to measure snoring. It is less obtrusive and less disruptive of sleep than either in-lab sleep assessments or in-home polysomnography. Moreover, the participant can use the device himself/herself, with simple instructions, which were provided via video on the iPod Touch and a laminated pamphlet.

Measures

Demographics

Participants reported their age, gender, marital/partnered relationship status, race and ethnicity, and highest education level achieved. Height and weight were measured at the first visit.

Primary Outcomes

Subjective Sleep Measures

Self-reports of insomnia, sleep quality, and functional outcomes because of sleep were measured at baseline and final assessment visits using the ISI, the PSQI, and the Functional Outcomes of Sleep Questionnaire-10 item (FOSQ-10). The ISI has been shown to be sensitive to changes in insomnia severity with CBTI interventions [19,52,55], possible scores ranged from 0 to 28, with a higher score indicating more severe insomnia. The PSQI, a global measure of perceived sleep quality, has also been extensively validated and shown to be sensitive to change after CBTI [19,32]. Scores on the PSQI can range from 0 to 21, with a higher score indicating worse sleep quality. In the current sample, Cronbach alphas were acceptable (ISI=.83; PSQI=.75). The FOSQ-10 [56], a brief version of the original 30-item FOSQ [57], was used to assess the impact of sleepiness on functioning in everyday activities (Cronbach alpha=.89). Possible scores ranged from 5 to 20, with higher scores indicating better functional status.

Objective Sleep Measures

Objective sleep was measured via the WatchPAT and included total sleep time, total and percent time spent in light, deep, and REM sleep stages, apnea-hypopnea index (AHI), respiratory disturbance index (RDI), and number of awakenings. The WatchPAT is an FDA-approved portable diagnostic device that assesses sleep stages and detects probable sleep apnea with well-established validity in comparison to polysomnography-based measures of sleep apnea (eg, AHI and RDI; [58-65]). The WatchPAT calculates the proportion of REM sleep using a genetic algorithm (ie, a machine learning technique) to determine REM sleep onset and offset. The WatchPAT's REM stage determination has also been validated against traditional polysomnography [66], as have its estimates of the duration of episodes of light versus deep sleep [67]. Participants were asked to wear the WatchPAT on their nondominant hand. Given the well-known "first night effect" in which sleep can be negatively impacted by sleep monitoring [68], especially in those with insomnia [69], participants recorded 2 nights of sleep at the beginning of the study, followed by 1 additional night at the end of the intervention, totaling 3 nights of objective sleep data.

Secondary Outcomes

Self-reported mental health and pain outcomes were measured at pre, mid, and postintervention visits. Mental health measures included assessments of nonspecific physical symptoms with the Patient Health Questionnaire-15 (PHQ-15) [70], depressive symptoms with the PHQ-9 [71], and PTSD symptoms with the PTSD Checklist (PCL-5) [63,72,73]. Each of these measures demonstrated good internal consistency (PHQ-15 alpha=.85, PHQ-9 alpha=.89, and PCL-5 alpha=.96). Pain severity was assessed using a 3-item subscale from the West Haven-Yale Multidimensional Pain Inventory (WHYMPI) [74], and pain-related functional health was measured using the Pain Disability Index (PDI) [75,76]. Both the WHYMPI and PDI demonstrated good internal consistency, with Cronbach alphas of .92 and .94, respectively.

Procedure

At the first visit, participants completed the self-report questionnaires. They were given a printed version of the self-management guide, they were shown how to use it, and they were shown where to find an electronic copy on the iPod Touch, which they were given to use during the study. The researcher demonstrated use of the iPod Touch, *CBT-i Coach*, and the worksheet app and answered any questions. Participants were guided through setting a reminder to complete a weekly ISI assessment, and then they were guided through completing the first sleep diary and ISI on the *CBT-i Coach*. Participants were shown the WatchPAT, and they viewed a 4-min video on the iPod Touch that demonstrated how to set up and use the WatchPAT, including where and how to attach the sensors. The WatchPAT pamphlet also contained instructions and technical support information. Participants were instructed to complete daily sleep diaries and the activities in the self-management guide and to return with both devices for the second visit within the next week. During the second visit, the researcher downloaded the data from the WatchPAT and printed the associated sleep report. Those with an AHI score greater than 15 (indicative of probable moderate-to-severe sleep apnea) were excluded from further participation, and results were provided to their primary care provider for further assessment and clinical management. The report, which included the AHI score, total sleep times, and percentage of time in REM sleep, was reviewed with the participant by the researcher. Participants who were not excluded were asked to complete another night of WatchPAT monitoring within the coming week and then mail the device back to the investigators using a prepaid stamped envelope. The AHI score was again used to determine whether significant sleep apnea was likely (ie, a positive screen at the moderate or greater range), and, if so, the participant was withdrawn from the intervention and referred for further testing. At the midpoint of the intervention (about week 3-4), participants completed questionnaires with a researcher by phone and were scheduled for their postintervention visit (about week 6-7) and mailed a Watch PAT device. At the third and final in-person visit (ie, postintervention), participants returned both devices, completed final self-report questionnaires and a qualitative interview, and received their final WatchPAT report. WatchPAT reports were also shared with the participant's primary care provider. Participants received US $15 each for the baseline, first WatchPAT, and midpoint visits, and they received US $40 at the completion of the postintervention visit, for a possible total of US $70 for completion of the entire study. Use of the *CBT-i Coach* app, specifically input of nightly sleep diaries, was high, with 9 of the 11 participants using the sleep diary portion of the app 85% of the time.

Data Analysis

To address our primary aim, we evaluated whether subjective sleep outcome variables (ISI, PSQI, and FOSQ) changed from pre to postintervention, using Wilcoxon signed rank tests (nonparametric paired 2-tailed *t* tests). In addition, using repeated measures analysis of variance, we evaluated whether objective sleep outcome variables recorded from the WatchPAT—sleep time, REM percent of total sleep time, and sleep efficiency (ie, percent of total time in bed spent asleep vs awake)—changed across the 3 nights (2 nights early in the intervention and 1 at intervention end). Furthermore, using bivariate correlations, we explored the relationship among mental health symptom measures (ie, PHQ-15, PHQ-9, and PCL-5), pain severity (subscale from the WHYMPI), and subjective sleep outcomes for the entire enrolled sample (N=38). Finally, we used a chi-square test of independence to conduct exploratory analyses of sleep apnea status with demographic and mental health–related correlates of sleep apnea severity (ie, AHI in the mild, moderate, or severe range).

Results

Subjective Sleep Outcomes

Wilcoxon signed rank tests were used to compare the average scores for subjective sleep measures between pre- and posttest. Participants (n=11) showed a significant decrease in ISI scores (*Z*=–2.68, *P*=.007) from pre (median=15.00) to postintervention (median=13.00); a total of 9 participants noticed a reduction in reported insomnia severity over the 6-week intervention, and 2 participants reported no changes. Similarly, there was a significant decrease in PSQI scores (*Z*=–2.37, *P*=.02) from pre (median=12.50) to postintervention (median=12.00); a total of 8 participants noticed an improvement in sleep quality, 2 participants noticed their sleep being worse, and 1 participant reported no difference. Finally, there was a significant increase in FOSQ scores (*Z*=2.675, *P*=.007) from pre (median=14.33) to postintervention (median=16.33), with 9 participants reporting better functional outcomes at the end of the intervention that they attributed to better sleep, and 2 participants reporting no changes in functioning (see Table 2 for group descriptive statistics).

Table 2. Primary subjective sleep outcome measures at pre-, mid-, and postintervention.

Outcome	Preintervention		Midintervention		Postintervention	
	Mean (SD)	Median	Mean (SD)	Median	Mean (SD)	Median
ISI[a]	16.63 (5.55)	15.00	12.17 (4.39)	12.50	12.82 (3.74)	13.00
PSQI[b]	12.82 (4.60)	12.50	10.50 (2.65)	9.00	10.73 (3.32)	12.00
FOSQ[c]	13.86 (3.69)	14.33	15.26 (3.61)	15.25	15.379 (2.94)	16.33

[a] ISI: Insomnia Severity Index.

[b] PSQI: Pittsburgh Sleep Quality Index.

[c] FOSQ: Functional Outcomes of Sleep Questionnaire.

Objective Sleep Outcomes

Friedman tests (nonparametric repeated measures) were also used to compare total sleep time recorded across 3 nights of objective sleep measurement provided by the WatchPAT, which revealed a statistically significant difference in total sleep time across the 3 monitored nights ($\chi^2_{2,11}$=7.8; *P*=.02). Post hoc tests with Bonferroni correction revealed that there was a significant decrease in total sleep time (*P*=.004) from the first sleep session (median= 6 hours 53 min) to the second sleep session (median=5 hours 52 min) and from the first sleep session (*P*=.02, median=6 hours 53 min) to the third sleep session (median=6 hours 35 min). There was no statistically significant difference from the second to the third sleep sessions. There was also no significant effect of sleep session for the other objective sleep measures, including the percent REM sleep ($\chi^2_{2,11}$=2.5; *P*=.29) or sleep efficiency, ($\chi^2_{2,11}$=0.1; *P*=.91; see Table 3).

Table 3. Primary outcomes of objective sleep measures at sleep sessions.

Outcome	First sleep session		Second sleep session		Third sleep session	
	Mean (SD)	Median	Mean (SD)	Median	Mean (SD)	Median
Sleep time, hours	6.93 (1.39)	6.53	5.65 (1.45)	5.52	6.03 (1.71)	6.35
REM[a] (%)	23.36 (7.67)	22.5	26.72 (6.94)	27.9	23.62 (7.06)	26.9
Sleep efficiency (%)	85.65 (6.02)	87.24	85.93 (6.69)	86.62	86.15 (8.43)	87.12

[a] REM: rapid eye movement.

Relationship of Sleep to Mental Health and Pain Outcomes

Table 4 shows the bivariate correlations among mental health measures, pain severity, and subjective sleep outcomes, and Table 5 details descriptive statistics from baseline, midpoint assessment, and postintervention.

Higher scores on the ISI (ie, greater insomnia symptoms) were significantly and positively correlated with higher scores on the PHQ-15 (ie, greater nonspecific physical symptoms), PHQ-9 (ie, greater depressive symptoms), PCL-5 (ie, greater posttraumatic stress symptoms), and WHYMPI subscale (ie, greater pain severity). Higher scores on the FOSQ (ie, better functioning attributed to sleep) were significantly negatively correlated with nonspecific physical symptoms, depressive symptoms, posttraumatic symptoms, and pain severity (PHQ-15, PHQ-9, PCL-5, and WHYMPI subscale, respectively). In addition, higher scores on the PSQI (ie, poorer sleep quality) were significantly positively correlated with higher scores on the PHQ-15 (ie, greater nonspecific physical symptoms), PHQ-9 (ie, depressive symptoms), PCL-5 (ie, posttraumatic stress symptoms), and WHYMPI subscale (ie, pain severity). Finally, there were significant positive correlations between scores on the ISI and PSQI, and there were negative correlations of both measures with the FOSQ. Nonparametric Wilcoxon signed rank tests were run to assess changes between pre- and midpoint of the intervention (n=11) for the PHQ-9 (*P*=.23), PHQ-15 (*P*>.99), WHYMPI subscale (*P*=.55), and PDI (*P*=.28), as well as pre- and postintervention for the PCL-5 (*P*=.45), PHQ-9 (*P*=.12), PHQ-15 (*P*=.67), WHYMPI subscale (*P*=.53), and PDI (*P*=.18), with all changes nonsignificant.

Table 4. Correlation between subjective sleep and mental health measures (N=38).

Outcome	1	2	3	4	5	6	7
ISI[a]	—[b]	—	—	—	—	—	—
FOSQ[c]	–0.60[d]	—	—	—	—	—	—
PSQI[e]	0.55[d]	–0.39[f]	—	—	—	—	—
PCL-5[g]	0.55[d]	–0.69[d]	0.47[d]	—	—	—	—
WHYMPI[h]	0.41[f]	–0.36[f]	0.38[f]	0.49[d]	—	—	—
PHQ-15[i]	0.55[d]	–0.51[d]	0.49[d]	0.63[d]	.64[d]	—	—
PHQ-9[j]	0.48[d]	–0.72[d]	0.43[d]	0.79[d]	.33[f]	0.55[d]	—
Mean (SD)	15.42 (4.69)	13.9 (3.31)	12.05 (4.24)	22.45 (17.11)	5.53 (4.93)	8.61 (5.13)	8.82 (5.79)
Range	7-26	7-20	3-19	3-61	0-14	2-19	1-27

[a] ISI: Insomnia Severity Index.

[b] Correlation not applicable or redundant.

[c] FOSQ: Functional Outcomes of Sleep Questionnaire.

[d] *P*<.01.

[e] PSQI: Pittsburgh Sleep Quality Index.

[f] *P*<.05.

[g] PCL-5: Posttraumatic Stress Disorder checklist for Diagnostic and Statistical Manual of Mental Disorders-5.

[h] WHYMPI: West Haven-Yale Multidimensional Pain Inventory.

[i] PHQ-15: Patient Health Questionnaire-15.

[j] PHQ-9: Patient Health Questionnaire-9.

Table 5. Secondary outcome measures at pre-, mid-, and postintervention for completers (n=11).

Outcome	Preintervention		Midintervention		Postintervention	
	Mean (SD)	Median	Mean (SD)	Median	Mean (SD)	Median
PCL-5[a]	18.45 (16.03)	14.00	—[b]	—[b]	15.18 (11.50)	11.00
PHQ-9[c]	7.64 (3.98)	7.00	6.58 (4.96)	4.00	6.18 (3.79)	6.00
PHQ-15[d]	7.91 (5.43)	6.00	7.33 (5.40)	5.50	7.64 (4.61)	6.00
WHYMPI[e]	6.73 (4.54)	7.00	6.92 (5.30)	9.00	6.36 (4.48)	6.00
PDI[f]	22.18 (17.93)	21.00	16.33 (15.55)	11.50	15.45 (13.56)	15.00

[a] PCL-5: PTSD Checklist for DSM-5 (not administered midintervention).

[b] Measure not collected midintervention.

[c] PHQ-9: Patient Health Questionnaire-9.

[d] PHQ-15: Patient Health Questionnaire-15.

[e] WHYMPI: West Haven-Yale Multidimensional Pain Inventory.

[f] PDI: Pain Disability Index.

Sleep Apnea Rates and Correlates

First or second night WatchPAT data revealed that 54.6% (18/33) of participants had possible moderate-to-severe apnea (ie, AHI score>15). These participants were withdrawn from further participation and were provided with a referral for further apnea testing. Of these participants, 89% (16/18) were male, 50% (9/18) had a BMI in the overweight (BMI=25-29 kg/m^2, per the World Health Organization) and 39% (7/18) had a BMI in the obese (BMI>30 kg/m^2) range. The mean age was 42.22 years (SD 11.58). More detailed demographic comparisons by enrollment status can be seen in Table 1. There were no significant relationships between having moderate-to-severe apnea and gender, χ^2_2=1.3, *P*=.25; BMI category, χ^2_2=3.9, *P*=.14; or age category, χ^2_2=0.2, *P*=.89; although greater BMI was significantly positively correlated with having a higher first-night AHI score, *r*=.39, *P*=.02. The number of female participants was small, likely meaning that the gender comparison was underpowered. Chi-square tests revealed that those with moderate or severe apnea were not more likely to have depressive symptoms, χ^2_1=0.7, *P*=.68, or posttraumatic symptoms, χ^2_1 =1.0, *P*=.31, than those with no or mild apnea. There was a marginal difference in nonspecific physical symptoms, such that those with moderate or severe apnea reported marginally more nonspecific physical symptoms, χ^2_1=3.3, *P*=.07. Finally, a chi-square test of independence was performed to examine the relationship between moderate-to-severe apnea and pain measures, which revealed no significant relationships—pain severity, χ^2_1=1.8, *P*=.67; PDI, χ^2_1=0.9, *P*=.81; and sleep-related functioning via the FOSQ, χ^2_1=0.5, *P*=.46.

Discussion

Principal Findings

We conducted a study of a CBTI mobile app (*CBT-i Coach*), with supplemental self-management guided use of worksheets in veterans reporting chronic insomnia and assessed changes in subjective and objective sleep measures. Veterans who completed the 6-week mobile sleep intervention reported improvements in insomnia severity, sleep quality, and sleep-related functional outcomes. No changes in objective sleep measures were observed, except that time asleep decreased, even though subjective sleep was rated as better. Improvements in subjective sleep were found on 3 common subjective sleep assessments, the ISI, PSQI, and the FOSQ. In the context of clinical guidelines, the ISI decreased from a score indicative of moderate clinical insomnia severity at pretest (≥15) to a score below the subthreshold level of insomnia at posttest (scores 8-14) [52]. A score on the PSQI of 5 or greater has been suggested as indicative of poor sleep quality, and it often prompts recommendation of further follow-up with a health care provider [43]. The PSQI did not decrease below the suggested cutoff score of 5, although it did significantly decrease. In addition, functional sleep outcomes, as measured by the FOSQ, significantly improved as well. The findings of this small pilot study support the possibility that veterans can use these technology-based tools to self-manage their chronic insomnia and experience improved sleep quality and functioning, with minimal clinical resources. Although subjective sleep measures improved in completers, objective sleep measures did not significantly improve at the end of the intervention. In fact, contrary to our hypothesis, total sleep time decreased significantly from the first night to subsequent nights (both early in the intervention and postintervention). This decrease in sleep time may be related to when the participants chose to wear the WatchPAT. Individuals typically completed their first night's sleep on the same day as their first lab visit or on a weekend to allow more time for sleep, suggesting the possibility that they completed their first night of sleep monitoring on a day when sleep time could be more easily extended; thus, it was perhaps atypical of their usual daily sleep schedule. Interestingly, there was no difference from the second to the third sleep sessions in total sleep time. It is possible that the first night was unusual for participants. As it was their first assessment with the WatchPAT, patients may have made more of a concerted effort during the first sleep session to sleep longer, to obtain a valid sleep assessment. Future research would benefit from several objective assessments to get a more accurate assessment of sleep time and possible first-night effects when using at-home sleep monitors.

Identifying Sleep Apnea Through At-Home Wearable Devices

The WatchPAT findings also revealed a relatively high prevalence of sleep apnea in this largely middle-aged, male sample of veterans. Moderate-to-severe sleep apnea was not predicted by psychological or health-related factors, except for higher BMIs being associated with high apnea scores, as assessed by AHI scores, consistent with previous studies [77-79]. More than 54% of the sample that completed at least one WatchPAT measurement showed probable moderate-to-severe sleep apnea. This was unexpectedly high as compared with epidemiologically based community samples indicating approximately 5% rates in the general population [80].

Our rate is high even when compared with other veteran samples, including a veteran epidemiological sample of those seeking health care in the VA, with a sleep apnea diagnosis rate of 4.5% [5] or an intervention sample where 38.5% of the sample screened positive for self-reported sleep apnea symptoms [19]. Apnea rates appear to have increased as prior veteran administrative data sources revealed a prevalence of diagnosed sleep apnea of only about 3.5% [48,81]. Thus, our prevalence rate for sleep apnea is considerably higher than expected, although the sample is small and self-selected; thus, the rate may not be representative of the larger population. Increasing age of the current veteran population may partially explain this increase, although Alexander et al [5] report an age-adjusted increase in apnea rates of 3.7% from 2000 to 2010. The higher prevalence rate for this study may also be because of apnea detection using a well-validated home sleep monitoring device rather than using self-reported sleep apnea symptoms or self-reported diagnoses, which is likely to lead to an underestimation of apnea in the population. Overall, these findings suggest that use of an at-home, nonobtrusive, wearable sleep device to screen for sleep apnea is not only practical, efficient, and effective but it also could be used to identify otherwise occult, and therefore untreated, sleep apnea.

Future Application of Mobile-App Supported Sleep Self-Management in the Veterans Affairs

The VA has undertaken a whole health approach to help providers focus on what matters most to their patients, and sleep is one of the main health behaviors included in these discussions. The VA is committed to the development of mobile apps to support patients with a broad range of health conditions. To further this goal, our findings suggest that use of a home sleep monitor in veterans with insomnia may also assist clinicians in detecting unrecognized cases of sleep apnea and, as a result, lead to better apnea detection and more appropriate treatment of sleep disturbance. Utilizing these technologies also may help veteran patients feel empowered to improve their sleep health, potentially become more engaged in managing their own health conditions and more invested in making positive health behavior changes. It will be useful for future studies to determine how behavioral and sleep health clinicians in the VA (and elsewhere) can better incorporate these technologies into their work with patients.

Limitations and Future Directions

Limitations of this pilot study include the small sample of intervention completers, the lack of a control group, the night-to-night variability in objective sleep measurement, the lack of follow-up data to see if the improvements in subjective sleep lasted after the intervention ended, the potential influence of social desirability, and the multicomponent nature of the intervention. The participants who withdrew of their own accord or discontinued in the study cited difficulties with continuing, including family and employment factors. This may have led to self-selection biases in the completer sample, as those individuals who completed the intervention may have been more likely to use mobile apps for health-related concerns relative to those who dropped out of the study. Among those participants who were not excluded because of sleep apnea (N=15), 73% (11/15) participants completed the intervention. Future research on the use of mobile sleep apps and at-home sleep monitors will need to consider participant factors that may lead to attrition. The small completer sample size may have impacted the effect sizes of the subjective and objective sleep findings; small sample sizes can drastically affect the ability to detect statistical power and reduce the likelihood that a statistically significant difference reflects a reproducible difference by overestimating effect size [82]. Consequently, the results here must be replicated in a larger sample before strong inferences can be made. Finally, future studies will need to utilize a control group to ensure that changes are because of the intervention effects and not because of potential demand characteristics or other nonspecific changes with the passage of time. We are currently conducting a randomized controlled 2-arm (experimental and control arms) study. Another limitation of this study includes the use of self-reported outcomes, which could have been influenced by social desirability. However, these well-validated measures were self-administered, which has been documented to reduce the effect of social desirability [83]. Finally, without a multiarm randomized controlled trial, the combined use of the *CBT-i Coach* app, self-management guide, and WatchPAT makes it impossible to attribute any changes in sleep measurements to the *CBT-i Coach* app itself. Other technologies may have emphasized the importance of paying attention to sleep patterns. Such feedback and symptom monitoring are critical components of CBTI, which likely enhanced the intervention [36].

Conclusions

This study contributes to growing literature on the efficacy of cognitive behavioral therapy–based approaches in the treatment of insomnia, suggesting that less resource-intensive treatment modalities, such as self-management–based mobile apps, can provide clinically useful tools in the management of insomnia in some patients. Furthermore, these findings align well with the call to treat sleep disturbance in veterans and minimize work-related sleep disruptions in veteran active-duty military personnel, alongside greater acknowledgment of the importance of sleep by the military [84]. The prevalence of previously undetected sleep apnea uncovered by our approach to assessment was striking, and it merits further investigation. Integrating mobile sleep apps and wearable devices constitutes a promising area for helping veterans manage their health needs, and it merits future research to inform best practices for integrating these self-management options into clinical care throughout the VA.

Acknowledgments

This material is based on work supported by the Department of VA, VHA, Office of Research and Development, Health Services Research and Development (PPO 14-144) to KSQ. The findings and interpretations of the data expressed in the paper are the sole responsibility of the authors and do not necessarily represent the views of the Department of VA.

Conflicts of Interest

None declared.

References

1. Babson KA, Blonigen DM, Boden MT, Drescher KD, Bonn-Miller MO. Sleep quality among US military veterans with PTSD: a factor analysis and structural model of symptoms. J Trauma Stress 2012 Dec;25(6):665-674. [doi: 10.1002/jts.21757] [Medline: 23225033]
2. Capaldi VF, Guerrero ML, Killgore WD. Sleep disruptions among returning combat veterans from Iraq and Afghanistan. Mil Med 2011 Aug;176(8):879-888. [doi: 10.7205/milmed-d-10-00440] [Medline: 21882777]
3. Germain A, James J, Insana S, Herringa RJ, Mammen O, Price J, et al. A window into the invisible wound of war: functional neuroimaging of REM sleep in returning combat veterans with PTSD. Psychiatry Res 2013 Feb 28;211(2):176-179 [FREE Full text] [doi: 10.1016/j.pscychresns.2012.05.007] [Medline: 23149024]
4. Lew HL, Pogoda TK, Hsu P, Cohen S, Amick MM, Baker E, et al. Impact of the 'polytrauma clinical triad' on sleep disturbance in a department of veterans affairs outpatient rehabilitation setting. Am J Phys Med Rehabil 2010 Jun;89(6):437-445. [doi: 10.1097/PHM.0b013e3181ddd301] [Medline: 20489391]
5. Alexander M, Ray MA, Hébert JR, Youngstedt SD, Zhang H, Steck SE, et al. The national veteran sleep disorder study: descriptive epidemiology and secular trends, 2000-2010. Sleep 2016 Jul 1;39(7):1399-1410 [FREE Full text] [doi: 10.5665/sleep.5972] [Medline: 27091538]
6. Caldwell JA, Knapik JJ, Lieberman HR. Trends and factors associated with insomnia and sleep apnea in all United States military service members from 2005 to 2014. J Sleep Res 2017 Dec;26(5):665-670 [FREE Full text] [doi: 10.1111/jsr.12543] [Medline: 28448689]
7. Riemann D, Voderholzer U. Primary insomnia: a risk factor to develop depression? J Affect Disord 2003 Sep;76(1-3):255-259. [doi: 10.1016/S0165-0327(02)00072-1] [Medline: 12943956]
8. Wright KM, Britt TW, Bliese PD, Adler AB, Picchioni D, Moore D. Insomnia as predictor versus outcome of PTSD and depression among Iraq combat veterans. J Clin Psychol 2011 Dec;67(12):1240-1258. [doi: 10.1002/jclp.20845] [Medline: 22065464]
9. Pigeon WR, Campbell CE, Possemato K, Ouimette P. Longitudinal relationships of insomnia, nightmares, and PTSD severity in recent combat veterans. J Psychosom Res 2013 Dec;75(6):546-550. [doi: 10.1016/j.jpsychores.2013.09.004] [Medline: 24290044]
10. Léger D, Morin CM, Uchiyama M, Hakimi Z, Cure S, Walsh JK. Chronic insomnia, quality-of-life, and utility scores: comparison with good sleepers in a cross-sectional international survey. Sleep Med 2012 Jan;13(1):43-51. [doi: 10.1016/j.sleep.2011.03.020] [Medline: 22093806]
11. Léger D, Scheuermaier K, Philip P, Paillard M, Guilleminault C. SF-36: evaluation of quality of life in severe and mild insomniacs compared with good sleepers. Psychosom Med 2001;63(1):49-55. [doi: 10.1097/00006842-200101000-00006] [Medline: 11211064]
12. Roth T, Jaeger S, Jin R, Kalsekar A, Stang PE, Kessler RC. Sleep problems, comorbid mental disorders, and role functioning in the national comorbidity survey replication. Biol Psychiatry 2006 Dec 15;60(12):1364-1371 [FREE Full text] [doi: 10.1016/j.biopsych.2006.05.039] [Medline: 16952333]
13. Katz DA, McHorney CA. The relationship between insomnia and health-related quality of life in patients with chronic illness. J Fam Pract 2002 Mar;51(3):229-235. [Medline: 11978233]
14. LeBlanc M, Beaulieu-Bonneau S, Mérette C, Savard J, Ivers H, Morin CM. Psychological and health-related quality of life factors associated with insomnia in a population-based sample. J Psychosom Res 2007 Aug;63(2):157-166. [doi: 10.1016/j.jpsychores.2007.03.004] [Medline: 17662752]
15. McAndrew LM, D'Andrea EA, Lu SE, Abbi B, Yan GW, Engel C, et al. What pre-deployment and early post-deployment factors predict health function after combat deployment?: a prospective longitudinal study of Operation Enduring Freedom (OEF)/Operation Iraqi Freedom (OIF) soldiers. Health Qual Life Outcomes 2013 Apr 30;11:73 [FREE Full text] [doi: 10.1186/1477-7525-11-73] [Medline: 23631419]

