Interactive Journal of Medical Research

Impact Factor (2022): 2.0
Volume 7 (2018), Issue 1 ISSN 1929-073X

Contents

Corrigenda and Addenda

Original Papers

Short Paper

Corrigenda and Addenda

Metadata Correction: Engaging a Community for Rare Genetic Disease: Best Practices and Education From Individual Crowdfunding Campaigns

Romina Alicia Ortiz[1], MHS; Steven Witte[1], MD, PhD; Arvin Gouw[1], PhD; Ana Sanfilippo[1], MBA; Richard Tsai[1], MSc; Danielle Fumagalli[1], MA; Christine Yu[1], MPH; Karla Lant[1], BA; Nicole Lipitz[1], BA; Jennifer Shepphird[1], PhD; Fidelia B Alvina[1], BA; Jimmy Cheng-Ho Lin[1], MPH, MD, PhD

Rare Genomics Institute, Downey, CA, United States

Corresponding Author:
Romina Alicia Ortiz, MHS
Rare Genomics Institute
8504 Firestone Blvd #198
Downey, CA, 90241
United States
Phone: 1 5624132581
Email: romina.ortiz@raregenomics.org

Related Article:

Correction of: http://www.i-jmr.org/2018/1/e3/

(Interact J Med Res 2018;7(1):e10707) doi:10.2196/10707

In the paper by Ortiz et al, “Engaging a Community for Rare Genetic Disease: Best Practices and Education From Individual Crowdfunding Campaigns” (Interact J Med Res 2018;7(1):e3), author Nicole Lipitz’s last name was incorrectly spelled as “Lipintz”. This has now been corrected.

The corrected article will appear in the online version of the paper on the JMIR website on April 27, 2018, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article also has been re-submitted to those repositories.

Edited by G Eysenbach; submitted 05.04.18; this is a non–peer-reviewed article;accepted 05.04.18; published 27.04.18.

Please cite as:
Ortiz RA, Witte S, Gouw A, Sanfilippo A, Tsai R, Fumagalli D, Yu C, Lant K, Lipitz N, Shepphird J, Alvina FB, Cheng-Ho Lin J
Metadata Correction: Engaging a Community for Rare Genetic Disease: Best Practices and Education From Individual Crowdfunding Campaigns
Interact J Med Res 2018;7(1):e10707
URL: http://www.i-jmr.org/2018/1/e10707/
doi:10.2196/10707
PMID:30578239

Original Paper

Characteristics of Adults’ Use of Facebook and the Potential Impact on Health Behavior: Secondary Data Analysis

Kelly Bosak[1], PhD, ANP-BC; Shin Hye Park[1], PhD, RN

[1] School of Nursing, University of Kansas Medical Center, Kansas City, KS, United States

Corresponding Author:
Kelly Bosak, PhD, ANP-BC
School of Nursing
University of Kansas Medical Center
3901 Rainbow Boulevard
Mailstop 4043
Kansas City, KS, 66160
United States
Phone: 1 913 588 1656
Fax: 1 913 588 1660
Email: kbosak@kumc.edu

Abstract

Background: Social influences are a primary factor in the adoption of health behaviors. Social media platforms, such as Facebook, disseminate information, raise awareness, and provide motivation and support for positive health behaviors. Facebook has evolved rapidly and is now a part of many individuals' daily routine. The high degree of individual engagement and low attrition rate of this platform necessitate consideration for a potentially positive impact on health behavior.

Objective: The aim of this paper is to investigate the use of Facebook by adults. Knowledge is limited to the unique characteristics of Facebook users, including time spent on Facebook by adults of various age groups. Characteristics of Facebook users are important to understand to direct efforts to engage adults in future health behavior interventions.

Methods: Institutional Review Board approval was obtained for this secondary analysis of existing de-identified survey data collected for the Pew Research Center. The sample included adults age 18-65 years and above. Binomial logistic regression was performed for the model of age group and Facebook use, controlling for other demographics. A multinomial logistic regression model was used for the variable of time spent on Facebook. Based on the regression models, we computed and reported the marginal effects on Facebook use and time spent of adults age groups, including age groups 18-29, 30-49, 50-64, and 65 and over. We discuss these findings in the context of the implications for promoting positive health behaviors.

Results: The demographics of the final sample (N=730) included adults age 18-65 years and above (mean 48.2 yrs, SD 18.3 yrs). The majority of the participants were female (372/730, 50.9%), white (591/730, 80.9%) and non-Hispanic (671/730, 91.9%). Bivariate analysis indicated that Facebook users and nonusers differed significantly by age group (χ^2=76.71, P<.001) and sex (χ^2=9.04, P=.003). Among subjects aged 50 and above, the predicted probability was 66% for spending the same amount of time, 10% with increased time, and 24% with decreased time.

Conclusions: The key findings of this study were Facebook use among midlife and older adults was more likely to stay the same over time, compared to the other age groups. Interestingly, the young adult age group 18-29 years was more likely to decrease their time on Facebook over time. Specifically, younger females were most likely to decrease time spent on Facebook. In general, male participants were most likely to spend the same amount of time on Facebook. These findings have implications for future health intervention research, and ultimately, for translation to the clinic setting to improve health outcomes.

(Interact J Med Res 2018;7(1):e11) doi:10.2196/ijmr.9554

KEYWORDS

social media; health promotion; health behavior; adults

Introduction

Background

Social media has evolved rapidly and has been used as a means of communication, social interaction, and social support with minimal consideration for wider use in health care. It is recognized that social influences are an important factor in the adoption of health behaviors [1,2]. Despite the growing interest in practical research and the development and implementation of health behavior interventions in real-world settings, social media is often overlooked as an effective method to deliver the active components of a health behavior intervention. To accomplish this requires knowledge of the characteristics of social media use among the population of interest for integration in a health behavior intervention.

Significance

The use of the social media is deeply embedded in everyday life for an increasing number of adults [3]. The widely recognized, accessed, and the culturally integrated Facebook platform has an extensive reach, with over one billion daily adult users and continues to grow [4].

The Facebook platform has been used effectively to recruit participants for health behavior research [5], to provide online health information and education [6,7] as well as the more familiar social networking functions. This platform has been used as an adjunct to multimodal interventions or in conjunction with investigator-initiated in-person visits and other methods of communication, such as mobile phones, text messages, and apps [8-11]. Facebook used as a means for patients to communicate with other patients is a trend referred to as peer-to-peer health care [12]. These functions of Facebook are widely recognized; however, peer support is only a small part of the potential of this platform to improve health outcomes.

Facebook offers the capability to create groups via the platform referred to as “Facebook Groups.” This group platform offers a convenient, reliable, private, members-only site that can be adapted to motivate and support health behavior change interventions and conduct research. This platform also allows integration into clinic processes, with health care professionals serving as moderators and contributors of the content to ensure its’ veracity. Efficacious health behavior interventions translated into real-life settings for effectiveness testing require the ability to appeal to and influence a broad and diverse population of participants. This necessitates an understanding of the characteristics of Facebook use of potential participants from the population of interest.

A distinct advantage of using the Facebook platform to facilitate a health behavior intervention is the high level of individual engagement. This is an important factor in promoting positive health behavior change. Engagement in an intervention is critical for participants to receive an adequate dose of the intervention [13], and thus, improve health outcomes. Attrition is very low on this platform, although some individual users occasionally take breaks from social media use [4]. Minimizing attrition may support long-term maintenance of health behaviors. The high levels of engagement and retention on Facebook over time may translate to the Facebook Groups platform making this an appealing method for delivering the active components of a health behavior intervention. Interventions facilitated by social media platforms hold promise to reduce the challenges for the growing population of midlife adults to prevent or delay the progression of chronic health conditions [14].

Few studies reported health behavior interventions facilitated by social media platforms to provide the active components of an intervention. A review found positive outcomes with 87.5% (7/8) of the studies involving physical activity interventions delivered using Facebook [15]. Other research found that Facebook user engagement and social support predicted a reduction in smoking [16]. Features of the Facebook Groups platform to promote cardiac rehabilitation included participant only access with a health care provider involved as group moderator to provide evidence-based information and assure “trustworthiness” of the group [17]. Greater Facebook capability may increase willingness to participate in the Facebook Groups platform. Trends show that adults across the age continuum are capable users of this technology [18]. Further, participants may be more likely to engage on the Facebook Groups platform upon the recommendation of their health care provider. Earlier studies reported that some older adults had concerns about the usefulness of combining health resources with social media [19]. Social media use has advanced since its’ inception and is now being used by younger adults as well as older adults to influence health. Therefore, we investigated the characteristics of adult Facebook users to inform our future research on the Facebook Groups platform.

Study Objective

The objective of this study was to investigate the use of the social media platform, Facebook by adults (aged 18-65 years and above). Knowledge is limited to the characteristics of Facebook use by adults across the age continuum, necessitating this investigation. Characteristics of Facebook users are important to understand to guide the development of a future clinical trial using the social media platform, Facebook Groups to deliver the active components of a health behavior intervention.

Methods

Design

The Institutional Review Board at the University where this study was conducted approved this study as exempt from full board review due to the use of de-identified data. The design was a secondary analysis of survey data. The sample included adults across the age continuum (18-65 years and above). Participants were included in our analysis based on the following: (a) they responded “yes” to the initial question: “Do you use the internet?” (b) they were able to provide informed verbal consent by phone, and (c) they were able to speak and understand English or Spanish, as the survey was not conducted in any other languages.

Data Collection

The Pew Research Center surveyed a representative sample of adults across all regions of the continental United States about their use of the social media platform, Facebook [20]. This national self-report survey was conducted between November 14, 2012 and December 9, 2012, on landlines and cell phones, and in English and Spanish by Princeton Survey Research Associates International. Multiple attempts were made to reach each phone number at a variety of times throughout the day. The survey consisted of 20 items rated on a Likert-type scale or a true and false format. We selected the sociodemographic and Facebook use the variables: user or nonuser, and time spent using Facebook stayed about the same, increased or decreased for this analysis.

Data Analysis

We described demographics by adults' use of Facebook and time spent on Facebook, using chi-squared tests for categorical variables. Fisher's exact test was used for two variables due to small cell counts as noted. Multivariate binomial logistic regression was performed for the model of Facebook use (users and nonusers) on age group, controlling for other demographics, including sex, income, employment, race and ethnicity, and education. A multivariate, multinomial logistic regression model was used for our secondary outcome variable, regressing time spent on Facebook (stayed about the same, increased, and decreased) on age groups and other demographic variables. We also questioned whether the effects of age group on Facebook use and time spent would differ by sex. Therefore, after performing multivariate regression, we calculated marginal effects to show the interactions between age group and sex in the models. The marginal effects were interpreted as the predicted probability of using Facebook or time spent on Facebook for participants of a specific age group and sex, holding all other demographics constant to the means. We discuss these findings in the context of the implications for health behavior change. All data were analyzed using the STATA Version 14.0 (StataCorp, LP, College Station, TX, USA).

Results

The demographics of the sample (N=730) are shown in Table 1 and included adults age 18-65 years and above with a mean of 50.3 (SD 19.7) years. The majority of the participants interviewed were white (591/730, 80.9%) and non-Hispanic or another ethnicity 91.9% (671/730). There were 50.9% (372/730) females and 49.0% (358/730) males responding to the survey. Some college or college graduation was reported by 69.0% (504/730) of the respondents. Half (365/730) of the sample reported full-time employment, with just over half (380/730, 52.0%) reporting an income between US $30,000 and US $100,000. The majority (462/730, 63.2%) surveyed were Facebook users, and 52.0% (380/730) reported their time spent on Facebook stayed the same over the past year.

The bivariate analysis (Table 2) indicated that Facebook users and nonusers differed significantly by age group (N=730, χ^2_3=76.7, *P*<.001), and sex (N=730, χ^2_1=9.0, *P*=.003). There were no significant differences in Facebook users and nonusers based on income, employment status, race and ethnicity, or education level.

The bivariate analysis (see Table 3) indicated that time spent on Facebook (ie, stayed about the same, increased, and decreased) differed significantly by age group (n=461, χ^2_6=31.8, *P*<.001). No differences were found in time spent on Facebook based on income, employment status, race, or education. Notably, ethnicity was significant (n=461, χ^2_2=5.3, *P*=.045) using Fisher's exact test.

Based on our bivariate results, we examined whether the effects of age group differed by sex. Table 4 shows significant variations in the marginal effects of age group by sex on Facebook use. Younger and female subjects were more likely to use Facebook than older and male subjects. The probability of using Facebook was 87% among participants aged 18-29 years and differed by sex. The probability of using Facebook among participants aged 18-29 was 91% for female, whereas it was 83% for male. The probability of using Facebook among participants aged 65 and over was 45% for female, whereas it was 27% for male. Figure 1 shows the marginal effects of age group by sex on Facebook decreasing across age groups.

Table 5 shows marginal effects of age group and sex across the three subgroups of Facebook time (stayed the same, increased, and decreased). Due to small cell counts in age group 65+ and Asian group (Table 3) we regrouped the age variable into three groups (18-29, 30-49, and 50+) and the race variable into three groups (white, black or African American, and others) for analysis of time spent using Facebook.

Overall, subjects in the group that increased amount of time on Facebook showed the lowest probabilities than those in the other groups. Among subjects aged 50 and above, the predicted probability was 66% for spending the same amount of time, 10% with increased time, and 24% with decreased time. Older subjects were more likely to spend the same amount of time and less likely to decrease time on Facebook. Younger participants were more likely to decrease time on Facebook.

Among older males, the predicted probabilities for time spent using Facebook were 70% stayed the same, 8% increased, and 22% decreased time on Facebook. The predicted probabilities for Facebook use among women in this age group were similar with 64% stayed the same, 11% increased, and 26% decreased time spent on Facebook. Figures 2-4 show the marginal effects of age group by sex on Facebook use for the variables: stayed about the same (Figure 2), increased (Figure 3), and decreased (Figure 4).

Table 1. Descriptive summary of demographic characteristics and outcome variables.

Variables	Value (N=730), n (%)
Age group (years)	
18-29	152 (20.8)
30-49	232 (31.8)
50-64	191 (26.2)
65+	155 (21.2)
Sex	
Male	358 (49.0)
Female	372 (51.0)
Income (US$)	
<$30,000	219 (30.0)
$30,000-$100,000	380 (52.1)
>$100,000	131 (17.9)
Employment	
Full-time	362 (49.6)
Part-time	104 (14.2)
Not employed	264 (36.2)
Race	
Caucasian	589 (80.7)
African-American	77 (10.5)
Asian or Pacific Islander	19 (2.6)
Other	45 (6.2)
Ethnicity	
Hispanic	61 (8.4)
Not Hispanic or other	669 (91.6)
Education	
≤High school	223 (30.5)
≥College	507 (69.5)
Facebook users	
Users	462 (63.3)
Nonusers	268 (36.7)
Time spent on Facebook[a]	
Stayed the same	382 (52.3)
Increased	96 (13.2)
Decreased	252 (34.5)

[a] n=461. This subgroup was from Facebook users only and those responding to the item of time spent on Facebook.

Table 2. Bivariate statistics by Facebook users and nonusers.

Variable	Full sample (N=730), n (%)	Facebook users (N=462), n (%)	Facebook nonusers (n=268), n (%)	Chi-square (df)	*P* value
Age group (years)				76.7 (3)	<.001
18-29	152 (20.8)	129 (27.9)	23 (8.6)		
30-49	232 (31.8)	163 (35.3)	69 (25.7)		
50-64	191 (26.2)	109 (23.6)	82 (30.6)		
65+	155 (21.2)	61 (13.2)	94 (35.1)		
Sex				9.0 (1)	.003
Male	358 (49.0)	207 (44.8)	151 (56.3)		
Female	372 (51.0)	255 (55.2)	117 (43.7)		
Income (US $)				0.6 (2)	.748
< $30,000	219 (30.0)	140 (30.3)	79 (29.5)		
$30,000-$100,000	380 (52.1)	236 (51.1)	144 (53.7)		
>$100,000	131 (17.9)	86 (18.7)	45 (16.9)		
Employment				5.6 (2)	.062
Full-time	362 (49.6)	237 (51.3)	125 (46.6)		
Part-time	104 (14.2)	73 (15.7)	32 (12.0)		
Not employed	264 (36.2)	153 (33.1)	111 (41.4)		
Race				6.9 (3)	.076
White	589 (80.7)	361 (78.1)	228 (85.0)		
Black or African-American	77 (10.5)	57 (12.3)	20 (7.5)		
Asian or Pacific Islander	19 (2.6)	15 (3.3)	4 (1.5)		
Other	45 (6.2)	30 (6.4)	16 (6.0)		
Ethnicity				3.2 (1)	.076
Hispanic	61 (8.4)	45 (9.7)	16 (6.0)		
Not Hispanic (other)	669 (91.6)	417 (90.3)	252 (94.0)		
Education				1.0 (1)	.307
≤High school	223 (30.5)	135 (29.2)	88 (32.8)		
≥College	507 (69.5)	327 (70.8)	180 (67.2)		

Table 3. Bivariate statistics by time spent on Facebook (n=461).

Variable	Stayed about the same (n=241), n (%)	Increased (n=61), n (%)	Decreased (n=159), n (%)	Chi-square (df)	*P* value
Age group (years)				31.8 (6)	<.001[a]
18-29	55 (22.8)	12 (19.7)	61 (38.4)		
30-49	75 (31.1)	29 (47.5)	59 (37.1)		
50-64	65 (27.0)	17 (27.9)	27 (17.0)		
65+	46 (19.1)	3 (4.9)	12 (7.5)		
Sex				1.5 (2)	.466
Male	112 (46.5)	23 (37.8)	72 (45.3)		
Female	129 (53.5)	38 (62.3)	87 (54.7)		
Income				7.4 (4)	.118
<$30,000	67 (27.8)	16 (26.2)	56 (35.2)		
$30,000-$100,000	134 (55.6)	28 (45.9)	74 (46.5)		
>$100,000	40 (16.7)	17 (27.9)	29 (18.3)		
Employment				5.3 (4)	.255
Full-time	123 (51.0)	26 (42.6)	88 (55.4)		
Part-time	32 (13.3)	13 (21.3)	26 (16.4)		
Not employed	86 (35.8)	22 (36.1)	45 (28.3)		
Race				6.4 (6)	.363[a]
White	196 (81.3)	46 (75.4)	118 (74.2)		
Black or African-American	24 (10.0)	7 (11.4)	26 (16.4)		
Asian or Pacific Islander	6 (2.5)	4 (6.7)	5 (3.1)		
Other	15 (6.2)	4 (6.7)	10 (6.4)		
Ethnicity				5.3 (2)	.045[a]
Hispanic	27 (11.2)	1 (1.6)	17 (10.7)		
Not Hispanic (Other)	214 (88.9)	60 (98.4)	142 (89.3)		
Education				3.2 (2)	.204
≤High school	75 (31.1)	12 (19.7)	48 (30.2)		
≥College	166 (68.9)	49 (80.3)	111 (69.8)		

[a] *P* values from Fisher’s exact test due to small cell counts.

Table 4. Marginal effects of age group and sex on the probability of using Facebook.

Variable	Facebook user[a], respondent probability (95% CI)	*P* value
Age group (years)		
18-29	0.87 (0.82-0.92)	<.001
30-49	0.71 (0.65-0.77)	<.001
50-64	0.58 (0.51-0.65)	<.001
65+	0.36 (0.27-0.44)	<.001
Sex		
Male	0.56 (0.51-0.62)	<.001
Female	0.73 (0.69-0.78)	<.001
Age group by sex (years)		
Male		
18-29	0.83 (0.76-0.89)	<.001
30-49	0.62 (0.54-0.70)	<.001
50-64	0.48 (0.40-0.57)	<.001
65+	0.27 (0.19-0.36)	<.001
Female		
18-29	0.91 (0.87-0.95)	<.001
30-49	0.78 (0.72-0.84)	<.001
50-64	0.67 (0.59-0.74)	<.001
65+	0.45 (0.35-0.54)	<.001

[a] Marginal effects from binary logistic regression, adjusted for income, employment status, ethnicity, race, and education.

Figure 1. Marginal effects of age group and sex on Facebook use.

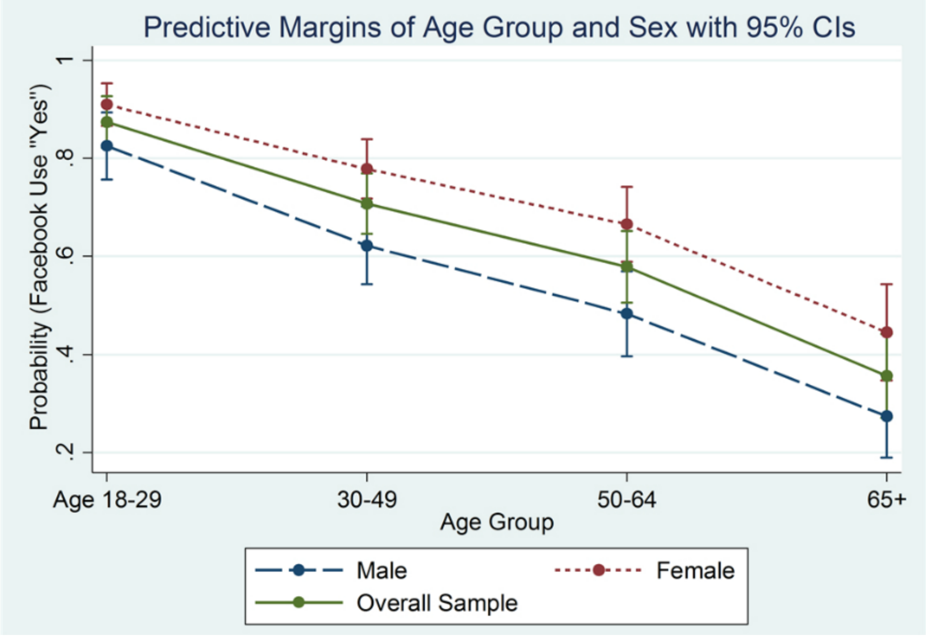


Table 5. Marginal effects of age group and sex on the probability of time on Facebook.

Variable	Stayed about the same[a], respondent probability (95% CI)	*P* value	Increased[a], respondent probability (95% CI)	*P* value	Decreased[a], respondent probability (95% CI)	*P* value
Age group (years)						
18-29	0.43 (0.33-0.52)	<.001	0.08 (0.03-0.14)	.001	0.49 (0.40-0.58)	<.001
30-49	0.49 (0.41-0.57)	<.001	0.15 (0.09-0.21)	<.001	0.36 (0.28-0.44)	<.001
50+	0.66 (0.59-0.74)	<.001	0.10 (0.05-0.14)	<.001	0.24 (0.17-0.31)	<.001
Sex						
Male	0.58 (0.50-0.65)	<.001	0.10 (0.05-0.14)	<.001	0.32 (0.26-0.39)	<.001
Female	0.51 (0.44-0.58)	<.001	0.12 (0.08-0.17)	<.001	0.37 (0.30-0.43)	<.001
Age group by sex (years)						
Male						
18-29	0.46 (0.36-0.56)	<.001	0.08 (0.03-0.13)	.003	0.46 (0.36-0.56)	<.001
30-49	0.53 (0.43-0.63)	<.001	0.13 (0.06-0.20)	<.001	0.34 (0.25-0.43)	<.001
Age 50+	0.70 (0.61-0.79)	<.001	0.08 (0.03-0.13)	.001	0.22 (0.14-0.30)	<.001
Female						
18-29	0.40 (0.29-0.51)	<.001	0.09 (0.03-0.15)	.003	0.51 (0.40-0.62)	<.001
30-49	0.46 (0.37-0.55)	<.001	0.16 (0.09-0.24)	<.001	0.38 (0.29-0.47)	<.001
50+	0.64 (0.55-0.72)	<.001	0.11 (0.05-0.16)	.001	0.26 (0.18-0.33)	<.001

[a] Marginal effects from multinomial logistic regression, adjusted for income, employment status, ethnicity, race, and education; 95% confidence intervals in parentheses.

Figure 2. Marginal effects of age group and sex by Facebook time "Stayed the Same".

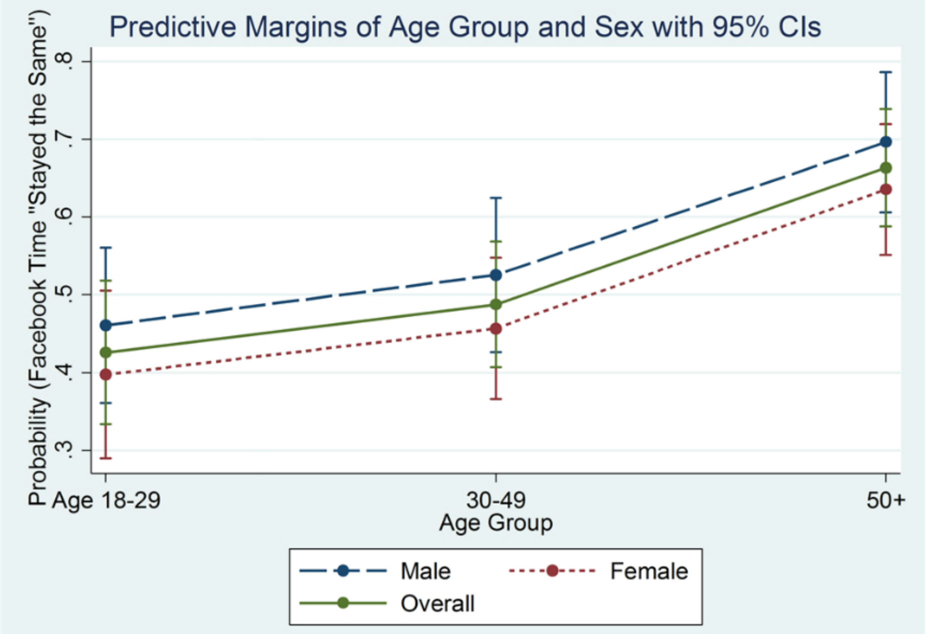


Figure 3. Marginal effects of age group and sex by Facebook time "Increased".

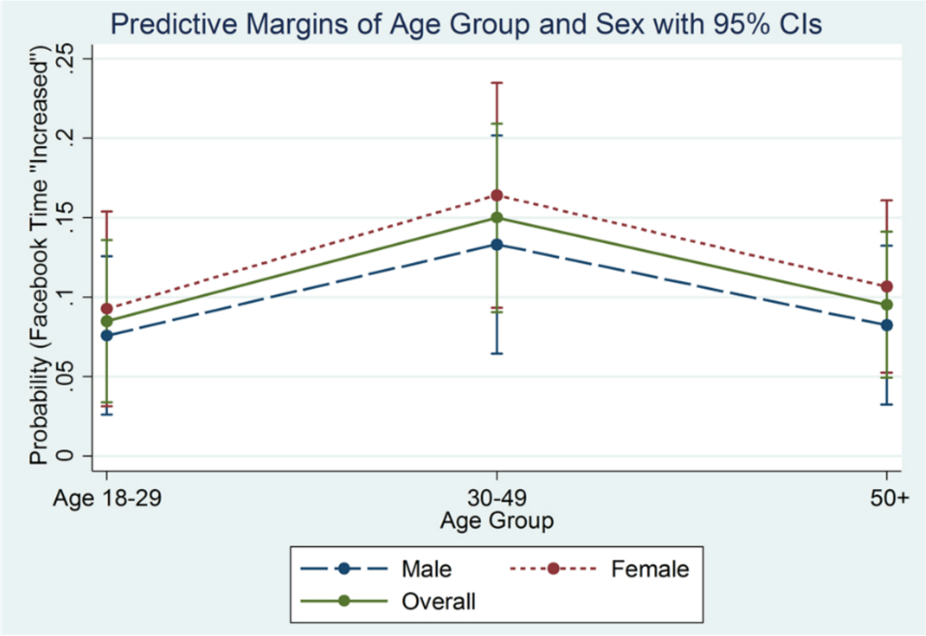
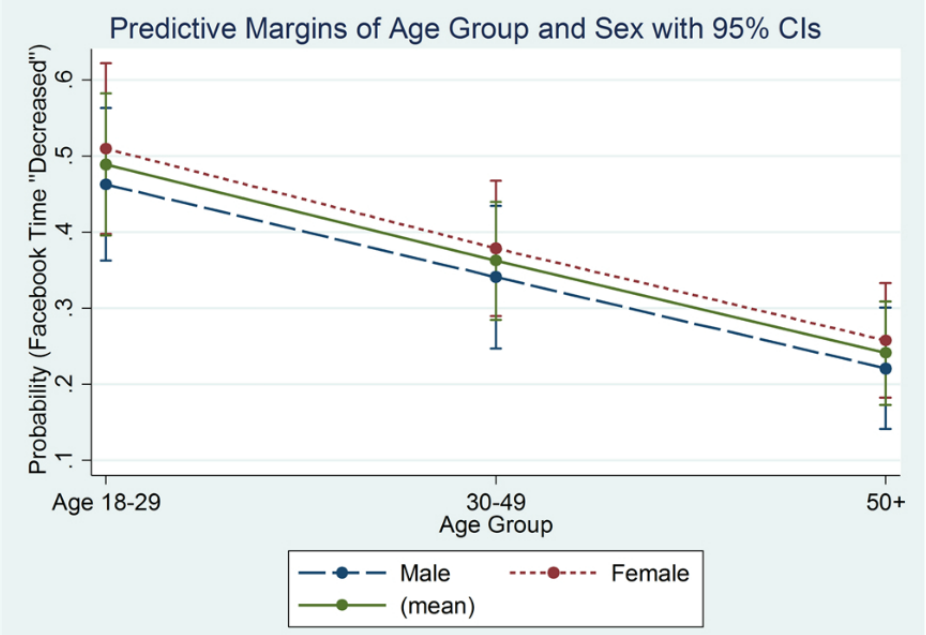


Figure 4. Marginal effects of age group and sex by Facebook time "Decreased".



Discussion

Principal Findings

The key findings of this study were Facebook use among midlife adults was more likely to stay the same over time, compared to the other age groups. Interestingly, in the young adult age group, 18-29 years was more likely to decrease their time on Facebook over time. Based on the available data, we are not able to explain this finding but it may be due to the use of other competing social media platforms (eg, Twitter, Instagram, Snapchat, etc) and lifestyle factors unique to younger adults. Younger females were most likely to decrease time spent on Facebook. In general, male participants were most likely to spend the same amount of time on Facebook. The probability of increasing time on Facebook was low for both males and females.

Earlier reports indicated that midlife and older adults were unlikely to use social media to seek or discuss health information as it was perceived as an intrusion of privacy [21]. This age group was reportedly reluctant to learn about using the latest technology [3]. Our study findings contradict these perceptions and show that older adults had the highest probability of maintaining Facebook use over time. Our results demonstrated that older adults' use of Facebook stayed about the same over time, compared to decreased time in the younger age adult groups.

The reluctance of adults of any age to participate in social media may be related to a lack of awareness or familiarity with the technology. Facebook is the most widely used social media platform, and the user base has grown more representative of the broader population [18]. While young adults were among the earliest social media adopters and had high rates of social media use, use among midlife and older adults has increased considerably in recent years [18]. An update of the Pew Research Center survey data reported in early 2018 indicated that ownership of mobile phones used to access social media platforms has continued to grow to nearly three-quarters of adults in the United States [22]. In this update, adults age 18 to 49 years were found to be close to saturation adoption. This is nearly double the percentage measured in the Pew survey in 2012. Notably, the fastest growing demographic regarding Facebook use was with the midlife and older adult population.

The Facebook Groups platform with high levels of engagement, low levels of attrition over time, and wide reach holds promise to motivate and support health behavior change for the growing population of midlife adults. Providing technical assistance for midlife and older adults with mobile device use and navigating the Facebook Groups platform must be considered to improve use for this age group when developing health behavior interventions.

Considerations for generalizing these findings to other populations include the representative sample of adults within the continental United States and the global reach of the Facebook platform. The wide geographical sampling distribution strengthens the Pew Research Center data for generalization to US adults, but may not be representative of the global population. We also acknowledge that secondary analyses of existing data are limited to the variables in the dataset, and this can be considered a limitation. We acknowledge that the variables used in this analysis may not be sufficient to account for other unmeasured factors related to our outcome variables. However, the sociodemographic variables examined, Facebook use and nonuse variables, and Facebook use over time support our study objective.

Conclusion

Overall, the findings of this study point to considerations for researchers to use the Facebook Groups platform to facilitate health behavior interventions. In addition to social communication and interaction, this platform is capable of providing health information, motivation and support for positive health behavior change. This platform can be integrated with clinic processes, involving health care professionals as moderators and contributors to ensure the veracity of the content. Providing a health behavior intervention on a social media platform is a unique approach to improve health outcomes, specifically, with the large and growing midlife adult population with escalating risk for chronic conditions. Our data indicate more consistent Facebook use among midlife adults than younger adults. Prospective research is needed to determine if the levels of engagement and retention of Facebook use for general social networking purposes translates to health behavior interventions facilitated by the Facebook Groups platform.

Conflicts of Interest

None declared.

References

1. Pampel FC, Krueger PM, Denney JT. Socioeconomic Disparities in Health Behaviors. Annu Rev Sociol 2010 Aug;36:349-370 [FREE Full text] [doi: 10.1146/annurev.soc.012809.102529] [Medline: 21909182]
2. Smith KP, Christakis NA. Annual Review of Sociology. Palo Alto, CA 94306 USA: Annual Reviews; 2008 Mar 24. Social Networks and Health URL: https://www.annualreviews.org/doi/10.1146/annurev.soc.34.040507.134601 [accessed 2018-05-09] [WebCite Cache ID 6zI8sBqD9]
3. Aaron S. Pew Research Center Internet & Technology. Washington, DC 20036 USA: Pew Research Center; 2014 Apr 03. Older adults and technology use URL: http://www.pewinternet.org/2014/04/03/older-adults-and-technology-use/ [accessed 2018-05-09] [WebCite Cache ID 6zI4nec2f]
4. Facebook. Facebook Newsroom. Menlo Park, California 94025 USA: Facebook; 2017. Company Info URL: https://newsroom.fb.com/company-info/ [accessed 2018-05-09] [WebCite Cache ID 6zI5xpi2e]
5. Whitaker C, Stevelink S, Fear N. The Use of Facebook in Recruiting Participants for Health Research Purposes: A Systematic Review. J Med Internet Res 2017 Aug 28;19(8):e290 [FREE Full text] [doi: 10.2196/jmir.7071] [Medline: 28851679]
6. Dagan N, Beskin D, Brezis M, Reis BY. Effects of Social Network Exposure on Nutritional Learning: Development of an Online Educational Platform. JMIR Serious Games 2015;3(2):e7 [FREE Full text] [doi: 10.2196/games.4002] [Medline: 26441466]
7. Nicolai L, Schmidbauer M, Gradel M, Ferch S, Antón S, Hoppe B, et al. Facebook Groups as a Powerful and Dynamic Tool in Medical Education: Mixed-Method Study. J Med Internet Res 2017 Dec 22;19(12):e408 [FREE Full text] [doi: 10.2196/jmir.7990] [Medline: 29273572]
8. Cavallo DN, Tate DF, Ries AV, Brown JD, DeVellis RF, Ammerman AS. A social media-based physical activity intervention: a randomized controlled trial. Am J Prev Med 2012 Nov;43(5):527-532 [FREE Full text] [doi: 10.1016/j.amepre.2012.07.019] [Medline: 23079176]
9. Cavallo DN, Tate DF, Ward DS, DeVellis RF, Thayer LM, Ammerman AS. Social support for physical activity-role of Facebook with and without structured intervention. Transl Behav Med 2014 Dec;4(4):346-354 [FREE Full text] [doi: 10.1007/s13142-014-0269-9] [Medline: 25584083]

10. Napolitano MA, Hayes S, Bennett GG, Ives AK, Foster GD. Using Facebook and text messaging to deliver a weight loss program to college students. Obesity (Silver Spring) 2013 Jan;21(1):25-31. [doi: 10.1002/oby.20232] [Medline: 23505165]
11. Valle CG, Tate DF, Mayer DK, Allicock M, Cai J. A randomized trial of a Facebook-based physical activity intervention for young adult cancer survivors. J Cancer Surviv 2013 Sep;7(3):355-368 [FREE Full text] [doi: 10.1007/s11764-013-0279-5] [Medline: 23532799]
12. Pagoto S, Waring ME, May CN, Ding EY, Kunz WH, Hayes R, et al. Adapting Behavioral Interventions for Social Media Delivery. J Med Internet Res 2016;18(1):e24 [FREE Full text] [doi: 10.2196/jmir.5086] [Medline: 26825969]
13. Barello S, Triberti S, Graffigna G, Libreri C, Serino S, Hibbard J, et al. eHealth for Patient Engagement: A Systematic Review. Front Psychol 2015;6:2013 [FREE Full text] [doi: 10.3389/fpsyg.2015.02013] [Medline: 26779108]
14. Wang J, Wang Y, Wei C, Yao NA, Yuan A, Shan Y, et al. Smartphone interventions for long-term health management of chronic diseases: an integrative review. Telemed J E Health 2014 Jun;20(6):570-583. [doi: 10.1089/tmj.2013.0243] [Medline: 24787747]
15. Maher CA, Lewis LK, Ferrar K, Marshall S, De BI, Vandelanotte C. Are health behavior change interventions that use online social networks effective? A systematic review. J Med Internet Res 2014;16(2):e40 [FREE Full text] [doi: 10.2196/jmir.2952] [Medline: 24550083]
16. Kim SJ, Marsch LA, Brunette MF, Dallery J. Harnessing Facebook for Smoking Reduction and Cessation Interventions: Facebook User Engagement and Social Support Predict Smoking Reduction. J Med Internet Res 2017 May 23;19(5):e168 [FREE Full text] [doi: 10.2196/jmir.6681] [Medline: 28536096]
17. Siegmund LA, Ahmed HM, Crawford MT, Bena JF. Feasibility of a Facebook Intervention for Exercise Motivation and Cardiac Rehabilitation Adherence: Study Protocol. JMIR Res Protoc 2017 Aug 18;6(8):e162 [FREE Full text] [doi: 10.2196/resprot.7554] [Medline: 28821473]
18. Pew Research Center. Pew Research Center Internet & Technology. Washington, DC 20036 USA: Pew Research Center; 2018 Feb 05. Social Media Fact Sheet URL: http://www.pewinternet.org/ [accessed 2018-05-09] [WebCite Cache ID 6zI6NMs1C]
19. Fox S, Duggan M. Pew Research Center Internet & Technology. Washington, DC 20036 USA: Pew Research Center; 2012 Nov 08. Mobile Health 2012 URL: http://www.pewinternet.org/2012/11/08/mobile-health-2012/ [accessed 2018-05-09] [WebCite Cache ID 6zI6cw96f]
20. Pew Research Center. Pew Research Center Internet & Technology Datasets 2012. Washington, DC 20036 USA: Pew Research Center; 2012 Dec 13. Facebook URL: http://www.pewinternet.org/datasets/2012/ [accessed 2018-05-09] [WebCite Cache ID 6zI7COM8v]
21. Fox S, Duggan M. Pew Research Center Internet & Technology. Washington, DC 20036 USA: Pew Research Center; 2013 Jan 15. Health Online 2013 URL: http://www.pewinternet.org/2013/01/15/health-online-2013/ [accessed 2018-05-09] [WebCite Cache ID 6zI9g5rSE]
22. Anderson M. Pew Research Center Internet & Technology. Washington, DC 20036 USA: Pew Research Center; 2015 Oct 29. Technology Device Ownership: 2015 URL: http://www.pewinternet.org/2015/10/29/technology-device-ownership-2015/ [accessed 2018-05-09] [WebCite Cache ID 6zIA7BSXW]

Edited by G Eysenbach; submitted 01.12.17; peer-reviewed by A Kingsnorth, SJ Kim, D Nault; comments to author 13.02.18; revised version received 01.03.18; accepted 10.04.18; published 14.06.18.

Please cite as:
Bosak K, Park SH
Characteristics of Adults' Use of Facebook and the Potential Impact on Health Behavior: Secondary Data Analysis
Interact J Med Res 2018;7(1):e11
URL: http://www.i-jmr.org/2018/1/e11/
doi:10.2196/ijmr.9554
PMID:29903698

Original Paper

A Social Media-Based Acute Alcohol Consumption Behavior (NekNomination): Case Series in Italian Emergency Departments

Stefania Barbieri[1,2,3*], MD; Paolo Feltracco[1*], MD; Vittorio Lucchetta[1], MD; Rosa Maria Gaudio[2], MD; Alberto Tredese[4], MD; Mauro Bergamini[3], MD, PhD; Gianna Vettore[1], MD; Vincenzo Pietrantonio[1], MD; Francesco Maria Avato[2], MD, PhD; Daniele Donato[5], MD; Deris Gianni Boemo[5], MD; Maria Vittoria Nesoti[5], MD; Rossella Snenghi[4], MD, PhD

[1]Department of Urgent and Emergency Care, University of Padova, Padova, Italy
[2]Forensic Medicine and Toxicology, University of Ferrara, Ferrara, Italy
[3]Preventive Medicine and Risk Assessment, University of Ferrara, Ferrara, Italy
[4]Forensic Medicine and Toxicology, University of Padova, Padova, Italy
[5]Department of Directional Hospital Management, Padova Hospital, Padova, Italy
[*]these authors contributed equally

Corresponding Author:
Stefania Barbieri, MD
Department of Urgent and Emergency Care
University of Padova
Via Giustiniani 2
Padova,
Italy
Phone: 39 3479812611
Fax: 39 0498218289
Email: stefibarbieri118@gmail.com

Abstract

Background: NekNomination, also known as NekNominate, Neck and Nominate, or Neck Nomination, is a social network–based drinking game which is thought to have originated in Australia and spread all over the world between 2013 and 2014. Individuals record videos of themselves while rapidly drinking excessive quantities of alcoholic drinks (necking) and then nominate friends to outdo them within 24 hours; the videos are then posted on social media such as Facebook or YouTube. The consequences of this drinking game have been very dangerous; at least 5 people under age 30 years have died after drinking deadly cocktails, and many others have suffered from alcohol intoxication.

Objective: The goal of the research is to evaluate data about clinically important acute alcohol intoxication among teenagers and young adults and inform and educate the general public, especially parents, teachers, and health workers, about the spreading craze of dangerous Internet-related behavior among today’s teenagers and young people up to the age of 23 years.

Methods: Patients aged 15 to 23 years with acute alcohol intoxication who came to the emergency department (ED) of 2 major hospitals in Italy from January 1, 2011, to June 30, 2014, were included in this study. Data were retrieved from prehospital and intrahospital medical records and included personal information, methods of intoxication, triage color code, date and time of access to the ED, any relevant signs and symptoms, blood alcohol concentration, and diagnosis at discharge.

Results: A total of 450 young patients (male 277/450, 61.5%, female 173/450, 38.5%; age 15 to 16 years 15/450, 3.3%, age 17 to 18 years 184/450, 40.9%, age 19 to 23 years 251/450, 55.8%) were recruited. The causes of intoxication were happy hour, binge drinking, NekNominate, eyeballing, other alcoholic games, or a mix of them. Happy hour was found to be more common among the older patients, whereas NekNominate accounted for almost half of the youngest group of hospitalizations. Eyeballing occurred in 1.6% (7/450) of cases; binge drinking and other alcoholic games caused 23.3% (105/450) and 23.8% (107/450) of hospitalizations, respectively. On admission, 44.2% (199/450) of patients were assigned a red or yellow color code requiring immediate medical attention; about 14% of them required additional medical assistance (after being in the ED) or hospitalization, some in semi-intensive care units.

Conclusions: Our study shows that the increased numbers of hospitalizations due to alcohol intoxication in the adolescent age group, as a consequence of NekNominate or other drinking games, is alarming and represents a serious public health issue. The potential markers of improper use of social networks must be clearly identified, including categories at risk of alcohol abuse, in order to develop intervention and prevention strategies in terms of education and awareness, which may help in averting potentially fatal episodes.

(Interact J Med Res 2018;7(1):e2) doi:10.2196/ijmr.6573

KEYWORDS

alcohol drinking; drinking behavior; underage drinking; binge drinking; alcoholic intoxication; adolescent; neknomination; binge drinking; alcoholic games; social network

Introduction

Background

Drinking games are thought to have originated thousands of years ago, when the ancient Romans played “ *la Passatella*,” a game in which the loser had to pay for the other players’ beer. These episodes very often ended up in brawls [1].

The common concept now is that drinking games are: (1) social drinking events, (2) played according to a set of rules specifying when and how much players should drink, (3) designed to promote the rapid consumption of large amounts of alcohol to facilitate drunkenness, and (4) involve the drinkers in performing cognitive and/or motor tasks [2].

Polizzotto et al [3] classified drinking games into 2 categories: competitive versus noncompetitive and skill-based versus chance-based. The authors noted that competitive games which do not require skill (like NekNominate) typically involve drinking the most or fastest in a short period of time and are therefore the most dangerous. These kinds of drinking games often involve external cues that dictate participants’ alcohol consumption (eg, media games such as drinking each time a character on television says a certain word or phrase).

Drinking games inevitably promote heavy alcohol consumption [2] and, compared with other high-risk drinking behaviors like prepartying (ie, drinking before going out to a social event or gathering [4]), drinking games are a unique high-risk activity because their rules are expressly designed to encourage drunkenness, leading to binge drinking (ie, 5 or more consecutive drinks within a 2-hour time-span on at least 1 day over the past 30 days [5]).

Since 1995, binge drinking has been increasing among teenagers aged 15 to 16 years in Europe [6] and accounts for more than 90% of alcohol consumed by 12- to 17-year-olds [7]. About 16.5% of males and 14% of females aged 12 to 20 years are binge drinkers, and many adolescents start to binge drink at very young ages [8]. Underage alcohol use in general contributes to the top 3 causes of mortality in this age group—injury, homicide, and suicide [9]—and is associated with other high-risk behaviors including suicide attempts, illicit drug use, sexual activity, increased number of sex partners, riding with a driver who has been drinking, and dating violence victimization [10]. In addition, very early drinking (under the age of 14 years) confers additional health risks, including a potential 4-fold increase in the likelihood of developing alcohol dependence [11].

Alternative drinking games practiced by young people have recently appeared. NekNominate is one of them, as well as “kings,” “flip cup,” “beer roulette,” “beer pool,” “tequila roulette,” “speed pennies,” “fuzzy duck,” and many others. “Butt chugging,” also known as “enema alcohol,” is the introduction of alcohol directly into the rectum and colon through the anus [12]; the recent increase in complications related to this drinking game is assumed to be due to alcohol abuse and the growing popularity of various erotic practices [13].

“Eyeballing” is performed by pouring alcoholic substances directly into the eyes; substances with high alcohol content are generally preferred, although cases of misuse of other substances (eg, cinnamon schnapps, 15% to 50% alcohol by volume [ABV]) have also been reported [14]. Drinking hand sanitizer consists of getting drunk by deliberately ingesting the amount of alcohol contained in hand disinfectants as a surrogate for drinkable alcohol [15].

Some of these practices are known; others have not been documented in medical journals or have only been discussed as case reports giving basic information [16]. The popularity of these behaviors is increasing [17] and poses a considerable challenge for emergency department (ED) physicians and workers [18].

Peer pressure is a well-known, influential fuel for excessive alcohol consumption, and Web-based social networks now offer a platform to expand the extent and impact of peer-to-peer coercion [19]. Mu et al [20] also demonstrated a significant dose-response relationship between Internet use and binge drinking, especially among adolescents.

NekNominate is a novel game that involves individuals posting videos of rapid alcohol consumption (“necking”) on social media sites and nominating an acquaintance to exceed their behavior within 24 hours. The named individual will then drink a larger quantity or a more extreme mixture with associated illegal, risky, embarrassing, or self-injurious behaviors. Other substances are often added to the alcohol and have included eggs, goldfish, insects, rodents, battery fluid, motor oil, de-icer, and human urine; associated behaviors have included driving, biking, skateboarding, and swimming while drinking [21]. What draws particular attention to this phenomenon is the fact that, although traditional extreme-consumption drinking games may also be fatal, they often occur in group settings with specific rules and more social control of excessive actions, something which is lacking in NekNominate. This current trend is particularly dangerous, given overt peer pressure with the potential for cyberbullying and online shaming if the nominee does not complete the dare.

Media reports have identified 5 NekNominate-related fatalities in the United Kingdom and Republic of Ireland, all occurring in February 2014 among young men aged 19 to 29 years; 4 from presumed acute alcohol poisoning and 1 from drowning after drinking [21]. However, it is difficult to assess the incidence of or resulting morbidity from this game; there are hundreds of NekNominate videos posted to various social media sites with unknown medical sequelae.

Objectives

The objective of our research is to evaluate data about clinically relevant (ie, requiring hospitalization) acute alcohol intoxication episodes among teenagers and young adults, while focusing on the methods of reaching drunkenness (competitions, games, occasional acute consumption during parties, binge drinking, or recreational alcohol ingestion) by analyzing ED admissions. In addition, we want to inform and educate the general public, especially parents, teachers, and health workers, about the rapid spread of dangerous Internet-related behaviors among today's teenagers.

Methods

This research is a retrospective longitudinal observational single-center study conducted at two major hospitals in Italy. Young patients who accessed the EDs of these hospitals from January 1, 2011, to June 30, 2014, were included in this study. Inclusion criteria are aged 15 to 23 years, diagnosis of acute intoxication by blood test, and positive history of alcohol intoxication due to happy hour, binge drinking, Neknominate, eyeballing, other alcoholic games, or a mix of them, with or without drunkenness-related injuries. Patients who are younger than 15 years or older than 23 years, have no diagnosis of acute alcohol intoxication, or have a negative history for above-mentioned alcohol-related activities are excluded from the study.

Patients included in this study were asked (directly or through their relatives or friends at the ED) for personal information and methods of intoxication. Also recorded for the sample of selected patients were triage color-code (red=emergency requiring immediate attention—zero waiting minutes; yellow=urgent—10 minutes; green=less urgent—30 minutes; and white=nonurgent—60 minutes), the date and time of access to the ED, any relevant signs and symptoms, blood alcohol concentration (BAC, determined with an enzymatic dehydrogenase method and related to the state of consciousness), and discharge diagnosis. Patients were classified by gender (M/F) and age group (15 to 16 years, 17 to 18 years, and 19 to 23 years).

Results

Considering the inclusion and exclusion criteria, 450 young patients who accessed the ED for acute alcohol intoxication due to specific drinking behaviors were recruited. The reasons for their intoxication were happy hour, binge drinking, NekNominate, eyeballing, other alcoholic games, or a mix of these. Among these patients, 15 (3.3%) were aged 15 to 16 years, 184 (40.9%) were 17 to 18 years, and 251 (55.8%) were 19 to 23 years; in total, there were 277 boys (61.5%) and 173 (38.5%) girls. The different distribution of alcoholic consumption classified by drinking behavior and age group in the general population is listed in Table 1.

Generally, young adults accessed the ED during the weekend at night, and alcohol-related hospitalizations therefore had an important effect on ED personnel, particularly on weekends. Conversely, Neknomination cases in our sample of adolescents accessed the ED during afternoons between Monday and Friday and during evening hours from Wednesday to Sunday. Half (226/450, 50.2%) of the patients were admitted between 12 AM and 8 AM.

None of the 15- to 16-year-old patients had alcohol intoxication while attending a happy hour event, but 30.4% (56/184) and 24.7% (62/251) of the 17- to 18- and 19- to 23-year-old groups, respectively, got drunk during happy hour. NekNominate-related hospitalizations totaled 5.8% (26/450) in the entire sample and represented almost half of the youngest group of hospitalizations. Eyeballing accounted for 1.6% (7/450) cases; binge drinking and other alcoholic games caused 23.3% (105/450) and 23.8% (107/450) of hospitalizations, respectively. A total of 53.2% (98/184) of the patients aged 19 to 23 years had been driving under the influence of alcohol during the late hours of the weekend.

Table 1. Distribution of alcohol consumption classified by drinking behavior and age group in the general population.

Behavior	Number by age group and sex						Total, n (%)
	15-16 (n=15)		17-18 (n=184)		19-23 (n=251)		(N=450)
	M (n=10)	F (n=5)	M (n=110)	F (n=74)	M (n=157)	F (n=94)	
Happy hour	0	0	32	24	41	21	118 (26.2)
Binge drinking	3	3	25	26	31	17	105 (23.3)
Neknomination	6	1	6	4	8	1	26 (5.8)
Other alcoholic games	1	1	23	6	40	36	107 (23.8)
Eyeballing	0	0	3	0	4	0	7 (1.6)
Mixed behaviors	0	0	21	14	33	19	87 (19.3)

BAC in the population varied between 1.5 and 2.4 g/L; 56% (252/450) of patients stated that they had consumed an excessive amount of alcohol in a short period of time. The most common signs and symptoms included heart palpitations (231/350, 66.0%), headache (119/215, 55.3%), diarrhea (45/102, 44.1%), nausea and vomiting (286/432, 66.2%), anxiety (2/7, 28.6%), problematic behavior (5/7, 71.4%), hallucinations (54/215, 25.1%), convulsions (5/225, 2.2%), coma (68/450, 15.1%), unconscious state (194/450, 43.1%), slow respiration (81/450, 18.0%), and rectal edema and severe acute colitis (1/10, 10.0%).

