Interactive Journal of Medical Research

Contents

Original Papers

Original Paper

Views of People With Epilepsy About Web-Based Self-Presentation: A Qualitative Study

Alison Ruth McKinlay[1], PhD; Leone Lorna Ridsdale[1], MD

Department of Basic & Clinical Neuroscience, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, United Kingdom

Corresponding Author:
Leone Lorna Ridsdale, MD
Department of Basic & Clinical Neuroscience
Institute of Psychiatry, Psychology & Neuroscience
King's College London
PO Box 57, Academic Neuroscience Centre
Denmark Hill
London, SE5 8AF
United Kingdom
Phone: 44 020 7848 0185
Email: leone.ridsdale@kcl.ac.uk

Abstract

Background: Web-based media, particularly social networking sites (SNSs), are a source of support for people with long-term conditions, like epilepsy. Living with epilepsy can reduce opportunities for accessing information and social support owing to transportation difficulties and stigma leading to self-isolation. However, some people with epilepsy (PWE) overcome these barriers using SNSs and other Web-based media. At present, little is known about Web-based identity and self-presentation of PWE; this study aims to address this gap.

Objective: This study aims to describe how the use of digital technologies, such as SNSs, impacts sense of identity in PWE.

Methods: We used qualitative research methods to examine Web-based media use and self-presentation in a group of 14 PWE (age range: 33-73 years; 7 men and 7 women). The median diagnosis duration was 25 years. Semistructured interviews ranged from 40 to 120 minutes, held at participants' homes or in a public place of their choice, in the United Kingdom. QSR Nvivo 11 software was used to perform an inductive thematic analysis.

Results: In this study, 9 participants used Web-based media to "silently" learn from other PWE by reading user posts on SNSs and epilepsy-related forums. When asked about self-presentation, 7 participants described feeling cautious about disclosing their epilepsy to others online. Six participants presented themselves in the same manner irrespective of the situation and described their identity as being presented in the same way both online and offline.

Conclusions: PWE can deploy SNSs and Web-based media to manage aspects of their condition by learning from others and obtaining social support that may otherwise be difficult to access. Some PWE share openly, whereas others silently observe, without posting. Both benefit from the shared experiences of others. Privacy concerns and stigma can act as a barrier to sharing using Web-based media and SNSs. For some, Web-based media offers a chance to experiment with identity and change self-presentation, leading to gradually "coming out" and feeling more comfortable discussing epilepsy with others.

(Interact J Med Res 2018;7(2):e10349)   doi:10.2196/10349

Introduction

Web-based media, including mobile phone apps and social networking sites (SNSs), enable the presentation of a *curated self*, where users can share selected aspects of their lives such as new relationships, pictures, and personal insights [1]. In June

2017, it was estimated that the number of monthly Facebook users had reached 2 billion [2] with 90% of older adults having used this SNS to access and share health-related information [3]. The phenomenon of using Web-based media as a health care aid is being referred to as *Health 2.0* [4]. Users can share their experiences of illness and learn about health care self-management in a Web-based environment [5]. Some users

may even modify their offline health care with information gathered online.

Web-based media may be used in place of a traditional face-to-face support group, with similar or added benefits, including the possibility to remain anonymous [6]. Anonymity may protect users against social stigma [6], where an individual is thought to have characteristics perceived as undesirable or unacceptable [7]. Through Web-based media, users can access information and support in real time that people with health conditions might have otherwise waited for a medical consultation to access. In one study of people with diabetes, Facebook groups were used to share experiences and gain knowledge through the experiences of others [8]. People with other long-term conditions may gain similar benefits from Web-based media use [9].

As a result of increased SNS use, where users can disclose information about themselves to others [10,11], topics such as Web-based identity and self-presentation have become a focus of research [12]. In 1959, Goffman described identity presentation as actors performing on stage; different aspects of an individual's identity are presented in different contexts [13]; this now also occurs in Web-based spaces [14]. SNSs provide a platform for identity performance, where practices in self-presentation can undergo continued maintenance and revision [15]. When Web-based users modify identity portrayal to fit with what is perceived to be an *ideal self*, these performances can be curated and reused in other social contexts [12,16].

Self-presentation is defined as the deliberate act of portraying identity in a certain, premeditated manner [17]. Practices in self-presentation, including deliberate nondisclosure, can mask users' daily struggles [18], creating bias and even misrepresentation of experiences shared online. In addition, self-presentation can be influenced by stigma, which has been associated with long-term health conditions such as psychosis, epilepsy, and HIV or AIDS [19]. Through the use of online support groups and user forums, people with stigmatized health conditions can share the aspects of their identity with others and may feel safer to do so than face-to-face interactions [6]. Past research has focused on the impact of Web-based self-presentation on long-term conditions, such as diabetes, chronic fatigue [20], and chronic cough [21], but a gap exists in the literature on epilepsy and Web-based identity.

Epilepsy is a neurological condition characterized by recurrent seizures, affecting 1% of the United Kingdom (UK) population [22] and approximately 50 million people worldwide [23]. Treatment consists of medication use and advice on self-management. Despite treatment, up to 40% of people continue to experience seizures [23] and low quality of life [24]. People with epilepsy (PWE) frequently use the internet for information about epilepsy [25], and the wider literature show the benefit of Web-based communities in the self-management of other long-term conditions [9]. The benefits of the SNS use

may be especially helpful for PWE, where unpredictability of seizures, transportation difficulty owing to a lack of driving license, and stigma, can limit access to face-to-face sources of information and support. This research aims to describe the Web-based practices of PWE and explore how SNSs use is associated with identity and self-presentation.

Methods

Study Design

We used a qualitative, cross-sectional research design, using purposive sampling methods and semistructured interviews. Ethical approval was granted by the Psychiatry, Nursing and Midwifery Research Ethics subcommittee at King's College London (PNM/13/14-18). All participants were given an information sheet and gave informed consent at the start of their interview.

Participants and Recruitment

Researchers advertised for volunteers through the webpage of the largest official UK epilepsy user group (Epilepsy Action, hereafter referred to as "user-group"), as well as their email list of 3500 members. The study eligibility criteria included the following: aged ≥17 years; self-reported epilepsy diagnosis; and experience of accessing user-group website information.

Overall, 35 PWE emailed the research team to register their interest in participating. Fifteen people attended interviews, with the final sample size of 14; this number of participants is acceptable in qualitative research [26], where the aim is to describe and explore rather than generalize. Of 20 who did not attend interviews, 2 respondents were based remotely and unable to attend interviews in-person, 2 respondents canceled their appointments owing to family emergency or holiday, and the remainder did not respond to further interview invitations. No incentives were offered to any participant involved in the study.

Data Collection and Analysis

Interviews followed a semistructured format. Questions were asked on various uses of Web-based media, which included SNSs, mobile phones, tablets, and computers. The interview topics (Table 1) were developed by LLR with input from a sociologist.

Interviews lasted between 40 and 120 minutes, were audiorecorded, and later transcribed by a professional transcription agency. All transcripts were stored on a secured server and, when printed, in a locked filing cabinet where only numeric identifiers were used to identify participants.

Interviews were conducted by a postdoctoral researcher (RR) between May and October 2015. By way of preparation, the researcher attended a self-management course for PWE [27] to gain an understanding of epilepsy-related challenges (ie, stigma and driving restrictions). Interviews took place either in participants' homes or a public place.

Table 1. Question guide.

Interview Topic	Questions
Diagnosis	- What kind of epilepsy do you have? - When were you diagnosed?
General experience of epilepsy offline	- Do you keep a symptoms or medication diary? Reminders? Family member? - Do you talk to people about your epilepsy? Who? - Do you know anyone else with epilepsy? Fictional characters? - Have you ever been to a support group? What was it like? Charity work? Advocacy? Why did you go?
General use of the internet and social media	- What sites do you use? Twitter, Facebook, Email, Tumblr, Blogs, Instagram, LinkedIn, Forums, or chat rooms? - How often? - When did you start? Why? Have you stopped using any? Why? - What devices do you use to access these sites or apps? - Do you change the way you present yourself on different sites or apps? Why? Do you try to? Why?
Broad experience of epilepsy online	- Do you talk about your epilepsy on these sites? In private messages? Why? Is privacy an issue? - How do you think epilepsy is portrayed in general online? For example, trolling, surgery games, "seizure" on Twitter, Photosensitive-provoking gifs. - Do you think your epilepsy affects the way you use these sites?
Epilepsy-specific websites	- How often do you use the Epilepsy Action forum? Why? Is it different from talking offline? - Do you use any other epilepsy forums or websites? Blogs? Why?—Information, friendship, advocacy? - Do you think these experiences have changed the way you think about your epilepsy? Yourself? The online world? - Has it changed the way you interact with doctors? Friends? Other people with epilepsy? - How could these websites be improved?

We used the thematic analysis when examining interview transcripts [28]. Thematic analysis is not tied to a specific methodological orientation, making it a flexible analytic method, suitable for health services research. The lead author (ARM) read each transcript twice and drafted a document of initial impressions, which was then reviewed by the second author (LLR). A set of codes were inductively produced. The codes were then organized into an set of themes. The lead author presented and gained feedback on the initial themes to 2 groups of researchers. One group had expertise in life writing and online identity, and another had expertise in neuroscience research. The themes and codes were further refined in light of this feedback. Finally, a series of definitions were developed for each theme and the final results report produced.

Results

Participants' Characteristics

There were 14 participants of an equal number of men and women (Table 2), with a median age of 50 (range: 33-73) years and the median number of 25 years since diagnosis (range: 13-63 years). Participants self-reported generalized tonic-clonic seizures (n=7), complex partial seizures (n=4), photosensitivity (n=4), and absences (n=3). Three participants disclosed a comorbid health condition, including diabetes, asthma, and physical disability.

Silent Observation of Others on Social Media and Websites

Nine participants described using Web-based media to learn from the shared accounts and posts from other PWE. Many described reading but not necessarily posting or interacting with others; this practice was explained as a desire to keep up-to-date with epilepsy-related information, without feeling the need to respond or contribute themselves.

> *I have used [user-group websites] for research on epilepsy medication. That's probably about it. I use forums for that, but I would never post on them.* [p13]

Social Comparison Assisting With Epilepsy Management

Seven participants spoke about using Web-based media for social support and to find narratives from other PWE as a way of learning about their condition. The use of social comparison in this context seemed to be an indirect form of social support. One participant (p11) was able to read how other PWE had coped with the same kind of neurosurgery they were planning to have. Another participant (p13) found an explanation for the medication side effects they had experienced from taking Keppra (generic name levetiracetam, an antiepileptic drug introduced in the United States in 2008 and the United Kingdom in 2011).

> *[I look for] personal stories of how they've got on with their medication, how they've got on with reactions to it and that kind of thing...I found myself getting angry, so I was finding out about the whole Keppra rage thing and seeing whether other people got it.* [p13]

XSL•FO

RenderX

Table 2. The demographic characteristics of participants.

Participant number	Sex	Age	Diagnosis length (years)	Social media platforms used
p1	Female	46	43	FB[a], Twitter, LinkedIn, user-group[b] forum
p2	Male	35	17	FB, FB messenger, Twitter, LinkedIn
p3	Female	50	15	LinkedIn
p4	Female	38	25	FB, Twitter, LinkedIn, user-group chat rooms
p5	Male	59	36	FB
p6	Female	43	25	Facetime, FB (including private epilepsy-related FB group)
p7	Female	43	24	FB, Twitter
p8	Male	56	35	FB (including messenger), Twitter, Instagram, foursquare, WhatsApp, Snapchat, Blackberry messenger, Skype
p9	Female	63	15	Twitter
p10	Male	52	47	FB, Twitter, private message on Epilepsy Action forum
p11	Male	65	Not specified	FB, user-group chat rooms, Skype
p12	Female	64	63	FB, user-group forums
p13	Male	33	13	FB, WhatsApp, Twitter
p14	Male	73	58	FB (including messenger), Twitter

[a]FB: Facebook.

[b]UK-based charity called Epilepsy Action.

Self-Identifying With Epilepsy

Six participants described the way they present themselves as being inseparable on websites and in real life. "Me" and "my epilepsy" were expressed as one singular identity.

> *I am what I am, if people don't like me, I don't care whether they do.* [p7]

One participant (p12) explained this in terms of wanting to be "as honest as possible" online and offline. One participant's (p5) openness about epilepsy came about after observing other people disclosing their epilepsy on a nonepilepsy-related website.

Self-Disclosure Dependent on the Audience

Five participants spoke about modifying their Web-based self-presentation depending on which SNS platform they were using, or whom they felt the audience was likely to be. When posting on a user-group chat room, as opposed to public epilepsy-related SNS pages, one participant felt he could talk more openly.

> *It changes the actual thing of what you want to say to the person. What they want to say to you, and obviously it makes a difference in the conversation and what it builds up to...In the epilepsy chat room, you can talk about more or less everything that goes on as on a day-to-day basis, your medication and everything...On Facebook [public epilepsy pages] you have to be careful what you are saying and what you put down.* [p10]

One of these participants (p7) said she did not talk about epilepsy online owing to concerns about privacy. Another (p3) was concerned about future employers learning through LinkedIn that she had left her previous job because of having a serious seizure. Several did not want the "world at large" to know about their epilepsy and, therefore, tended to only share with close friends privately or on closed epilepsy-specific Facebook groups.

Cautious About Web-Based Disclosure and Self-Presentation

Four participants described some level of apprehension about having epilepsy-related content on their personal social media pages, as though there was a separation of their "public" Web-based identity and epilepsy-related identity. Unless directly asked by others, they did not discuss the day-to-day elements of their condition openly.

> *So that's the only thing with Facebook, it highlights the good bits of what people do. It doesn't tend to show them at their worst. I think nobody would want to be my friend if I showed my worst.* [p6]

On the other hand, another participant (p1) had initially felt apprehension about identity and self-disclosure offline but had since become more confident in both domains after talking about epilepsy online.

> *I didn't make it public, but I didn't entirely hide it...I would tell them, but if they didn't ask, I wouldn't tell them...I'm only recently trying to be more open about my epilepsy online because, for years, I wasn't very public about it...It's a new thing for me to able to discuss it and, for so long, I couldn't.* [p1]

Presenting Epilepsy to Web-Based Social Contacts

Four participants used Web-based media to indirectly share or create awareness about the topic of epilepsy with their social

contacts. One of these participants (p1) described not feeling completely comfortable when talking about herself as someone with epilepsy but was happy to have her image on Facebook and a user-group website having a telemetry assessment. Two others described sharing images of themselves to their wider SNS networks (ie, through Twitter and Facebook) supporting epilepsy-related charities, while not openly disclosing that they have epilepsy.

> *The photo of me on there is the day I did my Epilepsy Action Run...I had the t-shirt on and I had my kids sat either side of me. So, you'd probably know that I had some connection to epilepsy that way.* [p2]

One participant spoke early in his interview of there being a shift from feeling concerned about self-disclosure of his epilepsy, to no longer being worried about the audience; this occurred once his employment situation changed, from working for someone else to running his own business:

> *I don't wear a little badge saying I'm epileptic...I put it on my Facebook to make it clear to people who support it. So it doesn't worry me now. Certainly, with the fact that I'm working for myself it doesn't worry me.* [p11]

Discussion

Principal Findings

This study demonstrates that PWE received personally significant support and value from their SNSs and Web-based media use. For some, Web-based media was used as a forum to overcome a stigmatized identity, as well as develop connections with other PWE.

As with previous research on the uses of digital technology [25,29], PWE in this study used Web-based media to access social support. Connecting with others online about the shared experience of a long-term condition can reduce distress and provide reassurance [21]. Such benefits are particularly helpful for people with recurrent seizures, where high levels of anxiety and low quality of life are commonly reported [30].

In addition, accessing support through SNSs and user-group websites can have additional advantages over traditional face-to-face interactions. Users can read highly personal accounts from other users in several clicks; while individuals may not be as forthcoming with personal information when conversing with new social contacts offline [31]. The perceived benefit of Web-based media used in this way is the constancy of support it provides. Posts in a Facebook group are available in an asynchronous format (ie, not real time but when a user consciously checks), while private direct messages and chat provide synchronous connections with others in real time [6]. SNSs are available 24 hours a day on mobile phone or computer; thus, PWE have the potential to be connected with continual support and social connection [14]. As many PWE experience difficulties with transport, and can have unpredictable seizures, the function of Web-based media as a support group can be especially valuable.

Almost half of the participants in our group said that they modified their Web-based self-presentation depending on the perceived or *imagined audience* [32]. While in face-to-face interactions, people may perceive the audience as those whom they can see, in Web-based spaces, the audience and privacy level can vary greatly [32]. With the advent of screenshots and Web-based archiving, content posted online may be viewed by those other than whom the poster intended at the time of disclosing their personal information [33]; this creates the potential for digital stigma, where discrimination or embarrassment results from private information being publicly posted and accessed [34].

Closed Facebook groups require further evaluation as a possible tool for chronic illness management [35]. Our results show there was more of a tendency to discuss epilepsy freely in a closed epilepsy group on Facebook or message (which is considered "more private"), compared with a public Facebook page with open privacy settings. This behavior might be conceptualized in terms of theory on felt stigma, previously associated with epilepsy [36]. The perception of being different or unaccepted by others (ie, "felt stigma") [37] may explain why some participants felt they could only share their epilepsy experiences in some areas of SNSs (ie, closed Facebook group) as opposed to others who shared openly in all settings. Modifying SNSs to address stigma-related barriers may help to improve the other benefits PWE gain from using Web-based media to self-manage their condition.

Participants reported the use of Web-based media that would be expected of individuals with general health concerns (ie, information seeking). However, the experience of epilepsy is long-term, and symptoms can be debilitating, which adds further complexity to how individuals manage their day-to-day life. PWE can experience the perception of having personal characteristics that are discrediting, a kind of "spoiled identity" [7]. As a consequence, individuals living with chronic health conditions like epilepsy can withdraw and isolate themselves from others; this creates a vicious cycle, whereby individuals do not receive support or feedback and do not learn about managing their condition. SNSs are uniquely placed to disrupt this loop and overcome a stigmatized or spoiled identity by providing feedback from other PWE, observing others online, while maintaining anonymity. Over time, this may translate into changes in individual self-presentation as perceptions of identity are modified.

Changes in self-presentation over time were described by several participants, through either modeling observed disclosures of other PWE or using Web-based media to improve self-confidence. One participant, in particular, felt more able to talk about her epilepsy offline after discussing her condition with others online. This could be explained as a form of *identity play*, whereby profound development and personal change are precipitated by the life one leads using Web-based media [38]. For this participant, her interactions online provided a place to monitor or experiment with epilepsy-related identity performance, which affected her offline life. This form of identity play or management has been identified with other patient groups using Web-based spaces, but with the added

consideration that if privacy is compromised throughout this process, this could cause great distress to an individual [20].

In this study, almost half of the participants said that their self-presentation was the same regardless of the Web-based platform. They were also more likely to be "openly out" about their epilepsy in their offline lives. These SNS users may perceive a kind of social currency with portraying their authentic selves as a way of developing connections or engaging others online [32]. Several participants carried on long-distance friendships with PWE they had initially spoken with, which could further explain personal motivations to self-present in a way they felt was congruent with their actual selves [39]. The possibility of meeting face-to-face may also increase the likelihood of SNS users presenting their actual rather than idealized selves [40].

Study Implications

Findings from this study demonstrate how a group of PWE used Web-based media to manage elements of their epilepsy, through silent observation and making connections with others. In addition, Web-based media can facilitate a dialogue between health care professionals and patients. By reading shared Web-based experiences, health care professionals may gain insight into patient responses after using their services [20]. Some suggest health care practitioners may further support PWE by increasing their Web-based presence and engaging in public discussion about epilepsy care [41]. Despite concerns about Web-based information accuracy, evidence from research into other health care conditions shows forum posts of this nature are likely to be accurate [8] and, therefore, valuable to PWE; however, professional moderation and fact-checking is necessary [21].

Many PWE reported looking at social media (ie, forums and Facebook groups) for epilepsy-related information but not interacting with other users; this is also an important finding for health care providers who post educational content online, as PWE may gain value from epilepsy-related information without necessarily interacting or participating. Social media is inexpensive to use and may be an effective means of delivering health care information to the public [42]. However, without safety protocols in place, this can restrict the development of Web-based media as a tool for health care purposes [43]. Further training and support for those using these platforms for health care purposes would be needed [5,44].

In the past decade, issues of user privacy and data use have become linked to SNSs and can affect how users present themselves online [45]. An awareness of these issues can influence an individual's propensity to self-disclose personal information online [20,46]. While many of our participants were willing to share their experiences and seemed to find value in sharing with other PWE, many were also concerned about privacy or oversharing information on public SNSs. The benefits of mutual self-disclosure and information sharing may be enhanced if SNS platforms offer functions that do more to protect the privacy and confidentiality of users [10]. Closed Facebook groups, private messages, and user-group forums seemed to provide such a function for PWE in this study.

Strengths and Limitations

To the best of our knowledge, this was the first study of its kind to include the perspective of PWE on their use of Web-based media relating to their identity and self-presentation. Although, in seeking to describe how PWE use Web-based media to understand and self-manage their epilepsy, the study has several limitations. The people in our study were volunteers recruited from a user-group charity and identified by their use of Web-based media most broadly. Our participants were at various tech levels, ranging from early adopters who have greater digital literacy to those who use the internet and social media infrequently; they were more likely to be aged in their forties and above and had been living with epilepsy for >25 years. Future research with participants earlier on in their illness trajectory may help identify additional themes not captured in this study. It is also likely that adults aged 18-29 years, who are more frequent users of SNSs [47], would offer a different perspective, such as the potential risks and consequences of Web-based media, than those described in this study. These demographic characteristics are, nevertheless, comparable to participants of National Health Service users, recruited into a trial of group education for PWE in England [30].

Conclusions

Communication facilitated through digital technology and Web-based media has the potential to powerfully reshape an individual's experience of living with a long-term condition [9]. Self-presentation and Web-based disclosure are complex processes, reflecting the challenges posed by stigmatized identities. Although Web-based identity may appear to be represented through "static" posts or tweets, it is a constant reconstructive and reflexive process [15], with the potential to be impacted by affiliation with health-related charities and causes, as was described by several in this study. The use of SNSs to gather information and support represents an opportunity for health care providers and researchers to utilize SNS platforms to improve the patient experience. Some participants changed their self-presentation for reasons such as employment consequences, with a fear of stigma modifying their SNS use. Learning to guard, present, and manage their image relationally was an important benefit of using Web-based media for people in coming to terms with the stigma of epilepsy.

XSL·FO

RenderX

Conflicts of Interest

None declared.

References

1. Holmes KM, O'Loughlin N. The experiences of people with learning disabilities on social networking sites. Br J Learn Disabil 2012 Sep 07;42(1):1-5. [doi: 10.1111/bld.12001]

2. Lanchester J. The London Review of Books. 2017. You Are the Product URL: https://www.lrb.co.uk/v39/n16/john-lanchester/you-are-the-product [accessed 2018-03-06] [WebCite Cache ID 6xidfwmK9]

3. Tennant B, Stellefson M, Dodd V, Chaney B, Chaney D, Paige S, et al. eHealth literacy and Web 2.0 health information seeking behaviors among baby boomers and older adults. J Med Internet Res 2015 Mar;17(3):e70 [FREE Full text] [doi: 10.2196/jmir.3992] [Medline: 25783036]

4. Sarasohn-Kahn J. California HealthCare Foundation. 2008. The wisdom of patients: Health care meets online social media URL: https://www.chcf.org/wp-content/uploads/2017/12/PDF-HealthCareSocialMedia.pdf [accessed 2018-03-03] [WebCite Cache ID 6xiipbI1y]

5. Moorhead SA, Hazlett DE, Harrison L, Carroll JK, Irwin A, Hoving C. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. J Med Internet Res 2013 Apr;15(4):e85 [FREE Full text] [doi: 10.2196/jmir.1933] [Medline: 23615206]

6. Wright KB, Bell SB, Wright KB, Bell SB. Health-related Support Groups on the Internet: Linking Empirical Findings to Social Support and Computer-mediated Communication Theory. J Health Psychol 2003 Jan;8(1):39-54. [doi: 10.1177/1359105303008001429] [Medline: 22113899]

7. Goffman E. Stigma: Notes on the Management of Spoiled Identity. New York: Penguin; 1990.

8. Greene JA, Choudhry NK, Kilabuk E, Shrank WH. Online social networking by patients with diabetes: a qualitative evaluation of communication with Facebook. J Gen Intern Med 2011 Mar;26(3):287-292 [FREE Full text] [doi: 10.1007/s11606-010-1526-3] [Medline: 20945113]

9. Allen C, Vassilev I, Kennedy A, Rogers A. Long-Term Condition Self-Management Support in Online Communities: A Meta-Synthesis of Qualitative Papers. J Med Internet Res 2016 Mar 10;18(3):e61 [FREE Full text] [doi: 10.2196/jmir.5260] [Medline: 26965990]

10. Bateman P, Pike J, Butler B. To disclose or not: publicness in social networking sites. Info Technology & People 2011 Mar;24(1):78-100. [doi: 10.1108/09593841111109431]

11. Pike J, Bateman P, Butler B. I didn't know you could see that: the effect of social networking environment characteristics on publicness and self-disclosure. In: AMCIS 2009 Proceedings. Presented at: Fifteenth Americas Conference on Information Systems; 6-9th August, 2009; San Francisco, California.

12. Mehdizadeh S. Self-presentation 2.0: narcissism and self-esteem on Facebook. Cyberpsychol Behav Soc Netw 2010 Aug;13(4):357-364. [doi: 10.1089/cyber.2009.0257] [Medline: 20712493]

13. Goffman E. The Presentation of Self in Everyday Life. London: Penguin; 1990.

14. Hodkinson P. Bedrooms and beyond: Youth, identity and privacy on social network sites. New Media & Society 2017;19(2):272-288. [doi: 10.1177/1461444815605454]

15. Cover R. Becoming belonging: performativity, subjectivity,the cultural purposes of social networking. In: Poletti A, Rak J, editors. Identity Technologies: Constructing the Self Online. Wisconsin: The University of Wisconsin Press; 2014:55-69.

16. Seidman G. Self-presentation and belonging on Facebook: How personality influences social media use and motivations. Pers Individ Dif 2013;54(3):402-407. [doi: 10.1016/j.paid.2012.10.009]

17. Schlenker BR, Leary MR. Social anxiety and self-presentation: a conceptualization and model. Psychol Bull 1982 Nov;92(3):641-669. [Medline: 7156261]

18. Steers M, Wickham R, Acitelli L. Seeing Everyone Else's Highlight Reels: How Facebook Usage Is Linked to Depressive Symptoms. J Soc Clin Psychol 2014;33(8):701-731.

19. Scambler G. Health-related stigma. Sociol Health Illn 2009 Apr;31(3):441-455 [FREE Full text] [doi: 10.1111/j.1467-9566.2009.01161.x] [Medline: 19366430]

20. Brady E, Segar J, Sanders C. "I Always Vet Things": Navigating Privacy and the Presentation of Self on Health Discussion Boards Among Individuals with Long-Term Conditions. J Med Internet Res 2016 Dec 13;18(10):e274 [FREE Full text] [doi: 10.2196/jmir.6019] [Medline: 27737819]

21. Sinha A, Porter T, Wilson A. The Use of Online Health Forums by Patients With Chronic Cough: Qualitative Study. J Med Internet Res 2018 Jan 24;20(1):e19 [FREE Full text] [doi: 10.2196/jmir.7975] [Medline: 29367181]

22. Ridsdale L, Charlton J, Ashworth M, Richardson MP, Gulliford MC. Epilepsy mortality and risk factors for death in epilepsy: a population-based study. Br J Gen Pract 2011 May;61(586):e271-e278 [FREE Full text] [doi: 10.3399/bjgp11X572463] [Medline: 21619751]

23. World Health Organisation. 2017. Epilepsy fact sheet URL: http://www.who.int/mediacentre/factsheets/fs999/en/ [WebCite Cache ID 6t55Xlegt]

XSL•FO

RenderX

24. Baker GA, Jacoby A, Buck D, Stalgis C, Monnet D. Quality of life of people with epilepsy: a European study. Epilepsia 1997 Mar;38(3):353-362 [FREE Full text] [Medline: 9070599]

25. Escoffery C, Diiorio C, Yeager KA, McCarty F, Robinson E, Reisinger E, et al. Use of computers and the Internet for health information by patients with epilepsy. Epilepsy Behav 2008 Jan;12(1):109-114. [doi: 10.1016/j.yebeh.2007.07.013] [Medline: 17959420]

26. Guest G, Bunce A, Johnson L. How Many Interviews Are Enough? Field Methods 2006;18(1):59-82. [doi: 10.1177/1525822X05279903]

27. Magill N, Ridsdale L, Goldstein LH, McCrone P, Morgan M, Noble AJ, et al. Self-management education for adults with poorly controlled epilepsy (SMILE (UK)): statistical, economic and qualitative analysis plan for a randomised controlled trial. Trials 2015 Jun 12;16:269 [FREE Full text] [doi: 10.1186/s13063-015-0788-9] [Medline: 26068449]

28. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]

29. Hixson JD, Barnes D, Parko K, Durgin T, Van Bebber S, Graham A, et al. Patients optimizing epilepsy management via an online community. Neurology 2015 Jun 17;85(2):129-136. [doi: 10.1212/wnl.0000000000001728]

30. Ridsdale L, Philpott SJ, Krooupa A, Morgan M. People with epilepsy obtain added value from education in groups: results of a qualitative study. Eur J Neurol 2017 Dec;24(4):609-616 [FREE Full text] [doi: 10.1111/ene.13253] [Medline: 28181344]

31. Ziebland S, Wyke S. Health and illness in a connected world: how might sharing experiences on the internet affect people's health? Milbank Q 2012 Jun;90(2):219-249 [FREE Full text] [doi: 10.1111/j.1468-0009.2012.00662.x] [Medline: 22709387]

32. Marwick AE, Boyd D. I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. New Media & Society 2010 Jul 07;13(1):114-133. [doi: 10.1177/1461444810365313]

33. Boyd D, Ellison N. Social Network Sites: Definition, History, and Scholarship. J Comput Mediat Commun 2007;13(1):210-230. [doi: 10.1111/j.1083-6101.2007.00393.x]

34. Trottier D. Identity Problems in the Facebook Era. New York: Routledge; 2013.

35. Partridge SR, Gallagher P, Freeman B, Gallagher R. Facebook Groups for the Management of Chronic Diseases. J Med Internet Res 2018 Jan 17;20(1):e21 [FREE Full text] [doi: 10.2196/jmir.7558] [Medline: 29343460]

36. Jacoby A. Felt versus enacted stigma: a concept revisited. Evidence from a study of people with epilepsy in remission. Soc Sci Med 1994 Jan;38(2):269-274. [Medline: 8140453]

37. Scambler G, Hopkins A. Being epileptic: coming to terms with stigma. Sociol Health & Illness 1986 Mar;8(1):26-43. [doi: 10.1111/1467-9566.ep11346455]

38. Turkle S. Cyberspace and Identity. Contemporary Sociology 1999 Nov;28(6):643. [doi: 10.2307/2655534]

39. Tifferet S, Vilnai-Yavetz I. Gender differences in Facebook self-presentation: An international randomized study. Comput Human Behav 2014;35:388-399.

40. Ellison N, Heino R, Gibbs J. Managing Impressions Online: Self-Presentation Processes in the Online Dating Environment. J Comp Mediated Comm 2006 Jan;11(2):415-441. [doi: 10.1111/j.1083-6101.2006.00020.x]

41. Meng Y, Elkaim L, Wang J, Liu J, Alotaibi N, Ibrahim G, et al. Social media in epilepsy: A quantitative and qualitative analysis. Epilepsy Behav 2017;71:79-84.

42. Gough A, Hunter RF, Ajao O, Jurek A, McKeown G, Hong J, et al. Tweet for Behavior Change: Using Social Media for the Dissemination of Public Health Messages. JMIR Public Health Surveill 2017 Mar 23;3(1):e14 [FREE Full text] [doi: 10.2196/publichealth.6313] [Medline: 28336503]

43. Hawn C. Take two aspirin and tweet me in the morning: how Twitter, Facebook, and other social media are reshaping health care. Health Aff (Millwood) 2009;28(2):361-368. [doi: 10.1377/hlthaff.28.2.361] [Medline: 19275991]

44. Antheunis ML, Tates K, Nieboer TE. Patients' and health professionals' use of social media in health care: motives, barriers and expectations. Patient Educ Couns 2013 Sep;92(3):426-431. [doi: 10.1016/j.pec.2013.06.020] [Medline: 23899831]

45. Rathore S, Sharma P, Loia V, Jeong YS, Park J. Social network security: Issues, challenges, threats, and solutions. Information Sciences 2017:43-69 [FREE Full text]

46. Nemec Zlatolas L, Welzer T, Heričko M, Hölbl M. Privacy antecedents for SNS self-disclosure: The case of Facebook. Comput Human Behav 2015 Apr;45:158-167 [FREE Full text] [doi: 10.1016/j.chb.2014.12.012]

47. Perrin A. Pew Research Centre. 2015. Social Media Usage: 2005-2015 URL: http://assets.pewresearch.org/wp-content/uploads/sites/14/2015/10/PI_2015-10-08_Social-Networking-Usage-2005-2015_FINAL.pdf [accessed 2018-03-03] [WebCite Cache ID 6xif8fHAN]

Abbreviations

PWE: people with epilepsy
SNS: social networking site
UK: United Kingdom

XSL•FO
RenderX

Original Paper

Use of Electronic Entertainment and Communication Devices Among a Saudi Pediatric Population: Cross-Sectional Study

Sami Omar Amawi[1], MBBS, MSc; Ahmed Hussein Subki[1], MBBS; Hazim Abdulkarim Khatib[1]; Omer Sameer Alkhateeb[1]; Rida Hashim Fida[1]; Omar Mohammed Saggaf[1]; Mohammed Mohammed Jan[1], MBChB, FRCP(C)

Department of Pediatrics, College of Medicine, King Abdulaziz University, Jeddah, Saudi Arabia

Corresponding Author:
Ahmed Hussein Subki, MBBS
Department of Pediatrics
College of Medicine
King Abdulaziz University
Jeddah,
Saudi Arabia
Phone: 966 560662735
Email: ahs.subki@gmail.com

Abstract

Background: Excessive use of various electronic entertainment and communication devices, particularly among children, has been associated with increased behavioral problems. Despite children's escalating use of these devices, parents' awareness about the impact thereof is still lacking.

Objective: The objective of this study was to assess the use of electronic entertainment and communication devices among children attending a health care facility in Jeddah, Saudi Arabia, as well as the parental impression regarding the impact of electronic devices use on the behavior of their child.

Methods: A focused 15-item questionnaire was designed for this cross-sectional study involving mothers of children attending the Well Baby Clinic of King Abdulaziz University Hospital, Jeddah, Saudi Arabia from July 1, 2016 to November 30, 2016.

Results: This study included 190 mothers. The mean ages of the children, mothers, and fathers were 7.3 (SD 3.5), 35 (SD 6.5), and 43 (SD 8.3) years, respectively. Most children were of Saudi Arabian nationality (106/190, 55.8%). The most used device in this study was television (154/190, 81.0%), followed by mobile phones (134/190, 70.5%), and tablets (116/190, 61.0%). Computers were the least used device in this study (59/190, 31.0%). In total, 24.7% (47/190) of children in this study used electronic entertainment and communication devices for more than 4 hours per day. Most mothers (129/190, 67.9%) felt that their child spends too much time on electronic devices. Hyperactivity or behavioral problems were reported by 20.0% (38/190) of mothers in this study. Children spending longer hours on electronic devices were much more likely to be perceived to suffer from hyperactivity or behavioral problems (*P*=.01).

Conclusions: Parental awareness is necessary to counteract the harmful effects of using electronic devices for a prolonged period. Parents require help to cope with this problem effectively.

Introduction

The continuous process of digitization of the world has led to an escalating use of various electronic entertainment and communication devices, including mobile phones, tablets, computers, and television (TV) [1,2]. A study in 2015 found that children who spent ≥2 hours daily on electronic devices were more likely to demonstrate an increased body mass index (BMI) [2]. A 2013 study compared the use of mobile phones among children in 4 different Arab countries and determined that around 80% of children aged between 8 and 18 years used mobile phones in Saudi Arabia. The study also ranked Saudi Arabia second after Egypt (87%) in high proportion of mobile phone use among children [3]. However, there is a scarcity of

XSL•FO
RenderX

data about the use of electronic entertainment and communication devices among children and parental impression of behavior in Saudi Arabia.

The impact of screen addiction on health has been the subject of many studies. The results of these studies showed that excessive use of electronic devices, especially mobile phones and tablets had a harmful impact not only on the eyesight but also on the psycho-affective and cognitive development of the child. Moreover, the associated inactivity could trigger the development of some diseases, such as obesity and hypertension, at early ages [4-6]. TV viewing is the most widely studied behavior across the globe. This habit is, to some extent, correlated with childhood obesity [7,8]. A study conducted among 197 school children with a mean age of 8.7 (SD 2.1) years in the United Arab Emirates found that TV viewing and video game use for >2 hours per day was associated with behavioral problems such as being withdrawn, inattentive, and having externalizing behaviors, including hyperactivity and aggressive behaviors [9]. Higher aggression levels have also been observed in children watching TV programs or playing video games, especially with violent content [9]. On the contrary, an intervention study on elementary school students showed a reduction in aggressive behavior on watching TV for lesser hours [9]. Besides, excessive use of electronic devices could adversely affect a child's sleep-wake cycle, resulting in further behavioral problems and occasionally stereotypical motor behaviors [1,10,11]. This is particularly important given the extent of use of such devices in our community and their potential negative impacts on children's well-being. Parental awareness about the harmful effects of electronic and communication devices on children is tremendously important for preventing their occurrence [12,13].

The objective of this study was to assess the level of use of electronic devices among children attending a health care facility in Saudi Arabia and the parents' perspectives about the impact of screen use on the behavior of their children.

Methods

A cross-sectional survey involving mothers of children attending the Well Baby Clinics of King Abdulaziz University Hospital (KAUH), Jeddah, Saudi Arabia was conducted from the 2016 July 1 to 2016 November 30. KAUH is the main university teaching hospital of the Western Region of Saudi Arabia. It is a large multi-specialty adult and pediatric (children aged 8-15 years) hospital that also provides primary medical care and immunization for all children born at the hospital. Three Well Baby Clinics per week provide parental support, education, and routine immunizations.

For this survey, 190 participants were chosen by convenience sampling and all 190 mothers responded to the questionnaires, thus providing a response rate of 100%. Inclusion criteria were as follows: mothers with one or more child who used electronic entertainment devices and was willing to participate in this survey. No exclusion criteria were set.

Because we could not find any validated questionnaires about the impact of electronic devices use on children behavior, we developed a focused 15-item questionnaire to examine the use of various electronic entertainment and communication devices. The questionnaire also assessed parental supervision and interaction. The questionnaire was not validated. The first part of the questionnaire collected sociodemographic characteristics such as the age of the child and the parents, the parents' education, occupation, family income, and nationality. The second part was relevant to the type of electronic devices used by the child, the time spent by children on those devices, and whether the parents noticed any change in the behavior of their children while using these devices (see Multimedia Appendix 1).

The study and questionnaire were explained to the mothers before they voluntarily participated in the study. The identity of the participants was not requested to ensure privacy and encourage accurate responses. The bioethical committee of KAUH approved the study design and questionnaire.

Participants filled a paper-pencil survey; later, we entered the data into an Excel sheet. Statistical analyses were performed using the Statistical Package for Social Sciences version 20 (SPSS Inc, Chicago, IL, USA). Frequencies and percentages were used for the descriptive analyses. Association between categorical variables was examined using chi-square test. Statistical significance was defined as P values of <.05.

Results

This study included 190 mothers. The mean ages of the children, mothers, and fathers were 7.3 (SD 3.5), 35 (SD 6.5), and 43 (SD 8.3) years, respectively. Most children were of Saudi Arabian nationality (106/190, 55.8%). Most mothers were housewives (145/190, 76.3%), and 40.5% (77/190) of the mothers in our study had school-level education. Most fathers were employed (135/190, 71.0%), and 53.1% (101/190) of the fathers in our study had a university-level education. Of the families in the study, 42.1% (80/190) had a monthly income of <5000 SAR (<US $1300; Table 1).

The most used device in this study was TV (154/190, 81.0%), followed by mobile phones (134/190, 70.5%) and tablets (116/190, 61.0%). Computers (59/190, 31.0%) were the least used device in this study. In total, 24.7% (47/190) of children in this study used electronic entertainment and communication devices for than 4 hours (Table 2).

When asked if the parents would limit the time spent on these devices, about 63.1% (120/190) answered "yes" (16% responded "sometimes" and 47% responded "always"). Most mothers (129/190, 67.9%) felt that their child spends too much time on electronic devices. Hyperactivity or behavioral problems were reported by 20.0% (38/190) of mothers in this study. Overall, 10.0% (19/190) of the mothers "sometimes" observed hyperactivity or behavioral problems in their children; similarly, 10.0% (19/190) of the mothers "always" observed hyperactivity or behavioral problems in their children (Table 3).

Table 1. Sociodemographic characteristics of the study families (N=190).

Characteristics	Value
Age (years), mean (SD), range	
Child	7.3 (3.5), 1-16
Mother	35 (6.5), 24-50
Father	43 (8.3), 28-63
Nationality of child, n (%)	
Saudi Arabia	106 (55.8)
Yemen	19 (10.0)
Sudan	17 (8.9)
Other nationalities	48 (25.3)
Mother's employment, n (%)	
Employed	45 (23.7)
Unemployed	145 (76.3)
Father's employment, n (%)	
Employed	135 (71.0)
Unemployed	55 (28.9)
Mother's educational level, n (%)	
None	30 (15.8)
School grade	83 (43.7)
College or university	77 (40.5)
Father's educational level, n (%)	
None	29 (15.3)
School grade	60 (31.6)
College or university	101 (53.1)
Monthly income of family, n (%)	
<5000 SAR (<US $1333)	80 (42.1)
5000-10,000 SAR (US $1333-2666)	57 (30.0)
10,000 SAR (>US $2666)	53 (27.9)

Table 2. Most frequently used devices and overall time spent on electronic entertainment and communication devices (N=190).

Most used device	n (%)
Television	
Yes	154 (81.0)
No	36 (18.9)
Mobile phones	
Yes	134 (70.5)
No	56 (29.5)
Tablets	
Yes	116 (61.0)
No	74 (38.9)
Computers/laptops	
Yes	59 (31.0)
No	131 (68.9)
Overall time spent on electronic entertainment and communication devices	
1 hr per day	61 (32.1)
2-4 hr per day	82 (43.1)
4-6 hr per day	25 (13.1)
>6 hr per day	22 (11.6)

Table 3. Parent-child interaction and behavioral impression (N=190).

Behavioral impression and parent-child interaction	n (%)
Do you try to limit the time your child spends on these devices?	
Never	70 (36.8)
Sometimes	30 (15.8)
Always	90 (47.4)
Do you think your child spends too much time on these devices?	
Yes	129 (67.9)
No	61 (32.1)
Does your child suffer from hyperactivity or behavior problems?	
Yes	38 (20.0)
No	152 (80.0)
How often do you observe hyperactivity or behavior problems problem in your child?	
Never	152 (80.0)
Sometimes	19 (10.0)
Always	19 (10.0)

Time spent on electronic devices correlated significantly with parental impression that their child was suffering from behavioral problems. In total, 8.4% (16/190) of mothers in our study whose children spent >2 hours per day on electronic devices reported that their children "always" exhibited behavioral problems. On the other hand, only 1.6% (3/190) mothers whose children spent 1 hour per day on electronic devices reported that their children "always" had behavioral problems. Children who spent longer hours on electronic devices were much more likely to be perceived to suffer from hyperactivity or behavioral problems (P=.01; Table 4).

Table 4. Association between time spent on electronic devices and parental impression that the child suffered from hyperactivity or behavioral problems (N=190; *P*=.01).

Time spent on electronic devices (hr per day)	Hyperactivity or behavioral problems, n (%)		
	None (n=152)	Sometimes (n=19)	Always (n=19)
1 (n=61)	54 (88)	4 (7)	3 (5)
2-4 (n=82)	67 (82)	8 (10)	7 (8)
4-6 (n=25)	16 (64)	6 (24)	3 (12)
>6 (n=22)	15 (68)	1 (5)	6 (27)

Discussion

Principal Findings

Our study confirmed that many families living in Saudi Arabia were affected by their children's excessive use of various electronic entertainment and communication devices. When asked if the parents would limit the time spent on these devices, about 63.1% (120/190) answered "yes" and mothers (129/190, 67.9%) felt that their child spends too much time on electronic devices. In fact, in 24.7% (47/190) of children, the time spent on electronic devices was >4 hours per day. After TV viewing, mobile phones were the second most frequently used device. Previous research identified that less time spent on various electronic and communication devices resulted in prevention of depression in young adults [14], thereby potentially improving health during adulthood. In addition, other investigators found correlations with several future neurological disorders, such as recurrent headaches and migraine [15].