16. Quigley KS, McAndrew LM, Almeida L, D'Andrea EA, Engel CC, Hamtil H, et al. Prevalence of environmental and other military exposure concerns in Operation Enduring Freedom and Operation Iraqi Freedom veterans. J Occup Environ Med 2012 Jun;54(6):659-664. [doi: 10.1097/JOM.0b013e3182570506] [Medline: 22588478]
17. Germain A, Shear MK, Hall M, Buysse DJ. Effects of a brief behavioral treatment for PTSD-related sleep disturbances: a pilot study. Behav Res Ther 2007 Mar;45(3):627-632. [doi: 10.1016/j.brat.2006.04.009] [Medline: 16777060]
18. Swanson LM, Favorite TK, Horin E, Arnedt JT. A combined group treatment for nightmares and insomnia in combat veterans: a pilot study. J Trauma Stress 2009 Dec;22(6):639-642. [doi: 10.1002/jts.20468] [Medline: 19908322]
19. Ulmer CS, Edinger JD, Calhoun PS. A multi-component cognitive-behavioral intervention for sleep disturbance in veterans with PTSD: a pilot study. J Clin Sleep Med 2011 Feb 15;7(1):57-68 [FREE Full text] [Medline: 21344046]
20. Pigeon WR, Funderburk J, Bishop TM, Crean HF. Brief cognitive behavioral therapy for insomnia delivered to depressed veterans receiving primary care services: a pilot study. J Affect Disord 2017 Dec 1;217:105-111. [doi: 10.1016/j.jad.2017.04.003] [Medline: 28395207]
21. Edinger JD, Olsen MK, Stechuchak KM, Means MK, Lineberger MD, Kirby A, et al. Cognitive behavioral therapy for patients with primary insomnia or insomnia associated predominantly with mixed psychiatric disorders: a randomized clinical trial. Sleep 2009 Apr;32(4):499-510 [FREE Full text] [doi: 10.1093/sleep/32.4.499] [Medline: 19413144]
22. Jungquist CR, O'Brien C, Matteson-Rusby S, Smith MT, Pigeon WR, Xia Y, et al. The efficacy of cognitive-behavioral therapy for insomnia in patients with chronic pain. Sleep Med 2010 Mar;11(3):302-309 [FREE Full text] [doi: 10.1016/j.sleep.2009.05.018] [Medline: 20133188]
23. Manber R, Carney C, Edinger J, Epstein D, Friedman L, Haynes PL, et al. Dissemination of CBTI to the non-sleep specialist: protocol development and training issues. J Clin Sleep Med 2012 Apr 15;8(2):209-218 [FREE Full text] [doi: 10.5664/jcsm.1786] [Medline: 22505869]
24. Currie SR, Wilson KG, Pontefract AJ, deLaplante L. Cognitive-behavioral treatment of insomnia secondary to chronic pain. J Consult Clin Psychol 2000 Jun;68(3):407-416. [doi: 10.1037/0022-006X.68.3.407] [Medline: 10883557]
25. Vitiello MV, Rybarczyk B, von Korff M, Stepanski EJ. Cognitive behavioral therapy for insomnia improves sleep and decreases pain in older adults with co-morbid insomnia and osteoarthritis. J Clin Sleep Med 2009 Aug 15;5(4):355-362 [FREE Full text] [doi: 10.1111/j.1532-5415.2009.02376.x] [Medline: 19968014]
26. Bishop TM, Britton PC, Knox KL, Pigeon WR. Cognitive behavioral therapy for insomnia and imagery rehearsal in combat veterans with comorbid posttraumatic stress: a case series. Mil Behav Health 2016;4(1):58-64 [FREE Full text] [doi: 10.1080/21635781.2015.1100564] [Medline: 27695657]
27. Manber R, Bernert RA, Suh S, Nowakowski S, Siebern AT, Ong JC. CBT for insomnia in patients with high and low depressive symptom severity: adherence and clinical outcomes. J Clin Sleep Med 2011 Dec 15;7(6):645-652 [FREE Full text] [doi: 10.5664/jcsm.1472] [Medline: 22171204]
28. Taylor DJ, Lichstein KL, Weinstock J, Sanford S, Temple JR. A pilot study of cognitive-behavioral therapy of insomnia in people with mild depression. Behav Ther 2007 Mar;38(1):49-57. [doi: 10.1016/j.beth.2006.04.002] [Medline: 17292694]
29. Christensen H, Batterham PJ, Gosling JA, Ritterband LM, Griffiths KM, Thorndike FP, et al. Effectiveness of an online insomnia program (SHUTi) for prevention of depressive episodes (the GoodNight study): a randomised controlled trial. Lancet Psychiatry 2016 Apr;3(4):333-341. [doi: 10.1016/S2215-0366(15)00536-2] [Medline: 26827250]
30. Kang HK, Bullman TA, Smolenski DJ, Skopp NA, Gahm GA, Reger MA. Suicide risk among 1.3 million veterans who were on active duty during the Iraq and Afghanistan wars. Ann Epidemiol 2015 Feb;25(2):96-100. [doi: 10.1016/j.annepidem.2014.11.020] [Medline: 25533155]
31. Epstein DR, Babcock-Parziale JL, Haynes PL, Herb CA. Insomnia treatment acceptability and preferences of male Iraq and Afghanistan combat veterans and their healthcare providers. J Rehabil Res Dev 2012;49(6):867-878 [FREE Full text] [doi: 10.1682/JRRD.2011.06.0114] [Medline: 23299258]
32. Epstein DR, Babcock-Parziale JL, Herb CA, Goren K, Bushnell ML. Feasibility test of preference-based insomnia treatment for Iraq and Afghanistan war veterans. Rehabil Nurs 2013;38(3):120-132. [doi: 10.1002/rnj.67] [Medline: 23319338]
33. Trockel M, Manber R, Chang V, Thurston A, Taylor CB. An e-mail delivered CBT for sleep-health program for college students: effects on sleep quality and depression symptoms. J Clin Sleep Med 2011 Jun 15;7(3):276-281 [FREE Full text] [doi: 10.5664/JCSM.1072] [Medline: 21677898]
34. Schooley BL, Horan TA, Lee PW, West PA. Rural veteran access to healthcare services: investigating the role of information and communication technologies in overcoming spatial barriers. Perspect Health Inf Manag 2010 Apr 1;7:1f [FREE Full text] [Medline: 20697468]
35. Hoffman JE, Taylor K, Maneber R, Trockel M, Gehrman P, Woodward S, et al. Mobile Behavior Design Lab. 2013. CBT-I Coach (Version 1.0) URL: http://www.myvaapps.com/cbti [accessed 2018-09-21] [WebCite Cache ID 72bFnZxOT]
36. Kuhn E, Weiss BJ, Taylor KL, Hoffman JE, Ramsey KM, Manber R, et al. CBT-I Coach: a description and clinician perceptions of a mobile app for cognitive behavioral therapy for insomnia. J Clin Sleep Med 2016 Apr 15;12(4):597-606 [FREE Full text] [doi: 10.5664/jcsm.5700] [Medline: 26888586]
37. Koffel E, Kuhn E, Petsoulis N, Erbes CR, Anders S, Hoffman JE, et al. A randomized controlled pilot study of CBT-I Coach: feasibility, acceptability, and potential impact of a mobile phone application for patients in cognitive behavioral therapy for insomnia. Health Informatics J 2018 Dec;24(1):3-13. [doi: 10.1177/1460458216656472] [Medline: 27354394]

38. Manber R, Friedman L, Siebern AT, Carney C, Edinger J, Epstein D, et al. Treatment Works For Vets. 2014. Cognitive Behavioral Therapy for Insomnia in Veterans: Therapist Manual URL: https://www.treatmentworksforvets.org/wp-content/uploads/2018/04/CBT-I-Therapist-Manual.pdf [accessed 2019-07-10]
39. Lauderdale DS, Knutson KL, Yan LL, Liu K, Rathouz PJ. Self-reported and measured sleep duration: how similar are they? Epidemiology 2008 Nov;19(6):838-845 [FREE Full text] [doi: 10.1097/EDE.0b013e318187a7b0] [Medline: 18854708]
40. Edinger JD, Fins AI, Glenn DM, Sullivan RJ, Bastian LA, Marsh GR, et al. Insomnia and the eye of the beholder: are there clinical markers of objective sleep disturbances among adults with and without insomnia complaints? J Consult Clin Psychol 2000 Aug;68(4):586-593. [doi: 10.1037/0022-006X.68.4.586] [Medline: 10965634]
41. Morin CM. Measuring outcomes in randomized clinical trials of insomnia treatments. Sleep Med Rev 2003 Jun;7(3):263-279. [doi: 10.1053/smrv.2002.0274] [Medline: 12927124]
42. Carney CE, Buysse DJ, Ancoli-Israel S, Edinger JD, Krystal AD, Lichstein KL, et al. The consensus sleep diary: standardizing prospective sleep self-monitoring. Sleep 2012 Feb 1;35(2):287-302 [FREE Full text] [doi: 10.5665/sleep.1642] [Medline: 22294820]
43. Buysse DJ, Reynolds 3rd CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989 May;28(2):193-213. [doi: 10.1016/0165-1781(89)90047-4] [Medline: 2748771]
44. Bastien CH, Vallières A, Morin CM. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. Sleep Med 2001 Jul;2(4):297-307. [doi: 10.1016/S1389-9457(00)00065-4] [Medline: 11438246]
45. Germain A, Buysse DJ, Nofzinger E. Sleep-specific mechanisms underlying posttraumatic stress disorder: integrative review and neurobiological hypotheses. Sleep Med Rev 2008 Jun;12(3):185-195 [FREE Full text] [doi: 10.1016/j.smrv.2007.09.003] [Medline: 17997114]
46. Pillar G, Malhotra A, Lavie P. Post-traumatic stress disorder and sleep-what a nightmare!. Sleep Med Rev 2000 Apr;4(2):183-200. [doi: 10.1053/smrv.1999.0095] [Medline: 12531165]
47. Sharafkhaneh A, Giray N, Richardson P, Young T, Hirshkowitz M. Association of psychiatric disorders and sleep apnea in a large cohort. Sleep 2005 Nov;28(11):1405-1411. [doi: 10.1093/sleep/28.11.1405] [Medline: 16335330]
48. Sharafkhaneh A, Richardson P, Hirshkowitz M. Sleep apnea in a high risk population: a study of Veterans Health Administration beneficiaries. Sleep Med 2004 Jul;5(4):345-350. [doi: 10.1016/j.sleep.2004.01.019] [Medline: 15222990]
49. Mellman TA, Bustamante V, Fins AI, Pigeon WR, Nolan B. REM sleep and the early development of posttraumatic stress disorder. Am J Psychiatry 2002 Oct;159(10):1696-1701. [doi: 10.1176/appi.ajp.159.10.1696] [Medline: 12359675]
50. Mellman TA, Pigeon WR, Nowell PD, Nolan B. Relationships between REM sleep findings and PTSD symptoms during the early aftermath of trauma. J Trauma Stress 2007 Oct;20(5):893-901. [doi: 10.1002/jts.20246] [Medline: 17955526]
51. Husain AM, Miller PP, Carwile ST. REM sleep behavior disorder: potential relationship to post-traumatic stress disorder. J Clin Neurophysiol 2001 Mar;18(2):148-157. [doi: 10.1097/00004691-200103000-00005] [Medline: 11435805]
52. Morin CM, Belleville G, Bélanger L, Ivers H. The Insomnia Severity Index: psychometric indicators to detect insomnia cases and evaluate treatment response. Sleep 2011 May 1;34(5):601-608 [FREE Full text] [doi: 10.1093/sleep/34.5.601] [Medline: 21532953]
53. Norton MC, Tschanz JA, Fan X, Plassman BL, Welsh-Bohmer KA, West N, et al. Telephone adaptation of the Modified Mini-Mental State Exam (3MS). The Cache County Study. Neuropsychiatry Neuropsychol Behav Neurol 1999 Oct;12(4):270-276. [Medline: 10527112]
54. Carney CE, Manber R. Quiet Your Mind and Get to Sleep: Solutions to Insomnia for Those With Depression, Anxiety or Chronic Pain. Oakland, CA: New Harbinger Publications; 2009.
55. Pigeon WR, Moynihan J, Matteson-Rusby S, Jungquist CR, Xia Y, Tu X, et al. Comparative effectiveness of CBT interventions for co-morbid chronic pain & insomnia: a pilot study. Behav Res Ther 2012 Nov;50(11):685-689 [FREE Full text] [doi: 10.1016/j.brat.2012.07.005] [Medline: 22982083]
56. Chasens ER, Ratcliffe SJ, Weaver TE. Development of the FOSQ-10: a short version of the Functional Outcomes of Sleep Questionnaire. Sleep 2009 Jul;32(7):915-919 [FREE Full text] [doi: 10.1093/sleep/32.7.915] [Medline: 19639754]
57. Weaver TE, Laizner AM, Evans LK, Maislin G, Chugh D, Lyon K, et al. An instrument to measure functional status outcomes for disorders of excessive sleepiness. Sleep 1997 Oct;20(10):835-843. [doi: 10.1093/sleep/20.10.835] [Medline: 9415942]
58. Pang KP, Gourin CG, Terris DJ. A comparison of polysomnography and the WatchPAT in the diagnosis of obstructive sleep apnea. Otolaryngol Head Neck Surg 2007 Oct;137(4):665-668. [doi: 10.1016/j.otohns.2007.03.015] [Medline: 17903588]
59. Choi JH, Kim EJ, Kim YS, Choi J, Kim TH, Kwon SY, et al. Validation study of portable device for the diagnosis of obstructive sleep apnea according to the new AASM scoring criteria: Watch-PAT 100. Acta Otolaryngol 2010 Jul;130(7):838-843. [doi: 10.3109/00016480903431139] [Medline: 20082567]
60. Collop NA, Tracy SL, Kapur V, Mehra R, Kuhlmann D, Fleishman SA, et al. Obstructive sleep apnea devices for out-of-center (OOC) testing: technology evaluation. J Clin Sleep Med 2011 Oct 15;7(5):531-548 [FREE Full text] [doi: 10.5664/JCSM.1328] [Medline: 22003351]

61. Onder NS, Akpinar ME, Yigit O, Gor AP. Watch peripheral arterial tonometry in the diagnosis of obstructive sleep apnea: influence of aging. Laryngoscope 2012 Jun;122(6):1409-1414. [doi: 10.1002/lary.23233] [Medline: 22522750]
62. Yuceege M, Firat H, Demir A, Ardic S. Reliability of the Watch-PAT 200 in detecting sleep apnea in highway bus drivers. J Clin Sleep Med 2013 Apr 15;9(4):339-344 [FREE Full text] [doi: 10.5664/jcsm.2584] [Medline: 23585749]
63. Weathers FW, Litz BT, Keane TM, Palmineri PA, Marx BP, Schnurr PP. National Center for PTSD. 2013. PTSD Checklist for DSM-5 (PCL-5) URL: https://www.ptsd.va.gov/professional/assessment/adult-sr/ptsd-checklist.asp
64. Weimin L, Rongguang W, Dongyan H, Xiaoli L, Wei J, Shiming Y. Assessment of a portable monitoring device WatchPAT 200 in the diagnosis of obstructive sleep apnea. Eur Arch Otorhinolaryngol 2013 Nov;270(12):3099-3105. [doi: 10.1007/s00405-013-2555-4] [Medline: 23708441]
65. Yalamanchali S, Farajian V, Hamilton C, Pott TR, Samuelson CG, Friedman M. Diagnosis of obstructive sleep apnea by peripheral arterial tonometry: meta-analysis. JAMA Otolaryngol Head Neck Surg 2013 Dec;139(12):1343-1350. [doi: 10.1001/jamaoto.2013.5338] [Medline: 24158564]
66. Herscovici S, Pe'er A, Papyan S, Lavie P. Detecting REM sleep from the finger: an automatic REM sleep algorithm based on peripheral arterial tone (PAT) and actigraphy. Physiol Meas 2007 Feb;28(2):129-140. [doi: 10.1088/0967-3334/28/2/002] [Medline: 17237585]
67. Bresler M, Sheffy K, Pillar G, Preiszler M, Herscovici S. Differentiating between light and deep sleep stages using an ambulatory device based on peripheral arterial tonometry. Physiol Meas 2008 May;29(5):571-584. [doi: 10.1088/0967-3334/29/5/004] [Medline: 18460762]
68. Agnew Jr HW, Webb WB, Williams RL. The first night effect: an EEG study of sleep. Psychophysiology 1966 Jan;2(3):263-266. [doi: 10.1111/j.1469-8986.1966.tb02650.x] [Medline: 5903579]
69. Toussaint M, Luthringer R, Schaltenbrand N, Carelli G, Lainey E, Jacqmin A, et al. First-night effect in normal subjects and psychiatric inpatients. Sleep 1995 Jul;18(6):463-469. [doi: 10.1093/sleep/18.6.463] [Medline: 7481418]
70. Kroenke K, Spitzer RL, Williams JB. The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms. Psychosom Med 2002;64(2):258-266. [doi: 10.1097/00006842-200203000-00008] [Medline: 11914441]
71. Kroenke K, Spitzer RL. The PHQ-9: a new depression diagnostic and severity measure. Psychiatr Ann 2002;32(9):509-515 [FREE Full text]
72. Blevins CA, Weathers FW, Davis MT, Witte TK, Domino JL. The posttraumatic stress disorder checklist for DSM-5 (PCL-5): development and initial psychometric evaluation. J Trauma Stress 2015 Dec;28(6):489-498. [doi: 10.1002/jts.22059] [Medline: 26606250]
73. Bovin MJ, Marx BP, Weathers FW, Gallagher MW, Rodriguez P, Schnurr PP, et al. Psychometric properties of the PTSD checklist for diagnostic and statistical manual of mental disorders-fifth edition (PCL-5) in veterans. Psychol Assess 2016 Dec;28(11):1379-1391. [doi: 10.1037/pas0000254] [Medline: 26653052]
74. Kerns RD, Turk DC, Rudy TE. The West Haven-Yale Multidimensional Pain Inventory (WHYMPI). Pain 1985 Dec;23(4):345-356. [doi: 10.1016/0304-3959(85)90004-1] [Medline: 4088697]
75. Tait RC, Chibnall JT, Krause S. The Pain Disability Index: psychometric properties. Pain 1990 Feb;40(2):171-182. [doi: 10.1016/0304-3959(90)90068-O] [Medline: 2308763]
76. Tait RC, Pollard CA, Margolis RB, Duckro PN, Krause SJ. The Pain Disability Index: psychometric and validity data. Arch Phys Med Rehabil 1987 Jul;68(7):438-441. [doi: 10.1016/0304-3959(90)90068-O] [Medline: 3606368]
77. Tuomilehto H, Seppä J, Uusitupa M. Obesity and obstructive sleep apnea--clinical significance of weight loss. Sleep Med Rev 2013 Oct;17(5):321-329. [doi: 10.1016/j.smrv.2012.08.002] [Medline: 23079209]
78. Hoffstein V, Szalai JP. Predictive value of clinical features in diagnosing obstructive sleep apnea. Sleep 1993 Feb;16(2):118-122. [Medline: 8446830]
79. Ancoli-Israel S, Kripke DF, Klauber MR, Mason WJ, Fell R, Kaplan O. Sleep-disordered breathing in community-dwelling elderly. Sleep 1991 Dec;14(6):486-495 [FREE Full text] [doi: 10.1093/sleep/14.6.486] [Medline: 1798880]
80. Young T, Peppard PE, Gottlieb DJ. Epidemiology of obstructive sleep apnea: a population health perspective. Am J Respir Crit Care Med 2002 May 1;165(9):1217-1239. [doi: 10.1164/rccm.2109080] [Medline: 11991871]
81. Sharafkhaneh A, Sharafkhaneh H, Bredikus A, Guilleminault C, Bozkurt B, Hirshkowitz M. Effect of atrial overdrive pacing on obstructive sleep apnea in patients with systolic heart failure. Sleep Med 2007 Jan;8(1):31-36. [doi: 10.1016/j.sleep.2006.06.012] [Medline: 17157066]
82. Leon AC, Davis LL, Kraemer HC. The role and interpretation of pilot studies in clinical research. J Psychiatr Res 2011 May;45(5):626-629 [FREE Full text] [doi: 10.1016/j.jpsychires.2010.10.008] [Medline: 21035130]
83. Bowling A. Mode of questionnaire administration can have serious effects on data quality. J Public Health (Oxf) 2005 Sep;27(3):281-291. [doi: 10.1093/pubmed/fdi031] [Medline: 15870099]
84. Pedersen ER, Troxel WM, Shih RA, Pinder E, Lee D, Geyer L. Increasing resilience through promotion of healthy sleep among service members. Mil Med 2015 Jan;180(1):4-6 [FREE Full text] [doi: 10.7205/MILMED-D-14-00264] [Medline: 25562849]

Abbreviations

AHI: apnea-hypopnea index
BMI: body mass index
CBTI: cognitive behavioral therapy for insomnia
FOSQ-10: Functional Outcomes of Sleep Questionnaire-10 item
ISI: Insomnia Severity Index
PCL-5: PTSD Checklist
PDI: Pain Disability Index
PHQ-9: Patient Health Questionnaire-9
PHQ-15: Patient Health Questionnaire-15
PSQI: Pittsburgh Sleep Quality Inventory
PTSD: posttraumatic stress disorder
RDI: respiratory disturbance index
REM: rapid eye movement
VA: Veterans Affairs
VHA: Veterans Health Administration
WHYMPI: West Haven-Yale Multidimensional Pain Inventory

Edited by C Dias; submitted 03.10.18; peer-reviewed by S Speedie, J Whealin, D Bradford; comments to author 05.11.18; revised version received 21.12.18; accepted 09.04.19; published 24.07.19.

Please cite as:
Reilly ED, Robinson SA, Petrakis BA, Kuhn E, Pigeon WR, Wiener RS, McInnes DK, Quigley KS
Mobile App Use for Insomnia Self-Management: Pilot Findings on Sleep Outcomes in Veterans
Interact J Med Res 2019;8(3):e12408
URL: http://www.i-jmr.org/2019/3/e12408/
doi:10.2196/12408
PMID:31342904

Original Paper

The Presence of Physical Symptoms in Patients With Tinnitus: International Web-Based Survey

Sarah Michiels[1,2,3], PhD; Stephen Harrison[4], MSc; Markku Vesala[4], MSc; Winfried Schlee[5], PhD

[1]Department of Rehabilitation Sciences and Physiotherapy, Faculty of Medicine and Health Sciences, University of Antwerp, Wilrijk, Belgium
[2]Department of Otorhinolaryngology, Antwerp University Hospital, Edegem, Belgium
[3]Department of Translational Neurosciences, Faculty of Medicine and Health Sciences, University of Antwerp, Wilrijk, Belgium
[4]Tinnitus Hub Ltd, Hemsworth, United Kingdom
[5]Department of Psychiatry and Psychotherapy, University of Regensburg, Bezirksklinikum Regensburg, Regensburg, Germany

Corresponding Author:
Sarah Michiels, PhD
Department of Rehabilitation Sciences and Physiotherapy
Faculty of Medicine and Health Sciences
University of Antwerp
Universiteitsplein 1
Wilrijk, 2610
Belgium
Phone: 32 485840758
Email: sarah.michiels@uantwerpen.be

Abstract

Background: Tinnitus, or ringing in the ears, is a phantom perception of sound in the absence of overt acoustic stimulation. Many patients indicate that the perception of their tinnitus is not constant and can vary from moment to moment. This tinnitus fluctuation is one of the diagnostic criteria for somatosensory tinnitus (ST), a tinnitus subtype that is influenced by cervical spine or temporomandibular dysfunctions, although various factors have been reported to cause fluctuations in tinnitus, such as stress, anxiety, and physical activity.

Objective: The aim of this study was twofold: (1) to investigate the presence of physical symptoms in a large group of participants with tinnitus and (2) to investigate if these physical symptoms are more frequently present in a subgroup of participants with ST.

Methods: A Web-based survey, questioning the presence of physical symptoms in a convenience sample of participants with tinnitus, was launched on the online forum, Tinnitus Talk, managed by Tinnitus Hub. After a general analysis of the physical symptoms present in our survey population, we further analyzed the group of participants who were diagnosed by a physician (n=1262). This subgroup was divided into 2 groups, one group diagnosed with ST and another group diagnosed with other types of tinnitus.

Results: In total, 6115 participants with a mean age of 54.08 years (SD 13.8) completed the survey. Physical symptoms were frequently present in our sample of participants with tinnitus: 4221 participants (69.02%) reported some form of neck pain, 429 (7.01%) were diagnosed with temporomandibular disorders, 2730 (44.64%) indicated they have bruxism, and between 858 and 1419 (14.03%-23.20%) participants were able to modulate their tinnitus by voluntary movements. ST was diagnosed in 154 out of 1262 (12.20%) participants whose tinnitus cause was diagnosed by a physician. Symptoms referring to the known diagnostic criteria were evidently more present in the ST group than in the non-ST group. Additionally, participants with ST more often indicated a negative effect of a bad night’s sleep (*P*=.01) and light intensity exercise (*P*=.01).

Conclusions: Physical activity and movement (disorders) frequently affect tinnitus severity. Head-neck related symptoms are more frequently reported in the ST group, as is the ability to modulate the tinnitus by head or jaw movements. Additionally, participants with ST more often report fluctuations of their tinnitus and reaction to sleeping difficulties and low intensity exercise.

(Interact J Med Res 2019;8(3):e14519) doi:10.2196/14519

KEYWORDS

tinnitus; self report; surveys

Introduction

Tinnitus is the phantom perception of sound in the absence of overt acoustic stimulation [1]. The perception of tinnitus can be tonal or noise-like and is often described as hissing, sizzling, or ringing [2]. If the perception of tinnitus is ongoing for more than 6 months, the condition is considered as *chronic tinnitus*. This occurs in about 10% to 15% of adults [2] and is often related to hearing loss or a noise trauma, where cochlear abnormalities are the initial source and neural changes in the central auditory system maintain the tinnitus [2]. Furthermore, for many, the perception of tinnitus is not constant and can vary from moment to moment [3]. This fluctuation of tinnitus can depend on various factors, such as stress [4], emotional states [5], anxiety [6], depression [6], cervical spine dysfunction [7], and temporomandibular disorders (TMDs) [8], but also, physical activity [9,10] has been reported to influence the perception of tinnitus.

In this study, we were interested in the influence of physical activity on the perception of tinnitus. It has been found that accelerometer-assessed physical activity correlates negatively with tinnitus severity [9,10] and correlates positively with health-related and global quality of life [10]. In contrast, patients in clinical practice often complain about an increase in tinnitus loudness during or immediately after physical activity. Additionally, it is still unclear if these effects are present in all patients with tinnitus.

Furthermore, cervical spine and TMDs can be a risk factor for tinnitus, which is often referred as somatosensory tinnitus (ST) [7,11-13]. In this type of tinnitus, changes in somatosensory afference from the cervical spine or temporomandibular area are causing or changing the tinnitus percept. ST can be diagnosed based on the presence of a series of characteristics of the tinnitus, such as the ability to change the tinnitus by certain movements, and accompanying symptoms, such as neck pain, headache, or jaw pain [12]. However, it is not known if other physical influences are also more common in patients with ST than in patients with other types of tinnitus.

The aim of this study was therefore twofold: first to investigate the presence of physical symptoms in a large group of participants with tinnitus and second to investigate if these physical symptoms are more frequently present in a subgroup of participants with ST.

Methods

Survey

A Web-based survey, questioning the presence of physical symptoms in a convenience sample of participants with tinnitus, was launched on the online forum, Tinnitus Talk, managed by Tinnitus Hub, in February 2017. The idea of the survey topic was conceived from talking to patients and moderating the online forum. Questions were designed on consultation with tinnitus researchers and a small pool of the forum's community and trialed with this group before launch. This was done to make sure that all questions were clear and unambiguous and that no technical issues were present. The final questionnaire consisted of 21 questions that asked for different physical symptoms that can accompany the tinnitus and for a set of tinnitus characteristics. The complete list of questions is displayed in Multimedia Appendix 1.

The survey was advertised on the Tinnitus Hub website and launched as a closed survey, open to all registered users of the online forum, Tinnitus Talk. No incentives were offered to participants. An internet protocol check was used to identify and block potential duplicate entries from the same user. All participants gave informed consent to use their anonymized data. No personal information was collected during the process.