On hospital admission, 44.2% (199/450) of the patients were assigned a red or yellow color code, requiring immediate medical attention; 4.4% (20/450) were transferred to the short-term medical observation ward because they required continuous clinical evaluation, diagnostic testing, and appropriate treatment protocols; 3.3% (15/450) were admitted to a semi-intensive care unit; 80.4% (362/450) were discharged within 72 hours; 6.0% (27/450) were hospitalized; 2.0% (9/450) were spontaneously discharged; and 4.3% (18/450) refused hospitalization.

Discussion

Principal Findings

To our knowledge, this is the first study analyzing the clinical consequences of drinking games, with particular attention to NekNominate as a social media–based phenomenon. Some of these practices are already known; others have not been documented in medical journals or have only been discussed as case reports giving basic information. Drinking behaviors including the alcoholic games Neknomination and eyeballing are still not well recognized as public health problems, but increasing numbers of adolescents are admitted to EDs for acute alcoholic intoxication.

Our sample contained 450 young patients aged 15 to 23 years who accessed the ED after ingesting excessive amounts of alcohol in an unconventional and dangerous way. Most of them were between 19 and 23 years old and male. Of all patients, happy hour (a special offer, in which drinks are cheaper at a certain time of day) was the most common way of reaching severe drunkenness (26% of all ED accesses), followed by various alcoholic games, binge drinking, mixed behaviors, NekNomination and eyeballing. Happy hour only affected patients aged 17 to 23 years, indicating that this practice is not frequent among younger people. Conversely, NekNomination was the most common cause of alcohol intoxication among the youngest patients (aged 15 to 16 years), representing 47% of all hospitalizations in this age group. This figure adds to the dangerousness of this practice in view of the fact that the earlier alcohol abuse begins, the worse the consequences (eg, problematic alcohol use in later life and increased susceptibility to high-risk styles of drinking [22]). In addition, as a social media–based drinking game, this highlights the importance of peer pressure, which is particularly influential in younger adolescents who are less able to resist this type of coercion.

Alcoholic games in general have caused increasing numbers of hospitalizations directly proportional to patients' ages: 13% in the 15- to 16-year-old group, 16% in those aged 17 to 18 years, and 30% in those aged 19 to 23 years, and this matches the most recent reviews about the correlation of age with drinking games [22]. Most of the older patients were hospitalized at night over weekends; conversely, younger adolescents who practiced NekNomination accessed the ED during working days in the afternoon, indicating different drinking patterns in the various age groups.

The abuse of alcoholic drinks is a significant source of risk for adolescents and is one of the primary causes of many accidental injuries, such as falls, accidents at home, and road accidents (38% of the patients aged 19 to 23 years were driving under the influence of alcohol, thus also representing a danger for others). A large proportion (44%) of patients admitted to the ED were assigned a red or yellow color code, meaning that they required immediate or almost immediate treatment; about 14% of them required additional medical assistance (after being in the ED) or hospitalization, sometimes in a semi-intensive care unit, representing a significant burden in terms of personnel requirements and costs for diagnosis and treatments. The most common presenting signs and symptoms were those typical of alcohol intoxication: heart palpitations (66%), headache (55%), diarrhea (44%), nausea and vomiting (66%), and problematic behavior (63%).

NekNominate seems to be an on-off phenomenon related to 2013-2014; nevertheless, Wombacher et al [23] wrote about it and its relation to social norms in June 2016. Although NekNomination is no longer shown in the media and is now not a search trend in Google or Twitter, this does not mean the practice has disappeared: modern video exchange methods, (eg, Snapchat or WhatsApp) are now more private than in the past and allow users to exchange videos between individuals or groups of selected friends. Thus, in addition to the above considerations, we decided to focus on NekNominate because of its peculiar aspects: (1) it was probably the first drinking game which spread enormously through social media websites, allowing all participants to confront anyone anywhere in the world and thus setting a milestone in dangerous Web-based challenges, (2) it belongs to the category of competitive games that do not require skill and are proven to be the most hazardous [3], (3) it has recently caused drastic consequences such as deaths and hospitalizations (as reported here), (4) the role of peer pressure is extremely significant, as people refusing to accept the challenge are potentially victims of cyberbullying and therefore tend to exaggerate their behaviors in order not to be excluded by their peers, and (5) literature investigating this high-risk practice is still not frequently available.

We present the various patterns of alcohol intoxication among adolescents with the image of an iceberg to highlight the magnitude of unknown behaviors compared with what is known in the scientific literature (Figure 1). There are many news reports on binge drinking and happy hour–related alcohol intoxication and alcohol (vodka) enema or eyeballing, especially concerning young males aged 24 to 27 years. Nevertheless, the scientific literature appears to be lacking in regard to these games, which are spreading at a dangerous rate and are rapidly affecting adolescents all over the world.

Figure 1. The iceberg of current knowledge of excessive drinking behaviors.

Most of the iceberg is below the waterline and represents people’s unawareness regarding young people’s acute alcohol consumption. Binge drinking, Neknomination, and alcoholic games are also often associated with abnormal behaviors such as not wearing a helmet while riding a motorbike, not fastening a seatbelt while driving a car (both illegal in Italy/Europe), having promiscuous sex, and consuming illicit drugs.

The widespread use of social media is having a severe effect on our society: its easy accessibility makes it one of the most commonly used tools among teenagers and young adults, and the changes induced in current behaviors are still under evaluation [24]. Social media certainly played an important role in the spread of NekNominate; posting videos on the Web in order to collect likes and visualizations seems to be a way of getting into a group or being socially accepted by peers, and inviting friends to do the same has made the game spread rapidly among young people.

Limitations

This study is limited by the small sample size, which is not representative of the entire Italian teenager population. However, it is still useful for preliminary data about drinking games.

Conclusions

Drinking games are still relatively undisclosed and “at the bottom of the iceberg” of health workers’ attention and knowledge. They encourage young people to consume large quantities of alcohol within a short period of time, putting them at risk of alcohol intoxication that can potentially cause injuries, later regretted sexual activities, suicides, and traffic accidents and deaths.

Social media is an important tool for connections among young people and may be associated with behavioral and social maladjustment among adolescents. It also represents a powerful source of peer pressure. Adolescents who consume alcohol in new and dangerous ways such as NekNomination reach EDs with high BAC, are classified as red or yellow codes at triage, present severe impairment of vital functions, and often require further intensive medical assistance.

In order to develop a prevention strategy which could help to avert potentially fatal episodes, it is necessary to highlight the potential markers of problematic social network use and to inform and educate health personnel, social workers, and school teachers about these new phenomena, which are spreading very rapidly among adolescents. Further studies to better describe and analyze the problem, particularly in its clinical aspects, are in progress.

Acknowledgments

The authors gratefully acknowledge the support of Michele Lazzari, Chiara Billato, Anna Dodi, Giorgia Milanopulo, Luca Marinuzzi, and Gloria Sguotti for their contribution to the manuscript and technical assistance throughout the process.

Conflicts of Interest

None declared.

References

1. Davis J. Passatella: an economic game. Brit J Sociol 1964;15(3):191-206.
2. Zamboanga BL, Pearce MW, Kenney SR, Ham LS, Woods OE, Borsari B. Are extreme consumption games drinking games? Sometimes it's a matter of perspective. Am J Drug Alcohol Abuse 2013 Sep;39(5):275-279 [FREE Full text] [doi: 10.3109/00952990.2013.827202] [Medline: 23968169]
3. Polizzotto MN, Saw MM, Tjhung I, Chua EH, Stockwell TR. Fluid skills: drinking games and alcohol consumption among Australian university students. Drug Alcohol Rev 2007 Sep;26(5):469-475. [doi: 10.1080/09595230701494374] [Medline: 17701509]
4. Zamboanga BL, Tomaso CC, Haas AL, Olthuis JV, Borish S, Borsari B. Trouble brewing: pregaming among high school and incoming college students. J Psychosoc Nurs Ment Health Serv 2013 Oct;51(10):14-17 [FREE Full text] [doi: 10.3928/02793695-20130828-01] [Medline: 24016135]
5. NIAAA Newsletter Issue 3. 2004. URL: https://www.niaaa.nih.gov/sites/default/files/newsletters/Newsletter_Number3.pdf [accessed 2017-12-12] [WebCite Cache ID 6vfVWLwGc]
6. Farke W, Anderson P. Binge drinking in Europe. Adicciones 2007;19(4):333-339 [FREE Full text] [Medline: 18173097]
7. Nelson JP. Binge drinking and alcohol prices: a systematic review of age-related results from econometric studies, natural experiments and field studies. Health Econ Rev 2015;5:6 [FREE Full text] [doi: 10.1186/s13561-014-0040-4] [Medline: 25853004]
8. Results from the 2013 National Survey on Drug Use and Health: summary of national findings. 2014. URL: http://www.samhsa.gov/data/sites/default/files/NSDUHresultsPDFWHTML2013/Web/NSDUHresults2013.pdf [accessed 2017-12-13] [WebCite Cache ID 6vgU9nwuc]
9. The Surgeon General's Call to Action To Prevent and Reduce Underage Drinking. 2007. URL: https://www.ncbi.nlm.nih.gov/books/n/sgdrinkt/pdf/ [accessed 2017-12-13] [WebCite Cache ID 6vgUTd3TQ]
10. Miller JW, Naimi TS, Brewer RD, Jones SE. Binge drinking and associated health risk behaviors among high school students. Pediatrics 2007 Jan;119(1):76-85. [doi: 10.1542/peds.2006-1517] [Medline: 17200273]
11. Hingson R, Heeren T, Winter M. Age at drinking onset and alcohol dependence: age at onset, duration, and severity. Arch Pediatr Adolesc Med 2006 Jul;160(7):739-746. [doi: 10.1001/archpedi.160.7.739] [Medline: 16818840]
12. El Mazloum R, Snenghi R, Barbieri S. Butt-chugging, a new way of alcohol assumption in young people. Eur J Public Health 2015;25(suppl 3).
13. Peterson T, Rentmeester L, Judge BS, Cohle SD, Jones JS. Self-administered ethanol enema causing accidental death. Case Rep Emerg Med 2014;2014:191237 [FREE Full text] [doi: 10.1155/2014/191237] [Medline: 25436159]
14. Bersani F, Corazza O, Albano G, Bruschi S, Minichino A, Vicinanza R, et al. The “Eyeballing” technique: an emerging and alerting trend of alcohol misuse. Eur Rev Med Pharmacol Sci 2015 Jun;19(12):2311-2317 [FREE Full text] [Medline: 26166661]
15. Gormley NJ, Bronstein AC, Rasimas JJ, Pao M, Wratney AT, Sun J, et al. The rising incidence of intentional ingestion of ethanol-containing hand sanitizers. Crit Care Med 2012 Jan;40(1):290-294 [FREE Full text] [doi: 10.1097/CCM.0b013e31822f09c0] [Medline: 21926580]

16. Baldwin JM, Stogner JM, Miller BL. It's five o'clock somewhere: an examination of the association between happy hour drinking and negative consequences. Subst Abuse Treat Prev Policy 2014 Apr 23;9:17 [FREE Full text] [doi: 10.1186/1747-597X-9-17] [Medline: 24758616]
17. Bosmia A, Griessenauer C, Tubbs R. Vodka eyeballing: a potential cause of ocular injuries. J Inj Violence Res 2014;6(2):93-94. [doi: 10.5249/jivr.v6i2.540]
18. Zamboanga BL, Bean JL, Pietras AC, Pabón LC. Subjective evaluations of alcohol expectancies and their relevance to drinking game involvement in female college students. J Adolesc Health 2005 Jul;37(1):77-80. [doi: 10.1016/j.jadohealth.2004.12.007] [Medline: 15963914]
19. Barbieri S, Feltracco P, Vettore G. Evolution and revolution of dangerous drinking games among adolescents and young people. Eur J Public Health 2015;25(suppl 3). [doi: 10.1093/eurpub/ckv175.074]
20. Zonfrillo M, Osterhoudt K. NekNominate: a deadly, social media-based drinking dare. Clin Pediatr (Phila) 2014 Oct;53(12):1215. [doi: 10.1177/0009922814536265] [Medline: 24872340]
21. Mu K, Moore S, LeWinn K. Internet use and adolescent binge drinking: findings from the monitoring the future study. Addict Behav Rep 2015 Dec 01:61-66. [doi: 10.1016/j.abrep.2015.09.001]
22. The Daily Mirror. 2014 Feb 16. Neknominate craze—the 'fun' internet drink challenge that WILL kill again URL: http://www.mirror.co.uk/news/uk-news/neknominate-fun-internet-drinking-craze-3119787 [accessed 2017-12-13] [WebCite Cache ID 6vgVgAL7y]
23. Wombacher K, Reno J, Veil S. NekNominate: social norms, social media, and binge drinking. Health Commun 2017 May;32(5):596-602. [doi: 10.1080/10410236.2016.1146567] [Medline: 27331576]
24. Suffoletto B, Kristan J, Callaway C. A text message alcohol intervention for young adult emergency department patients: a randomized clinical trial. Ann Emerg Med 2014 Dec;64(6):e4. [doi: 10.1016/j.annemergmed.2014.06.010]

Abbreviations

ABV: alcohol by volume
BAC: blood alcohol concentration
ED: emergency department

Edited by G Eysenbach; submitted 30.08.16; peer-reviewed by A Kucharski, C Rios-Bedoya, H Hendriks; comments to author 20.12.16; revised version received 24.01.17; accepted 19.11.17; published 31.01.18.

Please cite as:
Barbieri S, Feltracco P, Lucchetta V, Gaudio RM, Tredese A, Bergamini M, Vettore G, Pietrantonio V, Avato FM, Donato D, Boemo DG, Nesoti MV, Snenghi R
A Social Media-Based Acute Alcohol Consumption Behavior (NekNomination): Case Series in Italian Emergency Departments
Interact J Med Res 2018;7(1):e2
URL: http://www.i-jmr.org/2018/1/e2/
doi:10.2196/ijmr.6573
PMID:29386170

Original Paper

Weight Stigma Goes Viral on the Internet: Systematic Assessment of YouTube Comments Attacking Overweight Men and Women

Yongwoog Andrew Jeon[1], MA, MPhil; Brent Hale[2], MA; Eric Knackmuhs[3], PhD; Michael Mackert[1], PhD

[1]Center for Health Communication, Moody College of Communication, The University of Texas at Austin, Austin, TX, United States
[2]The Media School, Indiana University Bloomington, Bloomington, IN, United States
[3]School of Kinesiology, Recreation & Sport, Western Kentucky University, Bowling Green, KY, United States

Corresponding Author:
Yongwoog Andrew Jeon, MA, MPhil
Center for Health Communication
Moody College of Communication
The University of Texas at Austin
300 W Dean Keeton (A0900)
Austin, TX, 78712
United States
Phone: 1 8122233163
Email: yongwoog@utexas.edu

Abstract

Background: Anonymous verbal attacks against overweight individuals on social media are common and widespread. These comments often use negative, misogynist, or derogatory words, which stigmatize the targeted individuals with obesity. These verbal attacks may cause depression in overweight individuals, which could subsequently promote unhealthy eating behavior (ie, binge eating) and further weight gain. To develop an intervention policy and strategies that tackle the anonymous, Web-based verbal attacks, a thorough understanding of the comments is necessary.

Objective: This study aimed to examine how anonymous users verbally attack or defend overweight individuals in terms of 3 themes: (1) topic of verbal attack (ie, what aspects of overweight individuals are verbally attacked), (2) gender of commenters and targeted overweight individuals, and (3) intensity of derogation depending on the targeted gender (ie, the number of swear words used within comments).

Methods: This study analyzed the content of *YouTube* comments that discuss overweight individuals or groups from 2 viral videos, titled “Fat Girl Tinder Date” and “Fat Guy Tinder Date.” The twin videos provide an avenue through which to analyze discussions of obesity as they organically occurred in a contemporary setting. We randomly sampled and analyzed 320 comments based on a coding instrument developed for this study.

Results: First, there were twice as many comments verbally attacking overweight individuals (n=174) than comments defending them (n=89). Second, overweight women are attacked for their capacities (eg, laziness, maturity; 14/51, 28%), whereas overweight men are attacked for their heterosocial skills (eg, rudeness, annoyance; 24/29, 83%). Third, the majority of commenters who attacked overweight women are male (42/52, 81%). Fourth, attacking comments generated toward overweight women included more swear words (mean 0.44, SD 0.77) than those targeting men (mean 0.23, SD 0.48).

Conclusions: Our data elucidate a worrying situation of frequent disinhibited aggressive messages against overweight individuals online. Importantly, the patterns of verbal aggression differ depending on the gender of the targeted overweight individuals. Thus, gender-tailored intervention strategies that specifically tackle Internet users’ verbal aggression against overweight individuals need to be developed.

(Interact J Med Res 2018;7(1):e6) doi:10.2196/ijmr.9182

KEYWORDS

stigma; cyberbullying; gender; sex differences; verbal behavior

Introduction

Background

Weight stigma—negative attitudes and beliefs toward individuals with obesity—is pervasive across societies [1]. Weight stigma often manifests stereotypes, rejection, and prejudice toward individuals with obesity. One of the most common forms of expressing weight stigma is through verbal attack or verbal bullying defined as statements that attack the self-concept of the receiver, intending to deliver psychological pain (eg, teasing, ridiculing, derogating, devaluing, humiliating). Research shows that verbal attacks on overweight individuals are common on social media. A recent study, which analyzed 1.37 million posts collected from various social media channels including YouTube, found that 92 percent of the posts related obesity with negative, misogynist or derogatory words [2]. Additionally, researchers suggest that weight-based aggressive comments are highly likely to induce depression or anxiety in overweight individuals [3]. Depression could subsequently result in binge eating and weight gain, which may increase exposure to further weight-stigma and, at worst, even lead to suicide [4]. Due to these adverse effects of weight stigma, various authors called for more intervention efforts, including development of health messages, to reduce weight stigma and relieve mental burden for individuals with obesity [5,6]. Effectiveness of health messages to reduce weight stigma largely depends on precise knowledge of the audience—those being stigmatized (ie, obese individuals) and those who stigmatize (ie, individuals who verbally attack individuals with obesity) [7,8]. Particularly, men and women may experience weight stigma differently [9]. Taking this into account, gender-tailored messages need to be developed since they enhance persuasion by increasing perceived relevance [10]. To develop health messages tailored to female and male audiences, it is important to understand how overweight women and men are stigmatized *differently* on social media, and how female and male commenters stigmatize overweight individuals differently.

This study asks: Which negative trait is most dominantly associated with overweight males or females? As evolutionary psychological scholars suggest, verbal aggression against men generally addresses men's physical or mental strength (eg, laziness, physical weakness) while those against women tends to be about their appearance [11]. This study attempted to establish this topical difference in verbal attack as it applies to overweight individuals. Second, in terms of those who stigmatize, we attempted to find out what the major concerns of male and female *YouTube* users are when commenting on overweight individuals. Two similar videos (female and male Tinder videos), which were sequentially published to YouTube, provided an opportunity to minimize confounds due to the gender of the obese model in the video (ie, the female model may attract more comments targeting female overweight individuals or vice versa). Furthermore, these comments provide an opportunity to conduct a natural experiment to gain insight into how the commenters—which would not be made in a laboratory setting due to the social desirability issue—attack or defend individuals with obesity.

It is important to develop effective, gender-tailored intervention strategies that are fully informed by theoretical understanding and empirical findings on the verbal aggression targeting overweight individuals on the Internet. Thus, for this purpose, the study's coding-instrument, based on linguistic and evolutionary psychological theories, was developed to analyze comments posted in response to the *YouTube* videos. In the following section, the theoretical framework for the analysis of the aggressive comments is discussed. On the basis of this review, the specific research questions are proposed.

Gender Differences in Weight Stigma

Some cultures value obese human bodies as expressing beauty, marriageability, control of selfish desires, generosity, fertility, and closeness to God [1]. Yet across many contemporary cultures, obesity is strongly linked with various negative traits [1], which include laziness, unattractiveness, lack of will power, poor self-control, slowness, inactiveness, physical weakness, overeating, ineptitude, sexlessness, and unhealthiness [12,13]. Obese individuals have reported frequently experiencing verbal attacks that target these negative traits in their everyday lives [14] from strangers (eg, an overweight woman was called a "big fat pig" and physically assaulted by a stranger on a commuting train) [15] or from a spouse (eg, a husband called his wife "disgustingly fat") [16]. More evidence has been recently accumulated to suggest that the association between negative traits and obesity is becoming a global issue [1].

However, most research on weight stigma has not explicitly addressed possible gender variations in associations of negative traits with obesity. By understanding how male and female overweight individuals are attacked, one may inductively infer how people's attitude toward overweight males differs from that toward overweight females. For instance, in the context of gender differences, the following questions can be asked: Are overweight females attacked for laziness more frequently than obese males? Are obese males attacked for physical weakness more commonly than females? With a lack of current empirical research exploring these questions, this study draws upon evolutionary psychological views to provide theoretical grounding.

Evolutionary perspectives offer the basis for explaining (at a fundamental level) and predicting sex differences in verbal aggression. Scholars in evolutionary psychology claim that to maximize reproductive success (ie, the number of their offspring), males, on the one hand, evolved to infer a mate's reproductive success from physical appearance, more so than females [11,17-20]. Healthy skin, hair, and low hip-waist ratio are some of the visual cues for *reproductive power* [17] . On the other hand, females evolved to infer a mate's *capacities for resource acquisition* or the ability to provide protection or food [17]. Accordingly, females primarily sought mates with resource acquisition-signaling features most of which were not physically visible: including possession of territory, shelter, industriousness, or high social status [21]. These evolutionary psychological views have been supported by various studies examining contemporary people's feature preferences in mates [22-24]. For instance, in a survey that sampled more than 1000 undergraduates, women responded that they value males' education and financial prospect (ie, capacities for resource acquisition) more than “good looks” (ie, reproductive power), whereas men said that they value females’ physical attractiveness more than financial prospect [22].

Such gender difference in preferred features of mates may predict which personal characteristics will be targeted by aspects of verbal aggression. Buss and Dedden (1990) conducted an experimental study where they investigated what a man or woman would derogate about a perceived competitor for a mate. They found that when women verbally attacked other competing women, their reproductive power tended to be addressed in the verbal aggression (eg, “she is physically unattractive”) [11]. When men verbally attacked other competing men, their capacities for resource acquisition were likely to be addressed (eg, “the man lacked ambition” or “the other man is poor”). Informed by the evolutionary psychological perspective, the reproductive power and capacities for resource acquisition dichotomy is employed in this study. Utilizing this dichotomy, we examined verbally aggressive comments in relation to gender difference within the Tinder date videos. Videos and comments on *YouTube* addressing overweight individuals may show how they are stigmatized on social media [25,26]. While findings from previous studies are useful for understanding what generates stigmatic content on *YouTube*, various important questions regarding gender difference in weight stigma remain unanswered. Thus, the following research questions (RQ) were developed for this purpose:

> *RQ1a: To what extent are overweight women targeted regarding reproductive power more than men in aggressive comments within YouTube comments?*
>
> *RQ1b: To what extent are overweight men targeted regarding capacities for resource acquisition more than women in aggressive comments within YouTube comments?*

Other research suggests that people judge their romantic partners based on heterosocial skills [22], which generally refer to a person’s ability to carry a good conversation or express a good personality. Specifically, studies suggest that males are likely to succeed in finding a female mate if he is heterosocially skilled [27,28]. However, to our knowledge, little is known about how obese individuals’ heterosocial skills are perceived. Like other negative traits attributed to overweight individuals, heterosocial skills could be underestimated in order to maintain negative valence. To explore this possibility within weight-based verbal attacks in *YouTube* comments, the following research question is proposed:

> *RQ2: What gender differences exist in how overweight individuals’ heterosocial skills are attacked within YouTube comments?*

Finally, among the 3 explored topics of verbal attack (ie, resource acquisition, reproductive power, and heterosocial skills), we explore gender-based differences in topical prominence:

> *RQ3: What gender differences exist between the prominence of attacks on reproductive power, capacities for resource acquisition, and heterosocial skill?*

Gender Difference in Susceptibility to Weight Stigma

Many researchers have suggested that obese women are more susceptible to weight stigma than men [9,29]. In a study conducted by Chen and Brown (2005), 449 college students were asked to order 6 pictures of potential sexual partners according to their personal preference [30]. The pictures included an obese person, a nonobese person, and individuals with various disabilities. A rating comparison between subjects showed that men rated an obese partner less preferentially than women did. In addition, men are more likely to choose sexual partners based on a potential partner’s weight. Overall, their findings showed that women may be more vulnerable to weight stigma exposure than men. In contrast to these findings; however, Hussin et al (2011) found that across the 50 most viewed *YouTube* videos which devalued overweight individuals, overweight men are twice as likely to be targets than overweight women [25].

Men are primarily considered the perpetrators of weight stigma in a masculine society, as for women “becoming an attractive object is a role obligation” [31]. Thus, men are less likely to feel guilty when they fulfill masculinity by victimizing women whom they perceive as deviant [32]. However, in the context of romantic relationships, Buss and Dedden (1990) argued that women evolved to engage in verbal aggression against their own sex since members of a sex must compete for access to a mate with desired qualities (eg, reproductive power or physical strength) [11]. A question may be raised as to whether this is the case for overweight individuals of the same sex or not. Unlike slender women, overweight women may be perceived as less physically attractive (and less of a competitor), and thus would not require verbal aggression as frequently. Alternatively, it is possible that women will defend overweight women due to contemporary cultural or social factors. Therefore, we sought to answer the following research questions regarding the presence of attack and defense between genders:

RQ4: *Which gender more frequently verbally attacks or defends overweight individuals? (a) men or (b) women?*

Intensity of Derogation—Toxic Disinhibition Effects and Swearing

Many scholars claim that the Internet is an arena where people feel disinhibited to state what they would not ordinarily say face-to-face [33-36]. Suler (2004) argues that this is an affordance of anonymity online—more so than face-to-face communication— allowing Internet users to feel less restrained by the consequences of saying what is ordinarily socially undesirable, which includes swearing or expressing hatred against overweight people. This phenomenon is termed as “Toxic Disinhibition Effects” [36]. According to the perspective of Toxic Disinhibition Effects, people would feel less inhibited in expressing verbal aggression against individuals with obesity, potentially resulting in a plethora of derogatory or aggressive comments targeting obese people [2].

A body of evidence has supported the disinhibition effect [2,37,38]. For example, in a content analysis of 9376 Myspace pages by Thelwall (2008), it was found that the pages of nearly all young and half of middle-aged Myspace users contained swear words. Furthermore, Chou et al (2014) analyzed 2.2 million posts across Twitter and Facebook, and they found that keywords including "obesity," "overweight," and "fat" are associated with derogatory terms including "a**" or "b****." Also, qualitative analysis of these posts revealed that the most prevalent theme is weight stigma (eg, "you're an ugly fat b****. Kill yourself") [2]. Within *YouTube*, these verbal aggressions were found to be common as well [38,39]. On the basis of this literature, the following research questions are formulated:

> *RQ5a: To what comparative extent do overweight individuals receive more verbal attacks than defense?*
>
> *RQ5b: To what extent are verbal attacks accompanied by swear words?*
>
> *RQ5c: To what comparative extent do male and female overweight individuals receive verbal attacks which include swear words?*

Methods

Study Design

This study entailed a content analysis of *YouTube* comments that discuss overweight individuals or groups from 2 videos, titled "Fat Girl Tinder Date" and "Fat Guy Tinder Date." These 2 videos were released approximately 3 weeks before the beginning of this study, providing an avenue through which to analyze discussions of obesity as they organically occurred in a digital setting. Utilizing this window of opportunity, the authors sought to evaluate how attacks toward the stigmatized may occur in this context.

Background Information and Rationales for Choosing the Two Videos

On September 24, 2014, a video titled "Fat Girl Tinder Date" (hereafter the female Tinder video) [40] was published by a *YouTube* content creator. In this video, a slender woman set up a number of dates through Tinder, a matchmaking mobile app, using a profile picture which accurately depicted her appearance; however, before appearing for each date she disguised herself as an obese woman by wearing body adhesives. Hidden cameras videotaped how men reacted to the "obese" woman. Later, a similar video was released titled "Fat Guy Tinder Date" (hereafter the male Tinder video) [41], with a slender man attending Tinder-arranged dates wearing comparable body adhesives. After their release, these videos went viral, accumulating over 12 million views and approximately 4000 user-submitted comments within 3 weeks. One of the most popular comments (indicated by the number of "thumbs up" received) states "I f****** hate fat b****," and another one says "If you're fat just move your f****** a** and go to a gym". The majority of obesity-related *YouTube* videos explicitly use weight-based teasing (eg, verbally mocking individuals with obesity or blaming them for their weight statuses) [26]. These videos, therefore, may promote negative comments on individuals with obesity [25].

The 2 videos chosen for this study exemplify the obesity-related *YouTube* videos that contain such a negative theme, though less explicitly [26]. Also, in general, the narrative structure of the 2 videos are similar; yet everything is not equal (eg, contents of conversations). Despite the potential confounds coming from difference between the contents of the 2 videos, these videos best serve our research purpose: the examination of gender difference in comments on obese individuals. This is in that the gender of the person with obesity may result in more comments about the traits that are stereotypically linked with obese individuals of the same gender (eg, a male obese model in the video promotes discussion of traits typically linked with an obese man than woman). As in our data, the Male Tinder video attracts significantly more male comments than the Female Tinder video and *vice versa* (Table 1) . Thus, to minimize the confounds coming from the gender of the model in the video, we chose the 2 "twin" videos to sample comments from.

More importantly, the comments that may explicitly verbally attack obese individuals may not be obtained in a survey or a laboratory experiment due to social desirability [42]. Instead, the 2 Tinder videos provided a setting for "a natural experiment" [43]. A natural experiment has been increasingly employed for research in social media contexts [16,44]. The producer purposefully made these 2 videos highly similar to see how different genders would react to the obese blind-date partners. Then they posted these 2 videos on *YouTube*, which inadvertently provided a setting for a natural experiment.

Data Collection and Sampling

Approximately 3 weeks after the 2 videos were released, all available comments were captured and saved in 2 word processing files (1 for each video). The captured material included the commenter's profile photo, the commenter's username, the comment text, and the timestamp for the comment's creation. Upon completion of data collection, the authors processed a random sample of 320 comments from the data. The random sample was procedurally selected through the utilization of a random number generator, which selected pages of the word processing file to analyze for relevant comments. All comments were then subjected to our inclusion criteria (explained in the next section), and rejected comments were replaced by continuing the selection process until a sample of 320 comments was obtained. Once the sample was collected, the authors conjointly reviewed the comments to achieve agreement that each entry passed our inclusion criteria, which left us with 316 verified comments to analyze. A total of 145 comments were selected from the "Fat Guy Tinder Date" video, while 171 comments were selected from the "Fat Girl Tinder Date" video.

Inclusion Criteria

Each comment was subjected to a filter designed to allow analysis of only the relevant comments. The first criterion was that each comment needed to contain some reference to an individual or group being overweight or make a comment about overweight people generally. The second criterion was that each selected comment needed to make some statement about weight, including an opinion, criticism, or evaluation about being overweight or the experiences of being overweight. In addition, we intended to exclude those comments directed at the actor or actress, not overweight individuals in general. No comment in our dataset was found and excluded for this reason.

Variables of Interest

Each of the following variables was developed for the purpose of this study, and intercoder reliability for each variable is reported based on Krippendorff alpha, a statistical measure of the interreliability among the coders for our variables [45]. The overall Krippendorff alpha is .85.

Gender of Commenter (Alpha=.70)

This study used the commenters' names and profile pictures to infer gender. Although it could be posed that identification of true commenter gender cannot be accurately conducted within digital contexts due to potentially unidentifiable gender swapping, some research has shown that females and males tend to represent themselves accurately in their profile photo or name rather than engaging in gender swapping [46-48]. There may be other options (eg, directly contacting the commenters by leaving a message in their channel or "vlog" on *YouTube*). However, we realized that contacting more than 300 commenters was extremely difficult and time-consuming. Even if we contacted the commenters, perhaps due to potential social desirability issue, *YouTube* commenters would probably hesitate to provide their true identity (eg, gender). Thus, inference of gender from their uses of profile name and photos was the second best approach.

Commenter gender was coded when identifiable in the profile photo, username, or comment text. In many cases, gender was determinable in at least one of these locations, and no instances occurred where gender determination conflicted between locations (ie, male profile photo and female username occurring simultaneously). In terms of the profile photo, the 3 coders judged whether the face is male or female. In addition, we used the database on the names trending in America which is provided for researchers by US Social Security Administration [49]. This database matches names with gender (eg, Jake for male). When identifiable, commenters were coded as male or female; otherwise, indeterminable. Thus, we examined both usernames and profile photos to infer the gender of the commenter and compared this result with information within the comment text. To focus on our primary research purpose (examination of gender differences in the comments on overweight individuals), we excluded those comments whose profiles do not provide clear clues of gender identifications.

Gender of Target (Alpha=.88)

Comments were coded according to the gender of the target of the comment text. The target of the comment text was determined as the individual or group whose weight is being discussed or evaluated. In some instances many targets existed, while most included only a single target. When identifiable, this was coded as male, female, or both in cases where both genders were targeted by the comment; otherwise, neither.

Attack (Alpha=.93)

Comments were categorized as containing an attack when the comment text offended, belittled, or expressed disapproval of an individual or group being overweight. In cases where only portions of comment text concerned weight, only those sections that passed our inclusion criteria were evaluated for attack potential.

Defense (Alpha=.88)

Comments were categorized as containing a defense when the comment text protected or supported an individual or group's weight, or an overweight individual's actions, thoughts, or traits related to weight. As with attack, only those portions of comment text that passed the inclusion criteria were evaluated.

Evolutionary Mate Values (Alpha=.77)

This variable was operationalized in 2 distinct ways, resulting in 2 categories. A comment was categorized as an attack or defense of capacities for resource acquisition in cases where the following were targeted: (1) an overweight individual or group's physical or mental ability and (2) personal characteristics of responsibility, laziness, health, or maturity. Comments were categorized as an attack or defense of reproductive power in cases where an overweight individual or group's attractiveness, ability to find a mate, give birth, or raise a child are targeted. These categories were adapted from Buss and Dedden (1995).

Heterosocial Skills (Alpha=.71)

A comment was categorized as an attack or defense of relational skills when the comment text addressed an individual or group's ability to interact with others, which included attacks on social skills or behavior, ability to be selected as a mate, and desirability.

Intensity of Derogation (Alpha=.89)

The number of swear words were counted and recorded to indicate derogation intensity.

Results

Overview

Within the selected data from the female Tinder video (n=171), 75 (44%) commenters were male, 37 (22%) were female, and 59 (35%) were indeterminable. For the male Tinder data (n=145), 45 (31%) commenters were male, 52 (36%) were female, and 48 (33%) interminable. The descriptive figures of the 2 Tinder videos are reported in Table 1.

Table 1. Overview of descriptive results (those interminable are excluded in the analysis and this table).

Gender of the actor in the video	Purpose		Target gender		Commenter gender	
	Attack	Defense	Male	Female	Male	Female
Male video, n (%)	71 (49)	43 (30)	73 (50)	33 (23)	45 (31)	52 (36)
Female video, n (%)	103 (60)	46 (27)	7 (4)	98 (57)	75 (44)	37 (22)
Total, n	174	89	80	131	120	89

In terms of analysis, all the research questions focus on the difference in frequency of occurrence (eg, number of comments that verbally attack). Thus, to analyze this dataset, chi-square analyses were employed, with the exception of intensity of derogation, for which we employed a *t* test that examined the mean differences.

Evolutionary Mate Values—RQ1a and RQ1b

Our first research question asks up to what extent overweight women are verbally attacked according to reproductive power more than men. In addition, we ask up to what extent overweight men are attacked according to their capacities for resource acquisition more than women. We found no significant differences in the frequency of comments attacking physical appearance and resource acquisition between comments targeting male and female overweight individuals, χ^2_1(N=66)=2.8, *P*=.09. The majority of comments targeting overweight males attacked their reproductive power (14/19, 74%), while females were targeted for both resource acquisition (24/47, 51%) and reproductive power (23/47, 49%). These results indicate that when addressing evolutionary mate values, it is primarily reproductive power (ie, physical attractiveness) for which both genders are verbally attacked.

Heterosocial Skills—RQ2

An exploration of our second research question requires an examination of the comparative extent up to which overweight males and females receive verbal attacks addressing heterosocial skills. A significant difference was found between male and female overweight targets, χ^2_1(N=138)=9.9, *P*=.002. Our results indicate that male targets are more likely to receive comments that attack their heterosocial or interpersonal skills. Of male targets, 87% (40/46) received comments about relationship deficits. Of female targets, 61% (56/92) received comments of this type.

Topic of Verbal Attack Comparison—RQ3

To explore our third research question, we conducted a frequency comparison between attacks on capacities for resource acquisition, reproductive power, and heterosocial skills between male and female individuals. Initially, evolutionary mate values (ie, personal failings, which include resource acquisition and reproductive power) and heterosocial skills were coded separately, and therefore overlap between these 2 variables was allowed. That is, both evolutionary mate values and heterosocial skills could be expressed in a single comment. We combined evolutionary mate values and heterosocial skills into a single variable, Topic of Comment, to determine whether gender differences exist between the prominence of attacks on reproductive power, capacities for resource acquisition, and heterosocial skills.

Our results show that there exists a significant difference in the topic of aggressive comments between target gender, χ^2_2(N=80)=7.6, *P*=.02. On the basis of the standardized residuals of each comparison, we found that capacities for resource acquisition (residual=−4.8) and heterosocial skills (residual=4.6) are the two sources of significance, that is, female overweight targets are more likely to receive comments that attack capacities for resource acquisition (14/51, 28%) than male overweight targets (1/29, 3%). However, male overweight targets are more likely to receive comments that attack their heterosocial skills (24/29, 83%) than female overweight targets (29/51, 57%). Reproductive power in this case was not a source of difference between male and female targets (residual=−.2). When defending, there was no significant difference in the topic of comments between male and female overweight targets, χ^2_2(N=49)=1.1, *P*=.56.

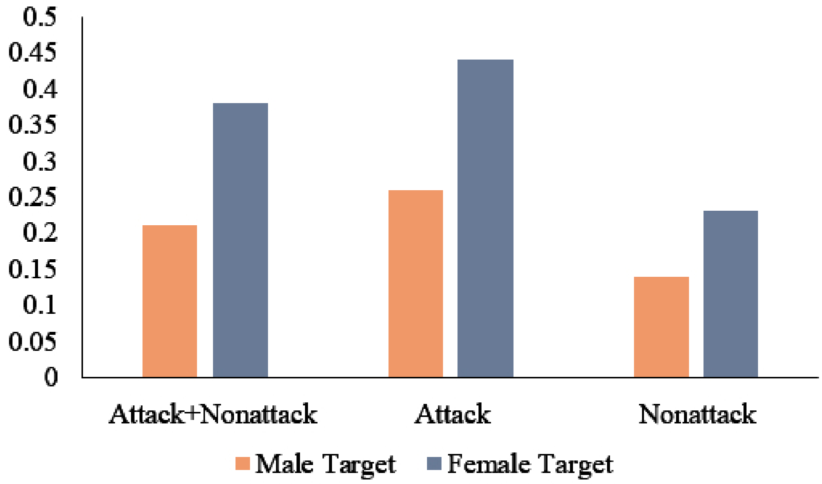
Target Gender—RQ4a and RQ4b

We asked which gender receives more verbal attacks within the 2 videos. No significant difference in the number of verbal attacks was found between the gender of targets, χ^2_1(N=211)=3.6, *P*=.06. To answer RQ4a and RQ4b, we explored which commenter gender tends to attack male or female overweight individuals. We found a significant difference between male and female commenters in attacking female overweight targets, χ^2_1(N=87)=5.7, *P*=.02 Male commenters are more likely to attack female overweight targets (42/52, 81%) than female commenters (20/35, 57%). Yet, when male overweight individuals are targeted, no difference in frequency of verbal attack was found between the commenter genders, χ^2_1(N=50)=1.2, *P*=.28. No significant difference between the commenter gender in the frequency of defense was found.

Intensity of Derogation—RQ5a, RQ5b and RQ5c

We asked how pervasive the verbal attacks against overweight individuals are. Our data show that the majority of comments verbally attacked overweight people (56%, 174/316). Only 28% percent of the comments (89/316) defended the overweight individuals. Our analysis of RQ5b assesses to what extent verbal attacks toward overweight individuals are accompanied by swear words. When attacking (mean 0.37, SD 0.74), comment intensity (ie, number of swear words) is significantly higher than nonattack comments (mean 0.17, SD 0.40), t_{314}=2.48, *P*=.01. RQ5c asks up to what comparative extent male and female overweight individuals receive comments that include swear words. A significant difference was found between comments that target overweight male individuals (mean 0.21, SD 0.50) and those targeting females (mean 0.38, SD 0.70), $t_{204.97}$=2.040, *P*=.04 (Figure 1). Also, when targeting overweight females, attacking comments accompanied more swear words (mean 0.44, SD 0.77) than nonattack comments (mean 0.23, SD 0.48), $t_{204.97}$=−2.040, *P*=.05. However, no difference between the attack and nonattack comments was found for those targeting males.

Figure 1. Average number of swear words (intensity of derogation). The y-axis indicates the number of swear words used in the comments. "Attack+Nonattack" refers to comments regardless of its valence. Attack and nonattack refer to the attack and nonattack comments, respectively.



Discussion

Results and Explanations

The primary purpose of this study was to understand how and up to what extent overweight individuals are verbally attacked online. Particularly, this study delved into patterned differences in verbal attacks depending on the genders of the commenters and the overweight target, especially in terms of (1) the topic of the comment, (2) the intensity of derogation (ie, swearing), and (3) the targeted genders. These results are important for the development of intervention strategies that can effectively reduce the social stigma of obesity, which can adversely influence overweight individuals' mental and physical health [4]. In our data, when addressing evolutionary mate values, both overweight men and women were verbally attacked for deficits in reproductive power, primarily signaled by physical attractiveness. This result is seemingly unexpected according to the evolutionary psychological view, which suggests that men will be attacked for deficits in men's primary mate features: capacities for resource acquisitions (eg, financial prospect and industriousness). However, one may need to apply the evolutionary psychological view with caution. Buss (1989) argued that mate feature preferences for men and women vary depending on time, culture, and physical environment [21]. Historically men's physical attractiveness may not have signaled economic or social status or ability to accumulate resources (eg, money) to raise offspring and protect a family. In contemporary society, however, men's physical attractiveness may be an important visual cue. Physical attractiveness may signal an ability to get resources for supporting and protecting a family, as a good-looking man is more likely to be hired, make more sales, and receive raises [50]. Our results show this change in how men's mate features are preferred. Another possible explanation for this finding may be that the comments were posted in response to videos about meeting a date through Tinder, a mobile app where men and women may look for a short-term relationship. In this instance, men's resource acquisition may not be as highly valued since females may not be looking for a long-term relationship (in which resource acquisition would be more useful).

While there were no significant gender differences in the dichotomous mate values, our data do show gender discrepancies in the prominence of the topic of the verbal attacks. First, we explored up to what extent weight stigma is perpetuated within the comments on the 2 Tinder videos. Our data show that the majority of the comments attack overweight individuals rather than defend them. This result generally supports the view that *YouTube* is a space where people feel disinhibited to express hatred against others [38]. Given that the 2 Tinder videos do not explicitly ridicule or tease obese individuals, our result also confirms that even in cases of indirect address of obesity, commenters choose to verbally attack obese individuals. Some may argue that people, especially bored young people, enjoy verbally attacking targets with hostile language (ie, flaming) [38]. Some users may even think that flaming is a "funny way of interacting that is not to be taken too seriously or that it is a necessary side effect of vivid debate and freedom of speech" [38]. Importantly, people are likely unaware of what their targets—overweight individuals—feel about their aggressive comments, and the possible subsequent effects induced by such stigmatizing comments. This may imply that to stop anonymous aggressive comments, we need stricter self-regulation of websites. For example, Reddit, a popular social network website, has already started stricter regulation, closing hatred forums within its website including "Fat People Hate," the biggest forum in Reddit at the time [51]. Of course, this restriction will likely prompt debates on freedom of expression. Nevertheless, our results, along with findings from other studies, elucidate a worrying situation in which the online world is filled with unfettered stigmatizing messages against overweight individuals.

Second, overweight men are more likely to be attacked for deficits in heterosocial skills than overweight women. For men, heterosocial skills may be seen as an important feature for success in mating [27], but overweight men are stereotyped as lacking in heterosocial skills just as other negative traits unjustly associated with overweight individuals. This stereotypical assumption may result in more verbal attacks targeting heterosocial skills of overweight men than that of overweight women. In addition, perhaps unexpectedly, overweight women were more likely to receive verbal attacks addressing capacities for resource acquisition (eg, being lazy, unhealthy) than men. This is also seemingly at odds with the evolutionary psychological perspective, in which men evolved to value female mates' reproductive power. One possible explanation can be that contemporary overweight women are blamed more for failures in individual responsibility or ability to control or discipline her body than overweight men, that is, overweight women may be perceived to be less disciplined and weaker than overweight men. Some research supports this view suggesting that overweight women are seen to lack in ability to do a strenuous job, whereas overweight men are seen as capable as nonoverweight men [52]. However, this argument requires further validation based on more empirical research, including more representative samples of comments from *YouTube* or other digital platforms. Also, a survey with a representative sample of the population may be needed to examine changes in attitudes about evolutionary mate values for overweight men and women.

The third gender difference examined in this paper is the susceptibility to weight stigma. Is weight stigma, as many studies suggest, primarily a women's issue even in a digital context? Our data suggests that overweight women are prevalently victimized by men. More specifically, men are more likely than women to attack overweight women, and women do not show any tendency of attacking overweight men more than women. It is possible that men may blame obese women for not conforming to their role of "being an attractive object" in a patriarchal society [32]. Also, we found that overweight women receive significantly more swear words within comments than overweight men. This is important as swearing can indicate anger, arousal, or intention to strengthen argument [53,54], and if being overweight for women is less "forgivable" than for men, then women would be attacked with higher levels of intensity, indicated by a higher number of swear words.

Limitations, Implications, and Future Directions

This study has 3 limitations. First, our study is exploratory, focusing on only 2 sets of comments posted in response to the 2 Tinder date videos. Our findings cannot be generalized to all *YouTube* videos, which may differ in terms of their content (eg, narrative structure). Therefore, this study should be expanded into comments on other videos and social media platforms. Also, a larger dataset (ie, big data) from diverse sources can be used to further delve into the gender differences in verbal aggression against overweight individuals, increasing generalizability of the findings [55].

Second, this study sampled only root comments—comments that are posted in direct response to the video content and first response comments to the root comments. Therefore, our sampling methodology may exclude second, third, or later response comments that may provide further counterarguments against comments that attack overweight individuals. To understand how commenters may defend or empathize with overweight men and women, future research should include more response comments as well.

Third, researchers should seek to identify how overweight individuals respond to verbal attacks on *YouTube*. This research question would be difficult to answer through content analysis, as commenters rarely reveal their weight within this context. Survey or focus group interviews which ask questions about overweight individuals' experience of verbal attacks on *YouTube* may be useful in exploring this question further. Also, our research focused on 2 binary genders (ie, male vs female). Thus, the impacts of nonconforming and marginalized gender identities of commenters or overweight targets (eg, transgender or nonbinary) await future research.

Despite these limitations, this study contributes to current research on weight-stigma by showing how weight-stigma is perpetuated *differently* for overweight men and women in a digital context. Theoretically, this study attempted to expand evolutionary psychological theories by examining gender differences within the context of weight stigma. Our results show that the evolutionary perspective on verbal aggression needs to be applied cautiously with overweight individuals in contemporary society.

Findings of this study should provide useful input in designing effective gender-tailored messages—customizing a message to different perceptions of overweight men and women—to maximize resources available to improve attitude toward obese individuals [56], and, in the long term, reduce depression caused by weight stigma in a digital context. For example, the message for overweight men should focus on the perception that overweight men lack in heterosocial skills by showing counterstereotypical images of overweight men (eg, overweight men being active and good at socializing with others). Conversely, messages for overweight women could better tackle the perception that overweight women lack in will power (eg, overweight women depicted as active in sports such as marathon that require greater will-power) rather than the negative traits primarily associated with overweight males. Also, based on findings of this study, we suggest that further research begin addressing gender discrepancies in weight stigma. For example, the widely used scale for weight bias, Fat Phobia Scale, was originally developed based on adjectives to describe overweight individuals, listed by people who entered a motor vehicle license bureau in a suburb of St. Paul/Minneapolis metropolitan area in 1984 [57]. This study identified 12 out of the total 14 items in this scale to be traits generally related to capacities for resource acquisition; only 2 items of this semantic differential scale (Item #3: "Unattractive vs Attractive"; Item#11: "Shapeless vs Shapely") were related to reproductive power. Thus, this scale may need revision to incorporate a more appropriate balance of negative traits associated with reproductive power and capacities for resource acquisition. Additionally, more explicit items measuring overweight individuals' heterosocial skills may need to be included as well.

In other words, more gender-balanced weight bias scales are needed to reflect the changes in gender differences in weight stigma.

In conclusion, the results of this study were indispensable in that they revealed how implicit beliefs stigmatize overweight men and women differently. Thus, based on the data, it is recommended that effective intervention strategies that tackle these specific beliefs and gender be developed to lift the mental burden for overweight individuals in fighting obesity.

Conflicts of Interest

None declared.