One out of 5 mothers (38/190, 20.0%) thought their child was experiencing hyperactivity or behavioral problems. This was queried in a direct manner and not examined in a structured way or through the use of a standardized diagnostic inventory. It was simply based on the mother's overall impression of her child's behavior. The accuracy of such global impressions has been proven to be effective in other studies [16]. Watching TV, which was most prevalent (154/190, 81.0%) in our sample, was previously noted by other investigators who found no association between watching TV and emotional symptoms, hyperactivity, inattention, or peer relationship problems [17]. However, previous research found an association between watching TV for >2 hours per day and antisocial behaviors, aggressive personalities, and criminal convictions [18]. This might be explained by cultural differences or inclusion of older children and adolescents in those studies.

Limitations

This study had several limitations that warrant discussions. Firstly, our sample size was relatively small. A larger study could provide more insight about potential correlations with sociodemographic variables such as parents' ages, educational levels, family incomes, and nationalities. Secondly, because we used convenience sampling from one region in Saudi Arabia (Jeddah), our findings might not be representative of other regions of Saudi Arabia. Moreover, the study participants comprised working-class people with average educational and economic levels. This, in turn, limited our ability to generalize our findings to the population at large. Hence, more detailed research, involving different regions of the country and different groups of people according to educational qualifications, income levels, and origins, should be conducted to attain a generalizable conclusion. Thirdly, our questionnaire was not used or validated previously, mainly owing to the limited availability of relevant literature. The questionnaire did not include specific behavioral problems that were observed. Therefore, the use of structured behavioral inventories is needed to confirm the link between hyperactivity or behavioral disorders and excessive use of different electronic entertainment and communication devices. Another possible limitation of this study, like most other cross-sectional studies, was self-reporting. This rendered the information obtained regarding the time spent by the children on electronic devices questionable [2]. Additionally, the information was collected from the mothers only, rather than the children themselves and there might be a probability that the mothers exaggerated their responses. Even though several previous population-based studies utilized mother-reported screen time, there are concerns regarding the validity and reliability of these reports [17]. Furthermore, this study did not assess the content of the electronic and communication devices. Earlier studies revealed that aggressive content might be responsible for psychological or behavioral problems in children, regardless of how much time was spent watching TV [6]. However, we did not assess whether the screen contents were age-appropriate or whether they negatively impacted the children's behaviors. Therefore, it is essential to survey the contents of the electronic and communication devices to refine the results of the study. Despite these limitations, the major strength of this study was that this was the first study that attempted to survey children in Jeddah, Saudi Arabia to determine how much time they spent on electronic entertainment and communication devices.

Conclusions

In conclusion, this study detected excessive use (2-6 hours or more per day) of electronic devices by the children in the study sample. The time spent is higher than the recommended upper limit of 2 hours per day according to American Academy of Pediatrics (AAP). In fact, using communication devices >2 hours per day is linked with significant behavioral problems [9]. Accordingly, the parents in this study reported increased behavioral problems or hyperactivity among children who spent more time on electronic and communication devices. Our results indicate that the parents should follow the recommendations of AAP to limit their children's exposure to electronic devices to no more than >2 hours per day, potentially reducing the risk of developing behavioral abnormalities [19]. Through increased

parental awareness, the possible detrimental effects of technology on children's psychological development and behaviors could be counteracted [9]. This in turn might help to reduce the overall time spent on electronic devices [6]. Nonetheless, to deal with this issue efficiently, parents require help. Public educational campaigns are imperative for addressing this growing problem and providing effective solutions. Finally, slower-paced, nonviolent educational programs are believed to have less negative impact on children's behaviors [19]. Therefore, besides monitoring screen time, parents should monitor screen content. Further studies must be attempted to detect the influence of the screen content on children's behavioral patterns.

Conflicts of Interest

None declared

Multimedia Appendix 1

Questionnaire.

[PDF File (Adobe PDF File), 21KB - ijmr_v7i2e13_app1.pdf]

References

1. Chahal H, Fung C, Kuhle S, Veugelers PJ. Availability and night-time use of electronic entertainment and communication devices are associated with short sleep duration and obesity among Canadian children. Pediatr Obes 2013 Feb;8(1):42-51. [doi: 10.1111/j.2047-6310.2012.00085.x] [Medline: 22962067]
2. Al-Agha AE, Nizar FS, Nahhas AM. The association between body mass index and duration spent on electronic devices in children and adolescents in Western Saudi Arabia. Saudi Med J 2016 Apr;37(4):436-439 [FREE Full text] [doi: 10.15537/smj.2016.4.15018] [Medline: 27052287]
3. Lynch S. Children's Use of Mobile Phones – An international Comparison 2013 URL: https://www.gsma.com/publicpolicy/wp-content/uploads/2012/03/GSMA_ChildrensMobilePhones2013WEB.pdf [accessed 2018-08-01] [WebCite Cache ID 71L2pK9MA]
4. Cash H, Rae C, Steel A, Winkler A. Internet Addiction: A Brief Summary of Research and Practice. Curr Psychiatry Rev 2012 Nov;8(4):292-298 [FREE Full text] [doi: 10.2174/157340012803520513] [Medline: 23125561]
5. Pantic I. Online social networking and mental health. Cyberpsychol Behav Soc Netw 2014 Oct;17(10):652-657 [FREE Full text] [doi: 10.1089/cyber.2014.0070] [Medline: 25192305]
6. Rosen LD, Lim A, Felt J, Carrier LM, Cheever NA, Lara-Ruiz JM, et al. Media and technology use predicts ill-being among children, preteens and teenagers independent of the negative health impacts of exercise and eating habits. Comput Human Behav 2014 Jun;35:364-375 [FREE Full text] [doi: 10.1016/j.chb.2014.01.036] [Medline: 25717216]
7. Al Shehri A, Al Fattani A, Al Alwan I. Obesity among Saudi children. Saudi J Obesity 2013;1(1):3-9. [doi: 10.4103/2347-2618.119467]
8. Council on CommunicationsMedia, Strasburger VC. Children, adolescents, obesity, and the media. Pediatrics 2011 Jul;128(1):201-208. [doi: 10.1542/peds.2011-1066] [Medline: 21708800]
9. Yousef S, Eapen V, Zoubeidi T, Mabrouk A. Behavioral correlation with television watching and videogame playing among children in the United Arab Emirates. Int J Psychiatry Clin Pract 2014 Aug;18(3):203-207. [doi: 10.3109/13651501.2013.874442] [Medline: 24329399]
10. Subki AH, Alsallum MS, Alnefaie MN, Alkahtani AM, Almagamsi SA, Alshehri ZS, et al. Pediatric motor stereotypies: an updated review. J Pediatr Neurol 2017;15:156. [doi: 10.1055/s-0037-1602256]
11. Kondo Y, Tanabe T, Kobayashi-Miura M, Amano H, Yamaguchi N, Kamura M, et al. Association between feeling upon awakening and use of information technology devices in Japanese children. J Epidemiol 2012;22(1):12-20 [FREE Full text] [Medline: 22041529]
12. Hawi NS, Rupert M. Impact of e-Discipline on Children's Screen Time. Cyberpsychol Behav Soc Netw 2015 Jun;18(6):337-342 [FREE Full text] [doi: 10.1089/cyber.2014.0608] [Medline: 26075921]
13. Canadian Paediatric Society. Impact of media use on children and youth. Paediatr Child Health 2003 May;8(5):301-317 [FREE Full text] [Medline: 20020034]
14. Grøntved A, Singhammer J, Froberg K, Møller NC, Pan A, Pfeiffer K, et al. A prospective study of screen time in adolescence and depression symptoms in young adulthood. Prev Med 2015 Dec;81:108-113. [doi: 10.1016/j.ypmed.2015.08.009] [Medline: 26303369]
15. Chiu CT, Chang YH, Chen CC, Ko MC, Li CY. Mobile phone use and health symptoms in children. J Formos Med Assoc 2015 Jul;114(7):598-604 [FREE Full text] [doi: 10.1016/j.jfma.2014.07.002] [Medline: 25115529]
16. Tasgin EC, Oner O, Yurtbasi P, Munir K. Effects of maternal symptom ratings and other clinical features on short-term treatment response to OROS methylphenidate in children and adolescents with ADHD in a naturalistic clinical setting.

Klinik Psikofarmakol Bulteni 2016;26(2):126-133 [FREE Full text] [doi: 10.5455/bcp.20150703013708] [Medline: 27746700]

17. Parkes A, Sweeting H, Wight D, Henderson M. Do television and electronic games predict children's psychosocial adjustment? Longitudinal research using the UK Millennium Cohort Study. Arch Dis Child 2013 May;98(5):341-348 [FREE Full text] [doi: 10.1136/archdischild-2011-301508] [Medline: 23529828]

18. Robertson LA, McAnally H, Hancox R. Childhood and adolescent television viewing and antisocial behavior in early adulthood. Pediatrics 2013 Mar;131(3):439-446 [FREE Full text] [doi: 10.1542/peds.2012-1582] [Medline: 23420910]

19. Swing EL, Gentile DA, Anderson C, Walsh D. Television and video game exposure and the development of attention problems. Pediatrics 2010 Aug;126(2):214-221. [doi: 10.1542/peds.2009-1508] [Medline: 20603258]

Abbreviations

AAP: American Academy of Pediatrics
BMI: body mass index
KAUH: King Abdulaziz University Hospital
TV: television

XSL•FO
RenderX

Original Paper

Basic Life Support Knowledge Among a Nonmedical Population in Jeddah, Saudi Arabia: Cross-Sectional Study

Ahmed Hussein Subki[1], MBBS; Hatan Hisham Mortada[1], MBBS; Mohammed Saad Alsallum[1], MBBS; Ali Taleb Alattas[1], MBBS; Mohammed Ali Almalki[1], MBBS; Muhab Mohammed Hindi[1], MBBS; Siham Hussein Subki[1]; Wesam Awad Alhejily[1], MD, FRCPC

Department of Internal Medicine, Faculty of Medicine, King Abdulaziz University, Jeddah, Saudi Arabia

Corresponding Author:
Ahmed Hussein Subki, MBBS
Department of Internal Medicine
Faculty of Medicine
King Abdulaziz University
Al Sulaymaniyah
Jeddah,
Saudi Arabia
Phone: 966 560662735
Email: ahs.subki@gmail.com

Abstract

Background: Providing basic life support (BLS) at the site of an accident is crucial to increase the survival rates of the injured people. It is especially relevant when health care is far away.

Objective: The aim of our study is to assess the BLS knowledge level of the Saudi Arabian population and identify influencing factors associated with level of knowledge about BLS.

Methods: Our study is a cross-sectional descriptive study, which was conducted using a self-administered online questionnaire derived from the BLS practice test. The Saudi population was the target population. The questionnaire was divided into two parts: one contained demographic data and the second part contained questions to test the population's perception about how to perform BLS techniques properly. The data were collected between July and August 2017. Statistically significant differences were defined as those with a P value $<.05$, and a score of five or more was considered a passing score on the second part. We used SPSS version 21 for data analysis.

Results: Our study included 301 participants. Our participants' BLS online exam scores ranged from 0 to 10, with a mean of 4.1 (SD 1.7). Only 39.2% (118/301) of the participants passed the test. The percentage of bachelor's degree or higher holders constituted 60.1% (181/301) of the study population. In addition, higher income was significantly associated with higher scores on the test ($P=.04$).

Conclusions: This study demonstrated that the theoretical knowledge level of BLS among the general population in Jeddah was below average. There is a critical need to increase the public's exposure to BLS education through raising awareness campaigns and government-funded training programs that aim to curb the incidence of out-of-hospital cardiac arrest mortalities in the Saudi community.

Introduction

Life-threatening emergencies can occur anytime and anywhere. The lack of training and incompetence to deal with these emergencies can have tragic and legal consequences. Basic resuscitation skills, including prompt and effective cardiopulmonary resuscitation (CPR), increases the survival rate following cardiopulmonary arrest. Theoretical knowledge with practical demonstrations and regular practice with up-to-date recommendations is important in maintaining the capability of basic life support (BLS) and advanced life support (ALS) providers [1]. Saving peoples' lives involves a sequence

of steps that constitute the chain of survival. This includes four stages: early approach to cohesive medical emergency, early initiation of BLS, early defibrillation, and ALS [2].

Sudden cardiac arrest is the most common cause of death worldwide with a large variation in survival rates between different communities [3,4]. Early detection of cardiac arrest and initiation of CPR have been shown to decrease morbidity and mortality [4-6]. A previous study conducted in Arizona showed that a statewide hands-only CPR public awareness campaign increased bystander CPR rates from 28.2% to 39.9% and improved out-of-hospital cardiac arrest (OHCA) survival rates from 3.7% to 9.8% [7]. The number of OHCA cases in the United States is approximately 300,000 and the mortality rate is 92% [8]. The chance of survival increases two-fold if BLS is conducted by the first person to intervene and with the use of automated external defibrillators [8].

Difficulties performing bystander CPR in more developed countries were due to inadequate knowledge or training, absence of skill, lack of confidence, and fear of litigation [9]. In order to provide proper care, the community needs an adequate amount of knowledge and insight on BLS by training either in person or online [10]. An important method to increase CPR and BLS success for any emergency case is to increase public knowledge and understanding of the practical applications of BLS intervention for a successful result after an emergency. An assessment of community knowledge and perceptions of BLS has not been performed in Saudi Arabia.

Our study aims to assess the BLS knowledge level of a Saudi population and to identify demographic factors such as age, gender, nationality, marital status, educational level, and monthly family income related to the level of knowledge about BLS.

Methods

Study Design and Data Collection

Our study is a cross-sectional descriptive study that was designed to assess the knowledge and perception about BLS among a Saudi population via an online questionnaire. The data were collected between July and August 2017 in Jeddah, Saudi Arabia.

Measures

Our questionnaire was developed by BLS provider National Health Care Provider Solutions, who provided permission to use the questionnaire [11]. The questionnaire adhered to the latest American Heart Association and Emergency Cardiovascular Care guidelines [12].

Our questionnaire was an online, multiple-choice, self-administered questionnaire in two parts. The first part focused on the participants' demographic data, including age, nationality, gender, marital status, education, household income, place of residence, and district. The second part was about the responders' perceptions about BLS, including the ratio for one rescuer giving CPR, how often should rescuers switch roles

when performing two-rescuer CPR, the proper steps for operating an automated external defibrillator, site to check a pulse in a child whose age is between 1 year to puberty, the initial BLS steps in adults, critical characteristics of high-quality CPR, the five steps in the adult chain of survival, the recommended BLS sequence of steps, signs of airway obstruction, and how often the breaths are administered in an adult with an advanced airway in place during two-rescuer CPR (see Multimedia Appendix 1 for the questionnaire). A score of five or more was considered a passing score.

The survey was created initially in English, then translated and back-translated into the Arabic language by two independent translators. Any discrepancy was resolved through consensus.

The inclusion criteria included all males and females in the Saudi population. Participants who refused to participate or had a medical background were excluded.

A link to the questionnaire was distributed to the participants through Facebook and WhatsApp groups. They were informed on the first page of the questionnaire about the study with a disclaimer on it that only those people without a medical background could fill it; if they agreed, they could continue. The definition of "medical background" was left up to the participants to decide, but it included anyone who works or had worked in the health care system or was student in medical, nursing, dental schools, and any health care-related colleges. Basic health knowledge taught in school was not included in our definition. Ethical approval was obtained from the Unit of Biomedical Ethics, Research Committee at King Abdul Aziz University, before data collection started.

Statistical Analysis

Data were presented as a mean for central tendency with standard deviation for variance. The chi-square test was used to compare different variables. Statistically significant differences were defined as those with a $P<.05$. We used SPSS version 21 for data analysis.

Results

Our study included 301 responders. Of these, 278 (92.4%) were Saudi Arabian. The response rate was not possible to determine, but this was due to the online nature of the questionnaire. If they did not want to participate they could simply discard the link. The mean age for participants was 24.1 (SD 8.9) years. More than two-thirds of participants were younger than 25 years of age (227/301, 75.4%). The majority of responders were also female (217/301, 72.1%) and single (240/301, 79.7%). Of the respondents, more than half held a bachelor's degree (174/301, 57.8%) and one-third had a total family income between 5000 SR to 10,000 SR. Most of the participants (245/301, 81.4%) lived in Jeddah, Saudi Arabia (see Table 1 for full demographic details of participants. Our participants' BLS online exam scores ranged from 0 to 10, with a mean of 4.08 (SD 1.71). Only 39.2% (118/301) of participants passed the test. The percentage of participants who passed the exam was compared across various factors (Table 2) using the chi-square test.

Table 1. Characteristics of participants (N=301).

Variable	n (%)
Age[a]	
<25 years	225 (75.3)
≥25 years	74 (24.7)
Nationality	
Saudi	278 (92.4)
Non-Saudi	23 (7.6)
Sex	
Male	84 (27.9)
Female	217 (72.1)
Marital status	
Single	240 (79.7)
Married	60 (19.9)
Divorced	1 (0.3)
Educational level	
Illiterate	2 (0.7)
High school or less	106 (35.2)
Professional certificate (noncollege issued)	11 (3.7)
Bachelor	174 (57.8)
Masters or PhD	8 (2.7)
Monthly income (SR)	
<5000	34 (11.3)
5000 to 10,000	100 (33.2)
10,000 to 20,000	98 (32.6)
>20,000	69 (22.9)
Residence	
Jeddah	245 (81.4)
Outside Jeddah	56 (18.6)
Area	
South	63 (20.9)
East	29 (9.6)
West	20 (6.6)
North	105 (34.9)
Center	38 (12.6)
Outside	46 (15.3)
Success	
Failed	183 (60.8)
Passed	118 (39.2)

[a]n=299. Our respondents completed all variables, except for 2 participants who did not fill the age variable. They were not excluded from the analysis.

There was no significant difference in the percentage of success between the following groups: age group younger than 25 years and 25 years or older; Saudis and non-Saudis; males and females; single, married, and divorced; different degrees of education; college educated or not; different levels of income; living in Jeddah or outside Jeddah; and the area living within Jeddah. There was no significant difference in the exam success rates among any of these groups (all P>.05).

Table 2. Success rate differences among different variables.

Variable	Success, n (%)		P value
	Failed	Passed	
Age group			.09
<25 years	130 (57.8)	95 (42.2)	
≥25 years	51 (68.9)	23 (31.1)	
Nationality			.38
Saudi	171 (61.5)	107 (38.5)	
Non-Saudi	12 (52.2)	11 (47.8)	
Sex			.19
Male	56 (66.7)	28 (33.3)	
Female	127 (58.5)	90 (41.5)	
Marital status			.41
Single	142 (59.2)	98 (40.8)	
Married	40 (66.7)	20 (33.3)	
Divorced	1 (100.0)	0 (0.0)	
Educational level			.21
Illiterate	1 (50.0)	1 (50.0)	
High school or less	67 (63.2)	39 (36.8)	
Professional certificate (noncollege issued)	9 (81.8)	2 (18.2)	
Bachelor	99 (56.9)	75 (43.1)	
Masters or PhD	7 (87.5)	1 (12.5)	
Monthly income (SR)			.25
less than 5000	26 (76.5)	8 (23.5)	
5000 to 10,000	58 (58.0)	42 (42.0)	
10,000 to 20,000	59 (60.2)	39 (39.8)	
>20,000	40 (58.0)	29 (42.0)	
Residence			.36
Jeddah	152 (62.0)	93 (38.0)	
Outside Jeddah	31 (55.4)	25 (44.6)	
Area			.85
South	42 (66.7)	21 (33.3)	
East	17 (58.6)	12 (41.4)	
West	12 (60.0)	8 (40.0)	
North	65 (61.9)	40 (38.1)	
Center	22 (57.9)	16 (42.1)	
Outside	25 (54.3)	21 (45.7)	
College			.26
Not college educated	77 (64.7)	42 (35.3)	
College educated	106 (58.2)	76 (41.8)	

Pearson correlation was done (Table 3) to test for significance between the total score and age and it showed a significant correlation between (P=.04). It was a very weak negative correlation (r=−.12), meaning the test total score tended to be lower if the participant was of older age. Spearman correlation was done between the total score and the educational level as well as the total score and income level, but there was no significant correlation in either case.

Table 3. Correlations between the total score and age, income level, and educational level (N=301).

Total score versus	Correlation coefficient	P value
Age[a]	–.119	.04
Income level	.066	.26
Educational level	.052	.37

[a]n=299. Our respondents completed all variables, except for 2 participants who did not fill the age variable. They were not excluded from the analysis.

Table 4. Distribution of college-educated participants according to different variables.

Variable	College education, n (%)		P value
	Not college educated	College educated	
Age group			<.001
<25 years	103 (45.8)	122 (54.2)	
≥25 years	15 (20.3)	59 (79.7)	
Nationality			.07
Saudi	114 (41.0)	164 (59.0)	
Non-Saudi	5 (21.7)	18 (78.3)	
Sex			.40
Male	30 (35.7)	54 (64.3)	
Female	89 (41.0)	128 (59.0)	
Marital status			<.001
Single	108 (45.0)	132 (55.0)	
Married	10 (16.7)	50 (83.3)	
Divorced	1 (100.0)	0 (0.0)	
Monthly income (SR)			.07
<5000	16 (47.1)	18 (52.9)	
5000 to 10,000	29 (29.0)	71 (71.0)	
10,000 to 20,000	44 (44.9)	54 (55.1)	
>20,000	30 (43.5)	39 (56.5)	
Residence			.97
Jeddah	97 (39.6)	148 (60.4)	
Outside Jeddah	22 (39.3)	34 (60.7)	
Area			.40
South	19 (30.2)	44 (69.8)	
East	9 (31.0)	20 (69.0)	
West	8 (40.0)	12 (60.0)	
North	44 (41.9)	61 (58.1)	
Center	18 (47.4)	20 (52.6)	
Outside	21 (45.7)	25 (54.3)	

Being college educated or not was tested among different factors (Table 4). It was only significant with the age group (P<.001), and the percentage college educated was found to be higher in the age group older than 25 years. It was also significant for marital status (P<.001), where college education was higher in percentage in married participants than in singles. There was no significant difference regarding college education for the other factors, including nationality, gender, income level, residence in Jeddah, and area of living within Jeddah.

Table 5. Binary logistic regression result for the score (passed/failed).

Variables	Odds ratio (95% CI)	P value
Nationality		
Saudi	0.499 (0.196-1.269)	.14
Non-Saudi (ref[a])		
Sex		
Male	0.732 (0.409-1.310)	.29
Female (ref)		
Marital status		
Single (ref)		.95
Married	1.163 (0.470-2.877)	.74
Divorced	0.000 (0.000)	>.99
Monthly income (SR)		
<5000 (ref)		.19
5000 to 10,000	2.631 (1.033-6.704)	.04
10,000 to 20,000	2.632 (1.023-6.772)	.05
>20,000	2.773 (1.045-7.355)	.04
Residence		
Jeddah	0.611 (0.327-1.143)	.12
Outside Jeddah (ref)		
Age group		
<25 years	1.933 (0.814-4.587)	.14
≥25 years (ref)		
College education		
College educated	0.724 (0.433-1.210)	.22
Noncollege educated (ref)		
Constant	0.554	.42

[a]Ref: reference.

Binary logistic regression was done to test for the effect of various factors on passing the test (Table 5). The Nagelkerke R^2 for the model was .064, meaning that the model could predict only 6.4% of cases. The only significant odds ratio was in the income groups compared to the group with income less than 5000 SR. The odds ratio for the income between 5000 and 10,000 was 2.63 (95% CI 1.03-6.70), for the group 10,000 SR to 20,000 SR was OR 2.63 (95% CI 1.02-6.77), and for the group with income greater than 20,000 SR, the odds ratio was 2.77 (95% CI 1.04-7.355), meaning that participants with income greater than 5000 SR had a higher chance of passing the exam.

Discussion

More than two-thirds of the study sample were female (72.1%), which is much higher than the latest census [13]. and even most reported studies [14-16]. The majority of this population (75.3%) were young adults younger than 25 years. The percentage of bachelor's degrees or higher holders was more than 60%; this

may be attributed to the samples being collected in Jeddah, a city with three major universities. Despite this, the sample provided us with insight into a young, educated group, who had more time to gain BLS and CPR knowledge in university as evident by another study done in Riyadh, Saudi Arabia, where the largest number of participants with any CPR knowledge gained it while they were enrolled in university [17].

The number of participants who passed the test was only 118 (39.2%), which is higher than a similar study done in Hong Kong [16], but still less than other studies done in Poland, the United States, and Turkey [18-20]. Even with the higher scores, it does not necessarily mean the ones who passed had any previous BLS training. This can be explained partially by the possibility that some people do have general knowledge without attending BLS courses through courses in school and mainstream media. The American Heart Association has stated that to reduce both morbidity and mortality from OHCA in a significant manner, 20% of adults need to have CPR training [21].

Higher income (more than 5000 SR/month) was associated with higher chance of passing the exam. This can be explained alongside the high education levels: the higher the education level was, the more proportionately higher was the income [22]. However, both income and education were not correlated with higher total score. Age, on the other hand, was negatively correlated with the total BLS score, indicating that younger age was associated with a higher score. Although, it is weakly related, it is nonetheless significant. This may be attributed to forgetting the theoretical part of BLS, especially if more years have passed since they have been exposed to this knowledge. Some hospitals recommend taking refresher courses at least every 3 years [16]. This is due to the diligent aspects of resuscitation and that mistakes in any of its steps may compromise the whole process.

Numerous studies have demonstrated the effects of incorporating BLS support measures by trained regular citizens/laypersons, reducing both the rates of mortality and morbidity [23-25]. Individuals receiving CPR from trained citizens have been found to be four times more likely to live and survive for a month in comparison to those who did not receive CPR [26]. Increasing BLS training centers should be essential to improve the survival rates from OHCA. It should even include schools because it provides an opportunity to expose as much of the population as possible to resuscitation techniques. One region in the United Kingdom introduced its schools to CPR training and it was well received with more than 99% of its students agreeing to its benefit to them [27].

One of the limitations of our study is the small number and scope due to including just one city (Jeddah) in Saudi Arabia, thus the study may not represent the whole population of Jeddah. In comparison to the total Saudi population, our study contained more females and had more participants with a higher education, with generally a younger age group in comparison to the total Saudi population [13]. Another limitation of our study was the inherit bias in how participants decided whether they had a medical background. In addition, it was not possible to determine the response rate of the participants due to using social media sites to distribute the questionnaire. Also, a response was only registered after full completion of the questionnaire, so any person who stopped midway could not be accounted for. We propose more studies to be done that are more comprehensive and include the entire region of Saudi Arabia. Another limitation of this study is that it only tested the participant's theoretical knowledge; testing it practically was not viable for our study population. Lastly, there was no way to ascertain if any of the population had prior BLS training or if they ever practiced any BLS measures in a real-life setting.

This study demonstrated that the theoretical knowledge level of BLS among the general population in Jeddah was below average. There is a need to increase the public's exposure to BLS education by raising awareness campaigns and government-funded training programs that aim to curb the incidence of OHCA in the Saudi community.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Questionnaire used in study.
[PDF File (Adobe PDF File), 30 KB - ijmr_v7i2e10428_app1.pdf]

References

1.   Kharsan V, Singh R, Madan R, Mahobia Y, Agrawal A. The ability of oral & maxillofacial surgeons to perform basic life resuscitation in chattisgarh. J Clin Diagn Res 2015 Feb;9(2):ZC58-ZC60 [FREE Full text] [doi: 10.7860/JCDR/2015/12630.5603] [Medline: 25866738]
2.   Hunter PL. Cardiac arrest in the dental surgery. Br Dent J 1991 Apr 20;170(8):284. [Medline: 2036267]
3.   Lombardi G, Gallagher J, Gennis P. Outcome of out-of-hospital cardiac arrest in New York City. The Pre-Hospital Arrest Survival Evaluation (PHASE) Study. JAMA 1994 Mar 02;271(9):678-683. [Medline: 8309030]
4.   Sasson C, Rogers M, Dahl J, Kellermann A. Predictors of survival from out-of-hospital cardiac arrest: a systematic review and meta-analysis. Circ Cardiovasc Qual Outcomes 2010 Jan;3(1):63-81. [doi: 10.1161/CIRCOUTCOMES.109.889576] [Medline: 20123673]
5.   Rajapakse R, Noč M, Kersnik J. Public knowledge of cardiopulmonary resuscitation in Republic of Slovenia. Wien Klin Wochenschr 2010 Dec;122(23-24):667-672. [doi: 10.1007/s00508-010-1489-8] [Medline: 21132393]
6.   Bobrow BJ, Spaite DW, Berg RA, Stolz U, Sanders AB, Kern KB, et al. Chest compression-only CPR by lay rescuers and survival from out-of-hospital cardiac arrest. JAMA 2010 Oct 06;304(13):1447-1454. [doi: 10.1001/jama.2010.1392] [Medline: 20924010]

7.   Roger V, Go A, Lloyd-Jones D, Adams R, Berry J, Brown T, American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics--2011 update: a report from the American Heart Association. Circulation 2011 Feb 01;123(4):e18-e209 [FREE Full text] [doi: 10.1161/CIR.0b013e3182009701] [Medline: 21160056]

8.   Berdowski J, Blom MT, Bardai A, Tan HL, Tijssen JG, Koster RW. Impact of onsite or dispatched automated external defibrillator use on survival after out-of-hospital cardiac arrest. Circulation 2011 Nov 15;124(20):2225-2232. [doi: 10.1161/CIRCULATIONAHA.110.015545] [Medline: 22007075]

9.   Mori S, Whitaker IY, Marin HD. [Evaluation of an educational website on First Aid]. Rev Esc Enferm USP 2013 Aug;47(4):950-957 [FREE Full text] [doi: 10.1590/S0080-623420130000400025] [Medline: 24310695]

10. Arbon P, Hayes J, Woodman R. First aid and harm minimization for victims of road trauma: a population study. Prehosp Disaster Med 2011 Aug;26(4):276-282. [doi: 10.1017/S1049023X11006522] [Medline: 22008363]

11. National Health Care Provider Solutions. BLS practice test URL: https://nhcps.com/bls-practice-quiz/

12. American Heart Association. 2005 American Heart Association (AHA) guidelines for cardiopulmonary resuscitation (CPR) and emergency cardiovascular care (ECC) of pediatric and neonatal patients: pediatric basic life support. Pediatrics 2006 May;117(5):e989-e1004. [doi: 10.1542/peds.2006-0219] [Medline: 16651298]

13. General Authority for Statistics. Demographic Survey 2016. Riyadh, Kingdom of Saudi Arabia: General Authority for Statistics, Saudi Arabian Government; 2016. URL: https://www.stats.gov.sa/sites/default/files/en-demographic-research-2016_2.pdf

14. Dixe MD, Gomes JCR. Knowledge of the Portuguese population on Basic Life Support and availability to attend training. Rev Esc Enferm USP 2015 Aug;49(4):640-649 [FREE Full text] [doi: 10.1590/S0080-623420150000400015] [Medline: 26353102]

15. Pergola AM, Araujo IE. Laypeople and basic life support. Rev Esc Enferm USP 2009 Jun;43(2):335-342. [Medline: 19655674]

16. Cheung BM, Ho C, Kou KO, Kuong EE, Lai KW, Leow PL, University Of Hong Kong Cardiopulmonary Resuscitation Knowledge Study Group. Knowledge of cardiopulmonary resuscitation among the public in Hong Kong: telephone questionnaire survey. Hong Kong Med J 2003 Oct;9(5):323-328 [FREE Full text] [Medline: 14530525]

17. Alharbi MM, Horaib YF, Almutairi OM, Alsuaidan BH, Alghoraibi MS, Alhadeedi FH, et al. Exploring the extent of knowledge of CPR skills among school teachers in Riyadh, KSA. J Taibah U Med Sci 2016 Oct;11(5):497-501. [doi: 10.1016/j.jtumed.2016.07.007]

18. Rasmus A, Czekajlo MS. A national survey of the Polish population's cardiopulmonary resuscitation knowledge. Eur J Emerg Med 2000 Mar;7(1):39-43. [Medline: 10839378]

19. Pane GA, Salness KA. A survey of participants in a mass CPR training course. Ann Emerg Med 1987 Oct;16(10):1112-1116. [Medline: 3662155]

20. Özbilgin Ş, Akan M, Hancı V, Aygün C, Kuvaki B. Evaluation of public awareness, knowledge and attitudes about cardiopulmonary resuscitation: report of İzmir. Turk J Anaesthesiol Reanim 2015 Dec;43(6):396-405 [FREE Full text] [doi: 10.5152/TJAR.2015.61587] [Medline: 27366536]

21. Selby ML, Kautz JA, Moore TJ, Gombeski WR, Ramirez AG, Farge EJ, et al. Indicators of response to a mass media CPR recruitment campaign. Am J Public Health 1982 Sep;72(9):1039-1042. [Medline: 7102855]

22. Gregorio J, Lee J. Education and income inequality: new evidence from cross-country data. Rev Income Wealth 2002 Sep;48(3):395-416. [doi: 10.1111/1475-4991.00060]

23. Pergola AM, Araujo IE. O leigo em situação de emergência. Rev Esc Enferm USP 2008 Dec;42(4):769-776. [doi: 10.1590/S0080-62342008000400021]

24. Murad M, Husum H. Trained lay first responders reduce trauma mortality: a controlled study of rural trauma in Iraq. Prehosp Disaster Med 2010;25(6):533-539. [Medline: 21181688]

25. Ashour A, Cameron P, Bernard S, Fitzgerald M, Smith K, Walker T. Could bystander first-aid prevent trauma deaths at the scene of injury? Emerg Med Australas 2007 Apr;19(2):163-168. [doi: 10.1111/j.1742-6723.2007.00948.x] [Medline: 17448104]

26. Wissenberg M, Lippert FK, Folke F, Weeke P, Hansen CM, Christensen EF, et al. Association of national initiatives to improve cardiac arrest management with rates of bystander intervention and patient survival after out-of-hospital cardiac arrest. JAMA 2013 Oct 02;310(13):1377-1384. [doi: 10.1001/jama.2013.278483] [Medline: 24084923]

27. McCluskey D, Moore P, Campbell S, Topping A. Teaching CPR in secondary education: the opinions of head teachers in one region of the UK. Resuscitation 2010 Nov;81(11):1601. [doi: 10.1016/j.resuscitation.2010.06.011] [Medline: 20727654]

Abbreviations

ALS: advanced life support
BLS: basic life support
CPR: cardiopulmonary resuscitation
OHCA: out-of-hospital cardiac arrest
OR: odds ratio

Original Paper

Importance and Presence of High-Quality Evidence for Clinical Decisions in Neurosurgery: International Survey of Neurosurgeons

Jill Martens[1*], BSc; Guido de Jong[1*], MSc; Maroeska Rovers[1*], PhD; Gert Westert[1*], PhD; Ronald Bartels[1*], MD, PhD

Neurosurgical Center Nijmegen, Department of Neurosurgery, Radboud University Medical Center, Nijmegen, Netherlands
[*]all authors contributed equally

Corresponding Author:
Jill Martens, BSc
Neurosurgical Center Nijmegen
Department of Neurosurgery
Radboud University Medical Center
Department of Neurosurgery, huispost 636
Postbus 9101
Nijmegen, 6500 HB
Netherlands
Phone: 31 681557132
Fax: 31 243635117
Email: jill.martens@radboudumc.nl

Abstract

Background: The publication rate of neurosurgical guidelines has increased tremendously over the past decade; however, only a small proportion of clinical decisions appear to be based on high-quality evidence.

Objective: The aim was to evaluate the evidence available within neurosurgery and its value within clinical practice according to neurosurgeons.

Methods: A Web-based survey was sent to 2552 neurosurgeons, who were members of the European Association of Neurosurgical Societies.

Results: The response rate to the survey was 6.78% (173/2552). According to 48.6% (84/173) of the respondents, neurosurgery clinical practices are based on less evidence than other medical specialties and not enough high-quality evidence is available; however, 84.4% (146/173) of the respondents believed neurosurgery is amenable to evidence. Of the respondents, 59.0% (102/173) considered the neurosurgical guidelines in their hospital to be based on high-quality evidence, most of whom considered their own treatments to be based on high-quality (level I and/or level II) data (84.3%, 86/102; significantly more than for the neurosurgeons who did not consider the hospital guidelines to be based on high-quality evidence: 55%, 12/22; *P*<.001). Also, more neurosurgeons with formal training believed they could understand, criticize, and interpret statistical outcomes presented in journals than those without formal training (93%, 56/60 and 68%, 57/84 respectively; *P*<.001).

Conclusions: According to the respondents, neurosurgery is based on high-quality evidence less often than other medical specialties. The results of the survey indicate that formal training in evidence-based medicine would enable neurosurgeons to better understand, criticize, and interpret statistical outcomes presented in journals.

KEYWORDS

Introduction

Evidence-based neurosurgery is a paradigm of neurosurgical practice in which the best available evidence is consulted to establish the principles of diagnosis and treatment. These principles are applied considering the neurosurgeon's training and experience, as well as being informed by the patient's individual circumstances and preferences, to produce the best possible health outcomes [1]. Although evidence-based medicine (EBM) is the gold standard in medicine [1-3], it is estimated that only 10% to 25% of clinical decisions are based on

high-quality evidence [4], defined as level I and level II evidence (see Table 1 for definitions) [5].

Table 1. Levels of evidence in neurosurgery.

Level of evidence	Studies
I	(1) Randomized controlled trial, (2) meta-analysis of randomized controlled trials with homogeneous results
II	(1) Prospective comparative study (therapeutic), (2) meta-analysis of level II studies or level I studies with inconsistent results
III	(1) Retrospective cohort study, (2) case-control study, (3) meta-analysis of level III studies
IV	(1) Case series
V	(1) Case report, (2) expert opinion, (3) personal observation

In 2011, a new rating system, the Grading of Recommendations Assessment, Development, and Evaluation (GRADE), was developed, which offers an outcome-centric system for rating the quality of evidence derived from different types of studies [6]. The GRADE guidelines enable the rating of the quality of evidence in systematic reviews and clinical guidelines, as well as a determination of the strength of recommendations made in these documents. To the best of our knowledge, neurosurgery is currently still using the levels of evidence more often than the GRADE guidelines.

Rothoerl et al [7]. and Yarascavitch et al [5] published investigations into the levels of evidence in the neurosurgical literature in 2003 and 2012, respectively. These studies assigned a level of evidence to all published clinical papers in three major neurosurgical journals for the years 1999 and 2009-2010, respectively, graded according to the study design shown in Table 1 [5,8,9]. The authors found that 22.8% and 10.3% of evidence was considered higher-level evidence (level I or level II). Level I evidence, from randomized controlled trials yielding homogeneous results, was only found in 3.8% and 2.1% of the papers evaluated in these two studies, respectively.

These studies suggest that surgeons are increasingly turning their backs on research. Further evidence indicates that compared with a decade or two ago, surgeons apply for and receive fewer grants, publish less, and—perhaps most perniciously—feel that research is not part of their role [10]. Involvement in research allows surgeons to develop rigor in their everyday work and to judge, maintain, and improve the quality of the work done by their peers.

The goal of this study is to investigate the opinion of neurosurgeons about the evidence available in neurosurgical practice and the extent to which this evidence is implemented in clinical practice.

Methods

Recruitment

Evaluation by an ethical committee was not necessary for this study. A survey was conducted among 2552 members of the European Association of Neurosurgical Societies (EANS). The survey asked the opinion of neurosurgeons regarding the levels of evidence generated in neurosurgical studies, their understanding of the levels of evidence, and to what extent neurosurgeons implement evidence in clinical practice. The survey was emailed directly to the members of the EANS by

the society's administrative personnel. Data were collected over a period of 3 weeks from the date of the first mailing. Two reminders, each 1 week apart, were sent to the cohort.

Survey

The survey (Multimedia Appendix 1) was made using a Google Inc program (Google Forms) and consisted of 13 sections containing a total of 22 questions. Sections with multiple questions within the survey were randomized to minimize the influence of the sequence of questions on the answers. Levels of evidence were used instead of ratings determined using the GRADE recommendations because the levels are still more commonly used by neurosurgeons to the best of our knowledge. Participants were asked for their opinions on high-quality evidence, the usability of the results of different research methods in clinical practice, the amenability of neurosurgery to evidence, the quality of guidelines in their hospital and of the guidelines used by the neurosurgeons themselves, and the most important factors for choosing between treatments. The guidelines mentioned in the questionnaire were selected in a previous study [11], in which a PubMed search was used to identify the most recent guidelines available. These guidelines were then characterized by the strength of their evidence [11]. This search covered the Agency for Healthcare Research and Quality (AHRQ) National Guidelines Clearinghouse and included both European guidelines and American guidelines. The participants were also asked whether they had received formal training in EBM, such as EU-ebm or CEBM, and if they considered themselves capable of understanding, criticizing, and interpreting statistical outcomes in journals.

Most questions consisted of a five-item Likert scale, which was chosen because each item is of equal value, so the respondents were scored rather than the items. The Likert scale was also likely to yield highly reliable answers and is easy to read and complete [12]. Some questions, for example about formal training in EBM and the ability to understand, criticize, and interpret statistical outcomes, were asked with a choice of Likert scale answers to enable the neurosurgeons to "rate" their training or ability to understand outcomes. The remaining questions had binary answers or were choices between statements. Participation was voluntary and completely anonymous, and the purpose of the survey was explained to the participants.

Statistical Analyses

IBM SPSS version 22 (Armonk, NY, USA) was used for the statistical analyses. For the continuous data, Student t tests were used, whereas chi-square tests were used to analyze categorical

data. Some values, for example the number of years as a neurosurgeon, were categorized before being statistically examined. Comparisons were made between different groups of neurosurgeons. Multiple groups were formed based on the answers to the questions asked and their opinions on different aspects, and consisted of respondents responding with a positive answer ("strongly agree," "agree," or "yes"), a negative answer ("strongly disagree," "disagree," or "no"), or an indecisive answer ("indifferent") to particular questions. Afterwards, comparisons were made between the answers and opinions of certain groups for different aspects of neurosurgery. Values are presented as a mean ±95% confidence interval (CI).

Results

The response rate was 6.78% (177 respondents) of the 2552 EANS members surveyed. All completed surveys contained complete data and had no partial or missing responses. Four respondents were excluded: three were still residents and one response was sent twice. A final total of 173 completed surveys (6.78%) were analyzed.

Respondent Statistics

Table 2 shows the demographics of the respondents. Their years of experience varied. Most respondents (98.3%, 170/173) were specialized in one or more subspecialties. Of these, 85.9% (146/170) were specialized in two or more subspecialties, with a mean of 3.2 per person (95% CI 2.97-3.43). A total of 57.2% (99/173) of the respondents had one or more academic qualifications, such as a professorship or a PhD. The majority of the respondents (79.2%, 137/173) worked in one of 29 European countries, mostly Germany (10.4%, 18/173), the Netherlands (7.5%, 13/173), Greece (6.4%, 11/173), or the United Kingdom (5.8%, 10/173).

Table 2. Demographics of the respondents (N=173).

Demographics	n (%)
Years working as a neurosurgeon	
1-5	30 (17.3)
5-10	46 (26.6)
10-15	38 (22.0)
15-20	13 (7.5)
20-25	13 (7.5)
25-30	16 (9.3)
>30	17 (9.8)
Academic qualifications[a]	
Yes[b]	99 (57.2)
Professor	28 (23.0)
PhD	67 (54.9)
MSPH	2 (1.6)
MPH	5 (4.1)
Other	20 (16.4)
Subspecialty	
Yes[c]	170 (98.3)
Neurocritical care	62 (11.7)
Cerebrovascular neurosurgery	70 (13.2)
Neuroendovascular surgery	14 (2.6)
Spinal neurosurgery	121 (22.9)
Neurosurgical oncology	125 (23.6)
Pediatric neurosurgery	57 (10.8)
Peripheral nerve neurosurgery	35 (6.6)
Stereotactic and functional neurosurgery	30 (5.7)
Other	15 (2.8)

[a]PhD: Doctor of Philosophy; MPH: Master of Public Health; MSPH: Master of Science in Public Health.

[b]28.6% of the neurosurgeons who answered "yes" had more than one academic qualification, with a mean of 1.2 per person (95% CI 1.11-1.29).

[c]85.9% of the neurosurgeons who answered "yes" had more than one subspecialty, with a mean of 3.2 per person (95% CI 2.97-3.43).