Data Analysis

General Characteristics and Physical Symptoms

Participant characteristics and the presence of physical symptoms in the entire group were analyzed using descriptive statistics and frequencies.

Somatosensory Tinnitus

Participants who indicated in question 6 that their tinnitus cause was diagnosed by a physician were separated from the rest of the patients. The group that was diagnosed by a physician was then used to create 2 subgroups: one diagnosed with ST (ST group) and another with other tinnitus diagnoses (non-ST group). Differences between both groups were analyzed using Fisher exact tests in case of dichotomous variables and via independent sample t tests for continuous variables. Correction for multiple comparison was made with the Benjamini-Hochberg false discovery rate procedure, using a false discovery rate of 5%. In the Results section, only the corrected Benjamini-Hochberg P values are presented. The significance level was set at P less than .05.

Only complete questionnaires, without missing data, were used for the analysis. All analyses were performed using IBM SPSS Statistics for Macintosh (version 25.0; IBM Corporation).

Results

General Characteristics

A group of 6115 participants (51.60% male participants, 47.80% female participants, and 0.30% transgender participants) with a mean age of 54.08 years (SD 13.8) filled out the Web-based survey (details are presented in Table 1). Participants originated from 62 different countries with the highest percentages from the United Kingdom (25.30%) and the United States (40.40%). In 4157 (67.90%) participants, the primary tinnitus cause corresponded to one of the causes listed in the questionnaire (Table 2) and 1618 (26.50%) participants indicated an unknown cause of their tinnitus. In all other cases, the cause of the tinnitus was marked as *not listed*. The most frequently mentioned causes were noise-induced hearing loss (748 or 12.20%) and noise trauma (624 or 10.20%). ST was identified as the primary cause of the tinnitus in 409 out of 6115 (6.60%) participants and as a secondary cause in 16.30%. Psychological problems, such as stress, anxiety, and depression, were listed as the primary cause in 4.40% of the participants and in 12.90% (996/6115) as a secondary cause.

Table 1. Demographics.

Variable	Value
Subjects, n	6115
Gender, n	
Male	3154
Female	2925
Transgender	17
Age (years), mean (SD)	54.07 (13.81)
VAS[a] loudness, mean (SD)	5.62 (1.92)

[a] VAS: visual analogue scale.

Table 2. Tinnitus causes (N=6115).

Primary tinnitus cause	Value, n (%)
Age-related hearing loss	250 (4.09)
Allergy to something	28 (0.46)
Barotrauma	65 (1.06)
Dental treatment	51 (0.83)
Ear wax build up	36 (0.59)
Ear wax procedure	69 (1.13)
Eustachian tube dysfunction	108 (1.77)
Head or neck injury	217 (3.55)
Menière’s disease	170 (2.78)
Metabolic disease	55 (0.90)
Noise induced hearing loss	748 (12.23)
Noise trauma	624 (10.20)
Otosclerosis	48 (0.78)
Ototoxic medication	304 (4.97)
Psychological	269 (4.40)
Sudden hearing loss	237 (3.88)
Temporomandibular dysfunction	141 (2.31)
Virus	476 (7.78)
Unknown	1618 (26.46)
Cause not listed	340 (5.56)

The duration of tinnitus ranges from very recent to more than 30 years with an average of 6.26 years (SD 6.81). The average tinnitus loudness over the last week on the visual analogue scale (VAS) was 5.62 (SD 1.92). The average tinnitus annoyance over the last week on the VAS was 5.12 (SD 2.24).

A total of 1763 (62.80%) participants were diagnosed with moderate-to-severe hearing loss and 3976 (65.40%) reported to experience a negative effect on their tinnitus from emotional stress and 3756 (62.00%) from anxiety.

Presence of Physical Symptoms

In our group of 6115 participants, 4221 (69.00%) pointed out that they regularly get neck pain, although only 649 (10.60%) participants have an actual medical condition related to this neck pain (Multimedia Appendix 2). Only 429 out of 6115 (7.00%) participants were diagnosed with TMDs, although 668 (10.90%) complained about jaw pain, 517 (8.50%) about a tired feeling in the jaw muscles, and 1126 (18.40%) about clicking of the jaw. Additionally, 2730 (44.60%) participants have bruxism and 2325 (38.00%) complain about frequent headache episodes.

Clinically, participants often complain about changes in their tinnitus pitch or loudness, which often makes it hard to keep the tinnitus out of the conscious perception. In our group, 4250 participants (69.50%) report daily fluctuations of their tinnitus and 2732 (44.70%) have fleeting episodes at least once a month. Apart from tinnitus changes after sound exposure (63.50%), tinnitus can also be modulated by different movements of the head, neck, or jaw or by pressure on different parts of the body. In our study, the largest group reported that they could modulate their tinnitus by clenching their teeth (1419 or 23.20%) or pushing their jaw outward (1304 or 21.30%). More details can be found in Multimedia Appendix 3.

With regard to the short-term influence of physical activity on tinnitus, we found that 902 participants (14.80%) reported positive effects of light exercise, whereas only 496 (8.10%) indicated a positive effect of an intense workout on their tinnitus. On the contrary, 1195 (19.50%) complained about a negative influence of an intense workout on their tinnitus, compared with only 574 (9.40%) that complained after light exercise (Multimedia Appendix 4).

Somatosensory Tinnitus

The cause of the tinnitus was diagnosed by a physician in 1262 participants (20.60%). For the following part of the results, we only used these 1262 participants who were divided into 2 groups based on the diagnosis: one group of participants diagnosed with ST (ST group) and another group with other diagnoses (non-ST group).

In 154 out of 1262 participants (12.20%), ST was diagnosed as either the primary (5.80%) or secondary (11.90%) cause of the tinnitus. The results of the comparison of the ST and non-ST group are shown in Table 3 and Multimedia Appendices 5-7. No significant differences in age and gender were found between both groups (*P*=.61). With regard to the tinnitus pitch, 61 out of 154 (42.40%) participants of the ST group describe their tinnitus as a mixture of tones, compared with 367 out of 1108 (32.10%) in the non-ST group. In the ST group, a significantly higher percentage reported that his or her tinnitus varied during the day. A feeling of fullness in the ears after activity was also more present in the ST group. With regard to the ability to modulate the tinnitus by head or jaw movements, all types of modulation are significantly more often present in the ST group, except for the modulation by pushing the jaw outward.

Head-neck–related symptoms are evidently more frequently reported in the ST group. For instance, participants with ST indicated more often that they suffered from stiff or sore neck muscles and from headache. Similarly, TMD symptoms are more often present in the ST group. Significant differences were found for the presence of bruxism and pain or discomfort in the jaw.

Table 3. Presence of diagnostic criteria for somatosensory tinnitus in participants with and without somatosensory tinnitus.

Characteristics referring to DC-ST[a]	ST[b] group (n=154), %	Non-ST group (n=1108), %	Corrected *P* value
Neck pain from medical condition	25.00	8.00	<.001
Restricted neck movement	27.00	13.00	<.001
Headaches	60.00	33.00	<.001
Bruxism	66.00	44.00	<.001
Tinnitus modulation while pressing jaw	33.00	16.00	<.001
Tinnitus modulation while pushing jaw backward	28.00	16.00	<.001
Tinnitus modulation while looking up	33.00	16.00	<.001
Jaw pain	28.00	9.00	<.001
Jaw blockage	10.00	1.00	<.001
Jaw tired feeling	21.00	7.00	<.001
Tense jaw muscles	36.00	11.00	<.001
Jaw clicking	38.00	14.00	<.001
TMD[c] diagnosis	33.00	4.00	<.001
Neck pain from bad posture	30.00	19.00	.01
Tinnitus modulation while pushing head forward against resistance	28.00	18.00	.01
Tinnitus changes during the day	81.00	70.00	.01
Neck pain after physical activity	22.00	13.00	.01
Tinnitus modulation while clenching teeth	34.00	24.00	.01
Tinnitus modulation while pushing jaw outwards	30.00	23.00	.08
Neck pain from lying in bed	24.00	20.00	.33

[a] DC-ST: diagnostic criteria for somatosensory tinnitus.

[b] ST: somatosensory tinnitus

[c] TMD: temporomandibular disorder.

Additionally, participants diagnosed with ST more often reported a negative effect of a bad night's sleep on their tinnitus. Negative effects of an intense workout and moderate exercise were present in 265 and 239 out of 1262 (20.00%-25.00%) participants, respectively, but no significant differences were found between both groups. The negative influence of light exercises, on the contrary, was more often present in the ST group.

Discussion

Principal Findings

The aim of this study was twofold: first to investigate the presence of physical symptoms in a large group of participants with tinnitus and second to investigate if these physical symptoms are more frequently present in a subgroup of participants with ST.

In general, our study population is similar to tinnitus populations in other studies with regard to age and average tinnitus loudness and annoyance [7,14-16].

In a study sample of 6115 participants, we found that physical symptoms are frequently present in participants with tinnitus: 69.00% of the tinnitus participants reported some form of neck pain. This number falls within the range of lifetime prevalence numbers for nonspecific neck pain [17], but it is still higher than the average lifetime prevalence of 48.50%. For comparison, the occurrence for headache in 38.00% of the cases corresponds roughly to the prevalence of an active headache disorder in the general population [18].

In our study, 7.00% of the participants were diagnosed with TMD and 10.90% complained about jaw pain, which corresponds to the prevalence of jaw pain in the general adult population [19]. However, we would have expected this prevalence to be higher, as tinnitus is a very common symptom in participants with TMD. In the literature, tinnitus prevalence between 30.40% and 64.00% are reported in the TMD population [8,20,21]. Jaw muscle tightness and jaw clicking, which are symptoms of TMD, were reported by 14.80% and 18.40% of our study population, respectively, and bruxism, a parafunction that can lead to TMD, was present in 44.60%. One explanation for the low number of TMD diagnosis in our study might be that, not everyone with TMD symptoms visited a health care provider for his or her symptoms, resulting in an underestimation of the actual number of participants with TMDs. Another explanation might be that the studies reporting the strong association between tinnitus and TMDs are always situated in a primary TMD population. Therefore, it is hard to tell if all these participants with TMDs who also perceive tinnitus are actually bothered by their tinnitus. Our results from a primary tinnitus population, where the prevalence of TMDs is not higher than in the general population, strongly suggest this statement.

Daily fluctuations of the tinnitus pitch or loudness are reported by 69.50% of our participants. These fluctuations are known to be very typical in patients with ST and they are included in the list of diagnostic criteria for ST [12]. Our study also pointed out that fluctuations are more often present in the ST group but the large percentage of participants reporting daily fluctuations in the non-ST group, however, suggests that this item should not be used as a single criterion for diagnosing ST. More specifically, 14.00% to 23.10% of our participants were able to modulate their tinnitus by one of the listed voluntary movements. Additionally, tinnitus modulation was more often reported in the ST group than in the non-ST group, although it must be noted that only 28.00% to 34.00% of the participants in the ST group reported tinnitus modulation. Therefore, the absence of tinnitus modulation should never be the only reason to exclude ST diagnosis [12].

As mentioned in the Introduction section, physical activity can positively influence tinnitus, as especially moderate physical activity reduces stress levels [22]. In our population, only 11.50% reported a positive effect of moderate exercise, whereas 18.30% contrarily reported a negative influence. Generally, our results show that the higher the intensity of the workout, the higher the percentage of participants who perceive a negative influence and the lower the percentage who experience a positive effect. These findings are in contrast with the findings of Carpenter-Thompson et al [10], who stated that higher levels of physical activity were significantly associated with lower levels of tinnitus severity on the Tinnitus Functional Index. This contradiction can be explained by the temporary increase in blood pressure associated with higher physical activity and the fact that tinnitus is associated with arterial hypertension [23-25]. On the contrary, physical activity shows positive long-term effects on arterial hypertension [26]. Despite the large sample size, our study is based on retrospective, self-reported data, which might have influenced the results. More prospective research is needed to investigate the relationship of physical exercise and its impact on tinnitus.

Of the 6115 included patients, 22.90% indicated that they had some form of ST. This corresponds to the 26.50% prevalence of ST in the tinnitus clinic of the Antwerp University Hospital (unpublished data) and to data from a cohort study of Ward et al [27] in the United Kingdom. Other studies have mentioned higher prevalences of ST, ranging from 43.00% to 83.00% depending on the used diagnostic criteria [7,28]. In 2018, a new set of diagnostic criteria for ST was agreed on by an international group of tinnitus experts [12]. As these criteria were not yet available when the survey was launched, it is not entirely clear how the ST diagnosis was made in our survey and some underdiagnosing might be present.

Items such as the patient's ability to modulate the tinnitus are often used for diagnosing ST, which reflects clearly in the higher percentages of modulation in the ST group, but it must be noted that tinnitus modulation is also present in 16.00% to 24.00% of the non-ST group. Other criteria that might be used for diagnosing ST, such as the presence of neck pain, headache, or temporomandibular joint problems, are similarly more often present in the ST group although small percentages are present in the non-ST group. One very specific criterion from the new list, *Tinnitus is reported to vary*, was also more often present in the ST group (81.00%), but it is also present in 70.00% of the non-ST group. Some caution with the use of this criterion is therefore needed. Additionally, patients in the ST group more often describe their tinnitus as a mixture of tones.

Interestingly we found a significantly higher percentage of participants in the ST group who complained about the "feeling of fullness in the ears after activity." This symptom has, to our knowledge, never been described as typical for patients with ST. On the contrary, fullness in the ears has been described as a common symptom in patients with TMDs [29]. Further analysis of our data showed that the participants who indicate they have a "feeling of fullness in the ears after activity" significantly more often suffer from pain or dysfunction of the jaw.

As mentioned earlier, physical activity, and especially the higher intensity physical activity, can have a negative effect on tinnitus severity. No significant differences were found between ST and non-ST groups for the negative effect of high or moderate intensity physical activity. However, we did find significantly more participants in the ST group indicating a negative effect of light intensity exercise.

Finally, participants in our ST group reported significantly more often a negative effect of a bad night's sleep compared with the non-ST group. These findings are logical if we consider that not only tinnitus is affected by sleeping difficulties but also neck and jaw pain that are causing the somatosensory influence in the ST group [30,31].

Limitations

Our study has 1 major limitation: the self-reported nature of the data, which is of course inextricably linked to survey-based studies. Patients might, for instance, not always remember the exact diagnosis the physician made, especially when the consult was years ago. Additionally, some items, such as "the effect of exercise on tinnitus severity" might be unknown in a sedentary population, resulting in an underestimation of the presence of this symptom. Additionally, the survey was advertised as investigating physical links in tinnitus. This might have introduced some bias, as people who experienced some physical influence on their tinnitus in the past would be extra motivated to participate. However, the large sample already corrects largely for this potential bias. Despite these limitations, our study was able to identify a set of physical symptoms that are often present in participants with tinnitus. Future studies should aim to investigate these symptoms in a more controlled and detailed way to draw any definitive conclusions.

Conclusions

Physical activity and movement (disorders) frequently affect tinnitus severity. Head-neck–related symptoms are more frequently reported in the ST group, as is the ability to modulate the tinnitus by head or jaw movements. Additionally, participants with ST more often report fluctuations of their tinnitus and reaction to sleeping difficulties and low-intensity exercise. Large prospective cohort studies are needed to confirm these findings and to address the limitations of this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Survey questions.

[DOCX File, 19KB - ijmr_v8i3e14519_app1.docx]

Multimedia Appendix 2

Head and neck dysfunctions (in % of N=6115). TMD: Temporomandibular disorder.

[PNG File, 101KB - ijmr_v8i3e14519_app2.png]

Multimedia Appendix 3

Changes in tinnitus pitch or loudness (in % of N=6115).

[PNG File, 105KB - ijmr_v8i3e14519_app3.png]

Multimedia Appendix 4

Short-term effect of physical activity on tinnitus severity (in % of N=6115).

[PNG File, 87KB - ijmr_v8i3e14519_app4.png]

Multimedia Appendix 5

Presence of hearing-related characteristics in participants with and without somatosensory tinnitus.

[DOCX File, 15KB - ijmr_v8i3e14519_app5.docx]

Multimedia Appendix 6

Different tinnitus sounds in participants with and without somatosensory tinnitus.

[DOCX File, 15KB - ijmr_v8i3e14519_app6.docx]

Multimedia Appendix 7

Negative effects of stress, anxiety, and physical activity in participants with and without somatosensory tinnitus.

[DOCX File, 14KB - ijmr_v8i3e14519_app7.docx]

References

1. Landgrebe M, Azevedo A, Baguley D, Bauer C, Cacace A, Coelho C, et al. Methodological aspects of clinical trials in tinnitus: a proposal for an international standard. J Psychosom Res 2012 Aug;73(2):112-121 [FREE Full text] [doi: 10.1016/j.jpsychores.2012.05.002] [Medline: 22789414]
2. Baguley D, McFerran D, Hall D. Tinnitus. Lancet 2013 Nov 09;382(9904):1600-1607 [FREE Full text] [doi: 10.1016/S0140-6736(13)60142-7] [Medline: 23827090]
3. Schlee W, Pryss RC, Probst T, Schobel J, Bachmeier A, Reichert M, et al. Measuring the moment-to-moment variability of tinnitus: the TrackYourTinnitus smart phone app. Front Aging Neurosci 2016;8:294 [FREE Full text] [doi: 10.3389/fnagi.2016.00294] [Medline: 28018210]
4. Mazurek B, Szczepek AJ, Hebert S. Stress and tinnitus. HNO 2015 Apr;63(4):258-265. [doi: 10.1007/s00106-014-2973-7] [Medline: 25862619]
5. Probst T, Pryss R, Langguth B, Schlee W. Emotional states as mediators between tinnitus loudness and tinnitus distress in daily life: Results from the "TrackYourTinnitus" application. Sci Rep 2016 Feb 08;6:20382 [FREE Full text] [doi: 10.1038/srep20382] [Medline: 26853815]
6. Bhatt JM, Bhattacharyya N, Lin HW. Relationships between tinnitus and the prevalence of anxiety and depression. Laryngoscope 2017 Dec;127(2):466-469 [FREE Full text] [doi: 10.1002/lary.26107] [Medline: 27301552]
7. Michiels S, De Hertogh W, Truijen S, Van de Heyning P. Cervical spine dysfunctions in patients with chronic subjective tinnitus. Otol Neurotol 2015 Apr;36(4):741-745. [doi: 10.1097/MAO.0000000000000670] [Medline: 25415466]
8. Buergers R, Kleinjung T, Behr M, Vielsmeier V. Is there a link between tinnitus and temporomandibular disorders? J Prosthet Dent 2014 Mar;111(3):222-227. [doi: 10.1016/j.prosdent.2013.10.001] [Medline: 24286640]
9. Loprinzi PD, Lee H, Gilham B, Cardinal BJ. Association between accelerometer-assessed physical activity and tinnitus, NHANES 2005-2006. Res Q Exerc Sport 2013 Jun;84(2):177-185. [doi: 10.1080/02701367.2013.784840] [Medline: 23930543]
10. Carpenter-Thompson JR, McAuley E, Husain FT. Physical activity, tinnitus severity, and improved quality of life. Ear Hear 2015;36(5):574-581. [doi: 10.1097/AUD.0000000000000169] [Medline: 25906172]
11. Michiels S, Van de Heyning P, Truijen S, De Hertogh W. Diagnostic value of clinical cervical spine tests in patients with cervicogenic somatic tinnitus. Phys Ther 2015 Nov;95(11):1529-1535. [doi: 10.2522/ptj.20140457] [Medline: 26045606]
12. Michiels S, Ganz Sanchez T, Oron Y, Gilles A, Haider HF, Erlandsson S, et al. Diagnostic criteria for somatosensory tinnitus: a delphi process and face-to-face meeting to establish consensus. Trends Hear 2018;22:2331216518796403 [FREE Full text] [doi: 10.1177/2331216518796403] [Medline: 30213235]
13. Ralli M, Greco A, Boccassini A, Altissimi G, Di Paolo C, Falasca V, et al. Subtyping patients with somatic tinnitus: modulation of tinnitus and history for somatic dysfunction help identify tinnitus patients with temporomandibular joint disorders. PLoS One 2018;13(8):e0202050 [FREE Full text] [doi: 10.1371/journal.pone.0202050] [Medline: 30102717]
14. Michiels S, Naessens S, Van de Heyning P, Braem M, Visscher CM, Gilles A, et al. The effect of physical therapy treatment in patients with subjective tinnitus: a systematic review. Front Neurosci 2016;10:545 [FREE Full text] [doi: 10.3389/fnins.2016.00545] [Medline: 27965530]
15. Kreuzer PM, Lehner A, Schlee W, Vielsmeier V, Schecklmann M, Poeppl TB, et al. Combined rTMS treatment targeting the anterior cingulate and the temporal cortex for the treatment of chronic tinnitus. Sci Rep 2015 Dec 15;5:18028 [FREE Full text] [doi: 10.1038/srep18028] [Medline: 26667790]
16. Jacquemin L, Shekhawat GS, Van de Heyning P, Mertens G, Fransen E, Van Rompaey V, et al. Effects of electrical stimulation in tinnitus patients: conventional versus high-definition tDCS. Neurorehabil Neural Repair 2018 Dec;32(8):714-723. [doi: 10.1177/1545968318787916] [Medline: 30019630]
17. Fejer R, Kyvik KO, Hartvigsen J. The prevalence of neck pain in the world population: a systematic critical review of the literature. Eur Spine J 2006 Jun;15(6):834-848 [FREE Full text] [doi: 10.1007/s00586-004-0864-4] [Medline: 15999284]
18. Stovner L, Hagen K, Jensen R, Katsarava Z, Lipton R, Scher A, et al. The global burden of headache: a documentation of headache prevalence and disability worldwide. Cephalalgia 2007 Mar;27(3):193-210. [doi: 10.1111/j.1468-2982.2007.01288.x] [Medline: 17381554]

19. Goulet JP, Lavigne GJ, Lund JP. Jaw pain prevalence among French-speaking Canadians in Québec and related symptoms of temporomandibular disorders. J Dent Res 1995 Nov;74(11):1738-1744. [doi: 10.1177/00220345950740110401] [Medline: 8530734]
20. Manfredini D, Olivo M, Ferronato G, Marchese R, Martini A, Guarda-Nardini L. Prevalence of tinnitus in patients with different temporomandibular disorders symptoms. Int Tinnitus J 2015;19(2):47-51. [doi: 10.5935/0946-5448.20150008] [Medline: 27186932]
21. Lam DK, Lawrence HP, Tenenbaum HC. Aural symptoms in temporomandibular disorder patients attending a craniofacial pain unit. J Orofac Pain 2001;15(2):146-157. [Medline: 11443826]
22. Fleshner M. Physical activity and stress resistance: sympathetic nervous system adaptations prevent stress-induced immunosuppression. Exerc Sport Sci Rev 2005 Jul;33(3):120-126. [Medline: 16006819]
23. Figueiredo RR, de Azevedo AA, Penido ND. Tinnitus and arterial hypertension: a systematic review. Eur Arch Otorhinolaryngol 2015 Nov;272(11):3089-3094. [doi: 10.1007/s00405-014-3277-y] [Medline: 25190255]
24. Palatini P, Mormino P, Mos L, Di Marco A, Munari L, Del Torre M, et al. Blood pressure changes during physical exercise (the beat phenomenon). J Hypertens Suppl 1988 Dec;6(4):S88-S90. [Medline: 3241285]
25. Palatini P, Mos L, Mormino P, Di Marco A, Munari L, Fazio G, et al. Blood pressure changes during running in humans: the "beat" phenomenon. J Appl Physiol (1985) 1989 Jul;67(1):52-59. [doi: 10.1152/jappl.1989.67.1.52] [Medline: 2759978]
26. Diaz KM, Shimbo D. Physical activity and the prevention of hypertension. Curr Hypertens Rep 2013 Dec;15(6):659-668 [FREE Full text] [doi: 10.1007/s11906-013-0386-8] [Medline: 24052212]
27. Ward J, Vella C, Hoare DJ, Hall DA. Subtyping somatic tinnitus: a cross-sectional UK cohort study of demographic, clinical and audiological characteristics. PLoS One 2015;10(5):e0126254 [FREE Full text] [doi: 10.1371/journal.pone.0126254] [Medline: 25996779]
28. Abel MD, Levine RA. Muscle contractions and auditory perception in tinnitus patients and nonclinical subjects. Cranio 2004 Jul;22(3):181-191. [doi: 10.1179/crn.2004.024] [Medline: 15293775]
29. Peng Y. Temporomandibular joint disorders as a cause of aural fullness. Clin Exp Otorhinolaryngol 2017 Sep;10(3):236-240 [FREE Full text] [doi: 10.21053/ceo.2016.01039] [Medline: 28103655]
30. Staniszewski K, Lygre H, Bifulco E, Kvinnsland S, Willassen L, Helgeland E, et al. Temporomandibular disorders related to stress and HPA-axis regulation. Pain Res Manag 2018;2018:7020751 [FREE Full text] [doi: 10.1155/2018/7020751] [Medline: 29854038]
31. Ortego G, Villafañe JH, Doménech-García V, Berjano P, Bertozzi L, Herrero P. Is there a relationship between psychological stress or anxiety and chronic nonspecific neck-arm pain in adults? A systematic review and meta-analysis. J Psychosom Res 2016 Dec;90:70-81. [doi: 10.1016/j.jpsychores.2016.09.006] [Medline: 27772562]

Abbreviations

DC-ST: diagnostic criteria for somatosensory tinnitus
ST: somatosensory tinnitus
TMD: temporomandibular disorder
VAS: visual analogue scale

Edited by G Eysenbach; submitted 29.04.19; peer-reviewed by MS Aslam, D Carvalho; comments to author 26.05.19; revised version received 03.06.19; accepted 12.06.19; published 30.07.19.

Please cite as:
Michiels S, Harrison S, Vesala M, Schlee W
The Presence of Physical Symptoms in Patients With Tinnitus: International Web-Based Survey
Interact J Med Res 2019;8(3):e14519
URL: http://www.i-jmr.org/2019/3/e14519/
doi:10.2196/14519
PMID:31364603

Original Paper

What the Health? Information Sources and Maternal Lifestyle Behaviors

Emilie Mass Dalhaug[1], MSc; Lene Annette Hagen Haakstad[1], MSc, PhD

The Norwegian School of Sport Sciences, Department of Sports Medicine, Oslo, Norway

Corresponding Author:
Emilie Mass Dalhaug, MSc
The Norwegian School of Sport Sciences
Department of Sports Medicine
PO Box 4014, Ullevål Stadion
Oslo, 0806
Norway
Phone: 47 23 26 20 00
Fax: 47 22 23 42 20
Email: emiliefmass@gmail.com

Abstract

Background: Regular physical activity (PA), adequate gestational weight gain (GWG), and healthy eating are important for the long-term health of both mother and baby. Hence, it is important that women receive current and updated advice on these topics and are encouraged to adopt a healthy lifestyle during pregnancy.

Objective: The aim of this study was to investigate the main information sources among pregnant women regarding PA, GWG, and nutrition as well as to evaluate how these information sources may affect their health behaviors.

Methods: A cross-sectional study design, comprising an electronic questionnaire, was distributed to 2 antenatal clinics, as well as pregnancy-related online chat forums and social media. The inclusion criteria were ≥18 years, ≥20 weeks gestation, and able to read and write Norwegian. In total, 150 pregnant women answered the questionnaire, which was a mix of 11-point Likert scales, close-ended questions, and semi–close-ended questions with the option to elaborate. The relationship between information sources and selected variables, including health behaviors and descriptive variables, were assessed by logistic regression, linear regression, or chi-square as appropriate ($P<.05$).

Results: Mean age (years), gestation week, and prepregnancy body mass index (kg/m^2) were 31.1 (SD 4.3), 30.6 (SD 5.9), and 24.2 (SD 4.2), respectively. More than eight out of 10 had received or retrieved information about nutrition (88.7%, 133/150) and PA (80.7%, 121/150), whereas 54.0% (81/150) reported information on GWG. When combining all 3 lifestyle factors, 38.5% had retrieved information from blogs and online forums and 26.6%, from their midwife or family physician. Women who reported the internet and media as their primary source of information on weight gain had increased odds of gaining weight below the Institute of Medicine (IOM) guidelines compared with gaining within the guidelines (odds ratio [OR] 15.5, 95% CI 1.4-167.4; P=.02). Higher compliance with nutritional guidelines was seen among those who cited the internet and media as their main source of information on nutrition (beta=.7, 95% CI 0.07-1.3; P=.03). On the other side, receiving advice from friends and family on weight gain was significantly associated with gaining weight above the IOM guidelines compared with gaining within the guidelines (OR 12.0, 95% CI 1.3-111.7; P=.03). No other associations were found between information sources and health behaviors.