References

1. Brewis AA, Wutich A, Falletta-Cowden A, Rodriguez-Soto I. Body norms and fat stigma in global perspective. Curr Anthropol 2011 Apr;52(2):269-276 [FREE Full text] [doi: 10.1086/659309]
2. Chou WS, Prestin A, Kunath S. Obesity in social media: a mixed methods analysis. Transl Behav Med 2014 Sep;4(3):314-323 [FREE Full text] [doi: 10.1007/s13142-014-0256-1] [Medline: 25264470]
3. Tomiyama AJ. Weight stigma is stressful. A review of evidence for the Cyclic Obesity/Weight-Based Stigma model. Appetite 2014 Nov;82:8-15. [doi: 10.1016/j.appet.2014.06.108] [Medline: 24997407]
4. Carpenter KM, Hasin DS, Allison DB, Faith MS. Relationships between obesity and DSM-IV major depressive disorder, suicide ideation, and suicide attempts: results from a general population study. Am J Public Health 2000 Feb;90(2):251-257. [Medline: 10667187]
5. Himmelstein MS, Incollingo BA, Tomiyama AJ. The weight of stigma: cortisol reactivity to manipulated weight stigma. Obesity (Silver Spring) 2015 Feb;23(2):368-374 [FREE Full text] [doi: 10.1002/oby.20959] [Medline: 25522347]
6. Jackson SE, Beeken RJ, Wardle J. Obesity, perceived weight discrimination, and psychological well-being in older adults in England. Obesity (Silver Spring) 2015 May;23(5):1105-1111 [FREE Full text] [doi: 10.1002/oby.21052] [Medline: 25809860]
7. Kreuter MW, Wray RJ. Tailored and targeted health communication: strategies for enhancing information relevance. Am J Health Behav 2003;27(Suppl 3):S227-S232. [Medline: 14672383]
8. Lang A. Using the limited capacity model of motivated mediated message processing to design effective cancer communication messages. J Communication 2006 Aug;56(s1):S57-S80 [FREE Full text] [doi: 10.1111/j.1460-2466.2006.00283.x]
9. Latner JD, Stunkard AJ, Wilson GT. Stigmatized students: age, sex, and ethnicity effects in the stigmatization of obesity. Obes Res 2005 Jul;13(7):1226-1231 [FREE Full text] [doi: 10.1038/oby.2005.145] [Medline: 16076992]
10. Jensen JD, King AJ, Carcioppolo N, Davis L. Why are tailored messages more effective? A multiple mediation analysis of a breast cancer screening intervention. J Commun 2012 Aug 27;62(5):851-868. [doi: 10.1111/j.1460-2466.2012.01668.x]
11. Buss DM, Dedden LA. Derogation of competitors. J Soc Pers Relat 1990;7(3):395-422. [doi: 10.1177/0265407590073006]
12. Becker A. Body, self, and society: The view from Fiji. Philadelphia, PA: University of Pennsylvania Press; 1995.
13. Stunkard A, Sobal J. Psychosocial consequences of obesity. In: Brownell K, Fairburn C, editors. Eating disorders and obesity: A comprehensive handbook. New York, NY: The Guilford Press; 1995:417-421.
14. Thomas SL, Hyde J, Karunaratne A, Kausman R, Komesaroff PA. “They all work...when you stick to them”: a qualitative investigation of dieting, weight loss, and physical exercise, in obese individuals. Nutr J 2008 Nov 24;7:34 [FREE Full text] [doi: 10.1186/1475-2891-7-34] [Medline: 19025661]
15. Winterman D. BBC. 2009 Oct 29. Why are fat people abused? BBC News Magazine URL: http://news.bbc.co.uk/2/hi/8327753.stm [accessed 2018-03-16] [WebCite Cache ID 6xxPfTKQR]
16. Chen Y, Wang Q, Xie J. Online social interactions: a natural experiment on word of mouth versus observational learning. J Mark Res 2011;48(2):238-254. [doi: 10.1509/jmkr.48.2.238]
17. Barber N. The evolutionary psychology of physical attractiveness: sexual selection and human morphology. Ethol Sociobiol 1995 Sep;16(5):395-424 [FREE Full text] [doi: 10.1016/0162-3095(95)00068-2]
18. Betzig L. Redistribution: Equity or exploitation. In: Betzig L, Monique M, Turke P, editors. Human reproductive behavior: A Darwinian perspective. Cambridge, CA: Cambridge University Press; 1988:49-63.
19. Ellis B. The evolution of sexual attraction: Evaluative mechanisms in women. In: Barkow BH, Cismides L, Tooby J, editors. The Adapted Mind: Evolutionary psychology and the generation of culture. New York, NY: Oxford University Press; 1995:267-288.
20. Renzetti C, Curran D, Maier S. Women, men, and society. London: Pearson; 1992.
21. Buss DM. Sex differences in human mate preferences: evolutionary hypotheses tested in 37 cultures. Behav Brain Sci 2010 Feb 4;12(01):1 [FREE Full text] [doi: 10.1017/S0140525X00023992]
22. Boxer CF, Noonan MC, Whelan CB. Measuring mate preferences: a replication and extension. J Fam Issues 2013;36(2):163-187. [doi: 10.1177/0192513X13490404]

23. Gangestad SW, Simpson JA. The evolution of human mating: trade-offs and strategic pluralism. Behav Brain Sci 2000 Aug;23(4):573-87; discussion 587. [Medline: 11301543]
24. Li NP, Bailey JM, Kenrick DT, Linsenmeier JA. The necessities and luxuries of mate preferences: testing the tradeoffs. J Pers Soc Psychol 2002 Jun;82(6):947-955. [Medline: 12051582]
25. Hussin M, Frazier S, Thompson JK. Fat stigmatization on YouTube: a content analysis. Body Image 2011 Jan;8(1):90-92. [doi: 10.1016/j.bodyim.2010.10.003] [Medline: 21126932]
26. Yoo JH, Kim J. Obesity in the new media: a content analysis of obesity videos on YouTube. Health Commun 2012 Jan;27(1):86-97. [doi: 10.1080/10410236.2011.569003] [Medline: 21809934]
27. Barlow DH, Abel GG, Blanchard EB, Bristow AR, Young LD. A heterosocial skills behavior checklist for males. Behav Ther 1977 Mar;8(2):229-239 [FREE Full text] [doi: 10.1016/S0005-7894(77)80272-4]
28. Twentyman CT, McFall RM. Behavioral training of social skills in shy males. J Consult Clin Psychol 1975 Jun;43(3):384-395. [Medline: 1159128]
29. Pescosolido BA, Martin JK, Lang A, Olafsdottir S. Rethinking theoretical approaches to stigma: a Framework Integrating Normative Influences on Stigma (FINIS). Soc Sci Med 2008 Aug;67(3):431-440 [FREE Full text] [doi: 10.1016/j.socscimed.2008.03.018] [Medline: 18436358]
30. Chen EY, Brown M. Obesity stigma in sexual relationships. Obes Res 2005 Aug;13(8):1393-1397 [FREE Full text] [doi: 10.1038/oby.2005.168] [Medline: 16129721]
31. Schur E. Labeling women deviant: Gender, stigma, and social control. New York, NY: McGraw-Hill Companies; 1984.
32. Prohaska A, Gailey JA. Fat women as “easy targets”: achieving masculinity through hogging. In: The fat studies reader. New York, New York: NYU Press; 2009:158-166.
33. Chiou W. Adolescents' sexual self-disclosure on the internet: deindividuation and impression management. Adolescence 2006;41(163):547-561. [Medline: 17225667]
34. Joinson A. Internet behaviour and the design of virtual methods. In: Virtual methods: Issues in social research on the internet. New York, NY: Berg; 2005:21-34.
35. Lapidot-Lefler N, Barak A. Effects of anonymity, invisibility, and lack of eye-contact on toxic online disinhibition. Comput Human Behav 2012;28(2):434-443. [doi: 10.1016/j.chb.2011.10.014]
36. Suler J. The online disinhibition effect. Cyberpsychol Behav 2004 Jun;7(3):321-326. [doi: 10.1089/1094931041291295] [Medline: 15257832]
37. Thelwall M. Fk yea I swear: cursing and gender in MySpace. Corpora 2008;3(1):83-107. [doi: 10.3366/E1749503208000087]
38. Moor PJ, Heuvelman A, Verleur R. Flaming on youtube. Comput Human Behav 2010;26(6):1536-1546. [doi: 10.1016/j.chb.2010.05.023]
39. Lange P. Publicly private and privately publicocial networking on YouTube. J Comput Mediat Commun 2007;13(1):361-380. [doi: 10.1111/j.1083-6101.2007.00400.x]
40. YouTube. 2014. Fat Girl Tinder Date (Social Experiment) [Video] URL: https://www.youtube.com/watch?v=2alnVIj1Jf8 [accessed 2017-10-14] [WebCite Cache ID 6uCVB8wij]
41. YouTube. Fat Guy Tinder Data (Social Experiment) [Video] URL: https://www.youtube.com/watch?v=iUy3_kBme4M [accessed 2017-10-14] [WebCite Cache ID 6uCUise57]
42. Dutton DG, Hemphill KJ. Patterns of socially desirable responding among perpetrators and victims of wife assault. Violence Vict 1992;7(1):29-39. [Medline: 1504031]
43. Shadish W, Cook T, Campbell D. Experimental and quasi-experimental designs for generalized causal inference. Round Rock, TX: Houghton Mifflin; 2002.
44. Claussen J, Kretschmer T, Mayrhofer P. The effects of rewarding user engagement: the case of facebook apps. Inf Syst Res 2013 Mar;24(1):186-200 [FREE Full text] [doi: 10.1287/isre.1120.0467]
45. Krippendorff K. Reliability in content analysis. Hum Commun Res 2004;30(3):411-433. [doi: 10.1111/j.1468-2958.2004.tb00738.x]
46. Bullingham L, Vasconcelos AC. 'The presentation of self in the online world': Goffman and the study of online identities. J Inf Sci 2013 Jan 04;39(1):101-112. [doi: 10.1177/0165551512470051]
47. Sussman N, Tyson D. Sex and power: gender differences in computer-mediated interactions. Comput Human Behav 2000 Jul;16(4):381-394 [FREE Full text] [doi: 10.1016/S0747-5632(00)00020-0]
48. Vismara MF, Toaff J, Pulvirenti G, Settanni C, Colao E, Lavano SM, et al. Internet use and access, behavior, cyberbullying, and grooming: results of an investigative whole city survey of adolescents. Interact J Med Res 2017 Aug 29;6(2):e9 [FREE Full text] [doi: 10.2196/ijmr.6231] [Medline: 28851675]
49. SSA. Popular Baby Names: Beyond the Top 1000 Names URL: https://www.ssa.gov/oact/babynames/limits.html [accessed 2018-01-17] [WebCite Cache ID 6wY63PqWK]
50. Shellenbarger S. freerepublic. 2011 Oct 07. On the Job, Beauty Is More Than Skin-Deep URL: http://www.freerepublic.com/focus/news/2798513/posts [accessed 2018-03-16] [WebCite Cache ID 6xxRJ74sj]
51. Dewey C. Washingtonpost. 2015 Jun 10. These are the 5 subreddits Reddit banned under its game-changing anti-harassment policy and why it banned them URL: https://www.washingtonpost.com/ [accessed 2018-03-16] [WebCite Cache ID 6xxRNXZXK]

52. Miller BJ, Lundgren JD. An experimental study of the role of weight bias in candidate evaluation. Obesity (Silver Spring) 2010 Apr;18(4):712-718 [FREE Full text] [doi: 10.1038/oby.2009.492] [Medline: 20111025]
53. Jay T, Janschewitz K. The pragmatics of swearing. JPR 2008;4:267-288. [doi: 10.1515/JPLR.2008.013]
54. Rassin E, Muris P. Why do women swear? An exploration of reasons for and perceived efficacy of swearing in Dutch female students. Pers Individ Dif 2005 May;38(7):1669-1674 [FREE Full text] [doi: 10.1016/j.paid.2004.09.022]
55. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. Health Inf Sci Syst 2014;2:3 [FREE Full text] [doi: 10.1186/2047-2501-2-3] [Medline: 25825667]
56. Schmid KL, Rivers SE, Latimer AE, Salovey P. Targeting or tailoring? Maximizing resources to create effective health communications. Mark Health Serv 2008;28(1):32-37.
57. Robinson BE, Bacon JG, O'Reilly J. Fat phobia: measuring, understanding, and changing anti-fat attitudes. Int J Eat Disord 1993 Dec;14(4):467-480. [Medline: 8293029]

Abbreviations

RQ: research question

Edited by G Eysenbach; submitted 14.10.17; peer-reviewed by N Bragazzi, R Beeken, Y Mejova, R Ashford; comments to author 23.11.17; revised version received 18.01.18; accepted 05.02.18; published 20.03.18.

Please cite as:
Jeon YA, Hale B, Knackmuhs E, Mackert M
Weight Stigma Goes Viral on the Internet: Systematic Assessment of YouTube Comments Attacking Overweight Men and Women
Interact J Med Res 2018;7(1):e6
URL: http://www.i-jmr.org/2018/1/e6/
doi:10.2196/ijmr.9182
PMID:29559426

Original Paper

Quality of Web Information About Palliative Care on Websites from the United States and Japan: Comparative Evaluation Study

Kouichi Tanabe[1], RPh, PhD; Kaho Fujiwara[1], BS; Hana Ogura[1], BS; Hatsuna Yasuda[2], BA; Nobuyuki Goto[3], RPh, PhD; Fumiko Ohtsu[1], RPh, PhD

[1]Drug Informatics, Faculty of Pharmacy, Meijo University, Nagoya, Japan

[2]Modern Languages, Southampton University, Kanagawa, Japan

[3]Department of Pharmacy, University of Fukui Hospital, Fukui, Japan

Corresponding Author:
Fumiko Ohtsu, RPh, PhD
Drug Informatics
Faculty of Pharmacy
Meijo University
150 Yagotoyama, Tempaku-ku
Nagoya,
Japan
Phone: 81 52 833 1063
Email: fohtsu@meijo-u.ac.jp

Abstract

Background: Patients and their families are able to obtain information about palliative care from websites easily nowadays. However, there are concerns on the accuracy of information on the Web and how up to date it is.

Objective: The objective of this study was to elucidate problematic points of medical information about palliative care obtained from websites, and to compare the quality of the information between Japanese and US websites.

Methods: We searched Google Japan and Google USA for websites relating to palliative care. We then evaluated the top 50 websites from each search using the DISCERN and LIDA instruments.

Results: We found that Japanese websites were given a lower evaluation of reliability than US websites. In 3 LIDA instrument subcategories—engagability ($P<.001$), currency ($P=.001$), and content production procedure ($P<.001$)—US websites scored significantly higher and had large effect sizes.

Conclusions: Our results suggest that Japanese websites have problems with the frequency with which they are updated, their update procedures and policies, and the scrutiny process the evidence must undergo. Additionally, there was a weak association between search ranking and reliability, and simultaneously we found that reliability could not be assessed by search ranking alone.

(Interact J Med Res 2018;7(1):e7) doi:10.2196/ijmr.9574

KEYWORDS

internet; website; reliability; quality; usefulness; palliative care; consumer health information; evaluation studies

Introduction

The World Health Organization defines palliative care as "an approach that improves the quality of life of patients and their families facing the problems associated with life-threatening illness, through the prevention and relief of suffering by means of early identification and impeccable assessment and treatment of pain and other problems, physical, psychosocial and spiritual" [1]. The incorporation of palliative care into cancer treatment from the early stages also leads to the necessity of enhancing patients' and their families' understanding of such care [2].

On the other hand, with the spread of the internet, it has become possible for patients and their families to easily collect treatment-related information through websites. The merit of enabling users to immediately obtain extensive information has made the internet an important means to collect medical information [3]. According to the National Telecommunication Survey 2015 in Japan, internet use through mobile phones and tablet terminals is increasing in all age groups, and the increase in the rate of such use is particularly marked among those aged 60 to 79 years [4]. The rise in the number of internet users suggests that an increasing number of patients are collecting medical information using this method.

However, information available on the internet also has drawbacks, such as not necessarily being understandable for patients without expertise and having low updating frequencies [5,6]. Furthermore, it has been reported that approximately 30% of those who browse webpages containing information regarding palliative care are unsatisfied due to insufficient information and detailed contents, indicating the necessity of improving the quality of such information as a challenge [2,7].

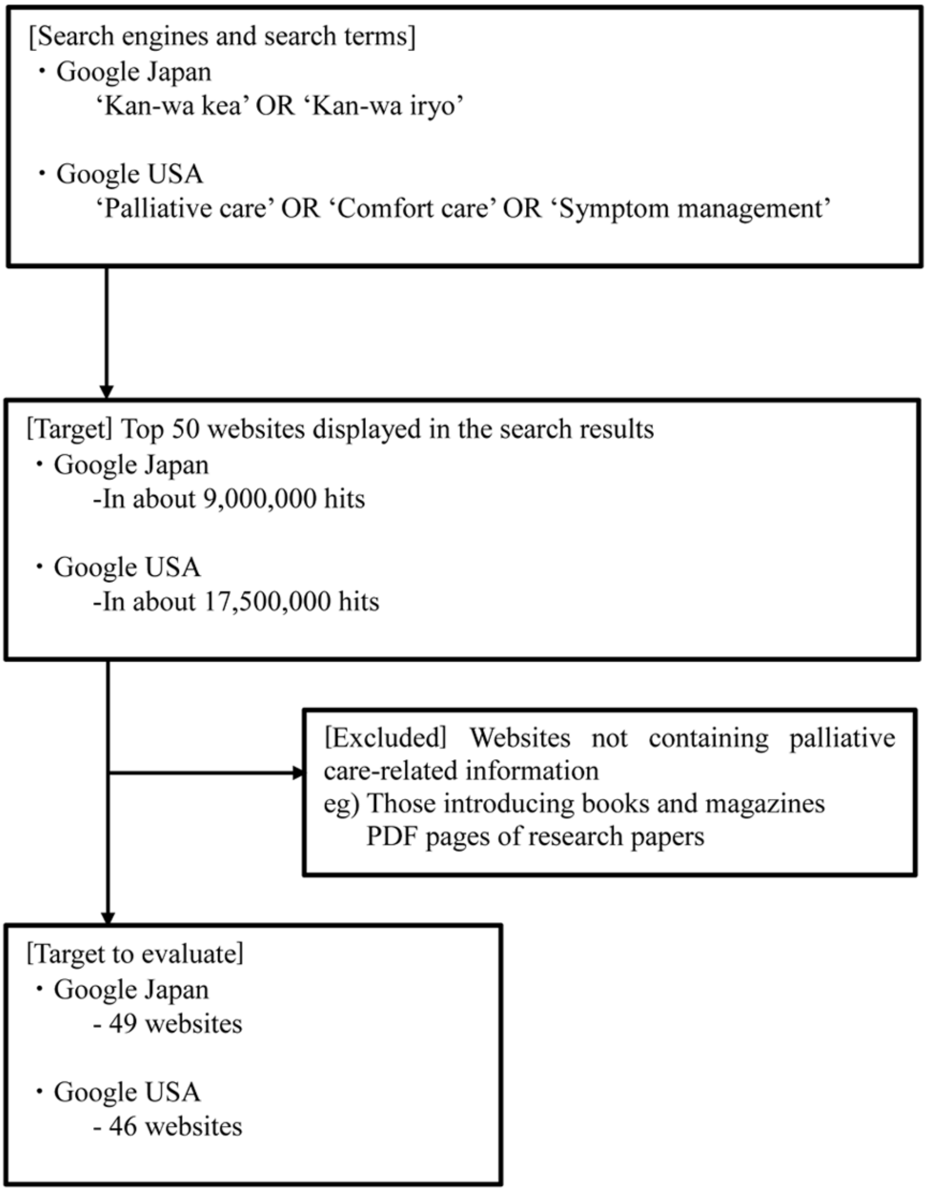
Therefore, in this study we focused on palliative care–related information available through websites and compared the quality of the information between Japan and the United States in order to identify problems related to medical information available through websites on the internet.

Methods

Website Selection

Figure 1 outlines the process of selecting target websites. We searched for websites using 2 search engines: Google Japan and Google USA (Google LLC). When conducting Web searches, we turned off both the positional information and history information reference functions of the personal computer. The search terms were (palliative care OR palliative medicine [in Japanese]) on Google Japan and (palliative care OR comfort care OR symptom management) on Google USA. We targeted the top 50 websites displayed in the search results. When selecting sites, we regarded webpages having the same domain as belonging to the same website. We excluded websites not containing palliative care–related information, such as those introducing books and magazines, and with PDF pages of research papers. We ascertained that the top 50 websites found on Google USA were US sites.

Figure 1. The process of selecting target websites.



Evaluation

We evaluated the information available on the target websites by examining the owner information included in each website and by using 2 evaluation indices: the DISCERN (Textbox 1) and LIDA instruments (Textbox 2).

The DISCERN instrument [8] is an index to evaluate the quality of information regarding treatment choices, developed by Shepperd and colleagues at Oxford University, Oxford, UK. We evaluated the reliability of information (reliability-1) and the quality of information regarding treatment choices (information quality) on a 5-point scale from 1 (no) to 5 (yes). Subsequently, we evaluated overall aspects on a 5-point scale from 1 (serious or extensive shortcomings) to 5 (minimal shortcomings).

The LIDA instrument is an index to evaluate health-related information available on websites, created by Minervation, a health care consultancy [9]. It consists of 3 domains: accessibility, usefulness, and reliability (reliability-2). In this study, we rated the evaluation items on a 4-point scale from 0 (never) to 3 (always). Among the 3 domains, we evaluated only usefulness and reliability-2, as accessibility is not closely associated with reliability.

The evaluation was conducted by multiple raters independently. When interrater agreement was poor, we discussed the results with a new supervisor and used the agreed-upon results.

We divided owner information into 2 categories for evaluation: personal and organizational. We further divided organizational into medical, administrative (public), general (profit), and general (nonprofit) for comparison between Japan and the United States. We defined each category as outlined in Textbox 3.

We divided the categories of the website contents into dedicated palliative care sites, general information sites about cancer (including palliative care–related information), general medical sites (including palliative care–related information), and nonmedical sites (including palliative care–related information), and we assessed those. Textbox 4 outlines the definitions of the website content categories.

To evaluate their reliability, we also divided the target websites into 2 groups—high rank and low rank—based on their ranks when displayed in the search results. We then compared their scores for each item of the 2 evaluation indices. With regard to the reliability-related items (reliability-1 and reliability-2) of both evaluation indices, we examined the correlation between the scores for these items and ranks.

Textbox 1. DISCERN instrument for judging the quality of written consumer health information. Section and questions are presented below. Scored as 1=no, 3=partially, 5=yes.

Reliability-1: Is the publication reliable?

- Are the aims clear?
- Does it achieve its aims?
- Is it relevant?
- Is it clear what sources of information were used to compile the publication (other than the author or producer)?
- Is it clear when the information used or reported in the publication was produced?

Information quality: How good is the quality of information treatment choices?

- Is it balanced and unbiased?
- Does it provide details of additional sources of support and information?
- Does it refer to areas of uncertainty?
- Does it describe how each treatment works?
- Does it describe the benefits of each treatment?
- Does it describe the risks of each treatment?
- Does it describe what would happen if no treatment is used?
- Does it describe how the treatment choices affect overall quality of life?
- Is it clear that there may be more than one possible treatment choice?
- Does it provide support for shared decision making?

Overall evaluation: Overall rating of the publication

- Based on the answers to all of the above questions, rate the overall quality of the publication as a source of information about treatment choices

Textbox 2. LIDA instrument for validation of health care websites. Level and questions are presented below. Scored from 0=never to 3=always.

Usefulness/usability

Clarity

- Is there a clear statement of who this website is for?
- Is the level of detail appropriate to their level of knowledge?
- Is the layout of the main block of information clear and readable?
- Is the navigation clear and well structured?
- Can you always tell your current location in the site?
- Is the color scheme appropriate and engaging?

Consistency

- Is the same page layout used throughout the site?
- Do navigational links have a consistent function?
- Is the site structure (categories or organization of pages) applied consistently?

Functionality

- Does the site provide an effective search facility?
- Does the site provide effective browsing facilities?
- Does the design minimize the cognitive overhead of using the site?
- Does the site support the normal browser navigational tools?
- Can you use the site without third-party plug-ins?

Engagability

- Can the user make an effective judgment of whether the site applies to them?
- Is the website interactive?
- Can the user personalize their experience of using the site?
- Does the website integrate nontextual media?

Reliability-2

Currency

- Does the site respond to recent events?
- Can users submit comments on specific content?
- Is site content updated at an appropriate interval?

Conflicts of interest

- Is it clear who runs the site?
- Is it clear who pays for the site?
- Is there a declaration of the objectives of the people who run the site?

Content production

- Does the site report a clear content production method?
- Is this a robust method?
- Can the information be checked from original sources?

Content production procedure

- Are the audience needs identified in advance?
- Is comprehensive literature searching conducted?
- Are retrieved documents critically appraised?

- Is content authored by subject experts?
- Is content reviewed by an independent expert or panel?

Output of content

- Has literature searching found the right information?
- Does the content check out?
- Is the content accurate?

Textbox 3. Definitions of owner information categories.

- Personal: personal websites
- Organizational: websites of organizations, such as companies and facilities
- Medical: websites of medical institutions or corporations, such as hospitals and care facilities
- General (profit): websites of for-profit organizations, such as joint-stock companies, excluding medical institutions
- General (nonprofit): websites of nonprofit or nongovernmental organizations, excluding medical institutions
- Administrative (public): websites of prefectures, states, ministries, and agencies
- Academic: websites of academic societies

Textbox 4. Definitions of website content categories.

- Dedicated palliative care sites: websites from medical conferences that are dedicated to palliative care, educational and informational sites about palliative care, hospice service program informational sites, and academic journal sites about palliative care
- General information sites about cancer: educational and informational sites about cancer and informational sites about hospitals specializing in cancer treatment
- General medical sites: informational sites about home visits and hospitals
- Nonmedical sites: featured articles from newspapers, nonmedical informational sites, and online dictionaries

Analysis

We conducted a chi-square test to compare owner information. For comparisons of reliability, quality, and usefulness of the information available on the target websites between Japan and the United States, we calculated the mean scores for the 2 evaluation indices. The *t* test was conducted (with the significance level set at .05%, and using Cohen *d* for effect size calculation) using the statistical software IBM SPSS Statistics 24 (IBM Corporation). We applied Cohen's standard setting to measure effect size, where no effect size is $0 \leqq d < 0.2$, small effect size is $0.2 \leqq d < 0.4$, medium effect size is $0.4 \leqq d < 0.6$, and large effect size is $0.6 \leqq d$. Furthermore, we examined the correlation between each score and ranks by calculating the Pearson product-moment correlation coefficient and conducting a test to confirm the absence of a correlation.

Results

Overview

We identified approximately 9,000,000 websites using Google Japan (date of search: December 13, 2016) and 17,500,000 websites using Google USA (December 14, 2016). After excluding 1 website from the Google Japan search and 4 from the Google USA search, we targeted 49 Japanese sites and 46 US sites for evaluation.

Owner Information

None of the targeted websites from either Japan or the United States were personal websites. We categorized organizational websites as follows: (1) from Japan, medical: 33; general (profit): 7; general (nonprofit): 3; administrative (public): 4; and academic: 2; and (2) from the United States, medical: 24; general (profit): 7; general (nonprofit): 4; administrative (public): 7; and academic: 4. In both cases, medical websites were the most frequent (Table 1). The chi-square test revealed no significant differences in the categorization of owner information between the Japanese and US websites (*P*=.55).

Website Content Categories

Most dedicated palliative care sites from the United States, but none of those from Japan, were hospice service program websites. Furthermore, most general medical sites from the United States were informational sites about home visits, while most of those from Japan were general hospital introductory sites.

Evaluation Index Scores

On comparing evaluation index scores, US websites had significantly higher scores related to reliability-1 and reliability-2 than did the Japanese websites (Table 2). Scores related to information quality and usefulness did not differ markedly.

Table 1. Comparison of the frequency of websites with owner information and content categories between Japan and the United States.

Information categories	Japan (n)	United States (n)	*P* value[a]
Owner information			.55
Medical	33	24	
General (profit)	7	7	
General (nonprofit)	3	4	
Administrative (public)	4	7	
Academic	2	4	
Personal	0	0	
Contents			.005
Palliative care related	5	20	
General information about cancer	6	6	
General medical	32	19	
Nonmedical	6	1	

[a] Chi-square test.

Table 2. Comparison of mean evaluation index scores between Japanese and US websites.

Instruments and items	Japan (n=49)	United States (n=46)	*P* value[a]	Effect size
DISCERN score, mean (SD)				
Reliability-1	12.1 (4.6)	14.5 (5.2)	.02	0.51
Information quality	18.6 (6.7)	19.9 (6.6)	.26	0.24
Reliability-1 plus information quality	30.0 (11.4)	34.4 (10.6)	.07	0.43
Overall evaluation	2.1 (0.9)	2.6 (0.9)	.007	0.59
LIDA score, mean (SD)				
Usefulness	44.1 (7.0)	44.5 (5.4)	.61	0.26
Clarity	14.2 (3.7)	13.6 (2.7)	.50	0.16
Consistency	9.4 (1.6)	8.9 (0.3)	.30	0.18
Functionality	12.5 (2.2)	12.4 (1.9)	.26	0.15
Engagability	7.9 (1.7)	9.5 (1.8)	<.001	0.95
Reliability-2	16.1 (6.3)	21.1 (8.9)	.004	0.60
Currency	2.8 (1.8)	3.9 (1.5)	.001	0.68
Conflicts of interest	5.0 (1.6)	5.4 (1.4)	.23	0.24
Content production	1.6 (1.5)	1.9 (2.8)	.54	0.28
Content production procedure	1.9 (1.9)	4.1 (3.7)	<.001	0.75
Output of content	4.8 (1.5)	5.8 (2.0)	.009	0.54
Usefulness plus reliability-2	60.2 (12.1)	65.3 (11.2)	.045	0.42

[a] Student *t* test.

Reliability of the Target Websites

Comparison Between High- and Low-Rank Websites

The reliability of high- and low-rank websites significantly varied in both Japan and the United States. In US websites showing larger effect sizes, the difference was more marked between low-rank websites (Table 3).

Similarly, information quality, but not usefulness, also significantly varied between high- and low-rank websites from both Japan and the United States.

A large number of US websites had marks indicating certification of compliance guarantees by third parties, represented by the Health on the Net (HON) Foundation. In contrast, none of the Japanese websites had such marks.

Table 3. Comparison of scores between high- and low-rank websites.

Instruments and items	Japan				United States			
	High (n=24)	Low (n=25)	*P* value[a]	Effect size	High (n=23)	Low (n=23)	*P* value[a]	Effect size
DISCERN score, mean (SD)								
Reliability-1	14.4 (5.0)	9.9 (2.9)	<.001	1.12	17.6 (3.4)	11.3 (4.9)	<.001	1.41
Information quality	21.8 (7.8)	15.4 (3.3)	<.001	1.06	24.3 (5.7)	15.3 (3.7)	<.001	1.83
Reliability-1 plus information quality	34.8 (14.0)	25.3 (4.8)	.004	0.90	41.9 (7.5)	26.6 (7.3)	<.001	1.96
Overall evaluation	2.5 (0.9)	1.6 (0.7)	<.001	1.14	3.1 (0.7)	2.0 (0.8)	<.001	1.47
LIDA score, mean (SD)								
Usefulness	44.6 (5.3)	45.3 (2.1)	.64	0.16	44.0 (5.7)	44.3 (5.7)	.56	0.17
Reliability-2	20.0 (6.9)	13.0 (2.8)	<.001	1.35	26.1 (9.0)	15.9 (9.0)	<.001	1.32
Usefulness plus reliability-2	62.0 (16.7)	58.3 (3.8)	.28	0.32	70.2 (11.7)	60.2 (11.7)	.02	0.66

[a] Student *t* test.

Figure 2. Correlation of DISCERN scores related to reliability-1 and ranks between Japanese and US websites.

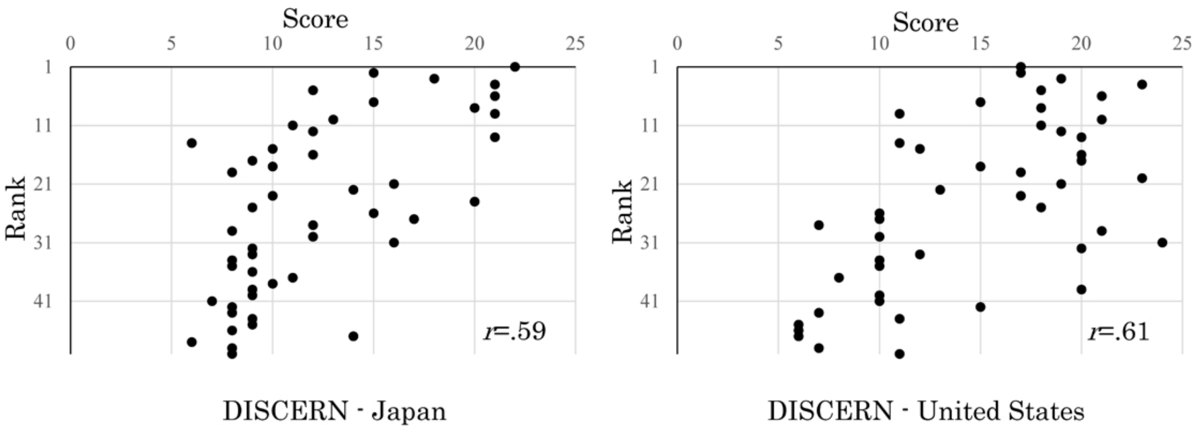
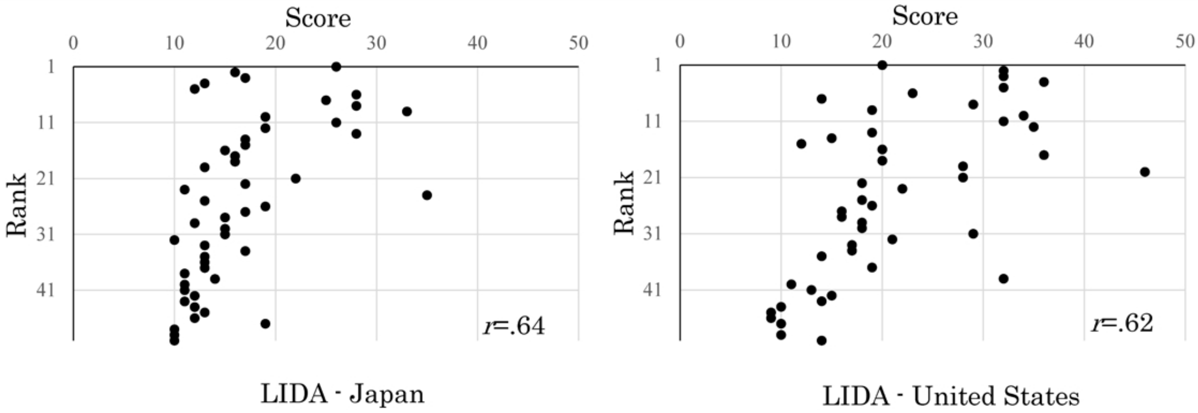


Figure 3. Correlation of LIDA scores related to reliability-2 and ranks between Japanese and US websites.



Correlation Between Reliability-Related Scores and Ranks

DISCERN (Figure 2) and LIDA (Figure 3) reliability scores of the target websites were weakly correlated with their ranks.

Discussion

Principal Findings

This study compared and evaluated the reliability, quality, and usefulness of palliative care information available on websites from Japan and the United States. Although the usefulness was similar between the 2 countries, the reliability and quality were lower in Japanese websites, revealing insufficient measures to provide such information with sufficient reliability in Japan. In previous studies, websites not specifying information sources and those showing profit-seeking behavior were noted as problematic [10,11]. Similarly, in this study, information sources and dates of publication were not appropriately described on some Japanese websites, revealing their insufficient usability for patients and their families. It has been suggested that Japanese palliative care sites have problems with their update frequency, their update policies and procedures, and the scrutiny process that evidence must undergo, since the scores for the US websites showed a significantly large effect size in 2 categories: currency and content production procedures, especially in the LIDA subcategory. There was no significant difference in the overall usefulness of the websites, but in the LIDA subcategory engagability, US websites had a higher score and were slightly more user friendly because many sites had interactive content and pictograms as an alternative to text. The subcategory engagability includes queries about website integration with nonverbal expression and, in the realm of drug information, the usefulness of pictograms as a tool for information communication has been reported [12,13]. This suggests that using pictograms could enhance a website's user friendliness for users whose mother tongue is different from the language of the website. Additionally, we do not know whether any Japanese palliative care–related websites have applied the HON Foundation code of conduct (HONcode), while many English-language websites do [14-16]. A certification system such as HONcode could be useful to help determine the reliability of a website because it would be possible to refine one's search by whether sites are certified.

Owner information categorization did not significantly vary between the 2 countries, indicating that websites belonging to medical institutions tended to be the most frequent information sources for patients in both countries. However, regarding each content category, dedicated palliative care sites ranked highest in the United States, and general medical information sites ranked highest in Japan (especially in cases where palliative care information was included in the hospital informational sites). Therefore, it may be necessary to create official sites that are dedicated to palliative care service and to provide information that is regularly reviewed and updated. Additionally, proactively applying a third-party certification system such as the HONcode is desirable because even medical experts often vary in their subjective assessments [17].

The significant difference in the level of reliability between high- and low-rank websites in both countries and the weak correlation between such levels and ranks suggest an association between the rank of each website and its reliability. The order in which websites are displayed in Google search results is based on an algorithm called PageRank [18]. This algorithm calculates original scores for individual websites based on keyword matching, the number of backlinks, updating frequency, information volume and consistency, browsing frequency, and coding appropriateness. PageRank determines where a website ranks in the search results. Therefore, the rank of a website in the search results does not necessarily reflect its reliability. However, the possibility of a website appearing high in the order as a result of being regarded as highly reliable by users and achieving a large number of accesses should also be considered. In this respect, detailed search orders themselves may not accurately reflect the reliability of websites, although those providing more reliable information tend to appear higher in the search results. In fact, in previous studies, search orders were reported to be inappropriate as a measure to examine the reliability of websites [19], and we obtained similar findings in this study. Thus, when using internet search services, the following points listed by some researchers should be noted: information contained in webpages may be inaccurate; such information is based on providers' intentions; and websites are not primary information sources [20]. Our study also suggested the necessity of exercising caution when using information available on the internet.

With the revision of Google' search logic, the individualization of information based on locality and access history has been promoted, but reliability-related issues have yet to be resolved. At this point, information available on websites visited by patients has been reported to be less accurate than that provided through websites targeting medical professionals [21,22]. As it may be difficult for patients without expertise to judge whether such information is sufficiently reliable, medical professionals should support patients in obtaining information that they need.

Study Limitations

This study had some considerable limitations. First, our searches may have missed some websites. Second, we might have obtained different results using other search engines. However, although search results can differ by adding keywords, we think that this difference would be negligible and unlikely to have a large influence on the results. Third, the results might not be applicable to palliative care websites from countries other than the United States and Japan. However, although there are some limitations to this study, few studies have assessed the reliability of website information in the palliative care realm. Therefore, our results can be considered to offer future possibilities for providing information.

Conclusions

While usefulness-related scores did not significantly vary, reliability-related scores were lower for Japanese websites. In 3 LIDA instrument subcategories, US scores were higher and had large effect sizes: engagability, currency, and content production procedure. This suggests that Japanese websites have problems in their frequency of being updated, update procedure and policy, and scrutiny process for evidence. We also clarified that the reliability of websites is weakly correlated with their ranks, but such ranks are not sufficient to judge whether the websites are sufficiently reliable. Based on these results, it may be necessary to evaluate information sources used by patients.

Conflicts of Interest

None declared.

References

1. World Health Organization. WHO definition of palliative care. Geneva, Switzerland: WHO; 2018. URL: http://www.who.int/cancer/palliative/definition/en/ [accessed 2018-02-06] [WebCite Cache ID 6x1fOcc03]
2. Takayasu M, Yonahara N, Matsumoto M, Sakurada T, Kobayashi E, Satoh N. Research on palliative care information provision. Regul Sci Med Prod 2016;6:11-20. [doi: 10.14982/rsmp.6.11]
3. Cullen RJ. In search of evidence: family practitioners' use of the Internet for clinical information. J Med Libr Assoc 2002 Oct;90(4):370-379 [FREE Full text] [Medline: 12398243]
4. Ministry of Public Management, Home Affairs, Posts, and Telecommunications. Results of the National Telecommunication Survey 2015. Tokyo, Japan: Ministry of Internal Affairs and Communications; 2009. URL: http://www.soumu.go.jp/menu_news/s-news/01tsushin02_02000099.html [accessed 2018-02-06] [WebCite Cache ID 6x1UD9Qmi]
5. Watanabe K. [Cancer information service for improving palliative care]. Gan Kanja To Taishoryoho 2013;24:61-66.
6. Yamada Y, Hirakata M, Todoroki K, Okazaki S, Ishiguro R, Nobutou A, et al. The level of the information and understanding state of patients and family members before and after palliative care unit admissionvaluation of the nurses. Palliat Care Res 2013;8:361-370. [doi: 10.2512/jspm.8.361]
7. Pithon MM, dos Santos ES. Information available on the internet about pain after orthognathic surgery: a careful review. Dental Press J Orthod 2014;19(6):86-92 [FREE Full text] [doi: 10.1590/2176-9451.19.6.086-092.oar] [Medline: 25628084]
8. 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] [Medline: 10396471]
9. Minervation. LIDA. Oxford, UK: Minervation Ltd; 2012 Oct 02. URL: http://www.minervation.com/category/lida/ [accessed 2017-09-30] [WebCite Cache ID 6x1UQd4CK]
10. Küçükdurmaz F, Gomez MM, Secrist E, Parvizi J. Reliability, readability and quality of online information about femoracetabular impingement. Arch Bone Jt Surg 2015 Jul;3(3):163-168 [FREE Full text] [Medline: 26213699]
11. Tieman J. Ensuring quality in online palliative care resources. Cancers (Basel) 2016 Dec 13;8(12):1-7 [FREE Full text] [doi: 10.3390/cancers8120113] [Medline: 27983592]
12. Liu Y, Chiu S, Lin Y, Chiou W. Pictogram-based method of visualizing dietary intake. Methods Inf Med 2014;53(6):493-500. [doi: 10.3414/ME13-01-0072] [Medline: 25301223]
13. Imanishi T, Takamatsu C, Takayama A. Creation of pictograms regarding pharmacological effects of medicine: the necessity and evaluation of created pictograms. Jpn J Pharm Health Care Sci 2017;43:640-647.
14. Boyer C, Selby M, Scherrer JR, Appel RD. The Health On the Net Code of Conduct for medical and health websites. Comput Biol Med 1998 Sep;28(5):603-610. [Medline: 9861515]
15. Brulet A, Llorca G, Letrilliart L. Medical wikis dedicated to clinical practice: a systematic review. J Med Internet Res 2015;17(2):e48 [FREE Full text] [doi: 10.2196/jmir.3574] [Medline: 25700482]
16. Boyer C, Gaudinat A, Hanbury A, Appel RD, Ball MJ, Carpentier M, et al. Accessing reliable health information on the Web: a review of the HON approach. Stud Health Technol Inform 2017;245:1004-1008. [Medline: 29295252]
17. Craigie M, Loader B, Burrows R, Muncer S. Reliability of health information on the Internet: an examination of experts' ratings. J Med Internet Res 2002;4(1):e2 [FREE Full text] [doi: 10.2196/jmir.4.1.e2] [Medline: 11956034]
18. Wikipedia: Google.: Wikimedia Foundation, Inc URL: https://ja.wikipedia.org/wiki/Google [accessed 2017-09-30] [WebCite Cache ID 6x25et370]
19. Kawazoe S, Tsutsui A, Kishimoto K, Fukushima N. A study of the health food information on Japanese websites. Shokumotu Gakkai Shi 2013;68:17-24.
20. Tanimizu M, Shinkai S, Hyodo I, Masumoto S, Nasu J, Hirazaki T. Search of the Internet cancer information in the patient care. Chiro 2005;87:1635-1639.
21. Knapp C, Madden V, Marcu M, Wang H, Curtis C, Sloyer P, et al. Information seeking behaviors of parents whose children have life-threatening illnesses. Pediatr Blood Cancer 2011 May;56(5):805-811. [doi: 10.1002/pbc.22674] [Medline: 21370415]
22. Kishimoto K, Yoshino C, Fukushima N. [Study of the health food information for cancer patients on Japanese websites]. Yakugaku Zasshi 2010 Aug;130(8):1017-1027 [FREE Full text] [Medline: 20686206]

Abbreviations

HON: Health on the Net
HONcode: HON Foundation code of conduct

Edited by G Eysenbach; submitted 05.12.17; peer-reviewed by R Pankomera, A Francke; comments to author 04.01.18; revised version received 08.02.18; accepted 25.02.18; published 03.04.18.

Please cite as:
Tanabe K, Fujiwara K, Ogura H, Yasuda H, Goto N, Ohtsu F
Quality of Web Information About Palliative Care on Websites from the United States and Japan: Comparative Evaluation Study
Interact J Med Res 2018;7(1):e7
URL: http://www.i-jmr.org/2018/1/e7/
doi:10.2196/ijmr.9574
PMID:29615388

Short Paper

The Validity of Online Patient Ratings of Physicians: Analysis of Physician Peer Reviews and Patient Ratings

Robert J McGrath[1*], BSc, MS, PhD; Jennifer Lewis Priestley[2*], BSc, MSc, PhD; Yiyun Zhou[2*], BSc, MA; Patrick J Culligan[3*], BSc, MD

[1]University of New Hampshire, Durham, NH, United States
[2]Graduate College, Kennesaw State University, Kennesaw, GA, United States
[3]Department of Urology, Weill Cornell Medicine, New York, NY, United States
[*]all authors contributed equally

Corresponding Author:
Yiyun Zhou, BSc, MA
Graduate College
Kennesaw State University
1000 Chastain Road, Kennesaw
Kennesaw, GA,
United States
Phone: 1 4043331288
Email: yzhou20@kennesaw.edu

Abstract

Background: Information from ratings sites are increasingly informing patient decisions related to health care and the selection of physicians.

Objective: The current study sought to determine the validity of online patient ratings of physicians through comparison with physician peer review.

Methods: We extracted 223,715 reviews of 41,104 physicians from 10 of the largest cities in the United States, including 1142 physicians listed as “America’s Top Doctors” through physician peer review. Differences in mean online patient ratings were tested for physicians who were listed and those who were not.

Results: Overall, no differences were found between the online patient ratings based upon physician peer review status. However, statistical differences were found for four specialties (family medicine, allergists, internal medicine, and pediatrics), with online patient ratings significantly higher for those physicians listed as a peer-reviewed “Top Doctor” versus those who were not.

Conclusions: The results of this large-scale study indicate that while online patient ratings are consistent with physician peer review for four nonsurgical, primarily in-office specializations, patient ratings were not consistent with physician peer review for specializations like anesthesiology. This result indicates that the validity of patient ratings varies by medical specialization.

(Interact J Med Res 2018;7(1):e8) doi:10.2196/ijmr.9350

KEYWORDS

physician review websites; online patient ratings; physician peer review

Introduction

In a 2016 study, the Pew Research Center found that 84% of all adults in the United States use online ratings sites to inform their product or service purchase decisions [1]. The same is true for health care: patients increasingly access online ratings sites to inform their health care decisions, with online ratings emerging as the most influential factor for choosing a physician. In a 2017 study by the National Institutes of Health, 53% of physicians and 39% of patients reported visiting a health care rating website at least once [2]. Overall, physicians indicated that the numerical results from these ratings websites were valid approximately 53% of the time, while patients indicated that they thought the ratings were valid 36% of the time [2].

RateMDs.com, HealthGrades.com, and Vitals.com are three frequently visited health care provider ratings websites, with over 2.6 million, 6.1 million, and 7.8 million reviews, respectively [3-5]. For these three sites, numeric rating scales range from 1 (poor) to 5 (excellent) and cover perceptions of physician knowledge, helpfulness, punctuality, and staff. Most patients give physicians positive ratings: one study reported that over 90% of all ratings were positive [6] and another reported that as the frequency of ratings increased, the average mean rating increased [7].

Extending the findings of the study by the National Institutes of Health, we sought to determine the validity of online patient ratings through comparison with physician peer review, defined in this study through Castle Connolly Medical. Specifically, we tested whether mean online patient ratings for physicians, by specialty, are higher for those physicians who have been nominated by their peers as one of “America’s Top Doctors” or not, as reported by Castle Connolly Medical. If online patient ratings were consistent with Castle Connolly Medical, ratings for physicians listed would be higher than for those not listed, thereby providing support for the validity of physician online review sites to inform health care-related decisions.

Methods

The basis for physician peer review selected for the current study is Castle Connolly Medical, a private consumer research firm that distinguishes top providers both nationally and regionally through a peer nomination process that involves over 50,000 providers and hospital and health care executives. Castle Connolly Medical receives over 100,000 nominations each year and a physician-led research team awards top providers from these nominations [8]. Lists are generated for each health care specialty as well as most subspecialties.

Several studies have similarly selected physician peer review through Castle Connolly Medical as a basis to assess the validity and role of patient online ratings sites, including an assessment for hand surgeons in the United States [9], as well as a more general correlation of physician attributes and ranking of hospital affiliations with peer review results [10]. Other studies have found alternative domain-specific objective measures to corroborate online review sites with relevant tangible outcomes, like restaurant ratings with patron visits [11].

Results

This study examined 223,715 reviews of 41,104 unique (nonduplicated) physicians from 10 of the largest cities in the United States (Atlanta, Boston, Chicago, Dallas, Washington DC, Los Angeles, Miami, New York, Philadelphia, and San Francisco). Reviews were extracted in January 2017. Of these physicians, 1142 were included as “America’s Top Doctors” in the Castle Connolly Medical rankings. The number of ratings and physicians evaluated makes this study the largest-scale evaluation of its kind, to date. The profile of the overall sample is provided in Table 1. Specific elements extracted included doctor name, rating (numeric), number of reviews, specialization, source (ratings site), city, and state. To mitigate issues related to “fake” reviews as well as influential observations, we excluded any physician with fewer than three reviews and specializations with fewer than five reviews. Of the total number of physicians with reviews, 16,525 had fewer than three reviews, making the final analyzed sample size of physicians 24,579 (Multimedia Appendix 1).

From Multimedia Appendix 1, four specializations demonstrated differences in online patient average ratings between those physicians included in Castle Connolly Medical’s listing of “America’s Top Doctors” and those not listed: allergists, family medicine, internists, and pediatricians. For each of these specializations, those physicians with a listing in Castle Connolly Medical received a higher rating than those physicians not listed. The remaining specializations exhibited little difference between physicians listed and those not listed.

Table 1. Rated physicians by source.

Ratings source	Number of physicians[a]	Number of reviews	Average rating (1-5)
HealthGrades	17,385	113,427	3.97
RateMDs	19,631	72,228	3.83
Vitals	4088	38,060	4.06
Total	41,104	223,715	3.91

[a] Nonduplicated, unique number of physicians.

Discussion

Principal Findings

This study sought to determine the validity of patient ratings for physicians by evaluating the mean online ratings for physicians, by specialty, between those who had been nominated by their peers as one of “America’s Top Doctors” or not, as reported by Castle Connolly Medical. We found that four specializations demonstrated differences in ratings between those physicians included in Castle Connolly Medical’s listing of “America’s Top Doctors” and those not listed: allergists, family medicine, internists, and pediatricians. Specifically, our study found that the validity of patient online reviews of physicians varies by specialization. This finding has implications related to how patients make choices related to health care.

Physicians have been inundated with mandates for attaining the “triple aim” of reducing costs and increasing patient experiences and quality [12]. In doing so, many have moved to a model of “patient-centered care” which seeks to form continuous patient-physician relationships [13]. Thus, some practices have simultaneously begun to direct attention at both the nature of the relationship and the quality of that encounter. Given that these efforts appear to be primarily directed at more “primary care” and “in-office” settings, our finding that patient reviews are valid for specializations that could be characterized as primarily “in-office” settings is not unexpected.

Within the context of promoting competition, information transparency needs to be both complete and understood. This review would suggest that online patient ratings accomplish neither of these market objectives. In fact, there may be implications for shopping behavior to negatively influence quality of care outcomes; care continuity is associated with many positive health outcomes including decreased hospitalizations, fewer emergency room visits, lower health care costs, and improvements in the use of preventative care services [14]. Conversely, evidence indicates that patients who experience more fragmented primary care services also have patterns of care that more significantly deviate from determined best practice guidelines and result in higher overall health care costs. Negative reviews could thus promote "doctor shopping" based on incomplete or nonfactual information and lead to more fragmented care continuity, and potentially less optimal health outcomes [15,16].

Health systems have called for more holistic approaches to treating patients and placing measurable value on attributes such as trust and continuity of care [17]. In a recent edition of the Journal of the American Medical Association, physicians discussed the role that standardized quality assessment tools have on care practice and the need to be thoughtful when constructing such measures [18]. Physician rating websites have utility, but are imperfect proxies for competence [19,20]. If such questions have arisen about standard best practice measurement, even greater questions exist about unstandardized and undefined open assessments such as online patient reviews, particularly in specialties where the patient has limited direct experience with their health care provider (eg, Anesthesiology).

Limitations

The selected basis for physician peer review for this study–Castle Connolly Medical–is not immune to challenge; while the organization does not receive payments or petitions, physicians have publicly questioned the "lobbying" efforts that some colleagues undertake to be included in their lists. However, no objective truth in the determination of a "good" or "bad" physician has been established. Other studies have explored alternative assessments for physician performance (eg, clinical outcomes, costs to treat, board certifications) and have acknowledged a variety of issues and limitations related to associating reviews with performance [21,22].

The current study only incorporated average numerical results for physicians (rather than an individual numeric rating for each review) from the three ratings sources; text from reviews was not analyzed. While the patterns and general findings would likely not change based upon text analysis, the text may provide additional insights regarding frequently occurring terms or relevant patterns for interested researchers.

We were not able to ascertain details about the individuals providing the ratings. Specifically, this study did not consider the patients' insurance type. This insurance type could affect how a patient experiences the service provided relative to perceived value; those with higher out-of-pocket direct costs via copays and/or high deductibles may be more cost sensitive and therefore more likely to "shop" for health care in the face these payments.

Conclusions

A deceptive review or set of reviews related to a hotel visit is an inconvenience, but decisions based on deceptive or poorly-informed patient reviews related to a health care provider could have dire consequences for an individual using these reviews to inform their health care-related decisions. The presence of online ratings sites will likely continue to grow and expand across all segments of the economy. The results of this large-scale study indicate that while patient ratings are consistent with physician peer review ratings for specialties like allergists and pediatricians, patient reviews were not consistent with medical peer review for specializations characterized by less patient contact (eg, anesthesiology). This result may indicate that patients are not sufficiently knowledgeable to provide informed physician ratings for some medical specializations, leading other information-seekers to potentially less-qualified providers.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Overall mean ratings by specialization for physicians listed and not listed in Castle Connolly Medical.