The remaining 36 respondents (20.8%, 36/173) worked in one of 18 non-European countries, particularly in countries in the Middle East (11.6%, 20/173), such as Saudi Arabia (2.9%, 5/173), Pakistan (2.3%, 4/173), and Iraq (1.7%, 3/173). The remaining respondents came mainly from Mexico (1.7%, 3/173), India (1.7%, 3/173), and the United States (1.2%, 2/173).

Evaluation Outcomes

Table 3 shows the opinions of the respondents regarding the levels of evidence and the use of EBM in clinical practice. Of the 173 respondents, 84 (48.6%) considered level I or level I and level II evidence to be of high quality. Figure 1 shows the levels of evidence used by neurosurgeons in clinical practice; most respondents implemented all levels of evidence into their clinical practice. The results of randomized controlled trials (RCTs) with inconsistent, but promising, results were used by

fewer than half of the respondents (45.7%, 79/173; Tables 4 and 5).

Every participant indicated the guidelines they used most often, to a maximum of three guidelines (Table 6). The level of evidence generated by the research underpinning each guideline is shown. Guidelines considering the surgical management of traumatic brain injury were most commonly used (39.3%, 68/173), followed by those for severe traumatic brain injury (38.2%, 66/173) and subarachnoid hemorrhage (38.2%, 66/173). Table 6 also shows the number of neurosurgeons using the guideline who subspecialized in the corresponding area of neurosurgery. The numbers of neurosurgeons specializing in the areas corresponding to the three most-used guidelines did not comprise a large percentage of the total number of neurosurgeons using these guidelines (48.5%, 48.5%, and 78.8% of neurosurgeons specializing in the surgical management of

traumatic brain injury, severe traumatic brain injury, and subarachnoid hemorrhage, respectively). This may be because these three areas are all critical conditions that require immediate care; therefore, it is likely that most neurosurgeons will use these guidelines even if it is not their subspecialty.

Table 3. Levels of evidence considered by neurosurgeons to be of high quality and usable in clinical practice.

Levels of evidence	Considered to be of high quality and usable in clinical practice, n (%)
None	3 (1.7)
Level I	15 (8.7)
Level I and level II	69 (39.9)
Level I, level II, and level III	53 (30.6)
Level I, level II, level III, and level IV	2 (1.2)
All levels (Level I-V)	31 (17.9)

Figure 1. Levels of evidence used by neurosurgeons in clinical practice. The results are presented on a five-item Likert scale: (1) strongly agree, (2) agree, (3) indifferent, (4) disagree, or (5) strongly disagree.
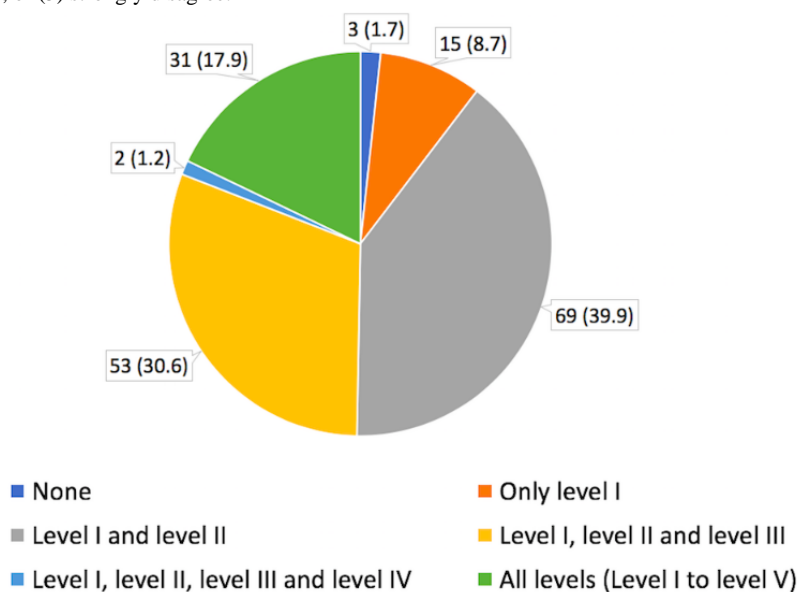
Table 4. Studies used by the participants in clinical practice. Scores were given from 1 (strongly disagree) to 5 (strongly agree).

Level of evidence	Studies	Studies used, mean (95% CI)
I	(1) RCT[a], (2) meta-analysis of RCTs with homogeneous results	3.8 (2.14-5.46)
II	(1) Prospective comparative study (therapeutic)	3.9 (2.48-5.32)
II	(2) Meta-analysis of level II studies	3.9 (2.34-5.46)
II	(3) Meta-analysis of level I studies with inconsistent results	3.3 (1.54-5.06)
III	(1) (Meta-analysis of) retrospective cohort study	3.8 (2.36-5.24)
III	(2) Case-control study	3.6 (2.06-5.14)
IV	(1) Case series	3.7 (2.02-5.38)
V	(1) Case report, (2) expert opinion, (3) personal observation	3.5 (1.64-5.36)

[a]RCT: randomized controlled trial.

Table 5. Summary of the overall survey results (N=173).

Survey item	Strongly agree or agree, n (%)	Indifferent, n (%)
Factors important for choosing a treatment		
Clinical experience is an important factor for choosing a treatment	172 (99.4)	0 (0.0)
Research is an important factor for choosing a treatment	160 (92.5)	10 (5.8)
Knowledge from patients and carers is an important factor for choosing a treatment	124 (71.7)	39 (22.4)
Local context and environment are important factors for choosing a treatment	123 (71.1)	42 (24.3)
Use of research in clinical practice		
I use prospective cohort studies in clinical practice	134 (77.5)	34 (19.7)
I use meta-analysis of prospective cohort studies in clinical practice	130 (75.1)	35 (20.2)
I use (meta-analysis of) retrospective cohort studies in clinical practice	126 (72.9)	39 (22.5)
I use (meta-analysis of) RCTs[a] with homogeneous results in clinical practice	122 (70.5)	41 (23.7)
I use case-control studies in clinical practice	113 (65.3)	42 (24.3)
I use case series in clinical practice	111 (64.2)	47 (27.2)
I use case reports, expert opinions, or personal observations in clinical practice	98 (56.6)	52 (30.1)
I use (meta-analysis of) RCTs with inconsistent, but promising, results in clinical practice	79 (45.7)	66 (38.2)
Guidelines and treatment options		
Treatment options I use are based on high-quality evidence	129 (74.5)	26 (15.0)
The neurosurgeons at my hospital are involved in the process of setting up the neurosurgical guidelines for my hospital	126 (72.8)[b]	3 (1.7)[c]
Guidelines at my hospital are based on high-quality evidence	102 (59.0)	49 (28.3)
Training		
I can understand, criticize, and interpret statistical outcomes in journals	87 (80.3)	19 (11.1)
I have received formal training in EBM[d]	60 (34.7)	29 (16.8)
Neurosurgery is amenable to evidence	146 (84.4)	19 (11.0)

[a]RCT: randomized controlled trial.

[b]Question was answered with "yes."

[c]Question was answered with "other."

[d]EBM: evidenced-based medicine.

Table 6. Most-used neurosurgical guidelines (N=173).

Guidelines	Level of evidence in research used to develop guideline [11]	Neurosurgeons using this guideline, n (%)	Neurosurgeons using this guideline subspecialized in this field, n (%)
Head injury	—[a]	150 (86.7)	—
Surgical management of traumatic brain injury	Moderate	68 (39.3)	33 (49)
Severe traumatic brain injury	Moderate	66 (38.2)	32 (49)
Pediatric traumatic brain injury	Moderate	9 (5.2)	7 (78)
Mild traumatic brain injury	High/moderate	7 (4.0)	4 (44)
Spine	—	136 (78.6)	—
Lumbar disk herniation	All levels	38 (22.0)	32 (84)
Cervical spine and spinal cord injury	Moderate	19 (11.0)	19 (100)
Degenerative lumbar spondylolisthesis	Moderate/low	17 (9.8)	15 (88)
Degenerative lumbar stenosis	Moderate/low	13 (7.5)	10 (77)
Degenerative cervical spine disease	Moderate	11 (6.4)	11 (100)
Lumbar spine fusion	All levels	11 (6.4)	10 (91)
Cervical radiculopathy and degenerative disease	All levels	10 (5.8)	10 (100)
Antibiotic prophylaxis in spine surgery	All levels	9 (5.2)	7 (78)
Intraoperative spinal monitoring	High	6 (3.5)	4 (67)
Somatosensory evoked potentials	Moderate	1 (0.6)	1 (100)
Vertebral osteomyelitis, diskitis, and epidural abscess	Moderate/low	1 (0.6)	1 (100)
Vascular	—	96 (55.5)	—
Subarachnoid hemorrhage	All levels	66 (38.2)	52 (79)
Intracerebral hemorrhage	High	23 (13.3)	19 (83)
Extracranial carotid disease	High/moderate	7 (4.0)	6 (86)
Tumor	—	81 (46.8)	—
Glioblastoma	Moderate	60 (34.7)	52 (87)
Brain metastases	High/moderate	21 (12.1)	18 (86)
Functional	—	11 (6.4)	—
Deep brain stimulation	High/moderate	10 (5.8)	9 (90)
Vagal nerve stimulation	Moderate	1 (0.2)	0 (0)
Other	—	22 (4.4)	—
Hydrocephalus	Moderate	18 (3.6)	14 (78)
Carpal tunnel syndrome	High/moderate	4 (0.8)	1 (25)

[a]Not applicable.

According to 84.4% of the neurosurgeons (146/173), neurosurgery is amenable to evidence (Table 5); however, nearly half of the respondents (48.6%, 84/173) believed that neurosurgery is less based on evidence than other medical specialties. Despite this, 74.6% of the respondents (129/173) consider their treatments to be based on level I and/or level II evidence. Of those who believed neurosurgery is amenable to evidence, 78.8% (115/146 respondents) considered their treatments to be based on level I and/or level II evidence, whereas significantly fewer (25%, 2/8) of those who did not believe neurosurgery is amenable to evidence considered their research to be based on such high-quality evidence (P=.048).

Of the most-used neurosurgical guidelines (Table 6), only 9.78% (242/2469) were based on level I evidence, whereas 20.43% (545/2668) were based on level II evidence.

Of the 129 respondents who believed their treatments were based on high-quality evidence, 50.4% (65/129) considered level I and/or level II to be high-quality evidence, whereas significantly fewer (39%, 7/18) of the respondents who did not consider their treatments to be based on high-quality evidence considered level I and/or level II to be high quality (P=.02).

Most respondents (72.8%, 126/173) were involved in the process of setting up the neurosurgical guidelines in their hospital. More

than half (59.0%, 102/173) of the respondents considered the neurosurgical guidelines of their hospital to be based on high-quality evidence. Of those who were involved in their establishment, 65.9% (83/126) considered the guidelines of their hospital to be based on high-quality evidence, compared with just 43% (20/47) of those who were not involved (P=.02).

Of the 59.0% (102/173) of respondents who considered the neurosurgical guidelines in their hospital to be based on high-quality evidence, 84.3% (86/102) considered their own treatments to be based on level I and/or level II evidence, whereas 55% (12/22) of the respondents who did not consider hospital guidelines to be based on high-quality evidence considered their own treatments to be based on level I and/or level II evidence (P<.001).

Only 34.7% (60/173) of the respondents said they had received formal training in EBM. Of those who received formal training, 76% (46/60) considered their own treatments to be based on level I and/or level II evidence, whereas 85% (71/84) of those without formal training believed their treatments were based on this level of evidence (P=.03).

The majority of respondents (80.3%, 139/173) said they could understand, criticize, and interpret statistical outcomes in medical research. This response was more common for respondents who received formal training (93%, 56/60) than for those without formal training (68%, 57/84; P<.001). There was no difference between the number of respondents with and without additional academic qualifications who stated that they could understand, criticize, and interpret statistical outcomes (85%, 84/99 and 74%, 55/74, respectively; P=.16).

All participants had the option to add their own comments at the end of the survey. The most frequent comment was that the lack of evidence is an important issue in neurosurgery. Neurosurgeons also said that RCTs are expensive and difficult to perform, although well-designed prospective comparative studies could be equally informative and easier to run. They therefore concluded that dismissing study designs other than RCTs when developing neurosurgical guidelines is holding back neurosurgery.

Discussion

This study is unique because it is, to the best of our knowledge, the first to evaluate the opinion of neurosurgeons in several countries regarding the use of evidence in neurosurgery. Level I and level II evidence is considered high quality; however, despite a worldwide acceptance of this classification, only 48.5% of the respondents (84/173) considered either level I or levels I and II to be high quality. Moreover, all levels of evidence seem to be used by the majority of neurosurgeons. Several neurosurgeons commented that the lack of evidence is an important issue in neurosurgery.

Indisputable advancements in neurosurgery have been traditionally based on technical innovations advocated by pioneers without rigid assessment in clinical trials; therefore, changes in clinical practice have frequently been technology-driven rather than strictly evidence-based [13]. Everyday clinical management in neurosurgery does not,

therefore, always seem to comply with the best available evidence.

Since 1970, the rate of increase in the publication of guidelines in all specialties has outpaced neurosurgery [11]; however, in the past 5 years, the number of guidelines published per year in neurosurgery has increased at the same rate as all specialties [11]. The available literature shows that neurosurgery uses a higher percentage of high-level evidence than some other specialties, including general plastic surgery [14] and maxillofacial surgery [15]; however, neurosurgery is still lagging behind many other specialties, including orthopedics [16], ophthalmology [17], otolaryngology [18], esthetic surgery [19], and urology [20]. The situation in other fields resembles that of neurosurgery; for example, 12.2% of the treatment for atrial fibrillation is based on level I and II evidence [21], although we could not find any data on the levels of evidence used in the treatment of cardiovascular disease as a whole. When comparing neurosurgery with oncology, we discovered that oncology uses the AGREE (Appraisal of Guidelines for Research and Evaluation) rating of guidelines [22-25]. The AGREE domains (scope and purpose, stakeholder involvement, rigor of development, clarity of presentation, applicability, and editorial independence) [26] are not comparable with the levels of evidence, in which studies are graded by study design.

Of the respondents who participated in this study, 25.4% (44/173) did not think that, or know whether, the treatment options they use are based on high-quality evidence. Ducis et al [11] investigated the quality of the guidelines used in neurosurgery clinical practice. In neurosurgery, 24.4% of the guidelines were based mainly on level I recommendations, whereas for vascular neurosurgery guidelines this percentage is significantly higher: 51.9%. Some other specialties have numbers of level I-based recommendations similar to neurosurgery, including endocrinology [27], infectious diseases [28], and hepatology [29]. Vascular neurosurgery is the subspecialty with the highest publication rate in neurosurgery [5], and vascular neurosurgery guidelines are the third most commonly used in clinical practice (according to Table 6). Guidelines relating to traumatic brain injuries are the most used according to the respondents, but this subspecialty accounted for just 6.4% of neurosurgery publications between 2009 and 2010 [5]. The level I-based recommendations for traumatic brain injury guidelines accounted for only 5.6% of all recommendations [11], significantly less than the level I-based recommendations for spine guidelines (10.0%) and vascular guidelines (51.0%), as assessed with a chi-square test (P<.001).

One participant commented that neurosurgery is currently based more on eminence than on evidence. Eminence refers to a clinical decision that is made solely by relying on the opinion of a medical specialist or any prominent health professional rather than the critical appraisal of the scientific evidence available [30]. Evidence is an integration of clinical knowledge and skills with the best critically appraised research available, as well as patient values and preferences, in order to make a clinical decision [30]. With the lack of evidence available to neurosurgeons, neurosurgery seems indeed to be based more on eminence than evidence in some cases.

Another neurosurgeon commented that the current definition of evidence-based neurosurgery is in dire need of rigorous update and expansion. A common misunderstanding of EBM is that a lack of available evidence means a lack of RCTs [31]. EBM evaluates the quality of evidence, based primarily on the likelihood that the evidence is biased. A powerful RCT is the best standard for evaluating this inherent bias, but it does not follow that only RCTs can be used to justify clinical practice in EBM; rather, EBM requires that we attempt to audit our decisions by obtaining the highest level of evidence ethically or logistically possible [31,32].

The differences in the confidence of neurosurgeons with and without formal training in EBM to adequately interpret the statistical outcomes presented in the literature was striking. Only 35% of the responding neurosurgeons (60/173) had received formal training in EBM and although it seems low, it is similar to that described in orthopedic surgery [33]. The lack of EBM training in neurosurgery has already been noted elsewhere; the University of Western Ontario in London, ON, Canada [34], and the American Accreditation Council for Graduate Medical Education [31] recently incorporated EBM into the curriculum for residency training programs in neurosurgery. Since EBM is based on the implementation of correctly interpreted research results, our findings may represent an argument for the introduction of more formal EBM training in the medical curriculum and supplementary training for neurosurgical departments.

A major strength of this study is the broad representation of the opinions of neurosurgeons worldwide through the involvement of the EANS. EANS has a large number of members all over the world. The membership of EANS is primarily located in Europe, but neurosurgeons from all countries are permitted to join. However, some potential limitations should also be discussed. First, due to the low response rate, selection bias cannot be precluded, which hampers the generalizability of our results. For external email surveys, a response rate between 10% to 25% is usually considered the average [35,36]. The response rate to our survey was a little lower than average, which might be caused by a general lack of interest by neurosurgeons or because they did not recognize the importance of this study. Bias could have been introduced because the survey may have been selectively answered by those who consider EBM important, specifically by those who were trained in EBM, involved in innovation and guidelines, or opinion leaders. Second, the participants of the survey may have chosen to provide socially desirable answers; however, this was counteracted by emphasizing the anonymity of the survey. A third potential issue is that neurosurgery is in the middle of transitioning from rating the guidelines using levels of evidence to rating them with the GRADE guidelines. We choose to use the levels of evidence here because they are currently more widely used by neurosurgeons; however, this could have been confusing for those who have already begun to use the GRADE guidelines.

The transition from using levels of evidence to the GRADE system is an important change in all medical specialties, including neurosurgery. The literature available about the evaluation of evidence in neurosurgery does not include the new system and is, therefore, in need of expansion and updating. The GRADE guidelines offer some major advantages over the levels of evidence, as they enable the rating of evidence in systematic reviews and guidelines and the grading of the strength of recommendations made in guidelines. Moreover, the GRADE system offers a transparent and structured process for developing and presenting summaries of evidence in systematic reviews and guidelines in health care, and for carrying out the steps involved in developing recommendations [6].

There are major gaps between the current definition of high-quality evidence (level I and level II) and neurosurgeons' opinions of evidence. Neurosurgeons are willing to base their clinical decisions on more than just RCTs, relying on lower-quality evidence. The transition to the GRADE system is one way to overcome this issue because these guidelines are outcome-centric and do not rate each study as a single unit. In the GRADE approach, RCTs are initially considered to be high-quality evidence and observational studies are initially considered low-quality evidence supporting estimates of intervention effects. Five factors may lead the ratings to be decreased, whereas three other factors may lead to an increased rating. Ultimately, the quality of evidence for each outcome falls into one of four categories, from high to very low [6].

EBM is important when choosing between treatments for patients; however, shared decision making (SDM) was also developed alongside the introduction of EBM. According to 71.7% (124/173) of the respondents of this study, knowledge gained from patients and carers was an important factor in the selection of a treatment, whereas 22.5% of the respondents (39/173) were indifferent and 5.8% (10/173) did not believe that the knowledge of patients and carers was an important factor when choosing a treatment. Evidence from trials has shown that engaged patients consume less health care resources [37,38]; furthermore, when doctors are too focused on EBM, preference misdiagnoses (also known as silent misdiagnoses) can be made, causing the patient to receive an unwanted treatment [39]. EBM is important but has to be combined with SDM to give patients the right treatment. Learning to combine these two sides of medicine could be an important development in all specialties, leading to the optimization of health care for patients.

Understanding EBM is key to using it correctly in practice. This study shows that relatively few neurosurgeons have received formal training in EBM; thus, more training in EBM, both in the medical curriculum for residents and at neurosurgical departments, would enable neurosurgeons to improve their abilities to better facilitate the implementation of statistical results into clinical practice. It would be interesting to perform this research on a larger scale, including a wider multinational sample in the future.

According to the respondents, neurosurgery is less commonly based on high-quality evidence than other medical specialties. The results of the survey suggest that providing more formal training in EBM is desirable, enabling neurosurgeons to better understand, criticize, and interpret the statistical outcomes presented in journals.

Acknowledgments

We would like to give our special thanks to Professor JA Grotenhuis for supporting the distribution of the Web-based survey by the EANS. We also express our gratitude to Radboud in'to Languages for performing the English language editing of this manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

"Evidence in Neurosurgery" survey.

[PDF File (Adobe PDF File), 20KB - ijmr_v7i2e16_app1.pdf]

References

1. Haines S, Walters B. Evidence-Based Neurosurgery: An Introduction. New York: Thieme Medical Publishers; 2006.
2. Rycroft-Malone J, Seers K, Titchen A, Harvey G, Kitson A, McCormack B. What counts as evidence in evidence-based practice? J Adv Nurs 2004 Jul;47(1):81-90. [doi: 10.1111/j.1365-2648.2004.03068.x] [Medline: 15186471]
3. Carnwell R. Essential differences between research and evidence-based practice. Nurse Researcher 2001 Jan;8(2):55-68. [doi: 10.7748/nr2001.01.8.2.55.c6150]
4. Goodman K. Ethics and Evidence-Based Medicine: Fallibility and Responsibility in Clinical Science. Cambridge: Cambridge University Press; 2002.
5. Yarascavitch BA, Chuback JE, Almenawer SA, Reddy K, Bhandari M. Levels of evidence in the neurosurgical literature: more tribulations than trials. Neurosurgery 2012 Dec;71(6):1131-1137. [doi: 10.1227/NEU.0b013e318271bc99] [Medline: 22986592]
6. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol 2011 Apr;64(4):383-394. [doi: 10.1016/j.jclinepi.2010.04.026] [Medline: 21195583]
7. Rothoerl RD, Klier J, Woertgen C, Brawanski A. Level of evidence and citation index in current neurosurgical publications. Neurosurg Rev 2003 Oct;26(4):257-261. [doi: 10.1007/s10143-003-0270-0] [Medline: 12827519]
8. Rutka JT. Editorial. Classes of evidence in neurosurgery. J Neurosurg 2017 Jun;126(6):1747-1748. [doi: 10.3171/2016.4.JNS161067] [Medline: 27367246]
9. Phillips B, Ball C, Sackett D, Badenoch D, Straus S, Haynes B. Centre for Evidence-based Medicine. 2009. Oxford Centre for Evidence-based Medicine-levels of evidence (March 2009) URL: https://tinyurl.com/y9vtoj38 [accessed 2017-12-03] [WebCite Cache ID 6vQ1GXklP]
10. Bellini S. More surgeons must start doing basic science. Nature 2017 Dec 21;544(7651):393-394. [doi: 10.1038/544393b] [Medline: 28447654]
11. Ducis K, Florman JE, Rughani AI. Appraisal of the quality of neurosurgery clinical practice guidelines. World Neurosurg 2016 Jun;90:322-339. [doi: 10.1016/j.wneu.2016.02.044] [Medline: 26947727]
12. Page-Bucci H. The Value of Likert Scales in Measuring Attitudes of Online Learners. 2003 Feb. URL: http://www.hkadesigns.co.uk/websites/msc/reme/likert.htm [accessed 2017-12-03] [WebCite Cache ID 6vQ1NHZt0]
13. Weber C, Jakola AS, Gulati S, Nygaard OP, Solheim O. Evidence-based clinical management and utilization of new technology in European neurosurgery. Acta Neurochir (Wien) 2013 Apr;155(4):747-754. [doi: 10.1007/s00701-013-1640-1] [Medline: 23440373]
14. Loiselle F, Mahabir RC, Harrop AR. Levels of evidence in plastic surgery research over 20 years. Plast Reconstr Surg 2008 Apr;121(4):207e-211e. [doi: 10.1097/01.prs.0000304600.23129.d3] [Medline: 18349600]
15. Lau SL, Samman N. Levels of evidence and journal impact factor in oral and maxillofacial surgery. Int J Oral Maxillofac Surg 2007 Jan;36(1):1-5. [doi: 10.1016/j.ijom.2006.10.008] [Medline: 17129707]
16. Obremskey WT, Pappas N, Attallah-Wasif E, Tornetta P, Bhandari M. Level of evidence in orthopaedic journals. J Bone Joint Surg Am 2005 Dec;87(12):2632-2638. [doi: 10.2106/JBJS.E.00370] [Medline: 16322612]
17. Lai TY, Leung GM, Wong VW, Lam RF, Cheng AC, Lam DS. How evidence-based are publications in clinical ophthalmic journals? Invest Ophthalmol Vis Sci 2006 May;47(5):1831-1838. [doi: 10.1167/iovs.05-0915] [Medline: 16638988]
18. Wasserman JM, Wynn R, Bash TS, Rosenfeld RM. Levels of evidence in otolaryngology journals. Otolaryngol Head Neck Surg 2006 May;134(5):717-723. [doi: 10.1016/j.otohns.2005.11.049] [Medline: 16647522]
19. Chuback JE, Yarascavitch BA, Eaves F, Thoma A, Bhandari M. Evidence in the aesthetic surgical literature over the past decade: how far have we come? Plast Reconstr Surg 2012 Jan;129(1):126e-134e. [doi: 10.1097/PRS.0b013e3182362bca] [Medline: 22186527]

20. Borawski KM, Norris RD, Fesperman SF, Vieweg J, Preminger GM, Dahm P. Levels of evidence in the urological literature. J Urol 2007 Oct;178(4 Pt 1):1429-1433. [doi: 10.1016/j.juro.2007.05.150] [Medline: 17706710]

21. Barnett AS, Lewis WR, Field ME, Fonarow GC, Gersh BJ, Page RL, et al. Quality of evidence underlying the American Heart Association/American College of Cardiology/Heart Rhythm Society Guidelines on the management of atrial fibrillation. JAMA Cardiol 2017 Mar 01;2(3):319-323. [doi: 10.1001/jamacardio.2016.4936] [Medline: 28002833]

22. Lei X, Liu F, Luo S, Sun Y, Zhu L, Su F, et al. Evaluation of guidelines regarding surgical treatment of breast cancer using the AGREE Instrument: a systematic review. BMJ Open 2017 Nov 14;7(11):e014883 [FREE Full text] [doi: 10.1136/bmjopen-2016-014883] [Medline: 29138191]

23. Irani S, Rashidian A, Yousefi-Nooraie R, Soltani A. Evaluating clinical practice guidelines developed for the management of thyroid nodules and thyroid cancers and assessing the reliability and validity of the AGREE instrument. J Eval Clin Pract 2011 Aug;17(4):729-736. [doi: 10.1111/j.1365-2753.2010.01561.x] [Medline: 21223456]

24. Chen Y, Wang Y, Li W, Chen L, Xu C, Lu T, et al. Critical evaluation of the quality and recommendations of clinical practice guidelines for nasopharyngeal carcinoma. J Natl Compr Canc Netw 2017 Dec;15(3):336-344. [Medline: 28275034]

25. Klaver CE, Groenen H, Morton DG, Laurberg S, Bemelman WA, Tanis PJ, Research committee of the European Society of Coloproctology. Recommendations and consensus on the treatment of peritoneal metastases of colorectal origin: a systematic review of national and international guidelines. Colorectal Dis 2017 Mar;19(3):224-236. [doi: 10.1111/codi.13593] [Medline: 28008728]

26. Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, AGREE Next Steps Consortium. AGREE II: advancing guideline development, reporting and evaluation in health care. CMAJ 2010 Dec 14;182(18):E839-E842 [FREE Full text] [doi: 10.1503/cmaj.090449] [Medline: 20603348]

27. Hazlehurst JM, Armstrong MJ, Sherlock M, Rowe IA, O'Reilly MW, Franklyn JA, et al. A comparative quality assessment of evidence-based clinical guidelines in endocrinology. Clin Endocrinol (Oxf) 2013 Feb;78(2):183-190. [doi: 10.1111/j.1365-2265.2012.04441.x] [Medline: 22624723]

28. Lee DH, Vielemeyer O. Analysis of overall level of evidence behind Infectious Diseases Society of America practice guidelines. Arch Intern Med 2011 Jan 10;171(1):18-22. [doi: 10.1001/archinternmed.2010.482] [Medline: 21220656]

29. Rowe IA, Parker R, Armstrong MJ, King AL, Houlihan DD, Mutimer D. Assessment of the quality of evidence underlying international guidelines in liver disease. Am J Gastroenterol 2012 Sep;107(9):1276-1282. [doi: 10.1038/ajg.2012.71] [Medline: 22951868]

30. Le HP. Students 4 Best Evidence. 2016 Jan 12. Eminence-based medicine vs evidence-based medicine URL: http://www.students4bestevidence.net/eminence-based-medicine-vs-evidence-based-medicine/ [accessed 2017-12-03] [WebCite Cache ID 6vQ1SMuhb]

31. Bandopadhayay P, Goldschlager T, Rosenfeld JV. The role of evidence-based medicine in neurosurgery. J Clin Neurosci 2008 Apr;15(4):373-378. [doi: 10.1016/j.jocn.2007.08.014] [Medline: 18249115]

32. Haines SJ. Evidence-based neurosurgery. Neurosurgery 2003 Jan;52(1):36-47; discussion 47. [Medline: 12493099]

33. Kumar M, Gopalakrishna C, Swaminath PV, Mysore SS. Evidence-based surgery--evidence from survey and citation analysis in orthopaedic surgery. Ann R Coll Surg Engl 2011 Mar;93(2):133-138 [FREE Full text] [doi: 10.1308/003588411X1285163910787] [Medline: 21073824]

34. Burneo JG, Jenkins ME, UWO Evidence-Based Neurology Group. Teaching evidence-based clinical practice to neurology and neurosurgery residents. Clin Neurol Neurosurg 2007 Jun;109(5):418-421. [doi: 10.1016/j.clineuro.2007.03.001] [Medline: 17418941]

35. Fryrear A. SurveyGizmo. 2015. 3 Ways to improve your survey response rates URL: https://www.surveygizmo.com/survey-blog/survey-response-rates/ [accessed 2017-12-03] [WebCite Cache ID 6vQ1fU41n]

36. Fincham JE. Response rates and responsiveness for surveys, standards, and the Journal. Am J Pharm Educ 2008 Apr 15;72(2):43 [FREE Full text] [Medline: 18483608]

37. Stacey D, Légaré F, Lewis K, Barry MJ, Bennett CL, Eden KB, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2011;4:CD001431. [doi: 10.1002/14651858.CD001431.pub5] [Medline: 28402085]

38. Wennberg DE, Marr A, Lang L, O'Malley S, Bennett G. A randomized trial of a telephone care-management strategy. N Engl J Med 2010 Sep 23;363(13):1245-1255. [doi: 10.1056/NEJMsa0902321] [Medline: 20860506]

39. Mulley AG, Trimble C, Elwyn G. Stop the silent misdiagnosis: patients' preferences matter. BMJ 2012 Nov 08;345:e6572. [Medline: 23137819]

Abbreviations

AGREE: Appraisal of Guidelines for Research and Evaluation
AHRQ: Agency for Healthcare Research and Quality
EANS: European Association of Neurosurgical Societies
EBM: evidence-based medicine
GRADE: Grading of Recommendations Assessment, Development, and Evaluation

RCT: randomized controlled trial
SDM: shared decision making

XSL•FO
RenderX

Original Paper

Body Mass Index Screening and Follow-Up: A Cross-Sectional Questionnaire Study of Pennsylvania School Nurses

Erica Francis[1], MS; Alicia Marie Hoke[1], MPH, CHES; Jennifer Lynn Kraschnewski[1], MD, MPH

Department of Pediatrics, Penn State University, Hershey, PA, United States

Corresponding Author:
Erica Francis, MS
Department of Pediatrics
Penn State University
90 Hope Drive, Suite 1103
Mail Code A145
Hershey, PA, 17033
United States
Phone: 1 717 531 1440 ext 8
Fax: 1 717 531 0942
Email: efrancis@psu.edu

Abstract

Background: Childhood overweight and obesity health concerns can affect a student's academic performance, so it is important to identify resources for school nurses that would help to improve self-efficacy, knowledge, and confidence when approaching parents with sensitive weight-related information and influence overall obesity prevention efforts in the school setting.

Objective: The purpose of this study was to conduct a Pennsylvania (PA) state-wide 29-item survey addressing school nursing barriers and practices, supplementing information already known in this area. Although the survey covered a range of topics, the focus was body mass index (BMI) screening and its related practice within the schools.

Methods: We conducted a state-wide Web-based survey of school nurses in PA to understand current areas of care, find ways to address child health through school BMI screenings and follow up, and identify current educational gaps to assist school nurses with providing whole child care within the realm of weight management. Chi-square test of independence was conducted to determine the relationship between BMI screening follow up and interest in a BMI toolkit.

Results: Nurse participants (N=210), with a 42% (210/500) response rate, represented 208 school districts across PA. Participants were asked about their current process for notifying parents of BMI screening results. The majority (116/210, 55.2%) send a letter home in the mail, while others (62/210, 29.5%) send a letter home with students. A small number (8/210, 3.8%) said they did not notify parents altogether, and some (39/210, 18.6%) notify parents electronically. More than one-third (75/210, 35.7%) of nurses reported receiving BMI screening inquiries from parents; however, under half (35/75, 46.7%) of those respondents indicated they follow up with parents whose child screens overweight or obese. Overall, the vast majority (182/210, 86.7%) do not follow up with parents whose child screens overweight or obese. The majority (150/210, 71.4%) of the nurses responded they would benefit from a toolkit with resources to assist with communication with parents and children about BMI screenings. A significant association between respondent follow up and interest in a BMI toolkit was observed (*P*=.01).

Conclusions: Schools must start recognizing the role school nurses play to monitor and promote children's health. This goal might include involving them in school-based preventive programs, empowering them to lead initiatives that support whole child health and ensuring opportunities for professional development of interest to them. Nonetheless, the first step in facilitating obesity prevention methods within schools is to provide school nurses with meaningful tools that help facilitate conversations with parents, guardians, and caregivers regarding their child's weight status and health through a BMI screening toolkit.

Introduction

Rising rates of childhood overweight and obesity contribute to increased risk of chronic disease and other health concerns [1,2] that directly affect a student's academic performance [3]. While many environmental factors play a role in the rising prevalence of obesity, school-based health services hold enormous potential to affect children's health outcomes given the fact that children spend the majority of their waking hours in school settings [4]. Unfortunately, the School Health Policies and Practices Study [5] indicates few services linked to obesity prevention despite school nurses' ability and interest to champion strategies to curb rising obesity rates [6]. School nurses are limited in the knowledge and resources available for addressing the topic of childhood obesity prevention [7]. As a result, there is an inconsistency in the involvement of school nurses in students' weight management education [8]. The National Association of School Nurses (NASN) recognizes the role of the school nurse in promoting the prevention of overweight and obesity and has developed a comprehensive toolkit with resources to assist nurses with the management of overweight and obese youth [9]. Despite existing resources, state-specific mandates to school nursing practices make a one-size-fits-all approach to providing educational opportunities and support for students with obesity difficult. School nurses who are introduced to a toolkit, like that of the NASN, are offered an identifiable and feasible role in that of childhood obesity prevention training [10].

One strategy that has been used since April 2000, when the Centers for Disease Control and Prevention released body mass index (BMI) screening tools for males and females 2-20 years old, is for schools to conduct their own BMI screenings among their student population. Consistent with early identification practice criteria, the Institute of Medicine and the American Academy of Pediatrics identified BMI screening as a practice that should be implemented by schools annually to help children and their parents understand and address healthy body weight [11-13]. Currently, 25 states across the United States mandate BMI screenings or weight-related assessments in their public schools in which only 44% of those states require parental notification [14]. The inconsistency in additional follow up by schools for parent inquiries may prevent affected children from receiving treatment. While school nurses serve as the liaison between the school and parents in communicating health concerns, they cite barriers to addressing obesity with their students and families for a variety of reasons, including lack of parental engagement, inadequate knowledge, lack of support, societal norms, and low perceived competency [7,15]. At the core of the NASN Framework for 21st Century School Nursing Practice [16] is student-centered health care in the context of the student's family and school community. This underscores the need to identify the resources and tools school nurses would find most helpful in improving communication and engagement with families and addressing weight-related health with their students.

Although BMI screening results are useful for educating parents and encouraging follow up with a health care provider when necessary, this information can also be used to monitor aggregate changes in BMI year to year and evaluate obesity prevention programs in the school setting. School nurses should be considered leaders for implementing effective school-based programs, as their involvement has resulted in improvements in BMI percentile in several studies [6]. Therefore, additional training opportunities to do so creates benefits beyond that of the role of the school nurse and, instead, expands to benefit students and the larger community [16,17]. Furthermore, nurse self-efficacy in performing childhood obesity prevention practices significantly influences involvement [18], underscoring the importance of determining professional development opportunities, specifically those that overcome traditional barriers related to time and financial resources [19]. Pennsylvania (PA) is 1 out of 11 states requiring BMI screening and BMI result reporting to parents or guardians [14]. Therefore, this study, focusing on PA school nurse BMI screenings, will provide insight into the efficacy and follow through of the school's role in childhood obesity prevention. We conducted a state-wide survey of school nurses in PA to understand current areas of care, find ways to address child health through school BMI screenings and follow up, and identify current educational gaps to assist school nurses with providing whole child care within the realm of weight management and education.

Methods

Design

This study included a cross-sectional Web-based survey delivered via email to 500 public school nurses in PA. Combined listservs for Penn State Health, PRO Wellness, and Penn*Link, an email listserv developed through the Pennsylvania Department of Education, were used to distribute the survey. All study and recruitment procedures were approved by the authors' institutional review board.

Procedure

School nurses were contacted in April 2016 to complete the survey through email invitation. The survey was open for 1 month and 2 reminders were sent to encourage participation. Inclusion criteria included school nurses in PA schools who were English-literate. Participants were prompted with consent language prior to entering the survey; thus, entering the survey implied consent to participate. The survey consisted of closed-ended, multiple choice, open-ended, and Likert-type scale questions and required 15 minutes to complete. Although survey respondents were anonymous, they were directed to a contact form on the Penn State PRO Wellness website to provide contact information for receipt of a gift card to compensate them for completing the survey.

Instrumentation

The purpose of this study was to conduct a PA state-wide 29-item survey addressing school nursing barriers and practices, supplementing information already known in this area. The study team worked collaboratively with the Director of the Division of School Health at the Pennsylvania Department of Health to design the survey. Questions were based upon identified information gaps. The survey was researcher developed consisting of 3 broad categories of information. The

XSL·FO

RenderX

categories included (1) demographic of the school; (2) primary areas of care; and (3) general development and methodology of school nurse practices. It aimed to understand current practices and resource preferences for addressing student health and discuss BMI screening results and professional development needs. Although the survey covered a range of topics, the focus was BMI screening and its related practice within the schools.

Within the primary areas of care section of the survey, nurses were asked to share common areas of care provided to students on a daily basis and partnerships in completing mandated health screenings. Additionally, this section addressed BMI screenings by asking about BMI letter usage, inquiries, and follow up. Specific questions asked of school nurses included how school nurses notify parents of BMI screening results, whether or not they commonly receive inquiries, whether or not school nurses would benefit from a toolkit with resources to help communicate with parents and children about BMI screenings, topics of greatest and least interest to include in the toolkit, and format of toolkit information (Figure 1 and Table 1). Further questions asked about topics of interest for professional development, methods of delivery of educational opportunities, and areas of need for assistance.

Figure 1. Body mass index (BMI) toolkit topics of interest.
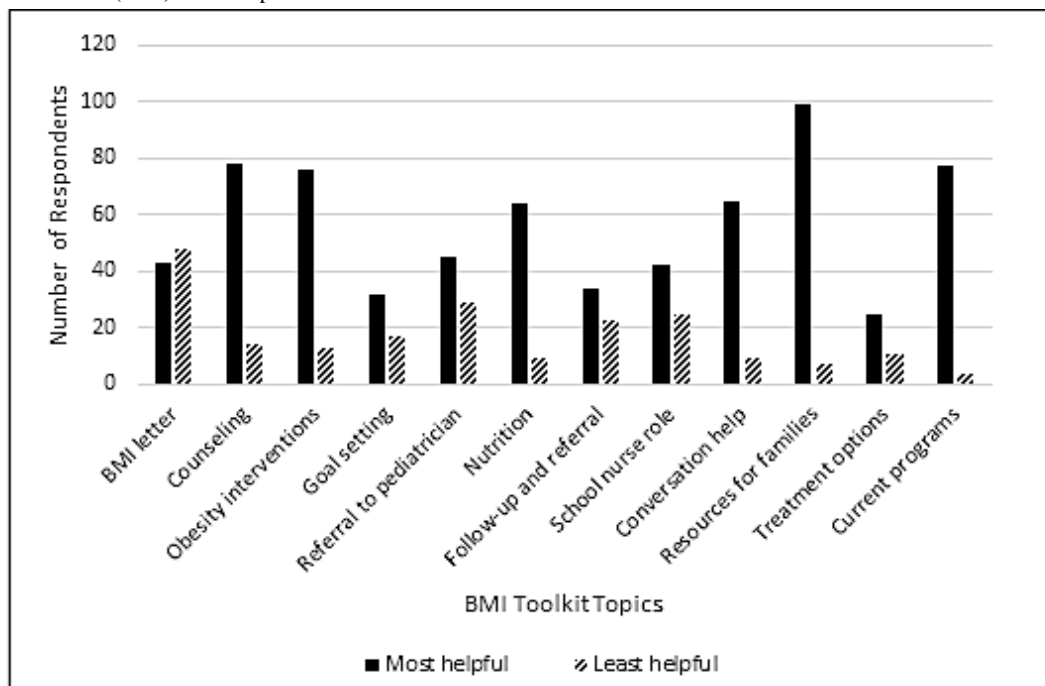
Table 1. Keywords and survey responses.

Keyword	Original survey response
BMI[a] letter	BMI screening letter
Counseling	Counseling and motivating overweight children and families
Obesity interventions	Current resources in obesity interventions (school-based)
Goal setting	Goal setting and follow up
Referral to pediatrician	Growth screening referral to pediatricians
Nutrition	Nutrition messaging
Follow up and referral	Protocol for ongoing follow up and referral
School nurse role	Role of the school nurse
Conversation help	How to talk about BMI with students and their parents
Resources for families	Tools for students and parents on nutrition and physical activity
Treatment options	Treatment options
Current programs	Wellness and weight management programs

[a]BMI: body mass index.

Participants

Nurse participants (N=210), with a 42% (210/500) response rate, represented 208 school districts from urban (14/208, 6.7%), suburban (102/208, 49.0%), and rural (92/44.2%) locales within the state of PA. This is a representative sample of PA schools according to data obtained from "The Center for Rural Pennsylvania" [20], which indicates that 47.6% (238/500) of the districts classify as rural and 52.4% (262/500) as suburban or rural. On average, these districts employ between 4 and 19 full-time nurses depending on urban or rural classification, with a mean student population of 4170 students. Depending on the size of the school district, some school nurses serve all grade levels. All grade levels were represented in the sample, including pre-K (46/210, 21.9%), elementary (132/210, 62.9%), and secondary (94/210, 44.7%) levels.

Data Analysis

Descriptive analysis was conducted using means and SDs. Chi-square test of independence was conducted to determine the relationship between BMI screening follow up and interest in BMI toolkit. The BMI screening follow up question "In addition to the BMI letter, how do you follow up with parents whose child screens overweight or obese?" was dichotomized, with responses of "phone call," "email," or "letter home" categorized as "yes" and "do not follow up" categorized as "no."

Results

Body Mass Index Screening and Results

Participants were asked about their current process for notifying parents of BMI screening results. The majority (116/210, 55.2%) send a letter home in the mail, while others (62/210, 29.5%) send a letter home with students. A small number (8/210, 3.8%) said they did not notify parents altogether, and some notify parents electronically, either by email (8/210, 3.8%) or by posting in a parent portal (31/210, 14.8%).

More than one-third (75/210, 35.7%) of the nurses reported receiving BMI screening inquiries from parents; however, under half (35/75, 47%) of those respondents indicated they follow up with parents whose child screens overweight or obese. Overall, the vast majority (182/210, 86.7%) do not follow up with parents whose child screens overweight or obese. The majority of the nurses (150/210, 71.4%) responded they would benefit from a toolkit with resources to assist with communication with parents and children about BMI screenings. Of those that answered no (60/210, 28.6%), most (31/60, 51%)

indicated that the parents are not interested in further discussion, do not take the screening seriously, or feel offended by the results. A significant association between respondent follow up and interest in a BMI toolkit was observed (*P*=.01).