Conclusions: The small number of health professionals giving information and the extensive use of internet- and media-based sources emphasize the need to address the quality of internet advice and guide women toward trustworthy sources of information during pregnancy. The association between information sources and PA, GWG, and nutrition requires further research.

(Interact J Med Res 2019;8(3):e10355) doi:10.2196/10355

KEYWORDS

pregnancy; physical activity; gestational weight gain; diet; prenatal care; behavior

Introduction

Regular physical activity (PA), adequate gestational weight gain (GWG), and healthy eating may lower the incidence and severity of serious conditions associated with pregnancy, including gestational diabetes mellitus [1-3], pregnancy-induced hypertension [2], preterm birth [4], macrosomia [1,4], and small for gestational age infants [4,5]. In addition, daily exercise for the pelvic floor muscles may prevent and treat urinary incontinence [6]. Hence, it is important that women receive current and updated advice on these topics and are encouraged to adopt a healthy lifestyle during pregnancy [7].

Research show that pregnant women retrieve health information from a variety of sources, including the internet, books, family and friends, parenting magazines, blogs, online forums and health professionals [8-10]. Of these, the internet and media and health professionals are often cited as the most helpful and informative sources of information [8,10]. On the contrary, a meta-analysis showed that health websites often lacked accuracy and that it was difficult to find high quality sites [11]. Others have found that pregnant women may perceive advice from family physicians and midwives confusing, vague, contradictory, and frequently changing [9,12-14]. In addition, research shows that most health care providers, regardless of medical training, lack knowledge and awareness of the American College of Obstetricians and Gynecologists [7] PA guidelines [13,15] and the Institute of Medicine (IOM) [16] GWG guidelines [15,17,18]. Furthermore, the majority of pregnant women being counselled about weight gain report that the advice is generally discordant with the guidelines [8,13,19,20].

In previous studies investigating pregnant women's information sources on PA, GWG, and/or nutrition [8-10,13,14,19,21-27], the population sizes were generally small (N=17-60) [9,13,14,21,23,26] and information sources were rarely the main outcome [9,13,19,22-26]. In addition, studies of moderate methodological quality and sufficient population sizes (N=350-368) have recruited women in the postpartum period, limiting the results to the women's memory [10,27], or halfway through pregnancy (mean 20.8 weeks) [8]. Hence, there is limited evidence investigating pregnant women's information sources regarding PA, GWG, and nutrition at late gestation. To our knowledge, this study is also the first to evaluate how different information sources may influence PA, GWG, and nutritional habits among pregnant women. Thus, the aim of this study was to investigate the main information sources among pregnant women in Scandinavia with regard to PA, GWG, and nutrition and evaluate how these information sources may affect their health behaviors.

Methods

Study Design

The present project was a cross-sectional study conducted in Oslo, Norway, from February to August 2016. Pregnant women were asked to fill an electronic questionnaire investigating their health behaviors, as well as information sources with regard to PA, GWG, and nutrition. The study was reviewed by the Regional Committee for Medical and Health Research Ethics (REK 2015/1941 A), who concluded that, according to the act on medical and health research (the Health Research Act 2008), the study did not require full review by REK. The study was approved by the Norwegian Social Science Data Service (NSD 45111).

Participants

Enrolment was limited to women living in Oslo, ≥18 years, ≥20 weeks gestation, and being able to read and write Norwegian.

Procedures

To ensure a representative sample with regard to different ethnicities, age groups, and socioeconomic backgrounds, all antenatal clinics in Oslo (n=18), both urban and rural, were invited to participate. However, due to ongoing research projects only 2 agreed to distribute questionnaires to their pregnant patients. Hence, we needed to recruit participants from other arenas. We chose to spread the link to the online survey through advertisements on Facebook and Instagram, as well as through various pregnancy-related online chat forums and the university website. The advertisement on Facebook and Instagram was not limited to pregnant women but targeted women living in Oslo. The internet-based questionnaire was active between June 1 and August 15, 2016.

Outcome Measures

The survey items used to answer our research questions were derived from a multidimensional survey that assessed pregnant women's information sources, PA level, nutritional habits, and GWG. The survey also investigated social support, motives, and barriers for being physically active, as well as pregnancy complaints and quality of life. The multidimensional electronic survey contained 101 questions and was developed using existing and validated questions [28-30], as well as some newly developed questions suitable to the purpose of this study. The current analysis focused on pregnant women's information sources and their lifestyle behaviors. Questions were a mix of 11-point Likert scales, close-ended questions, and semi–close-ended questions with the option to elaborate. The questionnaire was piloted for comprehensibility of questions and answer options among 23 pregnant women and was revised accordingly. Below are the questions used to answer this study's research questions (Table 1). A full questionnaire list in Norwegian may be provided upon request.

Table 1. Dimensions assessed and main variables and questions used to answer the research questions.

Dimensions assessed	Main variables and questions used	Reference
Sociodemographic characteristics	Age, gestation week, parity, marital status, place of residence, country of birth, educational level, occupation, and number of antenatal consultations.	Developed for this project
Anthropometry and knowledge of GWG[a] guidelines	Participants were asked to state their height, pre-pregnancy weight, and current GWG.[b] Also, women were asked whether they were familiar with the IOM[c] GWG table.	Developed for this project
Physical activity	Assessed using the question: The health authorities recommend all pregnant women to perform moderate-intensity aerobic physical activity (activities that take moderate physical effort and make you breathe somewhat harder than normal, such as brisk walking, housework etc) for a minimum of 30 minutes five days a week. With this in mind, would you characterize yourself as physically active a) prepregnancy and b) in your current gestation week? Response options: “Yes”, “No” or “I don’t know”.	Based on the American College of Obstetricians and Gynecologists recommendations [7]
Compliance with nutritional guidelines	Assessed using the question: “The Norwegian directorate of health recommend a balanced and varied diet, comprised of whole grain products, vegetables, fruits and berries, lean dairy products, fish, legumes and nuts, while also limiting the amount of processed meats, red meat and foods high in saturated fat, sugar and salt [31]. With this in mind, how would you characterize your diet in your current gestation week?” The participants rated their diet on a scale from 0-10, where 0 represented “Very poor” and 10 represented “Very good”.	Sagedal et al [28]
Information sources on PA[d]/GWG/nutrition	Assessed using the questions: “From where did you receive/retrieve information on PA/GWG/nutrition?” and: “Which of the information sources had the greatest impact on your PA/GWG/nutrition?” Response options (participants were able to choose more than one information source): “Midwife”, “Family physician”, “Blogs and Internet forums”, “Parenting magazines”, “Books and information pamphlets”, “Family and friends”, “I have not received/retrieved information/advice” and “Other”.	Developed for this project
Advice on PA	Assessed using the question: “Have you received any of the following advice on PA?” Response options: “Maintain the same level of PA as prepregnancy”, “Increase PA/exercise”, “Reduce PA/exercise”, “Avoid PA/exercise” and “Other”.	Developed for this project
Advice on GWG	Assessed using the question: “How much (in kg) have the information sources indicated that your total GWG should be?”	Developed for this project

[a] GWG: gestational weight gain.

[b] Prepregnancy height and weight were used to calculate prepregnancy body mass index (BMI). BMI categories and gestational weight gain ranges were consistent with the World Health Organization’s guidelines [32] and the guidelines from the IOM [16].

[c] IOM: Institute of Medicine.

[d] PA: physical activity.

Statistical Analyses

All statistical analyses were performed using IBM SPSS Statistical Software version 21.0 for Windows. Background variables and information sources are presented as frequencies, percentages, or means with standard deviation.

For the purpose of analysis, we divided the participants into 3 groups based on the information source perceived to mostly impact maternal health behavior:

1. Internet and media (including blogs and online forums, parenting magazines and books, and information pamphlets).
2. Health professionals (including midwives and family physicians).
3. Friends and family.

Whether a woman had gained weight below, within, or above the GWG guidelines was calculated using mean recommended weight gain in the first trimester (1.5 kg), adding the mean recommended number of grams per week multiplied by the number of weeks the woman was pregnant above the first trimester [16]. Good nutritional habits and compliance with nutritional guidelines was defined as a score ≥7 on an 11-point scale. The relationship between information sources and selected variables, including health behaviors and sociodemographic variables, was assessed by logistic regression, linear regression, or chi-square as appropriate. The level of statistical significance was set at $P<.05$.

Results

Participant characteristics are shown in Table 2. Responses were received from 275 pregnant women, 244 of whom were recruited through social media and 31, recruited through antenatal clinics. The analysis included only the results from the 150 participants who fully completed the questionnaire and provided data on information sources and health behaviors. The participants were residents in both urban and rural parts of Oslo. The age ranged from 19 to 45 years with a mean of 31.1 (SD 4.3) years. The mean gestation week was 30.6 (SD 5.9) and mean prepregnancy body mass index (BMI) was 24.2 (SD 4.2) kg/m^2. The mean number of antenatal consultations was 5.2 (SD 2.7; range 1-15).

Almost 90% of women reported that they were physically active for a minimum of 150 min of moderate intensity each week before pregnancy. This number decreased to less than 50% in the current gestation week (Table 2).

As shown in Table 3, nearly 65% of the women had gained weight outside the GWG guidelines. About half the participants (48.9%) had knowledge of the IOM table for recommended weight gain (Table 3). Binary logistic regression analyses revealed that weight gain was not significantly associated with knowledge of GWG guidelines (Table 4). The odds of gaining weight within recommended levels were not significantly different among participants knowing and not knowing about the guidelines (Table 4). Furthermore, the odds of knowing the GWG guidelines did not differ among those gaining weight below or above the recommended level as compared with those gaining within the recommended levels (Table 4). The low number of participants within the different categories limited adjustments of the statistical models. Nevertheless, the results in Table 4 did not change to any appreciable extent when we adjusted for parity, education level, and prepregnancy BMI (data not shown).

Perceived good nutritional habits and compliance with the nutritional guidelines was reported by 65.3% (98/150) of the women.

Table 2. Participant characteristics (N=150).

Characteristics	Value, n (%)
Parity	
Primiparous	91 (60.7)
Multiparous	59 (39.3)
Marital status	
Married or living together	147 (98.0)
Other	3 (2.0)
Country of birth	
Norway	130 (86.7)
Other	20 (13.3)
Education	
<4 years college or university	54 (36.0)
≥4 years college or university	96 (64.0)
Employment status	
Employed or student	144 (96.0)
Not employed	6 (4.0)
Physically active	
Prepregnancy	132 (88.0)
During pregnancy	73 (48.7)
Prepregnancy body mass index category	
Underweight	2 (1.3)
Normal weight	102 (68.4)
Overweight	28 (18.7)
Obese	17 (11.4)
Prenatal care provider	
Family physician	25 (16.7)
Midwife	43 (28.7)
Shared care[a]	75 (50.0)
Other	7 (4.7)

[a] Antenatal care shared between midwife and family physician.

Table 3. Women gaining within, below, or above the Institute of Medicine (IOM) recommendations (n=139). Data are presented in frequency (n), percentage (%), mean (SD) kg below and above recommendations and knowledge of IOM guidelines.

Variable	n (%)	mean (SD)	Knowledge of IOM[a] guidelines, n (%)
Within recommendations	51 (36.7)	—[b]	27 (19.4)
Below recommendations	37 (26.7)	−2.6 (2.2)	14 (10.0)
Above recommendations	51 (36.7)	+3.0 (2.4)	27 (19.4)

[a] IOM: Institute of Medicine.

[b] Not applicable.

Table 4. Statistics associated with (A) the odds of gaining weight within recommended levels for participants having knowledge of the Institute of Medicine (IOM) guidelines compared with participants without knowledge of the IOM guidelines and (B) the odds of having knowledge of the IOM guidelines for participants gaining weight below and above the recommended levels compared with participants gaining within the recommended levels.

Variable	n	Odds ratio (95% CI)	*P* value
A			
Not knowing the guideline	68	Ref[a]	—[b]
Knowing the guideline	68	1.29 (0.64-2.59)	.48
B			
Within recommendations	51	Ref	—
Below recommendations	37	0.54 (0.23-1.29)	.17
Above recommendations	51	1 (0.45-2.20)	1.00

[a] Ref: statistical reference group.

[b] Not applicable.

Information Sources

Most women reported multiple information sources on PA, GWG, and nutrition (Table 5). More than 8 out of 10 had received or retrieved information about nutrition (88.7%, 133/150) and PA (80.7%, 121/150), whereas 54.0% (81/150) reported information on GWG. When combining all 3 lifestyle factors, 38.5% had retrieved information from blogs and online forums and 26.7% from their midwife or family physician (Table 5). Books and information pamphlets were the most frequent information source on nutrition, reported by 48.0% (72/150).

Pearson's chi-square tests revealed that across all 3 lifestyle factors, significantly more women reported internet and media sources than health professionals as the information source with the most impact on their health behavior (PA: chi-square=23.25; *P*<.001; GWG: chi-square=38.13; *P*<.001; nutrition: chi-square=67.11; *P*<.001; Table 6). Significantly more women also reported family and friends to be the most important source of PA information compared with health professionals (chi-square=7.31; *P*=.007), but there was no difference between family and friends and health professionals with regard to GWG (chi-square=2.14; *P*=.14) and nutrition (chi-square=2.75; *P*=.097; Table 6).

Table 5. Pregnant women's information sources on physical activity, gestational weight gain, and nutrition.

Information sources	Physical activity, %	Gestational weight gain, %	Nutrition, %
Blogs and online forums	42.7	32.0	40.7
Books and information pamphlets	32.0	22.0	48.0
Parenting magazines	20.7	10.0	20.7
Friends and family	27.3	7.3	22.7
Midwife	30.7	18.0	35.3
Family physician	28.7	14.0	34.0
Other	10.0	2.7	7.3
I have not received or retrieved information	19.3	46.0	11.3

Table 6. Information sources with the most impact on health behavior. Participants were divided into 3 groups based on the information source perceived to mostly impact maternal health behavior.

Information sources	Physical activity, %	Gestational weight gain, %	Nutrition, %
Internet and media	54.1	63.2	68.0
Friends and family	27.1	8.8	8.0
Health professionals	18.8	28.1	24.0

When examining sources of information across sociodemographic groups, including age, education level, employment status, parity, number of children, marital status, and country of birth, the only significant association was between high education and friends and family as the main information source on nutrition. The odds of reporting friends and family as the most important source of information with regard to nutrition during pregnancy was 67% lower among women with ≥4 years of college or university education compared with women with <4 years of college or university education (odds ratio [OR] 0.33, 95% CI 0.13-0.85; *P*=.02).

Advice Consistent With Guidelines

Irrespective of the information source, 32.8% and 27.3% of participants had received advice inconsistent with the GWG guidelines [16] and PA guidelines [7], respectively. Less than one-third of women (29.3%, 44/150) had received specific advice on PA from a midwife or family physician. Of these, 25.0% (11/150) had received advice inconsistent with the present PA guidelines. With regard to GWG, 13.3% (20/150) reported receiving specific advice from a health professional, and half of these had received advice inconsistent with the GWG guidelines.

Impact of Information Sources on Women’s Health Behaviors

No significant associations were observed between the 3 groups of information sources and the odds of meeting the PA guidelines.

Binary logistic regressions on the impact of information sources on GWG revealed that choosing internet and media or health professionals as the most important source on GWG information increased the odds of gaining below the guidelines compared with gaining within the guidelines (internet and media: n=40, OR 15.5, 95% CI 1.4-167.4; *P*=.02; health professionals: n=40, OR 7.9, 95% CI 0.7-83.8; *P*=.08). Choosing friends and family as the most important source on GWG information increased the odds of gaining above the guidelines compared with gaining within (n=56, OR 12.0, 95% CI 1.3-111.7; *P*=.03). Otherwise, no significant relationships or interaction effects were observed.

Linear regressions revealed that choosing the internet and media as the most important information source on nutrition was associated with an increase of 0.7 on the 11-item scale representing compliance with nutritional recommendations (n=133; 95% CI 0.07-1.3; *P*=.03). This association remained significant when controlling for self-reported diet before pregnancy (*P*=.03). No significant associations were found between the categories, family and friends or health professionals and compliance with nutritional guidelines.

Discussion

Principal Findings

To our knowledge, this was the first study investigating the relationship between pregnant women’s information sources and their health behaviors. Consistent with previous research, the most common sources of information were the internet and media [8,33]. About one-fourth of women reported receiving information on PA, GWG, and nutrition from their midwife or family physician. Reporting the internet and media as the most important source increased the odds of gaining weight below the GWG guidelines but was also associated with higher compliance with nutritional guidelines. The category, friends and family, was significantly associated with gaining above the GWG guidelines.

Our recruitment method may be an important reason for the large proportion of women choosing the internet and media as preferred information sources and may have skewed the results. On the contrary, our results correspond with other studies, recruiting women in the first hospital antenatal visit [8,27]. A qualitative study has found that women turn to the internet to gather information before meeting with a health professional and afterward to obtain more information [23]. Social media was also considered an arena for socializing and sharing experiences with other pregnant women and mothers [23]. In this study, we have not thoroughly investigated the quality of the internet advice, but a meta-analysis of health website evaluations concluded that the advice often lacked accuracy and that it was difficult to find high quality sites [11]. Hence, it is important that women are guided toward trustworthy online resources during pregnancy.

According to the Norwegian guidelines for antenatal care, all prenatal patients should receive lifestyle counselling, including advice on PA, GWG, and nutrition on the first prenatal visit [34]. Consistent with other studies, only one-fourth of women in this study reported receiving advice from their midwife or family physician on these topics [8,9,13,14,19,20]. Moreover, only 10 participants received advice consistent with the GWG guidelines. Although others have also reported low numbers (5.2%-12.0%) with regard to information on GWG from health professionals [8,20], this is in contrast to research showing that receiving advice from a health professional increases the likelihood of gaining weight within the guidelines [27,35]. Hence, interventions are needed to increase the percentage of health professionals who accurately advise women on PA, GWG, and nutrition during pregnancy.

Impact on Women's Health Behaviors

Stating the internet and media as the most important sources on GWG information significantly increased the odds of gaining below the GWG guidelines. Hicks and Brown [36] found that time spent on social media was associated with body dissatisfaction among pregnant women. In this study, no association between negative body image and choosing the internet and media as the most important sources of information was found (data not shown). Yet, the way internet and media idealize a slender pregnant woman with a neat bump [36] may have influenced the women's GWG.

Reporting friends and family as the main information source on GWG was associated with gaining above the GWG guidelines. Public health guidelines on GWG have been revised several times over the past decades [16,37] and friends and family might not have knowledge of the current guidelines, and therefore give incorrect advice. However, only 10 women in this study considered friends and family to have the most impact on GWG.

The most frequently reported information source on nutrition was books and information pamphlets, highlighting the importance of this media for dietary information. The Norwegian Directorate of Health [31] distributes an information pamphlet with updated nutritional guidelines, available at antenatal clinics as well as on the internet. In this study, we observed that compliance with nutritional guidelines was significantly associated with reporting the internet and media as the primary information source. This association remained significant after controlling for self-reported diet before pregnancy.

Strengths and Limitations

This is one of the largest studies investigating pregnant women's information sources and the first to explore how different information sources may impact 3 distinct but importantly related health topics: PA, GWG, and nutrition. In addition, the questionnaire used covered a broad range of factors that could possibly explain women's health behaviors and was based on previously validated questions and questions used in similar studies [28-30,38]. Furthermore, the use of an electronic questionnaire is time efficient and cost-effective [39]. On the contrary, the internet recruitment made it difficult to determine the response rate. Moreover, the results of this study cannot be generalized owing to the recruitment of participants through social media and an overrepresentation of highly educated Nordic Caucasian women. In addition, a much higher percentage of participants reported meeting the PA guidelines compared with previous reports [30,40]. Pregnant women answering such a questionnaire may be more interested and more attentive to PA than nonparticipants, thus introducing the problem of selection bias. Our initial approach was to recruit participants through all 18 antenatal clinics in Oslo. However, only 2 decided to participate. Instead of recruiting women through social media and thus compromising the generalizability of the results, we should have explored the possibility of recruiting women through endocrine clinics or specialized clinics. Yet, the findings regarding women's information sources are consistent with other research [8,27]. Moreover, the use of a nonvalidated question to assess PA and dietary quality and the lack of questions evaluating predictors of health behavior change makes the results less reliable. All information was self-reported and therefore subjective to social desirability bias.

Conclusions

The small number of health professionals giving information and the extensive use of internet- and media-based sources highlight the need to address the quality of internet advice and guide women toward trustworthy sources of information during pregnancy. Even though internet and media sources seemed to have a positive impact on nutritional behavior, it was also associated with gaining below the GWG guidelines. Further research investigating how different information sources influence PA, GWG, and nutritional behaviors is needed.

Acknowledgments

The authors sincerely thank all the women who agreed to participate in this study. The authors also thank Knut Eirik Dalene, PhD, for assistance with the statistical analysis.

Conflicts of Interest

None declared.

References

1. da Silva SG, Ricardo LI, Evenson KR, Hallal PC. Leisure-time physical activity in pregnancy and maternal-child health: a systematic review and meta-analysis of randomized controlled trials and cohort studies. Sports Med 2017 Feb;47(2):295-317. [doi: 10.1007/s40279-016-0565-2] [Medline: 27282925]
2. di Mascio D, Magro-Malosso ER, Saccone G, Marhefka GD, Berghella V. Exercise during pregnancy in normal-weight women and risk of preterm birth: a systematic review and meta-analysis of randomized controlled trials. Am J Obstet Gynecol 2016 Nov;215(5):561-571. [doi: 10.1016/j.ajog.2016.06.014] [Medline: 27319364]
3. Shin D, Lee KW, Song WO. Dietary patterns during pregnancy are associated with risk of gestational diabetes mellitus. Nutrients 2015 Nov 12;7(11):9369-9382 [FREE Full text] [doi: 10.3390/nu7115472] [Medline: 26569302]
4. Goldstein RF, Abell SK, Ranasinha S, Misso M, Boyle JA, Black MH, et al. Association of gestational weight gain with maternal and infant outcomes: a systematic review and meta-analysis. J Am Med Assoc 2017 Jun 6;317(21):2207-2225 [FREE Full text] [doi: 10.1001/jama.2017.3635] [Medline: 28586887]

5. Knudsen VK, Orozova-Bekkevold IM, Mikkelsen TB, Wolff S, Olsen SF. Major dietary patterns in pregnancy and fetal growth. Eur J Clin Nutr 2008 Apr;62(4):463-470. [doi: 10.1038/sj.ejcn.1602745] [Medline: 17392696]
6. Mørkved S, Bø K. Effect of pelvic floor muscle training during pregnancy and after childbirth on prevention and treatment of urinary incontinence: a systematic review. Br J Sports Med 2014 Feb;48(4):299-310. [doi: 10.1136/bjsports-2012-091758] [Medline: 23365417]
7. American College of Obstetricians and Gynecologists. ACOG committee opinion no. 650: physical activity and exercise during pregnancy and the postpartum period. Obstet Gynecol 2015 Dec;126(6):e135-e142. [doi: 10.1097/AOG.0000000000001214] [Medline: 26595585]
8. Willcox JC, Campbell KJ, McCarthy EA, Lappas M, Ball K, Crawford D, et al. Gestational weight gain information: seeking and sources among pregnant women. BMC Pregnancy Childbirth 2015 Aug 7;15:164 [FREE Full text] [doi: 10.1186/s12884-015-0600-6] [Medline: 26249832]
9. Downs DS, Savage JS, Rauff EL. Falling short of guidelines? Nutrition and weight gain knowledge in pregnancy. J Womens Health Care 2014;3:pii: 1000184 [FREE Full text] [doi: 10.4172/2167-0420.1000184] [Medline: 25599012]
10. Grimes HA, Forster DA, Newton MS. Sources of information used by women during pregnancy to meet their information needs. Midwifery 2014 Jan;30(1):e26-e33. [doi: 10.1016/j.midw.2013.10.007] [Medline: 24246969]
11. Eysenbach G, Powell J, Kuss O, Sa ER. Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. J Am Med Assoc 2002;287(20):2691-2700. [doi: 10.1001/jama.287.20.2691] [Medline: 12020305]
12. Ferrari RM, Siega-Riz AM, Evenson KR, Moos MK, Carrier KS. A qualitative study of women's perceptions of provider advice about diet and physical activity during pregnancy. Patient Educ Couns 2013 Jun;91(3):372-377 [FREE Full text] [doi: 10.1016/j.pec.2013.01.011] [Medline: 23399436]
13. Stengel MR, Kraschnewski JL, Hwang SW, Kjerulff KH, Chuang CH. 'What my doctor didn't tell me': examining health care provider advice to overweight and obese pregnant women on gestational weight gain and physical activity. Womens Health Issues 2012;22(6):e535-e540 [FREE Full text] [doi: 10.1016/j.whi.2012.09.004] [Medline: 23122213]
14. Clarke PE, Gross H. Women's behaviour, beliefs and information sources about physical exercise in pregnancy. Midwifery 2004 Jun;20(2):133-141. [doi: 10.1016/j.midw.2003.11.003] [Medline: 15177856]
15. Whitaker KM, Wilcox S, Liu J, Blair SN, Pate RR. Patient and provider perceptions of weight gain, physical activity, and nutrition counseling during pregnancy: a qualitative study. Womens Health Issues 2016;26(1):116-122 [FREE Full text] [doi: 10.1016/j.whi.2015.10.007] [Medline: 26621605]
16. Institute of Medicine (US) and National Research Council (US) Committee to Reexamine IOM Pregnancy Weight Guidelines. In: Rasmussen KM, Yaktine AL, editors. Weight Gain During Pregnancy: Reexamining the Guidelines. Washington, DC: The National Academies Press; 2009.
17. Chang T, Llanes M, Gold KJ, Fetters MD. Perspectives about and approaches to weight gain in pregnancy: a qualitative study of physicians and nurse midwives. BMC Pregnancy Childbirth 2013 Feb 21;13:47 [FREE Full text] [doi: 10.1186/1471-2393-13-47] [Medline: 23433216]
18. Wilkinson SA, Poad D, Stapleton H. Maternal overweight and obesity: a survey of clinicians' characteristics and attitudes, and their responses to their pregnant clients. BMC Pregnancy Childbirth 2013 May 21;13:117 [FREE Full text] [doi: 10.1186/1471-2393-13-117] [Medline: 23692981]
19. McDonald SD, Pullenayegum E, Bracken K, Chen AM, McDonald H, Malott A, et al. Comparison of midwifery, family medicine, and obstetric patients' understanding of weight gain during pregnancy: a minority of women report correct counselling. J Obstet Gynaecol Can 2012 Feb;34(2):129-135. [doi: 10.1016/S1701-2163(16)35155-6] [Medline: 22340061]
20. McDonald SD, Pullenayegum E, Taylor VH, Lutsiv O, Bracken K, Good C, et al. Despite 2009 guidelines, few women report being counseled correctly about weight gain during pregnancy. Am J Obstet Gynecol 2011 Oct;205(4):333.e1-333.e6. [doi: 10.1016/j.ajog.2011.05.039] [Medline: 21784404]
21. Bianchi CM, Huneau JF, Le Goff G, Verger EO, Mariotti F, Gurviez P. Concerns, attitudes, beliefs and information seeking practices with respect to nutrition-related issues: a qualitative study in French pregnant women. BMC Pregnancy Childbirth 2016 Dec 12;16(1):306 [FREE Full text] [doi: 10.1186/s12884-016-1078-6] [Medline: 27729021]
22. Ledoux T, van den Berg P, Leung P, Berens PD. Factors associated with knowledge of personal gestational weight gain recommendations. BMC Res Notes 2015 Aug 13;8:349 [FREE Full text] [doi: 10.1186/s13104-015-1306-6] [Medline: 26268578]
23. Kraschnewski JL, Chuang CH, Poole ES, Peyton T, Blubaugh I, Pauli J, et al. Paging 'Dr. Google': does technology fill the gap created by the prenatal care visit structure? Qualitative focus group study with pregnant women. J Med Internet Res 2014 Jun 3;16(6):e147 [FREE Full text] [doi: 10.2196/jmir.3385] [Medline: 24892583]
24. Lima-Pereira P, Bermúdez-Tamayo C, Jasienska G. Use of the internet as a source of health information amongst participants of antenatal classes. J Clin Nurs 2012 Feb;21(3-4):322-330. [doi: 10.1111/j.1365-2702.2011.03910.x] [Medline: 22093043]
25. Cannella D, Lobel M, Monheit A. Knowing is believing: information and attitudes towards physical activity during pregnancy. J Psychosom Obstet Gynaecol 2010 Dec;31(4):236-242. [doi: 10.3109/0167482X.2010.525269] [Medline: 20969538]