[PDF File (Adobe PDF File), 95KB - ijmr_v7i1e8_app1.pdf]

References

1. Smith A, Anderson M. Pew Research Center. 2016 Dec 19. Online shoppng and e-commerce: online reviews URL: http://www.pewinternet.org/2016/12/19/online-reviews [WebCite Cache ID 6voI1mgm9]
2. Holliday AM, Kachalia A, Meyer GS, Sequist TD. Physician and patient views on public physician rating websites: a cross-sectional study. J Gen Intern Med 2017 Jun;32(6):626-631. [doi: 10.1007/s11606-017-3982-5] [Medline: 28150098]
3. RateMDs. 2017. Doctor reviews and ratings URL: http://www.ratemds.com [WebCite Cache ID 6voIOhwyV]
4. Vitals. 2017. URL: http://www.vitals.com [WebCite Cache ID 6vpTL9etv]
5. HealthGrades. 2017. Review your doctor URL: https://www.healthgrades.com/review [WebCite Cache ID 6vpTWhQlp]

6. Emmert M, Sander U, Pisch F. Eight questions about physician-rating websites: a systematic review. J Med Internet Res 2013 Feb 01;15(2):e24 [FREE Full text] [doi: 10.2196/jmir.2360] [Medline: 23372115]
7. Grabner-Kräuter S, Waiguny MKJ. Insights into the impact of online physician reviews on patients' decision making: randomized experiment. J Med Internet Res 2015 Apr 09;17(4):e93 [FREE Full text] [doi: 10.2196/jmir.3991] [Medline: 25862516]
8. Castle Connolly. 2018. Top Doctors URL: https://www.castleconnolly.com [accessed 2018-03-26] [WebCite Cache ID 6yCyReZap]
9. Trehan SK, DeFrancesco CJ, Nguyen JT, Charalel RA, Daluiski A. Online patient ratings of hand surgeons. J Hand Surg Am 2016 Jan;41(1):98-103. [doi: 10.1016/j.jhsa.2015.10.006] [Medline: 26710742]
10. Wiley MT, Rivas RL, Hristidis V. Provider attributes correlation analysis to their referral frequency and awards. BMC Health Serv Res 2016 Mar 14;16:90 [FREE Full text] [doi: 10.1186/s12913-016-1338-1] [Medline: 26975310]
11. Gadidov B, Priestley JL. Does Yelp matter? Analyzing (and guide to using) ratings for a quick serve restaurant chain. In: Srinivasan S, editor. Guide to Big Data Applications. New York: Springer International Publishing; 2018:503-522.
12. Bodenheimer T, Sinsky C. From triple to quadruple aim: care of the patient requires care of the provider. Ann Fam Med 2014;12(6):573-576 [FREE Full text] [doi: 10.1370/afm.1713] [Medline: 25384822]
13. Detsky AS. What patients really want from health care. JAMA 2011 Dec 14;306(22):2500-2501. [doi: 10.1001/jama.2011.1819] [Medline: 22166610]
14. Frandsen BR, Joynt KE, Rebitzer JB, Jha AK. Care fragmentation, quality, and costs among chronically ill patients. Am J Manag Care 2015 May;21(5):355-362 [FREE Full text] [Medline: 26167702]
15. Frost A, Newman D. Health Care Cost Institute. 2018. Spending on shoppable services in health care URL: http://www.healthcostinstitute.org/files/Shoppable%20Services%20IB%203.2.16_0.pdf [WebCite Cache ID 6vpjZ9I4X]
16. Desai S, Hatfield LA, Hicks AL, Chernew ME, Mehrotra A. Association between availability of a price transparency tool and outpatient spending. JAMA 2016 May 03;315(17):1874-1881. [doi: 10.1001/jama.2016.4288] [Medline: 27139060]
17. Friedberg M, Chen P, Van BK, Aunon F, Pham C, Caloyeras J, et al. Factors affecting physician professional satisfaction and their implications for patient care, health systems, and health policy. Rand Health Q 2014;3(4):1 [FREE Full text] [Medline: 28083306]
18. Goitein L, James B. Standardized best practices and individual craft-based medicine: a conversation about quality. JAMA Intern Med 2016 Jun 01;176(6):835-838 [FREE Full text] [doi: 10.1001/jamainternmed.2016.1641] [Medline: 27183319]
19. Murphy GP, Awad MA, Osterberg EC, Gaither TW, Chumnarnsongkhroh T, Washington SL, et al. Web-based physician ratings for California physicians on probation. J Med Internet Res 2017 Aug 22;19(8):e254 [FREE Full text] [doi: 10.2196/jmir.7488] [Medline: 28830852]
20. Okike K, Peter-Bibb TK, Xie KC, Okike ON. Association between physician online rating and quality of care. J Med Internet Res 2016 Dec 13;18(12):e324 [FREE Full text] [doi: 10.2196/jmir.6612] [Medline: 27965191]
21. Trzeciak S, Gaughan JP, Bosire J, Mazzarelli AJ. Association between medicare summary star ratings for patient experience and clinical outcomes in US hospitals. J Patient Exp 2016 Mar;3(1):6-9 [FREE Full text] [doi: 10.1177/2374373516636681] [Medline: 28725825]
22. Liu J, Matelski J, Cram P, Urbach D, Bell C. Association between online physician ratings and cardiac surgery mortality. Circ Cardiovasc Qual Outcomes 2016 Nov;9(6):788-791 [FREE Full text] [doi: 10.1161/CIRCOUTCOMES.116.003016] [Medline: 27803091]

Edited by G Eysenbach; submitted 07.11.17; peer-reviewed by S Nargundkar, M Aaron, S Bidmon, A Herrmann, F Rothenfluh, S Grabner-Kräuter; comments to author 17.12.17; revised version received 27.12.17; accepted 19.02.18; published 09.04.18.

Please cite as:
McGrath RJ, Priestley JL, Zhou Y, Culligan PJ
The Validity of Online Patient Ratings of Physicians: Analysis of Physician Peer Reviews and Patient Ratings
Interact J Med Res 2018;7(1):e8
URL: http://www.i-jmr.org/2018/1/e8/
doi:10.2196/ijmr.9350
PMID:29631992

Original Paper

The Patient Perspective on the Impact of Tenosynovial Giant Cell Tumors on Daily Living: Crowdsourcing Study on Physical Function and Quality of Life

Monique Josephine Mastboom[1], MD; Rosa Planje[1], BSc; Michiel Adreanus van de Sande[1], MD, PhD

Department of Orthopedics, Leiden University Medical Center, University of Leiden, Leiden, Netherlands

Corresponding Author:
Monique Josephine Mastboom, MD
Department of Orthopedics
Leiden University Medical Center
University of Leiden
Albinusdreef 2
Leiden, 2333 ZA
Netherlands
Phone: 31 715264088
Fax: 31 715266743
Email: mjlmastboom@lumc.nl

Abstract

Background: Tenosynovial giant cell tumor (TGCT) is a rare, benign lesion affecting the synovial lining of joints, bursae, and tendon sheaths. It is generally characterized as a locally aggressive and often recurring tumor. A distinction is made between localized- and diffuse-type. The impact of TGCT on daily living is currently ill-described.

Objective: The aim of this crowdsourcing study was to evaluate the impact of TGCT on physical function, daily activities, societal participation (work, sports, and hobbies), and overall quality of life from a patient perspective. The secondary aim was to define risk factors for deteriorated outcome in TGCT.

Methods: Members of the largest known TGCT Facebook community, PVNS is Pants!!, were invited to an e-survey, partially consisting of validated questionnaires, for 6 months. To confirm disease presence and TGCT-type, patients were requested to share histological or radiological proof of TGCT. Unpaired t tests and chi-square tests were used to compare groups with and without proof and to define risk factors for deteriorated outcome.

Results: Three hundred thirty-seven questionnaires, originating from 30 countries, were completed. Median age at diagnosis was 33 (interquartile range [IQR]=25-42) years, majority was female (79.8% [269/337]), diffuse TGCT (70.3% [237/337]), and affected lower extremities (knee 70.9% [239/337] and hip 9.5% [32/337]). In 299 lower-extremity TGCT patients (32.4% [97/299]) with disease confirmation, recurrence rate was 36% and 69.5% in localized and diffuse type, respectively. For both types, pain and swelling decreased after treatment; in contrast, stiffness and range of motion worsened. Patients were limited in their employment (localized 13% [8/61]; diffuse 11.0% [21/191]) and sport-activities (localized 58% [40/69]; diffuse 63.9% [147/230]). Compared with general US population, all patients showed lower Patient-Reported Outcomes Measurements Information System-Physical Function (PROMIS-PF), Short Form-12 (SF-12), and EuroQoL 5 Dimensions 5 Levels (EQ5D-5L) scores, considered clinically relevant, according to estimated minimal important difference (MID). Diffuse versus localized type scored almost 0.5 standard deviation lower for PROMIS-PF ($P<.001$) and demonstrated a utility score of 5% lower for EQ-5D-5L ($P=.03$). In localized TGCT, recurrent disease and ≥2 surgeries negatively influenced scores of Visual Analog Scale (VAS)-pain/stiffness, SF-12, and EQ-5D-5L ($P<.05$). In diffuse type, recurrence resulted in lower score for VAS, PROMIS-PF, SF-12, and EQ-5D-5L ($P<.05$). In both types, patients with treatment ≤1year had significantly lower SF-12.

Conclusions: TGCT has a major impact on daily living in a relatively young and working population. Patients with diffuse type, recurrent disease, and ≥2 surgeries represent lowest functional and quality of life outcomes. Physicians should be aware that TGCT patients frequently continue to experience declined health-related quality of life and physical function and often remain limited in daily life, even after treatment(s).

(Interact J Med Res 2018;7(1):e4) doi:10.2196/ijmr.9325

KEYWORDS

synovitis; pigmented villonodular; giant cell tumor of tendon sheath; rare diseases; crowdsourcing; social media; patient-reported outcome measures; quality of life; health-related quality of life; social participation; surveys and questionnaires

Introduction

Tenosynovial giant cell tumor (TGCT), previously pigmented villonodular synovitis (PVNS), is a rare, proliferative neoplasm affecting the synovial lining of joints, bursae, and tendons sheaths. According to growth pattern, a radiological distinction is made between a well-circumscribed lesion (localized type) and a locally more aggressive lesion (diffuse type) [1,2]. The incidence rate reveals its rarity: for localized type (excluding digits), 10.2 per million person-years and for diffuse type, 4.1 per million person-years. TGCT is a monoarticular disease, concerning large joints, typically about the knee: 46% in localized-type and 64% to 75% in diffuse-type. Male-female ratio is about 1:1.5 for both types, with a median age at the time of TGCT diagnosis of 30 to 50 years [1-3]. Most common initial symptoms are pain, stiffness, and swelling. Additional symptoms might be limited range of motion, instability, giving way, and locking complaints [4]. Due to these unspecific signs and the rarity of the disease, patients frequently experience a delay of years in diagnosis [3,5,6]. To treat these symptoms, current treatment of choice is surgical excision, either by arthroscopic or open synovectomy [7]. After surgical resection, high recurrence rates are known, with the localized type up to 50% and the diffuse type up to 92% [6].

Once TGCT is diagnosed, a high health care burden is identified with a significant increase in health care costs, ambulatory expenses, and physical therapy [8]. In describing treatment benefits and standard oncologic end points, patient-reported outcome instruments are increasingly used. Visual Analog Scale (VAS) for worst pain-stiffness and Patient-Reported Outcomes Measurement Information System-Physical Function (PROMIS-PF) questionnaires were identified as most valuable measures for TGCT symptoms in a relatively small TGCT patient cohort (n=22) [4].

The impact of TGCT symptoms following surgery(s) and recurrences on daily living, sports, and work activities is currently ill-described. Although TGCT is not considered lethal, this tumor is hypothesized to have major impact on daily living. Especially diffuse disease is notorious for its negative influence on both local recurrence risk and functional outcome [9].

Use of an e-survey is a unique possibility to reach a large elusive TGCT population and to globally evaluate impact of TGCT on patients’ daily life. This crowdsourcing study evaluates effect of TGCT on physical function, daily activities, societal participation (work, sports, and hobbies), and overall quality of life from a patient perspective. Secondary aim is to define risk factors for deteriorated outcome in TGCT.

Methods

Study Design

This cross-sectional crowdsourcing study was performed at Leiden University Medical Center, Leiden, The Netherlands, in accordance with good clinical practice [the Declaration of Helsinki (2000)]. This study was conducted from December 2016 until end of May 2017 (6 months), using the largest known online TGCT community on Facebook, PVNS is Pants!!, to gather participants for the Web-based questionnaire. The study was conducted conforming to the Checklist for Reporting Results of Internet E-Surveys (CHERRIES), the checklist focusing on Web-based surveys [10] (Multimedia Appendix 1). NetQuestionnaire (NetQ) was used to complete the TGCT questionnaire. NetQ is a professional Web-survey software, approved for (bio)medical research and supported by the Leiden University Medical Center (LUMC). Respondents were able to review and change their answers before submitting.

Patients and Recruitment

Members of PVNS is Pants!! were requested to participate in our international crowdsourcing study “Evaluation of Tenosynovial Giant Cell Tumor on daily living” (Multimedia Appendix 2). At the time of writing (December, 2016), this closed Facebook community contained 2179 members. A patient-friendly TGCT-research-related message was posted in the Facebook community every 4 weeks to encourage TGCT patients to complete the questionnaire. Additional study updates and easily understandable information on TGCT were posted on the page of a newly designed TGCT study Facebook account [11].

All members of the Facebook community had access to the questionnaire. Solely patients with TGCT diagnosis were requested to participate in this study. To achieve a higher level of evidence, confirmation of TGCT (histological or radiological) was requested after completing the questionnaire. Sending (anonymized) medical reports to our protected email account was highly desirable but left to the discretion of the participant.

Members of Facebook community PVNS is Pants!! have been notified that (research-minded) doctors are members of this closed Facebook community for several years. Participation in this study was voluntary, and no incentives were offered. Informed consent was given by completing the survey. This study was approved by the Institutional Review Board (CME) from our institution (registration number P16.232, December 5, 2016).

Unique site visitors were determined by Internet protocol (IP) addresses. When duplicate entries were detected, the most recent one was included in the analyses. All password-protected documents were only accessible to TGCT researchers and saved on the secured departmental drive of our hospital. Data of participants were anonymized when medical proof was received or when the participant did not respond to our third request for medical confirmation. To ascertain TGCT diagnosis and TGCT type, all medical reports were verified by 2 TGCT researchers (MJLM, RP). When in disagreement, medical reports were checked by the senior orthopedic surgeon (MAJS) for final conclusion.

Questionnaire

On the very active Facebook community PVNS is Pants!!, several patient-initiated questionnaires and polls were performed, for instance, about treatments, coping strategies, daily limitations, and emotional struggles. Members expressed their desire for studies regarding these topics, since the majority of TGCT studies concern physical function and recurrent disease as outcome parameters. Therefore, a Web-based questionnaire, using mostly validated questionnaires, was composed to describe impact of TGCT on health-related outcome and daily living from a patient perspective. A prerequisite was that the questionnaire would be relevant for the heterogeneous TGCT population: for different large joints, different ages, males or females, localized or diffuse type, and for patients at different treatment stages.

To assess relevance and completeness of our questionnaire, a pilot test with the composed questionnaire was performed. One dedicated orthopedic oncologic surgeon (MAJS), 2 medical doctors (MJLM, RP), and 5 TGCT patients in our outpatient clinic, all fluent in written and spoken English language, tested the e-survey. Validated questionnaires were used as published by the owners. After the pilot test, a few nonvalidated questions were added or rephrased (Multimedia Appendix 3).

Nonvalidated questions concerned patient and tumor characteristics, medical history, TGCT symptoms, performed treatments, recurrences, employment status, sports, and number of visits to general practitioner and orthopedic surgeon. The majority of questions had a multiple-choice character, including a *not applicable* or *other* answer option. The exact number of nonvalidated questions depended on given answers. For instance, patients with an extensive TGCT-related history were asked additional questions on their history, in contrast to the patients awaiting their initial treatment.

Validated questionnaires on physical function and quality of life included: VAS for worst pain and stiffness in the last 24 hours, PROMIS-PF items, Short Form-12 Health Survey (SF-12), and EuroQoL EQ-5D-5L (EQ-5D-5L Descriptive System and EQ-5D-5L VAS). A total of 32 validated questions were included. VAS for pain and stiffness was used to estimate patient's pain and stiffness intensity for the past 24 hours: *no pain/stiffness at all* (0) and *worst pain/stiffness imaginable* (10).

PROMIS-PF instruments were used to measure self-reported capability of physical activities. In this study, short forms of physical functioning for lower and upper extremity were used with 5 response options: without any difficulty (5), with a little difficulty (4), with some difficulty (3), with much difficulty (2), and unable to do (1). Raw score was calculated by summing up the values of the response to each question and was converted into a T score by the Assessment Center from PROMIS-PF. A mean of a standardized T score of 50 with a standard deviation of 10 reflects the general US population [12].

The SF-12, a generic measure of health status, functioned as a shorter alternative for the SF-36. Number of answer options differed per question. Physical component summary (PCS) score and mental component summary (MCS) score were calculated. Similar to PROMIS-PF, the general US population had a mean of 50 with a standard deviation of 10 [13].

The EQ-5D-5L is one of the most commonly used generic health status measures in the world. Its descriptive system comprises 5 dimensions of health: mobility, self-care, usual activities, pain or discomfort, and anxiety or depression, with the following 5 levels of problems per dimension: no problems (1), slight problems (2), moderate problems (3), severe problems (4), and extreme problems (5) [14]. For each participant, answers per dimension were combined into an EQ-5D-5L health state. This health state was converted into a single index value (so-called utility score) for quality of life, by using the Crosswalk Index Value Calculator version 1.0 from the EuroQoL Group [15]. Utility scores were measured on an ordinal scale of 0 to 1, with 0 indicating death and 1 indicating full health [16]. Crosswalk valuation set for US population was used for all participants, since majority of the patients originated from the United States (42.7% [144/377]). A specific analysis, called sensitivity analysis, was performed using the valuation set for UK population, the second largest patient population (20.2% [68/337]) in this study. Scores calculated with US valuation set were compared with scores obtained by using UK valuation set to assess representativeness of the scores from validated questionnaires [14] (Multimedia Appendix 4).

Statistical Analysis

NetQ automatically captured questionnaire answers into an SPSS 23 file. Evaluation of TGCT on daily living was mainly descriptive.

Chi-square tests were used to compare patient groups with and without medical proof regarding gender (male vs female), TGCT localization (knee vs other large lower extremity joints [hip, ankle, and foot]), initial surgery (arthroscopy vs [one- or two-staged] open synovectomy), recurrence (yes vs no), total number of surgeries (1 surgery vs ≥2 surgeries), and time since last treatment for TGCT (≤1 year ago vs >1 year ago) (Multimedia Appendix 5).

Independent *t* tests were used to compare the mean age at the time of diagnosis and continuous scores of validated questionnaires. All reported *P* values were two-tailed. Statistical significance level was defined at *P*<.05.

Effect size, as a quantitative measure of the strength of a phenomenon, was calculated for both PROMIS-PF and SF-12 scores in localized- and diffuse-type patients, compared with general US population score. Effect size, or Cohen *d*, is the ratio of difference between two means divided by the standard deviation, expressed in standard deviation units. An effect size between 0.2 and 0.5 is considered small, 0.5 and 0.8 medium, and above 0.8 large [17].

The minimal important difference (MID), a quality of life measure, represents the smallest difference or change beyond statistical significance in an outcome measure score that would be considered clinically relevant by the value patients place on change. MID for EQ-5D-5L Index Scores is estimated between .037 and .069, based on the simulation-based instrument-defined MID estimates [18]. MID for PROMIS-PF was determined by Yost et al in advanced-stage cancer patients [19]. Differences in T scores between 4.0 and 6.0 were considered clinical relevant. MID for SF-12 PCS and MCS scores were calculated by Díaz-Arribas et al in >450 patients with low back pain and were stated at >3.29 for PCS and >3.77 for MCS [20].

Results

The TGCT questionnaire was initiated by 445 participants within a time frame of 6 months. For the present analysis, only fully completed, unique questionnaires (337) were included (Figure 1). The majority of incomplete questionnaires were early dropouts with a great lack of information and therefore unsuitable for analysis.

Most patients were female (79.8% [269/337]) and median age at diagnosis was 33 (interquartile range [IQR]=25-42) years. Patients originated from 30 different countries (United States: 42.7% [144/337]; United Kingdom: 20.2% [68/337]; and the Netherlands: 12.8% [43/337]). TGCT was typically located in lower extremities: knee (70.9% [239/337]), hip (9.5% [32/337]), ankle (11.0% [37/337]), and foot (3.0% [10/337]). Diffuse TGCT was diagnosed in 237 of 337 (70.3%) patients (Table 1). According to few TGCT patients with TGCT located in the upper extremity, 12 out of 337 patients (3.6%) were excluded for further analyses. Additionally, 26 out of 337 lower-extremity patients (7.7%) with unknown TGCT type were also excluded (Figure 1). Questionnaires of 299 lower-extremity patients with localized or diffuse TGCT were analyzed.

Disease Confirmation

Confirmation of TGCT was sent by 32.4% (97/299) of lower-extremity participants. In 81% (78/97) TGCT type was in concordance with questionnaire answer, in 16/97 (16%) medical reports TGCT type did not match the answer and was therefore adjusted according to the report and 3/97 (3%) patients answered TGCT type unknown and TGCT type was added in consistence with the report.

No important differences between patients with and without medical proof were detected (Multimedia Appendix 5), neither for localized or diffuse type separately. Therefore, patients with medical proof were considered representative for the entire study population and additional analyses were performed for the entire patient group.

Medical History and Tenosynovial Giant Cell Tumor Symptoms

5/69 (7%) and 29/230 (12.6%) in localized- and diffuse-type patients, respectively, had an autoimmune disease, mostly diabetes mellitus type I, Hashimoto, psoriasis, and thyroid disease. In all, 22/69 (32%) of localized-type and 70/230 (30.4%) of diffuse-type patients experienced a trauma at TGCT-affected joint, before diagnosis; sports injuries or fall incidents leading to a sprain or rupture. In all, 5/69 (7%) and 12/230 (5.2%) of patients in localized and diffuse TGCT had surgery of the affected joint before TGCT diagnosis, respectively, for example, meniscus or anterior cruciate ligament (ACL) reconstructions. In all, 6/230 (2.6%) of diffuse-type participants experienced both trauma and surgery before TGCT diagnosis.

Figure 1. Flowchart Tenosynovial Giant Cell Tumor (TGCT) questionnaire. Q: Questionnaires; L: Localized-TGCT; D: Diffuse-TGCT; U: Unknown-type TGCT.

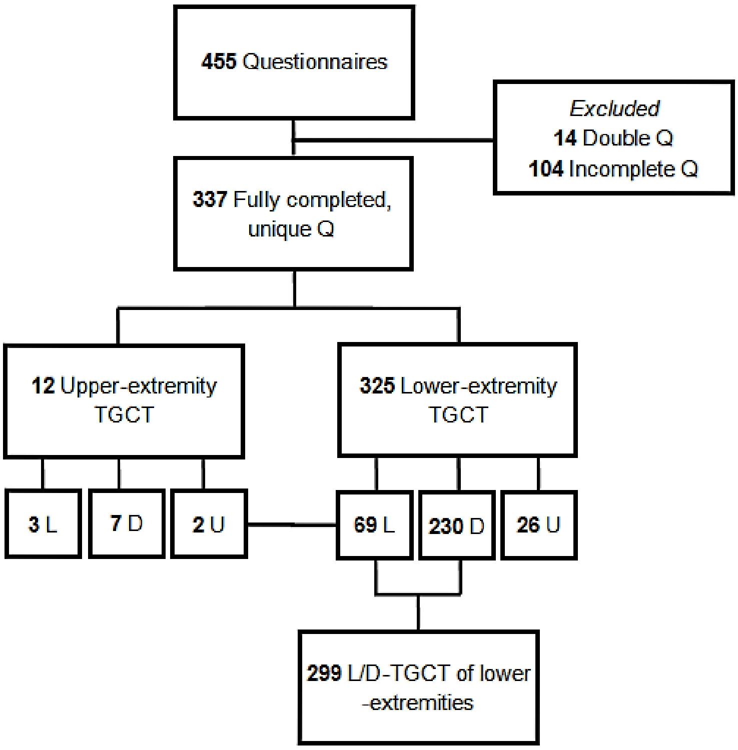


Table 1. Patient and tumor characteristics (n=337).

Characteristic	Value
Age at time of questionnaire (years), median (IQR[a])	41 (32-50)
Age at time of TGCT diagnosis (years), median (IQR)	33 (25-42)
Total, N (%)	337 (100)
Gender, n (%)	
Male	68 (20.2)
Female	269 (79.8)
Country of residence, n (%)	
United States of America	144 (42.7)
United Kingdom	68 (20.2)
The Netherlands	43 (12.8)
Australia	22 (6.5)
Canada	14 (4.2)
Other	46 (13.6)
TGCT[b] localization, n (%)	
Knee	239 (70.9)
Hip	32 (9.5)
Ankle	37 (11.0)
Foot	10 (3.0)
Shoulder	4 (1.2)
Elbow	6 (1.8)
Wrist	2 (0.6)
Other[c]	7 (2.1)
TGCT type, n (%)	
Localized	72 (21.4)
Diffuse	237 (70.3)
Unknown	28 (8.3)

[a] IQR: interquartile range (25-75%).

[b] TGCT: tenosynovial giant cell tumor.

[c] Other included multiple TGCT locations (all in lower extremity).

Table 2. Initial and current symptoms for localized and diffuse tenosynovial giant cell tumor (TGCT) (n=337).

TGCT[a]-related symptom	Localized TGCT (n=69)		Diffuse TGCT (n=230)	
	Initial, n (%)	Current, n (%)	Initial, n (%)	Current, n (%)
Pain	57 (83)	47 (68)	186 (80.9)	170 (73.9)
Swelling	53 (77)	29 (42)	190 (82.6)	139 (60.4)
Stiffness	38 (55)	41 (59)	128 (55.7)	148 (64.3)
Limited range of motion	38 (55)	29 (42)	140 (60.9)	149 (64.8)

[a] TGCT: tenosynovial giant cell tumor.

Majority of patients (92.6% [277/299]) were treated for TGCT. For both types, pain and swelling improved compared with initial situation. After treatment, more patients reported stiffness and limited range of motion (Table 2). A minority of the patients (<6%) currently experienced additional symptoms, including instability, buckling, hyperextension and/or hypermobility, clicking or locking or popping of joint, numbness, electric shocks, tingling, dull ache, heat of the affected joint, or hematoma.

Treatment(s)

Most performed initial surgery was arthroscopic synovectomy (57% [38/67] localized, 53.8% [113/210] diffuse) and open synovectomy, one- or two-staged (39% [26/67] localized, 42.9% [90/210] diffuse). In all, 5/67 (7%) localized-type and 53/210 (25.2%) diffuse-type patients had adjuvant therapies after initial surgery, mainly radiotherapy and 90-Yttrium. In all, 24/67 (36%) of localized type had recurrent disease after 1.5 (range 1-6) years, in contrast to 146/210 (69.5%) of diffuse type after 2.2 (range 1-23) years (Table 3). Additional surgery was performed in 23/67 (34%) of localized type and 125/210 (59.5%) of diffuse type, predominantly open synovectomy (one- or two-staged).

Impact of Tenosynovial Giant Cell Tumor on Daily Life

Due to TGCT, 8/61 (13%) and 21/191 (11.0%) of working population in localized and diffuse TGCT, respectively, was currently not able to (fully) perform their employment. Of these patients, 4/8 (50%) localized patients and 17/21 (81%) diffuse patients had recurrent disease. Majority of patients, 40/69 (58%) of localized and 147/230 (63.9%) of diffuse type, were unable to perform sport activities. In these patients, recurrent disease presented in 15/40 (38%) of localized type and 94/147 (63.9%) of diffuse type. Disease burden was estimated by mean number of visits to general practitioner (5.6 [range 1-50] visits for localized type, 7.1 [range 1-60] visits for diffuse type), and orthopedic surgeon (8.3 [range 1-97] visits for localized type, 11.9 [range 1-100] visits for diffuse type).

Results of validated questionnaires are shown in Table 4 (localized vs diffuse type), Table 5 (localized type), and Table 6 (diffuse type). Results with positive association are described in the text.

Worst Pain and Stiffness in Last 24 Hours: Visual Analog Scale Score

For localized type, best VAS pain score was 2.76 and VAS stiffness score was 2.80. In diffuse type, best scores for pain and stiffness were 3.04 and 3.08, respectively. Patients with recurrence of TGCT had deteriorated VAS score for pain and stiffness (P=.01 localized type and P<.001 diffuse type). In localized type, patients with ≥2 surgeries had higher VAS score for pain (P=.02) and stiffness (P=.01).

Patient-Reported Outcomes Measurements Information System-Physical Function: T Score

All TGCT patients had clinically relevant impaired T scores (44.5 and 41.3 for localized and diffuse type, respectively) compared with the general US population (T score of 50). Corresponding effect size was medium for localized type (d=0.55) and large for diffuse type (d=0.87). When comparing both types, diffuse-type patients scored lower (P<.001). In localized type, female patients scored lower (P=.04). Diffuse-type recurrent patients had decreased scores (P=.02).

Table 3. Treatment characteristics of 277 treated tenosynovial giant cell tumor (TGCT) patients.

Treatment	Localized-TGCT[a] (n=67)	Diffuse TGCT (n=210)
Initial surgery, n (%)		
Arthroscopic synovectomy	38 (57)	113 (53.8)
Open synovectomy (one- or two-staged)	26 (39)	90 (42.9)
Combined arthroscopic/open synovectomy	3 (4)	0 (0.0)
Total joint replacement/(tumor) prosthesis	0 (0)	5 (2.4)
Amputation	0 (0)	2 (1.0)
Adjuvant therapy, n (%)	5 (7)	53 (25.2)
Radiotherapy	4 (6)	18 (8.6)
90-Yttrium	1 (1)	14 (6.7)
Systemic	0 (0)	15 (7.1)
Other[b]	0 (0)	6 (2.9)
Recurrent disease, n (%)	24 (36)	146 (69.5)
Additional surgery, n (%)	23 (34)	125 (59.5)
Arthroscopic synovectomy	7 (10)	32 (15.2)
Open synovectomy (one- or two-staged)	10 (15)	74 (35.2)
Combined arthroscopic/open synovectomy	1 (1)	4 (1.9)
Total joint replacement/(tumor) prosthesis	2 (3)	12 (5.7)
Amputation	3 (4)	3 (1.4)

[a] TGCT: tenosynovial giant cell tumor.

[b] Other adjuvant therapies were cryosurgery, burning tools, steroid injections, or combination of multiple adjuvant therapies.

Table 4. Risk factor comparison of 69 localized versus 230 diffuse tenosynovial giant cell tumor (TGCT) of lower extremities.

TGCT[a] -type	Worst pain VAS[b] score, 0 best score and 10 worst score		Worst stiffness VAS score, 0 best score and 10 worst score		PROMIS-PF[c] T score, mean 50 (SD 10), MID[d] 4.0–6.0			SF-12[e], PCS[f] score, mean 50 (SD 10), MID >3.29			SF-12, MCS[g] score, mean 50 (SD 10), MID >3.77			EQ-5D-5L DS[h] utility score, 0 death and 1 full health, MID .037–.069	
	Mean	*P* value	Mean	*P* value	Mean	*d* [i]	*P* value	Mean	*d*	*P* value	Mean	*d*	*P* value	Mean	*P* value
Localized	3.36	.24	3.46	.14	44.5	0.55	*<.001*	40.5	0.95	.08	47.5	0.25	.40	0.76	*.03*
Diffuse	3.79		4.01		41.3	0.87		38.1	1.19		46.3	0.38		0.72	

[a] TGCT: tenosynovial giant cell tumor.

[b] VAS: Visual Analog Scale.

[c] PROMIS-PF: Patient-Reported Outcomes Measurement Information System-Physical Function.

[d] MID: minimal important difference represents the smallest difference or change beyond statistical significance in an outcome measure score that would be considered important by the value patients place on change ref [18-20].

[e] SF: Short-Form.

[f] PCS: physical component summary.

[g] MCS: mental component summary.

[h] DS: descriptive system.

[i] *d*: Cohen *d* or effect size, ratio of difference between 2 means divided by the standard deviation.

Short Form-12 Health Survey: Physical and Mental Component Summary Score

In comparison with general US population (score of 50), both types had impaired PCS (40.5 localized and 38.1 diffuse type) and MSC scores (47.5 localized and 46.3 diffuse type). In all patients in all compared groups, PCS score was clinically relevant declined, in contrast to MCS score which did not transcend the MID threshold in majority of patient groups. A large effect size was calculated for mean PCS scores (0.95 and 1.19 for localized and diffuse type, respectively) and a medium effect size (0.25 and 0.38 for localized and diffuse type, respectively) for MCS scores. In localized type, higher number of surgeries (≥2) affected PCS score negatively (*P*=.03). Localized- and diffuse-type patients who underwent treatment for TGCT ≤1 year ago, showed lower PCS score (*P*=.04 localized, *P*=.01 diffuse). In diffuse type, female patients demonstrated a decreased MCS score (*P*=.04), as well as patients with recurrence of TGCT (*P*=.04).

EuroQoL 5 Dimensions 5 Levels Health Questionnaire: Index Value

All patients, in all groups (Tables 4-6), presented declined EQ5D-5L utility scores compared with full health (1), and all scores transcended MID threshold. Overall, utility score was lower in diffuse patients compared with localized patients (*P*=.03). In localized type, participants with recurrence of TGCT and ≥2 surgeries scored lower (*P*=.01 and *P*=.02, respectively). Similarly, diffuse patients with recurrence had decreased scores (*P*=.02). Median health question VAS score was 75 (IQR 65-85) for localized and 75 (IQR 56.5-85) for diffuse type. No differences between scores calculated with US and UK valuation sets were detected in sensitivity analysis (Multimedia Appendix 4).

Table 5. Risk factor comparison of 69 localized tenosynovial giant cell tumor (TGCT) of lower extremities.

Risk-factors	Worst pain VAS[a] score, 0 best score and 10 worst score		Worst stiffness VAS score, 0 best score and 10 worst score		PROMIS-PF[b] T score, mean 50 (SD 10), MID[c] 4.0-6.0		SF-12[d], PCS[e] score, mean 50 (SD 10), MID >3.29		SF-12, MCS[f] score, mean 50 (SD 10), MID >3.77		EQ-5D-5L DS[g] utility score, 0 death and 1 full health, MID .037-.069	
	Score	*P* value	Score	*P* value	Score	*P* value	Score	*P* value	Score	*P* value	Score	*P* value
Gender												
Male (n=14)	2.93	0.53	3.29	.79	48.5	*.04*	43.3	.23	49.4	.42	.81	.18
Female (n=55)	3.47		3.51		43.5		39.8		47.0		.75	
Age of diagnosis												
<35 years (n=36)	3.36	.997	3.39	.82	44.1	.63	40.1	.74	45.4	.07	.76	.95
≥35 years (n=33)	3.36		3.55		45.0		40.9		49.8		.77	
TGCT localization												
Knee (n=53)	3.04	.08	3.13	.07	44.2	.57	41.0	.43	47.2	.63	.77	.41
Hip, ankle, foot, other (n=16)	4.44		4.56		45.5		38.8		48.6		.74	
Initial surgery												
Arthroscopy (n=38)	3.45	.58	3.26	.63	44.2	.52	40.7	.96	47.2	.73	.76	.92
Open surgery[h] (n=26)	3.04		3.62		45.6		40.8		48.0		.77	
Recurrence												
Yes (n=24)	4.50	*.01*	4.71	*.01*	42.8	.20	38.3	.17	45.7	.27	.70	*.01*
No (n=45)	2.76		2.80		45.4		41.7		48.5		.80	
Total no. of surgeries												
1 surgery (n=44)	2.77	*.02*	2.86	*.01*	45.4	.20	42.2	*.03*	48.4	.19	.79	*.02*
≥2 surgeries (n=23)	4.48		4.65		42.6		36.7		45.0		.71	
Last treatment for TGCT												
≤1 year ago (n=31)	3.77	.26	3.81	.38	42.4	.06	37.6	*.04*	46.0	.34	.74	.29
>1 year ago (n=36)	3.00		3.19		46.1		42.6		48.3		.78	

[a] VAS: Visual Analog Scale.

[b] PROMIS-PF: Patient-Reported Outcomes Measurement Information System-Physical Function.

[c] MID: minimal important difference represents the smallest difference or change beyond statistical significance in an outcome measure score that would be considered important by the value patients place on change ref [18-20].

[d] SF: Short-Form.

[e] PCS: physical component summary.

[f] MCS: mental component summary.

[g] DS: descriptive system.

[h] One- or two staged open synovectomy.

Table 6. Risk factor comparison of 230 diffuse tenosynovial giant cell tumor (TGCT) of lower extremities.

Risk factors	Worst pain VAS[a] score, 0 best score and 10 worst score		Worst stiffness VAS score, 0 best score and 10 worst score		PROMIS-PF[b] T score, mean 50 (SD 10), MID[c] 4.0-6.0		SF-12[d], PCS[e] score, mean 50 (SD 10), MID >3.29		SF-12, MCS[f] score, mean 50 (SD 10), MID >3.77		EQ-5D-5L DS[g] utility score, 0 death and 1 full health, MID .037-.069	
	Score	*P* value	Score	*P* value	Score	*P* value	Score	*P* value	Score	*P* value	Score	*P* value
Gender												
Male (n=51)	3.63	.70	4.13	.71	42.2	.23	39.9	.11	49.0	*.04*	.75	.17
Female (n=179)	3.84		3.98		41.0		37.5		45.6		.71	
Age of diagnosis												
<35 years (n=119)	3.70	.58	3.74	.09	42.1	.07	39.2	.07	46.9	.38	.73	.22
≥35 years (n=109)	3.89		4.32		40.4		36.8		45.6		.71	
TGCT localization												
Knee (n=170)	3.78	.93	3.92	.07	41.6	.29	38.1	.99	46.3	.98	.73	.40
Hip, ankle, foot, other (n=60)	3.82		4.55		40.5		38.1		46.3		.71	
Initial surgery												
Arthroscopy (n=113)	3.93	.82	4.19	.64	41.6	.66	38.3	.86	45.6	.64	.73	.25
Open surgery[h] (n=190)	3.84		4.01		41.2		38.0		46.4		.70	
Recurrence												
Yes (n=146)	4.23	*<.001*	4.55	*<.001*	40.5	*.02*	37.7	.49	45.1	*.04*	.70	*.02*
No (n=84)	3.04		3.08		42.7		38.7		48.2		.75	
Total no. of surgeries												
1 surgery (n=86)	3.79	.69	3.74	.09	42.0	.20	38.7	.48	46.1	.89	.73	.44
≥2 surgeries (n=124)	3.94		4.38		40.8		37.7		46.3		.71	
Last treatment for TGCT												
≤1 year (n=72)	4.10	.37	4.35	.37	40.4	.17	35.4	*.01*	45.6	.59	.70	.17
>1 year ago (n=138)	3.76		4.00		41.8		39.5		46.5		.73	

[a] VAS: Visual Analog Scale.

[b] PROMIS-PF: Patient-Reported Outcomes Measurement Information System-Physical Function.

[c] MID: minimal important difference represents the smallest difference or change beyond statistical significance in an outcome measure score that would be considered important by the value patients place on change ref [18-20].

[d] SF: Short-Form.

[e] PCS: physical component summary.

[f] MCS: mental component summary.

[g] DS: descriptive system.

[h] One- or two staged open synovectomy.

Discussion

Principal Findings

The name of the largest online community of patients with TGCT, PVNS is pants!!, suggests impact on quality of life. One of the community members motivated the name: “Pants is British slang for crap or garbage.” To date, it is unknown what the effect of TGCT on daily living is. A questionnaire was composed in consultation with TGCT patients to determine functional, socioeconomic, and health burden for TGCT patients. We intended to evaluate TGCT in the real world and concluded that TGCTs have a large impact on daily living, with declined health-related quality of life and limitations in daily activities, sports, work, and hobbies: especially the diffuse type of lower extremities and recurrent disease including multiple surgeries.

Limitations

The most important limitation to this study is selection bias. By using crowdsourcing to gather data, it is likely to have a higher number of patients with severe or recurrent diseases [21]. Consequently, when extrapolating these results to generally described populations of TGCT patients in literature, care should be taken not to overestimate the decreased physical function and additional socioeconomic limitations. TGCT usually affects young adults. Since younger patients are more likely to be on the World Wide Web, and our included patient population had a median age of 33 (25-42) years at time of diagnosis, also in concordance with the WHO classification [1,2] and Mastboom et al [3], we considered our participants representative for the heterogeneous disease TGCT. Additionally, the CHERRIES was completed. This checklist provides an understanding of the sample (self-)selection and its possible differences from a representative sample [10] (Multimedia Appendix 1). An additional limitation to this study is that patients in different stages of different treatments were included. To assess comparability within study population, we compared patients who had treatment less than a year ago with performed treatment over a year ago. No positive associations were discovered, except for the SF-12 PCS score in both types. This underlines the postoperative limitations during the first year of follow-up after treatment. As we set out to evaluate impact of TGCT on daily living in the real world heterogeneous TGCT population, the inclusion of patients in different treatment stages matched intention of our study. Furthermore, a known disadvantage of quality of life questionnaires (eg, SF-12) is the generalizability of the questions. Impaired overall quality of life could be attributed to TGCT but also to additional physical abnormalities or psychological problems. Also questionnaires may be completed by patients that have been ill-informed on their disease. In all, 28 patients filled out *unknown type of TGCT*, and 16% of patients who confirmed TGCT with medical proof filled out localized TGCT instead of diffuse TGCT or vice versa. Undeniably, differentiating in localized and diffuse TGCT is challenging even for (un)specialized physicians. The relatively high recurrence rate in this study could also be reflected by unawareness of disease specifics. Recurrence rates in our study were 36% and 70% for localized and diffuse type, compared with on average 4% to 6% (up to a maximum of 50%) and 14% to 40% (up to a maximum of 92%) according to van der Heijden et al [6], respectively. It is conceivable that residual disease or clinical symptoms were filled out as recurrent disease.

The use of self-reported questionnaires harbors the risk of incorrectly answered questions. One could argue that all patients should have been analyzed together, not subdividing into localized and diffuse type. However, differences between two types are major, and therefore separate analyses were necessary for a realistic view of impact of TGCT on daily living.

Crowdsourcing

The presumed definition of crowdsourcing is the practice of obtaining services, ideas, or content by collecting contributions from a comprehensive group from an online community rather than from traditional data suppliers [9]. However, the exact definition for crowdsourcing remains controversial, as 40 definitions originating from 32 unique articles, published between 2006 and 2011, were described by Estellés-Arolas [22]. It is therefore challenging to well define crowdsourcing coherently. After analyses of the 40 (sometimes contrasting) definitions, 8 characteristics common to any given crowdsourcing initiative were found: the crowd, the task at hand, the recompense obtained, the crowdsourcer or initiator of the crowdsourcing activity, what is obtained by them following the crowdsourcing process, the type of process, the call to participate, and the medium. First, in our study, the crowd is presented by patients with TGCT (preferably confirmed by medical reports). Second, the task at hand is completing a questionnaire about the effect of TGCT on daily living. Third, participating in this study was voluntary, therefore no recompenses were offered. Fourth, the initiators of this study are members of the Facebook group PVNS is Pants!! accompanied with the executors, known as the authors of this paper. Fifth, the researchers and subsequently the participants and TGCT patients gain more knowledge on the impact of TGCT on daily living. Sixth, the type of process is an evaluation process, aiming to evaluate effect of TGCT on daily living. Seventh, all patients with TGCT, fluent in English language, were invited to complete the questionnaire. Lastly, the medium Facebook was used to broadcast the questionnaire.

Facebook is the best applicable social network site for survey research, because it is continuously growing, internationally known and exceeds 2 billion users globally (June 2017). The Facebook community PVNS is Pants!!, created in 2009, is the largest TGCT online support group and mainly consists citizens of the United States. On this very active, closed Facebook community, patients are daily updating experiences on their disease, ask for advice from fellow TGCT patients, and comment on other posts to provide their knowledge or sympathy. By actively posting and commenting on research proposals, patients expressed their willingness to participate in research on TGCT. From these posts, we learned that adequate patient information on TGCT is lacking. Our crowdsourcing study stimulated patients' involvement in research and was an opportunity to align research questions with the public's interest [23,24]. TGCT is a rare disease and time to definitive diagnosis is prolonged due to unspecific symptoms and unfamiliarity of the disease [5]. A challenge in studying a rare disease is the lack of big data. Crowdsourcing is an effective and low-cost alternative to traditional methods of participant recruitment due to the possibility to reach large groups of individuals in a relatively short time frame [25]. Van der Heijden et al [9] concluded that crowdsourcing is a promising way for evaluation of rare diseases. Czajka et al [21] used crowdsourcing to efficiently recruit a global cohort and is the largest study on patients with multiple hereditary exostoses. Crosier et al [26] used Facebook to recruit patients with auditory hallucinations; within 6 weeks, over 250 patients had completed this survey. Pohlig et al [27] concluded that enrollment of patients in prospective studies is time-consuming and could be facilitated by use of crowdsourcing.

To obtain a higher level of scientific value, patients were requested for medical proof to ascertain TGCT diagnosis. To our knowledge, no other crowdsourcing studies considered disease confirmation. Patient data and outcome for validated questionnaires were comparable for patients with and without medical proof. Patients were not uniformly diagnosed and treated as they originated from 30 different countries globally. Neither was distinguished between treatment in peripheral or tertiary referral centers. Nevertheless, we consider our study group a reflection of the current worldwide situation and believe that declined impact on daily living is clinically relevant for all patients. In contrast to malignant diseases, survival rates are not of interest for TGCT with its benign character. According to high recurrence rates, quality of life (prior and after treatment) is essential to evaluate.

Patient-Reported Outcome Measures

Patient reported outcome measures (PROMs) are increasingly used in health policy, patient-centered care, and shared clinical decision making [28]. In the era of personalized medicine, patient involvement is increasing in shared decision making for different treatment strategies with functional outcome and quality of life. In our study, members of the largest online TGCT community were involved in establishing the questionnaire *Evaluation of TGCT on daily living*.

Functional outcome and health-related quality of life are only spars reported for TGCT. Four studies have reported on standardized PROMS [4,9,29,30]. Currently, validated PROMS for TGCT patients do not exist. In accordance with Gelhorn et al [4], VAS for worst pain and stiffness and PROMIS-PF questionnaires were used. Conform van der Heijden et al [9,29] and Verspoor et al [30], the SF-12, a quality of life questionnaire, was included, known as the shorter version of the SF-36. One study identified a high health care burden with a significant increase in health care costs, ambulatory costs, and physical therapy in 9328 TGCT patients [8].

In benign diseases, including TGCT, death is not an outcome variable. Besides tumor reduction, critical endpoint measures are clinical relevance and impact of treatment. Currently, clinical TGCT studies lack specific and validated PROMs to document treatment-induced symptomatic, functional, and economic (back to work) improvement [31]. To obtain an impression of physical function and quality of life in TGCT patients, participants in our study were requested to complete different validated questionnaires. In our experience, PROMIS-PF was most useful in determining these functional factors. To minimize the multitude of questions and include the most important components for clinical TGCT studies, we would propose a combination of PROMIS-PF and a short quality of life questionnaire, for instance EQ5D5L, in clinical practice.

Risk Factors for Deteriorated Outcome

Risk factors for deteriorated outcome in our study were diffuse-type TGCT, recurrent disease, and ≥2 surgeries performed. This is in concordance with current literature on risk factors for a high recurrence rate. According to the necessity of mutilating surgeries to treat recurrences, we considered risk factors for recurrent disease comparable to risk factors for deteriorated outcome.

Higher recurrence rate in diffuse TGCT compared with localized TGCT is exuberant described [2,5-7,30,32-34]. Bruns et al [34] described 173 patients treated in 10 orthopedic departments in Germany and Austria and reported higher recurrence rates in institutions treating less than 20 cases for TGCT, in diffuse disease, in the hip joint and after arthroscopy. Schwartz et al [35] described 99 patients with TGCT in the knee, hip, elbow, or shoulder. They concluded that localization in the knee, previous surgical procedures, and incomplete synovectomy were related significantly to higher number of subsequent recurrences. On the basis of current literature and to investigate possible risk factors for recurrent disease thoroughly, gender, age at time of diagnosis, TGCT localization, initial surgery, presence of recurrence, total number of surgeries, and time since last treatment for TGCT, were compared.

Conclusions

TGCTs have major impact on daily living in a relatively young, working population (median age at diagnosis, 33 years). Majority of symptoms improve after treatment, however, symptoms remain in about half of the TGCT patients; especially in patients with diffuse type, recurrent disease, and ≥2 surgeries. The high recurrence rate in diffuse TGCT results in clinically important deteriorated outcome in physical function and health-related quality of life. In preventing recurrent disease, and its deteriorated outcome, an extensive mutilating surgery might be necessary. Physicians should be aware that TGCT patients frequently experience symptoms and limitations in daily life and societal participation (work, sports, and hobbies), even after treatment(s). We deem it important for future research to evaluate treatment, including its effectiveness on improving quality of daily living. With this study, we hope to increase knowledge on TGCT among treating physicians, highlight the importance of quality of life, and to offer research-based information to patients.

Acknowledgments

The authors would like to thank all 5 TGCT patients of our outpatient clinic for their feedback in the pilot test to finalize the questionnaire. Sheila Piner, one of the most active members of PVNS is pants!!, was very helpful in advertising and motivating to complete the questionnaire. Administrators of PVNS is Pants!! have facilitated the gathering of participants for our study by accepting our posts. Lucianne Remijn was of great help in data gathering; Professor Dr TPM Vliet Vlieland for reviewing the first draft; CB Terwee, H van Dongen, and CS Leichtenberg for their advice and help regarding the used validated questionnaires. Finally, the authors would like to thank all the members of the Facebook group who have completed our questionnaire and enabled this study. No funding was received for this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Data reporting guidelines, checklist for reporting results of internet E-Surveys (CHERRIES).

[PDF File (Adobe PDF File), 41KB - ijmr_v7i1e4_app1.pdf]

Multimedia Appendix 2

Facebook introduction, invitation to complete questionnaire and thank you message.

[PDF File (Adobe PDF File), 259KB - ijmr_v7i1e4_app2.pdf]

Multimedia Appendix 3

Questions e-survey "Evaluation of TGCT on daily living".

[PDF File (Adobe PDF File), 66KB - ijmr_v7i1e4_app3.pdf]

Multimedia Appendix 4

Supplementary material—sensitivity analysis.

[PDF File (Adobe PDF File), 32KB - ijmr_v7i1e4_app4.pdf]

Multimedia Appendix 5

Comparing patients with medical proof to patients without medical proof.