For topics of inclusion in a BMI toolkit, participants were presented with a list of 12 options and asked to select 3 that would be most helpful to them and 3 that would be least helpful to them (Figure 1). The highest-ranking topics for a toolkit included tools for students and parents on nutrition and physical activity (99/210, 47.1%), counseling and motivating overweight children and families (78/210, 37.1%), wellness and weight management programs (77/210, 36.7%), and current resources in obesity interventions (76/210, 36.2%). School nurses prefer to receive this information via quick, one-page fact sheets (90/210, 42.9%), posters (85/210, 40.5%), or electronic or printable toolkit (81/210, 38.6%).

Professional Development Needs

To fully understand the breadth of a school nurse's priorities and development, our survey expanded beyond the topic of BMI screening. The survey asked about which areas of professional development were of interest. High-ranking topics included diabetes management (138/210, 65.7%), communicable disease control (138/210, 65.7%), and health education (135/210, 64.3%). Additionally, 101 respondents provided write-in topics of interest, including mental health or drug abuse (22/101, 21.8%), nutrition or health and wellness (13/101, 12.9%), concussions or sports injuries (10/101, 9.9%), allergies or asthma (8/101, 7.9%), dermatology (6/101, 5.9%), and other miscellaneous topics such as community resources, legal issues, sexual health, and emergency care (Figure 2). For preferred method of information distribution, respondents selected self-paced Web-based course as their top choice (154/210, 73.3%), followed by recorded webinars (128/210, 61.0%), email newsletters (111/210, 52.9%), an annual health summit (69/210, 32.9%), and live webinars (44/210, 21.0%).

Current Practices

Participants were presented with the question "Which of the following are the most common areas of care you provide to your students?" and were asked to respond to 7 different areas of care with either "very common," "somewhat common," or less common." Respondents answered "very common" with the highest frequency to the following areas of care: acute illnesses (205/210, 97.6%), student injury (159/210, 75.7%), and chronic illnesses (132/210, 62.8%).

Figure 2. Areas of professional development of interest to school nurses.
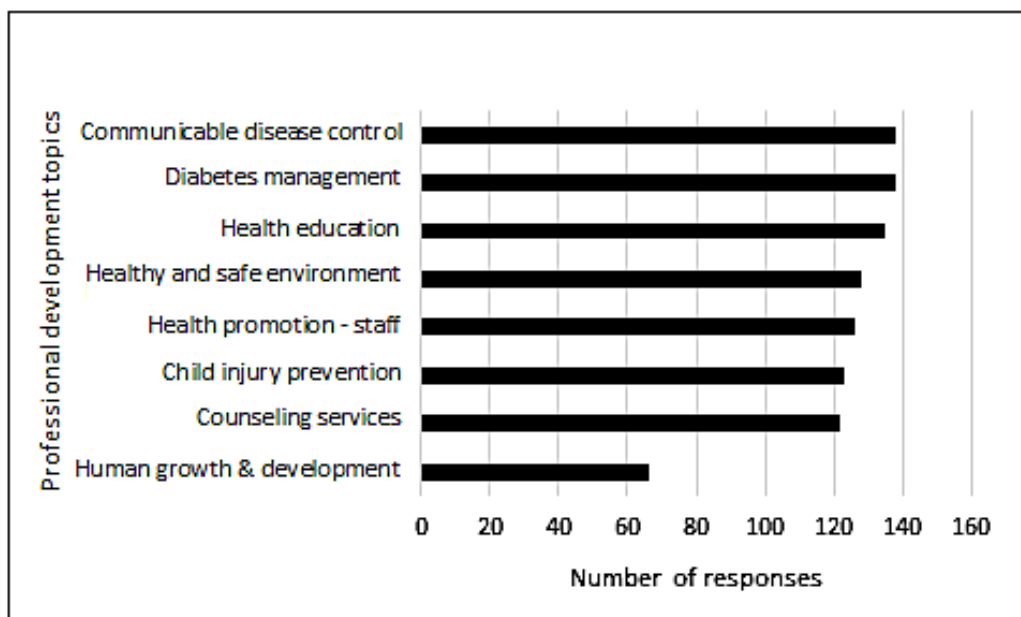
Discussion

Principal Findings

Overall, school nurses follow the PA mandate to complete BMI screenings with related communication to parents via a letter or email. However, though a large portion admitted to receiving inquiries from parents and students regarding BMI screening results, only half of those indicated they follow up on such inquiries. Also, overwhelmingly, the majority (182/210, 86.7%) did not follow up after initial notification of results to parents, indicating school nurses may require additional guidance for providing individualized support to students and parents about childhood overweight and obesity. Furthermore, the majority (150/210, 71.4%) indicated they could benefit from a BMI toolkit with resources to help communicate to parents about BMI screening results and to address weight-related concerns, and most (143/210, 68.1%) saw a need for more professional development opportunities that could aid them in serving as health champions and supporting whole child health in their districts. This study adds to the literature by examining educational needs and interests of PA school nurses, including specific topics of interest for a BMI toolkit and professional development.

Pennsylvania is 1 of only 25 states nationwide that require BMI screenings or weight-related assessments and only 1 of 11 states to require parent notification of results. Though most school nurses in the sample (201/210, 95.7%) responded favorably to communicating initial results to parents, an overwhelming majority (182/210, 86.7%) indicated that further follow up does not occur even if inquiries are received. This indicates a missed opportunity to provide guidance to families regarding lifestyle changes, improved nutrition, and other treatment options that could ultimately improve obesity rates. Though results from this study do not identify the reason for this lack of follow up, it has been well documented that school nurses face many barriers at individual, school, and societal levels that prevent

them from feeling confident and able to address weight-related concerns with their students and families [7].

Many resources to support school nurses with weight-related health already exist, including the NASN obesity toolkit [9], though the specific screening requirements of PA schools may require the development of additional guidance to more adequately support nurses in PA and other states with similar mandates. Most (151/210, 71.9%) of the respondents indicated they would benefit from a BMI toolkit; furthermore, 92.8% (195/210) of those that currently follow up on BMI screening results expressed interest in a toolkit, and 67.6% (142/210) of participants who do not currently follow up on BMI screening results expressed interest in a toolkit, indicating that regardless of current follow-up practices, the majority saw benefit in a toolkit for assistance with parental communication. A significant portion of responses indicated they would not benefit from a BMI toolkit (109/210, 52%) citing lack of parental engagement as a reason. However, other studies have shown parents are primarily concerned with confidentiality and message content (ie, no weight-labeling) and do see benefit in receiving additional follow up to support healthy eating, physical activity, and overall lifestyle change in the home environment [21]. It is important to identify resources for school nurses that would help to improve self-efficacy, knowledge, and confidence when approaching parents with sensitive weight-related information. Better integration of public schools into children's health care can result in systems of care that will more successfully manage and prevent obesity.

To our knowledge, there is little known regarding specific topics of interest that school nurses feel would assist them with improving weight-related communication with their students and families. To fill this gap, this study asked school nurses to identify 3 topics that would be most helpful for inclusion in a toolkit and 3 topics they feel would be least helpful for inclusion from a list of 12 topics. Based upon this information, school nurses prefer resources to counsel children and their families,

educational materials, and information about current weight management programs and preventive interventions. Least helpful is information regarding the screening letter and growth screening referral, indicating that their interest in serving as facilitators of obesity prevention efforts may be greater than their interest in simply providing information. These results further help to identify materials nurses would most likely utilize and promote to beneficially supplement this BMI screening process.

In addition to targeted follow up with families of students who screen as overweight or obese, districts should acknowledge and support school nurses as leaders in overall school health by involving them in obesity prevention efforts and providing them with the resources, both time and financial, to take advantage of professional development opportunities to improve knowledge and self-efficacy. One survey conducted in a sample of 221 school nurses found that most (76%) supported the use of school health services for obesity prevention, but the same study further calls for time and support from the school for success [17]. Although it may seem counterintuitive to provide education and training to health professionals, literature suggests many health care providers, such as school nurses, are not comfortable providing these services [6,7]. A 2015 survey conducted by NASN with 8006 school nurses nationwide determined that "staff and student wellness" was one of the top 6 identified educational needs of school nurses (30.8% of survey) [22]. The survey in this study broke this down further into 8 categories to determine specific topic areas related to staff and student wellness. This data should be utilized by health educators and local agencies to provide meaningful continuing education opportunities that are of greatest interest to school nurses.

Study Limitations

There are several limitations within the study design that should be addressed. For instance, the sample is limited to school nurses in the state of PA. As the data are derived from a single state, findings specific to BMI screenings and supporting resources may not be generalizable across other states; however, 25 states do require BMI screening or weight-related assessments. Furthermore, mandates on professional development requirements and required school screenings vary by state, so survey results from other populations may indicate different priorities. Questions posed in the survey were developed by the research team and Pennsylvania Department of Health after reviewing the more broadly disseminated NASN school nurse survey. Questions were created to supplement results already known regarding barriers and educational needs of school nurses, so this survey did not ask specific questions about barriers but rather inquired about specific topics of interest related to supplemental materials, resources, and trainings to help overcome known barriers. Furthermore, the reliability and validity of the measure is unknown.

Conclusions

Despite juggling a host of student health services, school nurses fill a critical need in each district throughout the state in keeping children's weight at a healthy level. It is the NASN's position that the school nurse supports the physical, mental, emotional, and social health of students and their success in the learning process. Recognizing the important role school nurses play to monitor and promote children's health, schools should involve them as key players in a role beyond anthropometric measurement when it comes to obesity prevention efforts. This goal might include involving them in school-based preventive programs, empowering them to lead initiatives that support whole child health and ensuring opportunities for professional development of interest to them. Nonetheless, the first step in facilitating obesity prevention methods within schools is to provide school nurses with meaningful tools that help facilitate conversations with parents, guardians and caregivers regarding their child's weight status and health through a BMI screening toolkit. In PA schools, where BMI screening with parent notification is required, there is an opportunity to turn the added responsibility of this process into an opportunity to curb rising obesity rates.

Conflicts of Interest

None declared.

References

1. Ogden CL, Carroll MD, Lawman HG, Fryar CD, Kruszon-Moran D, Kit BK, et al. Trends in Obesity Prevalence Among Children and Adolescents in the United States, 1988-1994 Through 2013-2014. JAMA 2016 Jun 07;315(21):2292-2299. [doi: 10.1001/jama.2016.6361] [Medline: 27272581]
2. Lohrmann D, YoussefAgha A, Jayawardene W. Trends in body mass index and prevalence of extreme high obesity among Pennsylvania children and adolescents, 2007-2011: promising but cautionary. Am J Public Health 2014 Apr;104(4):e62-e68. [doi: 10.2105/AJPH.2013.301851] [Medline: 24524514]
3. Basch CE. Healthier students are better learners: high-quality, strategically planned, and effectively coordinated school health programs must be a fundamental mission of schools to help close the achievement gap. J Sch Health 2011 Oct;81(10):650-662. [doi: 10.1111/j.1746-1561.2011.00640.x] [Medline: 21923878]
4. Glickman D, Parker L, Sim LJ, Del Valle Cook H, Miller EA. Accelerating Progress in Obesity Prevention: Solving the Weight of the Nation. In: Institute of Medicine. Washington, DC: National Academies Press; 2012.
5. Division of Adolescent and School Health. US Department of Health and Human Services, Centers for Disease Control and Prevention. 2015. Results from the School Health Policies Practices Study 2014 URL: https://www.cdc.gov/healthyyouth/data/shpps/pdf/shpps-508-final_101315.pdf [accessed 2018-11-30] [WebCite Cache ID 74K8Q0Wuc]

6.   Schroeder K, Travers J, Smaldone A. Are School Nurses an Overlooked Resource in Reducing Childhood Obesity? A Systematic Review and Meta-Analysis. J Sch Health 2016 May;86(5):309-321 [FREE Full text] [doi: 10.1111/josh.12386] [Medline: 27040469]

7.   Steele R, Wu Y, Jensen C, Pankey S, Davis A, Aylward B. School nurses? perceived barriers to discussing weight with children and their families: A qualitative approach. J Sch Health 2011;81:128-137.

8.   Quelly SB. Childhood obesity prevention: a review of school nurse perceptions and practices. J Spec Pediatr Nurs 2014 Jul;19(3):198-209. [doi: 10.1111/jspn.12071] [Medline: 24612520]

9.   Schantz S, Cygan H. School Nurse Childhood Obesity Toolkit. USA: National Association of School Nurses; 2016.

10.   Overweight and obesity in youth in schools-the role of the school nurse: position statement. NASN Sch Nurse 2014 May;29(3):152-153. [doi: 10.1177/1942602X14525569] [Medline: 24937904]

11.   Brener ND, Eaton DK, Lowry R, McManus T. The association between weight perception and BMI among high school students. Obes Res 2004 Nov;12(11):1866-1874 [FREE Full text] [doi: 10.1038/oby.2004.232] [Medline: 15601984]

12.   Koplan JP, Liverman CT, Kraak VI, Committee on Prevention of Obesity in ChildrenYouth. Preventing childhood obesity: health in the balance: executive summary. J Am Diet Assoc 2005 Jan;105(1):131-138. [doi: 10.1016/j.jada.2004.11.023] [Medline: 15635359]

13.   Krebs NF, Jacobson MS, American Academy of Pediatrics Committee on Nutrition. Prevention of pediatric overweight and obesity. Pediatrics 2003 Aug;112(2):424-430. [Medline: 12897303]

14.   Ruggieri DG, Bass SB. A comprehensive review of school-based body mass index screening programs and their implications for school health: do the controversies accurately reflect the research? J Sch Health 2015 Jan;85(1):61-72. [doi: 10.1111/josh.12222] [Medline: 25440454]

15.   Nauta C, Byrne C, Wesley Y. School nurses and childhood obesity: an investigation of knowledge and practice among school nurses as they relate to childhood obesity. Issues Compr Pediatr Nurs 2009;32(1):16-30 [FREE Full text] [doi: 10.1080/01460860802610186] [Medline: 19263291]

16.   National Association of School Nurses, Maughan E, Bobo N, Butler S, Schantz S, Schoessler S. Framework for 21st Century School Nursing Practice: An Overview. NASN Sch Nurse 2015 Jul;30(4):218-231. [doi: 10.1177/1942602X15589559] [Medline: 26118435]

17.   Kubik MY, Story M, Davey C. Obesity prevention in schools: current role and future practice of school nurses. Prev Med 2007 Jun;44(6):504-507 [FREE Full text] [doi: 10.1016/j.ypmed.2007.02.013] [Medline: 17451804]

18.   Quelly SB. Influence of perceptions on school nurse practices to prevent childhood obesity. J Sch Nurs 2014 Aug;30(4):292-302. [doi: 10.1177/1059840513508434] [Medline: 24128859]

19.   Anderson LS, Enge KJ. Education and information for practicing school nurses: which technology-supported resources meet their needs? J Sch Nurs 2012 Oct;28(5):358-369. [doi: 10.1177/1059840512443261] [Medline: 22495742]

20.   Johnson J. The Center for Rural Pennsylvania. 2017. School District Profiles URL: https://www.legis.state.pa.us/index.cfm [accessed 2018-11-13] [WebCite Cache ID 73tObD7og]

21.   Kubik MY, Story M, Rieland G. Developing school-based BMI screening and parent notification programs: findings from focus groups with parents of elementary school students. Health Educ Behav 2007 Aug;34(4):622-633. [doi: 10.1177/1090198105285373] [Medline: 16740517]

22.   Mangena A, Maughan E. The 2015 NASN School Nurse Survey: Developing and Providing Leadership to Advance School Nursing Practice. NASN Sch Nurse 2015 Nov;30(6):328-335. [doi: 10.1177/1942602X15608183] [Medline: 26515570]

Abbreviations

BMI: body mass index
NASN: National Association of School Nurses
PA: Pennsylvania

Original Paper

Health Information on Firefighter Websites: Structured Analysis

Mostin A Hu[1*]; Joy C MacDermid[1,2,3*], PhD; Shannon Killip[1*], BSc (Hons); Margaret Lomotan[1*], BA; FIREWELL[4*]

[1]School of Rehabilitation Science, McMaster University, Hamilton, ON, Canada

[2]School of Physical Therapy, Western University, London, ON, Canada

[3]Roth McFarlane Hand and Upper Limb Centre, St. Joseph's Hospital, London, ON, Canada

[4]Firefighter Injury/illness Remediation Enterprise: Work-participation that Enables Life & Livelihood, Hamilton, ON, Canada

[*]all authors contributed equally

Corresponding Author:
Joy C MacDermid, PhD
School of Physical Therapy
Western University
Elborn College, Room 1440
1201 Western Road
London, ON, N6G 1H1
Canada
Phone: 1 519 661 2111 ext 88912
Fax: 1 519 661 3866
Email: jmacderm@uwo.ca

Abstract

Background: Owing to the fact that firefighters have unique health risks, access to firefighter-specific internet-based health information is a potential mechanism for achieving better health and work outcomes.

Objective: The objective of our study was to identify the amount and nature of health information resources available on Canadian firefighter-specific websites and the extent to which resources are consistent across websites as a surrogate indicator of diffusion of information.

Methods: A search of health resources on firefighter websites (union and employer) for all Canadian provinces, major cities and a subset of smaller cities, and the International Association of Fire Fighters (IAFF) website was conducted on Google (July 2017). Content was identified and classified based on the type of resource, health focus, and location. The quantity and nature of the resources were summarized using descriptive statistics.

Results: Among all (N=313) websites reviewed, 41 websites had health information with a cumulative total of 128 resources that addressed firefighter mental (59/128, 46.1%), physical (43/128, 33.6%), and work health (26/128, 20.3%). The highest density of information was found on international and national websites (13 resources per website) and the least on local websites (1 resource per 7 websites). Three provinces (Ontario, Québec, and British Columbia) hosted 81% (65/80) of the provincial, territorial and local resources. General mental health (20/59, 34%), posttraumatic stress disorder (14/59, 24%), and suicide (14/59, 24%) were the most prevalent topics within the mental health resources, whereas half (21/43, 49%) of all physical health resources were on cancer. No resources from Northern Canada were found. Musculoskeletal health was not mentioned in any of the resources identified. There was minimal cross-linking of resources across sites (only 4 resources were duplicated across sites), and there was no clear indication of how the content was vetted or evaluated for quality.

Conclusions: There was wide variation in the amount and type of information available on different firefighter websites with limited diffusion of information across jurisdictions. Quality evaluation and coordination of resources should be considered to enhance firefighters' access to quality health information to meet their specific needs. Mental health and cancer information aligned with high rates of these health problems in firefighters, whereas the lack of information on musculoskeletal health was discordant with their high rate of work injury claims for these problems.

XSL•FO
RenderX

Introduction

Firefighting involves complex and uncontrollable situations during fire suppression, rescues, and medical calls. These risks carry many inherent risks to mental and physical health [1]. Notable occupational hazards include exposure to carcinogens [2-4], smoke [5], high physical demands [6], and traumatic incidents [7]. The three main areas of firefighters' health concerns are cancer [1,2], mental health issues [8-10], and musculoskeletal (MSK) injuries or cumulative trauma [11].

MSK injuries and cumulative disorders (MSKD) are not usually deadly; however, they present a predominant barrier to healthy work role participation. MSKD are the most common injuries and causes of firefighter work loss claims [1,6,12] with overexertion accounting for one-third of firefighter injury claims [13]. A study conducted by Frost and colleagues in 2013 found that MSK sprains and strains accounted for 64% of injuries reported in the Calgary Fire Department [6]. Tyakoff and colleagues [14] found that 44% of all firefighter injuries occurred on the fire scene and most were due to falls, contusions, and smoke inhalation. We found that almost half of firefighters have MSK pain localized to at least one body site [11].

A recent systematic review conducted by Jones looked at the relationship between duty-related trauma exposure and increased risk for mental health problems in first responders. The team found the accumulation of repeated traumatic exposures from calls to be a key risk factor for posttraumatic stress disorder (PTSD), depression, suicide, alcohol use, and sleep disturbances [7]. Firefighters from the United States and Canada [15] demonstrated PTSD prevalence rates of 22% and 17%, respectively, and associated the development of PTSD symptoms to traumatic incident exposures. "Dead on arrival" and "serious injury accidents" are common exposures among Canadian and US firefighters which may have cumulative effects, although the mechanisms are complex and multifactorial [16]. Firefighter-specific health information that could be reviewed online is particularly salient because firefighters have unique exposures and work within a male-dominated environment, where open discussion of mental health problems is challenging.

Another major area of health concerns for firefighters is cancer, with a meta-analysis demonstrating increased risk of multiple cancers [17], including testicular cancer, melanoma, brain cancer, esophageal cancer, and prostate cancer [18]. Owing to the fact that the increased risk of cancer is thought to be at least partially due to toxic exposures in the line of duty, awareness, prevention strategies, screening, and early management are important dimensions that can be supported by targeted health information.

Research from the Pew Internet Project conducted in 2013 discovered that 72% of internet users in the United States went online for health information [19]. Most firefighters have periods of down time and web access at their fire stations, making the web a feasible source of health information and independent learning. Further, because firefighters resist acknowledging the need for psychological help [20], the availability and anonymity of seeking health information online is especially attractive for health conditions that might have negative connotations [21].

Despite extensive research being conducted on the effects of occupational hazards on the physical and mental health of firefighters, very little attention has been paid to studying the access firefighters have to health information and how this information contributes to injury prevention and treatment. Web-based interventions for stress and anxiety management for the general population have been developed and evaluated [22,23]. However, these are unlikely to meet the needs of firefighters, given the unique and persistent exposures to critical incidents and high physical demands of urgent tasks [24,25].

Because we know that firefighters have unique information needs and are most likely to trust information coming from firefighter-specific websites, it would be helpful to understand their current access as a foundation for planning future improvements. The purpose of this study was to identify and characterize health information resources available on Canadian firefighter websites. A secondary purpose was to determine the extent to which the resources were concordant across sites as an indicator of diffusion of information across fire services.

Methods

Context

The population of Canada is 35,151,728, according to the 2016 Canada Census [26]. Canada is a large country geographically with a small number of large cities and large rural areas with low population. The number of professional firefighters in Canada was estimated to be 26,000 in 2016, which represents 0.72 firefighters per 1000 residents [27]. The number of volunteer firefighters was estimated to be 126,650 or 3.49 per 1000 residents [27].

Identifying the Research Question

This structured search explored the nature and extent of available resources pertaining to the physical and mental health of firefighters by searching fire department and firefighter union websites across Canada. Relevant websites were defined as any webpages on department or union websites that contained information on the prevention, treatment, or any course of action targeting physical and mental health injuries.

Identifying the Relevant Websites

Search Strategy

Google was used as the sole search engine for searches to explore resources on published fire department and firefighter union websites because this is the search engine most likely to be used by firefighters. A 3-layered search strategy (national, provincial or territorial, and local) was implemented to garner the widest scope possible while maintaining depth. In addition, because Canada is a subdivision that falls under the umbrella of the parent organization, the International Association of Fire Fighters (IAFF), the IAFF website was also included. For the national and provincial or territorial level searches, the following 3 types of unions or associations were explored: fire chiefs, professional firefighters, and volunteer firefighters. All searches were performed in July 2017.

The international and national search involved the IAFF, the Canadian Volunteer Fire Services Association, and the Canadian Association of Fire Chiefs websites. Once the association website was found, a detailed search of each page and any links on the website was conducted to identify any health-related information resources. Although many websites had information in a limited number of hierarchical levels, the IAFF website was more layered and required more extensive exploration of multiple levels to access all of the information on the website. A similar search occurred on the provincial and territorial level. Here we used Google to find the fire chiefs, professional firefighters, and volunteer firefighters association websites for the 10 provinces and 3 territories in Canada and then conducted a detailed search of each page of the website.

The initial search strategy and data extraction were developed using Ontario websites because Ontario is the largest province and has the largest proportion of the country's population. Thus, we expected that Ontario would have the greatest number of websites. We searched for each of the 444 municipalities in Ontario, as determined by the Association of Municipalities Ontario [28]. Upon further examination of the major cities in Ontario, the search strategy was revised to look at all "single tier" municipalities (eg, Toronto, Hamilton) and "upper tier, regional" municipalities (eg, the City of Mississauga, part of the Peel region). For the rest of the municipalities in Ontario, only the major cities under counties and districts were examined (see Multimedia Appendix 1 for a detailed list of Ontario municipalities included in the search). This strategy was then extended to other provinces, although each province had a different method of organizing its municipalities. To carry a similar search methodology across provinces, the search focused on the 20 most populated cities in each province (other than Ontario), as determined by the 2016 Canadian census [26]. In provinces with regional firefighter associations encompassing smaller municipalities (New Brunswick), the regional association websites were also included in the search (see Multimedia Appendix 1 for a full list of municipalities included in the search).

Search Terms

For the international and national level search, the following search terms were used: IAFF, IAFF Canada, Canadian Association of Fire Chiefs, and Canadian Volunteer Firefighters Association. For the provincial and territorial level search, the following search terms were used: province or territory fire chiefs association, province or territory professional firefighters association, and province or territory volunteer firefighters association. For the local level search, once the municipalities to be included were chosen (see Multimedia Appendix 1), the following search terms were used to find websites: city or town fire chiefs association and city or town firefighters association.

Selecting Websites

Inclusion and Exclusion Criteria

Official websites in English or French affiliated with a fire department or firefighter association were included. Resources were included if the firefighter website provided a link to the resource. Additional unlinked resources on the secondary website were not counted or classified. Websites under construction or inaccessible at the time of the search and social media accounts of the fire department or association were not included. This study also excluded links on firefighter websites that led to the homepages of other firefighter associations to avoid double-counting.

Charting the Data (Health Resources)

The team developed the initial data extraction sheet and revised it based on piloting. The coding was established to classify what health problem the resource addressed, focusing on the two domains of health (physical and mental health). Because some resources were generically related to work health that could cross both physical and mental health domains, this category was added. Iterative consultation with team members was used to check consistency and subgrouping of codes. Data from the IAFF website were coded separately because it is a unique international body that has a specified mandate for firefighter health and safety. The format of the resources was classified into, for example, checklists, apps, and forums. A supplementary table (see Multimedia Appendix 2) for tools was created to collect additional information on the tools' formats, health foci, and intended purposes.

Collating, Summarizing, and Reporting Results

Given the large numbers of resources and codes, the data were further synthesized to characterize the major types of information available. A descriptive numerical summary (frequency distributions or percentages) was used to describe the overall number of websites included, type of websites found, types of resources, topics or focus of the resource, and location of the website publication. Summary tables were used to describe the health information resources based on type, format, and content area. The differences in the total numbers or percentages are due to rounding.

Results

General Findings

Overall, we found large variability in the type of health information available and the density of the information on different websites. None of the websites mentioned a process for quality appraisal or relevance vetting by firefighters to determine which type of information was posted.

Distribution of Health Resources

It was found that despite the relatively large number of associations and departments having websites (N=313), only 41 of the 313 websites (13.1%) contained health information resources relating to firefighters' mental, physical, and work health. From these 41 websites, 128 resources were found with almost half (59/128, 46.1%) focused on mental health, a third (43/128, 33.6%) focused on physical health, and the rest (26/128, 20.3%) focused on work health, as seen in Figure 1.

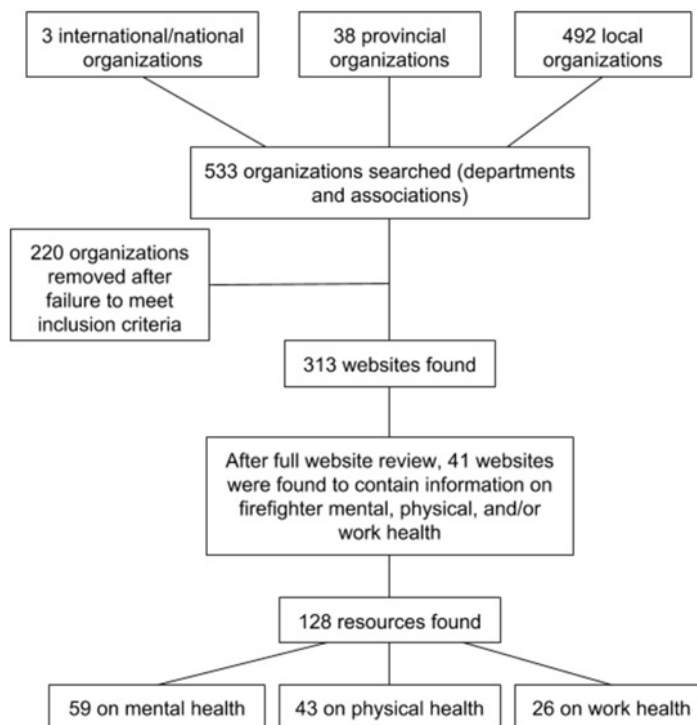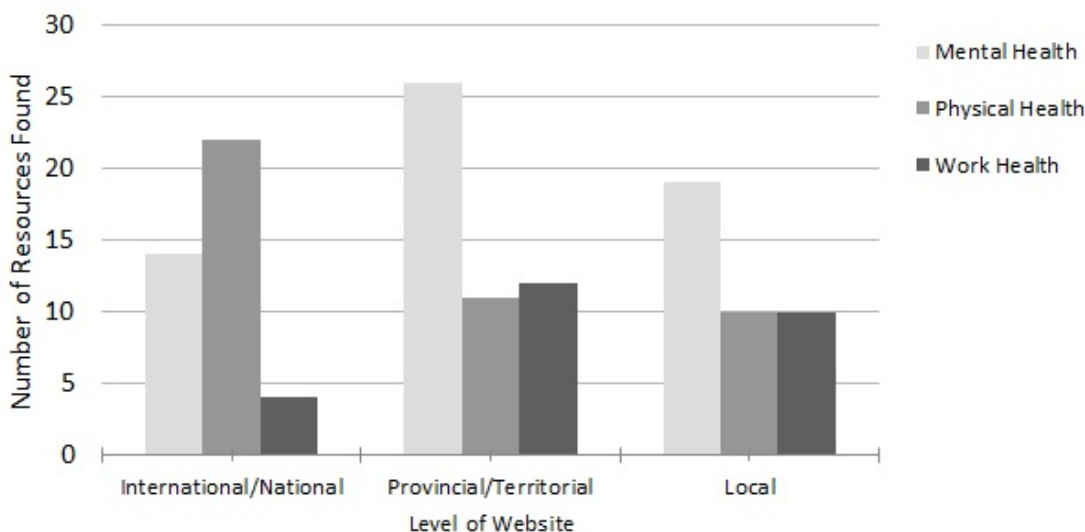**Figure 1.** Process of website resource selection.

Figure 2. Distribution of health resources on different levels of firefighter websites.

The international and national, provincial, and local firefighter websites each contained approximately one-third of the resources found. The IAFF, national fire chief, and firefighter websites accounted for more than half (22/43, 51%) of all the resources found on physical health. The provincial and territorial and local firefighter websites focused on mental health; approximately half of all resources found on these websites addressed topics such as PTSD, critical incident stress, and suicide, as seen in Figure 2 and Table 1.

The resources identified in this study were predominantly found on firefighter websites from Central Canada (Québec and Ontario) and the West Coast (British Columbia). Of the total number (n=80) of resources found on the provincial and territorial and municipal level, 65 were found in these 3

provinces; of these, 42 resources were from Ontario and Québec (Central Canada) and 23 from British Columbia (West Coast). In all the regions in Canada, mental health was the most common focus on firefighter health resources. No resources of any kind were found from the firefighter websites from the Canadian Territories in the North (Yukon, Northwest Territories, Nunavut), although fire department and association websites were identified, as seen in Figure 3.

Format and Topic of Focus

Resources on mental health (n=59) varied in format, most commonly existing as information sheets (13/59, 22%), information packets with multiple pages (9/59, 15%), and manuals (9/59, 15%; defined as larger and more structured

publications with a table of contents). Articles, videos, and links to website home pages were less common, each contributing 10% (6/59) to the total number of resources identified. Many (20/59, 34%) of the mental health resources were general in topic, and there was a relatively even split between more specific topics addressed, as seen in Table 1. Resources on physical and work health were equally diverse in format, as seen in Tables 2 and 3.

A common format for physical health resources (n=43) was links to webpages (15/43, 35%), whereas a large portion of work health resources (n=26) were forms (10/26, 38%), often relating to the administrative processes involved in work injury and return-to-work (eg, workplace safety course registration; work

injury or exposure claims; health insurance claims; physician statements; and return-to-work or work accommodation to track injured worker's progress). Predominant topics addressed in physical health resources (n=43) were cancer (21/43, 49%) and health hazards (10/43, 23%). Work health resources (n=26) were most commonly on safety (12/26, 46%) and protocols (7/26, 26%).

Density and Variation of Resources

The density of resources was highest on international and national level websites with 40 health information resources found on 3 websites (approximately 13.3 resources per website). Provincial websites averaged around 2 resources and websites with 49 resources found on 24 websites.

Table 1. Format and topic of focus for mental health resources (n=59).

Mental health resources	n (%)
Format	
Sheet	13 (22)
Manual	9 (15)
Info packet	9 (15)
Other	9 (15)
Article	6 (10)
Homepage	6 (10)
Video	6 (10)
Poster	1 (3)
Topic of Focus	
General	20 (34)
Posttraumatic stress disorder	14 (24)
Suicide	14 (24)
Critical incident stress	11 (18)

Figure 3. Distribution of health resources in different regions of Canada.
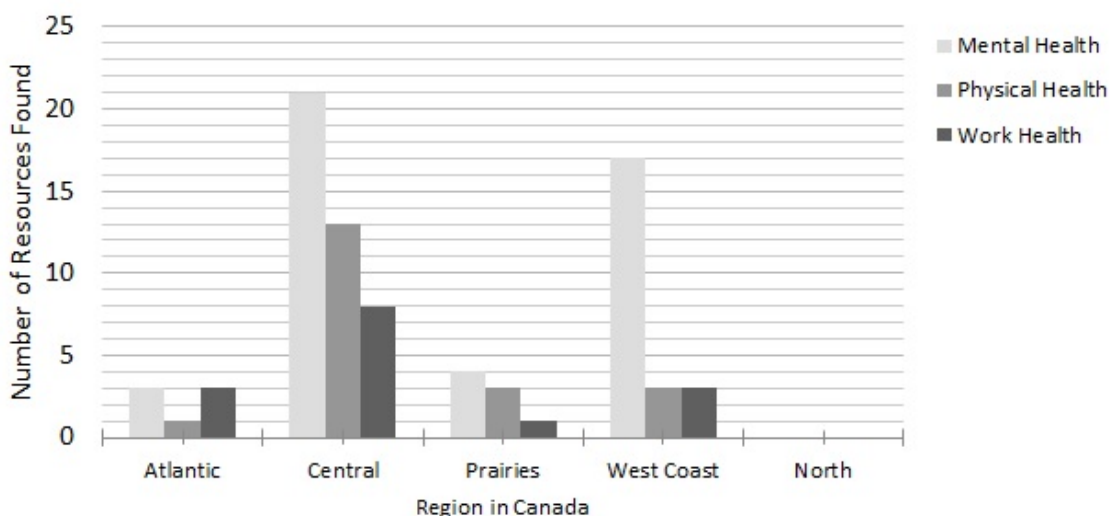
Table 2. Format and topic of focus for physical health resources (n=43).

Physical health resources	n (%)
Format	
Homepage	15 (35)
Sheet	11 (26)
Report	7 (16)
Video	8 (19)
Manual	2 (5)
Info packet	1 (2)
Other	3 (7)
Topic of Focus	
Cancer	21 (49)
Health hazards	10 (23)
Fitness	6 (14)
Other	3 (7)
Burns	2 (5)
Reproductive	1 (2)

Table 3. Format and topic of focus for work health resources (n=26).

Work health resources	n (%)
Format	
Form	10 (38)
Info packet	5 (20)
Sheet	4 (15)
Homepage	2 (8)
Poster	2 (8)
Video manual	1 (4)
Article	1 (4)
Topic of Focus	
Safety	12 (46)
Protocol	7 (26)
Accommodation	3 (12)
General	2 (8)
Prevention	1 (4)
Disability	1 (4)

Conversely, across 282 local websites, only an additional 39 health information resources were located (1 resource per 7 websites), indicating a much smaller density of information on these websites. Most of the 128 resources found were unique. Only 4 duplicate resources were found, which were mainly web links to other organizations providing relevant information (eg, asbestos.com and mesotheliomagroup.com).

Tools

In addition to cataloging information resources, we classified resources as tools if they were structured to implement or execute specific actions or serve as a platform for engagement (eg, checklists, surveys, and mobile applications). Only a small number of tools specific to firefighters were identified. The tools (n=12) found on local websites mostly focused on physical health (10/12, 83%) with the most common format being a members' forum (7/12, 60% of all local tools). Overall, the most common types of tools were generic or focused on work health.

XSL•FO
RenderX

Figure 4. Distribution of tools on different levels of firefighter websites.
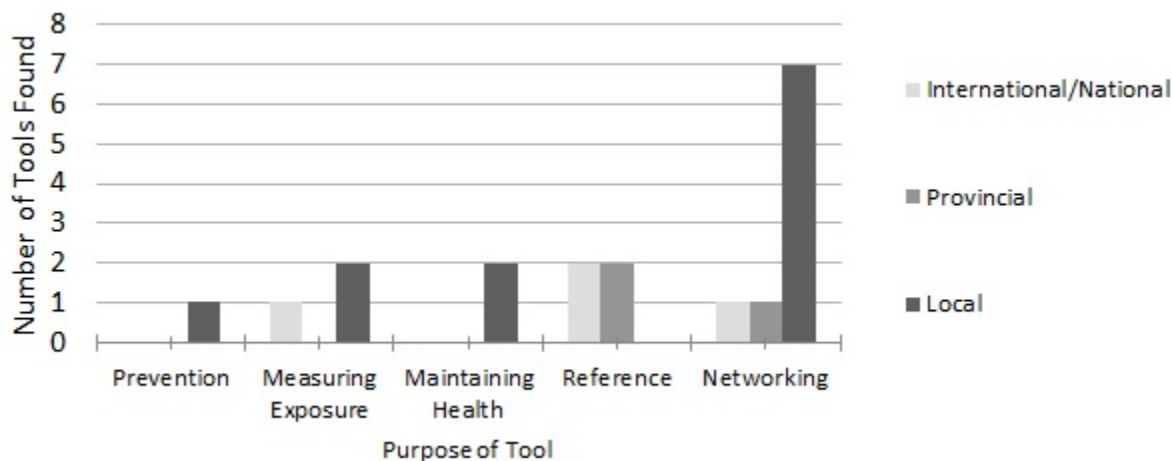
Work tools included operational resources that firefighters needed to implement workplace health and safety. The tool resources were provided in a variety of formats from forums to checklists and information packets. The majority (4/7, 57%) of national and provincial tools (n=7) were formatted as references, as seen in Figure 4.

Trends

The IAFF website was a major source of health information specific to firefighters. In addition to health information, it provided information about in-person training courses and programs on firefighter health and safety. Provincial and local websites frequently linked to IAFF health and safety resources, as seen in Multimedia Appendices 3 and 4 in tables on IAFF mental and physical health resources.

The British Columbia (BC) Professional Fire Fighters' website (Multimedia Appendix 5) was especially rich in mental health resources because it had a dedicated page to mental health resources that covered a wide focus, including critical incident stress management, suicide awareness, and maintaining mental health. Apart from BC's mental health resources, Ontario had the most resources on the provincial and local levels. It had 4 provincial websites (fire chief, professional, volunteer, and women). Overall, 17 out of 102 local websites contained information and 42 out of 128 resources found were from Ontario.

In general, health information resources were most often on union, rather than employer, websites. This was particularly apparent at the local level, where almost all resources found were on firefighter association websites and very few on the municipal fire department websites.

Discussion

Principal Findings

This study suggests that the process by which health information is provided to firefighters through their websites is highly variable and informal in terms of a process for vetting or sharing of information. This high variability might indicate that champions for a cause, expertise in different areas, experiences within the local environment, and other factors may influence what health resources are posted on websites. The one common link for health information is IAFF, which contains the most firefighter-specific health information and resources. However, the uniqueness of information with limited crossovers between different websites indicates that there is little diffusion of information across contexts.

The density of information and provincial differences indicates that resources may be a substantial issue in determining what is shared because smaller organizations and poorer provinces had the fewest health resources. Owing to the fact that the duplication of health resource development would be a poor utilization of public dollars, greater sharing of health resources across websites or a single repository of information might facilitate effective dissemination of firefighter health information resources. Because the majority of firefighters in Canada are volunteer firefighters, with few resources and little time for face-to-face knowledge sharing, sharing of internet resources would also enhance equity in access to information. Web-based information can be shared easily across contexts, including those without the resources to create their own websites.

Given the high rate of injuries and illnesses that firefighters incur, optimizing their access and use of health information are important. Firefighter-specific websites exist for many functions beyond sharing health information. This is reflected by the fact that only 13.1% (41/313) contained any health information. Many websites, particularly city department websites, were intended for public education, providing information on the fire department, fire prevention, and outreach programs in the community. Because these are all important aspects of the fire service mandate, it is important to recognize the important competing demands that exist in terms of website development and maintenance. The development, vetting, interpretation, and application of firefighter-specific health information requires substantial resources. Labor union and association websites were more likely to provide resources that are focused directly on firefighters, which is consistent with their mandate in comparison to the fire service that has a primary mandate for public safety.

The types of resources located reflect program development and priorities in different localities. The variation in content across different cities or provinces partially reflects the

difference in program development in those jurisdictions. If those programs diffused into other jurisdictions, we would have expected to find commonly shared resources on different websites. The lack of redundant information across websites indicates limited diffusion and the need for greater dissemination of information.

The fact that almost half (59/128, 46.1%) of all resources we found had a mental health focus was consistent with firefighters' need for support to deal with higher rates of exposures to traumatic events [25]. Their day-to-day job involves situations that the general public may only see once in a lifetime [25]. Owing to this, firefighters have often noted the need for occupation-specific resources adapted for their line of work. In particular, firefighters acknowledge the importance of camaraderie and peer support in dealing with job stressors [29]. The prominence of this content reflects the recent increased awareness of these problems related to public safety personnel [9,10]. Early program development, with respect to mental health in BC (mental health resiliency programs), is reflected in the higher density of information pertaining to that topic. Although not all the elements of the mental health resiliency program may be accessible on firefighter websites, greater focus on a topic within a jurisdiction can cause firefighters to access or share more health information on this topic.

Web-based discussion forums or resources may be particularly valuable in the fire service because a community of practice or virtual support group could be established, where people share common experiences or access contextually relevant information. The advantage of using the internet for sharing mental health experiences is that people can choose to remain anonymous. A study on university students with mental health issues found that about one-third of them used the internet to deal with their mental health problems. The single most common reason cited for valuing Web-based mental health resources was that it provided anonymity [30].

Similarly, our study found that physical health resources often addressed the cancer risks due to exposure to carcinogens. Some of the few linked websites that appeared recurrently were the home pages of mesothelioma information and cancer support organizations. Mesothelioma is a rare, aggressive form of cancer that develops in the lining of the lungs, abdomen, or heart; it has no known cure and has a very poor prognosis. Firefighters are particularly at risk because exposure to asbestos is the major risk factor. A study by Daniels conducted in 2013 found a two-fold increase in malignant mesothelioma mortality and incidence in firefighters compared with that found in the general US population [31]. Given the exposure to multiple carcinogens during fire suppression and clean-up [3-5], it was unsurprising to find such a large number of cancer resources. Further, this is directly aligned with substantial literature demonstrating that firefighters are at an increased risk of different cancers. A large registry study on males with cancer found that being a firefighter was associated with higher rates of testicular cancer, melanoma, brain cancer, esophageal cancer, and prostate cancer [32]. Health information resources that encourage awareness, screening, and protection strategies could be important in mitigating these risks, for example, decontamination procedures have been widely implemented [33].

The lack of resources on MSK health stands in stark contrast to the fact that MSK injuries are among the most common injuries incurred by firefighters [6,14]. Furthermore, the lack of any resources on injury prevention on websites contrasts with the fact that substantial efforts have focused on the creation of generic MSK injury programs [34,35]. Although studies have demonstrated high rates of MSK injuries during fire suppression and training, the websites did not identify any resources specifically designed to reduce these types of injuries. It is possible that some of these resources are integrated into training programs that are delivered face-to-face or in protected sections. However, there is no reason to think that there is preferential "hiding" of MSK health resources behind firewalls in comparison to other health topics. Therefore, the lack of resources about MSK health of firefighters indicates either a research gap in that area or a lack of perceived importance within the fire service. Because MSK problems are often the most common work injury claims within the fire service [1,13,14,36], the lack of health resources in this regard indicates an important gap. The need for health information to prevent or manage MSK problems specific to firefighters is emphasized by a number of factors, which suggest that this will continue to be a substantial problem. The fire service workforce is aging, consistent with our aging workforce. Aging is associated with more MSK problems. Further studies indicate that older firefighters tend to have more MSK problems [11]. Because firefighters often start their careers with good physical health and improve their health during their early training, it can be a challenge, given the variable nature of their physical activity to maintain that high level of fitness [37]. More Web-based resources for active-duty firefighters directed toward maintaining fitness may assist with mitigating this issue. In firefighters, fitness is particularly important for preventing injuries, given the high physical demands of the occupation and the unpredictable tasks performed during fire suppression [38].