26. Szwajcer EM, Hiddink GJ, Koelen MA, van Woerkum CM. Nutrition-related information-seeking behaviours before and throughout the course of pregnancy: consequences for nutrition communication. Eur J Clin Nutr 2005 Aug;59(Suppl 1):S57-S65. [doi: 10.1038/sj.ejcn.1602175] [Medline: 16052197]
27. Mercado A, Marquez B, Abrams B, Phipps MG, Wing RR, Phelan S. Where do women get advice about weight, eating, and physical activity during pregnancy? J Womens Health (Larchmt) 2017 Dec;26(9):951-956 [FREE Full text] [doi: 10.1089/jwh.2016.6078] [Medline: 28525293]
28. Sagedal LR, Øverby NC, Lohne-Seiler H, Bere E, Torstveit MK, Henriksen T, et al. Study protocol: fit for delivery - can a lifestyle intervention in pregnancy result in measurable health benefits for mothers and newborns? A randomized controlled trial. BMC Public Health 2013 Feb 13;13:132 [FREE Full text] [doi: 10.1186/1471-2458-13-132] [Medline: 23406306]
29. Haakstad LA, Gundersen I, Bø K. Self-reporting compared to motion monitor in the measurement of physical activity during pregnancy. Acta Obstet Gynecol Scand 2010 Jun;89(6):749-756. [doi: 10.3109/00016349.2010.484482] [Medline: 20450260]
30. Owe KM, Nystad W, Bø K. Correlates of regular exercise during pregnancy: the Norwegian mother and child cohort study. Scand J Med Sci Sports 2009 Oct;19(5):637-645. [doi: 10.1111/j.1600-0838.2008.00840.x] [Medline: 18627550]
31. A healthy lifestyle before and during pregnancy. Oslo: Norwegian Directorate of Health; 2016. URL: https://www.helsedirektoratet.no/brosjyrer/gode-levevaner-for-og-i-svangerskapet/_/attachment/download/73fa64c1-16f3-4052-851c-6a3b2e13ef12:33aa515ffb4a714584b93e4fad70568ec85f4fd0/Gode%20levevaner%20f%C3%B8r%20og%20i%20svangerskapet%20-%20engelsk.pdf
32. Global Recommendations on Physical Activity for Health. Geneva: World Health Organization; 2010.
33. Huberty J, Dinkel D, Beets MW, Coleman J. Describing the use of the internet for health, physical activity, and nutrition information in pregnant women. Matern Child Health J 2013 Oct;17(8):1363-1372. [doi: 10.1007/s10995-012-1160-2] [Medline: 23090284]
34. [Guidelines for Pregnancy Care]. Oslo: Norwegian Directorate for Health and Social Affaires; 2005. URL: https://www.helsedirektoratet.no/retningslinjer/svangerskapsomsorgen
35. Stotland NE, Haas JS, Brawarsky P, Jackson RA, Fuentes-Afflick E, Escobar GJ. Body mass index, provider advice, and target gestational weight gain. Obstet Gynecol 2005 Mar;105(3):633-638. [doi: 10.1097/01.AOG.0000152349.84025.35] [Medline: 15738036]
36. Hicks S, Brown A. Higher Facebook use predicts greater body image dissatisfaction during pregnancy: the role of self-comparison. Midwifery 2016 Sep;40:132-140. [doi: 10.1016/j.midw.2016.06.018] [Medline: 27428109]
37. Institute of Medicine. Nutrition During Pregnancy: Part I: Weight Gain, Part II: Nutrient Supplements. Washington, DC: National Academy Press; 1990.
38. Haakstad LA, Bø K. Exercise in pregnant women and birth weight: a randomized controlled trial. BMC Pregnancy Childbirth 2011 Sep 30;11:66 [FREE Full text] [doi: 10.1186/1471-2393-11-66] [Medline: 21961534]
39. Sinclair M, O'Toole J, Malawaraarachchi M, Leder K. Comparison of response rates and cost-effectiveness for a community-based survey: postal, internet and telephone modes with generic or personalised recruitment approaches. BMC Med Res Methodol 2012 Aug 31;12:132 [FREE Full text] [doi: 10.1186/1471-2288-12-132] [Medline: 22938205]
40. Gjestland K, Bø K, Owe KM, Eberhard-Gran M. Do pregnant women follow exercise guidelines? Prevalence data among 3482 women, and prediction of low-back pain, pelvic girdle pain and depression. Br J Sports Med 2013 May;47(8):515-520. [doi: 10.1136/bjsports-2012-091344] [Medline: 22904295]

Abbreviations

BMI: body mass index
GWG: gestational weight gain
IOM: Institute of Medicine
OR: odds ratio
PA: physical activity
REK: Regional Committee for Medical and Health Research Ethics

Edited by G Eysenbach; submitted 09.04.18; peer-reviewed by Y Gupta, SA Wilkinson; comments to author 07.08.18; revised version received 16.01.19; accepted 13.04.19; published 05.07.19.

Please cite as:
Dalhaug EM, Haakstad LAH
What the Health? Information Sources and Maternal Lifestyle Behaviors
Interact J Med Res 2019;8(3):e10355
URL: https://www.i-jmr.org/2019/3/e10355/
doi:10.2196/10355
PMID:31278731

Original Paper

Influence of Enthusiastic Language on the Credibility of Health Information and the Trustworthiness of Science Communicators: Insights From a Between-Subject Web-Based Experiment

Lars König[1], MSc; Regina Jucks[1], DrPhil

Department of Psychology, University of Münster, Münster, Germany

Corresponding Author:
Lars König, MSc
Department of Psychology
University of Münster
Fliednerstraße 21
Münster, 48149
Germany
Phone: 49 2518323685
Email: lars.koenig@uni-muenster.de

Abstract

Background: To decide whether online health information is reliable, information seekers apply 2 stretegies: first, information seekers can make credibility judgments by using their prior knowledge to evaluate the validity of the encountered health claim. Second, instead of evaluating the health claim itself, information seekers can make trustworthiness judgments by evaluating the character of the information source. In recent years, information givers from various professions have begun to use enthusiastic language to disseminate their information and persuade their audiences.

Objective: To systematically explore this phenomenon, the goal of this study was to answer the following research questions: (1) does an enthusiastic language style, in comparison with a neutral language style, increase the trustworthiness of a person arguing in an online health forum and the credibility of his or her information? (2) does working for a university, in comparison with working for a lobbying organization, increase the trustworthiness of a person arguing in an online health forum and the credibility of his or her information? (3) does working for a university in combination with using an enthusiastic language style result in especially high trustworthiness and credibility ratings?

Methods: In a 2x2 between-subject online experiment, 270 participants read a post from an online health forum and subsequently rated the trustworthiness of the forum post author and the credibility of his information. A total of 2 aspects of the forum post varied, namely the professional affiliation of the forum post author (whether the person introduced himself as a scientist or a lobbyist) and his language style (whether he used a neutral language style or an enthusiastic language style).

Results: When the forum post author used an enthusiastic language style, he was perceived as more manipulative ($P<.001$), less knowledgeable ($P<.001$), and his information was perceived as less credible ($P<.001$). Overall, scientists were perceived as less manipulative ($P=.04$) than lobbyists. Furthermore, language style and professional affiliation interacted: When the forum post author was a lobbyist, language style did not affect integrity ($P=.96$) and benevolence ($P=.79$) ratings. However, when the forum post author was a scientist, enthusiastic language led to lower integrity ($P=.002$) and benevolence ($P<.001$) ratings than neutral language.

Conclusions: The current findings illustrate that health information seekers do not just react to online health information itself. In addition, they are also sensitive to the ways in which health information is presented ("Which langue style is used to communicate health information?") and who presents it ("Who does the health information source work for?").

(Interact J Med Res 2019;8(3):e13619) doi:10.2196/13619

KEYWORDS

health communication; information seeking behavior; trust; language; occupations; deep learning; FMRI; source credibility; persuasiveness

Introduction

Background

How do information seekers decide whether they can rely on online health information? The importance of this question is stressed by 2 recent developments: First, information seekers have developed diverse ways to acquire online health information [1], and they rely on it frequently [2]. Second, online health information often contains misinformation [3-5] because the internet is not governed by professional editors [6,7]. Hence, information seekers constantly have to decide whether they should rely on the health information they encounter online. According to the content-source integration model [8], 2 strategies can be applied to make such decisions. First, information seekers can make credibility judgments (first-hand evaluations) by using their prior knowledge to evaluate the validity of the encountered health information claim. Second, instead of evaluating the health information claim itself, information seekers can make trustworthiness judgments (second-hand evaluations) by evaluating the character of the information source. As most information seekers have just a bounded understanding of science, they often lack the necessary expertise to make accurate credibility judgments [9,10]. Therefore, information seekers frequently have to turn to experts and evaluate their trustworthiness [11-14]. This development gives rise to another intriguing question: Which factors influence information seekers' credibility and trustworthiness judgments?

Language Styles and Their Relationship to Credibility and Trustworthiness

Various factors influence information seekers' credibility and trustworthiness judgments [15-17], but the language style of an information source seems to be an especially influential factor [18-22]. Thon and Jucks [18], for example, showed that the authors of health information forum posts were rated as more trustworthy and their information as more credible when they used an everyday language style (eg, "heart attack") instead of a technical language style (eg, "myocardial infarction"). Furthermore, Mayweg-Paus and Jucks [22] showed that participants accepted information from an online health article to a higher degree and processed it in more depth when the article was written in a tentative language style (eg, "is presumably similar") rather than a nontentative language style (eg, "is similar"). Tentativeness, however, does not just infuence trustworthiness and credibility judgments. In a study by Feinkohl et al [23], participants saw an online news article about the therapeutic application of deep brain stimulation. Depending on the experimental condition, the article was accompanied by different online forum comments or no comments at all. The results showed that participants were more likely to address the tentativeness of the findings in their own forum comments if they had previously seen forum comments that addressed this issue too (for additional information on scientific tentativeness and medical research, see the study by Flemming et al [24]). In another study, König and Jucks [20] showed that science communicators in scientific health debates were rated as more trustworthy and their information as more credible when they used a neutral language style (eg, "a series of methodological mistakes") instead of an aggressive language style (eg, "a series of really dumb methodological mistakes"). Besides influencing trustworthiness and credibility judgments, the use of emotional language also influences interactions in online forums and risk perceptions. For example, research showed that medical students incease their emotional language use when replying to emotional patient queries in medical forums [25]. Furthermore, a recent study showed that emotionalization can influence risk perceptions in a science communication context [26]. Aggressive, technical, tentative, and everyday language styles, however, are not the only language styles that are used to disseminate health information.

This Study: How Does Enthusiastic Language Influence the Credibility of Health Information and the Trustworthiness of Science Communicators?

In recent years, more and more people use enthusiastic language to disseminate information and persuade their audiences in diverse contexts. For example, Ghose et al [27] observed that, instead of just writing positive comments on online marketplace platforms, "buyers tend to use superlatives and highly enthusiastic language to praise a good merchant." In another study, Barry et al [28] pointed out that enthusiastic language is used as a persuasion tactic on commercial websites. Furthermore, enthusiastic language is nowadays often used to communicate scientific findings. Instead of using standard academic language, authors of scientific articles increasingly express their enthusiasm about their findings by including enthusiastic phrases in their article titles and texts, for example, "Fantastic beasts" [29] and "The Incredible Shrinking Spindle" [30].

So far, it is not clear why so many people use enthusiastic language when communicating information. One reason might be that enthusiasm is often encouraged in educational settings because it has been linked to various positive outcomes [31]. For example, teacher enthusiasm is linked to students' enjoyment [32], interest [33], achievement [34], motivation, and vitality [35]. Therefore, information givers might think that using enthusiastic language to communicate information makes them especially effective and follows best-practice examples. But is this true? Interestingly, no research so far has systematically explored whether enthusiastic language influences the trustworthiness of an information source and the credibility of his or her information. And even if it does, would using enthusiastic language be equally effective for people from different professions? Burgoon et al [36], for example, argue in the context of Language Expectancy Theory that "highly credible communicators have the freedom (wide bandwidth) to select varied language strategies and compliance-gaining techniques in developing persuasive messages, whereas low-credible communicators must conform to more limited language options if they wish to be effective." Following this argumentation, one would expect that high-credibility communicators such as scientists would benefit more from the use of an enthusiastic language style than low-credibility communicators such as lobbyists. To test these hypotheses, we developed a between-subject online experiment. In the experiment, participants read a post from an online health forum and subsequently rated the trustworthiness of the forum post author and the credibility of his information. We chose online health forums because they typically allow users from diverse educational and professional backgrounds to exchange information in an unrestricted way. Furthermore, diverse research findings suggest that information seekers frequently rely on online forums to aquire health information [37-40]. Within the forum post, we varied the professional affiliation of the forum post author (whether the person introduced himself as a scientist or a lobbyist) and his language style (whether he used a neutral language style or an enthusiastic language style). The goal was to test the following hypotheses:

> *Hypothesis 1: An enthusiastic language style, in comparision with a neutral language style, increases the trustworthiness of a person arguing in an online health forum and the credibility of his information.*
>
> *Hypothesis 2: Working for a university, in comparison with working for a lobbying organization, increases the trustworthiness of a person arguing in an online health forum and the credibility of his information.*
>
> *Hypothesis 3: Working for a university in combination with using an enthusiastic language style results in especially high trustworthiness and credibility ratings.*

Methods

Design and Material

We used a 2 (language style: neutral language vs enthusiastic language) x 2 (professional affiliation: scientist vs lobbyist) between-subject experimental design, resulting in 4 experimental conditions. In each experimental condition, participants saw 2 online forum posts: a question post and an answer post. The 2 posts were embedded in a screenshot of a website. The URL of the website indicated that it was an online health forum. In the question post, a woman stated that there currently is a controversial debate about whether functional magnetic resonance imaging and artificial intelligence technologies can be combined to improve medical diagnoses. Following this, she asked whether anybody could explain the technologies to her. The question post was written in a neutral language style and was the same in all 4 experimental conditions. Textbox 1 shows the text of the question post. The textbox shows an English translation of the German post. Therefore, the translated version may not appear as authentic to native English speakers as the original version appears to native German speakers. The original German version of the question post can be obtained from the authors upon request. In the answer post, a man introduced himself and explained the basics of the functional magnetic resonance imaging technology. Following this, he described the results of a study [41] that had combined functional magnetic resonance imaging and artificial intelligence technologies. The experimental manipulations were realized in the answer post. Depending on the experimental condition, the answer post was written either in (1) a neutral language style or (2) an enthusiastic language style. Furthermore, the author of the answer post introduced himself either as (1) a scientist who worked for an imaging technology institute at a university or (2) a lobbyist who worked at an imaging technology lobbying organization. Textbox 2 shows the text of the answer post with the experimental manipulation. The enthusiastic language style version of the answer post contained the words and phrases printed in italics and the neutral language style version did not contain these words and phrases. The textbox shows an English translation of the German post. Therefore, the translated version may not appear as authentic to native English speakers as the original version appears to native German speakers. The original German version of the post can be obtained from the authors upon request.

Textbox 1. Text of the question post.

Dear forum community,

Functional magnetic resonance imaging allows activated brain areas to be spatially displayed. Increasingly, this technology is combined with methods of artificial intelligence, creating new applications. There is currently a lot of controversy as to whether the emerging methods should be used to diagnose illnesses (such as depression).

The opinions are very different. Advocates argue that diagnoses would become clearer through the use of the new methods. Critics argue that the potential of the new methods is overestimated and could lead to devastating misdiagnoses.

Does anyone know this technology and can give me a brief introduction?

Thanks in advance!

Sabine Schneider

Textbox 2. Text of the answer post with the experimental manipulation.

Hello Mrs. Schneider,

[Scientist Manipulation:] My name is Johannes Becker and I work for the Institute for Imaging Technology at the University of Bochum.

[Lobbyist Manipulation:] My name is Johannes Becker and I work for the Association of Imaging Technology-Producing Industries in Bochum.

I can give you a brief insight into the functionalities and applications of functional magnetic resonance imaging. *And what I can tell you at the beginning: I think the topic is fascinating!*

First to the basics: How does functional magnetic resonance imaging work? Nerve cells require more nutrients when active than at rest. If a particular brain region is active, the metabolism in this region increases. In response to this increased metabolism, the body transports more oxygen-rich blood into the activated brain region. Functional magnetic resonance imaging measures these changes in blood flow and uses them as an indicator of activity in different brain regions. *In my opinion, this is a simple and ingenious method!*

Now to the possible applications: The possible applications of functional magnetic resonance imaging are manifold *and I personally think many of them are breathtaking*. A current example of use - *that absolutely excites me* - is this:

Researchers from the US have combined functional magnetic resonance imaging and deep-learning technologies to decipher various activities of the human brain. For this purpose, the researchers showed their participants a series of videos and simultaneously recorded their brain activity using functional magnetic resonance imaging. The collected data was then fed to algorithms and they learned how the video clips and the brain activities were related to each other.

Afterwards, the researchers examined what the algorithms were capable of and in the following I would like to focus briefly on three central results *that astonish me again and again.*

The first result: the researchers provided one of the algorithms with new videos, and the algorithm was able to predict what brain activity the new videos would trigger in the participants. *I think this result is simply groundbreaking!*

The second result - *which I find very interesting* - is the following: One of the algorithms could later read from the brain activities of the participants, what the participants saw at a certain moment. For example, the algorithm could say with high accuracy whether the subjects were seeing an airplane, a bird, or a face.

Now the third result *and this result amazes me again and again when I talk about it:* One of the algorithms was able to draw a picture of what the participants saw at certain moments. The images had not the best resolution, but you could see the outlines and contours.

I hope you learned something new about the functionalities and applications of functional magnetic resonance imaging *and who knows, maybe you are now just as excited as I am about this topic.*

Yours sincerely,

Johannes Becker

Procedure

The experiment was conducted online using the Questback Enterprise Feedback Suite Survey platform for data collection. Before the experiment started, participants were told that the experiment would address the communication of information in online forums. Furthermore, they were informed about the general procedure of the upcoming experiment and that they could end the experiment at any time. To start the experiment, participants had to indicate that they had read all provided information and that they agreed to take part in the experiment. On the remaining pages, participants indicated their age, gender, whether they studied at the bachelor’s or master’s level, the university where they studied, their study subjects, and the semester they were currently in. Furthermore, they answered the control measures (see section “Control Measures”). Following this, participants were randomly assigned to 1 of the 4 experimental conditions and were presented with the corresponding online forum posts (see section “Design and Material”). After reading the forum posts, participants answered the dependent measures (see section “Dependent Measures”). After answering the dependent measures, participants answered the manipulation check question (see section “Manipulation Check”) and were debriefed. They were told about the manipulations of the experiment and that they could contact the leading scientist if they had any further questions or comments. Furthermore, they could choose to leave their information to get reimbursed for their participation. The experiment was designed to comply with the ethical guidelines developed by the American Psychological Association and the German Psychological Society.

Sample

German university students enrolled in diverse majors from the humanities and sciences were contacted via email and social network sites and received 5 Euro for participating in the online experiment. Participants who indicated at the end of the study that they answered the questions honestly and completed the study without interruption and technical problems were included in data analyses. The sample contained 270 (207 female, 63 male) students (149 undergraduate students, 121 graduate students) with an average age of 23 years (mean 23.39, SD 3.04). Furthermore, the average participant was enrolled in their study program for 7 semesters (mean 6.73, SD 3.96) and took 12 min (mean 11.51, SD 5.09) to complete the study.

Control Measures

A total of 3 control measures were included to assess whether the experimental groups differed in regard to characteristics that could affect the study results. It is possible that people who frequently use online forums are better able to identify the quality of online forum posts and their content than people who just occasionally use online forums. Therefore, participants answered the questions "How often do you visit Internet forums?" (general use) and "How often do you visit Internet forums to learn something new or acquire new skills?" (educational use) on a scale ranging from 1 (very rarely) to 7 (very often). Furthermore, it is possible that people who are well informed about a topic make different credibility and trustworthiness judgments than people who are less informed about the same topic. Therefore, participants answered the question "How much do you know about functional magnetic resonance imaging and deep learning?" (prior knowledge) on a scale ranging from 1 (very little) to 7 (very much).

Manipulation Check

To check whether the language style manipulation was successful, participants answered the question "How would you describe Johannes Becker's choice of words?" on a scale ranging from 1 (neutral) to 7 (extremely enthusiastic). In addition, we assessed the strength of the experimental manipulations by asking the participants whether they remembered the language style and professional affiliation of the answer post author. To assess whether the participants remembered the language style, they were asked whether certain enthusiastic expressions were used in the answer post. Participants could chose between "Yes," "No," and "I do not know." To assess whether the participants remembered the professional affiliation of the answer post author, they were asked "For whom did Johannes Becker work?" Participants could choose between "Institute for Imaging Technology at the University of Bochum," "Association of Imaging Technology-Producing Industries in Bochum," and "I do not know."

Dependent Measures

For each dependent measure, a total score was generated by calculating the mean.

Message Credibility

As a credibility measure of the provided information, the Message Credibility Scale [42] was translated and adapted. Participants indicated their agreement with 3 statements, for example, "The provided information was believable."

Machiavellianism

The German version of the Machiavellianism Subscale from the Dirty Dozen Scale [43,44] was used to assess how manipulative the forum post author was perceived to be. Participants indicated their agreement with 4 statements, for example, "Johannes Becker has used deceit or lied to get his way," on a scale ranging from 1 (totally disagree) to 7 (totally agree).

Expertise, Integrity, and Benevolence

The Muenster Epistemic Trustworthiness Inventory [45] was used to assess how trustworthy the forum post author was perceived to be. Participants rated 15 items on a scale ranging from 1 (not trustworthy at all) to 7 (very trustworthy). The items measured expertise (eg, "unqualified-qualified"), benevolence (eg, "immoral-moral"), and integrity (eg, "insincere-sincere").

Results

General Procedure

For all analyses, the statistical software IBM SPSS Statistics version 25 was used. For the analyses of the dependent measures, 2-way between-subject analyses of variance were conducted with language style (neutral language vs enthusiastic language) and professional affiliation (scientist vs lobbyist) as independent variables. Type 3 sum of squares were used. For all analyses, the alpha level was set at .05. The dataset contains further variables that have not been described in this study and have not been analyzed yet because they exceed the scope of this study.

Control Measures

Before analyzing the dependent measures, 4 1-way between-subject analyses of variance were conducted with experimental condition as the independent variable and the control measures as dependent variables to analyze whether the participants in the 4 experimental groups differed in aspects that could influence the study results. The results showed that the participants in the 4 experimental groups did not significantly differ in regard to their general online forum use (general use: $F_{3,266}$=0.817; P=.49), online forum use for educational purposes (educational use: $F_{3,266}$=0.052; P=.98), and their prior knowledge (prior knowledge: $F_{3,266}$=1.071; P=.36). Therefore, the 3 control measures were not included in further analyses.

Manipulation Check

Participants in the enthusiastic language style condition (mean 6.04, SD 0.95) perceived the choice of words as more enthusiastic than participants in the neutral language style condition (mean 2.37, SD 1.26); $t_{244.992}$=-26.890; P<.001. Hence, the language style manipulation worked as expected. Furthermore, 252 (93%) participants correctly remembered the language style of the answer post author, and 231 (86%) correctly remembered his professional affiliation. A total of 216 (80%) participants remembered both correctly. The high remembrance rates suggest that the vast majority of participants recognized the experimental manipulations. As information seekers naturally differ in their attention to detail in real-world settings and the experimental manipulations might not need to be consciously remembered to have an effect, we included all participants in the data analyses.

Dependent Measures

Tables 1 and 2 show the means and SDs of the dependent measures.

Table 1. Main effects: means and SDs of the dependent measures by language style and professional affiliation.

Dependent measures[a]	Language style			Professional affiliation		
	Neutral (n=133), mean (SD)	Enthusiastic (n=137), mean (SD)	*P* value	Scientist (n=135), mean (SD)	Lobbyist (n=135), mean (SD)	*P* value
Credibility						
Message credibility	5.11 (1.04)	4.48 (1.23)	<.001	4.83 (1.16)	4.75 (1.19)	.56
Trustworthiness						
Machiavellianism	2.70 (1.17)	3.85 (1.23)	<.001	3.13 (1.28)	3.44 (1.37)	.04
Expertise	5.44 (0.98)	4.90 (1.19)	<.001	5.15 (1.11)	5.19 (1.13)	.75
Integrity[b]	4.67 (0.91)	4.42 (1.03)	.03	4.54 (0.92)	4.55 (1.04)	.95
Benevolence[b]	4.40 (0.82)	4.10 (1.02)	.006	4.21 (0.94)	4.29 (0.93)	.52

[a] General interpretation: On the Machiavellianism scale, a low score indicates high trustworthiness and a high score indicates low trustworthiness. On all other scales, a low score indicates low trustworthiness/credibility and a high score indicates high trustworthiness/credibility. All scales ranged from 1 to 7.

[b] Interaction is significant; see Table 2.

Table 2. Interaction effects: means and SDs of the dependent measures by language style and professional affiliation.

Dependent measures[a]	Scientist			Lobbyist		
	Neutral (n=67), mean (SD)	Enthusiastic (n=68), mean (SD)	*P* value	Neutral (n=66), mean (SD)	Enthusiastic (n=69), mean (SD)	*P* value
Integrity	4.79 (0.87)	4.28 (0.90)	.002	4.54 (0.94)	4.55 (1.13)	.96
Benevolence	4.50 (0.83)	3.93 (0.96)	<.001	4.31 (0.79)	4.26 (1.05)	.79

[a] On all scales, a low score indicates low trustworthiness and a high score indicates high trustworthiness. All scales ranged from 1 to 7.

Message Credibility

There was a significant main effect of language style ($F_{1,266}=20.300$; $P<.001$; $\eta_p^2=.071$) on message credibility, indicating that enthusiastic language led to lower message credibility ratings than neutral language. There was no main effect of professional affiliation ($F_{1,266}=0.334$; $P=.56$; $\eta_p^2=.001$) and no interaction effect ($F_{1,266}=0.178$; $P=.67$; $\eta_p^2=.001$).

Machiavellianism

There was a significant main effect of language style ($F_{1,266}=62.898$; $P<.001$; $\eta_p^2=.191$) on Machiavellianism, indicating that enthusiastic language led to higher Machiavellianism ratings than neutral language. Furthermore, there was a significant main effect of professional affiliation ($F_{1,266}=4.487$; $P=.04$; $\eta_p^2=.017$) on Machiavellianism, indicating that being a lobbyist led to higher Machiavellianism ratings than being a scientist. There was no interaction effect ($F_{1,266}=0.010$; $P=.92$; $\eta_p^2<.001$).

Expertise

There was a significant main effect of language style ($F_{1,266}=16.357$; $P<.001$; $\eta_p^2=.058$) on expertise, indicating that enthusiastic language led to lower expertise ratings than neutral language. There was no main effect of professional affiliation ($F_{1,266}=0.106$; $P=.75$; $\eta_p^2<.001$) on expertise and no interaction effect ($F_{1,266}=0.008$; $P=.93$; $\eta_p^2<.001$).

Integrity

There was a significant main effect of language style ($F_{1,266}=4.530$; $P=.03$; $\eta_p^2=.017$) on integrity, indicating that enthusiastic language led to lower integrity ratings than neutral language. There was no main effect of professional affiliation ($F_{1,266}=0.004$; $P=.95$; $\eta_p^2<.001$) on integrity. However, the interaction was significant ($F_{1,266}=4.863$; $P=.03$; $\eta_p^2=.018$). Simple effect analysis indicated that when the forum post author was a lobbyist, language style did not affect integrity ($F_{1,266}=0.003$; $P=.96$; $\eta_p^2<.001$). However, when the forum post author was a scientist, enthusiastic language led to lower integrity ratings than neutral language ($F_{1,266}=9.392$; $P=.002$; $\eta_p^2=.034$).