[PDF File (Adobe PDF File), 46KB - ijmr_v7i1e4_app5.pdf]

References

1. de St. Aubain S, van de Rijn M. Tenosynovial giant cell tumour, localized type. In: Fletcher CD, Bridge JA, Hogendoorn PC, Mertens F, editors. WHO Classification of Tumours of Soft Tissue and Bone. 4 edition. Lyon, France: IARC Press; 2013.
2. de St. Aubain S, van de Rijn M. Tenosynovial giant cell tumour, diffuse type. In: Fletcher CD, Bridge JA, Hogendoorn PC, Mertens F, editors. WHO Classification of Tumours of Soft Tissue and Bone. Lyon, France: lARC Press; Apr 30, 2013.
3. Mastboom MJ, Verspoor FG, Verschoor AJ, Uittenbogaard D, Nemeth B, Mastboom WJ, TGCT study group. Higher incidence rates than previously known in tenosynovial giant cell tumors. Acta Orthop 2017 Aug 08;88(6):688-694. [doi: 10.1080/17453674.2017.1361126] [Medline: 28787222]
4. Gelhorn HL, Tong S, McQuarrie K, Vernon C, Hanlon J, Maclaine G, et al. Patient-reported Symptoms of tenosynovial giant cell tumors. Clin Ther 2016 Dec;38(4):778-793 [FREE Full text] [doi: 10.1016/j.clinthera.2016.03.008] [Medline: 27041409]
5. Stephan SR, Shallop B, Lackman R, Kim TW, Mulcahey MK. Pigmented villonodular synovitis: a comprehensive review and proposed treatment algorithm. JBJS Rev 2016 Jul 19;4(7). [doi: 10.2106/JBJS.RVW.15.00086] [Medline: 27509331]
6. van der Heijden L, Gibbons CL, Hassan AB, Kroep JR, Gelderblom H, van Rijswijk CS, et al. A multidisciplinary approach to giant cell tumors of tendon sheath and synovium--a critical appraisal of literature and treatment proposal. J Surg Oncol 2013 Mar;107(4):433-445. [doi: 10.1002/jso.23220] [Medline: 22806927]
7. Gouin F, Noailles T. Localized and diffuse forms of tenosynovial giant cell tumor (formerly giant cell tumor of the tendon sheath and pigmented villonodular synovitis). Orthop Traumatol Surg Res 2017 Feb;103(1S):S91-S97. [doi: 10.1016/j.otsr.2016.11.002] [Medline: 28057477]
8. Burton T, Ye X, Parker E, Bancroft T, Healey J. Burden of illness in patients with tenosynovial giant cell tumors. 2017 Presented at: ISPOR (International Society for Pharmacoeconomics and Outcomes Research) 22nd Annual International Meeting; May 2017; Boston, MA, USA URL: https://www.ispor.org/ScientificPresentationsDatabase/Presentation/70633
9. van der Heijden L, Piner SR, van de Sande MA. Pigmented villonodular synovitis: a crowdsourcing study of two hundred and seventy two patients. Int Orthop 2016 Dec;40(12):2459-2468. [doi: 10.1007/s00264-016-3208-1] [Medline: 27169531]
10. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
11. Facebook. TGCT Study Facebook Webpage URL: https://www.facebook.com/unsupportedbrowser [accessed 2018-02-10] [WebCite Cache ID 6x7xTGZlj]
12. HealthMeasures. 2017. PROMIS (Patient-Reported Outcomes Measurement Information System) Physical Function Scoring Manual URL: http://www.healthmeasures.net/images/PROMIS/manuals/PROMIS_Physical_Function_Scoring_Manual.pdf [accessed 2018-02-16] [WebCite Cache ID 6xHGRU9le]

13. Ware JE, Kosinski M, Keller SD. Researchgate. 1995. SF-12: How to Score the SF-12 Physical and Mental Health Summary Scales (Second Edition) URL: https://www.researchgate.net/profile/John_Ware/publication/291994160_How_to_score_SF-12_items/links/58dfc467a6fdcc41bf920748/How-to-score-SF-12-items.pdf [accessed 2018-02-16] [WebCite Cache ID 6xHGbcrLN]
14. van Reenen M, Janssen B. Euroqol. 2015 Apr. EQ-5D-5L User Guide: Basic information on how to use the EQ-5D-5L instrument. Version 2.1 ed URL: https://euroqol.org/wp-content/uploads/2016/09/EQ-5D-5L_UserGuide_2015.pdf [accessed 2018-02-16] [WebCite Cache ID 6xHG7nFyg]
15. van Hout B, Janssen MF, Feng YS, Kohlmann T, Busschbach J, Golicki D, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. Value Health 2012;15(5):708-715 [FREE Full text] [doi: 10.1016/j.jval.2012.02.008] [Medline: 22867780]
16. Whitehead SJ, Ali S. Health outcomes in economic evaluation: the QALY and utilities. Br Med Bull 2010;96:5-21. [doi: 10.1093/bmb/ldq033] [Medline: 21037243]
17. Sawilowsky SS. New effect size rules of thumb. J Mod App Stat Meth 2009 Nov 01;8(2):597-599. [doi: 10.22237/jmasm/1257035100]
18. McClure NS, Sayah FA, Xie F, Luo N, Johnson JA. Instrument-defined estimates of the minimally important difference for EQ-5D-5L index scores. Value Health 2017 Apr;20(4):644-650. [doi: 10.1016/j.jval.2016.11.015] [Medline: 28408007]
19. Yost KJ, Eton DT, Garcia SF, Cella D. Minimally important differences were estimated for six patient-reported outcomes measurement information system-cancer scales in advanced-stage cancer patients. J Clin Epidemiol 2011 May;64(5):507-516 [FREE Full text] [doi: 10.1016/j.jclinepi.2010.11.018] [Medline: 21447427]
20. Díaz-Arribas MJ, Fernández-Serrano M, Royuela A, Kovacs FM, Gallego-Izquierdo T, Ramos-Sánchez M, et al. Minimal clinically important difference in quality of life for patients with low back pain. Spine (Phila Pa 1976) 2017;42(24):1908-1916. [doi: 10.1097/BRS.0000000000002298] [Medline: 28658040]
21. Czajka CM, DiCaprio MR. What is the proportion of patients with multiple hereditary exostoses who undergo malignant degeneration? Clin Orthop Relat Res 2015 Jul;473(7):2355-2361 [FREE Full text] [doi: 10.1007/s11999-015-4134-z] [Medline: 25582066]
22. Estellés-Arolas E, González-Ladrón-de-Guevara F. Towards an integrated crowdsourcing definition. J Inf Sci 2012 Mar 09;38(2):189-200. [doi: 10.1177/0165551512437638]
23. Truitt AR, Monsell SE, Avins AL, Nerenz DR, Lawrence SO, Bauer Z, et al. Prioritizing research topics: a comparison of crowdsourcing and patient registry. Qual Life Res 2018 Jan;27(1):41-50. [doi: 10.1007/s11136-017-1566-9] [Medline: 28382522]
24. Bartek MA, Truitt AR, Widmer-Rodriguez S, Tuia J, Bauer ZA, Comstock BA, et al. The Promise and Pitfalls of Using Crowdsourcing in Research Prioritization for Back Pain: Cross-Sectional Surveys. J Med Internet Res 2017 Oct 06;19(10):e341 [FREE Full text] [doi: 10.2196/jmir.8821] [Medline: 28986339]
25. Kuang J, Argo L, Stoddard G, Bray BE, Zeng-Treitler Q. Assessing pictograph recognition: a comparison of crowdsourcing and traditional survey approaches. J Med Internet Res 2015 Dec;17(12):e281 [FREE Full text] [doi: 10.2196/jmir.4582] [Medline: 26678085]
26. Crosier BS, Brian RM, Ben-Zeev D. Using Facebook to reach people who experience auditory hallucinations. J Med Internet Res 2016;18(6):e160 [FREE Full text] [doi: 10.2196/jmir.5420] [Medline: 27302017]
27. Pohlig F, Lenze U, Muhlhofer HM, Lenze FW, Schauwecker J, Knebel C, et al. IT-based psychosocial distress screening in patients with sarcoma and parental caregivers via disease-specific online social media communities. In Vivo 2017;31(3):443-450 [FREE Full text] [doi: 10.21873/invivo.11081] [Medline: 28438877]
28. Gagnier JJ. Patient reported outcomes in orthopaedics. J Orthop Res 2017 Oct;35(10):2098-2108. [doi: 10.1002/jor.23604] [Medline: 28513993]
29. van der Heijden L, Mastboom MJ, Dijkstra PD, van de Sande MA. Functional outcome and quality of life after the surgical treatment for diffuse-type giant-cell tumour around the knee: a retrospective analysis of 30 patients. Bone Joint J 2014 Aug;96-B(8):1111-1118. [doi: 10.1302/0301-620X.96B8.33608] [Medline: 25086129]
30. Verspoor FG, Zee AA, Hannink G, van der Geest IC, Veth RP, Schreuder HW. Long-term follow-up results of primary and recurrent pigmented villonodular synovitis. Rheumatology (Oxford) 2014 Nov;53(11):2063-2070. [doi: 10.1093/rheumatology/keu230] [Medline: 24917565]
31. Cassier PA, Italiano A, Gomez-Roca CA, Le Tourneau C, Toulmonde M, Cannarile MA, et al. CSF1R inhibition with emactuzumab in locally advanced diffuse-type tenosynovial giant cell tumours of the soft tissue: a dose-escalation and dose-expansion phase 1 study. Lancet Oncol 2015 Aug;16(8):949-956. [doi: 10.1016/S1470-2045(15)00132-1] [Medline: 26179200]
32. Chiari C, Pirich C, Brannath W, Kotz R, Trieb K. What affects the recurrence and clinical outcome of pigmented villonodular synovitis? Clin Orthop Relat Res 2006 Sep;450:172-178. [doi: 10.1097/01.blo.0000224051.01873.fb] [Medline: 16760805]
33. Ehrenstein V, Andersen SL, Qazi I, Sankar N, Pedersen AB, Sikorski R, et al. Tenosynovial giant cell tumor: incidence, prevalence, patient characteristics, and recurrence. A registry-based cohort study in Denmark. J Rheumatol 2017;44(10):1478-1483. [Medline: 28765257]

34. Bruns J, Ewerbeck V, Dominkus M, Windhager R, Hassenpflug J, Windhagen H, et al. Pigmented villo-nodular synovitis and giant-cell tumor of tendon sheaths: a binational retrospective study. Arch Orthop Trauma Surg 2013 Aug;133(8):1047-1053. [doi: 10.1007/s00402-013-1770-1] [Medline: 23681468]
35. Schwartz HS, Unni KK, Pritchard DJ. Pigmented villonodular synovitis. A retrospective review of affected large joints. Clin Orthop Relat Res 1989 Oct(247):243-255. [Medline: 2791393]

Abbreviations

ACL: anterior cruciate ligament
CHERRIES: Checklist for Reporting Results of Internet E-surveys
CME: Commissie Medische Ethiek (Institutional Review Board from the LUMC)
DS: descriptive system
EQ-5D-5L: EuroQoL 5 Dimensions 5 Levels
IP: Internet protocol
IQR: interquartile range
JMIR: Journal of Medical Internet Research
LUMC: Leiden University Medical Center
MCS: mental component summary
MID: minimal important difference
NetQ: NetQuestionnaire
PCS: physical component summary
PROMIS-PF: Patient-Reported Outcomes Measurements Information System-Physical Function
PROMS: Patient Reported Outcome Measures
PVNS: pigmented villonodular synovitis
SF-12: Short Form-12 Health Survey
SPSS: Statistical Package for the Social Sciences
TGCT: tenosynovial giant cell tumor
VAS: Visual Analog Scale

Edited by G Eysenbach; submitted 01.11.17; peer-reviewed by X Ye, C Kurz; comments to author 14.12.17; revised version received 21.12.17; accepted 21.12.17; published 23.02.18.

Please cite as:
Mastboom MJ, Planje R, van de Sande MA
The Patient Perspective on the Impact of Tenosynovial Giant Cell Tumors on Daily Living: Crowdsourcing Study on Physical Function and Quality of Life
Interact J Med Res 2018;7(1):e4
URL: http://www.i-jmr.org/2018/1/e4/
doi:10.2196/ijmr.9325
PMID:29475829

Original Paper

Factors Affecting Bone Mineral Density Among Snowy Region Residents in Japan: Analysis Using Multiple Linear Regression and Bayesian Network Model

Teppei Suzuki[1], MHS, MBA; Tomoko Shimoda[1], RN, MSN; Noriko Takahashi[1], RN, MSN; Kaori Tsutsumi[1], PhD; Mina Samukawa[1], PhD; Sadako Yoshimura[1], PhD; Katsuhiko Ogasawara[1], MBA, PhD

Faculty of Health Sciences, Hokkaido University, Sapporo, Japan

Corresponding Author:
Katsuhiko Ogasawara, MBA, PhD
Faculty of Health Sciences
Hokkaido University
Kita 12, Nishi 5, Kita-ku
Sapporo,
Japan
Phone: 81 11 706 3409
Email: oga@hs.hokudai.ac.jp

Abstract

Background: As the onset of osteoporosis leads to reduced activities of daily living and may result in patients being bedridden, efforts to prevent decreased bone density are necessary. Various studies on the relationship between sex, age, nutrients, and exercise habits and bone mineral density have been conducted to date. However, for snowy region residents, the magnitude of influence of various factors affecting bone mineral density and the influence level have not been clarified.

Objective: This study aimed to clarify the degree of influence and factors influencing bone mineral density based on survey results on health conditions and lifestyle habits in heavy snow areas.

Methods: A total of 354 citizens who visited a drugstore in the target area were included in a study that included using the brief-type self-administered diet history questionnaire on lifestyle and exercise habits. Height, weight, body composition, and bone densitometer values were analyzed using multiple regression to calculate their association with bone mineral density. In addition, a Bayesian network model was used to determine the influence level of each factor as a conditional probability.

Results: Multiple regression analysis revealed that age, sex, fracture, and calcium intake significantly influenced bone mineral density. In addition, the result of Bayesian network analysis suggested that age and sex affected bone mineral density, whereas nutrients and exercise habits might not have a direct impact. However, calcium intake and the T-score were significant factors affecting the presence or absence of fracture experiences, suggesting that adequate calcium intake is essential for preventing fractures.

Conclusions: In the multiple regression analysis, age, sex, fracture, and calcium intake were selected as factors; however, in the Bayesian analysis, only age and sex affected bone mineral density while nutrients did not. In addition, the fact that calcium intake and the T-score were shown to affect bone fracture history suggests that calcium intake is an important measure that can prevent bone fractures. Overall, these results suggest that measures such as ensuring a bone fracture–free environment and providing nutritional advice for calcium intake can be effective in preventing bone loss.

(Interact J Med Res 2018;7(1):e10) doi:10.2196/ijmr.8555

KEYWORDS

health care promotion; Bayesian network; health behavior change

Introduction

Background

The total population of Japan as of October 1, 2015, was 127.11 million. Seniors aged 65 years and older accounted for 26.7% (33.9 million) of the total population [1]. As the elderly population grows, nursing care problems increase. Approximately half of all individuals requiring nursing care need it for reasons including "weakening due to age," "joint disease," and "fractures or falls." Seniors are particularly prone to falls, which occur at least once annually in approximately 20% of seniors, 5% of which result in fractures. Therefore, fractures and falls should be prevented in a rapidly aging society [2]. Fractures caused by osteoporosis require nursing care and may reduce the quality of life in severe cases due to the patient becoming bedridden. To prevent osteoporosis, it is essential to achieve maximum bone mass while people are still young, and as people age, they should be mindful of their diet and exercise habits to prevent bone loss. Therefore, lifestyle-related factors such as nutrition, exercise, and muscle mass are believed to affect bone density [3].

This study was conducted in Iwamizawa City, located in midwestern Hokkaido, approximately 40 km from Sapporo City, the prefectural capital. One of its most notable characteristics is its heavy snowfall, with a very large average annual snowfall (337 cm), and the deepest single snowfall (133 cm) in 2011. Iwamizawa City experiences more snowfall than Sapporo City and approximately twice or more than that in Kushiro City and Hakodate City, making it the region with the heaviest snowfall in Hokkaido.

In regions with heavy snowfall, residents tend to exercise less and are highly at risk for bone fractures due to the snow- and ice-covered roads. Another characteristic of Iwamizawa City is that daylight hours are shorter than that of other regions in the country. The annual amount of daylight across Japan is 1978 hours compared with 1700 hours in Iwamizawa City. Humans need ultraviolet radiation in order to synthesize most of their vitamin D. The primary action of vitamin D is the promotion of calcium absorption in the small intestine, and reduced vitamin D activity inhibits calcium absorption. Persistently low calcium levels can impair bone calcification and reduce bone density [4]. In a previous study conducted in Iwamizawa City, Shimoda et al [5] investigated the body composition and nutritional intake of elderly residents in regions with heavy snowfall. They found that the basal metabolic rate and muscle mass decreased with age, although calories and nutritional intake remained sufficient, suggesting a limited effect of diet.

Last, Iwamizawa City has a high population aging rate at 32.4% as of January 1, 2016; a value approximately 6% higher than that of the whole country. Therefore, it is considered a geographical region with many residents suspected of having osteoporosis [6].

In earlier studies on osteoporosis in the region, Uchida et al [7] reported that measures preventing bone fractures are important in a society with an increasing number of bedridden patients and that maximum bone mass, which peaks between 20 and 30 years, is greatly affected by genetics. Therefore, Uchida et al investigated the factors affecting bone mineral density and found that environmental factors such as lifestyle were strongly associated with bone density. They proposed that studying the relationship between bone density and environmental factors such as nutrition is important from the standpoint of prophylactic medicine. Toda et al [8] conducted a study involving 654 healthy Japanese women aged 40 to 60 years to investigate the lumbar bone mineral density, anthropometric measurements, and 3-day meal, and a survey was also conducted to determine the effects of diet and physical activity on bone density, with the purpose of preventing osteoporosis. The results showed that appropriate exercise and a well-balanced diet were important factors in preventing osteoporosis. However, previous studies did not target regions with heavy snowfall nor did they provide details on nutritional intake. In addition, they did not assess the effects of various nutrients on bone density or whether an influence level existed.

Objectives

This study aimed to determine how nutritional intake, lifestyle and exercise habits, and body composition information affect bone density in heavy snowfall/cold regions and assess the extent of the effect on the related factors.

Methods

Participants and Data Collection

The data used in our study were obtained from a study conducted by a Center of Innovation program entitled, "Study of health status for the purpose of health innovation in Iwamizawa City." A total of 354 individuals (87 men and 267 women) who visited the city drugstore participated in the study (Table 1).

Measurement

The survey activities included completing a questionnaire related to lifestyle and exercise habits, completing the brief, self-administered diet history questionnaire so that daily nutritional intake could be determined, and measuring the body composition and bone density. In the nutrition survey, the brief-type self-administered dietary history questionnaire (BDHQ 58-item fixed-portion-type questionnaire) was used [9]. The equipment used consisted of an ultrasonic calcaneus measurement device, A-1000 EXPII (GE Healthcare Japan), for bone density measurements and a WELL-SCAN 900 (Canon Inc) for body composition measurements.

Data Analysis

Factors that largely affect bone density that were extracted as explanatory variables for the study analysis were bone fracture history; moderate exercise (such as sports or recreation that increase the heart rate for approximately 10 minutes); calcium, vitamin D, and cholesterol intakes per day (through meal); sex; age; body mass index; muscle mass; and degree of obesity (Table 2). A multiple regression analysis using bone density (the T-score) as the criterion variable was performed to analyze the effects of each explanatory variable on bone density. Statistical analyses were performed using JMP Pro version 12.2.0 (SAS Institute Inc).

Next, a Bayesian network was established to determine the influence level between bone density and each explanatory variable. It is a means of expressing the influence level of a target event as a graph model. Bayesian networks are expressed as noncircular directed graphs with directionality links that are interpreted as the direction of the causes and effects of events. The influence level between events is obtained by the conditional probability of a linked root event. Conditional probability is generally expressed as P(A|B) and represents the probability of event A occurring given that event B has occurred. A Bayesian network is a technology frequently used in the field of medicine. Velikova et al [10] designed a Bayesian network model that accounts for the changes of preeclampsia, a type of pregnancy-related disorder, for a certain period of time, and they also proposed a system that supported decision making at the individual level [10-13].

In this study, a Bayesian network was established using bone density (T-score) and explanatory variables extracted from a multiple regression analysis. Table 3 shows the criteria for the T-score. The influence level between each explanatory variable and bone density was determined, the conditional probability (prior probability) was calculated, and the factors affecting bone density status were analyzed. In addition, a sensitivity analysis, in which the parameters of each factor were changed to obtain the posterior probability, was performed, and the combination with the highest probability of osteoporosis or osteopenia occurrence was studied. The variables used in the multiple regression analysis were also used as factors in the Bayesian network model. The Bayesian network was established using BAYOLINK version 7.0.0 (NTT DATA Mathematical Systems Inc).

Table 1. Characteristics of study population by sex and age.

Characteristics	n (%)
Sex	
Women	267 (75.4)
Men	87 (24.6)
Age, years	
20-29	12 (3.4)
30-39	26 (7.3)
40-49	24 (6.8)
50-59	69 (19.5)
60-69	136 (38.4)
70-79	76 (21.5)
80+	11 (3.1)

Table 2. Variables with their values as used in the Baysian network model.

Factors	Survey method and contents	Values and ranges
Fracture	Questionnaire: fracture experience	No, yes
Exercise	Questionnaire: medium strength sports that last ≥10 min	No, yes
Calcium	Brief-type dietary history questionnaire	Adequate, slightly inadequate[a], inadequate[a]
Vitamin D	Brief-type dietary history questionnaire	Adequate, slightly adequate[a], inadequate[a]
Cholesterol	Brief-type dietary history questionnaire	Adequate, slightly high[a], high[a]
Sex	Body composition meter	Female, male
Age, years	Body composition meter	<65, ≥65[a]
BMI[c], kg/m^2	Body composition meter	Normal (<25), obesity[a] (≥25)
Muscle mass, kg	Body composition meter	<37.0, 37-44.9, ≥45.0[a]
Obesity, %	Body composition meter	Underweight (<–10), normal (–10 to 10), overweight[a] (10 to 20), obesity[a] (≤20)

[a] Risk factor.

[b] BMI: body mass index.

Table 3. Criteria for bone density status using T-score.

Bone density status	T-score
Normal	–1.0 and higher
Osteopenia	–1.0 to –2.5
Osteoporosis	–2.5 and lower

Figure 1. Formulas of precision, recall and F-measure.

$$\mathrm{Precision} = \frac{correct\ possitive\ predictions}{positive\ predictions} \quad (1)$$

$$\mathrm{Recall} = \frac{correct\ positive\ predictions}{positive\ examples} \quad (2)$$

$$\mathrm{F-measure} = \frac{2 * Precision * Recall}{Precision + Recall} \quad (3)$$

Precision, recall, and F-measure were used as indicators to evaluate the results of the Bayesian network. Precision represents data that match the verification value among the data predicted as the target state by reasoning. Recall represents the proportion of data coincident with the estimated value among the target state in the verification data. F-measure represents the harmonic mean of precision and recall.

Mutual information amount represents the changes in the amount of model entropy before and after inputting the observation. As the mutual information amount increases, the influence on bone density, which is the objective variable, increases (Figure 1).

Ethical Considerations

This study was conducted with the approval of the Ethics Review Committee of the Graduate School of Health Sciences, Hokkaido University (approval number 15-96). For subject recruitment, we presented government-distributed public magazines and information on research cooperation at the drugstore where the survey was conducted. We also explained that participation in the research was voluntary and that there were no disadvantages in declining the participation midway, and we obtained written consent. Data obtained were not used for any purpose other than research and were stored carefully.

Results

Table 4 shows the participants' bone density status based on sex, revealing that 43 men had normal bone density, 36 had suspected osteopenia, and 8 had suspected osteoporosis, while 117 women had normal bone density, 112 had suspected osteopenia, and 38 had suspected osteoporosis. Based on these results, approximately half of both men and women were healthy, and the other half were suspected to have osteopenia or osteoporosis Table 5.

Table 6 shows the results of multiple regression analysis. The fewer fractures men had when they were young and the higher their calcium intake, the higher their bone density tended to be. Nutrition, exercise habits, and muscle mass, which have been thought to affect bone density, were not selected for the model.

Figure 2 shows the Bayesian network in determining the influence level between bone density and each associated factor. The Bayesian network model expresses random variables as nodes and links to nodes with probabilistic dependencies. In the result shown in Figure 2, the T-score represents the state of bone density depending on sex and age. The T-score also shows its direct influence on the presence or absence of fracture. For the value of each factor, the prior probability is shown. The influence level between exercise or muscle mass and the T-score could not be confirmed. In addition, regarding vitamin D, although an influence level with calcium intake was confirmed, a direct influence level with the T-score cannot be confirmed.

Figure 3 shows a Bayesian network model limited to the effect on bone density using variables extracted from the results of multiple regression analysis. The T-score had an influence level with sex and age and has been affected by the history of bone fracture. In addition, an influence level between calcium intake and age was newly discovered.

Table 7 shows the results of estimated accuracy validation of the model, with the T-score as the criterion variable. When bone density was normal, prediction accuracy was highest. In the case of osteoporosis, prediction accuracy was lowest. Table 8 shows the mutual information of each variable, with the T-score as the criterion variable. Since the mutual information of explanatory variable "age" is the largest, the result that age has the largest influence on the state of bone density was obtained.

Table 4. Measurement results.

Characteristics	Male, n	Female, n	Total, n
Body mass index, kg/m^2			
Normal (<25)	57	213	270
Obesity (≥25)	28	42	70
Muscle mass, kg			
<37	2	180	182
37 to 45	28	69	97
≥45	55	6	61
Obesity, %			
Underweight (<–10)	8	60	68
Normal (–10 to 10)	43	138	181
Overweight (10 to 20)	15	31	46
Obesity (≥20)	19	26	45
Calcium			
Adequate	46	141	187
Slightly inadequate	9	40	49
Inadequate	32	77	109
Vitamin D			
Adequate	84	244	328
Slightly inadequate	0	9	9
Inadequate	3	5	8
Cholesterol			
Adequate	79	233	312
Slightly high	4	22	26
High	4	3	7
Fracture			
Yes	17	32	49
No	65	215	280
Exercise			
Yes	76	219	295
No	6	28	34

Table 5. Bone density by sex (χ^2=1.742, *P*=.418).

Bone density	Male (n=87), n (%)	Female (n=267), n (%)
Normal	43 (49)	117 (44.8)
Osteopenia	36 (41)	112 (41.9)
Osteoporosis	8 (9)	38 (14.2)

Table 6. Result of multiple regression analysis.

Variable	Regression coefficient	*P* value
Age	–0.67	<.001
Sex	–12.44	<.001
Fracture	–1.55	.009
Calcium	0.01	.027

Figure 2. Prior probability by Bayesian network.

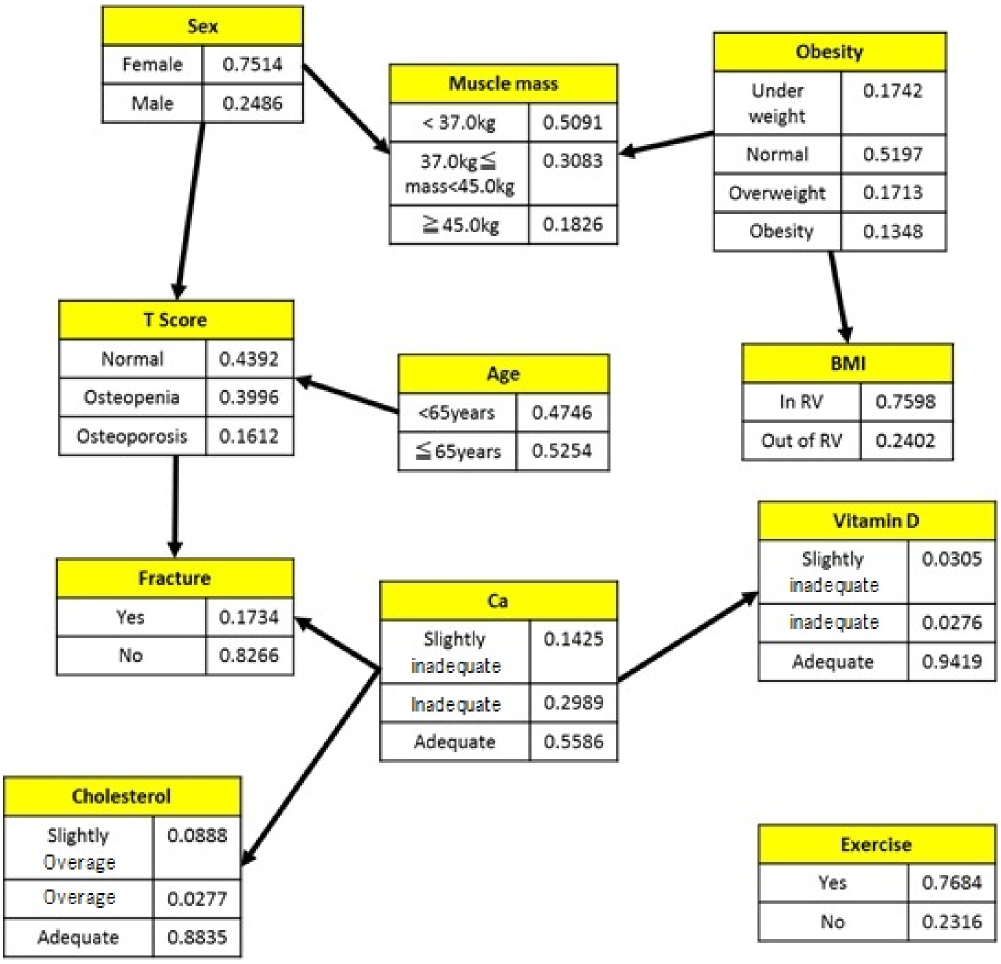


Figure 3. Bayesian network model of factors influencing bone density.

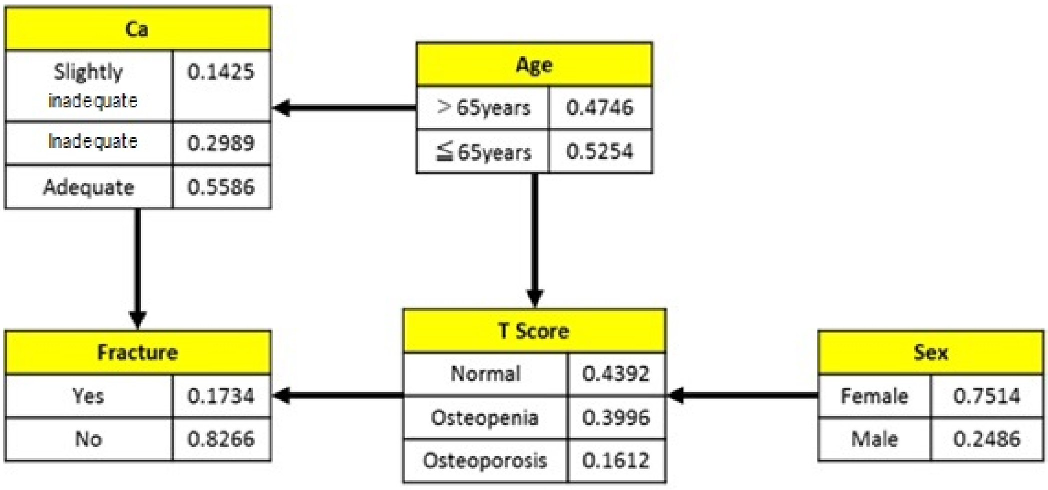


Table 7. T-score prediction accuracy.

Bone density	Precision	Recall	F-measure
Normal	0.56	0.68	0.61
Osteopenia	0.48	0.49	0.48
Osteoporosis	0.5	0.14	0.22

Table 8. Mutual information.

Variable	Mutual information
Age	0.065
Fracture	0.014
Sex	0.008
Calcium	0.002

Discussion

Multiple Regression Analysis Results

The explanatory variables selected for the model were age, sex, history of bone fracture, and calcium intake, but only age and sex were significant and are not something that an individual can control. However, in relation to bone fracture prevention and calcium intake, bone density can be increased by changing one’s awareness of daily lifestyle factors. According to Kubota et al [14], who studied approaches in preventing bone fracture and osteoporosis through calcium intake, one method to increase calcium intake is to take calcium supplements or consume calcium through food, which was proven to effectively reduce bone fracture risk due to osteoporosis in elderly populations. The degree to which bone density reduction can be prevented by supplements or calcium intake must be studied in the future.

Based on the results of this study, sex and age greatly affect bone density, and factors such as vitamin D and exercise habits, believed to affect bone density, were not included in the analysis. Kubota et al [14] reported that the coadministration of calcium and vitamin D in elderly men and women prevented bone loss in the femoral neck and throughout the whole body. In addition, a study by Matsubara et al [15] suggested the possibility that the rate of bone mineral loss throughout the whole body could be prevented in menopausal women. Therefore, physical exercise and nutrients such as vitamin D must be incorporated into the model, and the results obtained in the analysis must be considered.

Bayesian Network Analysis Results

The influence level between each factor was determined based on the results obtained through Bayesian network analysis, which are shown in Figures 2 and 3. These results showed that only age and sex affected bone density, and calcium intake and the T-score affected the occurrence of bone fractures; however, calcium intake was the most important measure that can be implemented to prevent bone fractures.

A sensitivity analysis performed on the prior probability obtained in the Bayesian network analysis revealed that several factors contributed to the highest probability of developing osteoporosis and osteopenia, as shown in Figures 4 and 5.

Figure 4. The combination with the highest probability of osteoporosis.

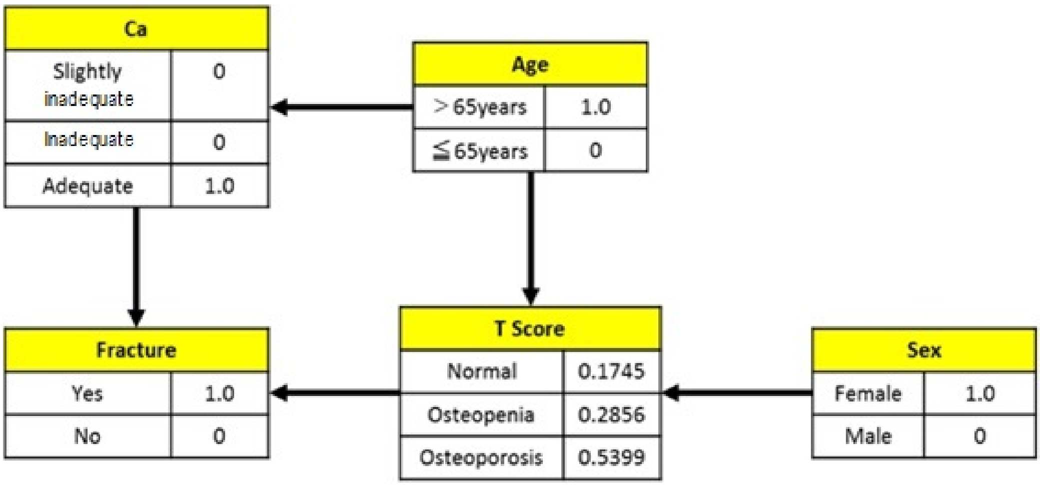


Figure 5. The combination with the highest probability of osteopenia.

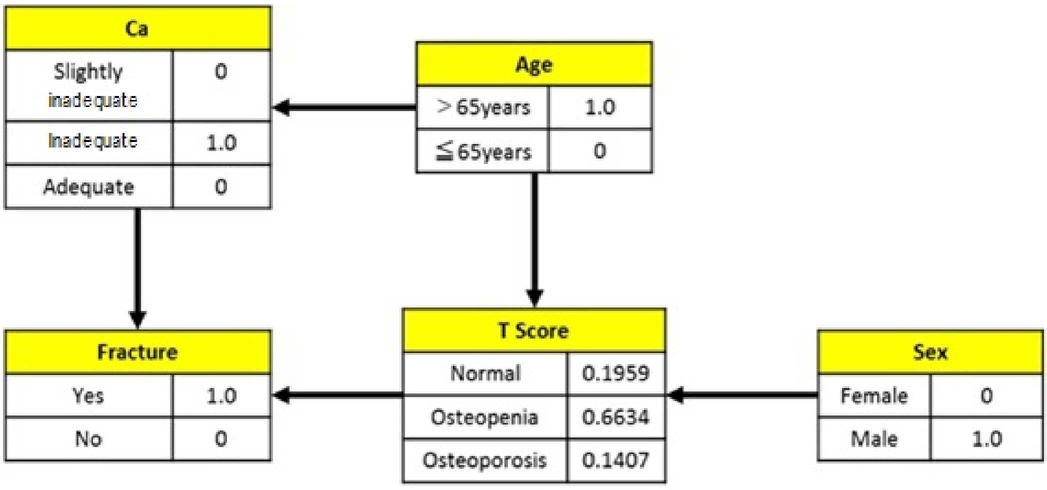
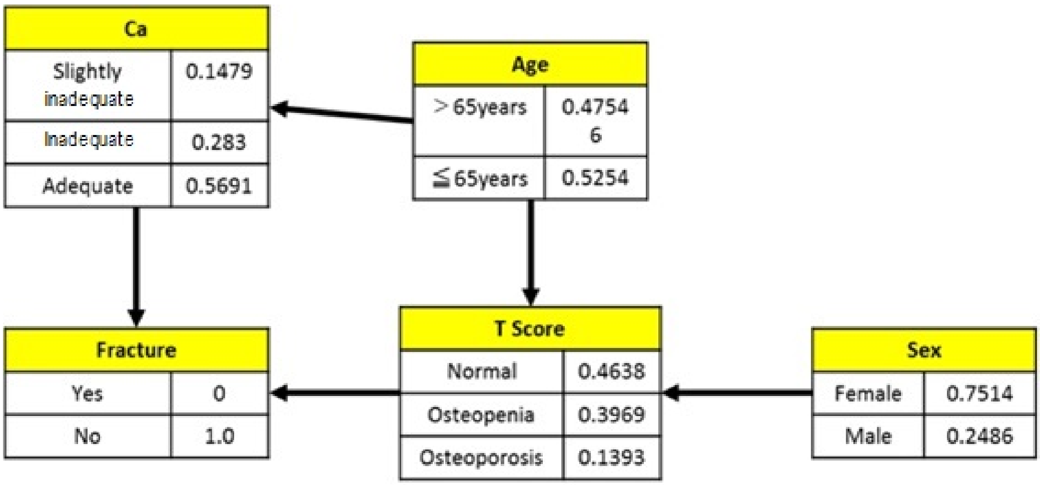


Figure 6. Posterior probability of calcium and T-score in the absence of fracture experience.



The combination with the highest probability to develop osteoporosis consisted of women aged 65 years and older with appropriate calcium intake and history of bone fracture, with the posterior probability of 0.5399 (3.35 times the prior probability). The combination with the highest probability to develop osteopenia consisted of men aged 65 years and older with insufficient calcium intake and history of bone fracture, with the posterior probability of 0.6634 (1.66 times the prior probability). Both results showed that the posterior probability increased when an individual had a history of bone fracture and was aged 65 years or older. However, in terms of sex differences, the results suggested that women had a higher probability of developing osteoporosis while men had a higher probability of developing osteopenia. In addition, the probability of developing osteopenia is suspected to increase when calcium intake is insufficient, as shown in Figure 5. Therefore, particularly in men aged 65 years and older, bone loss can be prevented by a periodic examination of calcium intake and providing nutritional advice.

The mutual information in Table 8 shows that age has the highest level of effect on bone density, followed by bone fracture history, sex, and calcium intake. Age and sex are uncontrollable factors. However, because bone fracture history and calcium intake can be controlled, bone loss can be prevented through measures such as ensuring a safe environment in which bone fracture does not occur and providing nutritional advice, including calcium intake.

Bayesian networks are also believed to be useful in establishing health care policies in the region by changing the posterior probability of each factor in the Bayesian network model and obtaining the factors associated with those changes. Miyauchi et al [16,17] established a Bayesian network using the results of a specific medical checkup and created an evaluation system for the specified medical checkup focusing on lifestyle-related diseases such as metabolic syndrome. As an example, a Bayesian network for the response “I have no history of bone fracture” is shown in Figure 6, with age and sex distributions set to the same conditions as in Figure 3, and the changes in the posterior probabilities are shown in Tables 9 and 10. Factors affecting bone fracture history are calcium intake and T-score. The percentage of individuals deficient in calcium intake should be decreased by 0.54%, and the percentage of persons with appropriate calcium intake must increase by 1.05%. The percentage of individuals within the appropriate T-score range must also increase by 2.46%, those with osteopenia must decrease by 0.27%, and those with osteoporosis must decrease by 2.19%. These methods are believed to be useful indices in establishing health care policies in the region.

Limitations and Future Research Directions

This study has some limitations. First, its reliability is possibly low because data on lifestyle and exercise habits were self-reported. To address this, an activity meter must be distributed to the participants to determine whether exercise was performed definitively. In addition, for items related to exercise in this survey, moderate exercise was defined as “sports or recreation that increases the heart rate for approximately 10 minutes.” However, the frequency and duration of exercise that effectively prevented bone loss are presently unknown. In the future, a more detailed survey must be conducted to determine the appropriate frequency and duration of exercise required to prevent bone loss.

Another limitation was the possibility that the amount of data for our study results was relatively small. In a cohort study conducted in Hisayama, Yoshida et al [18] analyzed data of 1550 participants aged 65 years and older to determine the cause of functional impairment in the elderly people. Kishimoto et al [19] conducted a study on 803 patients to determine their risk of cognitive impairment. Similarly, by continuing our study to gather more data for analysis, we believe that the analytical accuracy of our study can be improved. Table 7 shows the estimated accuracy of osteoporosis, in particular, to be greatly reduced. This reduction in accuracy resulted from a small sample size since only 46 individuals had osteoporosis out of 354 study participants. In order to collect more data, more residents should participate in the study through proactive activities in the region such as establishing a mechanism where participants can see a personal benefit in participating. In addition, an environment that is easily accessible for residents to participate and want to participate should be established such as providing them with incentives for study participation.

In this study, we conducted a survey on customers who came to the drugstore. Therefore, we think that the health condition of customers based on this drugstore may have an impact on the survey results. In the future, public facilities should be used as a survey site, and surveys should be conducted in a wide range of participants. This study covered a heavily snowy area, and many western residents have restricted activities. In addition, the arable land area ratio (arable land area/total land area) of the city is 41.2%, with many residents involved in agriculture.

Table 9. Changes in the posterior probability of calcium in the absence of fracture experience.

Calcium	Prior probability	Posterior probability	Variation (%)
Inadequate	0.30	0.28	–1.59
Slightly inadequate	0.14	0.15	0.54
Adequate	0.56	0.57	–1.05

Table 10. Changes in posterior probability of T-score in the absence of fracture experience.

T-score	Prior probability	Posterior probability	Variation (%)
Normal	0.44	0.46	2.46
Osteopenia	0.4	0.40	0.27
Osteoporosis	0.16	0.14	2.19

However, as farm work is limited during winter season compared with summer season, opportunities to move the body are expectedly reduced [19].

In this study, since the survey was conducted in winter, comparison with the summer season has not been done. Therefore, the influencing factors on the health condition should be analyzed by collecting data in summer and comparing it with the data obtained in winter.

Conclusion

Based on the multiple regression analysis results, age, sex, history of bone fracture, and calcium intake were selected as models, but the effects of other nutrients and exercise were not assessed. In the Bayesian analysis results, age and sex were the only factors that affected bone density; nutrients or exercise did not have an effect. However, the fact that calcium intake and T-score were shown to affect bone fracture history suggests that calcium intake is an important measure to prevent bone fractures. Overall, these results suggest that measures such as ensuring a safe environment to prevent bone fractures and providing nutritional advice for calcium intake are effective in preventing bone loss.

Conflicts of Interest

None declared.

References

1. Government of Japan. 2016. Annual report on the aging society (2016 fiscal year edition) URL: http://www8.cao.go.jp/kourei/whitepaper/w-2016/zenbun/28pdf_index.html [accessed 2018-04-26] [WebCite Cache ID 6yyPnYgHT]
2. Suzukawa M, Shimada H, Makizako H, Watanabe S, Suzuki T. Incidence of falls and fractures in disabled elderly people utilizing long-term care insurance. Jpn J Geriat 2009;46(4):334-340. [doi: 10.3143/geriatrics.46.334]
3. Japan Osteoporosis Society. Prevention and treatment guidelines for osteoporosis URL: http://jsbmr.umin.jp/pdf/GL2015.pdf [accessed 2018-04-26] [WebCite Cache ID 6yyPublcv]
4. Okazaki R, Ozono K, Fukumoto S, Inoue D, Yamauchi M, Minagawa M, et al. Assessment criteria for vitamin D deficiency/insufficiency in Japan: proposal by an expert panel supported by the Research Program of Intractable Diseases, Ministry of Health, Labour and Welfare, Japan, the Japanese Society for Bone and Mineral Research and the Japan Endocrine Society [Opinion]. J Bone Miner Metab 2017 Jan;35(1):1-5. [doi: 10.1007/s00774-016-0805-4] [Medline: 27882481]
5. Shimoda T, Suzuki T, Takahashi N, Tsutsumi K, Samukawa M, Yoshimachi S, et al. Nutritional status and body composition of independently living older adults in a snowy region of Japan. Gerontol Geriatr Med 2017;3:2333721417706854 [FREE Full text] [doi: 10.1177/2333721417706854] [Medline: 28516130]
6. The situation of the elderly population in Hokkaido. URL: http://www.pref.hokkaido.lg.jp/hf/khf/H28_1_1koureikaritsujyun.pdf [accessed 2018-04-26] [WebCite Cache ID 6yyQ5WGqc]
7. Uchida K, Tomonou M, Hayashi M, Shirota T. Relationships of the nutritional intake and other factors with the bone mineral density among elderly residents of Hisayama. Jpn J Nutr Diet 2003;61(5):307-315. [doi: 10.5264/eiyogakuzashi.61.307]
8. Ayumi T. Effects of dietary habit and physical activity on bone mineral density of pre- and postmenopausal Japanese women. J Nutrit Sci Vitaminol 1993;46:387-394 [FREE Full text]
9. Kobayashi S, Honda S, Murakami K, Sasaki S, Okubo H, Hirota N, et al. Both comprehensive and brief self-administered diet history questionnaires satisfactorily rank nutrient intakes in Japanese adults. J Epidemiol 2012;22(2):151-159 [FREE Full text] [Medline: 22343326]
10. Velikova M, van Scheltinga JT, Lucas PJ, Spaanderman M. Exploiting causal functional relationships in Bayesian network modelling for personalised healthcare. Int J Approx Reason 2014 Jan;55(1):59-73. [doi: 10.1016/j.ijar.2013.03.016]
11. Lucas PJF, van der Gaag LC, Abu-Hanna A. Bayesian networks in biomedicine and health-care. Artif Intell Med 2004 Mar;30(3):201-214. [doi: 10.1016/j.artmed.2003.11.001] [Medline: 15081072]
12. van Gerven MAJ, Taal BG, Lucas PJF. Dynamic Bayesian networks as prognostic models for clinical patient management. J Biomed Inform 2008 Aug;41(4):515-529 [FREE Full text] [doi: 10.1016/j.jbi.2008.01.006] [Medline: 18337188]
13. Velikova M, Samulski M, Lucas PJF, Karssemeijer N. Improved mammographic CAD performance using multi-view information: a Bayesian network framework. Phys Med Biol 2009 Mar 07;54(5):1131-1147. [doi: 10.1088/0031-9155/54/5/003] [Medline: 19174596]
14. Kubota M. [Optimization of calcium intake for the prevention of osteoporosis and osteoporotic fractures: a review of the evidence]. Nihon Eiseigaku Zasshi 2003 Sep;58(3):317-327. [Medline: 14533561]
15. Matsubara T, Sakai Y, Yanagawa M, Hijii C, Numata M, Egami K, et al. The effect of bench-stepping exercise training on BMD, BMC and bone metabolism in menopausal women. Jpn J Phys Fitness Sports Med 2013;62(1):95-103. [doi: 10.7600/jspfsm.62.95]
16. Miyauchi Y, Nishimura H. Bayesian network for healthcare of metabolic syndrome, Short paper No. 3164. 2013 Presented at: IEEE EMBC2013; 2013; Osaka. [doi: 10.11239/jsmbe.51.R-306]

17. Miyauchi Y, Nishimura H. Construction and evaluation of Bayesian networks related to the specific health checkup and guidance on metabolic syndrome. In: Innovation in Medicine and Healthcare. New York: Springer International Publishing; 2015:183-193.
18. Yoshida, Daigo. Prevalence and causes of functional disability in an elderly general population of Japanese: the Hisayama study. J Epidemiol 2012;22:222-229. [Medline: 22343328]
19. Kishimoto H, Ohara T, Hata J, Ninomiya T, Yoshida D, Mukai N, et al. The long-term association between physical activity and risk of dementia in the community: the Hisayama Study. Eur J Epidemiol 2016 Mar;31(3):267-274. [doi: 10.1007/s10654-016-0125-y] [Medline: 26857126]

Abbreviations

BDHQ: brief-type dietary history questionnaire
BMI: body mass index

Edited by G Eysenbach; submitted 28.07.17; peer-reviewed by T Tanikawa, JR Bautista; comments to author 23.11.17; revised version received 21.02.18; accepted 30.03.18; published 22.05.18.

Please cite as:
Suzuki T, Shimoda T, Takahashi N, Tsutsumi K, Samukawa M, Yoshimura S, Ogasawara K
Factors Affecting Bone Mineral Density Among Snowy Region Residents in Japan: Analysis Using Multiple Linear Regression and Bayesian Network Model
Interact J Med Res 2018;7(1):e10
URL: http://www.i-jmr.org/2018/1/e10/
doi:10.2196/ijmr.8555
PMID:29789280

Original Paper

Engaging a Community for Rare Genetic Disease: Best Practices and Education From Individual Crowdfunding Campaigns

Romina Alicia Ortiz[1], MHS; Steven Witte[1], MD, PhD; Arvin Gouw[1], PhD; Ana Sanfilippo[1], MBA; Richard Tsai[1], MSc; Danielle Fumagalli[1], MA; Christine Yu[1], MPH; Karla Lant[1], BA; Nicole Lipitz[1], BA; Jennifer Shepphird[1], PhD; Fidelia B Alvina[1], BA; Jimmy Cheng-Ho Lin[1], MPH, MD, PhD

Rare Genomics Institute, Downey, CA, United States

Corresponding Author:
Romina Alicia Ortiz, MHS
Rare Genomics Institute
8504 Firestone Blvd #198
Downey, CA, 90241
United States
Phone: 1 5624132581
Email: romina.ortiz@raregenomics.org

Related Article:

This is a corrected version. See correction statement: http://www.i-jmr.org/2018/1/e10707/

Abstract

Background: Genetic sequencing is critically important to diagnostic health care efforts in the United States today, yet it is still inaccessible to many. Meanwhile, the internet and social networking have made crowdfunding a realistic avenue for individuals and groups hoping to fund medical and research causes, including patients in need of whole exome genetic sequencing (WES).

Objective: Amplify Hope is an educational program designed to investigate what factors affect the success of medical crowdfunding campaigns. We conducted a needs assessment, a series of 25 interviews concerning crowdfunding, and provided training on best practices identified through our assessment for 11 individuals hoping to run their medical crowdfunding campaigns to raise money for patients to access trio WES to identify the mutated proteins that caused their apparent inherited disease.

Methods: The crowdfunding education was given in a 30-day training period with resources such as webinars, fact sheets and a crowdfunding training guide emailed to each participant. All campaigns were launched on the same date and were given 30 days to raise the same goal amount of US $5000. Reviewing the 4 crowdfunding campaigns that raised the goal amount within the 30-day period, we sought to identify features that made the 4 crowdfunding campaigns successful. In addition, we sought to assess which factors the resulting 75 donors report as influencing their decision to donate to a campaign. Finally, we investigated whether crowdfunding campaigns for exome sequencing had an impact on increasing applicant's and donors' knowledge of genomics.

Results: Of the 86 study inquiries, 11 participants submitted the required forms and launched their crowdfunding campaigns. A total of 4 of the 11 campaigns raised their goal amounts within 30 days.

Conclusions: We found that social media played an important role in all campaigns. Specifically, a strong social media network, an active outreach process to networks, as well as engagement within the study all correlated with a higher success rate. Amplify Hope donors were more likely to support projects that were near their fundraising goals, and they found video far more effective for learning about genomics than any other medium.

(Interact J Med Res 2018;7(1):e3) doi:10.2196/ijmr.7176

KEYWORDS

surveys and questionnaires; patient empowerment; fund raising; benchmarking; internet; genetic testing; molecular diagnostic techniques; social planning; health education

Introduction

Genetics in Society

DNA sequencing is at the forefront of many cutting-edge research and treatment programs, especially for genetic diseases. The diagnostic yield of whole exome genetic sequencing (WES) is reported to be 25% to 34% for children with undiagnosed genetic disease [1-3]. WES reduces the amount of time needed to reach a diagnosis compared with the current standard of care [1], and it has the advantage of being able to identify novel mutations [2]. Genetic sequencing also plays a central role in the implementation of personalized medicine, an increasingly important priority in our society. The Precision Medicine Initiative introduced by President Obama in 2015 aims to accelerate research efforts by enhancing data access and collaborations between researchers, doctors, and patients; a key component of the project includes genetic sequencing. In another nationwide initiative, the PrecisionFDA Consistency Challenge works to improve personalized care by achieving more consistent results in genetic tests.

Advances in genetic sequencing technology have substantially reduced its cost, making it more accessible to patients and researchers. Less than 20 years ago, sequencing the entire human genome cost billions of dollars, but today, the cost to sequence the most relevant parts of the genome and provide information as a clinical test option is in the thousands [4]. WES has become the gold standard of advanced genetic testing to determine the underlying cause of a patient’s undiagnosed illness. Majority of inherited diseases identified to date result from mutated proteins [5], and WES is the sequencing of the complete protein-coding regions (exome) of the human genome [6].

Although the cost to sequence a genome has been substantially reduced, it is still unaffordable for many patients because it is not routinely covered by health insurance. These out-of-pocket expenses are a financial strain for the patients that need testing for diagnosis. Looking just at financial strain for a typical person, a 2011 poll from the National Bureau of Economic Research found that nearly half of the respondents stated that if they were given a 30-day time frame, they would be unable to produce an extra US $2000 [7]. Beyond overtaxed household budgets, families dealing with undiagnosed or rare diseases incur countless unexpected health care costs that create tremendous financial burdens. These families shoulder the staggering health care expenses accumulated over years of seeking treatment.