Some aspects of health that were notably absent included sleep and nutrition which are potentially important to both physical and mental health. This is a notable absence because shift work potentially impacts both of these health behaviors.

This study provides an indirect insight into the knowledge translation processes that are currently taking place with respect to health information for firefighters. Effective dissemination would be reflected in concordance or as duplicate information across websites. However, we found only four duplicates across the websites from which we extracted information. This lack of overlap is a proxy for limited dissemination. Although firefighters might be motivated to go to the IAFF website for information, recognizing it as an umbrella organization, it is unlikely that they would go to multiple websites for information. A generalized and central firefighter-specific website that features structured information sharing processes might improve dissemination of information.

Another observation was the lack of transparent processes for how information was included on websites. Both the scientific quality and the relevance of the information should be important considerations. Although these might have been considered, it was unclear to what extent this was true. Therefore, the quality of the information that is being provided to firefighters is

uncertain. Of course, some resources did contain references. However, this was variable. Firefighter-specific health information is dependent on firefighter-specific health research. Generally, research that focuses on single occupations presents challenges for researchers and may not have the capacity that disease-specific health research does. This suggest that health researchers who focus on firefighter-specific topics have a heightened obligation to ensure that they engage in knowledge translation to optimize the dissemination and implementation of their work.

Regionally, Ontario was the province with the most resources available on both the provincial and local levels. This is likely because Ontario has the greatest number of firefighters in Canada with over 11,000 professional firefighters as members of the Ontario Professional Fire Fighters Association [39]. Another major contributor to the resources identified in this study was BC. This reflects substantial investments in the creation of mental health resources, despite the fact there are significantly fewer firefighters (4000) in the BC Professional Fire Fighters Association [40]. Larger organizations could be expected to have greater capacities for developing health information resources and investing in regular website updates to share those resources. Disparities in health can arise from unequal access to health information. The majority of firefighters in Canada are volunteers, and volunteer associations usually have limited resources. Because more rural areas of the country have more volunteer firefighters and small taxation bases, the likelihood of differential information access is high. Dissemination of well-developed health information in more associations has a potential to improve the health of the entire firefighter work force. Many firefighter association websites contained login-protected resources. The locked away resources potentially create a divide between volunteer and professional firefighters because many websites were members-only for firefighters who were part of the union or association. Most of the association websites identified in this study were for professional firefighters only, which means that volunteer firefighters were less likely to be able to access information on their mental, physical, and work health.

Strengths and Limitations

This paper provided a snapshot of the Web-based health information resources from a large number of websites that may be accessed by firefighters across multiple contexts and provinces in Canada. We used a structured process to extract available resources and classify them. Although this information provides insights into the health resources available, focus, knowledge translation outcomes and gaps in the field, the nature of the work also has substantial limitations that must be considered when interpreting the results.

First, we searched only Web-based information; we recognize that this represents a fraction of the health information available to firefighters. Firefighters engage in substantial training, conferences, and other learning opportunities that likely provide additional information. Second, some of the websites had password-protected areas which we were unable to access and these areas may have contained additional resources. Third, we did not evaluate the quality of the resources.

A further limitation is that we analyzed websites and content separately, not accounting for the interdependence between the two. Further, we could not ascertain from reading websites what processes for development and vetting of resources might actually be present across different organizations. We also do not know whether firefighters are not disseminating the information or just not posting the available information on their websites. There is a possibility that firefighters shared information on social media accounts which were not included in this study.

In Ontario, only the single districts, regions, and major cities were searched, and in other provinces, the 20 most populous cities' websites were searched; therefore, there is a chance that some resources from smaller cities or towns were missed. Although possible, this is unlikely because very little health information was found on websites for the smaller cities (eg, beyond the 10th most populated city in each province). In many of the smaller provinces, such as Prince Edward Island, New Brunswick, and Nova Scotia, there were fewer than 20 cities that existed within the entire province. Thus, all cities were included. This provided a comprehensive overview of the resources available to firefighters on fire department and association websites. Another limitation is the possibility of misclassification because a majority of the coding was conducted by a single person. To limit misclassification, the ongoing coding was examined regularly with another investigator and random rechecks of the classification were performed. Nevertheless, in classifying information, it is inevitable that the complexity of the information is lost and classifications are oversimplification of the information that was provided. An additional limitation arises from the fact that we only searched firefighter-specific websites and important information specific to firefighters could have been contained on other websites. We did this because we heard from firefighters that they tended to go to their local websites or provincial or IAFF for information. However, there are examples where substantial contributions to public safety health that would be relevant to firefighters would be contained elsewhere, for example, the Canadian Institute for Public Safety Research and Treatment recently published mental health resources (Web-based self-screening tools and research papers) for first responders [41]. The Vancouver firefighters have a critical incident stress management website [42] that is not linked to their fire department or association websites and contains information on a mental health resiliency program that was developed by Vancouver Fire and Rescue Services in collaboration with the Canadian Mental Health Association-Vancouver Fraser [43].

Implications

The key recommendations arising from this study are the need for better coordination and sharing of health information resources specific to firefighters. Greater linkage between firefighter-specific health research and health information resources for firefighters is needed. We suggest that formal processes for evaluating what health information is shared with firefighters, widespread dissemination of lay summaries of firefighter-specific research, collaboration between researchers and firefighters, and sharing of information on formal evaluation

of existing health programs would all be important contributions to effective knowledge translation. An emerging awareness of the important challenges facing firefighters with respect to firefighter health and programs was evident for mental health, whereas substantial gaps in MSK health resources were also evident. Given that some associations had few information resources, the need for greater sharing to decrease health inequities across regions is clear. Knowledge repositories and inventories that document firefighter-specific health information and those that are broadly shared might leverage existing knowledge for greater health impacts. Future research should focus on developing and evaluating the effectiveness of firefighter-specific health information resources and tools and knowledge translation plans should include widespread dissemination through firefighter websites.

Conflicts of Interest

None declared.

Multimedia Appendix 1

List of Ontario municipalities included in search and list of municipalities in Canada included in search (excluding Ontario).

[PDF File (Adobe PDF File), 49KB - ijmr_v7i2e12_app1.pdf]

Multimedia Appendix 2

Tools table.

[PDF File (Adobe PDF File), 86KB - ijmr_v7i2e12_app2.pdf]

Multimedia Appendix 3

The International Association of Fire Fighters (IAFF) mental health resources table.

[PDF File (Adobe PDF File), 48KB - ijmr_v7i2e12_app3.pdf]

Multimedia Appendix 4

The International Association of Fire Fighters (IAFF) physical health resources table.

[PDF File (Adobe PDF File), 55KB - ijmr_v7i2e12_app4.pdf]

Multimedia Appendix 5

British Columbia mental health resources table.

[PDF File (Adobe PDF File), 59KB - ijmr_v7i2e12_app5.pdf]

References

1.   Ramsden R, Smith J, Turcotte K, Garis L, Kunz K, Maxim P, Thomas L, Pike I. 2018. Determinants of Injury and Death in Canadian Firefighters A Case for a National Firefighter Wellness Surveillance System Internet URL: https://cjr.ufv.ca/determinants-of-injury-and-death-in-canadian-firefighters-a-case-for-a-national-firefighter-wellness-surveillance-system/ [accessed 2018-05-04] [WebCite Cache ID 6zANDSjO7]
2.   Tsai RJ, Luckhaupt SE, Schumacher P, Cress RD, Deapen DM, Calvert GM. Risk of cancer among firefighters in California, 1988-2007. Am J Ind Med 2015 Jul;58(7):715-729 [FREE Full text] [doi: 10.1002/ajim.22466] [Medline: 25943908]

3.   Keir JLA, Akhtar US, Matschke DMJ, Kirkham TL, Chan HM, Ayotte P, et al. Elevated Exposures to Polycyclic Aromatic Hydrocarbons and Other Organic Mutagens in Ottawa Firefighters Participating in Emergency, On-Shift Fire Suppression. Environ Sci Technol 2017 Nov 07;51(21):12745-12755. [doi: 10.1021/acs.est.7b02850] [Medline: 29043785]

4.   Fent KW, Alexander B, Roberts J, Robertson S, Toennis C, Sammons D, et al. Contamination of firefighter personal protective equipment and skin and the effectiveness of decontamination procedures. J Occup Environ Hyg 2017 Oct;14(10):801-814. [doi: 10.1080/15459624.2017.1334904] [Medline: 28636458]

5.   Fabian T, Borgerson JL, Kerber SI, Gandhi PD, Baxter CS, Ross CS, Lockey JE, Dalton JM. 2010. Firefighter Exposure to Smoke Particulates URL: https://www.ul.com/global/documents/offerings/industries/buildingmaterials/fireservice/WEBDOCUMENTS/EMW-2007-FP-02093%20-%20Executive%20Summary.pdf [accessed 2017-11-07] [WebCite Cache ID 6unq6Ounm]

6.   Frost DM, Beach TAC, Crosby I, McGill SM. Firefighter injuries are not just a fireground problem. Work 2015;52(4):835-842. [doi: 10.3233/WOR-152111] [Medline: 26409354]

7.   Jones S. Describing the Mental Health Profile of First Responders: A Systematic Review. J Am Psychiatr Nurses Assoc 2017 May;23(3):200-214. [doi: 10.1177/1078390317695266] [Medline: 28445653]

8.   McFarlane AC. Epidemiological evidence about the relationship between PTSD and alcohol abuse: the nature of the association. Addict Behav 1998;23(6):813-825. [Medline: 9801718]

9.   Carleton RN, Afifi TO, Turner S, Taillieu T, Duranceau S, LeBouthillier DM, et al. Mental Disorder Symptoms among Public Safety Personnel in Canada. Can J Psychiatry 2018 Jan;63(1):54-64. [doi: 10.1177/0706743717723825] [Medline: 28845686]

10.   Carleton R, Afifi T, Turner S, Taillieu T, LeBouthillier D, Duranceau S, et al. Suicidal ideation, plans, and attempts among public safety personnel in Canada. Canadian Psychology/Psychologie canadienne 2018 Feb 08. [doi: 10.1037/cap0000136]

11.   Negm A, MacDermid J, Sinden K, D'Amico R, Lomotan M, MacIntyre N. Prevalence and distribution of musculoskeletal disorders in firefighters are influenced by age and length of service. J Mil Veteran Fam Heal 2017 Dec 22;3(2):33-41.

12.   Taylor NAS, Dodd MJ, Taylor EA, Donohoe AM. A Retrospective Evaluation of Injuries to Australian Urban Firefighters (2003 to 2012): Injury Types, Locations, and Causal Mechanisms. J Occup Environ Med 2015 Jul;57(7):757-764. [doi: 10.1097/JOM.0000000000000438] [Medline: 26067214]

13.   Walton SM, Conrad KM, Furner SE, Samo DG. Cause, type, and workers' compensation costs of injury to fire fighters. Am J Ind Med 2003 Apr;43(4):454-458. [doi: 10.1002/ajim.10200] [Medline: 12645102]

14.   Tyakoff A, Garis L, Thomas L. Firefighter Injuries in British Columbia: An Examination of Frequency, Severity, Locational and Temporal aspects. 2015 May. URL: https://tinyurl.com/yc9a8y65 [accessed 2017-11-07] [WebCite Cache ID 6unqX3b3D]

15.   Corneil W, Beaton R, Murphy S, Johnson C, Pike K. Exposure to traumatic incidents and prevalence of posttraumatic stress symptomatology in urban firefighters in two countries. J Occup Health Psychol 1999 Apr;4(2):131-141. [Medline: 10212865]

16.   Murphy SA. Human responses to catastrophe. Annu Rev Nurs Res 1991;9:57-76. [Medline: 1756093]

17.   LeMasters GK, Genaidy AM, Succop P, Deddens J, Sobeih T, Barriera-Viruet H, et al. Cancer risk among firefighters: a review and meta-analysis of 32 studies. J Occup Environ Med 2006 Nov;48(11):1189-1202. [doi: 10.1097/01.jom.0000246229.68697.90] [Medline: 17099456]

18.   Armstrong JJ, Rodrigues IB, Wasiuta T, MacDermid JC. Quality assessment of osteoporosis clinical practice guidelines for physical activity and safe movement: an AGREE II appraisal. Arch Osteoporos 2016;11:6. [doi: 10.1007/s11657-016-0260-9] [Medline: 26759266]

19.   Fox S, Duggan M. 2013. Health online 2013, Pew Research Center's Internet & American Life Project URL: http://www.pewinternet.org/2013/01/15/health-online-2013/ [accessed 2018-05-04] [WebCite Cache ID 6zAOOtTFm]

20.   Hokanson M, Wirth B. The critical incident stress debriefing process for the Los Angeles County Fire Department: automatic and effective. Int J Emerg Ment Health 2000;2(4):249-257. [Medline: 11217156]

21.   Fox S, Rainie L. Vital decisions: how Internet users decide what information to trust when they or their loved ones are sick. Pew Research Center's Internet & American Life Project 2002 May 22 [FREE Full text]

22.   Ashford MT, Olander EK, Ayers S. Finding Web-Based Anxiety Interventions on the World Wide Web: A Scoping Review. JMIR Ment Health 2016 Jun 01;3(2):e14 [FREE Full text] [doi: 10.2196/mental.5349] [Medline: 27251763]

23.   Ryan C, Bergin M, Chalder T, Wells JS. Web-based interventions for the management of stress in the workplace: Focus, form, and efficacy. J Occup Health 2017 May 25;59(3):215-236 [FREE Full text] [doi: 10.1539/joh.16-0227-RA] [Medline: 28320977]

24.   Lourel M, Abdellaoui S, Chevaleyre S, Paltrier M, Gana K. Relationships between psychological job demands, job control and burnout among firefighters. N Am J Psychol 2008;10(3):489-495 [FREE Full text]

25.   Jahnke SA, Poston WSC, Jitnarin N, Haddock CK. Health concerns of the U.S. fire service: perspectives from the firehouse. Am J Health Promot 2012;27(2):111-118. [doi: 10.4278/ajhp.110311-QUAL-109] [Medline: 23113781]

26.   Statistics Canada. Census Profile, 2016 Census. 2017 Feb 08. URL: http://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/index.cfm?Lang=E [accessed 2017-11-07] [WebCite Cache ID 6unsG4Pd4]

27.   Haynes H. Canadian Fire Department Profile, 2014-2016. 2018. National Fire Protection Association Report URL: https://www.nfpa.org/-/media/Files/News-and-Research/Fire-statistics/Fire-service/oscanada.pdf [accessed 2018-05-04] [WebCite Cache ID 6zAVMnexV]

28.   Association of Municipalities of Ontario. Ontario Municipalities. URL: http://www.amo.on.ca/YLG [accessed 2017-11-07] [WebCite Cache ID 6unsAB8Y2]

29.   Tuckey M, Hayward R. Global and occupation-specific emotional resources as buffers against the emotional demands of fire-fighting. Appl Psychol 2011;60(1):1-23. [doi: 10.1111/j.1464-0597.2010.00424.x]

30.   Horgan A, Sweeney J. Young students' use of the Internet for mental health information and support. J Psychiatr Ment Health Nurs 2010 Mar;17(2):117-123. [doi: 10.1111/j.1365-2850.2009.01497.x] [Medline: 20465756]

31.   Daniels RD, Kubale TL, Yiin JH, Dahm MM, Hales TR, Baris D, et al. Mortality and cancer incidence in a pooled cohort of US firefighters from San Francisco, Chicago and Philadelphia (1950-2009). Occup Environ Med 2014 Jun;71(6):388-397 [FREE Full text] [doi: 10.1136/oemed-2013-101662] [Medline: 24142974]

32.   Bates MN. Registry-based case-control study of cancer in California firefighters. Am J Ind Med 2007 May;50(5):339-344. [doi: 10.1002/ajim.20446] [Medline: 17427202]

33.   Ontario Fire Service Section 21 Advisory Committee. Firefighters guidance note: Hygiene and decontamination. 2017. URL: https://tinyurl.com/ycv33rdq [accessed 2018-05-04] [WebCite Cache ID 6zAP4nuyp]

34.   Occupational Health and Safety Council of Ontario. MSD Musculoskeletal Disorders MSD Prevention Guideline for Ontario. 2007. URL: https://www.labour.gov.on.ca/english/hs/topics/pains.php [accessed 2017-11-07] [WebCite Cache ID 6untZeigd]

35.   Stock S, Baril R, Dion-Hubert C, Lapointe C, Paquette S, Sauvage J, et al. Work-related Musculoskeletal Disorders Guide and Tools for Modified Work. 2005 Jan. URL: http://www.irsst.qc.ca/media/documents/PubIRSST/OMRT-En.pdf [accessed 2018-05-04] [WebCite Cache ID 6zAPJFIxc]

36.   Jahnke SA, Poston WSC, Haddock CK, Jitnarin N. Injury among a population based sample of career firefighters in the central USA. Inj Prev 2013 Dec;19(6):393-398. [doi: 10.1136/injuryprev-2012-040662] [Medline: 23504995]

37.   Cornell DJ, Gnacinski SL, Meyer BB, Ebersole KT. Changes in Health and Fitness in Firefighter Recruits: An Observational Cohort Study. Med Sci Sports Exerc 2017 Nov;49(11):2223-2233. [doi: 10.1249/MSS.0000000000001356] [Medline: 28650351]

38.   Smith DL. Firefighter fitness: improving performance and preventing injuries and fatalities. Curr Sports Med Rep 2011;10(3):167-172. [doi: 10.1249/JSR.0b013e31821a9fec] [Medline: 21623308]

39.   Ontario Professional Fire Fighters Association. About Ontario's Fire Fighters. URL: https://www.ontariofirefighters.org/about [accessed 2018-05-04] [WebCite Cache ID 6zAQcaRdt]

40.   British Columbia Professional Fire Fighters Association. What We Do. URL: https://www.bcpffa.net/intro/ [accessed 2018-05-04] [WebCite Cache ID 6zAR9NRtt]

41.   The Canadian Institute for Public Safety Research and Treatment. Anonymous Self-assessment Mental Health Screening Tools. URL: https://ax1.cipsrt-icrtsp.ca/ [accessed 2017-11-07] [WebCite Cache ID 6unuGFtLU]

42.   Vancouver Critical Incident Stress Management. URL: http://www.vancouvercism.com/ [accessed 2018-05-04] [WebCite Cache ID 6zARVsXYs]

43.   Canadian Mental Health Association British Columbia. CMHA to work with BC firefighters to build resilient minds Internet. 2017. URL: https://cmha.bc.ca/news/cmha-to-work-with-bc-firefighters-to-build-resilient-minds/ [accessed 2018-05-04] [WebCite Cache ID 6zASbNCg7]

Abbreviations

BC: British Columbia
IAFF: International Association of Fire Fighters
MSK: musculoskeletal
MSKD: musculoskeletal disorders
PTSD: posttraumatic stress disorder

XSL•FO
RenderX

XSL•FO

RenderX

Original Paper

Trustworthiness, Readability, and Suitability of Web-Based Information for Stroke Prevention and Self-Management for Korean Americans: Critical Evaluation

Mikyoung A Lee[1], RN, PhD; Cha-Nam Shin[2], RN, PhD; Kyungeh An[3], RN, PhD

[1]College of Nursing, Texas Woman's University, Denton, TX, United States

[2]College of Nursing and Health Innovation, Arizona State University, Phoenix, AZ, United States

[3]Department of Adult Health and Nursing Systems, School of Nursing, Virginia Commonwealth University, Richmond, VA, United States

Corresponding Author:
Mikyoung A Lee, RN, PhD
College of Nursing
Texas Woman's University
PO Box 425498
Denton, TX, 76204-5498
United States
Phone: 1 940 898 2424
Fax: 1 940 898 2437
Email: mikyoungLee@twu.edu

Abstract

Background: Websites are common sources of health information to stroke survivors and caregivers for continual management of stroke and its long-term sequelae. The presence of risk factors and mortality rates related to stroke are high in Korean Americans. A vast majority of this group are active Web users and rely on the Web-based information due to lack of insurance and, thus, limited access to long-term stroke care. Thus, it is critical to evaluate existing stroke websites for their trustworthiness, readability, and suitability.

Objective: The objective of our study was to provide a systematic evaluation of stroke-related websites regarding (1) trustworthiness, (2) readability, and (3) suitability for stroke prevention and self-management for Korean Americans.

Methods: We selected a total of 156 websites using search terms "stroke," "CVA," "중풍 (jungpung)," and "뇌졸증 (noejoljung)" on Google and Yahoo. After eliminating duplicates and irrelevant websites (n=116), we evaluated a total of 42 websites (15 in English and 27 in Korean) using the National Library of Medicine's health website's evaluation tool for trustworthiness; Simple Measure of Gobbledygook for readability; and Suitability Assessment of Materials for suitability. All three instruments used the 3-point Likert scale: superior (=2), adequate (=1), or not suitable (=0).

Results: Of the 42 websites evaluated, we rated 62% (26/42) websites as "adequate" or above for trustworthiness. The information on 48% (20/42) websites had not been updated for more than a year, which indicates poor currency; 33% (14/42) websites failed to provide the publisher and contact information, which yields poor authority; 50% (21/42) websites did not cite sources of health information, which indicates lack of accuracy. Only 2 websites met the recommended readability (5th grade or lower reading level). The suitability was also suboptimal; only 1 website was rated as "superior"; 60% (25/42) websites were "adequate," and 38% (16/42) were "not suitable." Most websites were limited in graphical directions, interactive motivations for desired healthy behaviors, and multiple language translations.

Conclusions: The existing stroke-related websites in either English or Korean are trustworthy and suitable, yet precise citation of evidence-based information will improve trustworthiness. The contents requiring high reading level may set a barrier to the utilization of Web-based health information for Korean Americans with a lower level of education. In addition, supplementing graphical examples, interaction features, and culturally relevant information in multiple languages are the areas for improvement in suitability. The improved features can reduce the reading burden of stroke patients or caregivers and build their confidence when applying the information for stroke management in daily living. These strategies are especially crucial to Korean Americans, who inevitably seek Web-based information to fill the gap between their demand and access to health care for a long-term self-management after a stroke.

XSL•FO
RenderX

Introduction

With advancements in the internet services and communication technologies, both release and seeking of health information through the Web have been exponentially growing. In the United States, 89% of adults use the internet, and among them, 72% lookup online for health information [1,2]. Use of mobile phones has made the access to health information ubiquitous. Recent surveys have reported that 77% of US adults own a mobile phone, and among them, 62% have used their phone to get information about a health condition [3,4].

Use of Web-based health information is common among Korean Americans [5-8], who constitute about 1.8 million and are the fifth largest Asian American population in the United States [9]. One possible reason for this population using websites as an important resource for health information may be the limited access to health care professionals due to lack of health insurance [10]. Korean Americans record one of the lowest rates of health insurance coverage among all racial and ethnic groups living in the United States [11-13]. The limited English proficiency, which is predominant in first-generation (75.5%) and monolingual (37%) immigrants, has also contributed to limited health care access among this group [14,15]. Limited health care access increases Korean Americans' sense of self-responsibility and may lead individuals to be more active in seeking health information through available sources, such as Web-based information [10].

Universally, individuals with chronic or stigmatized diseases such as stroke are more likely to search for health information on the internet than those without the health conditions [16-19]. Stroke is the foremost cause of serious long-term disability, with high health care cost [20], and puts an increasing economic burden on health care resources [21]. The varying degrees of long-standing disability, as a result of stroke, lead to patients and their caregivers living with daunting long-term management. Stroke often results in emotional, cognitive, and physical impairments, which tend to be visible to others and cause stigmatizing social experiences after stroke [22,23]. With this burden, stroke survivors and caregivers often use websites as resources of information that they need [24].

Stroke is the third leading cause of death among Korean Americans [25]. The proportionate mortality ratios of stroke, especially hemorrhagic stroke, have been reported to be higher among Korean American women (2.07) and men (1.89) than among non-Hispanic white women (1.06) and men (0.94) [26]. Furthermore, the elderly have a higher prevalence of stroke risk factors, such as hypertension and dyslipidemia, than Caucasians, and the former lack knowledge about stroke [27]. Thus, prevention and appropriate long-term management of stroke and the sequelae with quality information is critical in this population.

Obtaining Web-based health information can be beneficial for or harmful to people. The explosion and proliferation of health information available online are on the promise that these online resources can confirm or broaden patients or families' understanding of diseases and treatment opinions that influence health care decisions and empower them to effectively self-manage health conditions [28,29]. However, the quality of the Web-based health information, generally in trustworthiness, readability, and suitability aspects, is often questioned. Online health information seekers could be at risk of finding unreliable or inappropriate information from the websites. Incorrect or inappropriate health information can be used in an improper way and can cause detrimental outcomes by negatively influencing health care-related decisions [30,31]. As patients have a wide range of health information literacy, some patients are unable to critically assess health information or might misinterpret it [31]. Online health information can be advantageous only if it is understandable to the consumers [29].

Considering the high demand of Web-based health information in native or English language among Korean Americans, who have limited English proficiency, it is critical for the websites to be *trustworthy* with up-to-date, reliable, and accurate information, which is *readable* and *suitable* for this specific population. Thus, it is important to examine the trustworthiness, readability, and suitability of health information available in native and English languages on the websites. The study findings will facilitate devising strategies to better design and implement stroke-related websites.

The purpose of this study was to address the following question: Is the Web-based stroke information adequate for Korean Americans to read, understand, and engage in stroke care? Specifically, we aimed to evaluate stroke-related websites in terms of their (1) trustworthiness, (2) readability, and (3) suitability for stroke prevention and self-management in this group.

Methods

Selection of the Websites for Stroke

Initially, we considered a question that how Korean Americans would seek stroke information on the internet. One of the common behaviors exhibited was bilingual searches in their health information seeking as a technique for coping with limited English proficiency [32]. In general, people selected health websites within only the first one or two pages of search results (10 results per page) when searching the internet; about half of the internet users entered a single query [32,33]. In van Deursen and van Dijk's [33] study, nobody used advanced search features (eg, Boolean operators like AND, OR, and NOT). In fact, Google and Yahoo have been the most popular search engines among Koreans and Americans [2,32].

Based on a review of these studies about online health information-seeking behaviors of Korean Americans, we

XSL•FO

RenderX

searched stroke-related websites in Google and Yahoo search engines using the following terms: "Stroke," "CVA," "중풍 (jungpung)," and "뇌졸증 (noejoljung)." The term, "Stroke" is health care consumers' preferred term listed in the Consumer Health Vocabulary Initiative [34]. The term, "CVA'" is another representative term for stroke. Furthermore, "중풍" and "뇌졸증" are the most commonly used Korean terms to indicate "stroke." We retrieved a total of 156 stroke-related websites that appeared on the 1st and 2nd pages only after searching. With the use of terms "Stroke" and "CVA," Google revealed 45 sites, while Yahoo led to 35 sites. Searching with the use of terms "중풍" and "뇌졸증" resulted in 36 sites from Google and 40 sites from Yahoo. Then, irrelevant websites, which were newsletters, commercial and noncommercial advertisements, and homonyms such as stroke motion in swimming or tennis, were excluded. After duplicates and the irrelevant sites (n=116) were eliminated, a total of 42 websites (15 in English and 27 in Korean), including those of the not-for-profit organizations, clinics, magazines, and blogs, were selected.

Evaluation Tools

Trustworthiness

We initially evaluated each website for its trustworthiness using the following three criteria for health websites' evaluation endorsed by the National Library of Medicine (NLM) [35]: currency or timeliness (when was the website last updated?); authority (who published the website?); and accuracy (are the sources cited reliable?). Each criterion was rated on a 3-point Likert scale, where 0=not suitable, 1=adequate, and 2=superior. The descriptions for each point are presented in Table 1. Thus, the possible total score per site ranged from 0 to 6, with higher scores indicating better trustworthiness.

Readability

We evaluated the readability of the information on selected websites using the Simple Measure of Gobbledygook (SMOG) Readability Test and the Reading Grade level of the Suitability Assessment of Materials (SAM) instrument. The SMOG has been validated as very easy to compute and provide a reasonably accurate measure of readability when evaluating consumer-oriented health care materials in many studies. We used the English version of the SMOG measure for evaluating the websites published in English [36] and the Korean version [37] for the websites in Korean. The formula requires counting 10 consecutive sentences at the beginning, middle, and end of website pages (30 total sentences). Second, it requires counting the number of words with ≥3 syllables in the 30-sentence sample. Then, with the total number of polysyllabic words counted, the grade level is determined using the SMOG conversion table. When using the SAM, the reading grade level can be measured by using different reading formulas, including the SMOG. Studies on the readability of patient education materials have often used both the SMOG and SAM tools together. For example, Rosenfeld et al [38] and Shieh and Hosei [39] computed the SMOG scores and integrated the scores to SAM 3-point categories. Similarly, in this study, the grade levels computed by the SMOG formula were converted to the SAM's 3-point scale, where 0=not suitable (9th-grade level and above),

1=adequate (6th-, 7th-, or 8th-grade level), and 2=superior (5th-grade level or lower).

Suitability

Each website was evaluated by the SAM instrument [40]. The SAM has been used in many studies to evaluate the suitability of online health information [41-46]. It consists of 22 items grouped under 6 factors, namely (1) Content, (2) Literacy Demands, (3) Graphics, (4) Layout and Typography, (5) Learning Stimulation and Motivation, and (6) Cultural Appropriateness. We added one item to the list, which was "multiple language translations," under the factor of Cultural Appropriateness. Each of the 23 items was rated on an ordinal scale, where 0=not suitable, 1=adequate, and 2=superior. Raw scores were summed to yield an overall score. This overall score was then converted to a percent of the possible total score for each website using the following formula: converted percent=total score/total possible score (46=23 items×2 maximum score per item)×100. A converted percent of 70%-100% indicates a superior website, 40%-69% indicates an adequate website, and 0%-39% indicates a not suitable website [40].

When previous studies assessed the suitability of health information across multiple sources, the converted percent was also applied to the present the quality of the health information per SAM criterion [47,48]. In this study examining which criterion was suitably met across the websites, the mean score per criterion was computed and then converted to a percent of the possible best score using the following formula: the converted percent=([n of websites×0 point]+[n of websites×1 point]+[n of websites×2 points])/Total possible best score (84=42 websites×2 points)×100.

Then, if the converted percent for a criterion was in the range of 70%-100%, the quality of the information regarding the criterion was considered as "superior" across websites. If the percentage was in the range of 40%-69%, the quality of the information per criterion was considered at an "adequate" level. If the percentage was in the range of 0%-39%, the information regarding a criterion was considered as "not suitable" across the websites.

Evaluation and Analyses

Three content experts evaluated each website independently in the first round and collected the ratings. In the second round, the three raters reviewed the websites synchronously to ensure the accuracy of understanding of each criterion and, then, shared their rationales for the ratings. The frequencies and intraclass correlation coefficients (ICCs) per the evaluation criteria were obtained using SPSS 24. The final interrater agreement levels for the three evaluation criteria were as follows: NLM, ICC=.969-.987; SMOG, ICC=.810; and SAM, ICC=.626-.994.

Results

General Characteristics

Out of the 42 appraised websites, 15 stroke websites were in English and 27 were in Korean. Of all, 12 websites were those of the not-for-profit organizations, 7 were blogs on stroke, 5

were published by clinics or hospitals, 5 were Wikipedia pages, 4 were medical magazine sites, 3 were general magazine sites, 2 were postings on broadcasting websites, 1 was a pharmacological company website, 1 was an insurance company website; 1 was a general online forum site, and 1 was a medical information website.

Trustworthiness

Overall, 62% (26/42) websites received a rating of >3.0, which presents "adequate" level. Regarding the currency or timeline, 79% (33/42) of the selected websites indicated published date of the information on their websites, whereas 21% (9/42) did not indicate a published date of the information. The updated dates of the information on 31% (13/42) websites were <1 year old, while on 48% (20/42) websites these were older than 1 year. In terms of authority, 21% (9/42) websites did not present the publisher or contact information. Regarding accuracy, 33% (14/42) websites posted health information based on medical research evidence along with citations, whereas 50% (21/42) of the websites did not cite any source of information (Table 1).

Readability

A total of 19 of 42 (45%) websites were presented at the 9th-grade level or above (0=not suitable). Twenty-one (50%) websites were presented at reading levels between 6th and 8th grade (1=adequate). Only 2 websites (5%) were presented at the 5th grade- or lower level (2=superior).

Suitability

Table 2 presents the frequencies per rating score, ICC of the three raters per the SAM criterion, and the converted suitability percent score. The overall converted suitability percent score of all 42 websites was 55.7%, which represents "adequate" suitability. Out of 42 websites, only 1 website of the American Stroke Association was rated as "superior"; 60% (25/42) websites were "adequate," and 38% (16/42) websites were "not suitable." The websites were superior in the quality of layout and typography (85.3%). Furthermore, the websites were at an adequate level regarding the Content (65.5%); Literacy Demands (68.6%); Graphic Illustration, Lists, Tables, and Charts (42.4%); and Learning Stimulation and Motivation (47.2%), even though there were some individual items with inadequate suitability. However, Cultural Appropriateness was not suitable (22.2%).

Under the Content category, the quality of purpose and scope was superior, with scores of 90.5% and 79.8%, respectively. Regarding the extent of the content topics, 26% (11/42) websites aimed at desirable behavior rather than at nonbehavioral facts. In addition, 60% (25/42) websites showed <40% of the content topics focusing on desirable behaviors or actions. However, 6 websites did not present such contents. Many of the appraised websites (23/42, 55%) did not present summaries or reviews well to convey key messages.

Regarding the Literacy Demands, the selected websites were suitable in writing style (81.0%), sentence construction (77.4%), vocabulary use (72.6%), and organization using road signs (82.1%). However, the reading grade level was not suitable across websites (29.8%).

Table 1. Trustworthiness of stroke websites by the National Library of Medicine criteria.

Evaluation criteria[a]		Frequency, n (%)	ICC[b]
Currency or timeliness: when was the website last updated?			.969
Superior (2)	The published date of the information is indicated and is less than 1 year old (last year)	13 (31)	
Adequate (1)	The published date of information is indicated, but is older than 1 year	20 (48)	
Not suitable (0)	No indication	9 (21)	
Authority: who published the website?			.980
Superior (2)	The publisher's information (individuals or organizations) and contact information can be easily found	28 (67)	
Adequate (1)	The publisher's information (individuals or organizations) can be found. But there is no contact information	5 (12)	
Not suitable (0)	No indication	9 (21)	
Accuracy: are the sources cited reliable?			.987
Superior (2)	The information is drawn based on sound medical research, and the information sources are cited	14 (33)	
Adequate (1)	The information sources are cited but not based on medical research	7 (17)	
Not suitable (0)	No indication	21 (50)	

[a]Score shown in parentheses.

[b]ICC: intraclass correlation coefficient.

Table 2. The evaluation of the stroke websites based on Suitability Assessment of Materials.

Category and criteria	Frequency, n (%)			ICC[a]	Converted %[b]
	Not suitable (0)	Adequate (1)	Superior (2)		
Content					
Purpose	0 (0)	8 (19)	34 (81)	.626	90.5
Scope	2 (5)	13 (31)	27 (64)	.896	79.8
Content topics	6 (14)	25 (60)	11 (26)	.875	56.0
Summary and review	23 (55)	8 (19)	11 (26)	.993	35.7[c]
Average					65.5
Literacy demands					
Reading grade level	19 (45)	21 (50)	2 (5)	.810	29.8[c]
Writing Style	1 (2)	14 (33)	27 (64)	.859	81.0
Sentence construction	1 (2)	17 (41)	24 (57)	.727	77.4
Vocabulary	2 (5)	19 (45)	21 (50)	.746	72.6
Learning enhancement by advance organizers	6 (14)	3 (7)	33 (79)	.962	82.1
Average					68.6
Graphic illustration, lists, tables, and charts					
Cover graphic	20 (48)	13 (31)	9 (21)	.950	36.9[c]
Type of illustrations	14 (33)	15 (36)	13 (31)	.950	48.8
Relevance of illustrations	15 (36)	11 (26)	16 (38)	.975	51.2
Graphical direction: lists, tables, charts, and forms	25 (60)	10 (24)	7 (17)	.956	28.6[c]
Captions are used to announce or explain graphics	14 (33)	17 (41)	11 (26)	.976	46.4
Average					42.4
Layout and typography					
Typography	1 (2)	9 (21)	32 (76)	.863	86.9
Layout	1 (2)	12 (29)	29 (69)	.813	83.3
Subheadings and chunking	4 (10)	4 (10)	34 (81)	.965	85.7
Average					85.3
Learning, stimulation, and motivation					
Interaction included in the text or graphics	16 (38)	15 (36)	11 (26)	.967	44.0
Desired behavior patterns are modeled	14 (33)	12 (29)	16 (38)	.975	52.4
Motivation	16 (38)	14 (33)	12 (29)	.965	45.2
Average					47.2
Cultural appropriateness					
Cultural match: logic, language, experience	24 (57)	8 (19)	10 (24)	.992	33.3[c]
Cultural image and examples	29 (69)	10 (24)	3 (7)	.987	19.0[c]
Multiple languages translation	36 (86)	0 (0)	6 (14)	.994	14.3[c]
Average					22.2[c]
Overall average					55.7

[a]ICC: intraclass correlation coefficient.

[b]Converted Percent=([n of websites×0 point]+[n of websites×1 point]+[n of websites×2 points])/total possible best score (84=42 websites×2 points)×100. Interpretation: 70%-100%=superior; 40%-69%=adequate; 0%-39% =not suitable.

[c]Information rated as "not suitable" per criterion across the evaluated websites.

The overall suitability of the Graphic Illustration, Lists, Tables, and Charts category was rated as "adequate" (42.4%). In this category, the type of illustrations (48.8%), relevance of illustrations (51.2%), and captions (46.4%) were "adequate" at the lower levels, but cover graphic (36.9%) and graphical direction (28.6%) were "not suitable." In more detail, 48% (20/42) of the websites did not present cover graphic to attract attention and to clearly portray the purpose of the website.

In addition, 60% (25/42) of the websites presented graphics without explanation and 24% (10/42) showed too brief "how to" directions with graphics for readers; only 17% (7/42) websites provided step-by-step directions with an example that will build self-efficacy (confidence). The layouts and typography were superiorly presented in most appraised websites, with the converted suitability scores of 83.3%-86.9%.

In the Learning Stimulation and Motivation category, the suitability of each criterion was rated as "adequate," even though the suitability percent scores were not high, ranging 44.0%-52.4%. In more details, 38% (16/42) websites did not provide interaction learning or stimulation and 36% (15/42) used passive interactions with the Questions & Answer format. Regarding the desired behavior pattern presentation, 38% (16/42) websites demonstrated instructions for specific behavior and skills by using specific, familiar instances with the rating of "superior." In addition, 29% (12/42) websites were rated as "adequate" as they provided the information in a mix of technical and common language that the reader may not easily link with daily living activities. In terms of Motivation, only 29% (12/42) websites reached the "superior" level, which means that complex topics are subdivided so that readers may experience small successes in understanding or problem solving, leading to self-efficacy (confidence). Furthermore, 38% (16/42) websites did not have features of motivation.

The Cultural Appropriateness across websites was "not suitable"; the converted suitability percent scores ranged from 14.3% to 33.3%. Regarding cultural match, 57% (24/42) websites did not present the information in a culturally similar logic, language, and experience of the target population culture. The cultural image and examples were rarely shown in 69% (29/42) websites. Moreover, 86% (36/42) of the websites were presented in only one language and could not be translated into other languages.

Discussion

Principal Findings

This study is the first to provide a systematic evaluation of the stroke-related websites considering Korean Americans. The findings of this study present valuable information about the trustworthiness, readability, and suitability of stroke-related English and Korean websites, which were trustworthy and marginally suitable, but not easily readable. In addition, areas that need improvement were identified in each criterion. This information can be utilized by clinicians and researchers in improving or designing stroke websites as well as by people who seek stroke information in their self-assessment of the quality of the information on websites.

Trustworthiness

Patients and caregivers look up the websites as they are easy to access, not expensive, and, in general, believed to be timely [32]. Patients are often unsure about which websites to trust and are concerned about accessing potentially misleading or illegitimate health information [49,50]. The outdated, inaccurate, and unreliable Web-based information can mislead the public in their understanding of stroke and stroke management. In the limited chance of education on which website and its information are trustworthy, Korean Americans consider merely the repetition of specific information as a criterion for evaluating the trustworthiness of websites; if the same information appears in several locations on the internet, Korean Americans are likely to simply believe that the information or the website is reliable [32]. The trustworthiness of the website information could be easily checked by the NLM's [35] three criteria of health website evaluations. These evaluation criteria need to be educated to online health information seekers, including the population in question.

In this study, the information on 20 websites was published more than 1 year ago and 9 websites did not indicate a published date of the information. Notably, half of the websites did not cite any source of information on their websites. Agarwal et al [30] also found in their evaluation of educational resources in 3 stroke-related websites that users criticized the lack of citations and references for the material and facts provided on all 3 websites. It is important to present the valid and reliable primary information with the published date and the information sources for health consumers to help them with right decision making regarding their health issues [16,30]. It is recommended to scrutinize or peer review before presenting the information on the website. Health care providers should be aware of, at least, these three criteria when providing health care information to patients and caregivers via papers or websites. Furthermore, health care providers should provide opportunities to educate patients and caregivers on how to check and whether to rely on health care information in any type of Web information platform.

Readability

Reading level is an essential component of health literacy and is included as one of the Literacy Demand criteria in the SAM tool. Concerned about health literacy, the Joint Commission [51] has recommended patient education materials to be written at or below the 5th-grade level and has developed an action plan to promote patient literacy. In this study, 19 out of 42 websites were written at the 9th-grade level or above, which is rated as "not suitable" and categorized as "difficult" in the US readability categories [52]. Only 2 websites were written at the 5th grade- or lower level. This result is consistent with those of previous studies, which evaluated some stroke-related websites: the mean readability level of stroke information was found to be 10th grade by Griffin et al [53] in their evaluation of 30 stroke educational websites; these websites were different from the set of websites appraised in our study, except 3 overlapping websites. Furthermore, Sharma et al [24] found that over half of 100 stroke websites were produced at the 12th-grade level or above when the readability was measured with the SMOG.

XSL·FO

RenderX

The high reading levels may be due to complex medical terminologies without explanations or translation into layperson's terms.

The health information is only useful if the consumer can comprehend the presented information. The consequences of not understanding health information can negatively affect both a person's health and his or her utilization of health care services. The use of consumer health vocabulary may help lower the reading grade level on both English and Korean stroke websites. In the United States, the Open-Access, Collaborative Consumer Health Vocabulary provides 156,826 consumer-friendly health phrases and synonyms for professional clinical terms [34,54]. In South Korea, a consumer vocabulary system for health information was developed in a study [55]. Also, the complex medical terms or sentences could be provided with additional understandable explanations so that readers become familiar with them and can use them appropriately [38]. Korean Americans' limited English proficiency should also be taken into consideration. Limited English proficiency was identified as a critical source of health vulnerability; it impacted health perceptions, led to higher health risk, and negatively influenced health care utilization [56-58]. Those with limited English proficiency scored lower on all measures of health conditions than their English-proficient counterparts [57]. Considering the limited English proficiency among Korean Americans, terms and sentence structures of stroke care information on English websites should be used at an appropriate reading level.

Suitability

The overall suitability of all 42 websites was rated as "adequate" (55.7%); however, there are needs for improvement in many areas. Regarding the Content, the majority of the websites explicitly showed the purpose and the scope of the content. However, they did not present summaries or reviews well to convey key messages. In order to ensure the delivery of important messages and to assist the readers' comprehension, it is necessary to retell key messages in different words or examples and provide their summaries.

The overall suitability of Graphic Illustration was marginally adequate (42.4%). Half of the websites did not present cover graphic to attract attention and clearly portray the purpose of the websites. In addition, 60% (25/42) of the websites presented graphics without explanation and 24% (10/42) showed too brief "how to" directions for readers. This finding is similar to the lack of graphics and other nontextual media covering patient education materials in Agarwal et al's [30] evaluation of educational resources in 3 stroke-related websites. Images and videos combined with the text can act as supplements to difficult topics covered on the site [30] and can increase the ability of the user to understand and retain the material [59]. Since the medical information is relatively unfamiliar to laypersons, and the reading level is high, step-by-step directions with graphical examples can reduce the reading burden of stroke patients and caregivers. Furthermore, use of strategies to enhance readers' understanding of the information on websites is especially important to Korean Americans, who inevitably seek online

information due to lack of insurance and limited English proficiency.