Benevolence

There was a significant main effect of language style ($F_{1,266}=7.633$; $P=.006$; $\eta_p^2=.028$) on benevolence, indicating that enthusiastic language led to lower benevolence ratings than neutral language. There was no main effect of professional affiliation ($F\ 1,_{266}=0.422$; $P=.52$; $\eta_p^2=.002$) on benevolence. However, the interaction was significant ($F_{1,266}=5.679$; $P=.02$; $\eta_p^2=.021$). Simple effect analysis indicated that when the forum post author was a lobbyist, language style did not affect benevolence ($F_{1,266}=0.072$; $P=.79$; $\eta_p^2<.001$). However, when the forum post author was a scientist, enthusiastic language led to lower benevolence ratings than neutral language ($F_{1,266}$=13.243; P<.001; η_p^2=.047).

Discussion

Principal Findings

A total of 3 hypotheses addressed the effects of language style (neutral vs enthusiastic) and professional affiliation (scientist vs lobbyist) in online health forums. We hypothesized that an enthusiastic language style, in comparison with a neutral language style, would positively affect the trustworthiness of a forum post author and the credibility of his information. Contrary to our hypothesis, we found that when the forum post author used an enthusiastic language style, he received higher Machiavellianism ratings, lower expertise ratings, and lower message credibility ratings. Furthermore, we hypothesized that working for a university, in comparison with working for a lobbying organization, would positively affect the trustworthiness of a forum post author and the credibility of his information. The results partly confirm this hypothesis: Scientists received lower Machiavellianism ratings than lobbyists. However, the professional affiliation did not affect the other trustworthiness measures, and it did not affect the perceived credibility of the provided information. Finally, we hypothesized that working for a university in combination with using an enthusiastic language style would result in especially high trustworthiness and credibility ratings. Although the results did reveal an interaction between language style and professional affiliation, it was in the opposite direction: Contrary to our hypotheses, we found that when the forum post author was a scientist, enthusiastic language led to lower integrity ratings and lower benevolence ratings than neutral language.

Overall, even though it was not hypothesized, the results show that the enthusiastic language style decreased the trustworthiness of the forum post author and the credibility of his information. One possible reason for this finding, derived from communication accommodation theory [46], might be the question-answer setting in which the information exchange took place. In the experiment, the help seeker asked for advice and formulated her question in a neutral language style. Therefore, the participants might have expected an answer in an equally neutral language style and perceived the enthusiastic language style as a violation to the introduced communication rule.

Another surprising finding is that the professional affiliation of the forum post author only affected the Machiavellianism measure because previous research has found that scientists are typically perceived as benevolent, sincere, and likable [20]. One reason for this finding might be that the author mentioned his professional affiliation just briefly at the beginning of his forum post. Hence, the manipulation might have been relatively weak.

The last unexpected finding was that scientists who used an enthusiastic language style got especially low integrity and benevolence ratings. One possible reason for this finding might be that scientists are typically perceived as being rational and objective and that this stereotypical image is not compatible with an enthusiastic language style. Lobbyists, on the other hand, might be perceived as people who relentlessly pursue a specific goal, and this stereotypical image might be more compatible with an enthusiastic language style.

Limitations

Although the findings of this study highlight the importance of the language style and professional affiliation of people who are communicating health information in online settings, there are limitations to the generalizability of the results. One limitation could be the geographical location in which the experiment took place. More specifically, countries have developed different civic epistemologies (ways in which societies evaluate and discuss knowledge claims; see [47,48]). Discussions in Germany, for example, typically focus on “building communally crafted expert rationales, capable of supporting a policy consensus,” whereas in the United States, “information is typically generated by interested parties and tested in public through overt confrontation between opposing, interest - laden points of view” [47]. Hence, information seekers in Germany may prefer neutral language styles as a constructive way of reaching a consensus. In the United States, however, informaiton seekers are used to emotionally laden discussions and therefore, may react differently to enthusiastic language styles. Another limitation could be the age of this study sample. Previous research has shown that age differences exist in regard to source monitoring and suggestibility to misinformation [49]. Young adults, for example, who grew up with modern information technologies and have been confronted with misinformation on the internet throughout their lives, may pay more attention to relevant source information when evaluating online information. Older adults, however, may not be as critical as younger adults because of their lack of experience with misinformation on the internet and therefore, pay less attention to relevant source information.

A further limitation concerns the topic of the forum posts and the chosen study sample. According to the Elaboration Likelihood Model of Persuasion [50], the personal relevance of a topic influences information processing: If a topic is personally relevant, peripheral cues become less important. Therefore, participants who do not considere the topic to be personally relevant may rely more heavily on peripheral cues such as the professional affilation of the forum post author when making trustworthiness and credibility judgments. Hence, to assess the generalizability of the results, future research needs to replicate this study in different communication settings and with different populations.

Conclusions

When health information seekers are confronted with enthusiastic language in online forums, they may judge the information source as less trustworthy (especially when the information source is a scientist) and deem the communicated information less credible. Furthermore, health information seekers may perceive an information source as more trustworthy when he or she works for a university rather than a lobbying organization. These findings illustrate that health information seekers do not just react to health information on its own. In addition, they are also sensitive to the ways in which health information is presented (“Which language style is used to communicate health information?") and who presents it ("Who does the health information source work for?").

Acknowledgments

The authors would like to thank Dr Celeste Brennecka for language editing and their research assistant Hannah Meinert for her support. The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: this work was supported by the German Research Foundation within the framework of the Research Training Group GRK 1712 Trust and Communication in a Digitized World. The funding was granted to the second author. The German Research Foundation had no involvement in study design, data collection, analysis and interpretation, and the decision to submit the study for publication.

Authors' Contributions

LK and RJ contributed to the conceptualization of the study. LK contributed to the study design. LK was responsible for data acquisition. LK and RJ were responsible for analysis and/or interpretation of data. LK was responsible for drafting the manuscript. LK and RJ were responsible for revising the manuscript critically for important intellectual content. LK and RJ approved of the revision of the manuscript to be published.

Conflicts of Interest

None declared.

References

1. Higgins O, Sixsmith J, Barry MM, Domegan C. A Literature Review on Health Information-Seeking Behaviour on The Web: A Health Consumer and Health Professional Perspective. Stockholm: ECDC; 2011.
2. Fox S, Duggan M. Pew Research Center. 2013. Health Online 2013 URL: http://www.pewinternet.org/2013/01/15/health-online-2013/ [accessed 2018-01-16] [WebCite Cache ID 75vv5Hdn7]
3. Miles J, Petrie C, Steel M. Slimming on the internet. J R Soc Med 2000 May;93(5):254-257 [FREE Full text] [doi: 10.1177/014107680009300510] [Medline: 10884770]
4. Keelan J, Pavri-Garcia V, Tomlinson G, Wilson K. YouTube as a source of information on immunization: a content analysis. J Am Med Assoc 2007 Dec 5;298(21):2482-2484. [doi: 10.1001/jama.298.21.2482] [Medline: 18056901]
5. Pandey A, Patni N, Singh M, Sood A, Singh G. YouTube as a source of information on the H1N1 influenza pandemic. Am J Prev Med 2010 Mar;38(3):e1-e3. [doi: 10.1016/j.amepre.2009.11.007] [Medline: 20171526]
6. Lewandowsky S, Ecker UK, Seifert CM, Schwarz N, Cook J. Misinformation and its correction: continued influence and successful debiasing. Psychol Sci Public Interest 2012 Dec;13(3):106-131. [doi: 10.1177/1529100612451018] [Medline: 26173286]
7. Metzger MJ, Flanagin AJ, Eyal K, Lemus DR, Mccann RM. Credibility for the 21st century: integrating perspectives on source, message, and media credibility in the contemporary media environment. Ann Int Commun Assoc 2003;27(1):293-335. [doi: 10.1080/23808985.2003.11679029]
8. Stadtler M, Bromme R. The content–source integration model: a taxonomic description of how readers comprehend conflicting scientific information. In: Rapp DN, Braasch JL, editors. Processing Inaccurate Information: Theoretical and Applied Perspectives from Cognitive Science and the Educational Sciences. Cambridge: MIT Press; 2014:379-402.
9. Bromme R, Goldman SR. The public's bounded understanding of science. Educ Psychol 2014 Jun 5;49(2):59-69. [doi: 10.1080/00461520.2014.921572]
10. Bromme R, Thomm E. Knowing who knows: laypersons' capabilities to judge experts' pertinence for science topics. Cogn Sci 2016 Jan;40(1):241-252 [FREE Full text] [doi: 10.1111/cogs.12252] [Medline: 25974834]
11. Keil FC, Stein C, Webb L, Billings VD, Rozenblit L. Discerning the division of cognitive labor: an emerging understanding of how knowledge is clustered in other minds. Cogn Sci 2008 Mar 1;32(2):259-300 [FREE Full text] [doi: 10.1080/03640210701863339] [Medline: 19759842]
12. Bromme R, Jucks R. Discourse and expertise: the challenge of mutual understanding between experts and laypeople. In: Schober MF, Rapp DN, Britt MA, editors. The Routledge Handbook of Discourse Processes. New York: Routledge; 2017:222-246.
13. Bromme R, Kienhues D, Porsch T. Who knows what and who can we believe? Epistemological beliefs are beliefs about knowledge (mostly) to be attained from others. In: Bendixen LD, Feucht FC, editors. Personal Epistemology in the Classroom: Theory, Research, and Implications for Practice. Cambridge: Cambridge University Press; 2010:163-194.
14. König L, Jucks R. When do information seekers trust scientific information? Insights from recipients' evaluations of online video lectures. Int J Educ Technol High Educ 2019 Feb 13;16:1. [doi: 10.1186/s41239-019-0132-7]
15. Choi W, Stvilia B. Web credibility assessment: conceptualization, operationalization, variability, and models. J Assoc Inf Sci Technol 2015 May 13;66(12):2399-2414. [doi: 10.1002/asi.23543]

16. Metzger MJ, Flanagin AJ. Psychological approaches to credibility assessment online. In: Sundar SS, editor. The Handbook of the Psychology of Communication Technology. Chichester: John Wiley & Sons; 2015:445-466.
17. Pornpitakpan C. The persuasiveness of source credibility: a critical review of five decades' evidence. J Appl Soc Psychol 2004 Feb;34(2):243-281. [doi: 10.1111/j.1559-1816.2004.tb02547.x]
18. Thon FM, Jucks R. Believing in expertise: how authors' credentials and language use influence the credibility of online health information. Health Commun 2017 Dec;32(7):828-836. [doi: 10.1080/10410236.2016.1172296] [Medline: 27466693]
19. Jucks R, Linnemann GA, Thon FM, Zimmermann M. Trust the words: insights into the role of language in trust building in a digitalized world. In: Blöbaum B, editor. Trust and Communication in a Digitized World. Heidelberg: Springer International Publishing; 2016:225-237.
20. König L, Jucks R. Hot topics in science communication: aggressive language decreases trustworthiness and credibility in scientific debates. Public Underst Sci 2019 Dec;28(4):401-416. [doi: 10.1177/0963662519833903] [Medline: 30843467]
21. Zimmermann M, Jucks R. How experts' use of medical technical jargon in different types of online health forums affects perceived information credibility: randomized experiment with laypersons. J Med Internet Res 2018 Dec 23;20(1):e30 [FREE Full text] [doi: 10.2196/jmir.8346] [Medline: 29362212]
22. Mayweg-Paus E, Jucks R. Evident or doubtful? How lexical hints in written information influence laypersons' understanding of influenza. Psychol Health Med 2015;20(8):989-996. [doi: 10.1080/13548506.2014.986139] [Medline: 25482286]
23. Feinkohl I, Flemming D, Cress U, Kimmerle J. The impact of personality factors and preceding user comments on the processing of research findings on deep brain stimulation: a randomized controlled experiment in a simulated online forum. J Med Internet Res 2016 Mar 3;18(3):e59 [FREE Full text] [doi: 10.2196/jmir.4382] [Medline: 26940848]
24. Flemming D, Cress U, Kimmerle J. Processing the scientific tentativeness of medical research: an experimental study on the effects of research news and user comments in online media. Sci Commun 2017 Nov 3;39(6):745-770 [FREE Full text] [doi: 10.1177/1075547017738091]
25. Bientzle M, Griewatz J, Kimmerle J, Küppers J, Cress U, Lammerding-Koeppel M. Impact of scientific versus emotional wording of patient questions on doctor-patient communication in an internet forum: a randomized controlled experiment with medical students. J Med Internet Res 2015 Nov 25;17(11):e268 [FREE Full text] [doi: 10.2196/jmir.4597] [Medline: 26607233]
26. Flemming D, Cress U, Kimmig S, Brandt M, Kimmerle J. Emotionalization in science communication: the impact of narratives and visual representations on knowledge gain and risk perception. Front Commun 2018 Jan 26;3:1-9. [doi: 10.3389/fcomm.2018.00003]
27. Ghose A, Ipeirotis PG, Sundararajan A. Opinion mining using econometrics?: a case study on reputation systems. In: Zaenen A, van den Bosch A, editors. Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics. Prague: Association for Computational Linguistics; 2007:416-423.
28. Barry C, Hogan M, Torres AM. Framing or gaming? Constructing a study to explore the impact of option presentation on consumers. In: Vogel D, Guo X, Linger H, Barry C, Lang M, Schneider C, editors. Transforming Healthcare Through Information Systems. Heidelberg: Springer; 2016:111-124.
29. Starling S. Evolutionary genetics: fantastic beasts - cephalopod RNA recoding. Nat Rev Genet 2017 Dec;18(6):329. [doi: 10.1038/nrg.2017.31] [Medline: 28420879]
30. Brownlee C, Heald R. The incredible shrinking spindle. Dev Cell 2018 Dec 21;45(4):421-423 [FREE Full text] [doi: 10.1016/j.devcel.2018.05.007] [Medline: 29787705]
31. Keller MM, Hoy AW, Goetz T, Frenzel AC. Teacher enthusiasm: reviewing and redefining a complex construct. Educ Psychol Rev 2015 Dec 23;28(4):743-769. [doi: 10.1007/s10648-015-9354-y]
32. Frenzel AC, Goetz T, Lüdtke O, Pekrun R, Sutton RE. Emotional transmission in the classroom: exploring the relationship between teacher and student enjoyment. J Educ Psychol 2009;101(3):705-716. [doi: 10.1037/a0014695]
33. Keller MM, Goetz T, Becker ES, Morger V, Hensley L. Feeling and showing: a new conceptualization of dispositional teacher enthusiasm and its relation to students' interest. Learn Instr 2014 Oct;33:29-38. [doi: 10.1016/j.learninstruc.2014.03.001]
34. Kunter M, Klusmann U, Baumert J, Richter D, Voss T, Hachfeld A. Professional competence of teachers: effects on instructional quality and student development. J Educ Psychol 2013;105(3):805-820. [doi: 10.1037/a0032583]
35. Patrick BC, Hisley J, Kempler T. 'What's everybody so excited about?': the effects of teacher enthusiasm on student intrinsic motivation and vitality. J Exp Educ 2000 Jan;68(3):217-236. [doi: 10.1080/00220970009600093]
36. Burgoon M, Pauls V, Roberts DL. Language expectancy theory. In: Dillard JP, Pfau M, editors. The Persuasion Handbook: Developments in Theory and Practice. Thousand Oaks: Sage Publications; 2002:117-136.
37. Barney LJ, Griffiths KM, Banfield MA. Explicit and implicit information needs of people with depression: a qualitative investigation of problems reported on an online depression support forum. BMC Psychiatry 2011 May 18;11:88 [FREE Full text] [doi: 10.1186/1471-244X-11-88] [Medline: 21592335]
38. Bhamrah G, Ahmad S, NiMhurchadha S. Internet discussion forums, an information and support resource for orthognathic patients. Am J Orthod Dentofacial Orthop 2015 Jan;147(1):89-96. [doi: 10.1016/j.ajodo.2014.08.020] [Medline: 25533076]

39. Kendal S, Kirk S, Elvey R, Catchpole R, Pryjmachuk S. How a moderated online discussion forum facilitates support for young people with eating disorders. Health Expect 2017 Dec;20(1):98-111 [FREE Full text] [doi: 10.1111/hex.12439] [Medline: 26725547]
40. Sinha A, Porter T, Wilson A. The use of online health forums by patients with chronic cough: qualitative study. J Med Internet Res 2018 Dec 24;20(1):e19 [FREE Full text] [doi: 10.2196/jmir.7975] [Medline: 29367181]
41. Wen H, Shi J, Zhang Y, Lu KH, Cao J, Liu Z. Neural encoding and decoding with deep learning for dynamic natural vision. Cereb Cortex 2018 Dec 1;28(12):4136-4160. [doi: 10.1093/cercor/bhx268] [Medline: 29059288]
42. Appelman A, Sundar SS. Measuring message credibility: construction and validation of an exclusive scale. Journal Mass Commun Q 2015 Oct 5;93(1):59-79. [doi: 10.1177/1077699015606057]
43. Küfner AC, Dufner M, Back MD. The dirty dozen and the wicked nine: short scales for the detection of narcissism, machiavellism and psychopathy. Diagnostica 2015 Jan;61(2):76-91. [doi: 10.1026/0012-1924/a000124]
44. Jonason PK, Webster GD. The dirty dozen: a concise measure of the dark triad. Psychol Assess 2010 Jun;22(2):420-432. [doi: 10.1037/a0019265] [Medline: 20528068]
45. Hendriks F, Kienhues D, Bromme R. Measuring laypeople's trust in experts in a digital age: the Muenster Epistemic Trustworthiness Inventory (METI). PLoS One 2015;10(10):e0139309 [FREE Full text] [doi: 10.1371/journal.pone.0139309] [Medline: 26474078]
46. Dragojevic M, Gasiorek J, Giles H. Communication accommodation theory. In: Berger CR, Roloff ME, editors. The International Encyclopedia of Interpersonal Communication. First Edition. Hoboken: John Wiley & Sons; 2015:1-21.
47. Jasanoff S. Cosmopolitan knowledge: climate science and global civic epistemology. In: Dryzek JS, Norgaard RB, Schlosberg D, editors. The Oxford Handbook of Climate Change and Society. Oxford: Oxford University Press; 2011.
48. Jasanoff S. Designs on Nature: Science and Democracy in Europe and the United States. Princeton: Princeton University Press; 2005.
49. Mitchell KJ, Johnson MK, Mather M. Source monitoring and suggestibility to misinformation: adult age-related differences. Appl Cogn Psychol 2002 Jan;17(1):107-119. [doi: 10.1002/acp.857]
50. Petty RE, Cacioppo JT. The elaboration likelihood model of persuasion. Adv Exp Soc Psychol 1986;19:123-205. [doi: 10.1016/S0065-2601(08)60214-2]

Edited by G Eysenbach; submitted 04.02.19; peer-reviewed by M Gotsis, J Kimmerle; comments to author 27.04.19; revised version received 14.06.19; accepted 14.06.19; published 12.08.19.

Please cite as:
König L, Jucks R
Influence of Enthusiastic Language on the Credibility of Health Information and the Trustworthiness of Science Communicators: Insights From a Between-Subject Web-Based Experiment
Interact J Med Res 2019;8(3):e13619
URL: http://www.i-jmr.org/2019/3/e13619/
doi:10.2196/13619
PMID:31411138

Original Paper

Readability and Quality of Online Information on Osteoarthritis: An Objective Analysis With Historic Comparison

Kieran Edward Murray[1], MB Bch BAOI; Timothy Eanna Murray[2], MB Bch BAOI; Anna Caroline O'Rourke[1], MB Bch BAOI; Candice Low[1], MB BCh BAOI; Douglas James Veale[1], MD

[1]St. Vincent’s University Hospital, University College Dublin, Dublin, Ireland
[2]Beaumont Hospital, Beaumont, Dublin, Ireland

Corresponding Author:
Kieran Edward Murray, MB Bch BAOI
St. Vincent’s University Hospital
University College Dublin
Merrion Road
Elm Park
Dublin
Ireland
Phone: 353 1 221 4000
Email: kemurray@hotmail.com

Abstract

Background: Osteoarthritis (OA) is the most common cause of disability in people older than 65 years. Readability of online OA information has never been assessed. A 2003 study found the quality of online OA information to be poor.

Objective: The aim of this study was to review the readability and quality of current online information regarding OA.

Methods: The term osteoarthritis was searched across the three most popular English language search engines. The first 25 pages from each search engine were analyzed. Duplicate pages, websites featuring paid advertisements, inaccessible pages (behind a pay wall, not available for geographical reasons), and nontext pages were excluded. Readability was measured using Flesch Reading Ease Score, Flesch-Kincaid Grade Level, and Gunning-Fog Index. Website quality was scored using the Journal of the American Medical Association (JAMA) benchmark criteria and the DISCERN criteria. Presence or absence of the Health On the Net Foundation Code of Conduct (HONcode) certification, age of content, content producer, and author characteristics were noted.

Results: A total of 37 unique websites were found suitable for analysis. Readability varied by assessment tool from 8th to 12th grade level. This compares with the recommended 7th to 8th grade level. Of the 37, 1 (2.7%) website met all 4 JAMA criteria. Mean DISCERN quality of information for OA websites was “fair,” compared with the “poor” grading of a 2003 study. HONcode-endorsed websites (43%, 16/37) were of a statistically significant higher quality.

Conclusions: Readability of online health information for OA was either equal to or more difficult than the recommended level.

(Interact J Med Res 2019;8(3):e12855) doi:10.2196/12855

KEYWORDS
osteoarthritis; arthritis; patient; readability; internet

Introduction

Background

Osteoarthritis (OA) causes significant morbidity. It is the most common cause of disability in people older than 65 years and a major contributor to health care cost worldwide [1]. A wide variety of treatments exist, with varying evidence bases. These include physiotherapy, pharmacotherapy, surgery, and alternative therapies [2].

Patient education forms a crucial role in decision making, medication adherence, and disease self-management [3,4]. Traditionally, physicians represented the primary, or even sole, source of patient information. However, the patient-doctor relationship and the flow of information are changing. This is partially because of increasing internet access and an increase in the depth and breadth of online content. The number of patients searching online for health information and the number of health-related websites continues to grow, and there are now over 70,000 websites providing health information [5,6]. Most patients now use the internet for health care information [7].

Despite an increase in the availability and usage of online health information, the readability and quality of online health information is variable [8]. At present, a number of standardized validated tools are available to assess both the readability (Flesch Reading Ease Score, FRES; Flesch-Kincaid Grade Level, FKGL; and Gunning-Fog Index, GFI) and quality (*Journal of the American Medical Association*, JAMA benchmark criteria; DISCERN instrument) of online health care information.

Guidelines state health information aimed at the general public should be at a 7th to 8th grade reading level (United States) [9]. Previous studies of other medical conditions have found most online health care information to be above this reading level, rendering it inaccessible to many patients [10-12]. However, the readability of online OA content has never been assessed.

A 2003 study, using the standardized and validated DISCERN tool, graded online information concerning OA as poor [2]. Since then, there has been a marked increase in the number of websites from 40,912,332 in 2003 to 1,329,189,590 at present [13]. A recent study of the *knee osteoarthritis treatment* showed a significant difference in the quality of online information between countries speaking different languages [14].

Objectives

Given the significant morbidity of OA, the lack of a previous study assessing readability, and the lack of any recent study (<10 years) assessing the quality of online information in relation to OA in general, the aims of this study were to assess both the readability and quality of current online OA information using 6 previously validated tools.

Methods

Internet Search Strategy

A total of 2 authors (KEM and TEM) familiar with the topic selected the 4 most appropriate commonly appearing disease-specific search terms for review: *osteoarthritis*, *osteoarthrosis*, *degenerative arthritis*, and *degenerative joint disease*. These were then searched across the 3 most popular UK search engines (Google, Bing, and Yahoo!), collectively representing over 98% of UK searches [15]. *Osteoarthritis* provided the most Web addresses, also known as URLs, and was thus chosen for analysis. As research has shown that patients are unlikely to search beyond 25 pages [16], the most-viewed (top ranking) 25 websites on each search engine were included [10,17].

Inclusion criteria were the first 25 pages from Google, Bing, or Yahoo! (n=75). Duplicate websites (n=31) and nonreadable websites (n=7) were excluded. The nonreadable websites were nontext pages (n=3), paywall-protected websites (n=2) and those inaccessible for geographic reasons (n=2). Of the 75 studied websites, 38 met exclusion criteria, and 37 were considered suitable for analysis. In cases of pagination of the webpage (where a single item was spread across sequential pages on the same website), the sequential pages were also assessed. All websites were reviewed in January 2018.

Website producer (the group responsible for hosting or publishing the website) was defined as health care provider, professional society, for-profit organization, or not-for-profit organization {including governmental organizations and nongovernmental organizations (NGOs)}. Where a website detailed dates for both content creation and last update, the most recent date was used when measuring website currency.

Website authors (and reviewers where specified) were categorized into doctor, other medical professional, nonspecified medical staff, nonmedical author, or not reported. Websites required explicit naming of authorship to comply with JAMA guidelines.

Readability

Using an online analysis tool, the readability of each website was evaluated for 3 validated scores: FRES, FKGL, and GFI [18].

Published in 1948, the FRES calculates readability using the formula 206.835 − 1.015 (total words/total sentences) − 84.6 (total syllables/total words). This generates a difficulty grading from 0 to 100, with higher scores indicating easier readablitiy [19]. FKGL was developed by the US Navy in 1975. It assesses readability with the following formula: 0.39 (total words/total sentences) + 11.8 (total syllables/total words) − 15.59 [20]. The GFI calculates readability with the formula 0.4 ([words/sentences] + 100 [complex words/words]). However, it also acknowledges a list of common words that are not considered complex, despite their syllable count. This forms an estimate as to the years of education required for readability [17,21].

Quality

The quality of each selected website was evaluated using Health On the Net Foundation (HON) Code of Conduct (HONcode) classification, JAMA benchmark criteria, and DISCERN score, all of which have been previously validated [10,17].

Founded in 1995, HON is a nonprofit, nongovernmental organization, accredited to the Economic and Social Council of the United Nations. It was created to promote the spread of quality health information around the world [22]. HONcode is perhaps the best-known quality label for medical and health websites. It was created to help standardize the reliability of medical and health information available online [23,24]. HONcode certification identifies websites with quality and nonbiased health information designed for patients [23]. It examines numerous factors including disclosure of authors' qualifications, attribution of sources, complementarity to the doctor-patient relationship, data protection, justifiability, transparency, and disclosure of funding sources and advertising. Over 8000 sites have been certified [25]. Each website was checked against the HONcode database.

Published in 1997, the JAMA benchmark criteria list 4 criteria that quality websites should fulfill. These are identification of authorship, identification of sources, specifying the date of creation/update, and disclosures (of ownership, advertising policy, sponsorship, and conflicts of interests) [26]. The presence or absence of each of these parameters was recorded. The content producer of the website was taken from the webpage itself or the Contact Us/About Us tab.

Published in 1999, DISCERN is an instrument created by British universities, the National Health Service, and the British Library [27]. It determines website quality and reliability by grading 16 items (concerning reliability, description of treatment choices, and overall rating) from 1 (inferior) to 5 (superior). The score ranges from 16 to 80, with a higher score indicating better-quality information [27]. Study grading was performed by a single author (KEM), with consensus joint scoring (TEM) in cases of uncertainty.

Statistical Methods

Mean website age, JAMA benchmark criteria, and DISCERN score for each website were reviewed with 1-way analysis of variance (ANOVA). Analysis was performed by Prism 7 (GraphPad software). Significance was set at $P<.05$.