Engaging a Community

The increasing reach of the internet and social networking has created a new channel for fundraising through internet-based crowdfunding, which asks many people via websites and social media to donate money and provide support to an individual project or a campaign. Organizers of crowdfunding campaigns use their online connections and offline activities to promote their campaigns. Unlike traditional funding that comes from a few large investors or campaigns that ask for a specific donation amount, crowdfunding allows anyone to donate any amount. In light of the diminishing funding for research grants, crowdfunding enables individuals and groups worldwide to bridge the funding gap and support medical and research causes [8]. Crowdfunding may help support patients in need of DNA sequencing by providing access to a test that might uncover the cause of their disease and potentially lead to a viable clinical treatment.

Although previous research has examined crowdfunding as a novel means for raising money and support, few studies have focused on the factors that improve crowdfunding outcomes. Our literature search identified just one study that focused exclusively on crowdfunding for personal medical expenses. Although several studies discuss the possibilities of crowdfunding for larger medical research projects that are not able to secure funding through traditional sources, the studies did not examine factors that contribute to successful campaigns. Burtch and Chan [9] found evidence that crowdfunding campaigns were correlated with a reduction in medical bankruptcies in the United States, concluding that crowdfunding helped reduce the rate of bankruptcy due to medical expenses by 3.7% in 2014. However, they do not go further beyond establishing this link to examine the factors responsible for the success of crowdfunding campaigns or the donors’ motivations for giving.

Two key drawbacks of the existing body of research limit its relevance to the funding of personal medical expenses. First, most studies have evaluated platforms that offer donors a return, such as a final product from a Kickstarter campaign, or interest on loans from the crowdfunding platform Prosper. Purely charitable crowdfunding projects have not been studied. Second, many of the platforms analyzed adopted an “all-or-nothing” policy, whereby if the goal goes unmet, the donations are returned. This feature along with the presence (or absence) of a reward almost certainly influences some individual decisions to contribute. Findings based on postcampaign surveys of donors are also limited.

With the gap in financial support for undiagnosed individuals, Rare Genomics (RG) began allowing patients and their families to create crowdfunding projects since 2012. These campaigns were on the RG website and set to raise the funds needed for clinical WES with partnering RG laboratories. Early campaigns varied in success, ranging from hours to days to months. Amplify Hope was then designed to investigate existing strategies for crowdfunding, educate participants preparing to crowdfund, and determine the effectiveness of crowdfunding strategies used as well as donor engagement.

Methods

Study Design

RG conducted the study in 5 phases, including a needs assessment to identify successful crowdfunding strategies, participant recruitment, a 30-day crowdfunding training period, a 30-day online crowdfunding period, and follow-up surveys given to participants and donors. Our research sought to (1) provide demographic information on the donor population; (2) identify common factors among successful medical crowdfunding campaigns; (3) identify factors that influenced people to donate, as reported by donors; and (4) describe the impact crowdfunding campaigns had on donors' self-reported knowledge of genomics.

CrowdRise, Indiegogo Life, and YouCaring were the crowdfunding platform partners. They provided the following information to be used for analytical purposes: the average donation to a fundraiser, unique page visits, conversion rates, and how many people had visited each page. We utilized 3 different crowdfunding sites to see if there was a measurable difference in terms of successful campaigns. Baylor Miraca Genetics Laboratories and Ambry Genetics were the sequencing partners in the study. They provided WES for all patients with successful crowdfunding campaigns.

Needs Assessment

A needs assessment was conducted to elucidate current crowdfunding practices and their effectiveness. The first phase consisted of 25 phone interviews with crowdfunding experts and founders of crowdfunding platforms as well as individuals who had run successful crowdfunding campaigns for personal medical care or scientific research projects. Ethan Austin, Breanna DiGiammarino, Adam Griff, Annette Hayswirth, Elizabeth Iorns, Nick Karolidis, Denny Luan, Molly Lindquist, Andrea Lo, Luke Miner, Jamie McDonald, Sandip Sekhon, Nick Sireau, Devin Thorpe, and Rob Wu are founders of crowdfunding platforms that shared their expertise. Sam De Brouwer, Zsuzsanna Darvai, Ignacio Garcia, Kimmie Ng, Glenn O'Neill, Susanne Shaw, and Jeneva Stone shared their insights about their own crowdfunding campaigns. The interviews were conducted for an average length of 30 min. The purpose of all interviews was to establish crowdfunding best practices, elicit recommendations, and develop materials for the training program phase of the study.

Participant Population

As the crowdfunding efforts were aimed to raise money needed for WES, inclusion in the study was dependent on 2 factors: that the participant had the desire to actively participate in the study to raise funds online, and that the participant intended to raise money for a patient that had a certified physician's request for WES.

A total of 13,542 recruitment emails were sent from May 20, 2015, to July 28, 2015, to rare disease advocacy groups and genetic counselors to inform them of the study opportunity for undiagnosed patients. Participants completed an interest form (Multimedia Appendix 1) asking who the intended sequencing was for, whether there was a physician referral, as well as their level of knowledge on WES and crowdfunding. To participants that indicated they had physician support, a Getting Started Worksheet (Multimedia Appendix 2) was sent allowing for a more comprehensive investigation into the patient and participant. The Getting Started Worksheet was designed to determine social media and community presence and activity. Participants who completed both forms were then asked to review a consent form (Multimedia Appendix 3) to enter the study, indicating they would participate in the training, launch their campaign, and provide campaign evaluation. All participants provided contact information for the patient's referring physician, and referring physicians confirmed the WES order by submission of a signed doctor's note (Multimedia Appendix 4).

The participants were randomly assigned to one of three crowdfunding sites. Participants in the study were given a 30 day training program. Educational resources on crowdfunding were emailed every week including webinars, fact sheets and a crowdfunding training guide. During the 30 day training program, participants received an eBook on crowdfunding best practices, educational worksheets on both exome sequencing and crowdfunding each week by email and were invited to attend webinars. The training included recommendations on engaging their network of contacts through email and phone calls, team building, creating a campaign video, and social media messaging. At the same time, participants received challenges designed to help them prepare for the campaigns. Incentives were included in the challenges. One challenge awarded prizes of US $200, US $100, and US $50 to the top 3 campaigns that raised the most money on the launch day. Another challenge awarded US $50 to every campaign that completed the campaign-page story by the deadline set before the campaign launch.

The fundraising goal for all participants was set to US $5000, the cost of a trio WES with partnering laboratories Ambry Genetics and Baylor Miraca Laboratories. The time frame was 30 days, with each participant randomly assigned to one of the 3 different crowdfunding sites. Challenge events after the campaign launch provided campaign organizers opportunities to engage with their networks by offering incentives to donate and share. A week-long challenge offered a first prize of US $200, a second prize of US $100, and a third prize of US $50 to the campaigns with the most new donors during the challenge week. A similar challenge offered prize awards for the campaigns with the most social shares for the week. If participants did not achieve their goals after 30 days, the campaigns were extended an additional 60 days. Overall, 2 participants exceeded the goal in 30 days and 4 met the goal within 60 days.

Effectiveness of Crowdfunding Strategies

To evaluate the effectiveness of the different crowdfunding strategies, the following metrics were assessed: time to complete fundraising, communication engagement with the participants, number of donors, average donation amount and donor education on crowdfunding and sequencing. To gauge communication engagement, we gave communication points for the following two instances: (1) when participants responded to our email or call communications about the study and (2) when participants initiated their own communication to us regarding the study. We subtracted communication points when a participant did not respond to our emails or calls regarding the study. We also noted when these communications had a quick response time (within the same day) and a long response time (over 3 days).

To examine whom the donors were and what factors contributed to their decision to donate to a campaign, demographic analyses of the donors were done utilizing surveys. After a campaign donation was received, donors were invited to respond to an anonymous survey asking about their demographics, such as sex, age range, and educational level, as well as gauging their self-reported baseline knowledge of genomics. Additionally, they were asked about the impact, if any, of the educational materials provided on their knowledge of genomics. They were asked what specifically they read of the information made available to them about genomics. They had the ability to access information from the campaign page and were asked what material did you receive, read, or watch from the campaign (check all that apply). The choices were read the campaign summary, looked at other similar campaigns, watched the TedxMidAtlantic video, read Genomics 101, read eBook(s), and other (please specify).

Education Through Crowdfunding

To determine the baseline self-reported knowledge of the participants, they were asked to rate their crowdfunding and sequencing knowledge from the initial interest form (SF1) from a scale of 0-10. Participants had the ability to access educational information from the campaign and were asked what material they did receive/read/watch from the campaign (check all that apply). The choices were as follows: (1) read the campaign summary, (2) looked at other similar campaigns, (3) watched the TedxMidAtlantic video, (4) read Genomics 101 pamphlet, (5) read eBook(s), and (6) other (please specify). We used postcampaign surveys to determine if the crowdfunding campaigns increased participants' and donors' knowledge of genomics or increased interest in learning about it.

Results

Needs Assessment

The takeaways from individual phone interviews with 25 crowdfunding platform founders and people who led successful campaigns were prepared in the document "Crowdfunding Best Practices" (Multimedia Appendix 5) and were used to train the participants in the study. Best practices identified include having a "pregame plan" before the start date of the campaign where you set a clear measurable goal, create a message and strategy for your target community, arrange initial commitments, and create media for the campaign including high-quality photos, a video, and written materials. Best practices during the campaign include holding a launch party, providing regular and detailed updates, use of social media daily and media contacts, linking of offline fundraising to the online work, and eliciting feedback and acting on it.

Participant Population

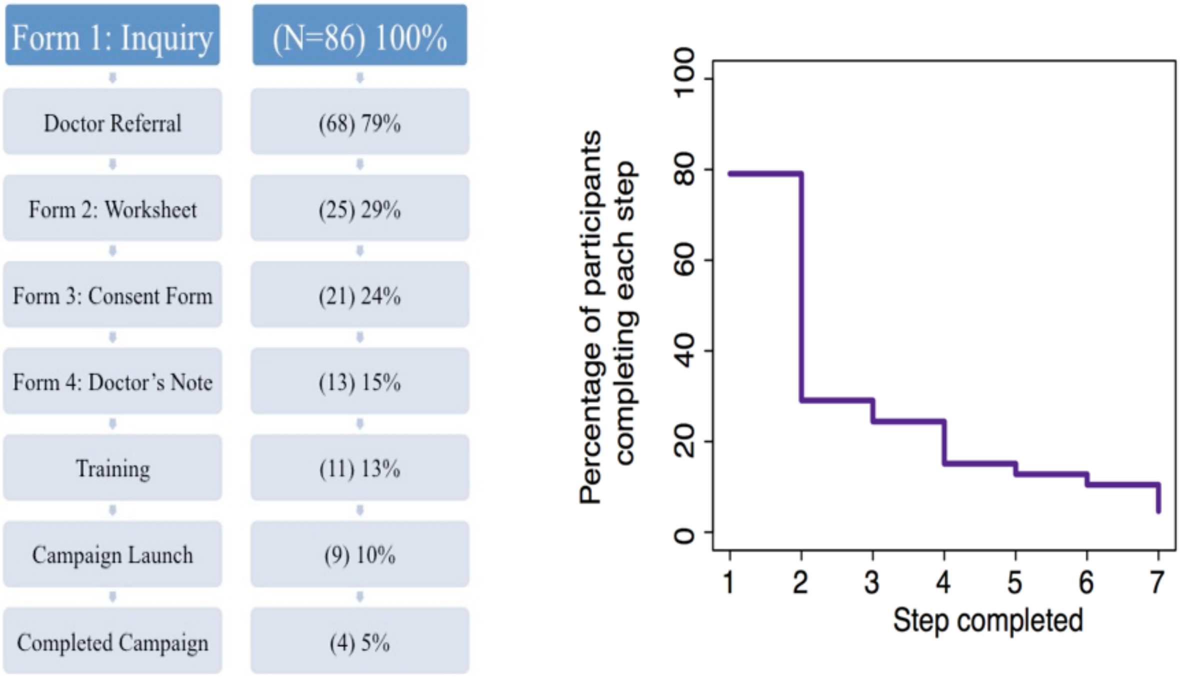
From the 86 completed inquiries, 68 applicants (79%) affirmed they had doctor referrals for the test (Figure 1). From the 68 that appeared to have medical support to request the test, 25 (less than 40%) completed Form 2 and 84% of those participants submitted signed consent forms (Multimedia Appendix 3) for the study, dropping the participant number to 21. The study team received signed doctor's notes for 13 of the 21 participants. The 8 referring physicians who did not submit doctor's notes were either unresponsive or stated that they did not recommend WES. All 13 cases with completed forms were invited to enter the study for training. One of the cases became unresponsive and did not enter training. Of the 12 that participated in the training, one participant dropped out during training due to lack of ability to commit time to the effort.

The principal sources of applicants were the patients themselves (38 out of 86) and parents of the patients (38 out of 86). A total of 65 (76%) of the 86 that applied were females and 21 (24%) were males. The ages were concentrated between 40 and 49 years. The second largest group was between 30 and 39 years. A total of 66 of the 86 (76%) applicants had a bachelor's degree or higher, but at the time of the study, 52 out of the 86 (60%) participants were either unemployed or disabled.

Effectiveness of Crowdfunding Strategies

From the 11 participants who completed training for the study, one secured complete sponsorship for their campaign during the training program before the launch, and one participant won a free WES test with first place in one of the study's training challenges. In total, 9 crowdfunding campaigns were launched, and 4 campaigns reached their funding goals of US $5500.

Figure 1. Of the 86 people that expressed interest in the study, 4 people (5%) successfully completed all subsequent steps, including the funding campaign (step 7).



One campaign was funded during preparation of the launch, and the other 3 campaigns that did launch each utilized a different crowdfunding platform—CrowdRise, Indiegogo Life, and YouCaring. Thus, all of the crowdfunding platforms had successful campaigns represented in our study.

We observed a few overlapping characteristics among the campaigns that achieved complete funding by the end of the study period in both internal engagement with the study coordinators as well as external engagement with social networks before, during, and at the end of the study. Higher levels of engagement were observed throughout the training before the start of the crowdfunding campaigns as compared with those that did not achieve their campaign goals. This included communication engagement as well as completion of study and campaign tasks.

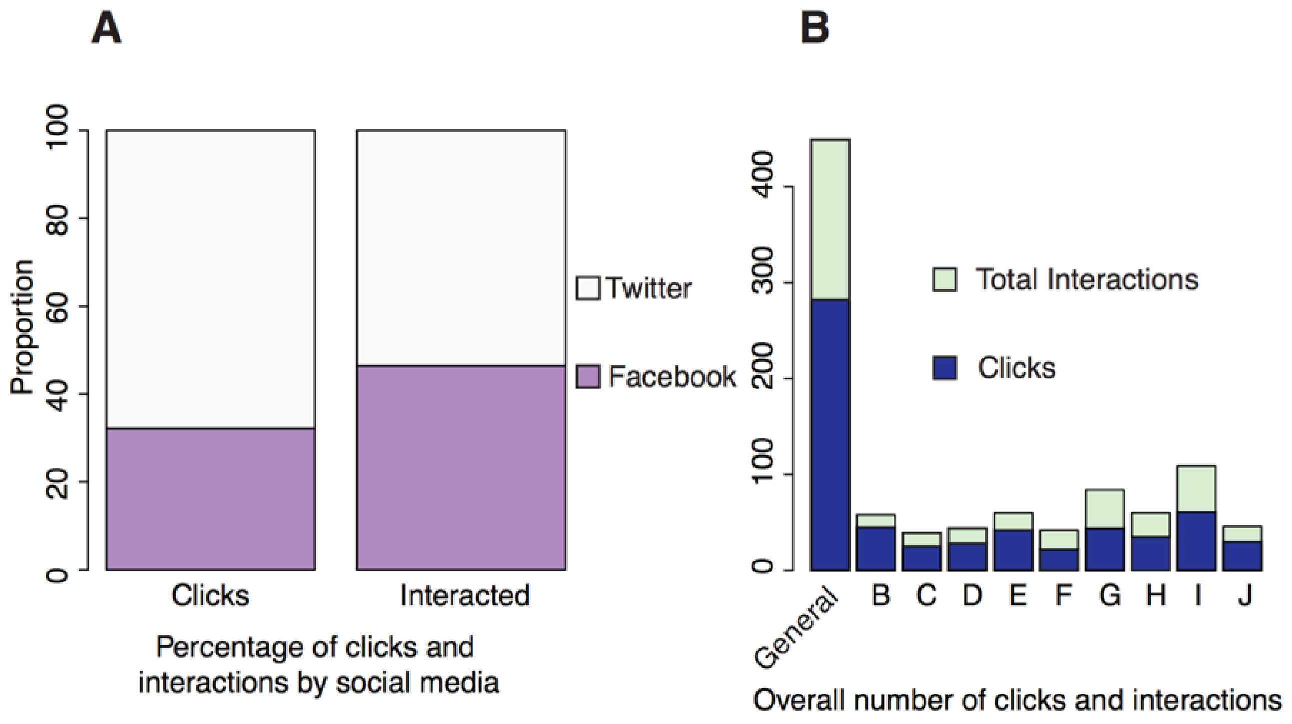
By the measure of internal study engagement, all successful crowdfunding campaigns had 4 or more total communication engagement points, whereas majority of the unsuccessful crowdfunding campaigns had 3 or less points total. Of note is participant #8 that had 12 engagement points total and still was unsuccessful. Interestingly, none of the successful crowdfunding campaigns had an instance where the participants did not respond to either email or phone communication from the study organizer (no responses). All successful campaigns had points for self-initiated engagement, and all but one unsuccessful campaign also had points for self-initiated engagement. However, successful campaigns had an average of 2.75 points, whereas unsuccessful campaigns had an average of 1.67 points. Moreover, successful campaigns had an average of 4 points for a quick response time of 1 day or less. Unsuccessful campaigns only had an average of 2.5 points. From both successful and unsuccessful campaigns, the number of instances with slow response times were insignificant at less than 0.5 points average for both.

Most participants used social media as a tool to engage their social network to raise funds, and most were already well connected through many platforms: 8 of the 9 reported they used Facebook (3 of the 9 had 500 or more Facebook friends), 6 of the 9 used Twitter, and 5 of the 9 used Google + and LinkedIn. In addition, many regularly shared updates about their rare disease journey on (8 of the 9) Facebook and Twitter (4 of the 9). Once the 30-day campaigns launched online, a total of 334 social media messages were shared on Twitter, Facebook, LinkedIn, and Google +. Of the 334 total social media posts, 217 (64.9%) were shared, mostly on Facebook and Twitter. Embedded links in the shared posts directed the user to an individual participant's crowdfunding site or to the Amplify Hope website. The click rate for embedded links was 70.0% (152/217) of shared messages, 67.8% (103/152) from Twitter posts and 32.2% (49/152) from Facebook posts. Of the total shared posts, 58.5% (127/217) garnered some form of interaction (eg, they were retweeted, liked, favorited, or shared) at least once (53.5% [68/127 on Twitter and 47% [59/127] on Facebook).

Messages promoting the Amplify Hope campaign through social media were divided into the following 2 categories: (1) general campaign messages to promote and encourage user participation and (2) campaign-specific messages tailored to drive traffic to an individual's campaign. Campaign-specific messages garnered slightly higher total click rates and interaction rates than did general messages on both Twitter (52% vs 44%) and Facebook (55% vs 44%), as shown in Figure 2. Similarly, individual-specific messaging appeared to be somewhat more effective than general messaging in reaching audiences on social media.

Separately, we sought to get a comprehensive understanding of characteristics and strategies used by the successful crowdfunding campaigns. All the 3 successful campaigns were organized by mothers of the patient in need of exome sequencing. In one of the cases, the mother completed the training program. She reported reading all the materials provided and engaged with the study coordinators by asking self-initiated questions throughout the process. She completed all aspects of the precampaign work, reaching out to her network of friends and family to inform them of her upcoming campaign, she wrote a campaign story on the CrowdRise crowdfunding platform, and included 5 photos and a video as recommended. She held a launch party the day the campaign went live online. The campaign won the first place in our challenge of the week for the most new donors and second place in the challenge for the most social shares. Her top source of messaging was Facebook, and during the 30-day online study period, the campaign had 1208 unique user sessions and a total of 74 donors. The campaign raised US $5680, exceeding the US $5500 goal. In the second case, the campaign was conducted from Tanzania using the YouCaring platform. Here too, the campaign organizer participated fully in the 30-day training and reported reviewing all the material provided. Of note, she had some delays in completing the study homework tasks as she had a child in the hospital. However, despite this, she completed the recommended work before launching her 30-day online campaign. She created a campaign story on YouCaring's platform, which included 7 photos and 1 video with words embedded in the video. She requested contributions at the end of the video and thanked supporters. Similar to the first case, the campaign won challenges. This campaign won first place for the most social shares in 1 week and second place for the most new donors in 1 week. The campaign had 2394 unique users and 3589 total sessions. The campaign exceeded the goal of US $5500 by raising US $5844 with 47 donors. One campaign received full funding during the training program. The campaign organizer was an adult patient in need of exome sequencing. During the training program, he was fully engaged in the training, conducting recommended outreach by email and phone to his network of friends, family, and contacts. He requested feedback on his campaign draft story from his network as well as the study team. During his outreach, one donor agreed to pay the entire cost of exome sequencing, and thus, he was fully funded for US $5500 before the online campaign period of the study.

Figure 2. Analysis of social media messages. (A) Breakdown of clicks and interactions of social media posts made via Twitter and Facebook. (B) Analysis of engagement of social media posts for individual campaigns. “General” refers to generic messages meant to promote user participation, and “B-J” refers to individual campaigns, which have been anonymized.



Data were also evaluated from a total of 75 donors who supported the Amplify Hope campaigns. After contributing to a campaign, donors were asked to complete a survey (Multimedia Appendix 6) about themselves and the factors that influenced them to donate. A total of 64 out of the 75 (85%) donors were females compared with males (11 out of the 75, 13%), and the majority had some college and graduate degree (Figure 3). Moreover, 30 out of the 75 (40%) donors said they were friends of the individual raising money. Family or relatives represented the next largest segment at 26 out of 75 (35%), and 9 out of the 75 (12%) donors reported that they did not know the individual raising money.

A total of 46 out of the 75 (61%) donors heard about the crowdfunding campaigns through social media with Facebook as the most prominent platform in 67 out of 75 (89%) responses for learning about the campaign. However, there was a dramatic increase in the contribution amount when the medium used was word-of-mouth, email, and phone. This may be due to an increased personal connection through delivery of the message. To assess the effects of social pressure on charitable giving, respondents were asked whether they prefer to donate at the beginning of a campaign to impact the initial momentum in a campaign or at the end to see a campaign reach its goal. Respondents were asked to note their agreement with a series of statements. The question, “How important is it for you to fund crowdfunding projects that have already received substantial donations from others?” relates to the crowd effect when respondents view other’s donations and make a decision to fund a campaign that has received money from others. Most respondents (58 out of 75, 77%) thought it was at least moderately important, indicating that social recognition is a factor and the performance of the campaigns in general is key to why someone donates. Furthermore, respondents were asked, “How important is it to fund projects that are close to meeting their fundraising goal?” In total, 64 out of 75 (85%) reported moderately to very important. The amount raised appears to have an impact on the willingness of donors to get involved in funding a campaign.

Education Through Crowdfunding

The 86 study applicants were asked to report the knowledge on exome sequencing and crowdfunding on a scale of 0 to 10, with 0 referring to no knowledge. Respondents were more familiar with sequencing than crowdfunding. Among 86 applicants, 13 reported they had no knowledge of crowdfunding compared with 6 who were unfamiliar with sequencing (see Figure 4). When the selected 10 participants were asked, “Participating in this crowdfunding campaign has helped me to better understand the genetic test,” campaign organizers reported either agreeing or strongly agreeing.

Donors were asked to report their knowledge on genomics before and after donating to the campaigns. In total, 58 responses were collected. Majority of the donors self-reported they had limited knowledge of genomics before donating to the campaigns. When asked before the campaign to rate their knowledge of genomics on a scale from 0 (no knowledge) to 10 (expert), the average score was 2. After participating in the campaign, the average score increased by 84% with about 9 out of 58 (16%) scoring more than a 7. There were 24 (43% of total) respondents who scored “0” before the campaign, and this decreased by 71% to only 7 respondents afterward. The average new score for those who scored “0” before donating to the campaign was approximately 2.2.

Figure 3. Donor demographic information. (A) Education level of donors. (B) Age of donors. (C). All data are self-reported by donors.

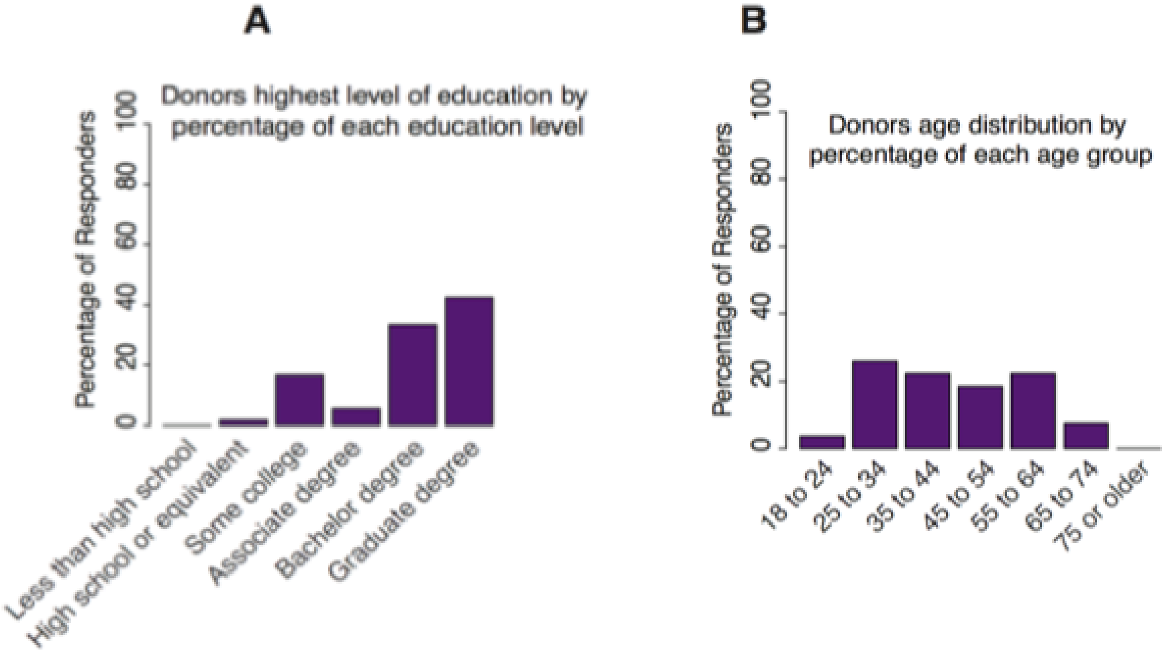
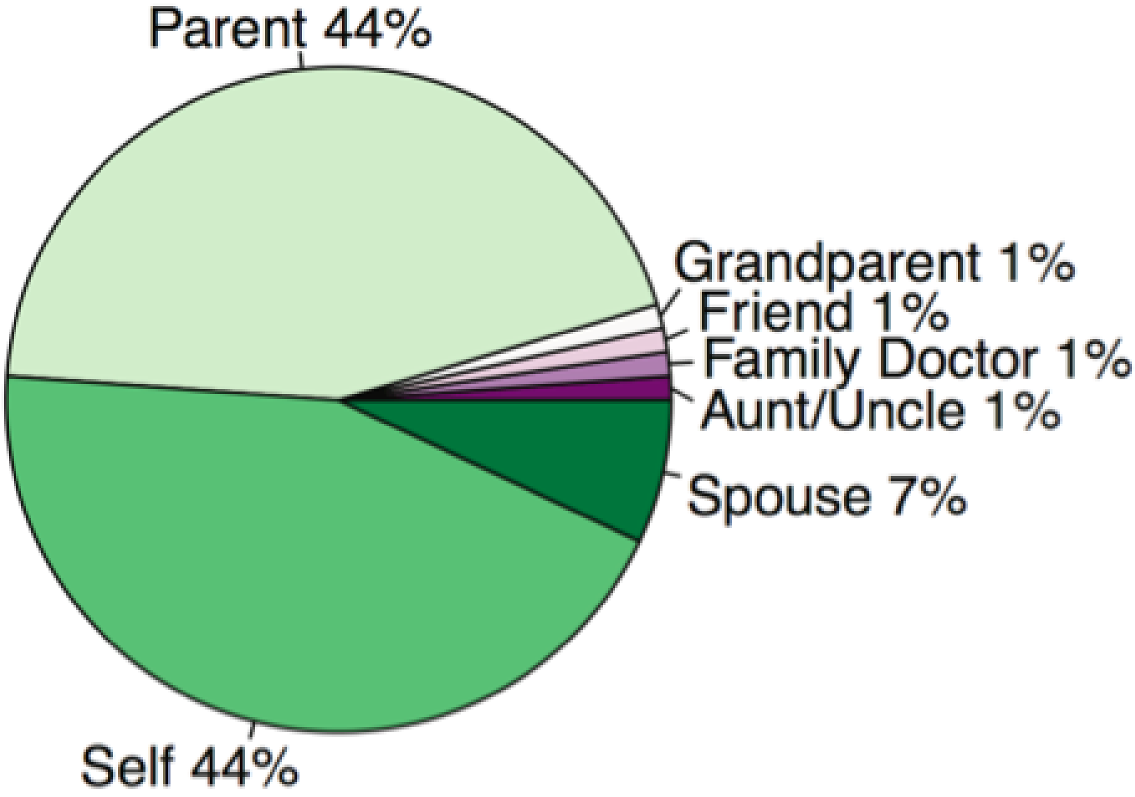


Figure 4. (A) Number of participants who reported each score for knowledge on exome sequencing. (B) Number of participants who reported each score for knowledge on crowdfunding.



Across all campaigns, all donors gained self-reported genomics knowledge by an average of 100% or more. To further examine the knowledge increase in genomics, respondents were asked if they understood what the funds of the project went toward and 51 out of 58 (87%) of the donors knew how the funds would be used, whereas 7 out of the 58 (7%) did not. When asked, “After donating to this campaign how interested are you in learning about genomics and personalized medicine,” 36 out of the 58 (62%) donors were interested in genomics and learning more about personalized medicine, whereas 7 out of the 58 (7%) were not interested. Lastly, donors were asked about their attitudes on the most effective ways to learn about genetic sequencing and genomics . Across all age groups, 46 of the 75 (61%) respondents felt videos were the most effective learning tool, followed by fact sheets/reports/technical papers and case studies.

Discussion

Limitations

Although the study showed several potential benefits of a crowdfunding approach to raise money for WES, there are also several limitations. One challenge inherent with rare or undiagnosed diseases is the small patient population. In addition, each participant in the Amplify Hope study had to have a genetic etiology consistent with rare undiagnosed diseases as confirmed by physician support for WES.

Other limitations of using a crowdfunding approach are the time and effort involved to reach individuals in a social network and to ask to support the campaign or share the campaign with their contacts. If a campaign organizer has a limited network of contacts, this could negatively affect the outcome. Another possible limitation is that individuals that are crowdfunding for WES may be the patient themselves and have medical limitations that prevent them from fully engaging in the crowdfunding process to achieve a successful outcome. A larger study using different milestones and improved methodologies would be needed to validate these conclusions with statistical power. With this, future researchers can continue to study this method or preference of instruction to provide information to educate the public on genomics and personalized medicine. Additional research will be needed, however, to assess the effectiveness of other instruction methods that were not included in this study.

Principal Findings

The campaigns that achieved complete funding during the 30-day online crowdfunding campaign study period shared several characteristics. Individual campaign organizers who reached their campaign goals engaged more during the training program before the start of the crowdfunding campaign, as compared with those that did not achieve their campaign goals. These organizers read our shared Amplify Hope educational materials, followed recommended guidelines, and perhaps most importantly initiated communication via phone calls and emails to their networks before the campaign launch. In other words, the successful campaigners frontloaded their campaigns. They also actively engaged their network and donors by providing updates throughout the campaigns rather than trailing off.

Social media was important in connecting donors to the crowdfunding campaigns. In particular, Facebook was reported as the way many donors learned about the crowdfunding campaigns. Friends and family were the top sources of donations to the campaigns, supporting research that a strong network of individuals and larger social network equates to a higher crowdfunding success rate. Mollick and Kuppuswamy [10] in other studies examining crowdfunding for Kickstarter ventures found that having a large number of friends on online social networks is correlated with success. Demographic data obtained from our surveys revealed that women are more likely to donate to this research. This is consistent with the literature, which shows that a higher proportion of crowdfunding campaign donors are female. Greenberg and Mollick [11] and Marom et al [12] also found that women were considerably more likely than men to successfully raise capital for business ventures on Kickstarter, and argue that this is primarily due to the tendency of female donors to help other women. Additionally, we found that this donor segment was highly educated, with 66 out of 86 (76%) reporting having a bachelor’s degree or higher.

We also sought to determine the factors that influenced donors to support the crowdfunding campaigns for WES. We found that if a campaign is close to a deadline and has not reached their goal, donors are more likely to give money to help them reach their goal before the deadline. This is consistent with Kuppuswamy and Bayus [13], who found that in terms of monetary targets, donors on Kickstarter are more likely to support projects that are near their goals, viewing these projects as more likely to be successful, and that nearly all projects on Kickstarter that reach 50% of their fundraising goal are eventually fully funded. They also find that successful projects on Kickstarter are likely to have a public or private update near their campaign’s target end date.

Through online campaign summaries and links to educational resources, crowdfunding campaigns may increase knowledge regarding genetic sequencing, particularly as it relates to undiagnosed and rare diseases. As most of our participants and donors rely on social and electronic media, we hypothesized that the crowdfunding campaign and social media use leads to an increase of knowledge. We used postcampaign surveys to determine if the crowdfunding campaigns increased participants’ and donors’ knowledge of genomics or increased interest in learning about it. It is likely that participants had better sequencing knowledge due to the prerequisite of the study of physician-recommended WES, given this parameter, the patients had probably discussed the need for this testing with their physician. Majority of the donors had limited knowledge of genomics before donating, and across all campaigns, donors gained self-reported genomics knowledge by an average of 100% or more. The campaigns appeared to spur an interest in learning more about personalized medicine—46 out of the 75 (61%) donors were interested in learning more about genomics and personalized medicine. Postcampaign surveys were used to ask organizers about education on genetic sequencing as a result of having organized and executed a campaign. When asked, “Participating in this crowdfunding campaign has helped me to better understand the genetic test,” campaign organizers reported either agreeing or strongly agreeing. Our findings are consistent with Facio et al [14], in that there is interest in learning about genome sequencing and a perceived value in that knowledge. In addition to reporting a desire to learn more about genomics, donors in Amplify Hope also stated a preference as to how they would like to learn. When asked what method(s) would be most effective to learn more about genomics—video, webinars, fact sheets/technical papers, e-newsletters, or case studies—46 of the 75 (65%) respondents reported videos as their first-choice learning format. This supports the notion that videos are effective at communicating messages in crowdfunding campaigns and are the preferred method for continuing education on genetic sequencing and genomics among respondents.

Although crowdfunding has become a more popular means of donating in recent years, there is still a stigma associated with asking friends and family for donations for personal medical causes. We found that to be successful, the campaign organizer should actively participate in the process of outreach and have a higher degree of comfort with social networking as reported by the study participants.

Future Works

This study showed that crowdfunding campaigns have the potential to benefit individuals with rare diseases seeking funding for various diagnostic and treatment options. The study highlights a self-reported donor population that is willing to fund research for friends, family, or strangers afflicted with a rare disease. Additionally, both donors and participants expressed a desire to not only learn more about personalized medicine and genomics, but there was additional specificity and preference on how to learn. The method of instruction through video is preferred by all respondents of the surveys. With this, future researchers can continue to study this method or preference of instruction to provide information to educate the public on genomics and personalized medicine. Additional research will be needed, however, to assess the effectiveness of other instruction methods that were not included in this study.

Crowdfunding offers a different approach to fundraising. The biggest difference between crowdfunding and traditional fundraising is that there are many smaller monetary donations in crowdfunding. Crowdfunding for scientific initiatives such as raising money for WES allows donors to choose the cause or person they would want to contribute to and they can view updates and continue to donate multiple times. There can be a sense of connectivity to the research by viewing updates on the campaign.

Acknowledgments

The authors would like to acknowledge and thank the John Templeton Foundation for funding this research.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Interest Form.

[PDF File (Adobe PDF File), 69KB - ijmr_v7i1e3_app1.pdf]

Multimedia Appendix 2

Getting started worksheet.

[PDF File (Adobe PDF File), 176KB - ijmr_v7i1e3_app2.pdf]

Multimedia Appendix 3

Consent form.

[PDF File (Adobe PDF File), 101KB - ijmr_v7i1e3_app3.pdf]

Multimedia Appendix 4

Doctor's note.

[PDF File (Adobe PDF File), 167KB - ijmr_v7i1e3_app4.pdf]

Multimedia Appendix 5

Crowdfunding best practices: steps & stories to help you launch a successful campaign.

[PDF File (Adobe PDF File), 4MB - ijmr_v7i1e3_app5.pdf]

Multimedia Appendix 6

Donor's survey.

[PDF File (Adobe PDF File), 542KB - ijmr_v7i1e3_app6.pdf]

References

1. Stavropoulos DJ, Merico D, Jobling R, Bowdin S, Monfared N, Thiruvahindrapuram B, et al. Whole-genome sequencing expands diagnostic utility and improves clinical management in paediatric medicine. NPJ Genom Med 2016 Jan 13;1:15012. [doi: 10.1038/npjgenmed.2015.12]
2. Yang Y, Muzny DM, Xia F, Niu Z, Person R, Ding Y, et al. Molecular findings among patients referred for clinical whole-exome sequencing. J Am Med Assoc 2014 Nov 12;312(18):1870-1879 [FREE Full text] [doi: 10.1001/jama.2014.14601] [Medline: 25326635]
3. Farwell KD, Shahmirzadi L, El-Khechen D, Powis Z, Chao EC, Tippin DB, et al. Enhanced utility of family-centered diagnostic exome sequencing with inheritance model-based analysis: results from 500 unselected families with undiagnosed genetic conditions. Genet Med 2015 Jul;17(7):578-586. [doi: 10.1038/gim.2014.154] [Medline: 25356970]
4. Hayden E. Nature. The $1,000 genome URL: http://www.nature.com/polopoly_fs/1.14901!/menu/main/topColumns/topLeftColumn/pdf/507294a.pdf, [accessed 2017-10-18] [WebCite Cache ID 6uJNIgOsO]
5. Makrythanasis P, Antonarakis SE. Pathogenic variants in non-protein-coding sequences. Clin Genet 2013 Nov;84(5):422-428. [doi: 10.1111/cge.12272] [Medline: 24007299]
6. Ezkurdia I, Juan D, Rodriguez JM, Frankish A, Diekhans M, Harrow J, et al. Multiple evidence strands suggest that there may be as few as 19,000 human protein-coding genes. Hum Mol Genet 2014 Nov 15;23(22):5866-5878 [FREE Full text] [doi: 10.1093/hmg/ddu309] [Medline: 24939910]
7. Lusardi A, Schneider D. NBER. 2011. Financially fragile households: evidence and implications URL: http://www.nber.org/papers/w17072.pdf [accessed 2017-12-26] [WebCite Cache ID 6vzgKc2Av]
8. Sharma A, Khan J, Khan J, Devereaux PJ. Is crowdfunding a viable source of clinical trial research funding? Lancet 2015 Jul 25;386(9991):338. [doi: 10.1016/S0140-6736(15)61407-6] [Medline: 26227465]

9. Burtch G, Chan J. Aisnet. 2014. Reducing medical bankruptcy through crowdfunding: evidence from giveforward URL: https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1463&context=icis2014
10. Mollick ER, Kuppuswamy V. SSRN. 2014 Jan 11. After the campaign: outcomes of crowdfunding URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2376997 [accessed 2017-12-26] [WebCite Cache ID 6vzgr7uPN]
11. Greenberg J, Mollick E. Leaning in or leaning on? Gender, homophily, and activism in crowdfunding. Acad Manage Proc 2015 Jan. [doi: 10.5465/AMBPP.2015.18365abstract]
12. Marom D, Robb A, Sade O. SSRN. 2016 Feb 23. Gender dynamics in crowdfunding (kickstarter): evidence on entrepreneurs, investors, deals and taste-based discrimination URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2442954 [accessed 2017-12-26] [WebCite Cache ID 6vzgXQiWs]
13. Kuppuswamy V. SSRN. 2017. Crowdfunding creative ideas: the dynamics of projects backers in kickstarter URL: http://ssrn.com/abstract=2234765 [accessed 2017-10-18] [WebCite Cache ID 6uJNUTceh]
14. Facio F, Eidem H, Fisher T, Brooks S. Intentions to receive individual results from whole-genome sequencing among participants in the ClinSeq study. Eur J Hum Genet 2013 Mar 23;21(3):261-265. [doi: 10.1038/ejhg.2012.179] [Medline: 22892536]

Abbreviations

RG: Rare Genomics
WES: whole exome genetic sequencing

Edited by A Keepanasseril; submitted 24.12.16; peer-reviewed by A Sharma, K Kaphingst; comments to author 12.04.17; revised version received 29.08.17; accepted 21.11.17; published 05.02.18.

Please cite as:
Ortiz RA, Witte S, Gouw A, Sanfilippo A, Tsai R, Fumagalli D, Yu C, Lant K, Lipitz N, Shepphird J, Alvina FB, Cheng-Ho Lin J
Engaging a Community for Rare Genetic Disease: Best Practices and Education From Individual Crowdfunding Campaigns
Interact J Med Res 2018;7(1):e3
URL: http://www.i-jmr.org/2018/1/e3/
doi:10.2196/ijmr.7176
PMID:29402763

Original Paper

Usage, Acceptability, and Effectiveness of an Activity Tracker in a Randomized Trial of a Workplace Sitting Intervention: Mixed-Methods Evaluation

Charlotte L Brakenridge[1], BPsySc (Hons); Genevieve N Healy[1,2,3], PhD; Elisabeth AH Winkler[1], PhD; Brianna S Fjeldsoe[1], PhD

[1]School of Public Health, The University of Queensland, Brisbane, Australia
[2]Baker Heart and Diabetes Institute, Melbourne, Australia
[3]School of Physiotherapy and Exercise Science, Curtin University, Perth, Australia

Corresponding Author:
Charlotte L Brakenridge, BPsySc (Hons)
School of Public Health
The University of Queensland
Herston Rd
Brisbane, 4006
Australia
Phone: 61 733655163
Email: c.brakenridge@uq.edu.au

Abstract

Background: Wearable activity trackers are now a common feature of workplace wellness programs; however, their ability to impact sitting time (the behavior in which most of the desk-based workday is spent) is relatively unknown. This study evaluated the LUMOback, an activity tracker that targets sitting time, as part of a cluster-randomized workplace sitting intervention in desk-based office workers.

Objective: Study objectives were to explore: (1) office workers' self-directed LUMOback use, (2) individual-level characteristics associated with LUMOback use, (3) the impact of LUMOback use on activity and sitting behaviors, and (4) office workers' perceived LUMOback acceptability.

Methods: Exploratory analyses were conducted within the activity tracker intervention group (n=66) of a 2-arm cluster-randomized trial (n=153) with follow-up at 3 and 12 months. The intervention, delivered from within the workplace, consisted of organizational support strategies (eg, manager support, emails) to stand up, sit less, and move more, plus the provision of a LUMOback activity tracker. The LUMOback, worn belted around the waist, provides real-time sitting feedback through a mobile app. LUMOback usage data (n=62), Web-based questionnaires (n=33), activPAL-assessed sitting, prolonged (≥30 min bouts) and nonprolonged (<30 min bouts) sitting, standing and stepping time (7-day, 24 h/day protocol; n=40), and telephone interviews (n=27) were used to evaluate study aims. LUMOback usage data were downloaded and described. Associations between user characteristics and LUMOback usage (in the first 3 months) were analyzed using zero-inflated negative binomial models. Associations between LUMOback usage and 3-month activity outcomes were analyzed using mixed models, correcting for cluster. LUMOback acceptability was explored using 3-month questionnaire data and thematic analysis of telephone interviews (conducted 6 to 10 months post intervention commencement).

Results: Tracker uptake was modest (43/61, 70%), and among users, usage over the first 3 months was low (1-48 days, median 8). Usage was greatest among team leaders and those with low self-perceived scores for job control and supervisor relationships. Greater tracker use (≥5 days vs <5 days) was significantly associated only with changes in prolonged unbroken sitting (−50.7 min/16 h; 95% CI −94.0 to −7.3; *P*=.02) during all waking hours, and changes in nonprolonged sitting (+32.5 min/10 h; 95% CI 5.0 to 59.9; *P*=.02) during work hours. Participants found the LUMOback easy to use but only somewhat comfortable. Qualitatively, participants valued the real-time app feedback. Nonuptake was attributed to being busy and setup issues. Low usage was attributed to discomfort wearing the LUMOback.

Conclusions: The LUMOback—although able to reduce prolonged sitting time—was only used to a limited extent, and its low usage may provide a partial explanation for the limited behavior changes that occurred. Discomfort limited the feasibility of the

LUMOback for ongoing use. Such findings yield insight into how to improve upon implementing activity trackers in workplace settings.

(Interact J Med Res 2018;7(1):e5) doi:10.2196/ijmr.9001

KEYWORDS

wearable electronic devices; fitness trackers; sedentary lifestyle; exercise; workplace; adult

Introduction

Desk-based workers engage in high levels of sedentary time [1,2] (ie, waking behaviors spent in a sitting, reclining, or lying posture with low energy expenditure [3]), putting themselves at risk of musculoskeletal [4] and cardio-metabolic issues [5]. Much of this sitting time is in prolonged, unbroken bouts [1], which have additional cardio-metabolic risks compared with sitting with regular breaks [5]. Furthermore, desk-based workers spend only a small percentage of their workday stepping (<10% of work time) [2] and engage in minimal amounts of moderate to vigorous activity both in and outside of work hours [6]. Given that sedentary work is increasing [7] and workers spend the majority of their day at work [8], there is a need to design and evaluate interventions that target the desk-based workplace.

In recent years, consumer-based activity trackers have become a popular component of workplace wellness programs to increase workers' physical activity [9], and it is expected that by 2021, 171.9 million activity trackers will be in workplaces worldwide [10]. Most trackers offer features known to be important for behavior change including self-monitoring, real-time feedback and prompting, guided goal setting, and rich and tailored feedback data [11]. Their popular use in workplace wellness programs indicates that they may be a feasible and acceptable intervention strategy for organizations to disseminate [9], and there is emerging evidence to suggest that consumer-based activity trackers can improve physical activity outcomes in desk-based workers [12,13].

Despite these findings, investigation concerning how desk-based workers use consumer-based trackers, which worker characteristics are associated with use, and whether workers find trackers acceptable and appealing has been scarce. The limited evidence suggests that although initial engagement is high [12,13], use may drop off in the long term [13,14]. Tracker users have been shown to fall into distinct groups, with high use seen among young and fit adults, as well as older adults with a desire to improve their health [15]. Acceptability may depend on user motivation and ease of use of the tracker [14]. Furthermore, it is also unknown whether use of a consumer-based activity tracker that targets and provides real-time *sitting* feedback can result in reductions in desk-based workers' sitting time. Strategies that target increases in activity may not lead to reductions in sitting time [16], and therefore, trackers that specifically target sitting time need to be evaluated.

To address these gaps, this study explored the usage and acceptability of an activity tracker (the LUMOback) that provided real-time feedback on sitting among desk-based office workers participating in a cluster-randomized workplace trial, targeting reductions in prolonged sitting time [17,18]. The LUMOback tracker was chosen over other activity trackers because it is one of the few commercially available trackers that specifically measures and provides real-time sitting feedback [19]. The aim of this study was to explore the following: (1) participants' LUMOback usage (ie, device wear); (2) which personal, health, job, and activity characteristics were associated with greater LUMOback usage; (3) the relationship between LUMOback usage and changes in sitting and activity outcomes both at work and across all waking hours; and (4) participants' acceptability and perceptions of the LUMOback.

Methods

Participants and Design

Details of the study [17] and the effectiveness outcomes [18] have been reported previously (trial registration: ACTRN12614000252617). Briefly, participants were desk-based office employees of a large international property and infrastructure group. Workers were recruited in teams from 2 locations (A and B). Teams were cluster-randomized to receive either organizational support strategies alone (Group ORG; 9 manager-led teams; 87 participants) or organizational support strategies plus the provision of a LUMOback tracker (Group ORG+Tracker; 9 manager-led teams; 66 participants). The intervention was ongoing; data collection occurred at baseline and 3 and 12 months post intervention commencement. Only the Group ORG+Tracker is examined in this study. The organizational support component, designed to be workplace driven, was delivered by a workplace champion (the organization's Head of Workplace Wellbeing) and consisted of an information booklet; emails with tips to stand up, sit less, and move more; and involvement of senior executives to support the study. Participants in both groups also received individualized feedback concerning their sitting time and physical activity as recorded by the activPAL after each assessment. The study was approved by the University of Queensland Behavioural and Social Sciences Ethical Review Committee (approval number: 2014000089). All participants provided written informed consent, which covered the 3-month evaluation, with reconsent required to participate in the 12-month evaluation [17].

The LUMOback Tracker

The activity tracker provided to participants was the LUMOback activity tracker (LUMO Bodytech, Mountain View, CA, USA). The LUMOback is worn as a belt around the lower back (over or under clothing) and provides real-time feedback (through a mobile app) on sitting, standing, sit-to-stand transitions, walking, running, step count, posture, and sleep. Users can set posture alerts that vibrate the belt when the user is in a poor lumbar posture (identified by pelvic tilt angle) and set sitting time push notifications to their phone when sitting for periods of 15 min, 30 min, 45 min, 1 hour, or 2 hours (user-defined). The LUMOback provides valid and reliable measures of sitting time and number of steps [20,21].

LUMOback trackers, along with a 4-page instruction booklet, were distributed by the workplace champion in study week 1. Participants set up their LUMObacks in their own time but had the contact details of the project coordinator and the workplace champion for any questions. LUMOback setup required calibration to the user's sitting and standing posture, and this was guided through the app and booklet instructions. The LUMOback was originally designed as a posture rather than as a sitting device [19]. As such, the default settings of the tracker were to have vibrating alerts (based on posture) on and sitting notifications off. To be consistent with evaluating the use of the tracker as it would be used outside of the research context, these settings were not preadjusted by project staff; however, the booklet instructions indicated that both these settings could be changed by the participant if the participant wished. LUMOback data were downloaded by project staff and monitored for usage during the study period. No specific instruction was given to participants on how much to wear the LUMOback; however, participants who did not have any LUMOback usage data were followed up by the project coordinator (via email or phone) or by the workplace champion (in person) to troubleshoot issues about setup during weeks 3 to 10. Participants were permitted to keep the LUMOback after the study ended.

Data Collection

Usage data were collected from LUMO Bodytech every 2 weeks during the initial 3 months and periodically thereafter. Participants completed Web-based questionnaires (via LimeService, used to address aims 2 and 4) at baseline, 3 months, and 12 months. The questionnaires collected data on sociodemographics (baseline), health and job characteristics (baseline, 3 months, and 12 months), LUMOback acceptability (3 months), and usage of behavior change techniques (eg, goal setting and prompting) in the LUMOback app (3 and 12 months). Height and weight measurements were collected objectively in person at the baseline assessment or via self-report (used in aim 2). Sitting and activity data (used to address aims 2 and 3) were collected at baseline, 3 months, and 12 months via the activPAL3 activity monitor (PAL Technologies Ltd., Glasgow, Scotland, UK). Telephone interviews, conducted 6 to 10 months after intervention commencement, were used to collect qualitative data regarding the LUMOback's acceptability to participants (used to address aim 4). Because so few participants reconsented for the 12-month assessment, aims 2 and 3 consider only usage and outcomes over the first 3 months.

Measures

Personal, Health, and Job Characteristics

Key participant characteristics, as measured at baseline, were age, gender, highest level of education completed, job category (senior or team leader, managerial staff, other general staff), use of other activity-promoting apps and wearable devices (yes or no), confidence using a smartphone (1-5 scale, least to most confidence), desired sitting at work (≥50%, <50% of work time) [22], knowledge of the health impacts of sitting (1-5 scale, least to most knowledge) [22], musculoskeletal symptoms (Nordic Musculoskeletal Questionnaire [23], categorized as problems over the last month [yes or no] in the upper body, back, and lower extremities), and physical and mental health-related quality of life (Short Form-12 version 1 [24], 0-100 scale) and several other health- and job-related scores, all measured on a 1-10 scale. Specifically, these were stress [25], job control [25], work satisfaction [25], supervisor relations [25] (all from the Health and Work Questionnaire), and job performance (self-rated job performance scale [26]). For all of the health- and job-related scores, apart from stress, higher values are desirable. Height and weight measurements were used to calculate body mass index at baseline.