The Learning Stimulation and Motivation category also received the rating of "adequate," but it was not high (47.2%). No interaction learning or stimulation was provided on 38% (16/42) websites. Regarding the desired behavior patterns presentation, the criterion of the "superior" rating probed if websites demonstrated instructions for specific behavior and skills by using specific, familiar instances; more than half of the websites failed to reach this level. Only 29% (12/42) websites ranked "superior" in the Motivation criterion. Mere access to Web-based health information does not necessarily empower consumers and patients. Providing actionable health information with accuracy is the top vision in the National Action Plan to Improve Health Literacy published by the US Department of Health and Human Services [60]. In addition, practical guidelines demonstrating specific behavior and skills applicable to daily living will motivate the readers to experience small success in understanding or problem solving, leading to self-efficacy. Multimedia and interactivity features can assist information users to build more confidence by tailoring to individual needs and progress of condition management and rehabilitation. The incorporation of self-assessments of risks or symptom changes and modeling of desirable health behaviors into the website will promote self-efficacy in learning. Miller and Leroy [54] suggested that it would be ideal if a consumer would go to a website, answer a few questions, and in a few seconds, a document written specifically for their needs and appropriate to their reading skill would appear in their Web browser. The features of dynamic discussions and communication will lead to co-learning with each other [28,61]. Chat rooms and email services with health care providers on websites are other recommendable features to satisfy patients' information needs.

Of all, 86% (36/42) websites were presented in only one language and could not be translated into other languages. More than half of the websites in English or Korean did not present information matched to the Korean culture in texts, images, or examples. The cultural factors influence health information acquisition and access to social support for ethnic minorities [62]. Korean consumers rely heavily on Korean-specific health information, regardless of education levels [32]. Korean Americans have a lower level of knowledge on stroke risk factors and symptoms than the general US population [63]. Websites are important sources and channels of health information for them. If websites provide more diverse and culturally appropriate information in both Korean and English, the health information can be delivered to a broader population of Korean Americans. Then, it may contribute to increase in knowledge about stroke prevention and self-management among this group.

Limitations

We evaluated the websites that appeared on the first and second pages only on Google and Yahoo browsers. Thus, the results cannot be generalized to reflect the features of all stroke-related websites. However, these websites can be the most popular or important ones because studies have found that 73%-95% of

Web users never view past the first search engine results page and have emphasized the value of the first page [64-67]. In addition, it may indicate the increasing ability of Web search engines to retrieve the relevant Web-based information more effectively [64]. The websites evaluated in this study were stroke specific, which may limit the application of this study results to websites of other specific health conditions.

Conclusions

The quality of health information for patients with stroke is vital to ensure the good recovery and to improve the quality of life. Easily obtainable, user-friendly, accurate, and reliable online resources could help people make appropriate decisions about how best to maintain or improve their health condition.

We found that the existing stroke-related websites in either English or Korean are trustworthy and suitable. However, it was notable that the accuracy of information on these websites was low due to lack of citations and references. Providing accurate, reliable, high-quality, and evidence-based information is a critical responsibility of health professionals. In addition, the study findings signified the need for diverse features to reduce the high reading level of the information on the websites. Graphical examples, multimedia, and interactive features can reduce the reading burden of stroke patients and caregivers, as

well as build more confidence or self-efficacy when applying the information for condition management and rehabilitation in daily living. Rather than posting the information in a nonspecific way, practical guidelines demonstrating specific behavior and skills applicable to daily living will motivate the reader to experience small success in stroke self-management, leading to self-efficacy.

These strategies to enhance readers' understanding of Web-based information are especially important to Korean Americans, who inevitably seek online information due to lack of familiarity with the US health care system, inadequate health insurance coverage, and language barriers. Culturally sensitive, high-quality health information relevant to this group through websites may have great potential to increase the knowledge of stroke self-management and make significant contributions to promote the health of Korean American stroke survivors. In the future, the evaluation of stroke websites by Korean American laypersons using the evaluation tools used in this study might be an educational health literacy intervention to learn how to evaluate and consider the design and content aspects in the websites. The examination of how this population navigates stroke websites and apply the health information obtained from these websites for their daily health behaviors should be considered for future studies.

Conflicts of Interest

None declared.

References

1. Pew Research Center. 2018 Feb 05. Internet/Broadband Fact Sheet URL: http://www.pewinternet.org/fact-sheet/internet-broadband/ [accessed 2018-03-15] [WebCite Cache ID 6xwL7IKQ3]
2. Fox S, Duggan M. Pew Research Center. 2013 Jan 15. Health Online 2013 URL: http://www.pewinternet.org/files/old-media/Files/Reports/PIP_HealthOnline.pdf [accessed 2018-03-15] [WebCite Cache ID 6xwM33Q2e]
3. Pew Research Center. 2018 Feb 05. Mobile Fact Sheet URL: http://www.pewinternet.org/fact-sheet/mobile/ [accessed 2018-03-15] [WebCite Cache ID 6xwM8JpT0]
4. Smith A. Pew Research Center. 2015 Apr 01. US Smartphone Use in 2015 URL: http://www.pewinternet.org/2015/04/01/us-smartphone-use-in-2015/ [accessed 2018-03-15] [WebCite Cache ID 6xwMJKZfs]
5. Kim S, Yoon J. The use of an online forum for health information by married Korean women in the United States. Information Research 2012 Jun;17(2):514 [FREE Full text]
6. Islam NS, Patel S, Wyatt LC, Sim SC, Mukherjee-Ratnam R, Chun K, et al. Sources of health information among select Asian American immigrant groups in New York City. Health Commun 2016;31(2):207-216 [FREE Full text] [doi: 10.1080/10410236.2014.944332] [Medline: 26266574]
7. Oh KM, Kreps GL, Jun J, Ramsey L. Cancer information seeking and awareness of cancer information sources among Korean Americans. J Cancer Educ 2011 Jun;26(2):355-364. [doi: 10.1007/s13187-010-0191-x] [Medline: 21210270]
8. Oh KM, Kreps GL, Jun J, Chong E, Ramsey L. Examining the health information-seeking behaviors of Korean Americans. J Health Commun 2012 Aug;17(7):779-801. [doi: 10.1080/10810730.2011.650830] [Medline: 22642692]
9. United States Census Bureau. 2017 Mar 14. FFF: Asian-American and Pacific Islander Heritage Month: May 2017 URL: https://www.census.gov/newsroom/facts-for-features/2017/cb17-ff07.html [accessed 2018-03-15] [WebCite Cache ID 6xwMi8BR9]
10. Oh KM, Zhou QP, Kreps G, Kim W. The influences of immigration on health information seeking behaviors among Korean Americans and Native Koreans. Health Educ Behav 2014 Apr;41(2):173-185. [doi: 10.1177/1090198113496789] [Medline: 23943681]

11.   Kao D. Factors associated with ethnic differences in health insurance coverage and type among Asian Americans. J Community Health 2010 Apr;35(2):142-155. [doi: 10.1007/s10900-009-9209-x] [Medline: 20013061]

12.   Yoo G, Kim B. Korean immigrants and health care access: Implications for the uninsured and underinsured. In: Kronenfeld JJ, editor. Inequalities and Disparities in Health Care and Health: Concerns of Patients, Providers and Insurers (Research in the Sociology of Health Care Vol 25). Bingley, United Kingdom: Emerald Group Publishing Limited; 2007:77-94.

13.   American Association of Retired Persons. 2015 Jan. The health and healthcare of Asian Americans and Pacific Islanders Age 50+: An AARP Report URL: https://www.aarp.org/content/dam/aarp/home-and-family/asian-community/2015-01/AARP-Report-Health-and-Healthcare-AAPIs-50-Jan2015.pdf [accessed 2018-03-15] [WebCite Cache ID 6xxZ8cvf2]

14.   Hoeffel E, Rastogi S, Kim MO, Shahid H. United States Census Bureau. 2012 Mar. The Asian Population: 2010 URL: https://www.census.gov/prod/cen2010/briefs/c2010br-11.pdf [accessed 2018-03-15] [WebCite Cache ID 6xxZNTocd]

15.   Pew Research Center. 2013 Apr 04. The rise of Asian Americans URL: http://www.pewsocialtrends.org/files/2013/04/Asian-Americans-new-full-report-04-2013.pdf [accessed 2018-03-15] [WebCite Cache ID 6xxZTXrnt]

16.   Amante DJ, Hogan TP, Pagoto SL, English TM, Lapane KL. Access to care and use of the Internet to search for health information: Results from the US National Health Interview Survey. J Med Internet Res 2015;17(4):e106 [FREE Full text] [doi: 10.2196/jmir.4126] [Medline: 25925943]

17.   Bansil P, Keenan NL, Zlot AI, Gilliland JC. Health-related information on the Web: Results from the HealthStyles Survey, 2002-2003. Prev Chronic Dis 2006 Apr;3(2):A36 [FREE Full text] [Medline: 16539777]

18.   Berger M, Wagner TH, Baker LC. Internet use and stigmatized illness. Soc Sci Med 2005 Oct;61(8):1821-1827. [doi: 10.1016/j.socscimed.2005.03.025] [Medline: 16029778]

19.   Bundorf MK, Wagner TH, Singer SJ, Baker LC. Who searches the internet for health information? Health Serv Res 2006 Jun;41(3 Pt 1):819-836 [FREE Full text] [doi: 10.1111/j.1475-6773.2006.00510.x] [Medline: 16704514]

20.   Benjamin EJ, Blaha MJ, Chiuve SE, Cushman M, Das SR, Deo R, American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Heart Disease and Stroke Statistics-2017 Update: A report from the American Heart Association. Circulation 2017 Mar 07;135(10):e146-e603. [doi: 10.1161/CIR.0000000000000485] [Medline: 28122885]

21.   Burridge JH, Lee ACW, Turk R, Stokes M, Whitall J, Vaidyanathan R, et al. Telehealth, wearable sensors, and the Internet: Will they improve stroke outcomes through increased intensity of therapy, motivation, and adherence to rehabilitation programs? J Neurol Phys Ther 2017 Jul;41 Suppl 3 Supplement, IV STEP Special Issue:S32-S38. [doi: 10.1097/NPT.0000000000000183] [Medline: 28628594]

22.   Anderson S, Whitfield K. Social identity and stroke: 'they don't make me feel like, there's something wrong with me'. Scand J Caring Sci 2013 Dec;27(4):820-830. [doi: 10.1111/j.1471-6712.2012.01086.x] [Medline: 23121474]

23.   Feigin VL, Forouzanfar MH, Krishnamurthi R, Mensah GA, Connor M, Bennett DA, Global Burden of Diseases, Injuries, and Risk Factors Study 2010 (GBD 2010) and the GBD Stroke Experts Group. Global and regional burden of stroke during 1990-2010: Findings from the Global Burden of Disease Study 2010. Lancet 2014 Jan 18;383(9913):245-254 [FREE Full text] [Medline: 24449944]

24.   Sharma N, Tridimas A, Fitzsimmons PR. A readability assessment of online stroke information. J Stroke Cerebrovasc Dis 2014 Jul;23(6):1362-1367. [doi: 10.1016/j.jstrokecerebrovasdis.2013.11.017] [Medline: 24389374]

25.   The Korean American Health Coalition. 2017. Korean American Health Coalition Strategic Plan 2016-2018 URL: http://kahealthcoalition.org/strategic-plan/ [accessed 2018-03-16] [WebCite Cache ID 6xxZeR7Cf]

26.   Jose PO, Frank AT, Kapphahn KI, Goldstein BA, Eggleston K, Hastings KG, et al. Cardiovascular disease mortality in Asian Americans. J Am Coll Cardiol 2014 Dec 16;64(23):2486-2494 [FREE Full text] [doi: 10.1016/j.jacc.2014.08.048] [Medline: 25500233]

27.   Song S, Song J, Park H, Chung J, The STROKE-COPE Community Advisory Panel, Ryan G, et al. Exploring stroke knowledge in Korean-American seniors: The strengthening our Korean elders through community-partnered education (STROKE-COPE) Project. Neurology 2013;80(7 Supplement) [FREE Full text]

28.   Du H, Ma J, Li M. High-quality health information provision for stroke patients. Chin Med J (Engl) 2016 Dec 05;129(17):2115-2122 [FREE Full text] [doi: 10.4103/0366-6999.189065] [Medline: 27569241]

29.   Walsh TM, Volsko TA. Readability assessment of internet-based consumer health information. Respir Care 2008 Oct;53(10):1310-1315 [FREE Full text] [Medline: 18811992]

30.   Agarwal N, Kommana SS, Hansberry DR, Kashkoush AI, Friedlander RM, Lunsford LD. Accessibility, reliability, and usability of neurosurgical resources. J Neurosurg 2017 Apr;126(4):1263-1268. [doi: 10.3171/2015.12.JNS151102] [Medline: 27257833]

31.   Tonsaker T, Bartlett G, Trpkov C. Health information on the Internet: Gold mine or minefield? Can Fam Physician 2014 May;60(5):407-408 [FREE Full text] [Medline: 24828994]

32.   Yi YJ, Stvilia B, Mon L. Cultural influences on seeking quality health information: An exploratory study of the Korean community. Libr Inf Sci Res 2012 Jan;34(1):45-51. [doi: 10.1016/j.lisr.2011.06.001]

33.   van Deursen AJAM, van Dijk JAGM. Using the Internet: Skill related problems in users' online behavior. Interact Comput 2009 Dec;21(5-6):393-402. [doi: 10.1016/j.intcom.2009.06.005]

34. The University of Utah, Department of Biomedical Informatics. Open Access, Collaborative Consumer Health Vocabulary Initiative. Consumer Health Vocabulary Wiki URL: http://consumerhealthvocab.chpc.utah.edu/CHVwiki/ [accessed 2018-03-16] [WebCite Cache ID 6xxlMNI7g]

35. National Network of Libraries of Medicine. Evaluating Health Websites URL: https://nnlm.gov/initiatives/topics/health-websites [accessed 2018-03-16] [WebCite Cache ID 6xxlTTALm]

36. McLaughlin G. SMOG grading: A new readability formula. J Reading 1969;12(8):639-646 [FREE Full text]

37. Sung N, Lee D, Park K. Suitability assessment of patients' education materials made by Korean Academy of Family Medicine. Korean J Fam Med 2004;25(9):669-677 [FREE Full text]

38. Rosenfeld L, Shepherd A, Agunwamba AA, McCray AT. Iterative evaluation of a web-based health information resource. J Health Commun 2013 Aug;18(8):974-990 [FREE Full text] [doi: 10.1080/10810730.2012.757395] [Medline: 23577665]

39. Shieh C, Hosei B. Printed health information materials: evaluation of readability and suitability. J Community Health Nurs 2008;25(2):73-90. [doi: 10.1080/07370010802017083] [Medline: 18449833]

40. Doak CC, Doak LG, Root JH. Assessing suitability of materials. In: Doak CC, Doak LG, Root JH, editors. Teaching patients with low literacy skills. 2nd ed. Philadelphia, PA: Lippincott Company; 1996:41-60.

41. Guan Y, Maloney KA, Roter DL, Pollin TI. Evaluation of the informational content, readability and comprehensibility of online health information on monogenic diabetes. J Genet Couns 2018 Jun;27(3):608-615. [doi: 10.1007/s10897-017-0155-y] [Medline: 28951986]

42. Tran BNN, Singh M, Lee BT, Rudd R, Singhal D. Readability, complexity, and suitability analysis of online lymphedema resources. J Surg Res 2017 Jun 01;213:251-260. [doi: 10.1016/j.jss.2017.02.056] [Medline: 28601323]

43. Robins S, Barr HJ, Idelson R, Lambert S, Zelkowitz P. Online health information regarding male infertility: An evaluation of readability, suitability, and quality. Interact J Med Res 2016 Oct 21;5(4):e25 [FREE Full text] [doi: 10.2196/ijmr.6440] [Medline: 27769954]

44. Tian C, Champlin S, Mackert M, Lazard A, Agrawal D. Readability, suitability, and health content assessment of web-based patient education materials on colorectal cancer screening. Gastrointest Endosc 2014 Aug;80(2):284-290. [doi: 10.1016/j.gie.2014.01.034] [Medline: 24674352]

45. Nasser S, Mullan J, Bajorek B. Assessing the quality, suitability and readability of internet-based health information about warfarin for patients. Australas Med J 2012;5(3):194-203 [FREE Full text] [doi: 10.4066/AMJ.2012862] [Medline: 22952566]

46. Finnie RK, Felder TM, Linder SK, Mullen PD. Beyond reading level: A systematic review of the suitability of cancer education print and Web-based materials. J Cancer Educ 2010 Dec;25(4):497-505 [FREE Full text] [doi: 10.1007/s13187-010-0075-0] [Medline: 20237884]

47. Garnweidner-Holme LM, Dolvik S, Frisvold C, Mosdøl A. Suitability assessment of printed dietary guidelines for pregnant women and parents of infants and toddlers from 7 European countries. J Nutr Educ Behav 2016 Feb;48(2):146-151.e1. [doi: 10.1016/j.jneb.2015.10.004] [Medline: 26603301]

48. Okuhara T, Ishikawa H, Okada H, Kiuchi T. Readability, suitability and health content assessment of cancer screening announcements in municipal newspapers in Japan. Asian Pac J Cancer Prev 2015;16(15):6719-6727 [FREE Full text] [Medline: 26434901]

49. Sbaffi L, Rowley J. Trust and credibility in Web-based health information: A review and agenda for future research. J Med Internet Res 2017 Jun 19;19(6):e218 [FREE Full text] [doi: 10.2196/jmir.7579] [Medline: 28630033]

50. Vega LC, Montague E, Dehart T. Trust between patients and health websites: A review of the literature and derived outcomes from empirical studies. Health Technol (Berl) 2011 Nov 18;1(2-4):71-80 [FREE Full text] [doi: 10.1007/s12553-011-0010-3] [Medline: 22288026]

51. The Joint Commission. Oakbrook Terrace, IL; 2010. Advancing Effective Communication, Cultural Competence, and Patient- and Family-Centered Care: A Roadmap for Hospitals URL: https://www.jointcommission.org/assets/1/6/ARoadmapforHospitalsfinalversion727.pdf [accessed 2018-03-16] [WebCite Cache ID 6xxu3aFiw]

52. The UCFM&MS. Toolkit for making written material clear and effective. 2012 URL: https://www.cms.gov/Outreach-and-Education/Outreach/WrittenMaterialsToolkit/index.html [accessed 2018-03-16] [WebCite Cache ID 6xxtznllU]

53. Griffin E, McKenna K, Worrall L. Stroke education materials on the World Wide Web: An evaluation of their quality and suitability. Top Stroke Rehabil 2004;11(3):29-40. [doi: 10.1310/JKJK-N2MP-9LQ1-FF2N] [Medline: 15480952]

54. Miller T, Leroy G. Dynamic generation of a Health Topics Overview from consumer health information documents. Int J Biomed Eng Technol 2008;1(4):395-414. [doi: 10.1504/IJBET.2008.020069]

55. Shin Y, Kim J. A study for building a system of consumer vocabulary for health information. J Korean Soc Med Inform 2009 Mar;15(1):31-40. [doi: 10.4258/jksmi.2009.15.1.31]

56. Jang Y, Kim G, Chiriboga DA. Health, healthcare utilization, and satisfaction with service: Barriers and facilitators for older Korean Americans. J Am Geriatr Soc 2005 Sep;53(9):1613-1617. [doi: 10.1111/j.1532-5415.2005.53518.x] [Medline: 16137296]

57. Jang Y, Yoon H, Park NS, Chiriboga DA. Health vulnerability of immigrants with limited English proficiency: A study of older Korean Americans. J Am Geriatr Soc 2016 Jul;64(7):1498-1502 [FREE Full text] [doi: 10.1111/jgs.14199] [Medline: 27305524]

58.   Sentell T, Braun KL. Low health literacy, limited English proficiency, and health status in Asians, Latinos, and other racial/ethnic groups in California. J Health Commun 2012;17 Suppl 3:82-99 [FREE Full text] [doi: 10.1080/10810730.2012.712621] [Medline: 23030563]

59.   Stegeman CA, Zydney J. Effectiveness of multimedia instruction in health professions education compared to traditional instruction. J Dent Hyg 2010;84(3):130-136. [Medline: 20579425]

60.   US Department of Health and Human Services, Office of Disease Prevention and Health Promotion. National Action Plan to Improve Health Literacy. 2010 URL: https://health.gov/communication/hlactionplan/pdf/Health_Literacy_Action_Plan.pdf [accessed 2018-03-16] [WebCite Cache ID 6xxuoq1qU]

61.   Smith J, Forster A, House A, Knapp P, Wright J, Young J. Information provision for stroke patients and their caregivers. Cochrane Database Syst Rev 2008 Apr;16(2):CD001919. [doi: 10.1002/14651858.CD001919.pub2] [Medline: 18425877]

62.   Kim W, Kreps GL, Shin C. The role of social support and social networks in health information-seeking behavior among Korean Americans: A qualitative study. Int J Equity Health 2015 Apr 28;14:40 [FREE Full text] [doi: 10.1186/s12939-015-0169-8] [Medline: 25927546]

63.   Hwang SY, Zerwic JJ. Knowledge of stroke symptoms and risk factors among Korean immigrants in the United States. Res Nurs Health 2006 Aug;29(4):337-344. [doi: 10.1002/nur.20144] [Medline: 16847912]

64.   Jansen BJ, Spink A. How are we searching the World Wide Web? A comparison of nine search engine transaction logs. Inform Process Manag 2006 Jan;42(1):248-263. [doi: 10.1016/j.ipm.2004.10.007]

65.   Jacobson M. Leverage Marketing. How far down the search engine results page will most people go? URL: https://www.theleverageway.com/blog/how-far-down-the-search-engine-results-page-will-most-people-go/ [accessed 2018-03-18] [WebCite Cache ID 6xzx2vnWm]

66.   Karaivanov D. Inet Solutions. 2016 Sep 08. 25 Informative SEO Facts & Statistics You Probably Don't Know About URL: https://www.inetsolutions.org/25-informative-seo-facts-statistics-probably-dont-know/ [accessed 2018-03-18] [WebCite Cache ID 6xzxCwljv]

67.   Hodgdon M. Infront Webworks. 2015 Jul 17. Value of Organic First-Page Results URL: https://www.infront.com/blog/the-infront-blog/2015/06/17/value-of-first-page-google-results [accessed 2018-03-16] [WebCite Cache ID 6xzxPA7WH]

Abbreviations

CVA: cardiovascular accident
ICC: intraclass correlation coefficient
NLM: National Library of Medicine
SAM: Suitability Assessment of Materials
SMOG: Simple Measure of Gobbledygook

XSL•FO
RenderX

Original Paper

Health Information Discrepancies Between Internet Media and Scientific Papers Reporting on Omega-3 Supplement Research: Comparative Analysis

Daryl Nault[1,2], MS; Ariel Beccia[1,3], MS; Haruka Ito[1], MS; Sarah Kashdan[1], MS; Angela Senders[1], ND, MCR

[1]Helfgott Research Institute, School of Research and Graduate Studies, National University of Natural Medicine, Portland, OR, United States
[2]Research Department, Maryland University of Integrative Health, Laurel, MD, United States
[3]Quantitative Health Sciences, University of Massachusetts Medical School, Worcester, MA, United States

Corresponding Author:
Daryl Nault, MS
Helfgott Research Institute
School of Research and Graduate Studies
National University of Natural Medicine
049 SW Porter Street
Portland, OR, 97201
United States
Phone: 1 971 704 1172
Email: dnault@muih.edu

Abstract

Background: Dietary supplements are the most used complementary and alternative health modality in the United States, and omega-3 supplements continue to be the most popularly used nonvitamin or nonmineral supplements by adults. Users of dietary supplements report that they obtain health guidance from internet media resources, but there is question as to whether or not these resources provide the necessary evidence to guide health decisions. Current evidence suggests that there is a mistranslation occurring somewhere between researchers and the media.

Objective: The aim of this study was to conduct a comparative cross-sectional analysis to identify areas of discordance created when science is translated from the laboratory to Web-based news media.

Methods: A Google news search provided our convenience sample of 40 omega-3 supplement–based media reports stratified by the years 2009 to 2012. Media reports (n=17) were compared with the corresponding scientific papers for content. Report and scientific paper content were extracted using commonly accepted reporting guideline domains, and domains were then compared for detecting underlying omissions or mistranslations in reporting. Mean scores for all of the scientific papers and media reports were assessed for each domain.

Results: Scientific papers (n=14) generally maintained a mean close to complete for each reporting domain. The only domain where there was not a significant difference between media and scientific reporting match was within the objectives domain ($\chi^2_1=$ 0.8, P=.36). Media reports (n=17) more frequently reported potential caveats and warnings for consumers with a mean domain for caveat reporting of 0.88, with possible scores falling between 0 and 1.

Conclusions: There are inherent differences in the intended audience, structure, and goals in scientific and media communications. These differences should be explored further, and consumers should be made aware of them. Additional considerations for balanced reporting and reader accessibility are also necessary to take into account and are explored further in this analysis.

KEYWORDS

XSL•FO
RenderX

Introduction

Background

More than half of the US adults take dietary supplements, making supplementation the most used complementary and alternative health modality [1-5]. Since 2007, omega-3s have been the most used nonvitamin and nonmineral dietary supplement according to the National Health Interview Survey (NHIS) data [1,6]. In one large survey, 40% to 45% of respondents said that they had used internet and news media as resources to obtain dietary supplement information [5]. Conversely, less than a quarter (23%) of respondents using dietary supplements on the most recent NHIS Alternative Health Survey reported that their use was because of the advice of a health care professional [6]. What is possibly more concerning is that about 25% of these same users of dietary supplement had not disclosed their use of supplement to their doctor [7]. Patients attempting to improve their health through the use of dietary supplements may have the best intentions in mind, but they may potentially be causing more harm than good. For this reason, it is important to identify the information obtained from Web-based resources so that we might improve the ways we inform these proactive health information consumers.

Despite how frequently their advice is followed, Web-based news resources may contain inaccurate information. It has been suggested in the past that Web-based health information can be misleading or exaggerated, especially when compared with scientific evidence [8-13]. These mistranslations may be more likely to occur between health research and media reporting [10-13]. This identifies *what* the problem is, but there is no current definitive answer to *where* these mistranslations may occur. Therefore, we sought to investigate potential discrepancies between publicly available Web-based health information (Web media reports) and the corresponding scientific evidence upon which those reports were based. Using a ubiquitous search engine fitted with specific parameters, we identified and extracted scientific evidence presented in Web-based media reports of omega-3 supplement research. Media-reported information was then compared with the scientific paper that was referenced. An analysis of the content provided in each resource will hopefully clarify where scientific and media dissemination differences occur. These findings may serve to better educate consumers of the health information on the information they may miss out on, especially when they do not fully communicate decisions with health care professionals.

Omega-3 Supplement Reports From 2009 to 2012

We chose to focus on a commonly used supplement during the time when use was most prevalent. The number of peer-reviewed scientific papers on omega-3 supplements began to rise in the early 2000s [14]. With an increase in omega-3 research, there was a corresponding rise in public interest. Between 2007 and 2012, popularity of use of omega-3 supplement nearly doubled (23% increase), whereas other supplements saw little to no increases in use during the same period [1]. The public has also expressed a growing interest in finding information about omega-3s. In early 2010, Google term-frequency reports show searches for *fish oil* and *omega-3* increased up until their peak popularity [15]. To capture a sample of media reports that were most likely influential, we limited our search to news media reports on omega-3 supplements posted between 2009 and 2012.

Objectives

The descriptive objective of this comparative analysis was to summarize how omega-3 supplement guidance was provided through Web-based news media (2009-2012). Our comparative objective was to identify where reporting gaps occurred between applicable media reports and corresponding scientific publications. As exploratory objectives, we also took balance of reporting, readability, and accessibility into account to demonstrate additional factors that might impact comprehension and decision making of a health information consumer.

Methods

Study Design

One author (DN) performed the Google news search and assessed inclusion criteria for Web-based health blogs and news stories on omega-3 dietary supplements, hereafter referred to as *media reports* with agreement from AS. Individual search pages were saved and archived for future use and transparency. Four authors (AB, HI, SK, and DN) independently extracted data in duplicate from media reports, using a standardized data collection form. Before data extraction, authors met for a brief training session on using the data collection form to improve the consistency of data collection. Next, the same authors independently extracted data in duplicate from each of the corresponding omega-3 dietary supplement scientific papers referenced by the media reports. Finally, DN compared the data extracted from media reports and scientific papers to assess match of information between media and corresponding scientific resources.

Search Strategy for Identification of Media Reports

We searched Google News between January 1 and December 31 for each year from 2009 to 2012, inclusive (Figure 1). This search was performed through an incognito window, without a user log-in, and in a newly downloaded Google Chrome app. All search return pages were archived to preserve the exact returns given at the time they were searched (December 2015). The search query was run once in Google News, and filters were then used to restrict dates of publication. Google PageRank has changed since its inception, but the basic idea behind Google search ranking remains the same: returns are ranked higher based on keyword relevancy and page popularity, as measured by user interest and recurring links [16]. Prior evidence suggests that the first 10 (or first page of) search returns are typically the most viewed, in some cases, receiving close to 100% of link traffic [17-23], and are therefore most likely to have the highest viewership for any particular search. Thus, we used a convenience sample consisting of the top 10 search returns for each year (2009-2012), totaling 40 media reports to be assessed for inclusion criteria.

Figure 1. Media report and corresponding scientific paper search and screening flow diagram. Media reports were initially excluded on the basis that they did not refer to dietary supplements specifically, were editorials or reviews, were based on nonpeer-reviewed evidence, or if their referencing information was not clear enough to link back to a specific scientific paper. Scientific papers were excluded if they were duplicates, they reported solely on dietary supplement use demographic statistics, or were only based on dietary interventions.
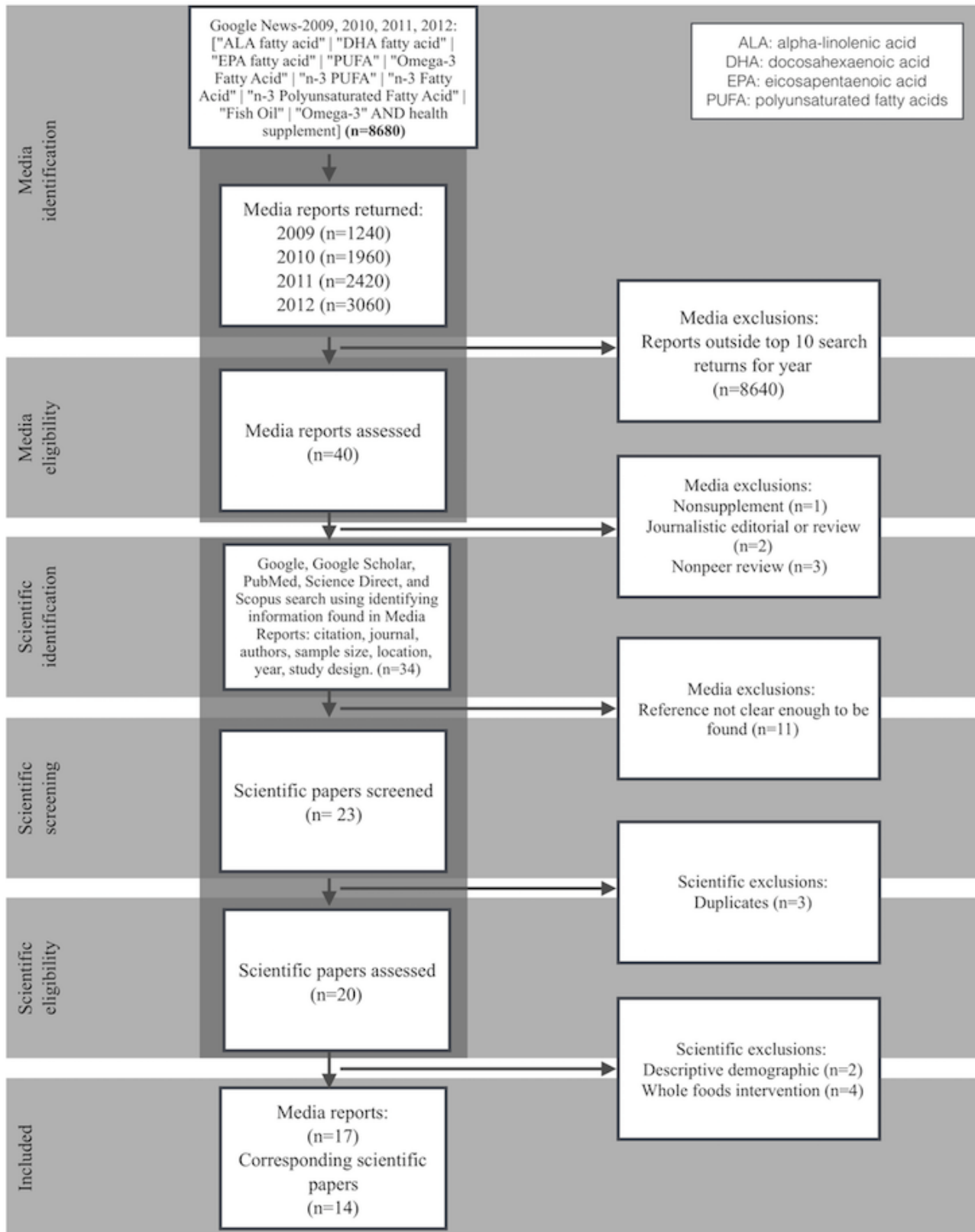
The search query used to obtain relevant media reports was as follows: ["ALA fatty acid" OR "DHA fatty acid" OR "EPA fatty acid" OR "PUFA" OR "Omega-3 Fatty Acid" OR "n-3 PUFA" OR "n-3 Fatty Acid" OR "n-3 Polyunsaturated Fatty Acid" OR "Fish Oil" OR "Omega-3" AND health supplement]. This search query was determined by both the terminology used

in the NHIS to define these health supplements: "Fish oil (or omega-3, DHA, or EPA fatty acid)" and the use of PubMed medical subject headings to find additional descriptors for this class of supplements such as polyunsaturated fatty acids (PUFA), n-3, and alpha-linolenic acid (ALA).

We included health blogs and news reports that were published on Web, in English, and fell within the top-10 returns for each year between January 1, 2009 and December 31, 2012. Media reports about omega-3 supplements that referenced a peer-reviewed scientific paper were included. Media reports that were simply *expert opinion* or *reviews not published in peer-reviewed literature* of omega-3 evidence were excluded, as they did not address 1 primary scientific research publication and therefore would be inappropriate for this appraisal method.

Search Strategy for Identification of Scientific Papers

Any information found within a media report that might identify the corresponding scientific paper was used to locate the full-text manuscript through Google, Google Scholar, PubMed, ScienceDirect, and Scopus. This information included but was not limited to any part or combination of the paper: full citation, author names, journal name, year of publication, sample sizes, location of the study, and study design.

We included papers that reported on controlled and uncontrolled intervention studies, observational studies, and narrative reviews of intervention studies (Figure 1). Both animal and human studies regarding the interventions of omega-3 dietary supplement were included. Cross-sectional reports on use of supplement *descriptive demographics* were excluded. Studies that aimed to alter or quantify a participant's *whole food* dietary consumption of omega-3s were also excluded. There were no restrictions on type or outcome measure of participants, as our purpose was to assess how outcomes were presented to the public through media reports, independent of the outcome type.

Data Extraction

We developed a standardized data extraction form using the EQUATOR Network's guidelines for research reporting (Multimedia Appendix 1; [24]). The following domains were extracted: study objectives, design, study population characteristics, intervention or exposure, comparison or control, outcome measure, participant attrition, statistical analyses, limitations, caveats (or warnings for consumer or clinical use), and results. The following data points were also collected: date of posting, URL, title, and author names. Two review authors independently extracted data from each media report and scientific paper. We used the same data extraction form for both resource types; disagreements were discussed by the authorship team and resolved by consensus.

Primary Analysis

Scientific papers were assessed first, according to whether or not they provided the relevant information from the EQUATOR guidelines. Each data point was coded as either a score of 0 to represent that the scientific paper omitted information that should have been present, or it was given a 1 when information was provided. To assess media report translation of scientific research, we compared the data extracted from media reports

with their corresponding scientific paper. For media, if the content of the report matched the information in the scientific paper, a score of 1 was assigned, and if the media report omitted or did not provide clear enough information to match the scientific paper, a score of 0 was assigned. This resulted in proportions of matching content for each domain (Multimedia Appendices 2 and 3). These proportions were used in subsequent chi-square tests.

Balance Analysis

Assessors provided their interpretation of balance for each report and paper. Balance was defined as, *whether or not alternative explanations for findings were provided to readers*. First, balance was individually assessed at the time of the data collection and then discussed among authors. Consensus was reached with the aid of supporting quotations extracted from each resource independently. We used a multiassessor and consensus approach to validate potentially subjective assessments [25,26]. To view balance in light of the reporting quality, we averaged all domain means for each piece and compared corresponding media and scientific papers.

Readability and Accessibility

During the data collection process, assessors extracted direct quotations to provide support for their assessments during consensus discussions. This resulted in a collection of quotations describing the *results* reported in both the media reports and scientific papers. Using these direct quotations, we assigned a Flesch-Kincaid readability score to each paper and report. The Flesch-Kincaid grade reading level indicates how difficult a passage is to comprehend based on word and sentence length; a lower score indicates an easier-to-read passage. Ideally, papers for a general audience would be written below a twelfth grade reading level [27-30]. We chose to assess readability of the Results section of the reports and papers, as this section is generally the most technically written. Inclusion of a citation to the scientific paper within media reports and the full-text accessibility of the paper was also recorded.

Changes to Original Protocol

We had not anticipated to come across media reports with unclear details on the type of intervention administered. When this was realized, we applied the intervention exclusion criteria of the media reports to the scientific articles as well. In response to peer-review considerations, we also chose to add an additional comparative analysis using chi-square tests of proportions.

Results

Media Report Characteristics

Overall, 40 media reports were screened from the Google search returns. Of the 40, 17 (17/40, 43%) met inclusion criteria for this study and were retained for further analysis to referenced scientific papers (Table 1). However, many of the scientific papers referenced (11/17, 65%) could not be found with the information these media reports provided. In total, 5 of the media reports returned were on omega-3 supplementation, but they lacked the scientific rigor of peer review, so they were not included in the primary analysis. Only one of the media reports

returned was clearly not based on the use of omega-3 supplements. When corresponding scientific papers were read in full, additional 6 were removed according to inclusion or exclusion criteria.

Scientific Paper Characteristics

Table 2 provides the journals and study designs for each of the 14 scientific papers referenced by 17 media reports within this study. In total, 2 (2/14, 14%) scientific papers were referenced multiple times by different media outlets [31,32]. More than half (9/14, 64%) of the scientific papers summarized by included media reports were randomized controlled trials (RCTs; 7/9, 78%) or secondary data analyses derived from RCTs (2/9, 22%). Overall, 3 (3/14, 21%) of the papers were systematic reviews and meta-analyses or systematic reviews or meta-analyses. The remaining 2 (2/14, 14%) papers were narrative literature reviews.

The study characteristics of the scientific papers are presented in Multimedia Appendix 4. Participants varied in age from infants to those in their 80s and included both males and females located in a variety of different countries. Method of exposure to omega-3 fatty acids varied across studies from capsules to enriched products. Outcomes assessed included cognitive function, mood disorders, cardiovascular function, cancer, infant morbidity, and adult mortality.

Primary Analysis

For our primary analysis, we calculated the proportion of parameters that were met within each reporting domain (Table 3) for each media report and scientific paper reporting on omega-3 dietary supplements (Multimedia Appendices 2 and 3). The only difference in proportions that was not statistically significant between media and scientific reporting was within the objectives domain ($\chi^2_1=0.8$; $P=.36$). All other domain comparisons were significant ($P<.05$).

We then generated an average domain score for all media reports and for all scientific papers (Figure 2). Scores closer to 1 signify that more parameters were met, and reporting was more complete. For each reporting domain, domain mean values for scientific papers were above 0.80, with the exception of reporting caveats mean of 0.36. Media reports show a wider variation in domain means, with study objectives mean of 0.98, limitations mean of 0.71, and caveats mean of 0.88 as the highest mean reporting scores. Media reports reviewed in this study had the lowest domain means for reporting statistical analysis with a mean of 0.18 and participant attrition with a mean of 0.20.

Table 1. Descriptions of all media reports obtained.

Descriptions	Type of report, n (%)							
	All media reports (N=40)	Nonsupplement (n=1)	Nonpeer-review (n=3)	Journalistic editorial or review (n=2)	Reference unclear (n=11)	Descriptive demographic report (n=2)	Whole foods (n=4)	Included reports (n=17)
Year of media publication								
2009	10 (25)	0 (0)	1 (33)	0 (0)	3 (27)	1 (50)	1 (25)	4 (24)
2010	10 (25)	0 (0)	0 (0)	1 (50)	5 (46)	0 (0)	1 (25)	3 (18)
2011	10 (25)	0 (0)	2 (67)	1 (50)	2 (18)	1 (50)	2 (50)	2 (12)
2012	10 (25)	1 (100)	0 (0)	0 (0)	1 (9)	0 (0)	0 (0)	8 (47)
Evidence types included[a]								
Evidence-based information	30 (75)	1 (100)	1 (33)	2 (100)	3 (27)	2 (100)	4 (100)	17 (100)
Expert opinion	31 (78)	1 (100)	2 (67)	2 (100)	8 (73)	2 (100)	2 (50)	14 (82)
Anecdotal evidence	2 (5)	0 (0)	1 (33)	0 (0)	1 (9)	0 (0)	0 (0)	0 (0)
Journalistic opinion	6 (15)	0 (0)	1 (33)	0 (0)	3 (27)	1 (50)	0 (0)	1 (6)
Topics discussed[a]								
Effectiveness	22 (55)	1 (100)	0 (0)	2 (100)	4 (36)	2 (100)	2 (50)	11 (65)
Quality	5 (13)	0 (0)	2 (67)	2 (100)	1 (9)	0 (0)	0 (0)	0 (0)
Safety	8 (20)	0 (0)	2 (67)	1 (50)	4 (36)	1 (50)	0 (0)	0 (0)
Use for multiple specific health conditions	10 (25)	0 (0)	0 (0)	2 (100)	2 (18)	0 (0)	1 (25)	5 (29)
Use for one specific health condition	17 (43)	1 (100)	0 (0)	0 (0)	2 (18)	0 (0)	3 (75)	11 (65)
Use for overall wellness	5 (13)	0 (0)	0 (0)	1 (50)	4 (36)	0 (0)	0 (0)	0 (0)

[a]More than one response possible for category.

Table 2. Scientific papers and corresponding media reports.

Journal	Media reports
Journal of the American Medical Association (May 2012) [31][a]	• Fish oil delivers few heart benefits, study finds, ABC News (April 2012) [33] • Weighing the Evidence on Fish Oils for Heart Health, NY Times (April 2012) [34]
Journal of the American Medical Association (September 2012) [32][a]	• Fish oil's heart benefits may be overstated, CNN Health (September 2012) [35] • *Flawed* omega-3 meta-analysis harms public health: GOED, NutraIngredients (September 2012) [36] • Questioning the Superpowers of Omega-3 in Diets, Wall Street Journal (October 2012) [37]
Cochrane Database Systematic Reviews (June 2012) [38][a]	• Fish Oil Fail: Omega-3s May Not Protect Brain Health After All, Time (June 2012) [39]
Pediatrics (September 2011) [40][b]	• Omega-3 can reduce risk of colds in babies, Telegraph UK (August 2011) [41]
Journal of the American Medical Association (January 2009) [42][b]	• Study: Fish Oil for Preemies May Boost Cognition, Time (January 2009) [43]
Expert Review of Cardiovascular Therapies (July 2009) [44][b]	• Try Fish Oil Instead of Drugs, To Your Health (February 2009) [45]
Journal of the American Medical Association (October 2010) [46][b]	• Fish oil doesn't benefit new moms, babies, CNN Health (October 2010) [47]
Journal of the American Medical Association (November 2010) [48][b]	• Fish oil ingredient doesn't slow Alzheimer's, CNN Health (November 2010) [49]
Alzheimer's & Dementia (January 2010) [50][b]	• Fish Oil Supplements Boost Memory DHA Supplements Help Stave Off *Senior Moments*, WebMD Health News (July 2009) [51]
Research in Developmental Disabilities (February 2010) [52][b]	• Fallacy of fish oil revealed as study finds supplements DON'T boost children's brain power, Daily Mail UK (April 2010) [53]
The American Journal of Clinical Nutrition (July 2012) [54][c]	• Omega 3 or B vitamins fail to benefit depressive symptoms, but low doses may be at fault, NeutraIngredients (June 2012) [55]
Cancer Prevention Research (October 2013) [56][c]	• Prostate/Prostate Cancer Nutrition/Diet Low-Fat Diet With Fish Oil Supplements Slows Growth Rate of Prostate Cancer Cells, Medical News Today (October 2011) [57]
Journal of the American College of Cardiology (August 2009) [58][d]	• Daily Omega-3s Recommended for Heart Benefits of Omega-3 Fatty Acids Prompt New Dosage for Heart Health, WebMD Health News (August 2009) [59]
The Journal of Lipid Research (August 2012) [60][d]	• Evidence is strong for omega-3's heart health benefits: Linus Pauling Review, NutraIngredients (November 2012) [61]

[a]Systematic review or meta-analysis.