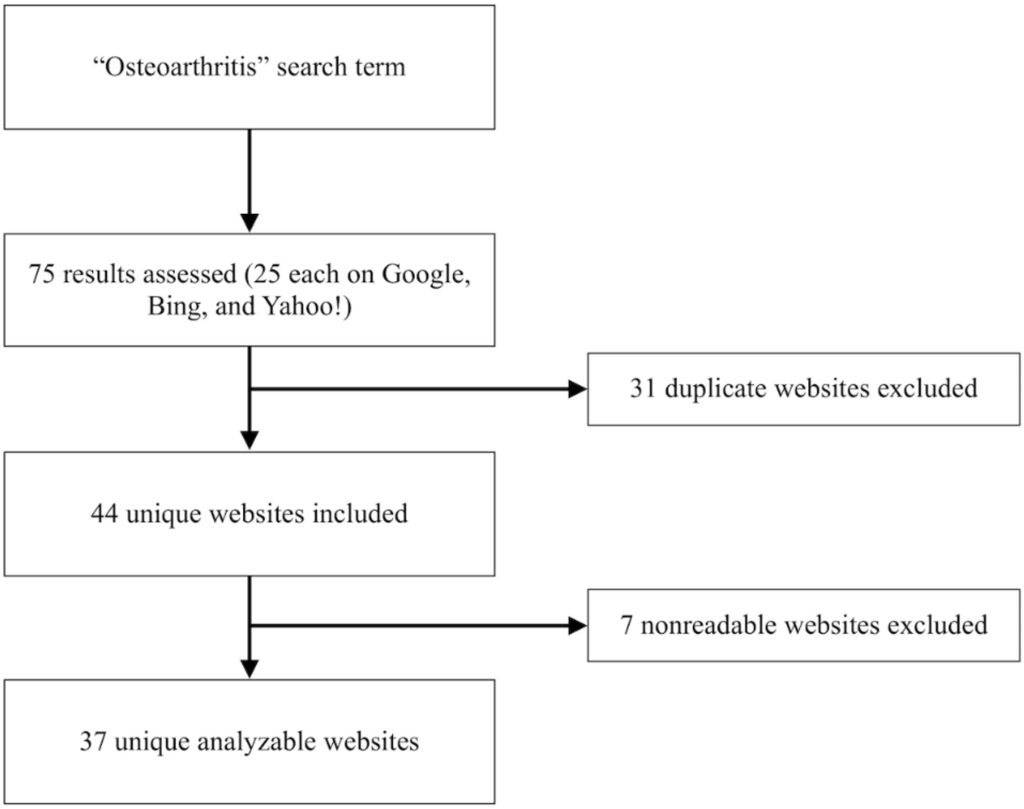
Results

Internet Search Strategy

Osteoarthritis was the most-searched term, with 33,960,000 results across Google, Bing, and Yahoo!. Thus, it was selected for further analysis. *Degenerative arthritis* (21,790,000), *degenerative joint disease* (7,720,000), and *osteoarthrosis* (1,139,000) yielded fewer results and were, therefore, disregarded.

The internet search strategy is summarized in Figure 1. Of 75 articles, 38 met exclusion criteria. Of these, 31 were duplicate websites and 7 were nonreadable (nontext pages (n=3), paywall-protected websites (n=2), or inaccessible for geographic reasons (n=2). A total of 37 websites were considered suitable for analysis.

Figure 1. Internet search strategy (January 2018).



Readability

The mean GFI of websites was 9.0, indicating a 9th grade reading level, mean FRES score was 51.4 (10th-12th grade reading ability), and mean FKGL score was 7.8 (8th grade). Readability scores by website content producer are shown in Tables 1-3.

There was no significant correlation between the type of organization publishing the website and readability, as measured by FRES (ANOVA r^2=0.01; P=.93), FKGL (r^2=0.01; P=.13), or GFI (r^2=0.02; P=.88). Similarly, there was no significant correlation between the author type and website readability, as measured by FRES (ANOVA r^2=0.18; P=.09), FKGL (r^2=0.20; P=.06), or GFI (r^2=0.14; P=.16).

Table 1. Quality and readability of online information on osteoarthritis.

Producer	Mean age (years)	HONcode[a] certified, n (%)	Mean DISCERN score			
			Readability	Treatment choices	Quality	Total
All	1.4	16 (43)	23.1	16.6	2.6	42.3
Not-for-profit (governmental and NGOs; n=14)	0.9	3 (21)	24.1	17.2	2.6	43.9
Professional society (n=4)	1.3	0 (0)	22.6	16.3	2.5	41
For-profit organization (n=15)	1.5	12 (80)	23	16.9	2.6	42.5
Health care providers (n=4)	0.4	1 (25)	20.5	14	2.3	36.8

[a] HONcode: Health On the Net Foundation Code of Conduct.

Table 2. Quality and readability of online information on osteoarthritis.

Producer	Fulfill JAMA[a] benchmark criteria, n (%)			
	Authorship	Attribution	Currency	Disclosure
All	11 (30)	9 (24)	22 (59)	9 (24)
Not-for-profit (governmental and NGOs; n=14)	0 (0)	3 (21)	7 (50)	3 (21)
Professional society (n=4)	1 (50)	2 (25)	3 (75)	1 (25)
For-profit organization (n=15)	8 (53)	4 (27)	11 (73)	5 (33)
Health care providers (n=4)	0 (0)	1 (25)	1 (25)	0 (0)

[a] JAMA: Journal of the American Medical Association.

Table 3. Quality and readability of online information on osteoarthritis.

Producer	Mean readability score		
	FRES[a]	FKGL[b]	GFI[c]
All	51.4	7.8	9.0
Not-for-profit (governmental and NGOs; n=14)	50.8	7.9	9.4
Professional society (n=4)	49.4	8.2	8.8
For-profit organization (n=15)	53.2	7.6	8.6
Health care providers (n=4)	48.5	8.2	9.2

[a] FRES: Flesch Reading Ease Score.

[b] FKGL: Flesch-Kincaid Grade Level.

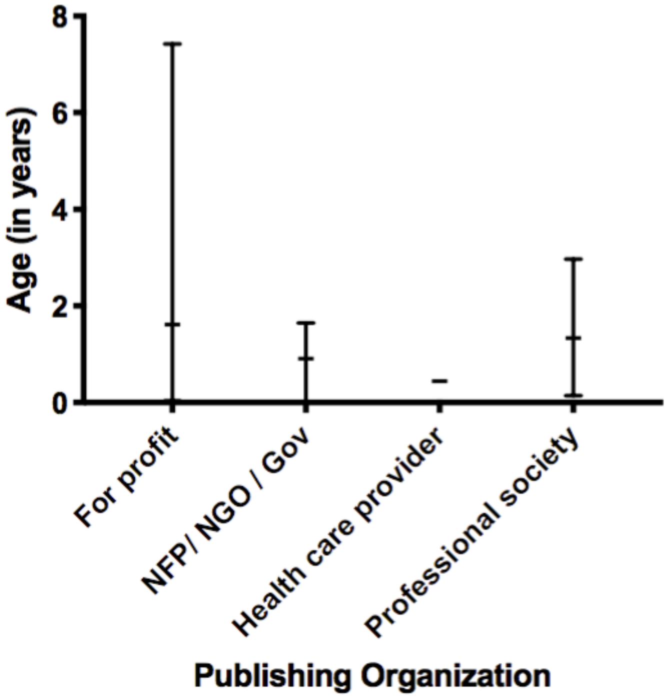
[c] GFI: Gunning-Fog Index.

Journal of the American Medical Association Benchmark Criteria

Each website was checked for compliance with JAMA benchmark criteria. Overall, 1 of 37 websites (3%) detailed author, attribution, currency, age, and disclosures. This was written by a doctor for a for-profit organization. For 23 websites (62%, 23/37) the author was not reported. Reported authors/reviewers were doctors (n=8), other health professionals (n=3), nonmedical author (n=2), and nonspecified medical staff (n=1). A total of 59% (22/37) of websites recorded the date of publication or update. The mean content age was 522 days (1.4 years).

The currency of the websites (time of last update or creation, whichever was most recent) did not significantly vary between publishing organization types (ANOVA r^2=0.05; P=.79; Figure 2). There was no correlation between the search engine ranking of the website and JAMA score (r=−0.004; P=.97).

Figure 2. Website age by publishing organization type. Horizontal dash indicates mean and bar represents range. NFP: not for profit; NGO: nongovernmental organization.

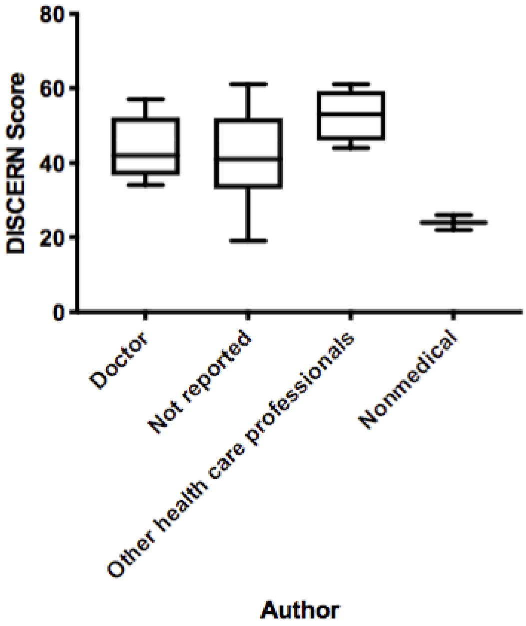


DISCERN Score

The mean DISCERN scores for each of the 16 assessments of quality are shown in Multimedia Appendix 1. Overall, mean DISCERN score was 42.3. Thus, the quality of online health information for OA was “fair” [28], comparing favorably with the “poor” grading in 2003 [2]. The website with the highest DISCERN score (61) was *Mayo Clinic Patient Care and Health Information on Osteoarthritis* [29]. This website is also HONcode certified.

There was a significant difference in quality among author types (ANOVA r^2=0.24; *P*=.03), with nondoctor health professional authors scoring the highest (mean 52.8 DISCERN score, SD 7.0) and nonmedical authors scoring the lowest (mean 24.0, SD 2.8; Figure 3). There was no correlation between the search engine ranking of the website and quality, as measured by either the DISCERN score (Spearman rank-order correlation *r*=0.05; *P*=.78). There was no significant difference in website quality (by DISCERN score) among publishing organization (ANOVA r^2=0.14; *P*=.32).

Figure 3. Relationship between website quality and author type. Box and whisker plot with horizontal line representing the median value, the box representing the interquartile range, and whiskers representing the range.



Health On the Net Foundation Code of Conduct Certification

Of 37 websites, 16 (43%) were HONcode certified. HONcode certification was significantly correlated with website quality as measured by DISCERN (unpaired 2-tailed *t* test; *P*=.001) and JAMA criteria (*P*=.02). HONcode certification was not, however, correlated with readability, as measured by FRES (*P*=.32), FKGL (*P*=.28), or GFI (*P*=.63) scores.

Rates of HONcode certification varied significantly among different publishing organization types (ANOVA r^2=0.34; *P*=.002). This was highest among for-profit organizations (80%), with health care providers (25%), not-for-profit organizations (21%), and professional societies (0%) faring worse (Multimedia Appendix 2).

Discussion

The mean readability of the websites differed somewhat by scoring technique. GFI indicated a 9th grade reading level. Mean FRES readability score was at the level of 10th to 12th grade reading ability, and FKGL score suggested an 8th grade reading ability. Thus, a large section of these websites exceeded the recommended 7th to 8th grade levels.

The quality of online OA health information (as per DISCERN score) has improved from "poor" in a 2003 study to "fair." HONcode certification significantly correlated with website quality. There was a significant difference in quality between author types. Interestingly, doctors did not rank highest (although it is possible that they may have authored websites where the author was not reported). Nondoctor health professionals scored highest, followed by doctors and nonmedical authors.

Previous studies have shown that patients are much more likely to view pages with a higher search engine ranking [16]. We found no correlation between the search engine ranking of the website and quality as measured by either the DISCERN score or JAMA benchmark criteria. This indicates that a higher search engine ranking is not predictive of higher content quality. There was also no significant difference in website quality (by DISCERN score) among publishing organization.

HONcode certification was high relative to similar studies for other conditions [10,30]. This may reflect the high prevalence and widespread public interest in OA. HONcode certification significantly correlated with website quality (as measured by DISCERN and JAMA criteria), suggesting this certification may provide both clinicians and patients with a useful aid to predict high-quality online information.

There were several limitations to this study. Website search strategy was different from that of the comparative study performed in 2003 (reflecting the wider variety of website analytic tools now available and employed in this study). This reflects the much-changed nature of Web-searching behavior over the intervening 15 years. This difference in study modality must be considered when comparing study findings. Although the number of pages analyzed in this study was limited to those appearing high up in the search engine order, this reflects previously researched patient-research patterns, where pages beyond 25 are rarely viewed [16]. Analyzing additional pages may provide additional statistical certainty, but this would not reflect patient search patterns. A further limitation is the analysis of only 3 search engines, which could feasibly limit the applicability of the results. These combined search engines represent 98% of the target audience's internet searches; however, it is unlikely that additional search engine inclusion in this study would add value [15].

JAMA benchmarks and DISCERN criteria use a methodical approach to assess quality. Despite this, all such grading introduces some degree of subjectivity. Ultimately, patients are the intended target audience for this study. In this study, website scoring was performed by doctors. Regardless, the DISCERN criteria was developed for use by either health professionals or the general public, and numerous peer-reviewed studies have demonstrated high interrater agreement [27,31,32].

The readability assessment tools are objective and precise for written text, owing to their computerized calculation. Other website components, however, such as videos and images, can also affect understanding, and these are not analyzed using the readability assessment tools. This limitation has been noted in previous similar studies [17,33].

Internet use is widespread. As health care providers, it is important to develop and direct patients toward readable, high-quality online health information. This study suggests health care professionals should direct patients to HONcode-certified websites written by health care professionals, as these websites were of significantly higher quality. The single highest quality source is noted to be a patient information website from the Mayo Clinic [29]. From a policy perspective, readability remains an important issue. Online health information for the general public released by US governmental websites should comply with the National Library of Medicine guidelines (7th-8th grade reading level) [9]. When creating websites to provide patient information on OA, authors may use the JAMA benchmark criteria and DISCERN criteria to ensure compliance with quality standards. Authors should consider testing readability levels to ensure the content is appropriately targeted at a 7th to 8th grade reading level. Finally, authors should consider HONcode certification to provide an external assessment of website quality.

Authors' Contributions

KEM, TEM, and DJV contributed to the conception and design. KEM and TEM were responsible for the analysis of quality and readability. TEM, KEM, and DJV were responsible for the analysis and interpretation of data. KEM, TEM, and DJV were responsible for revising it critically for important intellectual content. KEM, TEM, CL, ACOR, and DJV were involved in drafting the manuscript. KEM, TEM, CL, ACOR, and DJV read and approved the final version to be published.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Mean DISCERN scores.
[PDF File (Adobe PDF File)87 KB - ijmr_v8i3e12855_app1.pdf]

Multimedia Appendix 2
Health On the Net Foundation Code of Conduct (HONcode) certification by website publishing organization type. NFP: not for profit; NGO: nongovernmental organization.
[PDF File (Adobe PDF File)42 KB - ijmr_v8i3e12855_app2.pdf]

References

1. Brooks P. Managing osteoarthritis and other chronic diseases. Lancet 2004 Feb;363(9407):498. [doi: 10.1016/S0140-6736(04)15514-1]
2. Maloney S, Ilic D, Green S. Accessibility, nature and quality of health information on the internet: a survey on osteoarthritis. Rheumatology (Oxford) 2005 Mar;44(3):382-385. [doi: 10.1093/rheumatology/keh498] [Medline: 15572390]
3. Abbott SA. The benefits of patient education. Gastroenterol Nurs 1998;21(5):207-209. [Medline: 9830962]
4. Lorig KR, Mazonson PD, Holman HR. Evidence suggesting that health education for self-management in patients with chronic arthritis has sustained health benefits while reducing health care costs. Arthritis Rheum 1993 Apr;36(4):439-446. [doi: 10.1002/art.1780360403] [Medline: 8457219]
5. Cline RJ, Haynes KM. Consumer health information seeking on the internet: the state of the art. Health Educ Res 2001 Dec;16(6):671-692. [doi: 10.1093/her/16.6.671] [Medline: 11780707]
6. Grandinetti DA. Doctors and the web. Help your patients surf the net safely. Med Econ 2000 Mar 6;77(5):186-8, 194. [Medline: 10848409]
7. Pew Research Center. 2019. Internet/Broadband Fact Sheet URL: http://www.pewinternet.org/fact-sheet/internet-broadband [accessed 2018-02-04]
8. Berland GK, Elliott MN, Morales LS, Algazy JI, Kravitz RL, Broder MS, et al. Health information on the internet: accessibility, quality, and readability in English and Spanish. J Am Med Assoc 2001;285(20):2612-2621 [FREE Full text] [doi: 10.1001/jama.285.20.2612] [Medline: 11368735]
9. MedlinePlus. 2017. How to Write Easy-to-Read Health Materials URL: https://medlineplus.gov/etr.html [accessed 2018-02-03]
10. Murray TE, Mansoor T, Bowden DJ, O'Neill DC, Lee MJ. Uterine artery embolization: an analysis of online patient information quality and readability with historical comparison. Acad Radiol 2018 May;25(5):619-625. [doi: 10.1016/j.acra.2017.11.007] [Medline: 29331361]
11. Akinleye SD, Krochak R, Richardson N, Garofolo G, Culbertson MD, Erez O. Readability of the most commonly accessed arthroscopy-related online patient education materials. Arthroscopy 2018 Apr;34(4):1272-1279. [doi: 10.1016/j.arthro.2017.09.043] [Medline: 29287948]
12. Daraz L, Morrow AS, Ponce OJ, Farah W, Katabi A, Majzoub A, et al. Readability of online health information: a meta-narrative systematic review. Am J Med Qual 2018;33(5):487-492. [doi: 10.1177/1062860617751639] [Medline: 29345143]
13. Internet Live Stats. 2018. Total number of websites URL: http://www.internetlivestats.com/total-number-of-websites/2018 [accessed 2018-02-03]
14. Varady NH, Dee EC, Katz JN. International assessment on quality and content of internet information on osteoarthritis. Osteoarthritis Cartilage 2018 Aug;26(8):1017-1026 [FREE Full text] [doi: 10.1016/j.joca.2018.04.017] [Medline: 29758353]
15. Statista. 2018. Market Share Held by the Leading Search Engines in the United Kingdom (UK) as of June 2019 URL: https://www.statista.com/statistics/280269/market-share-held-by-search-engines-in-the-united-kingdom/ [accessed 2018-01-24]
16. Eysenbach G, Köhler C. How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. Br Med J 2002 Mar 9;324(7337):573-577 [FREE Full text] [doi: 10.1136/bmj.324.7337.573] [Medline: 11884321]
17. O'Neill SC, Baker JF, Fitzgerald C, Fleming C, Rowan F, Byrne D, et al. Cauda equina syndrome: assessing the readability and quality of patient information on the internet. Spine (Phila Pa 1976) 2014 May 1;39(10):E645-E649. [doi: 10.1097/BRS.0000000000000282] [Medline: 24583736]
18. WebFX. Readability Test Tool URL: https://www.webpagefx.com/tools/read-able/ [accessed 2018-02-04]

19. Kincaid JP, Fishburne Jr RP, Rogers RL, Chissom BS. ERIC - Education Resources Information Center. 1975. Derivation of New Readability Formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy Enlisted Personnel URL: https://eric.ed.gov/?id=ED108134 [accessed 2019-08-09]
20. McClure GM. Readability formulas: useful or useless? IEEE Trans Profess Commun 1987 Mar;PC-30(1):12-15. [doi: 10.1109/TPC.1987.6449109]
21. Walsh TM, Volsko TA. Readability assessment of internet-based consumer health information. Respir Care 2008 Oct;53(10):1310-1315 [FREE Full text] [Medline: 18811992]
22. Health On the Net. 2017. The Commitment to Reliable Health and Medical Information on the Internet URL: https://www.hon.ch/HONcode/Pro/Visitor/visitor.html [accessed 2018-02-04]
23. Boyer C, Selby M, Scherrer JR, Appel R. The Health On the Net code of conduct for medical and health websites. Comput Biol Med 1998 Sep;28(5):603-610. [Medline: 9861515]
24. Wilson P, Risk A. How to find the good and avoid the bad or ugly: a short guide to tools for rating quality of health information on the internet. Br Med J 2002 Mar 9;324(7337):598-602 [FREE Full text] [doi: 10.1136/bmj.324.7337.598] [Medline: 11884329]
25. Health On the Net. 2016. Discover the HONcode Certification URL: https://www.hon.ch/20-years/en/ [accessed 2018-02-04]
26. Silberg WM, Lundberg GD, Musacchio R. Assessing, controlling, and assuring the quality of medical information on the internet: caveant lector et viewor--let the reader and viewer beware. J Am Med Assoc 1997 Apr 16;277(15):1244-1245. [doi: 10.1001/jama.1997.03540390074039] [Medline: 9103351]
27. Charnock D, Shepperd S, Needham G, Gann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. J Epidemiol Community Health 1999 Feb;53(2):105-111 [FREE Full text] [doi: 10.1136/jech.53.2.105] [Medline: 10396471]
28. Hargrave DR, Hargrave UA, Bouffet E. Quality of health information on the internet in pediatric neuro-oncology. Neuro Oncol 2006 Apr;8(2):175-182 [FREE Full text] [doi: 10.1215/15228517-2005-008] [Medline: 16533758]
29. Staff MC. Mayo Clinic. 2017. Osteoarthritis: Overview URL: https://www.mayoclinic.org/diseases-conditions/osteoarthritis/symptoms-causes/syc-20351925 [accessed 2018-01-23]
30. Thakurdesai PA, Kole PL, Pareek R. Evaluation of the quality and contents of diabetes mellitus patient education on internet. Patient Educ Couns 2004 Jun;53(3):309-313. [doi: 10.1016/j.pec.2003.04.001] [Medline: 15186868]
31. Charnock D, Shepperd S. Learning to DISCERN online: applying an appraisal tool to health websites in a workshop setting. Health Educ Res 2004 Aug;19(4):440-446. [doi: 10.1093/her/cyg046] [Medline: 15155597]
32. Ademiluyi G, Rees CE, Sheard CE. Evaluating the reliability and validity of three tools to assess the quality of health information on the internet. Patient Educ Couns 2003 Jun;50(2):151-155. [doi: 10.1016/S0738-3991(02)00124-6] [Medline: 12781930]
33. Sabharwal S, Badarudeen S, Kunju SU. Readability of online patient education materials from the AAOS web site. Clin Orthop Relat Res 2008 May;466(5):1245-1250 [FREE Full text] [doi: 10.1007/s11999-008-0193-8] [Medline: 18324452]

Abbreviations

ANOVA: analysis of variance
FKGL: Flesch-Kincaid Grade Level
FRES: Flesch Reading Ease Score
GFI: Gunning-Fog Index
HON: Health On the Net Foundation
HONcode: Health On the Net Foundation Code of Conduct
JAMA: Journal of the American Medical Association
OA: osteoarthritis

Edited by G Eysenbach; submitted 19.11.18; peer-reviewed by L Daraz, K Sward; comments to author 31.03.19; revised version received 25.05.19; accepted 12.06.19; published 16.09.19.

Please cite as:
Murray KE, Murray TE, O'Rourke AC, Low C, Veale DJ
Readability and Quality of Online Information on Osteoarthritis: An Objective Analysis With Historic Comparison
Interact J Med Res 2019;8(3):e12855
URL: https://www.i-jmr.org/2019/3/e12855
doi:10.2196/12855
PMID:31538953

Original Paper

Physician Decision-Making Patterns and Family Presence: Cross-Sectional Online Survey Study in Japan

Kenji Tsuda[1], MD, PhD; Asaka Higuchi[1], RN; Emi Yokoyama[2], RN, MA; Kazuhiro Kosugi[3], MD; Tsunehiko Komatsu[4], MD, PhD; Masahiro Kami[1], MD, PhD; Tetsuya Tanimoto[5], MD

[1]Medical Governance Research Institute, Minatoku, Japan
[2]Graduate School of Education, Seisa University, Yokohama, Japan
[3]National Cancer Center Hospital East, Kashiwa, Japan
[4]Jyoban Hospital of Tokiwa Foundation, Iwaki, Japan
[5]Navitas Clinic, Tachikawa, Japan

Corresponding Author:
Kenji Tsuda, MD, PhD
Medical Governance Research Institute
108-0074
Minatoku,
Japan
Phone: 81 364557401
Email: thedod3@hotmail.com

Abstract

Background: Due to a low birth rate and an aging population, Japan faces an increase in the number of elderly people without children living in single households. These elderly without a spouse and/or children encounter a lack of caregivers because most sources of care for the elderly in Japan are not provided by private agencies but by family members. However, family caregivers not only help with daily living but are also key participants in treatment decision making. The effect of family absence on treatment decision making has not been elucidated, although more elderly people will not have family members to make surrogate decisions on their behalf.

Objective: The aim is to understand the influence of family absence on treatment decision making by physicians through a cross-sectional online survey with three hypothetical vignettes of patients.

Methods: We conducted a cross-sectional online survey among Japanese physicians using three hypothetical vignettes. The first vignette was about a 65-year-old man with alcoholic liver cirrhosis and the second was about a 78-year-old woman with dementia, both of whom developed pneumonia with consciousness disturbance. The third vignette was about a 70-year-old woman with necrosis of her lower limb. Participants were randomly assigned to either of the two versions of the questionnaires—with family or without family—but methods were identical otherwise. Participants chose yes or no responses to questions about whether they would perform the presented medical procedures.

Results: Among 1112 physicians, 454 (40.8%) completed the survey; there were no significant differences in the baseline characteristics between groups. Significantly fewer physicians had a willingness to perform dialysis (odds ratio [OR] 0.55, 95% CI 0.34-0.80; *P*=.002) and artificial ventilation (OR 0.51, 95% CI 0.35-0.75; *P*<.001) for a patient from vignette 1 without family. In vignette 2, fewer physicians were willing to perform artificial ventilation (OR 0.59, 95% CI 0.39-0.90; *P*=.02). In vignette 3, significantly fewer physicians showed willingness to perform wound treatment (OR 0.51, 95% CI 0.31-0.84; *P*=.007), surgery (OR 0.35, 95% CI 0.22-0.57; *P*<.001), blood transfusion (OR 0.45, 95% CI 0.31-0.66; *P*<.001), vasopressor (OR 0.49, 95% CI 0.34-0.72; *P*<.001), dialysis (OR 0.38, 95% CI 0.24-0.59; *P*<.001), artificial ventilation (OR 0.25, 95% CI 0.15-0.40; *P*<.001), and chest compression (OR 0.29, 95% CI 0.18-0.47; *P*<.001) for a patient without family.

Conclusions: Elderly patients may have treatments withheld because of the absence of family, highlighting the potential importance of advance care planning in the era of an aging society with a declining birth rate.

(Interact J Med Res 2019;8(3):e12781) doi:10.2196/12781

KEYWORDS

physician decision making; family; cross-sectional study; cardiopulmonary resuscitation; artificial ventilation

Introduction

With the progress of a low birth rate and an aging population, Japan faces an increase of elderly people without children living in single households. In 2016, 4% of people aged 60 years or older did not have children and were living alone [1]. These elderly without spouses and/or children result in a lack of caregivers because most care for the elderly in Japan is not provided by private agencies but by family members. Sixty percent of care is provided by family members living together, half of which is done by the spouse and the other half by children or spouses of children [2].

Family caregivers are not only helpers of daily living but are key participants in the treatment decision making [3]. More than one-quarter of hospitalized patients lack the capacity to make decisions about their care because of the illness itself or chronic cognitive disorders such as dementia. As such, family members often play the role of surrogates to make decisions on behalf of the patients [4]. For intensive care unit physicians, family directives were the fifth most important factor in withdrawal of life support and valued more than age and premorbid physical function of patients [5]. However, 16% of patients in intensive care units and 3% of nursing home residents had no designated surrogate and no identifiable family members to speak on their behalf [4,6,7]. In such cases, decisions to limit life support are generally made by physicians [6,8].

The effect of family absence on physician decision making has not been elucidated. Therefore, we conducted a cross-sectional online survey with three hypothetical vignettes of patients to estimate the influence of family absence on treatment decision making by Japanese physicians.

Methods

On February 2018, we distributed an online survey via email to all Japanese physicians registered to the mailing list offering job information provided by Mediwel Co, Ltd. In Japan, approximately 60% of physicians take on a part-time job in addition to their full-time job, and many physicians register to this mailing list as a source of information. The coverage of the database is not open to the public, but it is rated as one of the largest platforms for job-seeking in Japan. This is the first survey using this mailing list for research. Participation in our survey was voluntary without any incentives.