Activity

Sitting and activity data were collected via the activPAL3 monitor. At each assessment, participants were asked to wear the monitor for 7 days, 24 hours per day. Further information on monitor processing has been provided in previous trial publications [17,18]. Activity outcomes considered here are time per 10 hours at work and time per 16 hours awake spent engaged in sitting, prolonged sitting (≥30 min continuously), nonprolonged sitting (<30 min continuously), standing, and stepping. To take into account variations in work times between assessments, and between individuals, work hour outcomes were standardized to a 10-hour workday by multiplying by 10 hours and dividing by work wear time. Similarly, overall waking hour outcomes were standardized to a 16-hour day by multiplying by 16 hours and dividing by waking wear time. Work and waking times were collected by self-report.

LUMOback Usage

For participants who set up the LUMOback, data on LUMOback usage were downloaded as comma-separated value files from LUMO Bodytech. Each file contains date and time-stamped summaries per 5 min time interval on activity and posture in various metrics, and an indicator of whether the LUMOback was not worn or charging. Daily usage (hours/day) was estimated by calculating the time elapsed between the first and last 5-min time window each day during which the LUMOback recorded an activity other than lying down, and then subtracting from this usage time any probable device nonwear time (or sleep time) occurring during this period. Only days with ≥1 hour of usage were counted as days of usage. Participants with at least 1 day of usage were classed as LUMOback "users"; those with no days of usage, or who self-reported never using the LUMOback, were classed as "nonusers." Others who reported using the LUMOback but lacked usage data were classed as having an "unknown" degree of usage.

LUMOback usage over the initial 3 months of intervention was considered: continuously (days), as a binary classification (≥5 days or <5 days), and as a 4-category classification (nonuser, limited user, infrequent user, or frequent user). The thresholds used for the 4-category classification were 0, 1-4, 5-15, and ≥16 days, respectively, over 3 months. Long-term (12-month) usage was also assessed for the subset of participants who had opted for the 12-month assessment. LUMOback usage during the activity monitor assessment (yes or no) was also considered to provide some insight as to whether the timing of LUMOback wear (not just how much it was worn) was important for activity changes. Usage/nonusage during the activity monitor assessment was determined by the presence/absence of ≥1 day with ≥1 hour of usage over this time frame, or by self-report when LUMOback data were not available.

LUMOback Acceptability

All questionnaire items assessing LUMOback acceptability were created for the trial. Acceptability questionnaire items were closed response (1, not at all; 2, somewhat; 3, not sure; 4, comfortable/easy; and 5, extremely comfortable/easy) and assessed for the following: comfort of the LUMOback, the ease of setup of the LUMOback and app, ease of navigation and use of the app, and ease of LUMOback calibration to the user's posture. Participants' perceived usefulness of the LUMOback overall and of specific features such as the graphs, sitting notifications, and vibrating alerts were also assessed in closed-response items: 1, not at all; 2, somewhat; 3, not sure; 4, useful; 5, extremely useful; and, when relevant, 6, did not notice/use. Frequency of use of the behavior change features was also assessed (eg, "When wearing your LUMOback, how often did you have it set to vibrate?"), with response options from 1 (never) to 6 (all the time). The full list of questions is provided in Multimedia Appendix 1 of the study protocol [17].

All participants who had not withdrawn after the 3-month assessment were eligible to take part in the telephone interviews. To ensure representation across teams, it was planned that at least 2 participants per team would be interviewed, plus the team managers. Interviews continued until the data were saturated and at least 2 members per work team were interviewed, unless there were more than 5 unsuccessful attempts to interview. The lead author (CLB, in location B) conducted and recorded (using Audacity) one-on-one telephone interviews with participants (predominantly in location A), ranging in duration from 9 to 28 min (mean duration 14 min). These semi-structured interviews covered topics of general thoughts on the LUMOback, participant likes and dislikes, barriers to usage, use of the sitting notification, influence of work team on LUMOback use, and general activity tracker preferences.

Sample Size

The sample size was determined based on the needs of the cluster-randomized trial [17]. This study was exploratory and not powered on study aims. Minimum differences of interest (MDI) in this study were 30 min of sitting and standing and 15 min of stepping between LUMOback usage categories (≈per 10 days of usage), and 2 days of LUMOback usage between groups (categorical variables), per amount of activity equivalent to the above, or per 1 standard deviation of other continuous variables.

Analysis

Statistical analyses were performed in STATA version ≥13 (StataCorp, College Station, TX, USA). Significance was set at $P<.05$ 2-tailed. Analyses were of all evaluable cases with missing or "unknown" data excluded. Mean and standard deviation were reported for continuous variables, and percentage was reported for categorical variables.

To address aim 1, usage was described. In the case of non-normally distributed variables, the median and the 25th and 75th percentiles were reported in addition to the mean and SD. To address aim 2, associations of characteristics with LUMOback use (adjusting for age, sex, and location) were statistically tested using zero-inflated negative binomial models (which accounted for the excessive zeros, as indicated by the Vuong test). Results are reported as contrasts of marginal mean outcomes. To address aim 3, associations of LUMOback usage with activity outcomes were examined using linear mixed models, adjusting for baseline value of the outcome, age, and sex and correcting for team cluster. Models were checked for nonlinearity, non-normality of residuals, multicollinearity, heteroscedasticity, and influential cases. Outcomes were transformed when this improved models, with the results from these models presented back-transformed to regular units. Results are presented with and without influential cases, if these cases were found to change the interpretation of the results.

To address aim 4, acceptability of the LUMOback was described quantitatively, based on the questionnaire items, and qualitatively, based on thematic analysis of interview data. Qualitative interviews were transcribed verbatim using F4 software (audiotranskription, Marburg, Germany). Participant idiosyncrasies (eg, "um") were removed after transcription. Author CLB coded initial themes (using NVivo version 10 and Microsoft Office Word). These were then compared and contrasted with a second researcher and discussed and decided upon. Themes were not preidentified; however, some of the questions in the semi-structured interview had a narrow focus that heavily contributed to some of the themes. Themes were then compared and contrasted with the questionnaire data when these were available, with findings from both sources presented together.

Results

Participant Characteristics

Baseline characteristics of participants in the ORG+Tracker group (overall and by usage) are presented in Multimedia Appendix 1. Participants (35 female and 31 male) on average were aged 37.6 (SD 7.8) years and engaged in high levels of sitting time during work hours (453.0 [SD 55.9] min/10 h, ie, 76% of work hours). The majority were university educated (54/63, 86%); either in managerial (38/66, 58%), team or senior leader (8/66, 12%), or general staff (20/66, 30%) positions; and confident to extremely confident using a smartphone (52/54, 96%).

Aim 1: LUMOback Uptake and Usage

Uptake

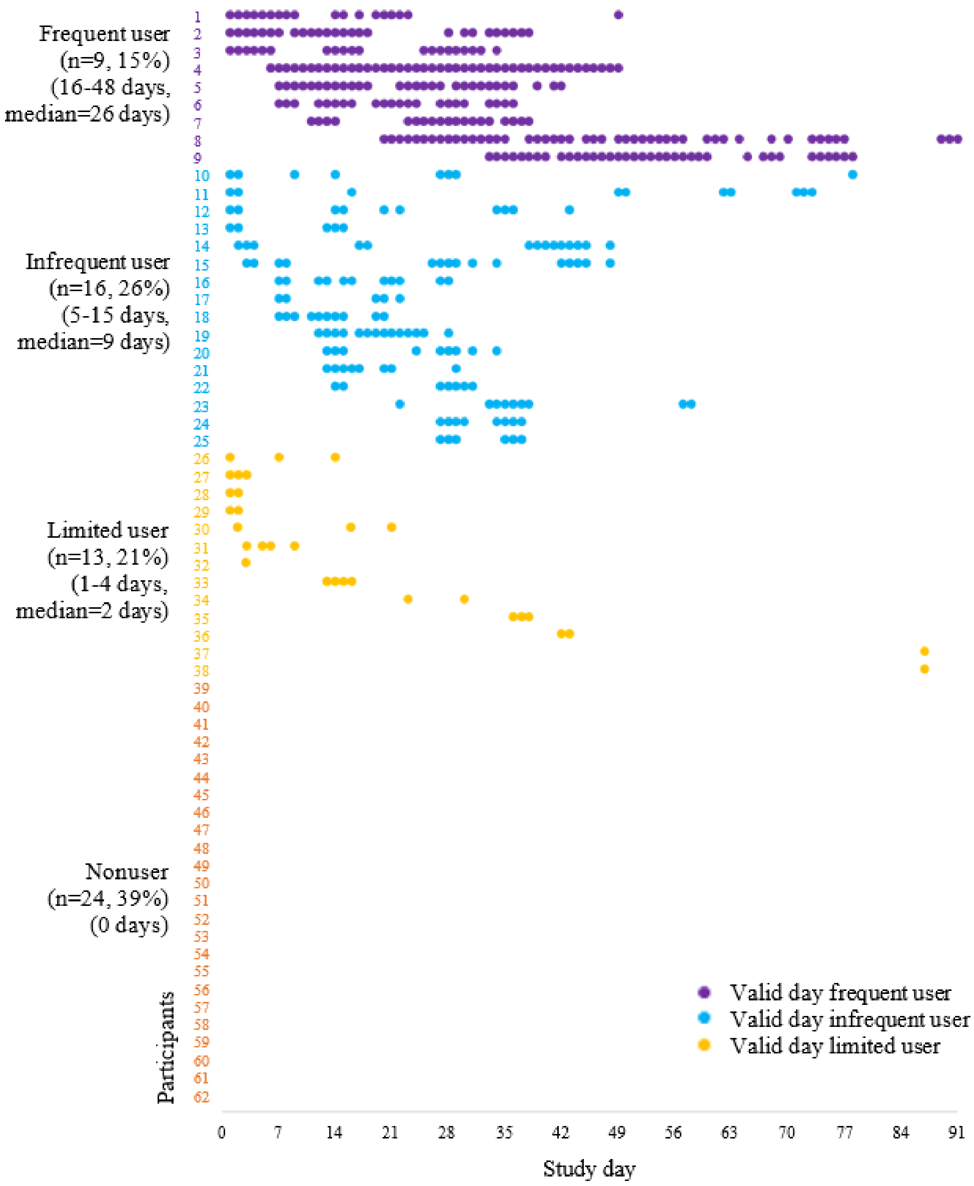
The LUMOback was provided to 61 participants. Overall, 5 participants did not receive the LUMOback because of either ineligibility (incompatible phone, not ambulatory, left organization; n=3) or refusal (n=1), or for an unknown reason (n=1; see Multimedia Appendix 2 for flow diagram). Uptake was modest overall (43/66, 65%) and within those provided a LUMOback (43/61, 70%). Nonusers' (n=18) reported reasons for nonuse were as follows: being too busy to set up (6/18, 33%), personal disruptions during the study period (4/18, 22%), having technical difficulties setting up (3/18, 17%), pregnancy (1/18, 6%), and unknown (4/18, 22%). An additional "user" was later classed as a nonuser as they only had a very small and invalid amount of usage data (<1 hour), bringing the total number of nonusers to 19. Nonusers of the LUMOback were less likely to provide 3-month follow-up activity data than LUMOback users (see Multimedia Appendix 2).

Short-Term Usage

Out of the remaining 42 users, 4 participants claimed to have used the LUMOback, but usage data could not be obtained from LUMO Bodytech. Among those whose usage data could be collected (n=38), the LUMOback was used over the first 3 months for 1 to 48 of a possible 91 days, with a mean of 11.7 (SD 11.8) days of use and median (25th, 75th percentile) of 8 (3, 15) days. As usage can be intermittent, the time from first to last recorded usage within the first 3 months was slightly longer than the days of usage, ranging from 1 to 78 days, with a mean of 24.5 (SD 21.3) days and median (25th, 75th percentile) of 17.5 (7, 38) days. Duration of usage on wear days averaged 9.8 (SD 3.7) h per day.

Figure 1 demonstrates the diverse patterns of usage during the first 3 months of the study by usage group. Only 9 participants (15%) were classed as frequent users, whereas the remaining participants were either infrequent users (16/62, 26%), limited users (13/62, 21%), or nonusers (24/62, 39%). Duration of usage on wear days (h/day) appeared to be longest in frequent users (mean 12.8, SD 1.6), followed by infrequent users (mean 10.0, SD 3.2), and then by limited users (mean 7.4, SD 3.8).

Figure 1. LUMOback use by study day over the first 3 months (91 days), ordered by first study day of use and total length of use within each usage category. Percentages do not add to 100% due to rounding.



Long-Term Usage

LUMOback usage after the 3-month assessment was evaluated in the 21 LUMOback users who reconsented to 12-month follow-up (reasons for nonconsent were not related to LUMOback usage; see Multimedia Appendix 2). Only 1 consenting participant used their LUMOback after the 3-month assessment, for 5 consecutive additional days in total, with the last day of use being day 185 (ie, approximately 6 months after intervention commencement).

Aim 2: Participant Characteristics Associated With Short-Term LUMOback Usage

Figure 2 shows the number of days of LUMOback use compared across personal, health, job, and activity characteristics. Job control and supervisor relations were the only statistically significant predictors of usage. For both variables, higher perceived scores were associated with fewer days of LUMOback usage (~3 days). Some of the nonsignificant differences were substantial (≥the MDI). Specifically, being a senior or team leader rather than managerial staff was meaningfully associated with more days (~14 days) of LUMOback usage, whereas having higher perceived job performance, being a smoker, being overweight or obese, and having back or lower extremity musculoskeletal problems were meaningfully associated with fewer days (~2 to 4 days) of LUMOback usage. Other differences were small (<the MDI).

Aim 3: Impact of Short-Term LUMOback Usage on Behavior Change

LUMOback usage for at least 5 days during the intervention period was significantly associated with more nonprolonged sitting time during work hours (+32.5 min/10 h workday, 95% CI 5.0 to 59.9, *P*=.02). An examination of Cook’s distance (*D*) identified 2 influential cases (Cook’s *D* ≥.14), and after removal of these cases, the association with nonprolonged sitting time was no longer statistically significant, although still tended in the same direction (+19.6 min/10 h, 95% CI −3.6 to 42.7, *P*=.10). Increases in nonprolonged sitting time were coupled with less prolonged sitting time (−38.1 min/10 h, 95% CI −80.1 to 3.9, *P*=.08), and only very small differences in sitting, standing, and stepping time changes during work hours (see Table 1).

During waking hours, those using the LUMOback for at least 5 days reduced their prolonged sitting time by significantly more than their lower usage counterparts (−50.7 min/16 h awake, 95% CI −94.0 to −7.3, *P*=.02). This comparatively greater shift in waking hours away from prolonged sitting time was coupled with a shift toward more nonprolonged sitting time (+25.4 min/16 h, 95% CI −5.7 to 56.6, *P*=.11), which became significant after the removal of 1 influential case (Cook’s *D*=.25; +31.9 min/16 h, 95% CI 3.3 to 60.4, *P*=.03). In addition, there remained some small differences in sitting (−29.4 min/16 h, 95% CI −67.2 to 8.4, *P*=.13), standing (+21.1 min/16 h, 95% CI −10.7 to 52.9, *P*=.19), and stepping (+4.6 min/16 h, 95% CI −9.2 to 18.3, *P*=.51) changes, all favoring those who used the LUMOback for at least 5 days.

Figure 2. Predictors of number of days of LUMOback usage over the first 3 months (91 days) of the intervention. (Gray indicates personal characteristics, green indicates health characteristics, blue indicates job characteristics, orange indicates work activity characteristics, and red indicates overall activity characteristics. Boldface refers to difference ≥ the minimum difference of interest. Marginal mean [95% CI] days of LUMOback use [zero inflated negative binomial models], adjusted for sex, age, and location contrasted between groups [categorical variables] or mean versus “mean + dose” [continuous variables]. Dose was 1 SD for most continuous variables, 6 min for sitting and standing, and 3 min for stepping, per day or workday. The CI for senior or team leader continues off figure.).

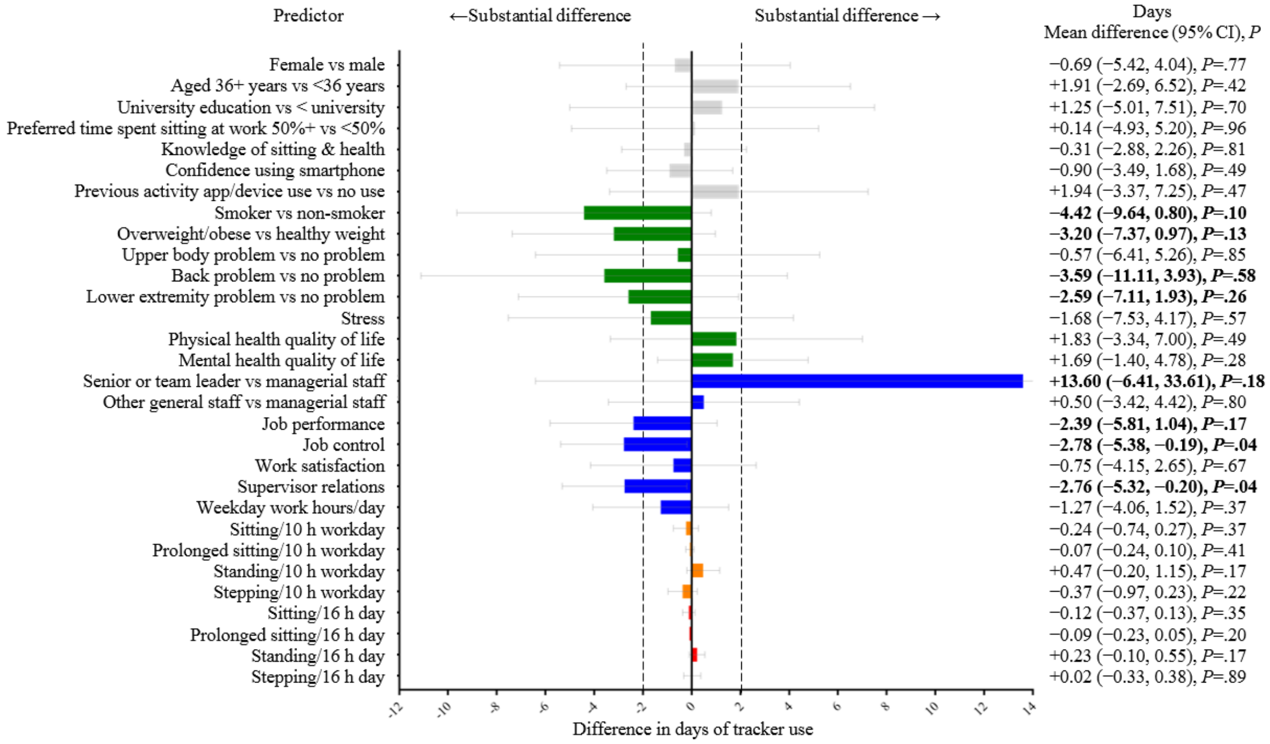


Table 1. Associations of LUMOback usage with 3-month activity changes during the intervention.

Outcome	Used during the intervention for ≥5 days, yes vs no (adjusted mean difference, 95% CI)[a]	*P* value	Usage during activity assessment, yes vs no (adjusted mean difference, 95% CI)[a]	*P* value
Work hours, min/10 h workday[b]				
Sitting	−1.9 (−22.0 to 18.1)[c]	.85	−7.5 (−27.8 to 12.9)[c]	.47
Prolonged sitting	−38.1 (−80.1 to 3.9)	.08	−28.3 (−74.0 to 17.4)	.23
Nonprolonged sitting	*32.5 (5.0 to 59.9)*	*.02*	24.7 (−8.0 to 57.4)	.14
Standing	0.7 (−17.3 to 18.7)	.94	−1.5 (−18.8 to 15.7)	.86
Stepping	2.6 (−7.5 to 12.8)	.61	5.3 (−4.4 to 15.0)	.29
All waking hours, min/16 h day[b]				
Sitting	−29.4 (−67.2 to 8.4)	.13	−2.0 (−40.0 to 36.0)	.92
Prolonged sitting	*− 50.7 (−94.0 to −7.3)*	*.02*	−35.1 (−79.4 to 9.2)	.12
Nonprolonged sitting	25.4 (−5.7 to 56.6)	.11	*36.5 (7.5 to 65.6)*	*.01*
Standing	21.1 (−10.7 to 52.9)	.19	4.2 (−27.1 to 35.5)	.79
Stepping	4.6 (−9.2 to 18.3)	.51	−1.0 (−14.2 to 12.2)	.89

[a] Table shows adjusted mean difference and 95% CI from linear mixed models, adjusting for baseline value of the outcome, age, and sex, and correcting for cluster. Italics indicate statistically significant association (set at *P*<.05 2-tailed).

[b] Analyzes are of evaluable cases among the 66 ORG+Tracker participants, excluding those with missing activity outcomes at work (n=30), during waking hours (n=26), or with unknown usage during the intervention (n=3). Work hours analyses include 18 users and 15 nonusers during the intervention and 21 users and 15 nonusers during the assessment while waking hours analyses include 19 users and 18 nonusers during the intervention and 22 users and 18 nonusers during the assessment.

[c] Sitting time during work hours was modeled as natural log of (600 minutes minus sitting minutes); results are back-transformed to original units for presentation in the table.

Examining usage of the LUMOback specifically over the time frame that activity was being measured revealed that concurrent usage was significantly associated with more nonprolonged sitting time during waking hours (+36.5 min/16 h, 95% CI 7.5 to 65.6, *P*=.01) and, accordingly, less prolonged sitting time (−35.1 min/16 h, 95% CI −79.4 to 9.2, *P*=.12; which became significant after the removal of 2 influential [Cook's *D* ≥.11] cases: −42.0 min/16 h, 95% CI −81.4 to −2.6, *P*=.04). Otherwise, there were only very small differences in sitting, standing, and stepping time changes. The size of the observed differences in work hour activities were all small (<the MDI) in addition to not reaching statistical significance.

Aim 4: Acceptability and Perceptions of the LUMOback

At 3 months, questionnaire data were collected from 33 LUMOback users (13 male; aged 26-55 years; 4 team leaders, 1 senior leader, 15 managerial staff, and 13 general staff), and qualitative interviews were conducted with 22 LUMOback users (11 male; aged 26-57 years; 5 team leaders, 11 managerial staff, and 6 general staff) and 5 nonusers (3 male; aged 36-45 years; 1 team leader and 4 managerial staff). Several themes emerged from these 2 sources of data on LUMOback acceptability: (1) comfort of the LUMOback; (2) acceptability of the app, real-time feedback, and graphs; (3) use of the LUMOback features; (4) additional barriers; (5) perceived barriers in nonusers; (6) the perceived influence of others on LUMOback use; and (7) suggested improvements in an activity tracker. Supportive quotes are provided below and in Multimedia Appendix 3.

Comfort of the LUMOback

A majority of survey respondents (21/33, 64%) reported the LUMOback as only being "somewhat comfortable," 9 participants (27%) reported the LUMOback to be "not at all comfortable," and only 3 participants (9%) reported the LUMOback to be "comfortable" (see Multimedia Appendix 4). Comfort was identified as an issue in the telephone interviews. Reported issues included being conscious of wearing the LUMOback, the waist placement, sweating when wearing it, and, for women in particular, being hard to wear with clothing.

> *I think it was something that was hard to wear with some types of clothes. Not something that you could just put on and not think about. It was certainly thought about all the time while you were wearing it.* [#3, female managerial staff, limited user]

Acceptability of the App, Real-Time Feedback, and Graphs

Most users perceived the LUMOback and app to be "easy" to "extremely easy" to set up (n=22/33, 67%), use and navigate (25/33, 76%; see Multimedia Appendix 4), and the feedback graphs to be at least "somewhat useful" for creating awareness of sitting, standing, stepping, and good posture behaviors (27/33 (82%), 27/33 (82%), 26/33 (79%), and 26/33 (79%), respectively; see Multimedia Appendix 5). In the interviews, many participants reported that it was informative to get the real-time feedback from the LUMOback app:

> *I liked that you could track your results and see how well you are doing, I think that was the most benefit*

> *that you could constantly check where you're at.* [#25, female general staff, infrequent user]

Use of the LUMOback Features

Use of the sitting notification was modest, with half (16/33, 49%) not using it at all (Multimedia Appendix 5). Encouragingly though, most notification users (14/17, 82%) found the prompts to be at least "somewhat useful" for creating sitting awareness. Participants were most likely to select to be prompted for a sitting break every 30 min (11/17, 65%), compared with 15 min (2/17, 12%), 1 hour (3/17, 18%), or 2 hours (1/17, 6%).

The minimal use of the sitting notifications was echoed in the findings from the telephone interviews. When prompted, it appeared that not all participants were aware of the option to set sitting notifications or had forgotten. Out of those interviewees who did report using the sitting notification, some reported that the sitting notification was a "nice feature," with one participant going on to say:

> *...[the sitting notification] actually became a useful conversation piece, you would be in a meeting and I would get a prompt and I'd stand up and I would explain to people what was going on, people thought it was interesting, so you are not just doing something weird like suddenly popping up in the middle of a meeting.* [#21, male team leader, infrequent user]

However, other users reported that the notifications were annoying and they turned them off, or they were easy to ignore.

The vibrating posture alerts were widely used and were highly discussed by participants. Some participants found the vibrating alerts beneficial:

> *...just the intrusiveness of it was actually quite good because it was always front of mind, you kind of couldn't forget that it was there, especially when you had the pulse thing on.* [#19, female managerial staff, limited user]

However, many participants found that the vibrating alerts were a key barrier to wear as they were annoying and/or caused distraction in meetings. Although some participants were happy once disabling the vibrating alerts, others felt the LUMOback had lost its utility and purpose without the vibrations:

> *I did [turn the vibrating alerts off], but then I thought it's a bit pointless because, not doing anything.* [#17, male general staff, infrequent user]

Additional Barriers

Some participants reported frustration having to repeatedly calibrate the LUMOback so that it was correctly measuring activity and posture and providing correct prompting:

> *You needed to recalibrate it every now and then and it would buzz as a result or it would be buzzing when you are in a right position and it was just a bit annoying.* [#20, male managerial staff, frequent user]

This was congruent with the questionnaire findings, with only just over half of participants (17/33, 52%) reporting the LUMOback was easy to calibrate (see Multimedia Appendix 4). Forgetting to wear the LUMOback was another barrier, with some participants reporting that the monitor used for assessments (ie, the activPAL) was easier to wear because "it just sat there all the time" and they did not need to remember to take it off and put it on each day (as it was taped to their thigh). Other barriers reported by a small number of participants included not wanting to wear the LUMOback at home because of the desire to sit and slouch and not be prompted during this time, having to charge the LUMOback, and the LUMOback app draining phone battery.

Perceived Barriers in Nonusers

Participants who did not wear the LUMOback were also interviewed and elaborated on their perceived barriers to uptake. A small number of participants reported set-up or syncing issues with the LUMOback. One of these participants mentioned that it would have been helpful having in-person assistance in setting up. A lack of interest, being too busy, or being forgetful also played a part for some participants not using the LUMOback.

The Perceived Influence of Others on LUMOback Use

Some participants reported that support from their team manager, the workplace champion, or other team members was related to uptake of the LUMOback. However, a majority of participants reported that their LUMOback use was not influenced by others because they did not talk about using the LUMOback or they could not see if other team members were wearing it.

Suggested Improvements in an Activity Tracker

Participants were also asked about their suggested improvements in activity trackers. The key desires for trackers included (1) being comfortable to wear, (2) being discreet, (3) being easy to use, (4) being waterproof, and (5) providing a range of accurate data.

> *Something I could just leave on that I wouldn't need to take on and off all the time, something that is waterproof, something that is maybe not so visible.* [#3, female managerial staff, limited user]

Discussion

Principal Findings

This is one of the first studies to comprehensively evaluate both the usage and acceptability of a wearable activity tracker when delivered as part of a workplace intervention. The real-time feedback was valuable to users, and greater use of the LUMOback appeared to assist with reducing the amount of daily prolonged sitting during the intervention. However, only 70% (43/61) of participants provided with a LUMOback used it. Among those who used the LUMOback, the amount of usage (left to the participant's discretion) was very limited, with a median of 8 days of usage, and a maximum of 48 days of usage out of a possible 91 days. Being too busy, technology issues during setup, and a lack of interest were possible reasons for nonuse. Within users, comfort was a key reason for low usage of the LUMOback. Having lower self-perceived job control and weaker supervisor relationships were significantly associated with more LUMOback usage; while participants who were senior staff and team leaders, and participants with a better health profile also tended to use the LUMOback more than their counterparts.

Workplace tracker interventions have previously seen high levels of tracker uptake (96-99%) [13,27]. However, in line with the findings observed in this study (where uptake was 70%), uptake of trackers when offered in real-world workplace wellness programs can be more modest (25-100%) [9]. In this study, some nonusers reported a lack of interest in their LUMOback tracker. Tracker use is a personal decision, and some workers simply may not want or need a device to track their behavior. As such, offering trackers to everyone is not necessarily going to be a catch-all solution, but just one of many strategies a workplace could offer its workers. Issues with setup of the device may also have impeded its uptake. Although the participants in this study felt confident using their smartphone, work demands and technological difficulties may have meant that in-person set-up support was still needed. This finding, in line with a previous study reporting technology barriers to tracker uptake [28], suggests that investing in additional support at the initiation stage may be warranted. Indeed, in one study where such support was provided (eg, downloading the app and setting up the tracker for the participant), uptake was 100% and usage was high across the study (median use of 95% days across 16 weeks of intervention) [29].

Even among those who proceeded to use the belt-worn LUMOback, usage was low. Workplace studies evaluating trackers clipped on to the hip or shoe [12,13,27], worn on the wrist [14,30], or worn in a range of hip, wrist, and chest locations [31] have reported usage ranging from a mean use of 79% days across 6 weeks of intervention [12] to a mean use of approximately 69% days across 3 months of intervention [27], and 15% to upward of 89% of participants still using their tracker after 2 months [13,14,30,31]. All studies reporting high usage incorporated monetary incentives [13,27] or text message/email prompts [12] to encourage usage. Usage in the studies without these additional features [14,30,31] was typically lower (15-60% of participants still using their tracker after 2 months). These findings suggest that additional strategies may be needed to achieve high tracker usage.

Social support strategies, such as being part of a team, sharing results with others, and participating in workplace competitions, may also be associated with increased uptake and use of activity trackers [9,32]. In this study, senior and team leaders tended to use the LUMOback more than other employees. Participants in these roles likely had a higher responsibility for managing health and safety issues, had more direct contact with the workplace champion (who was at a similar managerial level), and as such may have been more motivated to use their LUMOback tracker. Senior and team leader support facilitated LUMOback uptake for some participants, but they did not discuss ongoing LUMOback use with their teams. A greater extent of role modeling and discussion might be useful to promote tracker usage and could provide a means to alleviate issues with setup and calibration.

Greater users of the tracker tended to be healthier, which is consistent with previous research [15,33]. A novel finding was that participants with high self-perceived job control and supervisor relationships used the LUMOback for significantly fewer days. It may have been that these participants felt more freedom and confidence in their work roles to discontinue use of the tracker. Participants who had musculoskeletal problems at baseline used the LUMOback for fewer days than those without problems. It is possible that minor back pain following the posture advice, which was reported by 3 participants in the adverse outcomes of the trial [18], was a reason for less use in those participants who already had musculoskeletal problems at baseline. The waist-worn belt may have also been particularly troublesome for those with pre-existing back problems.

Comfort of the LUMOback appeared to be a key contributor to the low overall usage identified in this study. Participants reported dislike of the belt placement, and some noted that the LUMOback was difficult to wear with clothing. Similar findings have been reported with the Actigraph GT3X+ when attached via a belt [34] and for trackers worn tightly secured to the body [31]. Alternatives to belts, such as wristbands and clips, may be especially useful for long-term use with office clothing. However, the wrist and hip wear positions suited to wristbands and clips are not ideal for measuring sitting time [20] as they can lead to standing being misidentified as sitting. Other barriers identified by participants in this study and across other tracker research include a lack of perceived data accuracy, forgetting to use or charge the tracker, and difficulty using the tracker [14,30,31]. These findings suggest that, even with additional support, trackers need to be comfortable, accurate, discreet, and easy to use to facilitate ongoing use.

Consistent with other workplace tracker studies [14,31], we saw very limited usage of the LUMOback after 3 months (ie, only 1 out of the 21 participants whom we could follow up). However, the need for continuous self-monitoring of behavior is debatable [35]. Disuse of a tracker after a short period could indicate that participants have learned about their behavior and made changes rather than indicating failure to change [31]. Alternatively, the intermittent usage observed for some participants may reflect that self-monitoring provides a way to reidentify and adjust behavior as needed. Our findings did not suggest that the behavioral impact of the LUMOback is limited to the time frame over which it was worn; usage during the activity monitoring did not appear to predict behavior change any better than usage over the intervention generally, much of which occurred early on. A period of initial tracking may be sufficient to facilitate 3-month behavior changes. However, further research is needed to understand the interplay between tracker use and behavior.

Encouragingly, even within a group who used the LUMOback to a limited degree, larger favorable behavior changes during the intervention were seen in those with high usage compared with those with low or no usage. Greater usage during the intervention was significantly associated with overall time per day spent in prolonged sitting and, although there should be caution in interpreting this result, time spent in nonprolonged sitting during work hours. Unlike the majority of commercially available activity trackers [20], the LUMOback used posture measurement to detect and prompt on sitting time, which may be why this particular tracker was able to facilitate improvements in prolonged sitting time during the intervention.

However, the LUMOback is no longer commercially available. Thus, it is recommended that current and future commercial activity trackers include features to distinguish between sitting and upright postures so that they can appropriately measure and prompt on sitting time.

Notably, use of the LUMOback was not significantly associated with changes in sitting, standing, and stepping time. Greater changes in these behaviors may require greater use of the LUMOback and/or additional behavioral strategies. Environmental modifications to the workplace (eg, sit-stand workstations) and/or changes to the design of work tasks (eg, activity-based working, where workers can change workspace depending on work task) may also be needed to support regular shifts between sitting, standing, and moving.

Strengths and Limitations

A key study strength was that use of the LUMOback was self-directed by the participant and was provided as part of a broader workplace program. Thus, the findings are likely to be indicative of how a tracker may be implemented and used as part of a real-world worksite-driven program. The evaluation over the longer term enabled the examination of how the LUMOback was being used over time (and by whom), whereas the qualitative investigation added depth to the findings by exploring the reasons for lack of uptake and ceased use. It is recommended that future tracker evaluations take this mixed-methods approach.

This study was not powered a priori on the aims of this paper. Confidence intervals around estimates indicated that substantial differences in sociodemographic and health- and job-related predictors of usage and substantial differences in activity changes by usage may have been missed due to inadequate precision. Moreover, the limited differences in activity changes by usage may in part be due to a mismatch between when participants used their LUMOback and when we captured their behavior. Evaluating usage and behavior over a more similar time frame (the assessment period vs the intervention) did not tend to strengthen associations; however, our examination would not detect very short-term effects of usage on behavior (eg, on the same day or at the same time). This may be an important question to address in a sample with sufficient usage. Biases may have had an impact on the results. Due to the small sample size, it was not possible to adjust for all potential confounders. Information was not collected on those who declined participation or who were never approached. LUMOback nonusers were also more likely to drop out of the study than LUMOback users. This attrition bias suggests our results, if anything, are likely to underestimate the associations of usage with behavior change. Although only some participants participated in the telephone interviews, interviews were conducted in both LUMOback users and nonusers, and a range of perspectives was collected. Generalizability is limited to the type of tracker used (belt-worn tracker focused on posture and sitting) and to the context of its use. Notably, it was allocated and provided by others (not chosen or purchased by the user) in the context of a team-based workplace-delivered organizational intervention. Use of a tracker that is self-selected by participants might lead to greater uptake and/or longer usage. Self-selection may especially encourage usage in those with high job control who may be accustomed to having a high level of autonomy in their working life.

Conclusions

A tracker that provides real-time feedback on sitting time was associated with the reduction of prolonged sitting time during a workplace-delivered sitting-reduction intervention. The tracker evaluated was not particularly suitable for ongoing use in office workers, but this study has provided insight into barriers and facilitators for uptake and ongoing tracker use in this population.

Acknowledgments

This work was supported by a Vanguard Grant (Award ID: 100216) from the National Heart Foundation of Australia, a grant from the Office Ergonomics Research Committee, and Lendlease. The National Health and Medical Research Council (NHMRC) of Australia, through a Centre of Research Excellence Grant (#1057608) to GNH, provides salary support to EAHW and a top-up scholarship to CLB. CLB is also supported by an Australian Government Research Training Program Scholarship. GNH is supported by a NHMRC Career Development Fellowship (#108029). The authors would like to thank Lendlease staff for their participation in the study.

Authors' Contributions

CLB participated in the design, led the coordination and data collection of the study, conducted all analyses, and drafted the manuscript. GNH participated in the design of the study and helped draft the manuscript. EAHW and BSF participated in the design of the study, assisted in the analyses, and helped draft the manuscript. All authors contributed to, read, and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Characteristics of participants randomized to the ORG+Tracker intervention overall and by usage group.

[PDF File (Adobe PDF File), 42KB - ijmr_v7i1e5_app1.pdf]

Multimedia Appendix 2

Participant flow diagram.

[PDF File (Adobe PDF File), 83KB - ijmr_v7i1e5_app2.pdf]

Multimedia Appendix 3

Additional participant quotes.

[PDF File (Adobe PDF File), 27KB - ijmr_v7i1e5_app3.pdf]

Multimedia Appendix 4

Acceptability of the LUMOback.

[PDF File (Adobe PDF File), 23KB - ijmr_v7i1e5_app4.pdf]

Multimedia Appendix 5

Use and perceived usefulness of the LUMOback features.

[PDF File (Adobe PDF File), 23KB - ijmr_v7i1e5_app5.pdf]

References

1. Ryan CG, Dall PM, Granat MH, Grant PM. Sitting patterns at work: objective measurement of adherence to current recommendations. Ergonomics 2011 Jun;54(6):531-538. [doi: 10.1080/00140139.2011.570458]
2. Healy GN, Eakin EG, Owen N, Lamontagne AD, Moodie M, Winkler EA, et al. A cluster randomized controlled trial to reduce office workers' sitting time: effect on activity outcomes. Med Sci Sports Exerc 2016 Sep;48(9):1787-1797. [doi: 10.1249/MSS.0000000000000972] [Medline: 27526175]
3. Tremblay MS, Aubert S, Barnes JD, Saunders TJ, Carson V, Latimer-Cheung AE, SBRN Terminology Consensus Project Participants. Sedentary Behavior Research Network (SBRN) - terminology consensus project process and outcome. Int J Behav Nutr Phys Act 2017 Jun 10;14(1):75 [FREE Full text] [doi: 10.1186/s12966-017-0525-8] [Medline: 28599680]
4. Ariëns GA, van Mechelen W, Bongers PM, Bouter LM, van der Wal G. Physical risk factors for neck pain. Scand J Work Environ Health 2000 Feb;26(1):7-19 [FREE Full text] [Medline: 10744172]
5. Healy GN, Matthews CE, Dunstan DW, Winkler EA, Owen N. Sedentary time and cardio-metabolic biomarkers in US adults: NHANES 2003-06. Eur Heart J 2011 Mar;32(5):590-597 [FREE Full text] [doi: 10.1093/eurheartj/ehq451] [Medline: 21224291]
6. Clemes SA, O'Connell SE, Edwardson CL. Office workers' objectively measured sedentary behavior and physical activity during and outside working hours. J Occup Environ Med 2014 Mar;56(3):298-303. [doi: 10.1097/JOM.0000000000000101] [Medline: 24603203]
7. Church TS, Thomas DM, Tudor-Locke C, Katzmarzyk PT, Earnest CP, Rodarte RQ, et al. Trends over 5 decades in U.S. occupation-related physical activity and their associations with obesity. PLoS One 2011;6(5):e19657 [FREE Full text] [doi: 10.1371/journal.pone.0019657] [Medline: 21647427]
8. U.S. Bureau of Labor Statistics. American time use survey URL: http://www.bls.gov/tus/charts/ [WebCite Cache ID 6xDp6Wyto]
9. Health Enhancement Research Organization. Wearables in wellness: employer use of wearable tracking devices in wellness programs URL: http://hero-health.org/wp-content/uploads/2015/06/HERO-Wearables-in-Wellness-Report-FINAL.pdf [WebCite Cache ID 6xDpE5tX1]
10. Tractica. More than 66 million wearable devices will be shipped annually for use in enterprise and industrial environments by 2021 URL: https://www.tractica.com/newsroom/press-releases/more-than-66-million-wearable-devices-will-be-shipped-annually-for-use-in-enterprise-and-industrial-environments-by-2021/ [accessed 2017-02-09] [WebCite Cache ID 6xDpT0UNC]
11. Lyons EJ, Lewis ZH, Mayrsohn BG, Rowland JL. Behavior change techniques implemented in electronic lifestyle activity monitors: a systematic content analysis. J Med Internet Res 2014;16(8):e192 [FREE Full text] [doi: 10.2196/jmir.3469] [Medline: 25131661]
12. Poirier J, Bennett WL, Jerome GJ, Shah NG, Lazo M, Yeh H, et al. Effectiveness of an activity tracker- and internet-based adaptive walking program for adults: a randomized controlled trial. J Med Internet Res 2016;18(2):e34 [FREE Full text] [doi: 10.2196/jmir.5295] [Medline: 26860434]

13. Finkelstein EA, Haaland BA, Bilger M, Sahasranaman A, Sloan RA, Nang EE, et al. Effectiveness of activity trackers with and without incentives to increase physical activity (TRIPPA): a randomised controlled trial. Lancet Diabetes Endocrinol 2016 Dec;4(12):983-995. [doi: 10.1016/S2213-8587(16)30284-4] [Medline: 27717766]
14. Boulard Masson C, Martin D, Colombino T, Grasso A. “The device is not well designed for me” on the use of activity trackers in the workplace? Cham, Switzerland: Springer; 2016 Presented at: COOP 2016 the 12th International Conference on the Design of Cooperative Systems; May 23-27, 2016; Trento, Italy p. 39-55.
15. Ledger D, McCaffrey D. Endeavour. 2014. Inside wearables: how the science of human behavior change offers the secret to long-term engagement URL: https://blog.endeavour.partners/inside-wearable-how-the-science-of-human-behavior-change-offers-the-secret-to-long-term-engagement-a15b3c7d4cf3?gi=db661960715 [accessed 2018-02-19] [WebCite Cache ID 6xMd2fmTS]
16. Martin A, Fitzsimons C, Jepson R, Saunders DH, van der Ploeg HP, Teixeira PJ, EuroFIT consortium. Interventions with potential to reduce sedentary time in adults: systematic review and meta-analysis. Br J Sports Med 2015 Aug;49(16):1056-1063. [doi: 10.1136/bjsports-2014-094524] [Medline: 25907181]
17. Brakenridge CL, Fjeldsoe BS, Young DC, Winkler EA, Dunstan DW, Straker LM, et al. Organizational-level strategies with or without an activity tracker to reduce office workers' sitting time: rationale and study design of a pilot cluster-randomized trial. JMIR Res Protoc 2016 May 25;5(2):e73 [FREE Full text] [doi: 10.2196/resprot.5438] [Medline: 27226457]
18. Brakenridge CL, Fjeldsoe BS, Young DC, Winkler EA, Dunstan DW, Straker LM, et al. Evaluating the effectiveness of organisational-level strategies with or without an activity tracker to reduce office workers' sitting time: a cluster-randomised trial. Int J Behav Nutr Phys Act 2016 Nov 04;13(1):115 [FREE Full text] [doi: 10.1186/s12966-016-0441-3] [Medline: 27814738]
19. Sanders JP, Loveday A, Pearson N, Edwardson C, Yates T, Biddle SJ, et al. Devices for self-monitoring sedentary time or physical activity: a scoping review. J Med Internet Res 2016 May 04;18(5):e90 [FREE Full text] [doi: 10.2196/jmir.5373] [Medline: 27145905]
20. Rosenberger ME, Buman MP, Haskell WL, McConnell MV, Carstensen LL. 24 hours of sleep, sedentary behavior, and physical activity with nine wearable devices. Med Sci Sports Exerc 2016 Mar;48(3):457-465. [doi: 10.1249/MSS.0000000000000778] [Medline: 26484953]
21. Kooiman TJ, Dontje ML, Sprenger SR, Krijnen WP, van der Schans CP, de Groot M. Reliability and validity of ten consumer activity trackers. BMC Sports Sci Med Rehabil 2015;7:24 [FREE Full text] [doi: 10.1186/s13102-015-0018-5] [Medline: 26464801]
22. Dunstan DW, Wiesner G, Eakin EG, Neuhaus M, Owen N, LaMontagne AD, et al. Reducing office workers' sitting time: rationale and study design for the Stand Up Victoria cluster randomized trial. BMC Public Health 2013;13:1057 [FREE Full text] [doi: 10.1186/1471-2458-13-1057] [Medline: 24209423]
23. Kuorinka I, Jonsson B, Kilbom A, Vinterberg H, Biering-Sørensen F, Andersson G, et al. Standardised nordic questionnaires for the analysis of musculoskeletal symptoms. Appl Ergon 1987 Sep;18(3):233-237. [Medline: 15676628]
24. Ware J, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996 Mar;34(3):220-233. [Medline: 8628042]
25. Shikiar R, Halpern MT, Rentz AM, Khan ZM. Development of the Health and Work Questionnaire (HWQ): an instrument for assessing workplace productivity in relation to worker health. Work 2004;22(3):219-229. [Medline: 15156087]
26. Sundstrom E, Town JP, Rice RW, Osborn DP, Brill M. Office noise, satisfaction, and performance. Environ Behav 1994 Mar;26(2):195-222. [doi: 10.1177/001391659402600204]
27. Thorndike AN, Mills S, Sonnenberg L, Palakshappa D, Gao T, Pau CT, et al. Activity monitor intervention to promote physical activity of physicians-in-training: randomized controlled trial. PLoS One 2014;9(6):e100251 [FREE Full text] [doi: 10.1371/journal.pone.0100251] [Medline: 24950218]
28. Gilson ND, Pavey TG, Vandelanotte C, Duncan MJ, Gomersall SR, Trost SG, et al. Chronic disease risks and use of a smartphone application during a physical activity and dietary intervention in Australian truck drivers. Aust N Z J Public Health 2016 Feb;40(1):91-93. [doi: 10.1111/1753-6405.12501] [Medline: 26713400]
29. Cadmus-Bertram L, Marcus BH, Patterson RE, Parker BA, Morey BL. Use of the fitbit to measure adherence to a physical activity intervention among overweight or obese, postmenopausal women: self-monitoring trajectory during 16 weeks. JMIR Mhealth Uhealth 2015;3(4):e96 [FREE Full text] [doi: 10.2196/mhealth.4229] [Medline: 26586418]
30. Schrager JD, Shayne P, Wolf S, Das S, Patzer RE, White M, et al. Assessing the influence of a fitbit physical activity monitor on the exercise practices of emergency medicine residents: a pilot study. JMIR Mhealth Uhealth 2017 Jan 31;5(1):e2 [FREE Full text] [doi: 10.2196/mhealth.6239] [Medline: 28143805]
31. Lazar A, Koehler C, Tanenbaum J, Nguyen DH. Why we use and abandon smart devices. ACM. 2015 Presented at: UbiComp '15 the 2015 ACM International Joint Conference on Pervasive Ubiquitous Computing; September 7-11, 2015; Osaka, Japan.
32. Harrison D, Marshall P, Bianchi-Berthouze N, Bird J. Activity tracking: barriers, workarounds and customisation. ACM. 2015 Presented at: UbiComp '15 the 2015 ACM International Joint Conference on Pervasive Ubiquitous Computing; September 7-11, 2015; Osaka, Japan.

33. Yingling LR, Mitchell V, Ayers CR, Peters-Lawrence M, Wallen GR, Brooks AT, et al. Adherence with physical activity monitoring wearable devices in a community-based population: observations from the Washington, D.C., Cardiovascular Health and Needs Assessment. Transl Behav Med 2017 Dec;7(4):719-730 [FREE Full text] [doi: 10.1007/s13142-016-0454-0] [Medline: 28097627]
34. Huberty J, Ehlers DK, Kurka J, Ainsworth B, Buman M. Feasibility of three wearable sensors for 24 hour monitoring in middle-aged women. BMC Womens Health 2015 Jul 30;15:55 [FREE Full text] [doi: 10.1186/s12905-015-0212-3] [Medline: 26223521]
35. Fjeldsoe B, Neuhaus M, Winkler E, Eakin E. Systematic review of maintenance of behavior change following physical activity and dietary interventions. Health Psychol 2011 Jan;30(1):99-109. [doi: 10.1037/a0021974] [Medline: 21299298]

Abbreviations

MDI: minimum difference of interest

Edited by TR Soron; submitted 20.09.17; peer-reviewed by M Whelan, D Ummels, J Di; comments to author 11.10.17; revised version received 06.12.17; accepted 21.12.17; published 02.03.18.

Please cite as:
Brakenridge CL, Healy GN, Winkler EAH, Fjeldsoe BS
Usage, Acceptability, and Effectiveness of an Activity Tracker in a Randomized Trial of a Workplace Sitting Intervention: Mixed-Methods Evaluation
Interact J Med Res 2018;7(1):e5
URL: http://www.i-jmr.org/2018/1/e5/
doi:10.2196/ijmr.9001
PMID:29500158

Original Paper

Representation of Tinnitus in the US Newspaper Media and in Facebook Pages: Cross-Sectional Analysis of Secondary Data

Vinaya Manchaiah[1], PhD; Pierre Ratinaud[2], PhD; Gerhard Andersson[3], PhD

[1]Department of Speech and Hearing Sciences, Lamar University, Beaumont, TX, United States

[2]LERASS Laboratory, University of Toulouse, Toulouse, France

[3]Department of Behavioral Sciences and Learning, Linköping University, Linköping, Sweden

Corresponding Author:
Vinaya Manchaiah, PhD
Department of Speech and Hearing Sciences
Lamar University
PO Box 10076
Beaumont, TX, 77710
United States
Phone: 1 4098808927
Fax: 1 4098808927
Email: vinaya.manchaiah@lamar.edu

Abstract

Background: When people with health conditions begin to manage their health issues, one important issue that emerges is the question as to what exactly do they do with the information that they have obtained through various sources (eg, news media, social media, health professionals, friends, and family). The information they gather helps form their opinions and, to some degree, influences their attitudes toward managing their condition.

Objective: This study aimed to understand how tinnitus is represented in the US newspaper media and in Facebook pages (ie, social media) using text pattern analysis.

Methods: This was a cross-sectional study based upon secondary analyses of publicly available data. The 2 datasets (ie, text corpuses) analyzed in this study were generated from US newspaper media during 1980-2017 (downloaded from the database US Major Dailies by ProQuest) and Facebook pages during 2010-2016. The text corpuses were analyzed using the Iramuteq software using cluster analysis and chi-square tests.

Results: The newspaper dataset had 432 articles. The cluster analysis resulted in 5 clusters, which were named as follows: (1) brain stimulation (26.2%), (2) symptoms (13.5%), (3) coping (19.8%), (4) social support (24.2%), and (5) treatment innovation (16.4%). A time series analysis of clusters indicated a change in the pattern of information presented in newspaper media during 1980-2017 (eg, more emphasis on cluster 5, focusing on treatment inventions). The Facebook dataset had 1569 texts. The cluster analysis resulted in 7 clusters, which were named as: (1) diagnosis (21.9%), (2) cause (4.1%), (3) research and development (13.6%), (4) social support (18.8%), (5) challenges (11.1%), (6) symptoms (21.4%), and (7) coping (9.2%). A time series analysis of clusters indicated no change in information presented in Facebook pages on tinnitus during 2011-2016.

Conclusions: The study highlights the specific aspects about tinnitus that the US newspaper media and Facebook pages focus on, as well as how these aspects change over time. These findings can help health care providers better understand the presuppositions that tinnitus patients may have. More importantly, the findings can help public health experts and health communication experts in tailoring health information about tinnitus to promote self-management, as well as assisting in appropriate choices of treatment for those living with tinnitus.

(Interact J Med Res 2018;7(1):e9) doi:10.2196/ijmr.9065

KEYWORDS

tinnitus; chronic condition; health communication; health information; media; news media; social media

Introduction

Background

Tinnitus can be defined as the conscious perception of an auditory sensation in the absence of a corresponding external stimulus [1]. It is a common medical symptom, which can be debilitating to some individuals. Risk factors for tinnitus include hearing loss, use of ototoxic drugs, exposure to loud noise, and head injury. Various types of treatments are being offered to help tinnitus sufferers, although no effective curative drug treatment exists. Some main treatment and management approaches for tinnitus include sound therapy (ie, external noise to counteract the perception of tinnitus), hearing aids, and counseling [1,2]. However, cognitive behavioral therapy has the largest evidence base for the management of the distress associated with tinnitus [2].