[b]Randomized controlled trial.

[c]Secondary data from a randomized controlled trial.

[d]Narrative literature review.

Table 3. Percent match comparisons by domain.

Domain[a]	Media (n=17), n (%)	Scientific (n=14), n (%)	χ^2 (df)	P value
Analysis match	12 (18)	52 (93)	69.6 (1)	<.001
Attrition match	17 (20)	63 (90)	68.8 (1)	<.001
Design match	8 (47)	14 (100)	10.4 (1)	.001
Population match	61 (46)	108 (97)	69.6 (1)	<.001
Exposure match	48 (48)	80 (95)	45.9 (1)	<.001
Results match	19 (56)	28 (100)	16.3 (1)	<.001
Comparison match	40 (42)	70 (86)	31.0 (1)	<.001
Limitations match	12 (71)	14 (100)	4.9 (1)	.03
Outcomes match	39 (59)	45 (80)	5.3 (1)	.02
Objectives match	50 (98)	42 (100)	0.8 (1)	.36
Caveats match	15 (88)	5 (36)	9.3 (1)	.002

[a]Multiple domains had more than one parameter, as shown in Multimedia Appendix 2. Therefore, it was possible for a domain to have a number of matches higher than the actual sample size.

Figure 2. Domain mean match scores within reporting guideline domains for omega-3 dietary supplement media reports and scientific papers. Scores represent the proportion of the domain that was matched across all media reports or scientific papers. Points are connected by lines for visualization purposes.
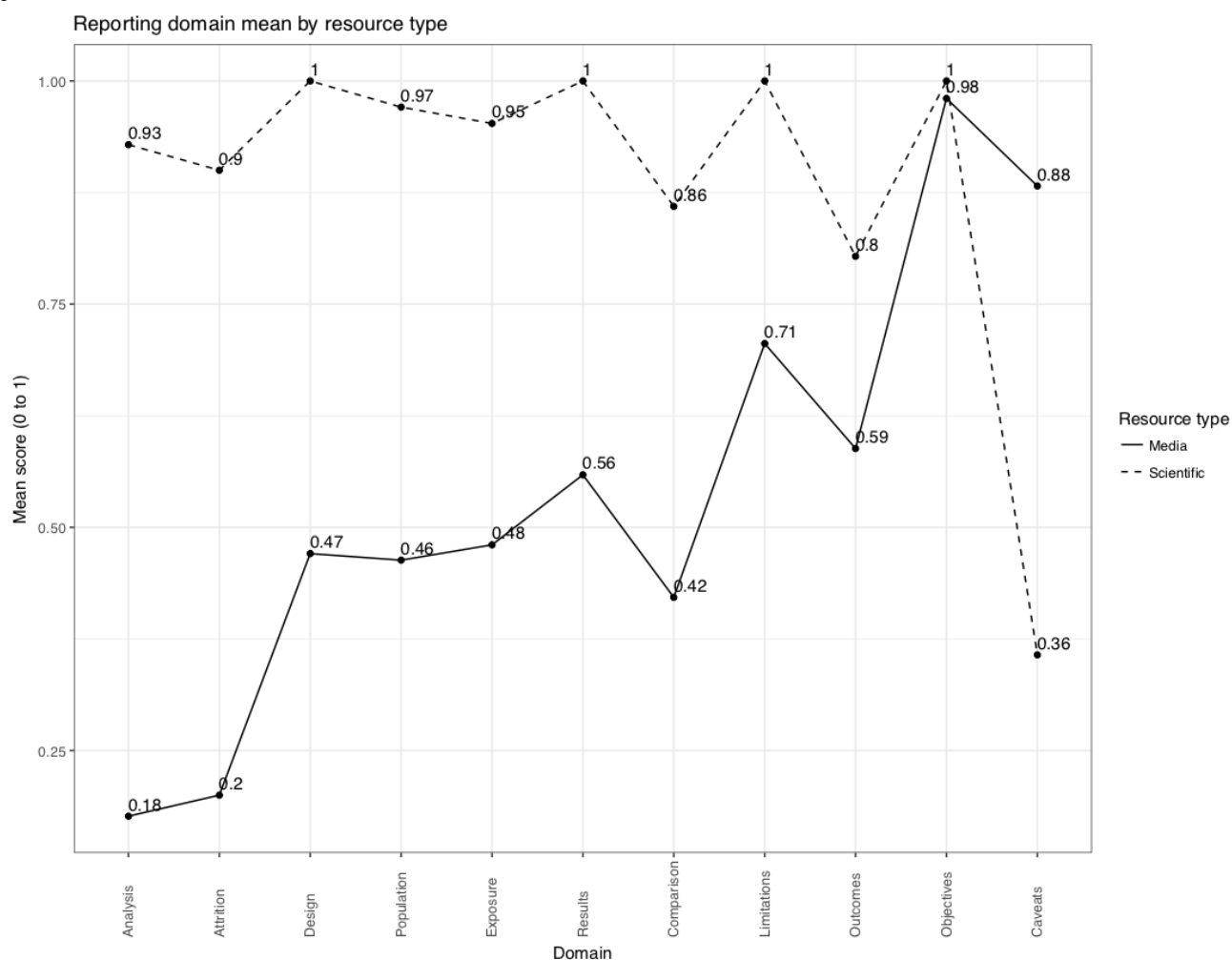
Balance Analysis

A total of 13 of the 17 (13/17, 76%) omega-3 dietary supplement report or paper pairings were considered to be balanced in reporting evidence (Figure 3, top panel), meaning that they presented the reader with multiple perspectives on the evidence at hand. The scientific paper that scored lowest in this group (at 0.58) was a narrative review; a corresponding media report or paper for that scored 0.35 across domain means. Although this pairing did report on multiple perspectives and therefore was found to be balanced in perspective, it contained far less detail than other report or paper pairings, including other nonsystematic reviews.

In 2 of the 17 (2/17, 12%) pairings, we considered neither media report nor scientific paper to be balanced (Figure 3, middle panel). These media reports held the lowest domain mean (0.40) and reported on RCTs. Their corresponding scientific papers scored 0.90 and 0.88, respectively, across domains. Finally, in 2 of the 17 (2/17, 12%) report or paper pairings (Figure 3, bottom panel), the scientific paper was found to be balanced, whereas the media report was assessed as unbalanced. The scientific papers that these unbalanced media reports were based on held domain means of 0.83 and 0.84. Corresponding media report domain means for these 2 pairings were 0.47 and 0.65, respectively. In none of the cases, the media report was assessed as balanced with a corresponding unbalanced scientific paper.

Readability and Accessibility

Readability was determined through Flesch-Kincaid grade level readability scores (Figure 4). Accessibility was examined through the use of clear references in the media reports focused on omega-3 supplement, and the public availability of the scientific paper in full text. A total of 76% (13/17) of the media reports used language above a twelfth grade reading level. In addition, 59% (10/17) of the media reports did not supply a reference for the corresponding scientific paper. Moreover, 64% (9/17) of the scientific papers were not available in full text to the public. In addition, 41% (7/17) of the media reports were posted before the corresponding scientific paper was published.

Figure 3. Resource balance and mean score: comparison of media report and corresponding scientific paper overall mean score, including whether or not resources were balanced in their reporting.
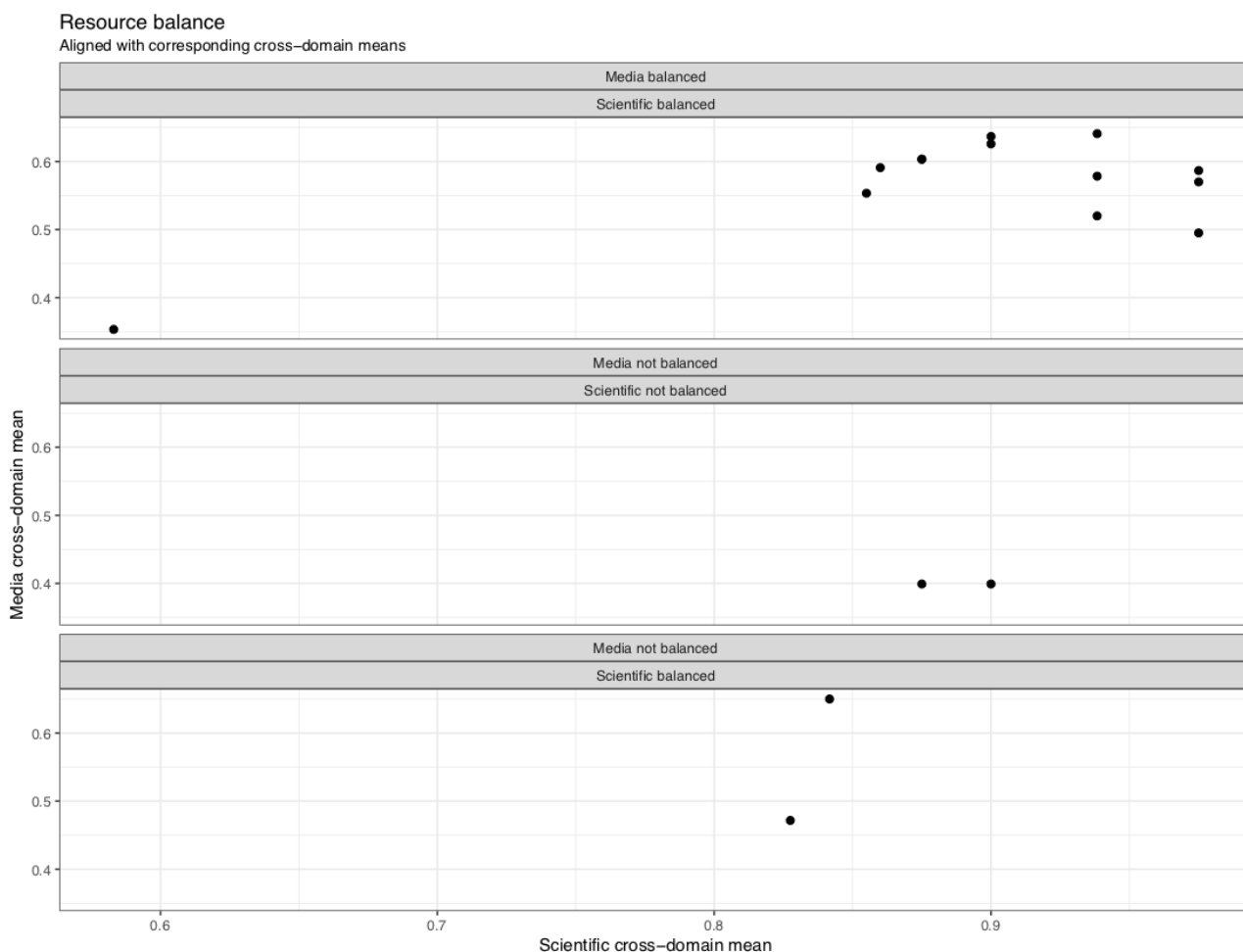
Figure 4. Comparison of media report and corresponding scientific paper reading grade level by whether or not a reference was provided in media report. Plot point represents whether or not scientific paper was available in full text to the general public or not.
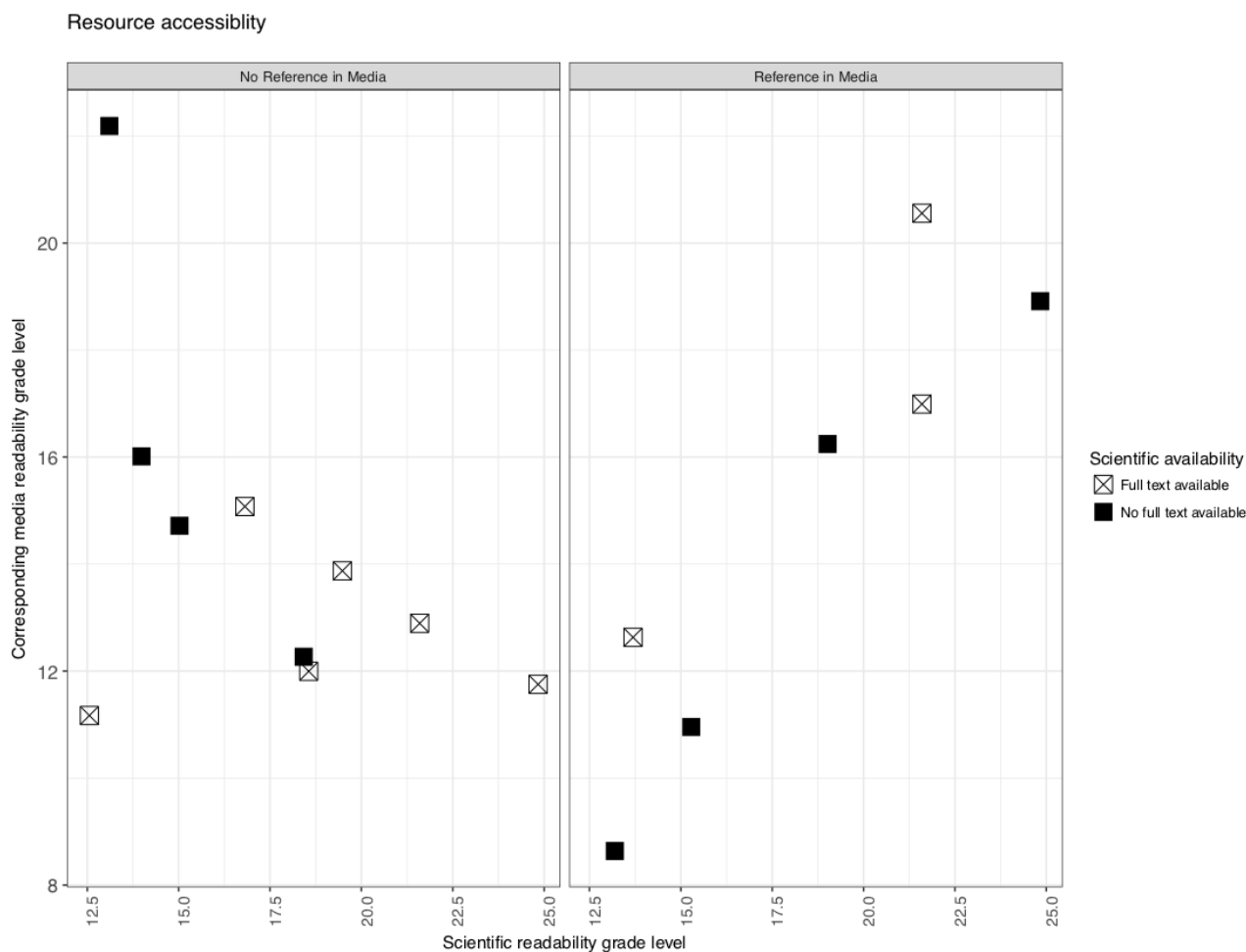
Discussion

Principal Findings

The reporting domains where media and scientific resources converge and diverge in this study speak to the intended audience and purpose of each resource. These results suggest that more technical aspects of a study, such as statistical analysis, participant attrition, participant characteristics, and even specific study design were more frequently omitted from media reports. The intended audience for the report itself may explain these omissions. Media reports, targeting a consumer audience, reported caveats or warnings about the direct use of the supplement more frequently than scientific papers did. It is important to recognize the intentions of each type of resource, as this likely dictates consumer's takeaway knowledge of the health information contained within.

Health information may be obtained via numerous Web-based outlets. However, not all consumers possess the capabilities to detect what is relevant or credible, nor do they necessarily always discuss the information with their physician [62,63]. Consumers have expressed mistrust in and confusion with Web-based health information in the past [18,62,64]. This confusion may be because, in part, of the format of the information presented. Media reports and scientific papers differ

when it comes to the structure in which they are presented. Scientific papers are pyramidically structured, building on background, setting, and methods information, and finally presenting the results and conclusions of data analysis last. In contrast, media reports are typically presented in the opposite format. To attract the most viewers at once, media reports often present the *exciting* results and conclusions first, followed by the relevant details of the study they are representing [65]. This approach may serve to capture more readers, but there could be a detrimental trade-off in quality. If readers are given the most interesting evidence in lieu of pertinent details, the story remains incomplete. For these reasons, it was expected that media reports would contain less detailed information than the scientific papers would.

As a result of these discrepancies, both science and media may see better reader comprehension if their reporting goals were more apparent to consumers. Consumers should also be better informed on the potential dangers of using a single study to make a health care decision without input from a health care professional. One way of clarifying goals of a resource and reducing bias would be to provide balanced information. The results of our exploratory balance analysis suggest that when either science or media reported alternative viewpoints and evidence, their mean reporting score was higher than those that only reported one perspective (equating to unbalanced

reporting). These findings indicate that there may be more to the relationship between reporting quality and balance overall. Including multiple viewpoints may serve to further educate consumers and instill trust in the scientific process through transparency. Consumers may not inherently notice when a piece is unbalanced or one-sided, but according to these findings, balance may be somewhat related to reporting quality, overall regardless.

When it comes to readability, we saw a stark contrast between the intended audience of media reports and scientific papers emerge. The readability of scientific papers never fell below a twelfth grade level, indicating that the information was geared toward a more educated demographic. This was to be expected, as scientific papers are intended to be read by the scientific community. It is interesting though that a majority of the media reports also scored above a twelfth grade reading level. There are limitations to the Flesch-Kincaid readability score here, which should be mentioned too. For instance, the inability to use less technical words when describing scientific principles in the Results section may have increased readability score of a media report more so than if they were reporting on a nonscientific topic.

Accessibility to scientific papers was somewhat limited for these resources. Direct references were not provided in many of the media reports (10/17, 59%). Whether that was because of the scientific paper itself not yet being published at the time the report was written or because of omission by the media report is unclear. However, the lack of reference still restricts the information consumers are able to access. Failure to include an appropriate reference to consumers places obstacles for those who want to and know they should cross-check reported findings of a study. This reduces the accessibility of scientific information for a general audience. At the time this study was run, full-text (public) availability was present for only about half of the scientific papers (n=8). Without access to a full-text paper, consumers do not have the option to verify the information within media reports if they want to bring it to their health care provider for further guidance.

Strengths and Limitations

Prior research has taken a similar approach to ours in looking at the information presented by both press releases and corresponding print newspapers [10,12,66]. One of these studies also used Consolidated Standards of Reporting Trials reporting guidelines to grade their assessments of news reporting on press releases of RCTs, which supports the methodology that we used here [66]. These studies align with our findings in the sense that the scientific research in many ways relates to the corresponding media report, whether it is in the tone, quality, or overall coverage. Another similar paper compared print media with scientific guidelines as well [13], but to our knowledge, this is the first study to view quality comparisons between *Web-based* news media and corresponding health research. It is certainly the first to focus on this popular health supplement as a topic of interest. One weakness in comparing our work with prior literature was that in the aforementioned studies, which viewed scientific claims made in print media, researchers found a tendency for media reports to under-report limitations and

overlook potential risks [25,26]. We found nearly the opposite, in that limitations were fairly well represented in the media reports (mean=0.71), and these sources provided risks, warnings, or caveats for consumers more frequently than the corresponding scientific papers did. This may be dependent on the type of scientific reporting.

The largest limitation in this study was the small sample of papers that could actually be compared. A 4-year retrospective sample consisting of 40 media reports is by no means comprehensive for all of the literature available. Knowing this, we sought to find the papers that would be considered the *most* influential to increased omega-3 supplement use trends. Google's PageRank Algorithm sifts the most influential returns to the top of the news search [16,17]. By choosing the top 10 returns in our stratified convenience sample, we were selecting the returns that most users were likely to have had exposure to, and therefore, influence from [16-23,67]. Several media reports were initially excluded (n=11), simply because they lacked the necessary information to clearly find a corresponding scientific paper. This was a large limitation to the number of papers and reports that could be compared with one another. It does also demonstrate another issue that health information consumers might experience with regard to accessibility though. Even if consumers know to cross-check a media report for missing information, there is no way for them to accurately do so when they cannot identify the scientific paper.

Further sample size restrictions were made when papers or reports failed to meet inclusion criteria. This inclusion criterion was set in place to examine only those report or paper combinations that were specifically about the effects of omega-3 dietary supplementation on health. We had expected that some papers may be vague on details of an intervention, and perhaps some might mention supplementation and dietary changes together. What we had not expected to see were vague media reports about whether or not participants had been given a supplement or dietary intervention (n=4). In these instances, it was not until we reviewed the scientific paper that we were able to discern whether or not the pairing fits inclusion criteria. Providing consumers with vague information on the type of intervention used is potentially misleading, especially considering that these consumers might make health care decisions based off of it. Further work in this area of study could be improved with alterations to the inclusion and exclusion criteria that would allow for more papers to be obtained for the comparison. This might include incorporating more dietary supplements or even incorporating more study designs.

A strength of this study was that it was focused on Web-based news media, rather than print media, which appears to be the primary focus of much of the prior work in this field [10,12,13,25,26,66]. Previous studies have indicated that there is a difference in the information presented by these 2 forms of media reporting, in that Web-based media may present a wider variety of themes when a scientific topic is searched [20]. This could also be seen as a drawback because different demographic groups may be more or less likely to obtain their information from print versus Web-based media sources. Therefore, the applicability of these results may be limited to those who tend to use Web-based news media resources.

The final limitation of this study was the use of somewhat subjective assessments of balance. Although the quality of reporting and accessibility could be determined by either a match of information or well-known readability scoring, balance is a bit more nuanced. The assessors did, however, individually decide how balanced they felt each piece was, and they were also required to support their judgment with specific examples from the resource to come to a consensus. To decrease subjectivity further, assessors were provided with specific definitions of *balance* for assessing it in a piece. Future work could include a deeper look into the ways that balance the impacts of the health information consumer's interpretation and application of Web media information to their health care decision making.

Authors' Contributions

Authors AS and DN were responsible for the conception and design of the study as well as data analysis and interpretation. DN drafted the paper, and AS provided supervision of the project from start to finish. Data collection and independent assessment were performed by DN, AB, HI, and SK. All authors discussed results and provided critical revisions to the final piece.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Guidelines used for data extraction: parameters collected under each reporting domain in this study according to respective EQUATOR reporting guideline's item numbers.

[PDF File (Adobe PDF File), 120KB - ijmr_v7i2e15_app1.pdf]

Multimedia Appendix 2

Percent of scientific resources matching. This figure illustrates the percent of the sample where scientific articles matched to the appropriate EQUATOR guideline.

[PNG File, 121KB - ijmr_v7i2e15_app2.png]

Multimedia Appendix 3

Percent of media resources matching. This figure illustrates the percent of the sample where media reports matched to their corresponding scientific article. If the scientific article did not contain relevant information, media reports also not containing the information were marked as "matched;" therefore, media reports could only be held accountable for reporting what scientific articles presented.

[PNG File, 122KB - ijmr_v7i2e15_app3.png]

Multimedia Appendix 4

Descriptive characteristics of peer-reviewed scientific articles included in this analysis.

[PDF File (Adobe PDF File), 188KB - ijmr_v7i2e15_app4.pdf]

References

1. Wu CH, Wang CC, Tsai MT, Huang WT, Kennedy J. Trend and pattern of herb and supplement use in the United States: results from the 2002, 2007, and 2012 national health interview surveys. Evid Based Complement Alternat Med 2014;2014:872320 [FREE Full text] [doi: 10.1155/2014/872320] [Medline: 25574184]
2. Barnes PM, Powell-Griner E, McFann K, Nahin RL. Complementary and alternative medicine use among adults: United States, 2002. Adv Data 2004 May 27(343):1-19. [Medline: 15188733]
3. Blackwell DL, Lucas JW, Clarke TC. Summary health statistics for U.S. adults: national health interview survey, 2012. Vital Health Stat 10 2014 Feb(260):1-161 [FREE Full text] [Medline: 24819891]
4. Bardia A, Nisly NL, Zimmerman MB, Gryzlak BM, Wallace RB. Use of herbs among adults based on evidence-based indications: findings from the National Health Interview Survey. Mayo Clin Proc 2007 May;82(5):561-566 [FREE Full text] [doi: 10.4065/82.5.561] [Medline: 17493422]
5. Rozga MR, Stern JS, Stanhope K, Havel PJ, Kazaks AG. Dietary supplement users vary in attitudes and sources of dietary supplement information in East and West geographic regions: a cross-sectional study. BMC Complement Altern Med 2013 Jul 30;13:200 [FREE Full text] [doi: 10.1186/1472-6882-13-200] [Medline: 23899171]
6. Bailey RL, Gahche JJ, Miller PE, Thomas PR, Dwyer JT. Why US adults use dietary supplements. JAMA Intern Med 2013 Mar 11;173(5):355-361. [doi: 10.1001/jamainternmed.2013.2299] [Medline: 23381623]

7.   Jou J, Johnson PJ. Nondisclosure of complementary and alternative medicine use to primary care physicians: findings from the 2012 National Health Interview Survey. JAMA Intern Med 2016 Apr;176(4):545-546. [doi: 10.1001/jamainternmed.2015.8593] [Medline: 26999670]

8.   Alnemer KA, Alhuzaim WM, Alnemer AA, Alharbi BB, Bawazir AS, Barayyan OR, et al. Are health-related tweets evidence based? Review and analysis of health-related tweets on twitter. J Med Internet Res 2015 Oct 29;17(10):e246 [FREE Full text] [doi: 10.2196/jmir.4898] [Medline: 26515535]

9.   Chang T, Verma BA, Shull T, Moniz MH, Kohatsu L, Plegue MA, et al. Crowdsourcing and the accuracy of online information regarding weight gain in pregnancy: a descriptive study. J Med Internet Res 2016 Apr 7;18(4):e81 [FREE Full text] [doi: 10.2196/jmir.5138] [Medline: 27056465]

10.   de Semir V, Ribas C, Revuelta G. Press releases of science journal articles and subsequent newspaper stories on the same topic. J Am Med Assoc 1998;280(3):294-295. [Medline: 9676688]

11.   Goldacre B. Preventing bad reporting on health research. Br Med J 2014 Dec 9;349:g7465. [Medline: 25498123]

12.   Schwartz LM, Woloshin S, Andrews A, Stukel TA. Influence of medical journal press releases on the quality of associated newspaper coverage: retrospective cohort study. Br Med J 2012 Jan 27;344:d8164 [FREE Full text] [Medline: 22286507]

13.   Cooper BE, Lee WE, Goldacre BM, Sanders TA. The quality of the evidence for dietary advice given in UK national newspapers. Public Underst Sci 2012 Aug;21(6):664-673. [doi: 10.1177/0963662511401782] [Medline: 23832153]

14.   Elsevier. Scopus URL: https://www.elsevier.com/solutions/scopus [accessed 2017-06-29] [WebCite Cache ID 6rc3rWCs2]

15.   Google Trends. Omega-3 fish oil URL: https://trends.google.com/trends/explore?cat=45&date=all&geo=US&q=omega-3%20%7C%20fish%20oil [WebCite Cache ID 70OrJBhtW]

16.   Google. How search algorithms work URL: https://www.google.com/search/howsearchworks/algorithms/ [accessed 2018-06-22] [WebCite Cache ID 70OjG9Qr2]

17.   Page L, Brin S, Motwani R, Winograd T. The pagerank citation ranking: bringing order to the web. Stanford InfoLab; 1998. URL: http://webcache.googleusercontent.com/search?q=cache:http://ilpubs.stanford.edu/422/1/1999-66.pdf [WebCite Cache ID 72BkD6ra9]

18.   Eysenbach G, Köhler C. How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. BMJ 2002 Mar 9;324(7337):573-577 [FREE Full text] [Medline: 11884321]

19.   Chitika Insights: The Value of Google Result Positioning. Westborough, MA: Chitika; 2013. URL: http://info.chitika.com/uploads/4/9/2/1/49215843/chitikainsights-valueofgoogleresultspositioning.pdf [WebCite Cache ID 70OjQPlgY]

20.   Cacciatore MA, Anderson AA, Choi D, Brossard D, Scheufele DA, Liang X, et al. Coverage of emerging technologies: a comparison between print and online media. New Media Soc 2012 Mar 21;14(6):1039-1059. [doi: 10.1177/1461444812439061]

21.   Cho J, Roy S. Impact of search engines on page popularity. ACM Press; 2004. Presented at: Proceedings of the 13th international conference on World Wide Web; May 17-20, 2004; New York, New York, USA p. 20-29. [doi: 10.1145/988672.988676]

22.   Hindman K, Tsioutsiouliklis K, Johnson JA. Googlearchy: how a few heavily-linked sites dominate politics on the web. 2003. URL: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.160.8347 [accessed 2018-06-22] [WebCite Cache ID 70OjiBjSY]

23.   Pan B, Hembrooke H, Joachims T, Lorigo L, Gay G, Granka L. In Google we trust: users' decisions on rank, position, and relevance. J Comput Mediat Commun 2007 Apr;12(3):801-823. [doi: 10.1111/j.1083-6101.2007.00351.x]

24.   UK EQUATOR Centre. The EQUATOR Network URL: http://www.equator-network.org/ [accessed 2017-06-29] [WebCite Cache ID 6rc3iOH6E]

25.   Bubela TM, Caulfield TA. Do the print media "hype" genetic research? A comparison of newspaper stories and peer-reviewed research papers. Can Med Assoc J 2004 Apr 27;170(9):1399-1407. [Medline: 15111473]

26.   Mcinnes N, Haglund BJ. Readability of online health information: implications for health literacy. Inform Health Soc Care 2011 Dec;36(4):173-189. [doi: 10.3109/17538157.2010.542529] [Medline: 21332302]

27.   Ley P, Florio T. The use of readability formulas in health care. Psychol Health Med 1996 Feb;1(1):7-28. [doi: 10.1080/13548509608400003]

28.   Cheng C, Dunn M. Health literacy and the internet: a study on the readability of Australian online health information. Aust N Z J Public Health 2015 Aug;39(4):309-314. [doi: 10.1111/1753-6405.12341] [Medline: 25716142]

29.   Cotugna N, Vickery CE, Carpenter-Haefele KM. Evaluation of literacy level of patient education pages in health-related journals. J Community Health 2005 Jun;30(3):213-219. [Medline: 15847246]

30.   Watad A, Bragazzi NL, Brigo F, Sharif K, Amital H, McGonagle D, et al. Readability of wikipedia pages on autoimmune disorders: systematic quantitative assessment. J Med Internet Res 2017 Jul 18;19(7):e260 [FREE Full text] [doi: 10.2196/jmir.8225] [Medline: 28720555]

31.   Kwak SM, Myung SK, Lee YJ, Seo HG, Korean Meta-analysis Study Group. Efficacy of omega-3 fatty acid supplements (eicosapentaenoic acid and docosahexaenoic acid) in the secondary prevention of cardiovascular disease: a meta-analysis of randomized, double-blind, placebo-controlled trials. J Am Med Assoc 2012 May 14;172(9):686-694. [doi: 10.1001/archinternmed.2012.262] [Medline: 22493407]

32. Rizos EC, Ntzani EE, Bika E, Kostapanos MS, Elisaf MS. Association between omega-3 fatty acid supplementation and risk of major cardiovascular disease events: a systematic review and meta-analysis. J Am Med Assoc 2012 Sep 12;308(10):1024-1033. [doi: 10.1001/2012.jama.11374] [Medline: 22968891]

33. Gann C, ABC News Medical Unit. Fish Oil Delivers Few Heart Benefits, Study Finds. ABC News; 2012 Apr 9. URL: https://abcnews.go.com/Health/fish-oil-delivers-heart-benefits-study/story?id=16103420 [accessed 2018-06-23] [WebCite Cache ID 70OiyADhc]

34. O'Connor A. Weighing the Evidence on Fish Oils for Heart Health. The New York Times; 2012 Apr 11. URL: https://well.blogs.nytimes.com/2012/04/11/weighing-the-evidence-on-fish-oils-for-heart-health/ [WebCite Cache ID 70Oj0X4Z2]

35. Gardner A. Fish oils heart benefits may be overstated. CNN; 2012 Sept 12. URL: https://www.cnn.com/2012/09/11/health/health-fish-oil/index.html [WebCite Cache ID 70OixPYuj]

36. Daniells S. 'Flawed' omega-3 meta-analysis harms public health: GOED. Nutraingredients-usa; 2012 Sept 11. URL: https://www.nutraingredients-usa.com/Article/2012/09/12/Flawed-omega-3-meta-analysis-harms-public-health-GOED [WebCite Cache ID 70OirubqB]

37. Beck M. Questioning the superpowers of omega-3 in diets. The Wall Street Journal; 2012 Oct 1. URL: https://www.wsj.com/articles/SB10000872396390444592404578030391175619824 [WebCite Cache ID 70OinoV1K]

38. Sydenham E, Dangour AD, Lim WS. Omega 3 fatty acid for the prevention of cognitive decline and dementia. Cochrane Database Syst Rev 2012 Jun 13(6):CD005379. [doi: 10.1002/14651858.CD005379.pub3] [Medline: 22696350]

39. Sifferlin A. Fish oil fail: omega-3s may not protect brain health after all. Time; 2012 Jun 13. URL: http://healthland.time.com/2012/06/13/fish-oil-fail-omega-3s-may-not-protect-brain-health-after-all/ [accessed 2018-06-23] [WebCite Cache ID 70OijXnX3]

40. Imhoff-Kunsch B, Stein AD, Martorell R, Parra-Cabrera S, Romieu I, Ramakrishnan U. Prenatal docosahexaenoic acid supplementation and infant morbidity: randomized controlled trial. Pediatrics 2011 Sep;128(3):e505-e512 [FREE Full text] [doi: 10.1542/peds.2010-1386] [Medline: 21807696]

41. Collins N. Omega-3 can reduce risk of colds in babies. The Telegraph; 2011 Aug 1. URL: https://www.telegraph.co.uk/news/health/news/8671525/Omega-3-can-reduce-risk-of-colds-in-babies.html [WebCite Cache ID 70Oil8rpD]

42. Makrides M, Gibson RA, McPhee AJ, Collins CT, Davis PG, Doyle LW, et al. Neurodevelopmental outcomes of preterm infants fed high-dose docosahexaenoic acid: a randomized controlled trial. J Am Med Assoc 2009 Jan 14;301(2):175-182. [doi: 10.1001/jama.2008.945] [Medline: 19141765]

43. Park A. Study: fish oil for preemies may boost cognition. Time; 2009 Jan 13. 2009 Jan 13. URL: http://content.time.com/time/health/article/0,8599,1871514,00.html [WebCite Cache ID 70OiKpFM7]

44. Marchioli R, Levantesi G, Silletta MG, Barlera S, Bernardinangeli M, Carbonieri E, GISSI-HF Investigators. Effect of n-3 polyunsaturated fatty acids and rosuvastatin in patients with heart failure: results of the GISSI-HF trial. Expert Rev Cardiovasc Ther 2009 Jul;7(7):735-748. [doi: 10.1586/erc.09.70] [Medline: 19589110]

45. Toyourhealth. 2009 Mar. Try fish oil instead of drugs URL: http://www.toyourhealth.com/mpacms/tyh/article.php?id=1157 [accessed 2018-06-23] [WebCite Cache ID 70OiIPqyk]

46. Makrides M, Gibson RA, McPhee AJ, Yelland L, Quinlivan J, Ryan P, DOMInO Investigative Team. Effect of DHA supplementation during pregnancy on maternal depression and neurodevelopment of young children: a randomized controlled trial. J Am Med Assoc 2010 Oct 20;304(15):1675-1683. [doi: 10.1001/jama.2010.1507] [Medline: 20959577]

47. Gardner A. Cnn. 2010 Oct 19. Fish oil doesn't benefit new moms, babies URL: http://www.cnn.com/2010/HEALTH/10/19/health.fish.oil.moms/ [WebCite Cache ID 70OiF14F3]

48. Quinn JF, Raman R, Thomas RG, Yurko-Mauro K, Nelson EB, Van Dyck C C, et al. Docosahexaenoic acid supplementation and cognitive decline in Alzheimer disease: a randomized trial. J Am Med Assoc 2010 Nov 3;304(17):1903-1911 [FREE Full text] [doi: 10.1001/jama.2010.1510] [Medline: 21045096]

49. Klein S. Cnn. 2010 Nov 2. Fish oil ingredient doesn't slow Alzheimer's URL: http://www.cnn.com/2010/HEALTH/11/02/health.fish.oil.alzheimers/ [WebCite Cache ID 70OhcYbak]

50. Yurko-Mauro K, McCarthy D, Rom D, Nelson EB, Ryan AS, Blackwell A, MIDAS Investigators. Beneficial effects of docosahexaenoic acid on cognition in age-related cognitive decline. Alzheimers Dement 2010 Nov;6(6):456-464. [doi: 10.1016/j.jalz.2010.01.013] [Medline: 20434961]

51. Laino C. Webmd. 2009 Jul 13. Fish oil supplements boost memory URL: https://www.webmd.com/brain/news/20090713/fish-oil-supplements-boost-memory [accessed 2018-06-23] [WebCite Cache ID 70OgsbKfU]

52. Kirby A, Woodward A, Jackson S, Wang Y, Crawford MA. A double-blind, placebo-controlled study investigating the effects of omega-3 supplementation in children aged 8-10 years from a mainstream school population. Res Dev Disabil 2010;31(3):718-730. [doi: 10.1016/j.ridd.2010.01.014] [Medline: 20171055]

53. Macrae F. Fallacy of fish oil revealed as study finds supplements DON'T boost children's brain power. The Daily Mail; 2010 Apr 29. URL: http://www.dailymail.co.uk/health/article-1269489/Fallacy-fish-oil-revealed-study-finds-boost-childrens-brain-power.html [WebCite Cache ID 70OgouTyI]

54. Andreeva VA, Galan P, Torrès M, Julia C, Hercberg S, Kesse-Guyot E. Supplementation with B vitamins or n-3 fatty acids and depressive symptoms in cardiovascular disease survivors: ancillary findings from the SUpplementation with FOLate,

vitamins B-6 and B-12 and/or OMega-3 fatty acids (SU.FOL.OM3) randomized trial. Am J Clin Nutr 2012 Jul;96(1):208-214. [doi: 10.3945/ajcn.112.035253] [Medline: 22648722]

55. Daniells S. Nutraingredients-usa. 2012 Jun 1. Omega 3 or b vitamins fail to benefit depressive symptoms, but low doses may be at fault URL: https://www.nutraingredients-usa.com/Article/2012/06/01/Omega-3-or-B-vitamins-fail-to-benefit-depressive-symptoms-but-low-doses-may-be-at-fault [WebCite Cache ID 70OhMsxBd]

56. Galet C, Gollapudi K, Stepanian S, Byrd JB, Henning SM, Grogan T, et al. Effect of a low-fat fish oil diet on proinflammatory eicosanoids and cell-cycle progression score in men undergoing radical prostatectomy. Cancer Prev Res (Phila) 2014 Jan;7(1):97-104 [FREE Full text] [doi: 10.1158/1940-6207.CAPR-13-0261] [Medline: 24169960]

57. Rattue G. Prostate / prostate cancer nutrition / diet low-fat diet with fish oil supplements slows growth rate of prostate cancer cells. Med News Today; 2011 Oct 28.

58. Lavie CJ, Milani RV, Mehra MR, Ventura HO. Omega-3 polyunsaturated fatty acids and cardiovascular diseases. J Am Coll Cardiol 2009 Aug 11;54(7):585-594 [FREE Full text] [doi: 10.1016/j.jacc.2009.02.084] [Medline: 19660687]

59. Warner J. Webmd. 2009 Aug 3. Daily omega-3s recommended for heart URL: https://www.webmd.com/heart/news/20090803/daily-omega-3s-recommended-heart [accessed 2018-06-23] [WebCite Cache ID 70OggfbZ0]

60. Jump DB, Depner CM, Tripathy S. Omega-3 fatty acid supplementation and cardiovascular disease. J Lipid Res 2012 Dec;53(12):2525-2545 [FREE Full text] [doi: 10.1194/jlr.R027904] [Medline: 22904344]

61. Daniells S. Evidence is strong for omega-3's heart health benefits: Linus Pauling review. Nutraingredients-usa; 2012 Nov 29. URL: https://www.nutraingredients-usa.com/Article/2012/11/29/Evidence-is-strong-for-omega-3-s-heart-health-benefits-Linus-Pauling-review [WebCite Cache ID 70OpBm5N8]

62. Lee K, Hoti K, Hughes JD, Emmerton L. Dr Google and the consumer: a qualitative study exploring the navigational needs and online health information-seeking behaviors of consumers with chronic health conditions. J Med Internet Res 2014 Dec 2;16(12):e262 [FREE Full text] [doi: 10.2196/jmir.3706] [Medline: 25470306]

63. Tan SS, Goonawardene N. Internet health information seeking and the patient-physician relationship: a systematic review. J Med Internet Res 2017 Jan 19;19(1):e9 [FREE Full text] [doi: 10.2196/jmir.5729] [Medline: 28104579]

64. Sbaffi L, Rowley J. Trust and credibility in web-based health information: a review and agenda for future research. J Med Internet Res 2017 Jun 19;19(6):e218 [FREE Full text] [doi: 10.2196/jmir.7579] [Medline: 28630033]

65. Lidwell W, Holden K, Butler J, Elam K. Universal Principles of Design: 125 Ways to Enhance Usability, Influence Perception, Increase Appeal, Make Better Design Decisions, and Teach Through Design. Beverly, MA: Rockport Publishers; 2010:1-272.

66. Yavchitz A, Boutron I, Bafeta A, Marroun I, Charles P, Mantz J, et al. Misrepresentation of randomized controlled trials in press releases and news coverage: a cohort study. PLoS Med 2012 Sep 11;9(9):e1001308 [FREE Full text] [doi: 10.1371/journal.pmed.1001308] [Medline: 22984354]

67. Allam A, Schulz PJ, Nakamoto K. The impact of search engine selection and sorting criteria on vaccination beliefs and attitudes: two experiments manipulating Google output. J Med Internet Res 2014 Apr 2;16(4):e100 [FREE Full text] [doi: 10.2196/jmir.2642] [Medline: 24694866]

Abbreviations

ALA: alpha-linolenic acid
NHIS: National Health Interview Survey
PUFA: polyunsaturated fatty acids
RCT: randomized controlled trial

Original Paper

Characterizing Websites That Provide Information About Complementary and Integrative Health: Systematic Search and Evaluation of Five Domains

Annie T Chen[1], PhD, MSIS; Lisa Taylor-Swanson[2], PhD; Ronald W Buie[1], BS; Albert Park[3], PhD, MS; Mike Conway[4], PhD

[1]Department of Biomedical Informatics and Medical Education, University of Washington School of Medicine, Seattle, WA, United States

[2]College of Nursing, University of Utah, Salt Lake City, UT, United States

[3]Department of Software and Information Systems, College of Computing and Informatics, University of North Carolina at Charlotte, Charlotte, NC, United States

[4]Department of Biomedical Informatics, University of Utah School of Medicine, Salt Lake City, UT, United States

Corresponding Author:
Annie T Chen, PhD, MSIS
Department of Biomedical Informatics and Medical Education
University of Washington School of Medicine
850 Republican Street, Box 358047, C238
UW Medicine South Lake Union
Seattle, WA, 98109
United States
Phone: 1 2062219218
Email: atchen@uw.edu

Abstract

Background: In recent years, there has been an increase in the utilization of complementary and integrative health (CIH) care, and an increase in information-seeking behavior focused on CIH. Thus, understanding the quality of CIH information that is available on the internet is imperative. Although there have been a limited number of studies evaluating the quality of websites providing information about specific CIH-related topics, a broad evaluation of CIH websites has not been conducted.