For the baseline characteristics, physicians provided their age, sex, years since graduation, type of institution that they worked in, number of times they have explained a patient's condition to his or her family per year, and specialty. No other personal information was collected.

Using the online survey, anonymity was secured, and physicians could answer honestly without fear of reputational risk.

Participants were informed that the aim of the study was to investigate physicians' decision making. To avoid biases, we did not mention the survey purpose was to investigate the effect of family absence. Participants were randomly assigned to either of two versions of the vignettes: without family (pattern A) or with family (pattern B). Each pattern included three vignettes; these were identical except for the description of family presence or absence. Allocation was randomized through the A/B testing function of the mail delivery system, which is often used in internet marketing to compare two versions of a webpage. The deadline for answers was one week later.

Details of the vignettes are shown in Multimedia Appendices 1 and 2. Briefly, the first vignette described a 65-year-old man with alcoholic liver cirrhosis, and the second described a 78-year-old woman with advanced dementia. Both of these patients developed pneumonia with impairment of consciousness and were taken to a hospital by an ambulance. In vignettes 1 and 2, patients did not have an advanced directive, and their preferences for treatment were unknown. The third vignette described a 70-year-old woman with necrosis of her lower limb, which needed amputation to save her life. The patient did not want amputation nor cardiopulmonary resuscitation. However, in pattern B, the woman's family wanted all possible treatments even though the patient refused to receive the procedures. To each vignette, respondents chose yes or no answers to questions for approximately 10 medical procedures about whether they would perform procedures necessary for the patient.

The primary endpoint was not set because there were no similar previous studies. This was intended as an exploratory survey, but we targeted for at least 100 participants in each pattern. All acquired data were analyzed, and no kinds of sampling were performed. Statistical analyses were performed using R version 3.3.3 (R Foundation for Statistical Computing, Vienna, Austria). Between patterns A and B, differences in age and years since graduation were tested by Student *t* tests; other baseline characteristics were tested by chi-square tests. Physicians' choices were compared between patterns A and B for each procedure by chi-square tests and Fisher exact tests. Subanalyses investigating differences in sex, experience, and type of institution were performed for procedures that showed significant differences between patterns A and B. Experienced physicians were defined as aged 41 years or older. Univariate analysis was performed, and factors related to withholding of artificial ventilation were identified. All variables with a *P* value <.10 in the initial univariate analysis were considered a potential influencing factor in the multivariate logistic regression model. A *P* value of <.05 was considered statistically significant. This study was approved by the institutional review board of the Medical Governance Research Institute, Tokyo, Japan.

Results

Among 1112 physicians who opened the link attached to the invitation email, 454 (40.8%) physicians completed the survey. There were no significant differences in the baseline characteristics between groups (Table 1). In vignette 1, for a patient without family compared with a patient with family, significantly fewer physicians were willing to perform dialysis (odds ratio [OR] 0.55, 95% CI 0.34-0.80; *P*=.002) and artificial ventilation (OR 0.51, 95% CI 0.35-0.75; *P*<.001). In vignette 2, fewer physicians were willing to perform artificial ventilation (OR 0.59, 95% CI 0.39-0.90; *P*=.02; Figure 1). Similarly, in vignette 3, significantly fewer physicians were willing to perform wound treatment (OR 0.51, 95% CI 0.31-0.84; *P*=.007), surgery (OR 0.35, 95% CI 0.22-0.57; *P*<.001), blood transfusion (OR 0.45, 95% CI 0.31-0.66; *P*<.001), vasopressor (OR 0.49, 95% CI 0.34-0.72; *P*<.001), dialysis (OR 0.38, 95% CI 0.24-0.59; *P*<.001), artificial ventilation (OR 0.25, 95% CI 0.15-0.40; *P*<.001), and chest compression (OR 0.29, 95% CI 0.18-0.47; *P*<.001) for a patient without family.

As shown in Table 2, male physicians significantly withheld all 10 procedures in patients without family. For female physicians, the odds ratios of all 10 procedures were less than 1, but only a few procedures were statistically significant. This was probably because there were fewer female participants compared with male participants in this study (n=84 and n=370, respectively). Table 3 shows that experienced physicians significantly withheld all 10 procedures in patients without family. For nonexperienced physicians, odds ratios of all 10 procedures were less than 1, but only a few procedures were statistically significant. As shown in Table 4, physicians working at a non-acute care hospital significantly withheld all 10 procedures in patients without family. In physicians working at an acute care hospital, the odds ratios of all 10 procedures were less than 1, but only a few procedures were statistically significant.

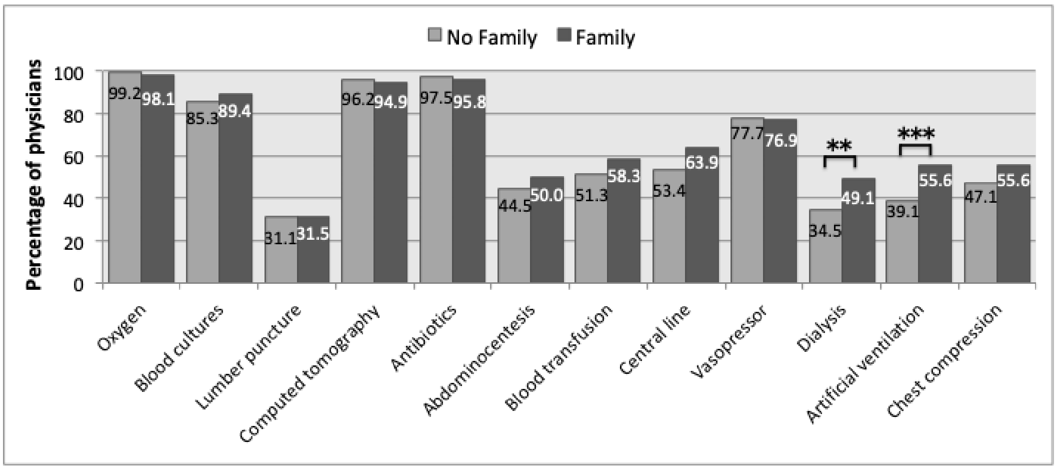
As shown in Table 5, in univariate analysis, years since graduation and pattern were found to be associated with withholding artificial ventilation in vignette 1. Institution and pattern were significant in vignette 2, and only pattern was significant in vignette 3. In a multivariate analysis, pattern was the only factor that was associated with withholding artificial ventilation through all vignettes.

Table 1. Baseline characteristics of study participants (N=454).

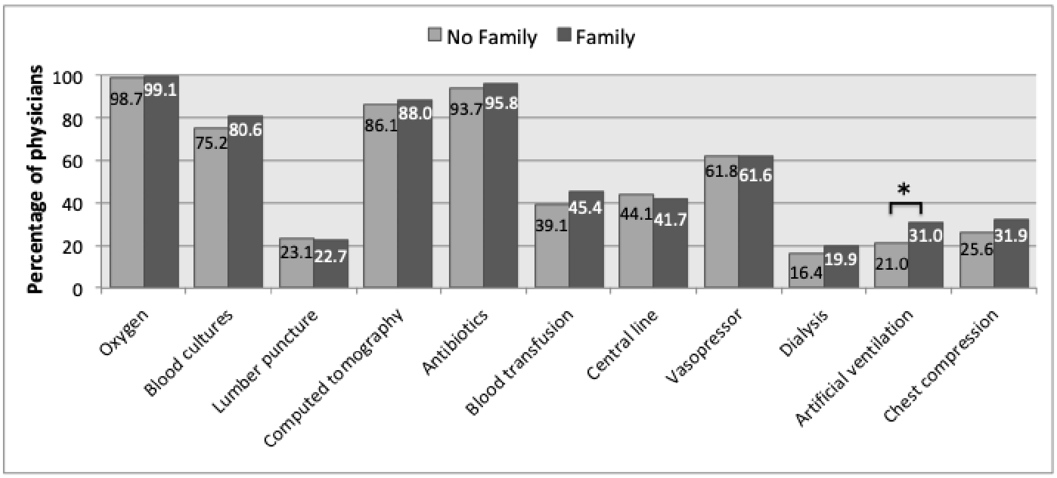
Variable	No family (n=238)	Family (n=216)	*P* value
Male, n (%)	194 (81.5)	176 (81.5)	>.99
Age (years), mean (SD)	45.7 (10.2)	45.1 (9.5)	.54
Years since graduation, mean (SD)	19.8 (9.9)	19.1 (9.1)	.42
Institution, n (%)			**.20**
Acute care hospital	132 (55.5)	132 (61.1)	
Clinic	51 (21.4)	48 (22.2)	
Chronic care hospital	32 (13.4)	26 (12.0)	
Other	23 (9.7)	10 (4.6)	
Number of times they have explained patient’s condition to his or her family per year, n (%)			**.75**
0	13 (5.5)	11 (5.1)	
1-6	32 (13.4)	37 (17.1)	
7-12	22 (9.2)	19 (8.8)	
13 or more	171 (71.8)	149 (69.0)	
Specialty, n (%)			**.11**
Internal medicine	95 (39.9)	88 (40.7)	
Surgery	27 (11.3)	14 (6.5)	
Orthopedics	24 (10.1)	11 (5.1)	
Psychiatry	18 (7.6)	14 (6.5)	
Others	74 (31.1)	89 (41.2)	

Figure 1. Differences in physicians' decision-making patterns between the presence and absence of family for critically ill patients. *P<.05, **P<.01, ***P<.001.

1) Vignette 1



2) Vignette 2



3) Vignette 3

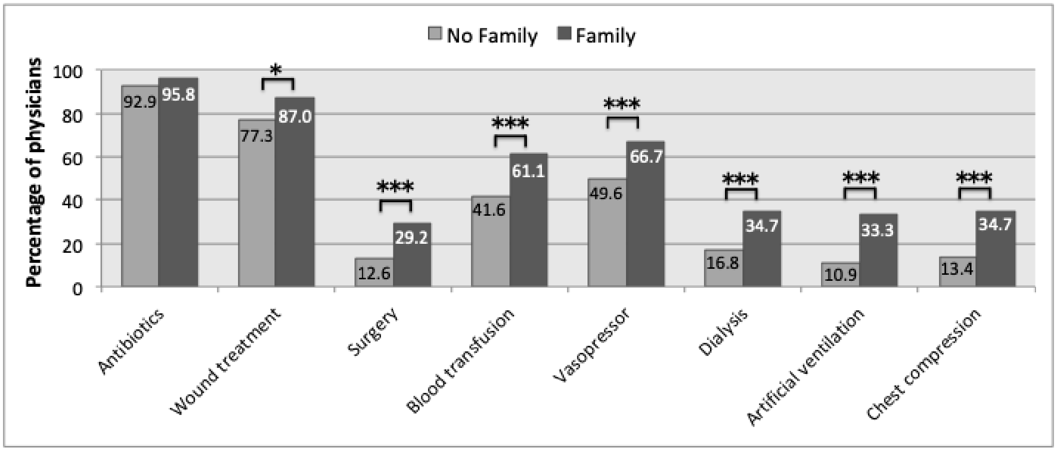


Table 2. Odds ratios of performing procedures in pattern A (without family present) versus pattern B (with family present) by sex (N=454).

Vignette and procedure	Male (n=370)		Female (n=84)	
	Odds ratio (95% CI)	*P* value	Odds ratio (95% CI)	*P* value
Vignette 1				
Dialysis	0.52 (0.34-0.78)	.002	0.69 (0.29-1.64)	.51
Artificial ventilation	0.48 (0.32-0.73)	<.001	0.67 (0.28-1.58)	.39
Vignette 2				
Artificial ventilation	0.57 (0.35-0.91)	.02	0.70 (0.28-1.76)	.49
Vignette 3				
Wound treatment	0.48 (0.28-0.84)	.01	0.64 (0.19-2.16)	.55
Surgery	0.36 (0.21-0.62)	<.001	0.30 (0.09-0.95)	.06
Blood transfusion	0.48 (0.31-0.72)	<.001	0.36 (0.14-0.88)	.03
Vasopressor	0.52 (0.34-0.79)	.002	0.37 (0.14-0.92)	.04
Dialysis	0.34 (0.2-0.55)	<.001	0.61 (0.23-1.61)	.34
Artificial ventilation	0.23 (0.14-0.41)	<.001	0.30 (0.09-0.95)	.06
Chest compression	0.29 (0.17-0.49)	<.001	0.29 (0.1-0.86)	.04

Table 3. Odds ratios of performing procedures in pattern A (without family present) versus pattern B (with family present) by experience (N=454).

Vignette and procedure	Experienced (n=289)		Nonexperienced (n=165)	
	Odds ratio (95% CI)	*P* value	Odds ratio (95% CI)	*P* value
Vignette 1				
Dialysis	0.43 (0.26-0.71)	<.001	0.80 (0.44-1.48)	.53
Artificial ventilation	0.46 (0.29-0.74)	.001	0.63 (0.34-1.16)	.16
Vignette 2				
Artificial ventilation	0.52 (0.30-0.90)	.02	0.73 (0.37-1.44)	.39
Vignette 3				
Wound treatment	0.40 (0.21-0.76)	.004	0.78 (0.34-1.78)	.67
Surgery	0.29 (0.16-0.53)	<.001	0.49 (0.22-1.08)	.08
Blood transfusion	0.39 (0.24-0.63)	<.001	0.59 (0.32-1.10)	.12
Vasopressor	0.43 (0.26-0.69)	<.001	0.62 (0.33-1.16)	.16
Dialysis	0.30 (0.17-0.53)	<.001	0.56 (0.28-1.13)	.12
Artificial ventilation	0.19 (0.10-0.35)	<.001	0.39 (0.18-0.87)	.02
Chest compression	0.24 (0.13-0.43)	<.001	0.42 (0.19-0.93)	.03

Table 4. Odds ratios of performing procedures in pattern A (without family present) versus pattern B (with family present) by type of institution (N=454).

Vignette and procedure	Acute care hospital (n=264)		Non-acute care hospital (n=190)	
	Odds ratio (95% CI)	*P* value	Odds ratio (95% CI)	*P* value
Vignette 1				
Dialysis	0.65 (0.40-1.06)	.11	0.43 (0.24-0.78)	.01
Artificial ventilation	0.59 (0.36-0.97)	.05	0.41 (0.23-0.74)	.003
Vignette 2				
Artificial ventilation	0.67 (0.38-1.17)	.21	0.49 (0.26-0.95)	.04
Vignette 3				
Wound treatment	0.85 (0.44-1.63)	.74	0.25 (0.11-0.59)	.001
Surgery	0.39 (0.21-0.71)	.003	0.31 (0.14-0.67)	.004
Blood transfusion	0.53 (0.32-0.88)	.02	0.37 (0.21-0.68)	.002
Vasopressor	0.65 (0.39-1.08)	.13	0.35 (0.19-0.63)	<.001
Dialysis	0.45 (0.26-0.79)	.01	0.30 (0.15-0.60)	<.001
Artificial ventilation	0.23 (0.12-0.46)	<.001	0.25 (0.12-0.52)	<.001
Chest compression	0.33 (0.18-0.61)	<.001	0.25 (0.12-0.51)	<.001

Table 5. Factors associated with withholding of artificial ventilation.

Vignette and factors	Univariate analysis		Multivariate analysis	
	Odds ratio (95% CI)	*P* value	Odds ratio (95% CI)	*P* value
Vignette 1				
Sex		.11		—[a]
Female	1		—	
Male	1.47 (0.91-2.37)		—	
Age		.08		—
<41	1		—	
≥41	1.44 (0.98-2.12)		—	
Years since graduation		.03		.03
<17	1		1	
≥17	1.55 (1.06-2.25)		1.54 (1.05-2.25)	
Institution		.13		—
Acute care hospital	1		—	
Clinic	0.76 (0.48-1.2)		—	
Chronic care hospital	1.66 (0.93-3.05)		—	
Others	0.82 (0.39-1.7)		—	
Number of explanations		.89		—
0	1		—	
1-6	1.09 (0.43-2.78)		—	
7-12	1.41 (0.51-3.93)		—	
≥13	1.12 (0.84-2.25)		—	
Department		.30		—
Internal medicine	1		—	
Surgery	0.82 (0.41-1.61)		—	
Orthopedics	2.07 (0.98-4.62)		—	
Psychiatry	1.38 (0.65-3.03)		—	
Others	1.03 (0.68-1.58)		—	
Pattern		<.001		<.001
A (without family)	1		1	
B (with family)	0.51 (0.35-0.74)		0.51 (0.35-.75)	
Vignette 2				
Sex		.29		—
Female	1		—	
Male	1.37 (0.81-2.29)		—	
Age		.37		—
<41	1		—	
≥41	1.25 (0.81-1.92)		—	
Years since graduation		.28		—
<17	1		—	
≥17	1.29 (0.85-1.97)		—	
Institution		.009		—
Acute care hospital	1		1	

Vignette and factors	Univariate analysis		Multivariate analysis	
	Odds ratio (95% CI)	*P* value	Odds ratio (95% CI)	*P* value
Clinic	0.67 (0.40-1.11)		0.66 (0.40-1.10)	.10
Chronic care hospital	2.89 (1.27-7.78)		2.84 (1.25-7.67)	.02
Others	0.64 (0.3-1.43)		0.57 (0.26-1.29)	.16
Number of explanations		.11		—
0	1		—	
1-6	0.67 (0.22-1.83)		—	
7-12	2.40 (0.64-9.38)		—	
≥13	0.95 (0.34-2.36)		—	
Department		.97		—
Internal medicine	1		—	
Surgery	0.89 (0.42-1.99)		—	
Orthopedics	1.10 (0.48-2.75)		—	
Psychiatry	0.83 (0.37-2.02)		—	
Others	0.88 (0.54-1.43)		—	
Pattern		.02		.01
A (without family)	1		1	
B (with family)	0.59 (0.39-0.90)		0.56 (0.36-0.87)	
Vignette 3				
Sex		.85		—
Female	1		—	
Male	0.91 (0.49-1.60)		—	
Age		.79		—
<41	1		—	
≥41	0.91 (0.57-1.45)		—	
Years since graduation		.68		—
<17	1		—	
≥17	0.88 (0.56-1.39)		—	
Institution		.42		—
Acute care hospital	1		—	
Clinic	0.74 (0.43-1.28)		—	
Chronic care hospital	1.43 (0.69-3.28)		—	
Others	0.79 (0.35-1.96)		—	
Number of explanations		.21		—
0	1		—	
1-6	1.84 (0.67-4.93)		—	
7-12	2.47 (0.8-7.86)		—	
≥13	2.4 (0.97-5.65)		—	
Department		.65		—
Internal medicine	1		—	
Surgery	2.01 (0.80-6.15)		—	
Orthopedics	0.94 (0.41-2.37)		—	
Psychiatry	1.00 (0.42-2.65)		—	

Vignette and factors	Univariate analysis		Multivariate analysis	
	Odds ratio (95% CI)	*P* value	Odds ratio (95% CI)	*P* value
Others	0.92 (0.56-1.53)		—	
Pattern		<.001		<.001
A (without family)	1		1	
B (with family)	0.25 (0.15-0.40)		0.25 (0.15-0.4)	

[a]Not applicable.

Discussion

Physicians withheld artificial ventilation for a patient without family through all three vignettes. Males, both experienced and those working at non-acute care hospitals, had a tendency to withhold medical procedures. Multivariate analysis showed the absence of family was the only significant factor associated with physicians' decision making regarding withholding artificial ventilation.

Withholding of artificial ventilation for a patient without family may be because discontinuation of mechanical ventilation is not legally authorized in Japan, and physicians are concerned about unresponsive wakefulness syndrome. Further, physicians may apply their preference to a lonely, incapacitated patient. In Japan, only 5% of people have an advance directive, and 73% of people die in hospital [9,10]. The attending physician at the end of life is often different from the usual family doctor and has little chance to know the preference of the incapacitated patient without relatives. Fewer physicians would want to receive cardiopulmonary resuscitation (CPR) at the end of life compared with the general public. A survey by the Japanese government in 2018 reported that the percentage of people in the general public who did not want CPR in case of terminal cancer, severe heart disease, and advanced dementia was 69.2%, 67.4%, and 71.9%, respectively; however, this was 89.3%, 87.6%, and 90.1%, respectively, among physicians [10].

There is also a possibility that physicians provide unnecessary medical intervention to patients with family to satisfy their sentiment and avoid litigation. In vignette 3, 30% of physicians answered that they would prioritize the will of the family over the will of the patient and perform invasive procedures such as surgery or CPR. Under traditional Japanese culture, patients' self-determination is sometimes in subordination to the will of the family; the Act on Organ Transplantation in Japan prohibits using organs of patients whose family members refused to offer, even when the patient themselves have a will to donate and have a donor card. As well, a previous study reported that among 934 patients with end-stage heart failure, only 4.7% of patients participated in end-of-life conversations with attending physicians and declared their preferences; surrogates made medical care decisions in 95.3% of cases [11].

Our study has some limitations. First, the patients were hypothetical; thus, this study does not demonstrate effects in daily clinical practice. Second, the response rates of online surveys tend to be lower than the response rates of face-to-face surveys [12-14]. Third, compared with the entire population of physicians in Japan, we assumed that fewer physicians in our study were working in a clinic, aged 20 to 29 years, or 60 years and older. However, this limitation would not considerably affect our main findings because approximately 74% of Japanese die in a hospital, and the last decisions regarding patient care are made in the hospital [15]. Further, physicians aged 20 to 29 years in training and 60 years or older in a supervising position do not usually work as an attending physician. Fourth, the majority of respondents were male, and this deviation can influence decision-making patterns [16]. Fifth, the view of death depends on the culture, and our results cannot be generalized to other countries directly. However, in the United States, it is estimated that there will be more than 2 million adults aged 70 years and older who have outlived all their friends and family members by 2030, so the implications of this survey do cross borders [4].

In conclusion, our findings suggest that elderly patients may sometimes be submitted to unwanted treatment due to the influence of family members or may have treatments withheld due to the absence of family. This highlights the potential importance of advance care planning in this era of an aging society with a declining birth rate.

Acknowledgments

The research team acknowledges the valued contributions of Motoaki Kojima and Michio Ando at Mediwel Co, Ltd. These individuals were not compensated for their contribution to this project.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Vignette A.
[PDF File (Adobe PDF File), 60KB - ijmr_v8i3e12781_app1.pdf]

Multimedia Appendix 2
Vignette B.
[PDF File (Adobe PDF File), 62KB - ijmr_v8i3e12781_app2.pdf]

References

1. Cabinet Office, Government of Japan. Economy and Living Environment of Elderly People by Type of Households [in Japanese]. URL: http://www8.cao.go.jp/kourei/ishiki/h28/sougou/zentai/pdf/sec_3_3.pdf [accessed 2019-08-07] [WebCite Cache ID 74HS5xj7R]
2. Ministry of Health, Labour and Welfare. Comprehensive Survey of Living Conditions for FY 2018 [in Japanese]. URL: https://www.mhlw.go.jp/toukei/saikin/hw/k-tyosa/k-tyosa16/dl/16.pdf [accessed 2018-11-29] [WebCite Cache ID 74HSEMR6g]
3. Laidsaar-Powell R, Butow P, Charles C, Gafni A, Entwistle V, Epstein R, et al. The TRIO Framework: conceptual insights into family caregiver involvement and influence throughout cancer treatment decision-making. Patient Educ Couns 2017 Nov;100(11):2035-2046. [doi: 10.1016/j.pec.2017.05.014] [Medline: 28552193]
4. Weiss BD, Berman EA, Howe CL, Fleming RB. Medical decision-making for older adults without family. J Am Geriatr Soc 2012 Nov;60(11):2144-2150. [doi: 10.1111/j.1532-5415.2012.04212.x] [Medline: 23110513]
5. Cook DJ, Guyatt GH, Jaeschke R, Reeve J, Spanier A, King D, et al. Determinants in Canadian health care workers of the decision to withdraw life support from the critically ill. Canadian Critical Care Trials Group. JAMA 1995 Mar 01;273(9):703-708. [Medline: 7853627]
6. White D, Curtis J, Lo B, Luce J. Decisions to limit life-sustaining treatment for critically ill patients who lack both decision-making capacity and surrogate decision-makers. Crit Care Med 2006 Aug;34(8):2053-2059. [doi: 10.1097/01.CCM.0000227654.38708.C1] [Medline: 16763515]
7. Fader A, Gambert S, Nash M, Gupta K, Escher J. Implementing a "Do-Not-Resuscitate" (DNR) policy in a nursing home. J Am Geriatr Soc 1989 Jun;37(6):544-548. [doi: 10.1111/j.1532-5415.1989.tb05687.x] [Medline: 2715562]
8. White D, Curtis J, Wolf L, Prendergast T, Taichman D, Kuniyoshi G, et al. Life support for patients without a surrogate decision maker: who decides? Ann Intern Med 2007 Jul 03;147(1):34-40. [doi: 10.7326/0003-4819-147-1-200707030-00006] [Medline: 17606959]
9. Ministry of Health, Labour and Welfare. e-Stat. Vital statistics [in Japanese] URL: https://tinyurl.com/y328jrr3 [accessed 2019-08-06] [WebCite Cache ID 74HSHnYyl]
10. Ministry of Health, Labour and Welfare. Survey of Attitudes Towards End-of-Life Care for FY 2017 [in Japanese]. URL: https://www.mhlw.go.jp/toukei/list/dl/saisyuiryo_a_h29.pdf [accessed 2018-11-07] [WebCite Cache ID 73kGLFCpI]
11. Nakamura K, Kinugasa Y, Sugihara S, Hirai M, Yanagihara K, Haruki N, et al. Sex differences in surrogate decision-maker preferences for life-sustaining treatments of Japanese patients with heart failure. ESC Heart Fail 2018 Dec;5(6):1165-1172 [FREE Full text] [doi: 10.1002/ehf2.12352] [Medline: 30264449]
12. So R, Shinohara K, Aoki T, Tsujimoto Y, Suganuma AM, Furukawa TA. Effect of recruitment methods on response rate in a web-based study for primary care physicians: factorial randomized controlled trial. J Med Internet Res 2018 Dec 08;20(2):e28 [FREE Full text] [doi: 10.2196/jmir.8561] [Medline: 29422450]
13. Shinjo T, Morita T, Kiuchi D, Ikenaga M, Abo H, Maeda S, et al. Japanese physicians' experiences of terminally ill patients voluntarily stopping eating and drinking: a national survey. BMJ Support Palliat Care 2019 Jun;9(2):143-145. [doi: 10.1136/bmjspcare-2017-001426] [Medline: 29118101]
14. Asch D, Jedrziewski M, Christakis N. Response rates to mail surveys published in medical journals. J Clin Epidemiol 1997 Oct;50(10):1129-1136. [Medline: 9368521]
15. Ministry of Health, Labour and Welfare. Survey of Physicians, Dentists and Pharmacists for FY 2016 [in Japanese]. URL: http://www.mhlw.go.jp/toukei/saikin/hw/ishi/16/dl/kekka_1.pdf [accessed 2019-08-06] [WebCite Cache ID 77cK9xJys]
16. Tsugawa Y, Jena AB, Figueroa JF, Orav EJ, Blumenthal DM, Jha AK. Comparison of hospital mortality and readmission rates for medicare patients treated by male vs female physicians. JAMA Intern Med 2017 Dec 01;177(2):206-213 [FREE Full text] [doi: 10.1001/jamainternmed.2016.7875] [Medline: 27992617]

Abbreviations

OR: odds ratio
CPR: cardiopulmonary resuscitation

Edited by G Eysenbach; submitted 28.11.18; peer-reviewed by E de Quincey, A Rodríguez-Molinero; comments to author 31.03.19; revised version received 13.05.19; accepted 21.07.19; published 06.09.19.

Please cite as:
Tsuda K, Higuchi A, Yokoyama E, Kosugi K, Komatsu T, Kami M, Tanimoto T
Physician Decision-Making Patterns and Family Presence: Cross-Sectional Online Survey Study in Japan
Interact J Med Res 2019;8(3):e12781
URL: http://www.i-jmr.org/2019/3/e12781/
doi:10.2196/12781
PMID:31493327

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org