Self-Management of Chronic Conditions

Medical therapy alone is not sufficient to address the distress caused by various chronic symptoms and conditions [3], which can also be true for tinnitus. For this reason, there is a move toward developing self-management initiatives in which individuals with a health condition take responsibility and ownership in managing their health conditions and symptoms [4]. Self-management can facilitate acceptance, coping, learning to live well with chronic health conditions despite the adverse consequences of such conditions and symptoms, and more importantly can result in improved health outcomes [5]. However, when people with health conditions start to manage these health issues, a significant factor that determines how effectively they manage their issues involves the type and quality of information that they obtain through various sources (eg, news media, social media, health professionals, friends, and family). The information they gather helps form their opinions, and will also to some degree influence their attitudes toward managing their condition. Literature from health psychology, particularly the Self-Regulatory Model of Illness, provides a framework for understanding how individual symptoms and emotions during a health threat or diagnosis influence perception of illness and hence guide subsequent behavior [6,7]. An individual may be confronted with the problem of a potential illness through 2 channels, which include: (1) symptom perception and (2) social messages. The individual experiences and symptom perception can be influenced by the social messages they receive (ie, input from friends, family, and media). For this reason, health communication is an important element of facilitating self-management in chronic disease management from the public health viewpoint [8].

Influence of Media on Health Behavior

The ubiquitous nature of mass media, in particular news media, makes it one of the most powerful sources of influence on specific issues (eg, climate change, political party), including health issues [9]. News media plays an important role in determining the content of our thoughts and opinions [10]. For example, the more coverage a specific topic receives on a particular news media, more likely are the people to take it into serious consideration. On the other hand, a topic that does not get much attention in the news media, although important, may not receive much public attention. Although society traditionally has relied heavily on the news media, in recent years, social media has radically transformed the way in which people communicate and gather information around the world [11]. Social media has changed the news business on a fundamental level [12]. For example, in social media content, providers are the first ones to report any given particular news, whereas editors and journalists are involved in content development and decision-making of what appears in the news media. People use social media to share and exchange perspectives and opinions about social, cultural, economic, and religious aspects; furthermore, there is a growing trend to also exchange information about health issues and concerns [13]. In keeping with this recent development, there is a critical need to better understand the influence of social media on health behavior and outcomes [14].

Although the media can help place specific health issues on a local and national agenda, there may be a discrepancy in what issues are covered and in their social and economic impact on society. For instance, we can assume that the issues more frequently raised in the media are those which are more likely to draw more attention and readership. Moreover, Hartz et al reports that there is a growing tension between media reporters, scientists, and health professionals [15]. This may be because of discrepancies in understanding specific health issues and also because of the importance placed on specific health issues by these key stakeholders [15].

There is some literature on how the deaf population is represented in the media [16,17]. For example, researchers have explored how and why hearing loss continues to be stigmatized through a study of media messages about hearing loss [16]. Another study focused on the positive and negative media framing toward political movement Deaf President Now [17]. Moreover, some recent literature also focuses on newspapers' representation of workers with hearing loss [18] and the use of social media in the hearing aid community [19]. We recognize that both mass media and social media can have much influence on people's knowledge and in the shaping of individual attitudes [14]. Although this is of importance, we were unable to find any studies focusing specifically on what information is presented about tinnitus in the news media, and also in social media.

This study aimed to understand how tinnitus is represented in the US newspaper media and in Facebook pages (ie, social media) using text pattern analysis.

Methods

Study Design

This study used a cross-sectional design based upon secondary analyses of publicly available data. The 2 datasets analyzed in this study were generated from US newspaper media and Facebook pages. The study did not require ethical approval as the data were gathered from publicly available sources. Only the publicly available Facebook pages (not personal pages) were included in the data extraction, thus maintaining the anonymity of the responses, and no personally identifiable information was included. Moreover, no individual dataset was discussed in the paper, again maintaining the anonymity of the data. Considering the minimal or no potential risk to individual participants, no ethical approval was required [20].

Data Extraction

To develop the newspaper media text corpus (ie, large and structured set of texts), we first explored the databases with the newspaper collection available at Lamar University, United States. Major Dailies by ProQuest was the database that had the largest newspaper collection. We then searched for articles related to tinnitus in this database between the years 1980 and 2017 and downloaded the results as text corpus. A python script was written to convert the text corpus to a format that was needed for data analysis and to preserve the metadata (ie, newspaper name, year of publication).

A different python script was written to extract posts with Facebook pages dealing with tinnitus (during 2010-2016). In total, 20 Facebook pages with more than 100 likes were identified, and the postings were downloaded as a text corpus. It is important to note that the data extraction was limited to what data were available publicly (ie, newspaper data during 1980-2017 and Facebook posts during 2010-2016).

Data Analysis

The text corpuses were analyzed using the Iramuteq software [21], which is an open source software. This software can perform various types of analyses on text data.

The text corpus is composed of multiple newspaper articles. The software treats each of these articles as text (ie, it's the first unit). The first step of the analysis is to segment each article (ie, text) into smaller units called text segments (ie, each text is split into multiple text segments based on criteria of size and punctuation). The split of the text into segments decreases the granularity of the units and thus makes it possible to increase the precision of certain analysis, in particular the search of themes within the text corpus. The goal is to create segments of consistent size while trying to maintain the natural segmentation of texts marked by punctuation. The segmentation process is based on a cutoff criterion that weighs the segment size by punctuation. The procedure does not fully respect the parameterized segment size if a strong punctuation (eg, a period, a question mark, and an exclamation mark) is present within a 15% margin of the planned cutoff. In the next step, the text corpus is lemmatized (ie, words are sorted by grouping inflected or variant forms of the same word) to their simplest forms, which are called lemmas (ie, group of words in a single form). Moreover, the software makes distinctions between "full words" (eg, verbs, noun, adjectives, and adverbs) and "tool words" (eg, pronoun, determent, and useful verbs such as—to be and to have). This distinction is done so that only full words are included in the main analysis. These steps are necessary to convert the large corpus into a more manageable dataset that is essential for further analysis. To specifically analyze the text that is closely related to "tinnitus," the text segments related to this object were extracted and new corpuses were formed. It is important to note that the expression "tinnitus" was inside each of the text segments extracted. The subcorpus with more directed text segments was used for all further analyses.

This was followed by cluster analysis made with the Reinert method used for textual data analysis [22-24]. The Reinert method has been used frequently in French media studies. This method facilitates the investigation of links between topics and the metadata associated with the text (eg, dates to study chronology). A recent comparison between the Reinert method and Latent Dirichlet Allocation showed that topics from the Iramuteq software that uses the Reinert method (ie, a divisive hierarchical clustering) are more accurate [25]. Although the traditional qualitative methods (eg, content analysis, thematic analysis) can provide in-depth understanding of the data, the automated text pattern analysis using the cluster analysis can provide a broader understanding of the data [22]. Automated analysis has the advantage of analyzing large amounts of data (ie, big data) as used in this study, whereas using the traditional qualitative methods for analyzing large quantities of data would not have been feasible.

In this analysis, the program initially builds a binary matrix with text segments in rows and full words in columns, and it then performs hierarchical divisive clustering based on a series of bipartitions made with correspondent analysis. At each step of the classification, the larger remaining cluster is cut into 2 parts by computing the information extracted while cutting after each line of the matrix along the first factor of the correspondent analysis. The remaining cut is the one that maximizes the information extracted. In a second step, each line is reversed from one cluster to the other. If this reversion increases the information extracted, it is kept. This step loops until no inversion increases the information extracted for the whole table. This cluster analysis groups the text segments based on co-occurrence of lemmas. Each of the clusters aim to be homogeneous (regrouping text segments with the common pattern of lemmas). Moreover, clusters have to be as heterogeneous as possible between them (pattern of lemmas between groups should be as different as possible). The results are presented in a dendrogram that characterizes the clustering. For each cluster, the program computes profiles of lemmas, which are overrepresented (ie, significantly in a higher proportion within the cluster when compared with the rest of the text corpus based on chi-square analysis). Finally, the same text corpus was subjected to a time series analysis using the metadata. For example, in this corpus, we analyzed how the patterns of clusters change over time (see Figures 1-6). Figures 2 and 5 present proportions of cluster for each year. In these figures, the width of the bar is proportional to the number of text segments each year (ie, the higher the width, the higher the number of text segments), and the height of the clusters (indicated in different colors) represents the frequency of text segments within clusters. Figures 3 and 6 present which clusters are significantly overrepresented (ie, based on chi-square analysis) in each year during 1990-2016. In this width of the bar, it is proportional to the number of text segments each year, and the height of the bar represents the size of clusters. For example, Figure 3 shows that cluster 5 is overrepresented during the years 2013-2017.

Figure 1. Dendrogram (ie, classification of clusters), size of clusters as percentage of the text segments, and overrepresented words in each cluster in tinnitus newspaper corpus. (Note: The words are ordered by chi-square value with words at the bottom having a lower value).

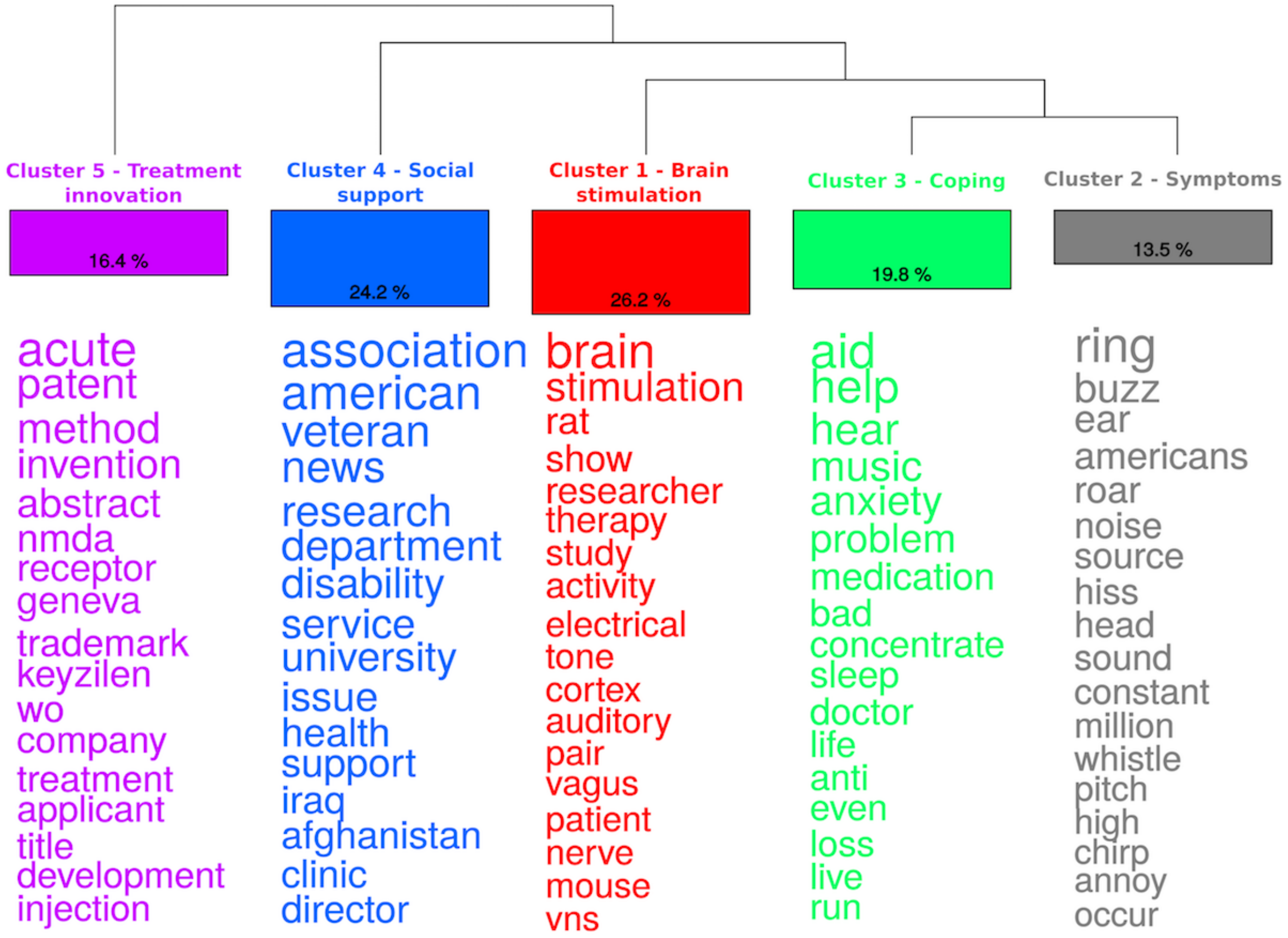


Figure 2. Chronological bar showing proportion of each cluster for each year in tinnitus US newspaper media corpus (Note: Width of the bar is proportional to the number of text segments each year, and the height of the clusters represents the frequency of text segments within clusters).

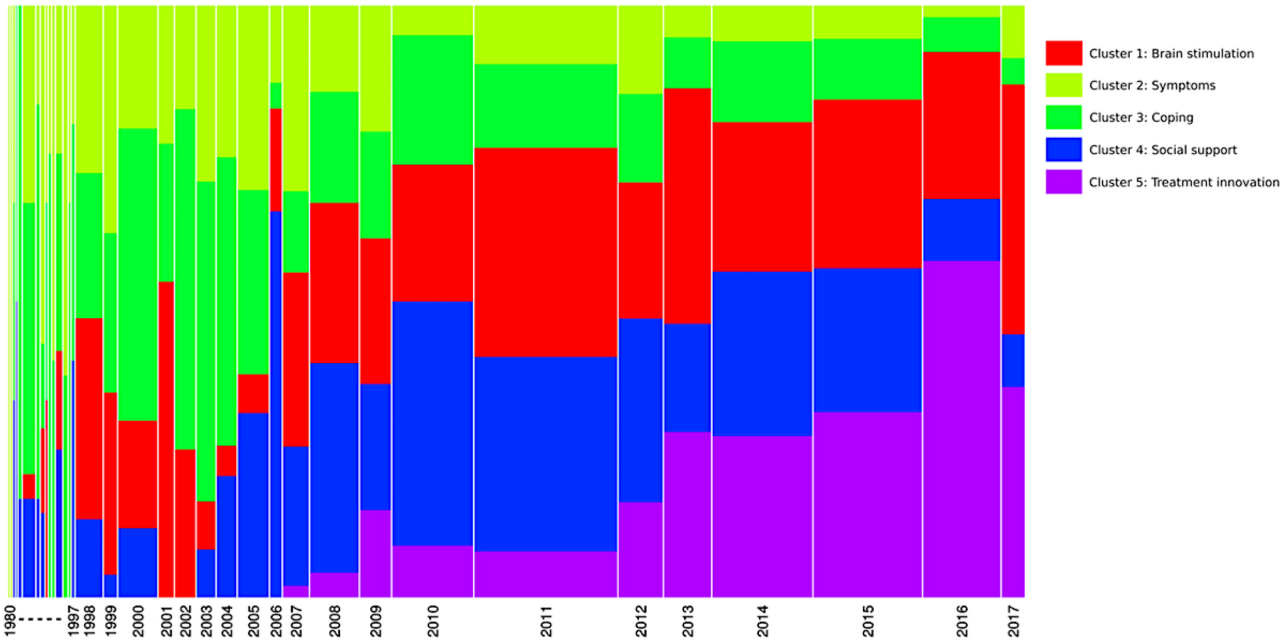


Figure 3. Chronological bar based on chi-square analysis showing proportion of each cluster for each year in tinnitus US newspaper media corpus (Note: Width of the bar is proportional to the number of text segments each year, and the height of the bar represents the size of clusters).

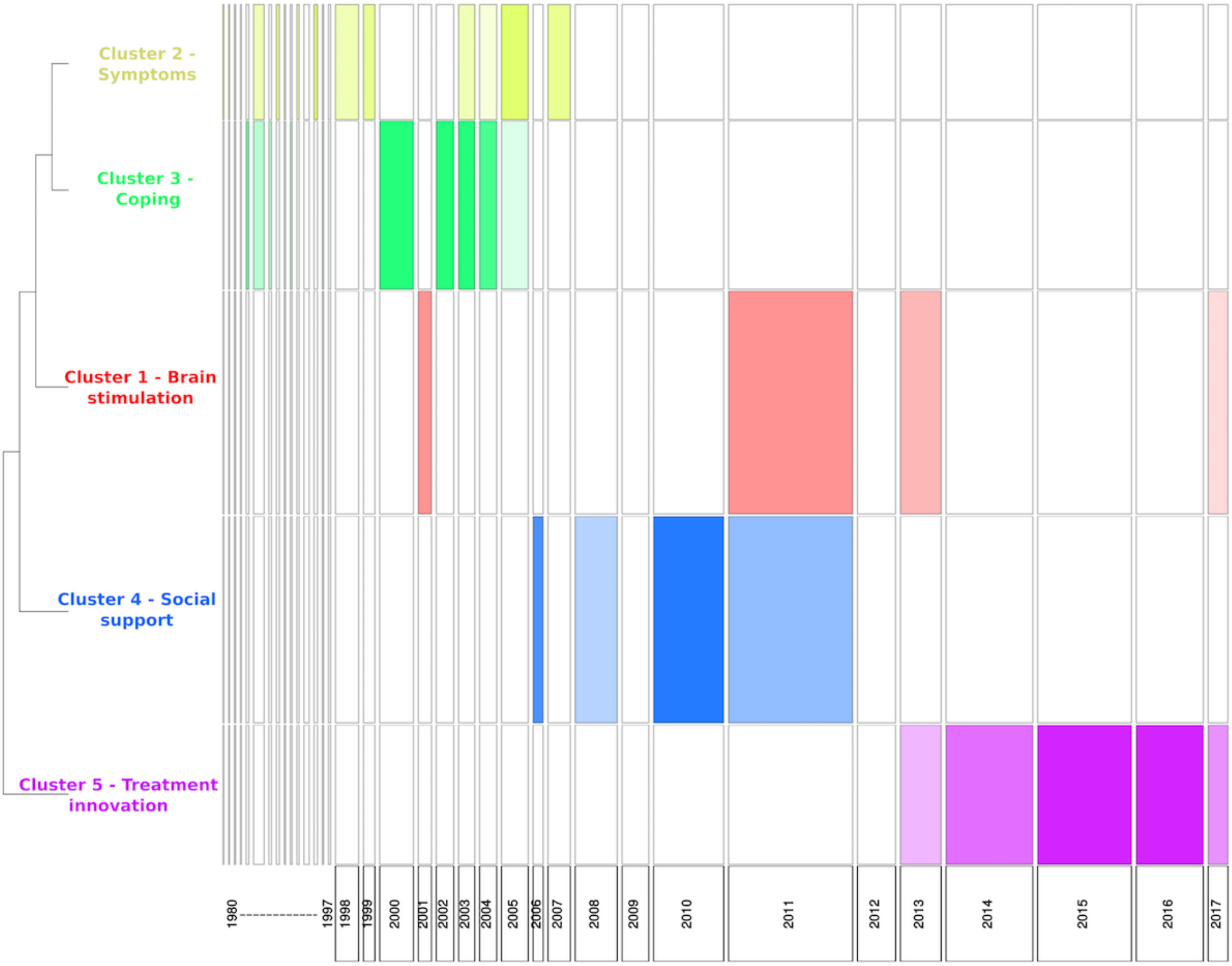


Figure 4. Dendrogram (ie, classification of clusters), size of clusters as percentage of the text segments, and overrepresented words in each cluster in tinnitus on Facebook pages corpus. (Note: The words are ordered by chi-square value with words at the bottom having a lower value).

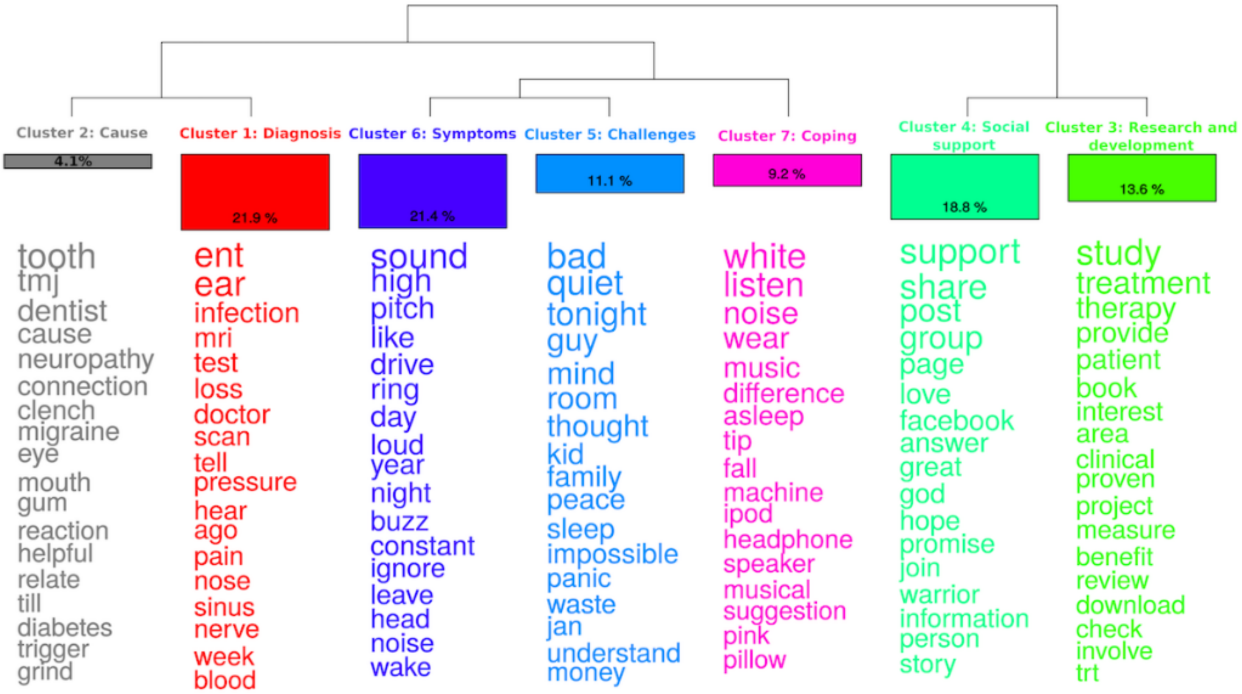


Figure 5. Chronological bar showing proportion of each cluster for each year on tinnitus Facebook pages corpus (Note: Width of the bar is proportional to the number of text segments each year, and the height of the clusters represents the frequency of text segments within clusters).

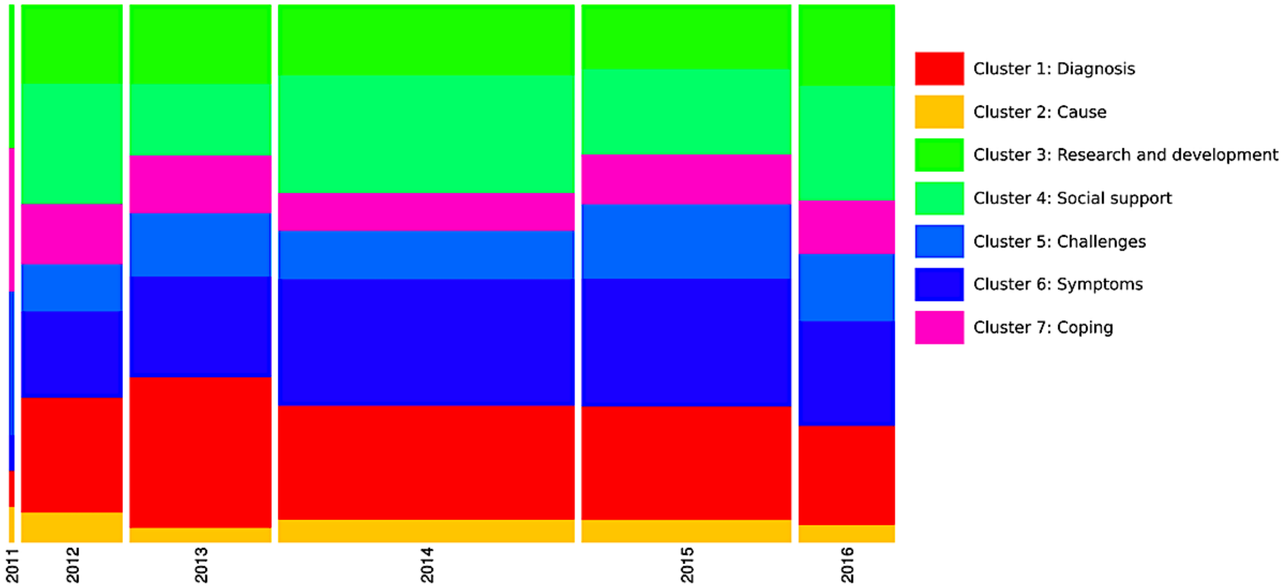
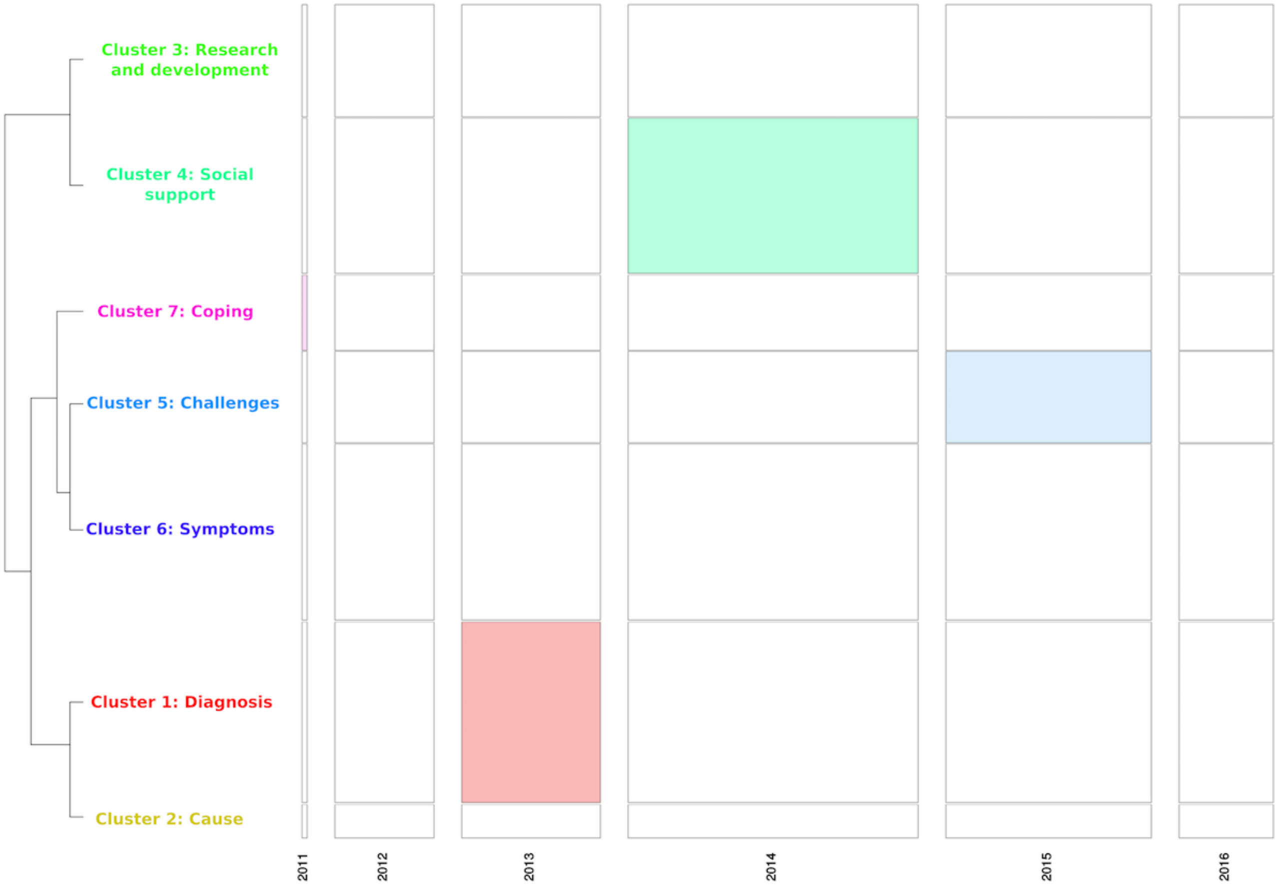


Figure 6. Chronological bar based on chi-square analysis showing the proportion of each cluster for each year on tinnitus Facebook pages corpus (Note: Width of the bar is proportional to the number of text segments each year, and the height of the bar represents the size of clusters).



Results

Newspaper Data

Context Analysis

The initial analysis text corpus had 433 texts (ie, each article from a newspaper is considered a text), 9176 text segments (ie, each text is split into multiple text segments based on criteria of size and punctuation), and 309,524 occurrences or tokens (ie, number of words). After extraction of a text segment related to "tinnitus," the text corps was reduced to 432 texts, 2173 text segments, and 79,684 occurrences or tokens.

Table 1 provides information about the newspaper source and the frequency of texts from each source. Most of the texts are from the US Fed News Services (30.79%), followed by The Washington Post (16.67%), Targeted News Service (13.19%), and the Chicago Tribune (10.19%). Table 2 provides information about the frequency of texts related to tinnitus over time. The frequency of texts published about tinnitus increases over time. The reason for this can be due to an increase in attention and publication of articles on tinnitus and/or the number of newspapers publishing articles about tinnitus. However, it is difficult to distinguish the relative contribution regarding these 2 reasons due to the increase in the frequency of texts related to tinnitus over time.

Content Analysis

Figure 1 provides results of cluster analysis on the newspaper data. The cluster analysis resulted in 5 main clusters that were named based on their characteristics. Cluster 1 was the largest cluster, which included 26.2% of texts, and mainly focused on aspects related to auditory and electrical stimulation in the brain as a cure for tinnitus, which was named as a *brain stimulation* cluster. Cluster 2 was the smallest cluster that included 13.5% of texts, mainly focusing on the symptoms of tinnitus such as buzzing, ringing, hiss, and was named as a *symptoms* cluster. Cluster 3 included 19.8% of texts and included aspects related to psychological problems (eg, anxiety), relaxation, and coping, and hence was named as a *coping* cluster. Cluster 4 was the second largest cluster and included 24.2% of the texts. This cluster had elements related to health and social support from government and charity organizations, and hence was named as a *social support* cluster. Cluster 5 included 16.4% of texts, and was focused on new and innovative treatment methods, and hence was named as a *treatment innovation* cluster. Table 3 provides examples of text segments that typically represent each of these 5 clusters.

Table 1. Frequency and percentage of articles containing at least one text-segment related to tinnitus in the subcorpus from different newspapers.

Newspaper	Frequency of articles, n (%)
Chicago Tribune	46 (11)
Farm Weekly	3 (0.7)
Investor's Business daily	8 (1.9)
Journal of Commerce	1 (0.2)
Journal Record	2 (0.5)
Los Angeles Times	25 (5.8)
Marine Corps Times	1 (0.2)
Miami Daily Business Review	1 (0.2)
Missouri Lawyers Media	1 (0.2)
NASDAQ OMXs News Release Distribution Channel	30 (7.0)
New York Times	27 (6.3)
Roll Call	2 (0.4)
Targeted News Service	57 (13.1)
The Village Voice	2 (0.5)
The Washington Post	72 (16.7)
The Weekly Times	1 (0.2)
US Fed News Services	133 (30.8)
Wall Street Journal	20 (4.6)
Total	432 (100.0)

Table 2. Frequency and percentage of articles containing at least one text-segment related to tinnitus in the subcorpus based on timescales from different newspapers.

Year	Frequency of articles, n (%)
1980	1 (0.2)
1981	1 (0.2)
1985	2 (0.5)
1986	2 (0.5)
1987	3 (0.7)
1988	3 (0.7)
1989	3 (0.7)
1990	4 (0.9)
1991	1 (0.2)
1992	2 (0.5)
1993	3 (0.7)
1994	4 (0.9)
1995	3 (0.7)
1996	1 (0.2)
1997	1 (0.2)
1998	10 (2.3)
1999	7 (1.6)
2000	7 (1.6)
2001	3 (0.7)
2002	6 (1.4)
2003	7 (1.6)
2004	14 (3.2)
2005	43 (10.0)
2006	6 (1.4)
2007	21 (4.9)
2008	18 (4.2)
2009	14 (3.2)
2010	24 (5.6)
2011	46 (10.7)
2012	28 (6.5)
2013	20 (4.6)
2014	42 (9.7)
2015	41 (9.5)
2016	33 (7.7)
2017	8 (1.9)
Total	432 (100.0)

Table 3. Example of a text segment for each cluster in newspaper media text corpus.

Cluster	Example of a text segment
Cluster 1: *Brain stimulation*	Past research has *shown* that the *severity* of chronic *pain* and tinnitus is tied to the degree of *plasticity* in the *brain's cortex* as a *previous study showed* that *repeatedly pairing sensory stimuli* with *electrical stimulation* of a *brain structure* called *nucleus* basalis *generates powerful* and *long-lasting* changes in *cortical* organization.
Cluster 2: *Symptoms*	*Commonly called ringing* in the *ears* but for *millions* of *people* who *suffer* tinnitus the *sound* that plagues them actually may be a *buzz hiss chirp* or *high-pitched whistle.*
Cluster 3: *Coping*	Playing *soft background music* or other *white noise* can be quite effective in *masking* the *noise* and short term use of *antidepressant medications* sometimes *helps* and often a *hearing aid helps* suppress tinnitus.
Cluster 4: *Social support*	A *senior member* of the *Senate Armed Services Committee* recently *received* the *senate* legislative *champion award* by the *American Tinnitus Association* for his dedicated *efforts* to include tinnitus *research* and *funding* to the *Annual National Defense* Authorization Act.
Cluster 5: *Treatment innovation*	As of *today* neither a *universal standard* of *care* for *acute* inner *ear* tinnitus nor a truly *proven* and *effective treatment method* is available about *Keyzilen TM* (AM-101). *Keyzilen TM* is a *small molecule N-methyl-D-aspartate receptor antagonist formulated* in a *biocompatible gel for intratympanic injection*.

Table 4. Example of a text segment for each cluster in Facebook pages text corpus.

Cluster	Example of a text segment
Cluster 1: *Diagnosis*	I went to an *ENT doctor 2 weeks ago*. He checked my *eustachian tubes* because my *ears* were *feeling* a bit *full* that's all, no *ringing* no *pain***.** He placed a probe up my *nose* but first *sprayed 2* pumps of lidocaine anesthetic.
Cluster 2: *Causes*	I have temporo mandibular *joint (TMJ),* which has left me with very *limited jaw movement,* only *open mouth* to less than an inch between *teeth. I have* been referred to a *hospital,* but the waiting time is up to 2 years. Tinnitus can be caused by *TMJ*.
Cluster 3: *Research and development*	*Go Hearing*: We *provide tinnitus* and hearing *therapy* and *treatment* for Bay *area* and *Canada, providing* tinnitus lab pax 100 and tinnitus treater sound *conditioning programs,* which does not cost *thousands*. Affordable *effective programs* and *treatment, check* us out at www.go-hearing.com.
Cluster 4: *Social support*	Greetings, please *share* our *group page* in your *page* it is sometimes not about *money,* we want to *share* and *spread love* to all those who are *suffer* from tinnitus. Please do *support* us, we *need* your help.
Cluster 5: *Challenges*	It was *torture* and a *lot* of *bad thoughts* were running through my *mind*. I used *sleep* machines. We think it was from coming off a *drug* seroquel since *Jan* and Feb this year. I have the ringing 3-4 *times* a week.
Cluster 6: *Symptoms*	*Complete* cure to tinnitus: I had tinnitus in both *ears* for *15 years* with a *high-pitched* two-*tone sound.* The *noises* are *constant* and I have *learned* to *ignore* the *ringing.* Later another *sound* was *added***.**
Cluster 7: *Coping*	Can a thyroid affect your tinnitus and my thyroid is coming out soon. I am hoping the hissing in my ear will subside a bit. I had to *sleep last night listening* to *white noise* for the first time in years. Usually I *listen* to the radio *playing music.*

Analysis of Trends Over Time

Figure 2 provides a chronological bar, and Figure 3 a chronological bar with chi-square for tinnitus newspaper data. These 2 figures provide information on how the information presented in newspaper media changes over time, and helps us understand media trends. For example, in Figure 2, it is evident that cluster 5 (ie, treatment innovation) was not discussed in newspapers before the year 2006, but the number of texts increased substantially from 2006 to 2017. The chi-square analysis (see Figure 3) shows that some clusters are more likely to occur on some timescale. For example, cluster 5 (ie, treatment innovation) was statistically significant and more frequent during the years 2013-2017, whereas the cluster 3 (ie, coping) was statistically significant and more frequent during the years 2000-2005. Overall, the time series analysis of clusters indicated changes in the pattern of information presented in newspaper media during the years 1980-2017.

Facebook Pages Data

Context Analysis

Facebook pages text corpus had 1569 texts, 2747 text segments, and 78,218 occurrences or tokens. We have removed the URL and also texts that contained only the URL before analysis.

Content Analysis

Figure 4 provides results of cluster analysis on the Facebook pages data, which identify 7 main clusters. Cluster 1 (21.9%) was the largest and had words related to ear disorders and its diagnosis, and hence was named as *diagnosis* cluster. Cluster 2 (4.1%) was the smallest and was focused on causes and trigger factors of tinnitus (eg, tooth problems, neuropathy) and was named as the *cause* cluster. Cluster 3 included 13.6% texts focusing on studies that focused on tinnitus and was named as the *research and development* cluster. Cluster 4 (18.8%) had elements related to social support from various sources and was named as the *social support* cluster. Cluster 5 was named as challenges and had 11.1% of texts. Cluster 6 (21.4%) focused on symptoms of tinnitus such as buzzing, ringing, and hiss and was named as the *symptoms* cluster. Cluster 7 (9.2%) was focused on relaxation and coping, and hence was named as the *coping* cluster. Table 4 provides examples of text segments that typically represent each of these 7 clusters.

Analysis of Trends Over Time

Figure 5 provides a chronological bar and Figure 6 a chronological bar with chi-square for tinnitus Facebook pages data. As indicated in Figure 5, there is not much variation in what clusters are discussed in each year over time. The chi-square analysis (see Figure 6) shows that cluster 1 (ie, diagnosis), cluster 4 (ie, social support), and cluster 5 (ie, challenges) occurred more significantly during the years 2013, 2014, and 2015, respectively.

Discussion

Principal Findings

The media has played a prominent role in the dissemination of health information [9]. This study explores the way tinnitus is represented in the media, specifically in the US newspaper media and in Facebook pages using text pattern analysis. Overall, the study results suggest that newspaper media and Facebook pages offering information on tinnitus focus on various elements including symptoms, coping, diagnosis, social support, research and development, and treatment innovations.

The analysis of text corpus extracted from the US newspaper media suggested that the information disseminated via newspaper media focuses mainly on 2 elements: (1) new developments in treatments (ie, brain stimulation, treatment innovations) and (2) disease information (ie, symptoms, coping, and social support). It is interesting to note that there is a fairly equal spread of information among all of these elements. Moreover, the analysis trends over time regarding information indicated a change in patterns of information presented in the newspaper media during 1980-2017. For example, there is an increasing emphasis of information focusing on treatment inventions (cluster 5 in Figure 2), whereas the information on coping appears to be steady over the years (cluster 3 in Figure 2). We were unable to find previous studies on newspaper media on tinnitus to compare this study’s results. However, a recent study explored the Canadian newspaper representation of hearing-impaired workers [18]. The study results reveal that the media focuses on a wide range of issues (eg, hearing-impaired people taking action to find employment, and being determined to find success in sought after positions of employment).

The analysis of content in Facebook pages related to tinnitus in this study suggests that tinnitus sufferers use social media for various purposes, including gaining symptoms and diagnostic information, social support, learning to cope, and also to obtain information about research in this area. It is important to note that nearly half of the discussions in Facebook pages were related to diagnosis and symptoms (ie, 43.3%), suggesting that this platform is used by tinnitus sufferers for self-assessment of their condition. Facebook pages are fairly recent when compared with news media, and we explored Facebook pages information only during the years 2011-2016. Time series analysis showed no change in patterns of information during this time, unlike newspaper media. The results of this study relate well to previous studies on social media and other health conditions; for example, a recent study was conducted on the hearing aid community who used social media sources for advice and support, information sharing, and service-related information [19]. Analysis of information related to the level of usefulness of diabetes foot care-related Facebook groups showed a significant association with related posts [26]. Social media is also reported to be useful in creating indicators of the food environment that are associated with area-level mortality, health behaviors, and chronic conditions [27].

One would expect that the media would provide publicity for an intervention that is evidence-based. Although there are various treatments and/or management strategies available for tinnitus sufferers, psychological interventions such as cognitive behavioral therapy have the best evidence base for alleviating tinnitus distress [2]. However, it is surprising to see that there is a lack of publicity regarding psychological intervention, which becomes clear from the analysis of the information in the media. This is perhaps one of the reasons contributing to a relatively low uptake of evidence-based counseling interventions by tinnitus sufferers in the US population [28].

Overall, there is growing literature about the importance of health communication in chronic condition management, particularly concerning the role of mass media and social media in forming individuals’ opinions, and its bearing upon health behaviors. For this reason, it is important for health care professionals to be aware of the type of information that is being provided by the media on specific health conditions such as hearing loss and/or tinnitus.

Study Implications

This study has several practical implications. As the media plays an important role in formulating people’s knowledge and opinions [14], it is important for health care professionals to understand the importance of media on health decision and behavior. This, we believe, may assist health care professionals in better understanding the presuppositions patients may have about various health aspects. It is suggested that when used wisely and prudently, social media can help promote individual and public health [29]. This study helps tinnitus researchers to understand the ways in which the public opinion about tinnitus is formed over time. Stakeholders including health care professionals, patient organizations, and public health specialists can use the media to ensure that appropriate and accurate information is being disseminated to promote positive health behavior and health outcomes [30]. However, the first step in this process is to understand the trends or patterns of information being presented in the media, as explored in this study.

Study Limitations and Future Directions

This study is the first of its kind in the area of tinnitus. However, it has a few limitations. First, the text pattern analysis using software helps analyze big data, and although the analysis provides more of a macro view of the data, it may only provide a superficial understanding of its content. The automated data analysis may help us understand “what” aspects of tinnitus are represented in the media, rather than “how” tinnitus and these aspects are represented. For example, questions such as *Is brain stimulation portrayed positively or negatively in the media?*, or, *Is tinnitus portrayed as a curable condition?*, and finally, *How are the causes of tinnitus portrayed?* cannot be answered through this kind of analysis. Hence, in-depth qualitative analysis may be useful in future studies. Second, the newspaper data extracted from the ProQuest database may not be a true representation of all the US newspapers. Although ProQuest has some of the most widely circulated newspapers in the United States, using a different search engine may have potentially resulted in a different corpus and perhaps different results patterns. Furthermore, some selection bias may have occurred while choosing Facebook pages, as we only included 20 pages with 100 “likes.” Different criteria may have resulted in a different text corpus. This study mainly focuses on the patterns of information found in newspapers and in Facebook pages but not about how appropriate or accurate the information actually is. Hence, future studies should explore the appropriateness and accuracy of content. Moreover, future studies should aim to validate the results of automated text analysis using predictive validity methods, which confirms results by examining the data looking at the identified trends over time [31]. It would be interesting to study how the change over time may have occurred as a result of other metadata (eg, types of newspaper, focus of Facebook pages), which was not possible in this analysis due to limitations with software. Future studies should also consider studying the commonalities and differences among different types of media (eg, newspaper media, social media) using automated text analysis through more powerful software. Moreover, there is a great need to explore the role of the media on health behaviors, specifically in the tinnitus population.

Conclusions

This study explored how tinnitus is represented in the media, specifically in the US newspaper media and in Facebook pages. The information in the newspapers regarding tinnitus was mainly about brain stimulation, symptoms, coping, social support, and treatment innovations. Time series analysis showed that there were some changes in the patterns of information presented in newspaper media during 1990-2017. The information in Facebook pages about tinnitus was mainly related to diagnosis, cause, research and development, social support, challenges, symptoms, and coping. However, no changes in the patterns of information presented in Facebook pages between 2011 and 2016 were noted. These findings can help clinicians to better understand presuppositions tinnitus patients may have in regard to their condition. In addition, the findings can also be of interest to public health and health communication experts, allowing them to tailor health information about tinnitus to promote self-management and appropriate treatment choices for tinnitus sufferers.

Acknowledgments

The authors would like to acknowledge Gayatri Bhatta and Amit Kumar for help with data extraction from the US Newspaper database and Facebook pages and Prof William Harn for initial thoughts about text pattern analysis.

Conflicts of Interest

None declared.

References

1. 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]
2. Hesser H, Weise C, Westin VZ, Andersson G. A systematic review and meta-analysis of randomized controlled trials of cognitive-behavioral therapy for tinnitus distress. Clin Psychol Rev 2011 Jun;31(4):545-553. [doi: 10.1016/j.cpr.2010.12.006] [Medline: 21237544]
3. Østbye T, Yarnall KS, Krause KM, Pollak KI, Gradison M, Michener JL. Is there time for management of patients with chronic diseases in primary care? Ann Fam Med 2005;3(3):209-214 [FREE Full text] [doi: 10.1370/afm.310] [Medline: 15928223]
4. Grady PA, Gough LL. Self-management: a comprehensive approach to management of chronic conditions. Am J Public Health 2014 Aug;104(8):e25-e31. [doi: 10.2105/AJPH.2014.302041] [Medline: 24922170]
5. Coleman MT, Newton KS. Supporting self-management in patients with chronic illness. Am Fam Physician 2005 Oct 15;72(8):1503-1510 [FREE Full text] [Medline: 16273817]
6. Leventhal H, Meyer D, Nerenz D. The common sense representation of illness danger. In: Rachman S, editor. Contributions to Medical Psychology. New York: New York Pergamon Press; 1980:17-30.
7. Diefenebach MA, Leventhal H. The common-sense model of illness representation: theoretical and practical considerations. J Soc Distress Homeless 1996;5(11):11-38. [doi: 10.1007/BF02090456]
8. Rimal RN, Lapinski MK. Why health communication is important in public health. Bull World Health Organ 2009;87(4):247-247a [FREE Full text] [doi: 10.2471/BLT.08.056713] [Medline: 19551226]
9. Schwitzer G, Mudur G, Henry D, Wilson A, Goozner M, Simbra M, et al. What are the roles and responsibilities of the media in disseminating health information? PLoS Med 2005 Jul;2(7):e215 [FREE Full text] [doi: 10.1371/journal.pmed.0020215] [Medline: 16033311]

10. McCombs ME, Shaw DL. The agenda-setting function of mass media. Public Opin Q 1972;36(2):176-187. [doi: 10.1086/267990]
11. ComScore. 2008. Social networking increases worldwide as sites increase their focus on cultural relevance URL: https://www.comscore.com/Insights/Press-Releases/2008/08/Social-Networking-World-Wide?cs_edgescape_cc=US [accessed 2018-04-04] [WebCite Cache ID 6yRRuVYpk]
12. Stelter B. The New York Times. 2009 Jun 28. Journalism rules are bent in news coverage from Iran URL: https://www.nytimes.com/2009/06/29/business/media/29coverage.html [accessed 2018-04-04] [WebCite Cache ID 6yRS6tkZm]
13. McWay DC, Smith KJ. Social media exchanges are sometimes part of the health record. J AHIMA 2014 Mar;85(3):28-32; quiz 33. [Medline: 24707642]
14. McNab C. What social media offers to health professionals and citizens. Bull World Health Organ 2009 Aug 1;87(8):566. [doi: 10.2471/BLT.09.066712]
15. Hartz J, Chappell C. Nashville, TN: First Amendment Center; 1997. Worlds apart URL: http://www.newseuminstitute.org/freedom-forum/ [accessed 2018-04-04] [WebCite Cache ID 6yRSq8s7j]
16. Foss KA. (De)stigmatizing the silent epidemic: representations of hearing loss in entertainment television. Health Commun 2014;29(9):888-900. [doi: 10.1080/10410236.2013.814079] [Medline: 24299187]
17. Kensicki LJ. Deaf president now! Positive media framing of a social movement within a hegemonic political environment. J Commun Inq 2001;25(2):147-166. [doi: 10.1177/0196859901025002005]
18. Koerber R, Jennings MB, Shaw L, Cheesman M. Representations of workers with hearing loss in Canadian newspapers: a thematic analysis. Int J Audiol 2017 Dec 14;56(4):260-266. [doi: 10.1080/14992027.2016.1265155] [Medline: 27967271]
19. Choudhury M, Dinger Z, Fichera E. The utilization of social media in the hearing aid community. Am J Audiol 2017 Mar 01;26(1):1-9. [doi: 10.1044/2016_AJA-16-0044] [Medline: 28241238]
20. Moreno MA, Goniu N, Moreno PS, Diekema D. Ethics of social media research: common concerns and practical considerations. Cyberpsychol Behav Soc Netw 2013 Sep;16(9):708-713 [FREE Full text] [doi: 10.1089/cyber.2012.0334] [Medline: 23679571]
21. Ratinaud P. 2014. IRaMuTeQ: Interface de R pour les Analyses Multidimensionnelles de Texteset de Questionnaires (Version 0.7 alpha 2) [R Interface for Multidimensional Analyzes of Texts and Questionnaires. Free software built with free software] URL: http://www.iramuteq.org/ [accessed 2018-04-04] [WebCite Cache ID 6yRTzfIEU]
22. Ratinaud P, Marchand P. Recherche improbable d'une homogène diversité : le débat sur l'identité nationale [Improbable search of a homogeneous diversity: the debate on national identity]. Languages 2012;187(3):93-107. [doi: 10.3917/lang.187.0093]
23. Reinert M. Une méthode de classification descendante hiérarchique: application à l'analyse lexicale par contexte [A top-down hierarchical classification method: application to lexical analysis by context]. Cah Anal Donnees 1983;8(2):187-198 [FREE Full text]
24. Reinert M. Alceste : une méthodologie d'analyse de données textuelles et une application: Aurélia de G. de Nerval. Bull Methodol Sociol 1990;26(1):24-54. [doi: 10.1177/075910639002600103]
25. Sbalchiero S, Tuzzi A. Web scraping, topics, time and classification: comparing methods. 2017 Presented at: ITACOMS2017; 2017; Bologna, Italy URL: https://events.unibo.it/itacosm2017/...of.../Abstract_Tuzzi%20Sbalchiero.pdf
26. Abedin T, Al Mamun M, Lasker MA, Ahmed SW, Shommu N, Rumana N, et al. Social media as a platform for information about diabetes foot care: a study of Facebook groups. Can J Diabetes 2017 Feb;41(1):97-101. [doi: 10.1016/j.jcjd.2016.08.217] [Medline: 28126155]
27. Nguyen QC, Meng H, Li D, Kath S, McCullough M, Paul D, et al. Social media indicators of the food environment and state health outcomes. Public Health 2017 Jul;148:120-128. [doi: 10.1016/j.puhe.2017.03.013] [Medline: 28478354]
28. Tunkel DE, Bauer CA, Sun GH, Rosenfeld RM, Chandrasekhar SS, Cunningham Jr ER, et al. Clinical practice guideline: tinnitus executive summary. Otolaryngol Head Neck Surg 2014 Oct;151(4):533-541. [doi: 10.1177/0194599814547475] [Medline: 25274374]
29. Bernhardt JM, Alber J, Gold RS. A social media primer for professionals: digital dos and don'ts. Health Promot Pract 2014 Mar;15(2):168-172. [doi: 10.1177/1524839913517235] [Medline: 24396119]
30. Ventola CL. Social media and health care professionals: benefits, risks, and best practices. P T 2014 Jul;39(7):491-520 [FREE Full text] [Medline: 25083128]
31. Grimmer J, Stewart BM. Text as data: the promise and pitfalls of automatic content analysis methods for political texts. Political Anal 2013;21(3):267-297 [FREE Full text]

Edited by G Eysenbach; submitted 27.09.17; peer-reviewed by R Kriahns, D Hoare, K Greenwell, G Searchfield, N Limsopatham; comments to author 03.12.17; revised version received 07.02.18; accepted 16.03.18; published 08.05.18.

Please cite as:
Manchaiah V, Ratinaud P, Andersson G
Representation of Tinnitus in the US Newspaper Media and in Facebook Pages: Cross-Sectional Analysis of Secondary Data
Interact J Med Res 2018;7(1):e9
URL: http://www.i-jmr.org/2018/1/e9/
doi:10.2196/ijmr.9065
PMID:29739734

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