Objective: This study was designed to fill that gap. We set out to assess website quality in 5 CIH domains: (1) acupuncture, (2) homeopathy, (3) massage, (4) reiki, and (5) yoga. This study aimed to 1) characterize the websites by type and quality; 2) evaluate website characteristics which may affect readers' perceptions, specifically message content, structural features, and presentation style, and 3) investigate the extent to which harms, benefits and purposes of use are stated on websites.

Methods: This study employed a systematic search strategy to identify websites in each of the target domains to be evaluated. The websites were then classified by type, and a set of checklists focusing on quality, message content, structural features, and presentation style was used to evaluate the websites. Lastly, we performed content analysis to identify harms, benefits, and perceived purposes of use.

Results: There were similarities across domains regarding their overall quality and their message content. Across all domains, a high proportion of websites received strong scores in terms of ownership, currency, interactivity and navigability. Scores were more variable concerning authorship, balanced presentation of information and the use of sources of information. However, there were differences regarding their structural features and presentation style. Acupuncture and reiki sites tended to include more external links, and yoga, fewer. There was variation across domains in the extent to which the websites contained domain-specific terminology. Websites tended to provide an extensive list of potential benefits, while reporting of harms was scarce.

Conclusions: This is the first study to perform a multidimensional assessment of websites in multiple CIH domains. This review showed that while there are similarities among websites of different CIH domains, there are also differences. The diverse distribution of website types suggests that, regardless of CIH domain, the public encounters information through many different types of media, and it would be useful to consider how the presentation of this content may differ depending on the medium. The characteristics for which variability exist are areas that warrant greater attention from researchers, policy makers, clinicians and patients. There is also a need to better understand how individuals may interact with CIH websites, and to develop tools to assist people to interpret the CIH-related information that they encounter.

XSL•FO
RenderX

Introduction

Increasing numbers of people have sought CIH care in the US [1] and worldwide [2]. Complementary and integrative health (CIH) and complementary and alternative medicine (CAM) are often used to refer to nonmainstream health care practices. The term CIH is now being increasingly employed to emphasize that these complementary modalities are being used alongside, rather than in place of, conventional medicine [3]. In this article, we refer to CAM whenever the term was used in a cited information source; otherwise, we employ the term CIH.

The Centers for Disease Control reported that 38% of adults in the US used one or more CAM practices in a 12-month period [4]. CAM use has been reported to be as high as 87% among cancer patients [5]. Previous research has also reported that there is substantial use of CAM for particular conditions including lower back pain [6], gastrointestinal diseases [7], and cardiovascular disease [8]. Alongside this, an increasing number of people seek web content to inform decision-making. In 2009, 35% of internet users looked online for information about alternative treatments, up from 28% in 2002 [9]. People also access websites, forums, blogs, online communities, and social networks to discuss CIH [10].

Various factors contribute to patients' use of CIH. First, patients seek health care that is more in line with their values, beliefs, and orientations. This may include a valuation of more holistic approaches to health, an orientation towards wellness, and a desire to participate in treatment decisions [11-17]. Relying on personal experiences [16,18-20], use of social networks for information and decision making [17,19-22], and social relationships may also influence individuals' decisions to seek CIH care [23-26]. Lastly, dissatisfaction with conventional practitioners and the lack of effectiveness of conventional treatments may also influence this shift towards CIH [14,20,27-29,17].

Given the increasing utilization of CIH modalities and increase in CIH-related information seeking, understanding the quality of CIH information that is available on the internet is imperative. In recent years, there have been concerns about the quality of CIH-related information on the internet, including attention to and concern about CIH-related news coverage [30-32]. Additionally, a review of randomized controlled trials for CIH has reported that the quality of safety reporting needs improvement [33].

However, to date, there have been few studies of the quality of CIH online health information. Those that have been conducted focus on the evaluation of websites presenting CIH information relating to specific topics, such as the utilization of CIH for cancer [34,35], and ginseng, ginkgo biloba, and St John's wort[36]. For example, one study of 150 websites containing information on ginseng, ginkgo biloba, and St John's wort

reported that most CIH websites had poor technical criteria compliance (authorship, references, editorial process) [36]. Additionally, websites often contained information that could result in harm if acted upon and omitted vital information. In another study of websites selling St John's wort, websites with references to information sources were more likely to give correct indications and to list drug interactions [37]. A study of websites of National Cancer Institute-designated comprehensive cancer care centers reported that, while more sites are providing information about CIH modalities, the quality and ease of navigability of these sites was variable [34]. Existing research has also examined the portrayal of CIH information in other media including Wikipedia and news media. It has been concluded that 56% of the reviewed articles in Wikipedia needed substantial improvements in content [38]. Also, there is considerable variability in news reporting practices about CIH, and much of the information that the public receives is inaccurate and incomplete [30].

When we consider the increased public interest and consumption of CIH against this backdrop of concern regarding online CIH information, a clear need emerges to better understand the characteristics of CIH-related websites. This project was designed to fill that gap. To do so, we aimed to characterize websites based on multiple dimensions. First, we categorized websites by type and evaluated them regarding quality using a checklist based on extant literature [39,40]. Though website quality may influence users' perceptions and inclination to use CIH, users may also be influenced by their perceptions of the credibility of these websites. Previous studies have identified various factors that can influence perception of website credibility, including the site type [41], design look [42], information structure [42], information focus [42], message content [43], and structural features such as the inclusion of a navigation menu, privacy policy, and links to external sites [43]. Thus, our evaluation of websites also considered characteristics that have been associated with perceived website credibility. Lastly, given the concern regarding safety and the potential for harm, we investigated the extent to which benefits, harms, and the purposes of websites were stated, to contextualize the information that was provided. This study had the following aims:

- Characterize the websites by their type and evaluate overall quality
- Assess the message content, structural features, and presentation style of websites that may influence the viewers' perceptions apart from quality
- Examine the extent to which harms, benefits, and website purposes are stated

Given the great diversity of CIH modalities, we reviewed several domains to provide a broader characterization of the quality of online CIH information. Previous studies of CIH have employed a classification system with 5 types [5,30,34]. We selected 1

XSL•FO

RenderX

modality in each of the 5: acupuncture (whole systems), homeopathy (biologic), massage therapy (manipulative), reiki (energy medicine), and yoga (mind-body).

Methods

Overview

This study employed a systematic search strategy to identify websites in each of the target domains to be evaluated. The websites were then classified by type, and a quality checklist and other checklists of factors that may influence credibility were used to evaluate the websites. Lastly, the websites were coded to identify harms, benefits, and perceived purposes of use.

Sample

Search Strategy

A systematic search strategy was developed to identify websites with relevant content for each of the 5 domains. Three popular search engines [9] including Google, Yahoo, and Bing were used to search for the following terms "reiki," "homeopathy," "massage," "acupuncture," and "yoga." People tend to look at the first page of the search engine results and occasionally go to the second search engine results page (SERP) [44], and SERPs yield an average of 8-10 number of sites [45]. Thus, the first 20 results, excluding ads, from each search engine were viewed and saved. This yielded 300 results (ie, the product of 5 domains searched, 3 search engines, and 20 results). Two blogs were added to the search results for each domain to ensure that they were represented in the websites that we examined. To do so, the search for each domain was rerun, adding the search term "blog", and the first two appropriate entries added. Duplicates were omitted, yielding a total of 165 sites (Figure 1). Browser and cookie information was erased before conducting the searches, and sponsored links and advertisements were excluded. Websites that aggregate sources of information rather than provide information itself were excluded, including thumbtack.com and yelp.com.

Quantitative Content Analysis

Each of the research aims was based on quantitative content analysis, defined as a research method in which content is assigned to categories according to rules, and the analysis of relationships involving these categories is conducted using statistical methods [46]. Each website was viewed by at least one trained coder and evaluated for type, quality of online information, message content, structural features, and presentation style, per the criteria to be discussed in subsequent subsections. Thirty-five of the 165 websites (21.2%) were double-coded, and interrater reliability was evaluated by calculating a Cohen kappa statistic [47] within each domain and across domains. The Cohen kappa across all 5 domains was .82, indicating a high level of interrater reliability. The agreement for each of the individual domains is reported in Multimedia Appendix 1. After interrater reliability was calculated, the 2 coders discussed and resolved disagreements on the websites that were double-coded.

Each coder viewed the number of pages of a website necessary to evaluate it based on each of the evaluation criteria. Some websites may have required the coder to view every single page. For others, particularly extensive ones, it might not have been necessary for the coder to view every single page to evaluate the website for all criteria. However, to ensure consistency, both coders reviewed at least three pages for each website: "about," "home," and "contact." In the case of blogs, coders answered the scale items based on the most recent blog entries (minimum of 10).

Due to the size of some of the larger websites, it was necessary to limit the amount of time that coders spent on reviewing them. Fifteen minutes was selected in order to standardize the procedure that coders followed. Coders noted if more than 15 minutes was necessary to evaluate and code a site correctly. In these cases, a code of "not present" does not necessarily mean the topic was not present, but rather, that it was not located within 15 minutes of reading and that it was hard to find. Twenty-four (14.5%) of the websites took longer than 15 minutes for a coder to review. Our reviews of media sites included not only the particular article that came up on our search but also other articles that were topically related and were available on the same website. Reading multiple articles ensured that our coding would reflect the overall media website content for the domain topic.

Aim 1: Website Type and Quality

Website Type

We categorized websites according to a typology that has been presented by Sillence and colleagues [39]. This typology was comprised of 10 different types of health advice websites: (1) web providers and portal sites, (2) support groups, (3) charity sites, (4) government websites, (5) companies or clinicians promoting the sale of CIH services, (6) sales sites specific to CIH products, (7) personal sites such as blogs, (8) information services, (9) media sites, and (10) clinician sites. This schema was revised to make it more suitable for CIH-related subject matter, and to increase the inclusivity and clarity of the terminology used. We present the schema, along with sample websites classified into each category in the "Results" section.

Quality Assessment

To perform a general assessment of quality, we employed the Sandvik scale, a well-established quality scale originally developed for the analysis of urinary incontinence websites [42]. Since its development in the late 1990s, the scale has been applied to multiple domains including CIH [35,40,48-50]. It is comprised of 7 items (see Table 1), each rated as 0 (absent), 1 (partially present), or 2 (present).

In addition to the Sandvik instrument, we also considered many other instruments for assessing website quality including DISCERN, Journal of the American Medical Association, and Health on the Net certification. We ultimately decided that the Sandvik instrument, along with the other checklists that we describe in aim 2, were most appropriate to perform a well-rounded assessment of different factors that might influence a person's interaction with a CIH website. In particular, we decided not to employ the widely used DISCERN instrument

because many of the questions from DISCERN relate to communication of different treatment options for a given condition, and CIH-related websites are not necessarily disease-specific. Many CIH modalities focus on wellness and preventive health care practices, rather than treatment for a particular condition.

Aim 2: Message Content, Structural Features, and Presentation Style

Aside from the characteristics that were examined in the Sandvik scale, other aspects of quality can also be related to users'

credibility perceptions and potentially, their subsequent intent to act upon the information presented by websites. In this aim, we characterize websites further based on 3 categories: (1) message content (inclusion of statistics, testimonials, and quotes), (2) structural features, such as the inclusion of a navigation menu, privacy policy, and links to external sites, and (3) presentation style (Table 2).

Figure 1. Search strategy.
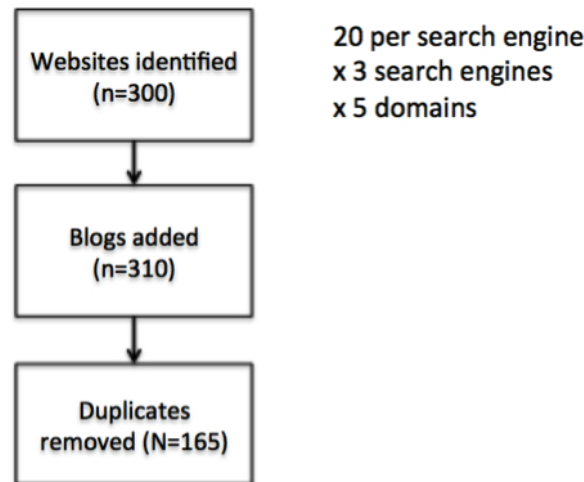
Table 1. Sandvik scale items and rating criteria.

Quality	Rating criterion
Ownership	• 2 (name and type of owner clearly stated on the "contact us," "about" or similar page) • 1 (other indications of ownership are present but hard to find, requiring several clicks through to various pages and not found on the contact/about pages) • 0 (no indication of ownership)
Authorship	• 2 (author's name and qualification clearly stated on the "contact us," "about" or similar page) • 1 (other indications of authorship are present but hard to find, requiring several clicks through to various pages and not found on the contact/about page) • 0 (no indication of authorship)
Source	• 2 (references given to scientific literature) • 1 (indications of source to nonscientific sources) • 0 (no indication of source)
Currency	• 2 (date of publication or update clearly stated on all pages) • 1 (indications of currency are not found on every page) • 0 (no indication of currency, including timestamps based on standard templates for the entire website)
Interactivity	• 2 (clear invitation to comment or ask questions by an email address or link to a form) • 1 (any other email address on the site) • 0 (no possibility for interactivity)
Navigability	• 2 (information easily found by following links from the home page) • 1 (information found only with difficulty by following links, search engine provided if information widely scattered on site) • 0 (information scattered around, no search engine)
Balance	• 2 (offers balanced information) • 1 (biased in favor of own products or services) • 0 (only promoting own products or services)

Table 2. Message content, structural features, and presentation style definitions.

Website characteristic	Definition
Message content	
Statistics	Pages with numeric information
Testimonials	Pages with stories, narratives or accounts
Quotations	Pages with a quote regarding the health topic/modality appearing in the main text of the website
Structural features	
Navigation menu	Pages with a menu or list that serve as a directory
Privacy policy	Pages with a link to the website's privacy policy
Links to external sites	Pages that include a link to external websites
Physical address	Pages with contact information (address or phone number)
Third party	Pages with an accreditation or third-party endorsement (Health on the Net code or otherwise)
Presentation style	
Prominence	Is CIH[a] information prominently displayed?
Visual appeal	Colors, images, and animation are pleasing and professional looking (high-quality website)
Advertisements	Absence of external advertisements for special offers and commercial sales
Terminology	Use of CIH language and terms specific to that domain (eg, "chi" or "Qi" for acupuncture, "like cures like" for homeopathy, "friction" for massage therapy, "Ki" for reiki, and "prana" for yoga)

[a]CIH: complementary and integrative health.

The scales for message content and structural features were based on those used in a study investigating the effect of website characteristics on perceptions of website credibility [43]. The scale for presentation style was based on items that were included in a study of the perceptions of CIH information conveyed by websites of medical institutions [51]. These 3 scales were revised to prevent duplication of items among the instruments, to minimize the need for redundant coding. The items in each of these scales were assigned 1 of 2 codes: 1 (present), and 0 (absent).

Examining Associations Between Domain and Website Attributes

We employed inferential methods for contingency tables to examine associations between specific domains, and the scores that websites were assigned for quality characteristics, message content, structural features, or presentation styles. We conducted chi-square tests of independence to examine whether there was association between domain and the presence of external links, and domain and the use of specialized terminology. In the case of domain versus source attribution, due to the presence of expected frequencies of less than 5, we employed Fisher's exact test [52]. A chi-square test of independence or Fisher's exact test on its own merely shows that there is an association between 2 nominal variables; it does not show which cells are contributing to the lack of fit of the model [53,54]. To further examine the nature of the dependence, we calculated standardized Pearson residuals. A standardized Pearson residual that exceeds 2 in absolute value indicates a lack of fit of the model in a given cell [53,54]. Lastly, we followed up with pairwise chi-square comparisons using the approach recommended in Benjamini and Hochberg to control for the false discovery rate [55], and to further contextualize differences between the domains.

Aim 3: Harms, Benefits, and Website Purposes

In previous CIH-related research, potential harms and benefits have often been a source of concern. For example, it has been reported that media coverage of CAM is inconsistent [30], and studies have argued for the need for increased awareness of harms and benefits [56]. Thus, we also investigated the extent to which harms and benefits were reported in our sample. We also noted the purposes of these websites to provide additional context regarding the nature of websites in each domain.

We reviewed literature on the 5 domains for harms and benefits. We started our coding scheme based on this literature and then added relevant codes from the process of coding the websites. Two coders independently coded the same 2 websites from each domain for harms, benefits, and purposes. They then came together to discuss and agree upon a coding scheme (Table 3). The remaining websites were then divided between the 2 coders, and each coder proceeded to code the rest of the websites that they were assigned. The code "misinformation" was applied to situations in which parts of the website text argued that misinformation caused by the modality had the potential to cause harm. For example, in one case, a yogi argued that Westerners should not be doing traditional yoga due to cultural differences in normal resting postures, such as sitting in chairs or on the ground, and that the expectation that yoga should be the same resulted in higher levels of injury among his advanced students. The websites were not mutually exclusive, meaning, that a website could indicate more than one harm, benefit, or purpose.

Table 3. Definitions of harms, benefits, and purposes.

Content dimension	Definition
Harms	
Contamination	Person could be harmed by contaminated supplements
Misinformation	Misinformation about or caused by the modality that has the potential to cause harm
Harmful if improper	Harm could occur if the patient saw an untrained provider, or if a supplement was inappropriately tried without seeing a clinician
Syncope	Mention of syncope
Bleeding	Mention of bleeding
Bruising	Mention of bruising
Other harm	Other harm not noted in this schema
No harm	No harm was mentioned
Benefits	
Psychological	Any psychological symptoms (ie, anxiety, depression)
Pain relief	The site discusses how a modality decreases pain or provides pain relief
Circulatory	The site discusses how a modality improves circulation
Neurologic	Any neurologic disease symptom (multiple sclerosis, dementia, neuralgia, paresthesia)
Gynecologic	Any gynecologic disease or symptoms (eg, dysmenorrhea, menopausal transition symptoms, hot flashes, premenstrual syndrome)
Digestive	Any digestive disease or symptom (eg, irritable bowel syndrome, Crohn's disease, diarrhea, constipation)
Endocrine	Any endocrine disease or symptom (eg, diabetes mellitus, Cushing's syndrome)
Cardiovascular	Any cardiovascular disease or symptom (eg, congestive heart failure, palpitations, tachycardia)
Posture	Any mention of benefits for posture
Multiple benefits	More than one benefit was mentioned or the website stated that there were multiple benefits
No benefit	No benefits were mentioned or the website stated that there were no benefits
Other benefit	Other benefit mentioned that does not fit the categories listed
Website purpose	
Information	The site provides information for the public about a modality
Training	The site provides training to clinicians and teachers of a modality
Sales	The site offers sales of services or products
History	The site provides historical context regarding the modality
Quality	The site provides evidence or information regarding the quality of a modality, or how to locate a quality provider, supplement or class
Professionalization	The site provides information regarding the professional status of a modality or aims to corroborate the professional nature of the modality
Promote research	The site promotes research on a modality
Present learned consensus	The site presents consensus of clinicians or teachers of a modality
Other motivations	Any other motivation not listed

Results

Sample

Based on the procedure described in the "Methods" section, 165 websites across the 5 CIH domains of acupuncture, homeopathy, massage, reiki, and yoga were selected for this review. The domains, their size, and example sites, are presented in Table 4. As can be seen, this review included websites that were specifically focused on a domain, as well as websites with a more general focus that include content relating to CIH domains.

Aim 1: Website Type and Quality

Website Type

We classified websites by type, such that websites could be considered as having more than one type (Table 5). For example, a website could simultaneously be coded as a clinician's website,

a blog, and promoting the sales of products and services. Almost half of the websites reviewed were classified into more than one type (74/165, 44.8%), and approximately one-fifth (34/165, 20.6%), into two or more types.

With respect to website categories across the 5 domains, there were no support sites and few sites that were classified as government (9/165, 5.4%) and nonprofit organizations (8/165,

4.8%) as shown in Table 6 and Figure 2. The most common government source of information was the National Center for Complementary and Integrative Health (NCCIH). Overall, sites promoting the sale of services (50/165, 30.3%) were the most common. High frequencies of clinician sites (42/165, 25.4%), media (40/165, 24.2%), information services (35/165, 21.2%), and sales of products (34/165, 20.6%) were also observed.

Table 4. Number of websites by domain, along with examples (N=165).

Domain	Example	Sites, n (%)
Acupuncture	American Academy of Medical Acupuncture	37 (22.4)
Homeopathy	Homeopathic Educational Services	33 (20.0)
Massage	"Massage Therapy Styles and Health Benefits" article on WebMD	31 (18.8)
Reiki	"What is Reiki" article on US News and World Report website	37 (22.4)
Yoga	"Yoga" page on Wikipedia	27 (16.4)

Table 5. Definitions and examples of website types.

Website type	Definition	Examples[a]
Web providers and portal sites	• Information and advice supplied by web provider rather than a physical organization • Portals act as catalogs of information providing a gateway to many other sites providing information and advice	• Medical News Today • Acupuncture.com: Gateway to Chinese Medicine, Health and Wellness
Support groups	• Often run by individuals or on behalf of support groups • May be local, national or global in scale • Often contains forums where consumers can read comments and contribute to discussions	• None
Nonprofit organizations	• An organization that is aimed at providing a service that is free and not for profit	• National Center for Homeopathy • Society of Homeopaths
Government websites	• Provide patient information in the form of news, features and fact sheets	• National Center for Complementary and Integrative Health • Medline Plus
Companies or clinicians promoting sales of CIH[b] services	• A corporate, institutional, or clinician website that promotes the sale of its own services	• University of Maryland Medical Center • Simply Hot Yoga Wellness Center
Sales sites that include the sale of CIH products	• Sales sites promote and sell drugs, medical devices or health plans often in addition to providing information	• mindbodygreen • 1-800 Homeopathy
Personal sites (blogs)	• Contains personal experiences of illnesses and health issues	• Quackwatch • Homeopathic Medicine Blog
Information services	• Websites that provide articles on health and illness issues	• Science-Based Medicine • WebMD
Media sites	• Extensions of print or television media sites that provide the latest news and commentary on health features	• Vanity Fair • Reader's Digest Best Health
Clinician sites	• Information on specific health issues or specialist clinics run by medical professionals	• To protect the privacy of the practitioners, examples are not provided.

[a]Examples reflect the parent website of a search result. For example, the page of the University of Maryland Medical Center that appeared in the page was about acupuncture.

[b]CIH: complementary and integrative health.

Table 6. Frequencies of website types by domain (N=165).

Website type	Domain				
	Acupuncture (n=37), n (%)	Homeopathy (n=33), n (%)	Massage (n=31), n (%)	Reiki (n=37), n (%)	Yoga (n=27), n (%)
Web providers or portal sites	8 (21.6)	4 (12.1)	7 (22.6)	3 (8.1)	1 (3.7)
Support groups	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Nonprofit organizations	1 (2.7)	4 (12.1)	1 (3.2)	1 (2.7)	1 (3.7)
Government websites	4 (10.8)	1 (3.0)	2 (6.5)	2 (5.4)	0 (0)
Sale of CIH[a] services	10 (27.0)	4 (12.1)	10 (32.3)	15 (40.5)	11 (40.7)
Sale of CIH products	8 (21.6)	6 (18.2)	4 (12.9)	10 (27.0)	6 (22.2)
Personal sites[b]	6 (16.2)	6 (18.2)	5 (16.1)	10 (27.0)	5 (18.5)
Information services	16 (43.2)	7 (21.2)	7 (22.6)	5 (13.5)	0 (0)
Media sites	8 (21.6)	9 (27.3)	5 (16.1)	6 (16.2)	12 (44.4)
Clinician sites	10 (27.0)	5 (15.2)	6 (19.4)	14 (37.8)	7 (25.9)

[a]CIH: complementary and integrative health.

[b]Totals include the additional blogs that were added to each domain as specified in the "Methods" section.

Figure 2. Modalities by site type. ACU: acupuncture, HOM: homeopathy, MAS: massage, REI: reiki, YOG: yoga.
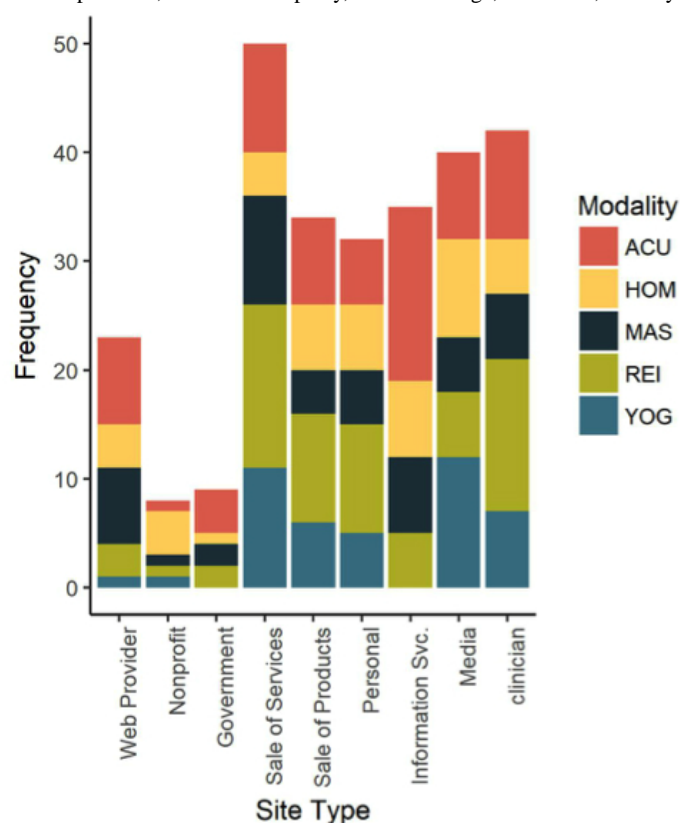
Quality Assessment

We first considered the websites based on the Sandvik scale. Across all domains, a high proportion of websites received a score of 2 on ownership, currency, interactivity, and navigability (Figure 3). Scores were more variable with respect to authorship, balanced presentation of information, and the use of sources of information. Acupuncture demonstrated a unique pattern in terms of balance of information, with a tendency for ratings to be at the two extremes: 0 (only promoting own products or services), and 2 (balanced information). There was also a greater tendency to include source attribution than would have been expected, had all domains been equal (P<.001; see Multimedia Appendix 2 for standardized Pearson residuals and Multimedia Appendix 3 for pairwise comparisons of source attribution).

Aim 2: Message Content, Structural Features, and Presentation Style

We also examined message content, structural features and presentation style of websites. With regard to message content, there was considerable consistency across domains. Websites were divided on the inclusion of statistics, testimonials and quotations (Figure 4). A notable exception was yoga websites' paucity of statistics.

With regard to structural features, nearly all websites had navigation menus, and most had privacy policies (Figure 5).

The presence of links to external websites differed across domains (N=165, X^2_4=16.4, P=.002) with acupuncture sites tending to include more external links, and yoga, fewer, based on standardized Pearson residuals greater than an absolute value of 2 (see Multimedia Appendix 2 for standardized Pearson residuals, and Multimedia Appendix 3 for pairwise chi-square tests of presences of external links). The presence of physical locations was more common than not among homeopathy and massage websites, and more evenly split in the case of the other domains. In general, websites did not have a third-party accreditation.

Figure 3. Sandvik quality indicators. ACU: acupuncture, HOM: homeopathy, MAS: massage, REI: reiki, YOG: yoga.
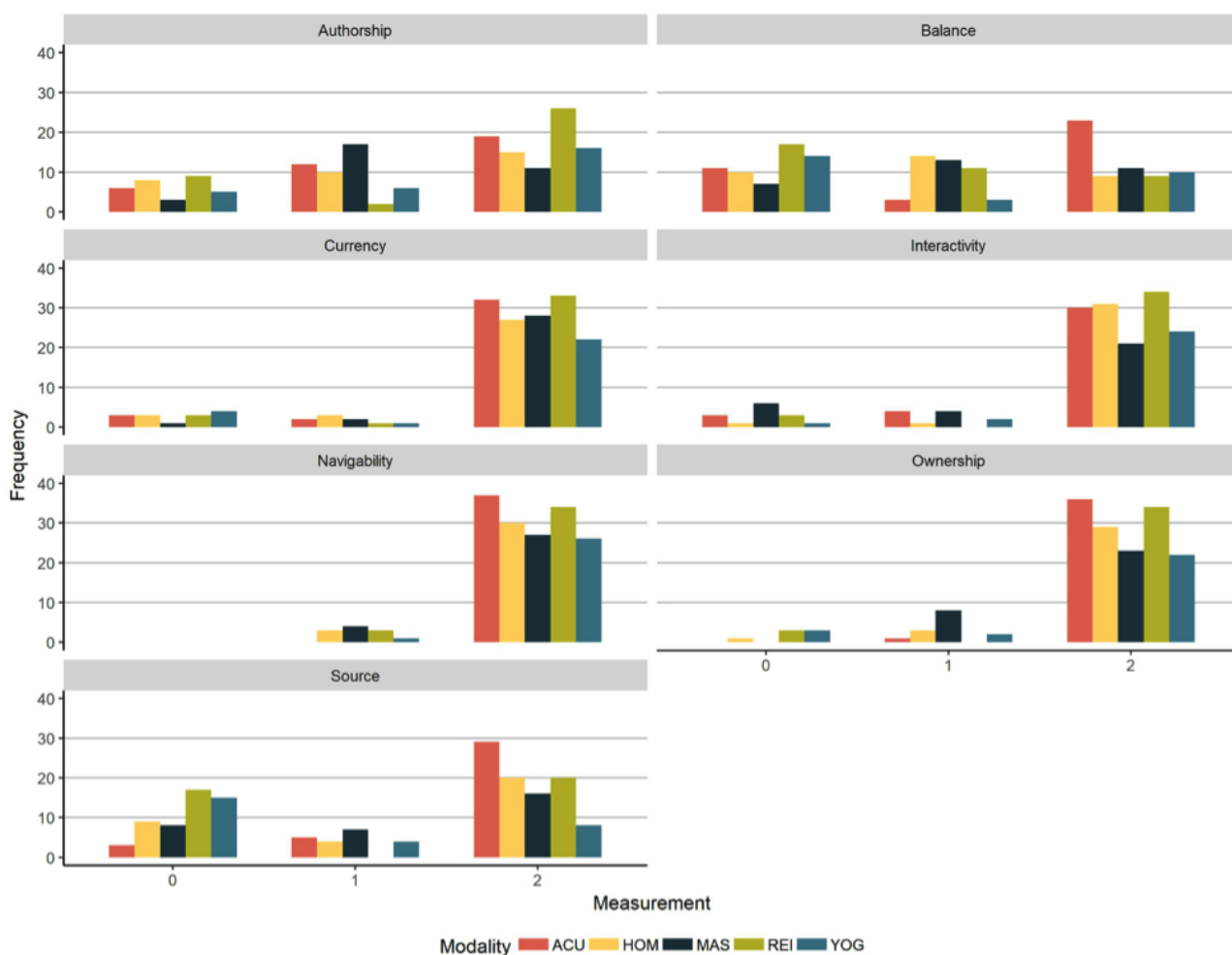

Figure 4. Message characteristics. ACU: acupuncture, HOM: homeopathy, MAS: massage, REI: reiki, YOG: yoga.
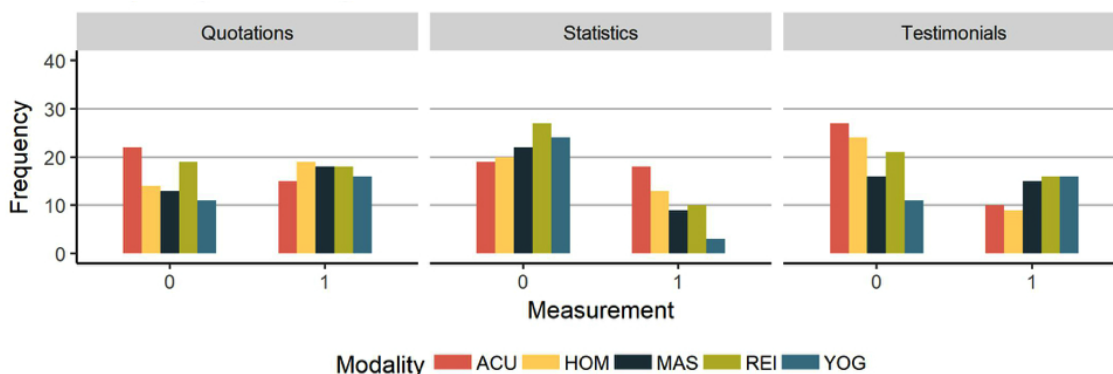
Figure 5. Structural characteristics. ACU: acupuncture, HOM: homeopathy, MAS: massage, REI: reiki, YOG: Yoga.
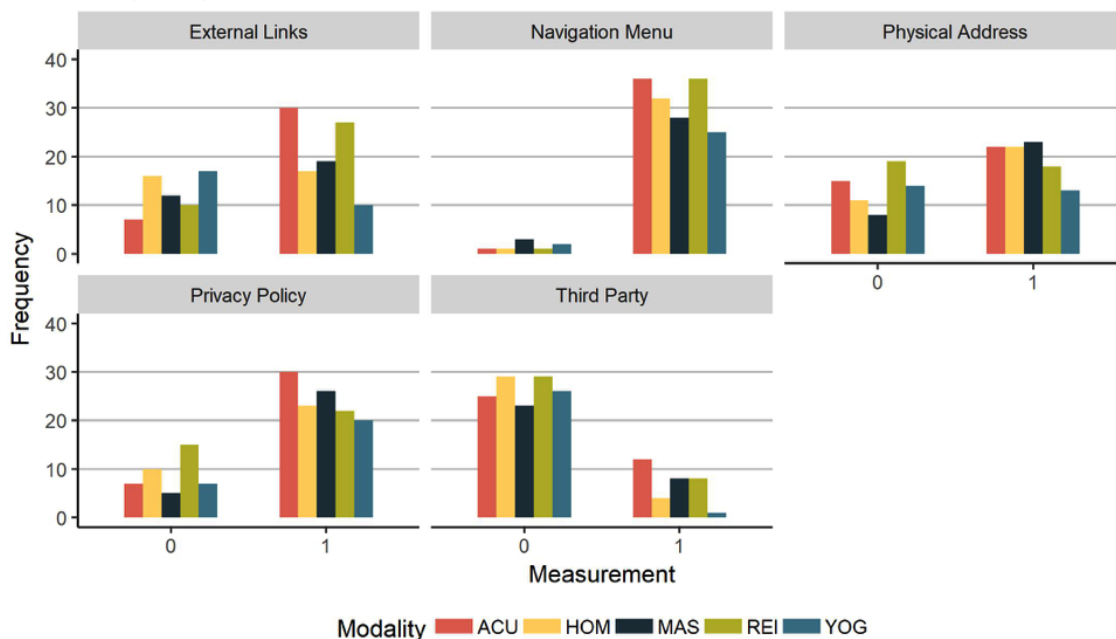
Figure 6. Presentation style. ACU: acupuncture, HOM: homeopathy, MAS: massage, REI: reiki, YOG: yoga.
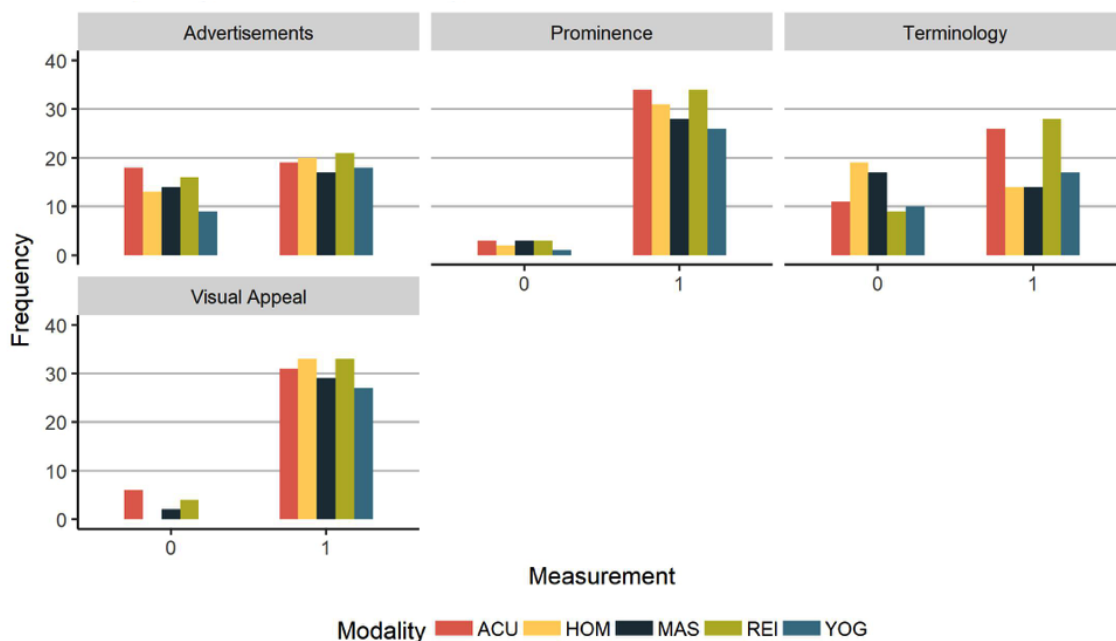
We now consider presentation style of the websites. Almost all had CIH information prominently displayed and presented information in a visually appealing fashion (Figure 6). They were generally mixed in terms of whether they contained advertisements. There was variation across domains in the extent to which the websites contained domain-specific terminology (N=165, X^2_4=12.6, P=.01). Reiki was more likely to have been rated as having domain-specific terminology, homeopathy was less likely, based on standardized Pearson residuals of greater than an absolute value of 2 (see Multimedia Appendix 2 for standardized Pearson residuals, and Multimedia Appendix 3 for pairwise chi-square tests of presences of domain-specific terminologies).

Aim 3: Harms, Benefits, and Website Purpose

We examined the extent to which websites reported harms and benefits (Table 7). The number of websites not providing information about harms was high (119/165, 72.1%), as was the number of websites stating that a given modality offered multiple benefits (90/165, 54.5%). Pain reduction and psychological benefits were also frequently mentioned, in 69 (41.8%) and 51 (30.9%) of websites, respectively. While almost all websites that were coded "no benefit" did not mention any benefits, there were a few cases, such as Wikipedia and Quackwatch, in which the websites stated that there were no benefits. The benefits have been described as categories, as opposed to specific harms, to be concise. Examples of other

harms that were noted include pneumothorax, broken needle, organ puncture, pain, drowsiness, dizziness, contamination, soreness, flu-like symptoms, blood sugar irregularities, and muscle strain. Examples of other benefits included benefits for sleep, fatigue, energy, relaxation, stress, immune function, anti-aging, wellness, and more.

With regard to purposes, providing information (126/165, 76.4%) and sales (61/165, 37.0%) were the most common (Table 8). Sites that were identified as providing information to support the professional nature of the modality included (1) the British Acupuncture Council [57], (2) Acupuncture Board, an autonomous body under the umbrella of the Department of Consumer Affairs, which licenses and regulates acupuncturists in California [58], (3) the British Homeopathic Association [59], and (4) the International Association of Reiki Professionals [60].

Table 7. Prevalence of harms and benefits (N=165).

Category	n (%)
Harms	
No harm	119 (72.1)
Misinformation	12 (7.3)
Improper	8 (4.8)
Contamination	4 (2.4)
Bleeding	3 (1.8)
Bruising	3 (1.8)
Syncope	1 (0.6)
Other	26 (15.8)
Benefits	
Multiple benefits	90 (54.5)
Pain reduction	69 (41.8)
Psychological benefit	51 (30.9)
No benefit	48 (29.1)
Digestive	29 (17.6)
Neurologic	25 (15.2)
Endocrine	20 (12.1)
Gynecologic	12 (7.3)
Cardiovascular	11 (6.7)
Circulatory	10 (6.1)
Posture	9 (5.5)
Other	51 (30.9)

Table 8. Prevalence of website purposes (N=165).

Website purpose	n (%)
Providing information	126 (76.4)
Sales	61 (37.0)
Training	12 (7.3)
Professionalization	11 (6.7)
Quality	8 (4.8)
History	7 (4.2)
Promotion of research	6 (3.6)
Present learned consensus	1 (0.6)
Other	18 (10.9)

Many of these sites were also noted as providing information about the history of the modality and promoting research on the modality. Though it was observed earlier that there were few government websites, sites that provide information about the profession may serve as regulatory bodies and provide varying degrees of quality control; thus, these sites may complement the role served by government websites such as that of the NCCIH.

Discussion

Principal Findings

We employed a systematic search strategy to identify top search results across 5 CIH domains and then performed quantitative content analysis to characterize the nature, quality, and assertions made by these websites. The most common types of websites were those selling services, and clinician, media, information services, and product websites. Overall, the diverse distribution of website types suggests that, regardless of CIH domain, the public encounters information through many different types of media, and it is important to consider how the presentation of this content may differ depending on the medium.

Regarding the overall quality, we observed similar patterns across domains on specific dimensions but not others. Websites received high scores on ownership, currency, interactivity, and navigability, but were inconsistent with respect to authorship, balanced presentation of information, and source attribution. These results suggest that guidelines and other efforts to improve information quality relating to CIH should place greater emphasis on recommendations for these dimensions. Concerning message content, we again observed similar patterns across domains in that websites were split regarding their inclusion of statistics, testimonials, and quotations.

The findings for structural features and presentation style appeared to differ across domains. Websites generally had a navigation menu and privacy policy. They also included information on CIH topics in a prominent way and were visually appealing. The listing of a physical address was more common among homeopathy and massage websites. Acupuncture and reiki sites tended to include more external links, and yoga, fewer. Also, there was variation across domains in the extent to which the websites contained domain-specific terminology. A substantive portion of websites reported either multiple benefits or no benefits, and reporting of harms was scarce. Though information provision was an important purpose for most of the websites that we reviewed (126/165, 76.4%), sales were also an important function (61/165, 37.0%).

The characteristics for which variability exist are areas that warrant greater attention from researchers, policy makers, clinicians, and patients. On their own, these characteristics may not necessarily be positive or negative. For example, testimonials and quotations could potentially provide misleading information, but they could also help readers to better understand the nature of the services being offered and particularly for those with more experience, to distinguish between different entities offering the same service. Similarly, references to external

material and domain-specific terminology can provide helpful information for individuals that are new to a modality, but they can also confuse readers if the explanatory text is insufficient. Thus, this review on its own is not intended to provide a clear set of guidelines to be followed, but rather, identify dimensions for which there is a need to develop guidelines and patient education materials, and to assist individuals in the consumption and evaluation of CIH content that is publicly available on the internet.

Comparison With Prior Work

This study adds to the extant literature concerning CIH in multiple ways. First, the study provides a characterization of the types, quality, and other characteristics of websites providing CIH-related information that may influence how health consumers interact with these websites. Structural features and presentation style can be associated with perceptions of website credibility [43,61], and design can play an important role in credibility. Thus, the study provides a basic understanding of these websites and differences across CIH domains.

However, the study does not answer the question of how people might react to online CIH information. In this study, we found that websites providing CIH-related information often did not report harms and reported multiple benefits. Additionally, the Sandvik scores showed that the degree to which the information was balanced varied in the domains reviewed. However, individuals might be influenced regardless of the degree to which websites present balanced information, and even when presented with the same information, people may make different decisions.

There is a need to understand the variability in individuals' decision-making processes concerning CIH, especially differences in how people may respond to online information about CIH. Previous research has reported that beliefs and attitudes are associated with CAM use [12,14], and some theoretical models have been used to conceptualize the process of choosing to utilize CAM [62]. However, none of these models have emphasized the role or presentation of information. Thus, more research is needed to understand how people interact with CIH-related websites and information, and how this influences their decisions to use CIH and their selection of modalities.

We might also consider the potential implications of the Elaboration Likelihood Model of Persuasion, which argues that individuals respond to source and message characteristics in different ways depending on the extent to which they focus on topics [63]. Prior research has also reported that individuals' evaluation of, and interactions with, information can change over the course of a chronic condition [64,65]. Given that individuals might respond differently based on a combination of personal, situational and source factors, there is a need for additional research to understand how individuals may interact with the CIH information that they encounter online, and in turn, assist them to make informed decisions. This may be particularly important given that extant literature has reported that patients often do not report CIH use to their health care providers [17,66,67].

It may also be of interest to compare the results of this study to prior research on the quality of websites relating to particular conditions. Two conditions for which CIH use is common are chronic pain and inflammatory bowel disease [7,68]. Websites providing health information for these conditions have been shown to be variable in quality and possess shortcomings in terms of source attribution, links to additional information, and balanced reporting [69,70]. Thus, these issues are not necessarily unique to CIH. However, in a subject matter area where much remains to be understood, continual work to improve the quality of online information and efforts to educate patients to make informed decisions is vital.

Lastly, it is interesting to consider the full range of benefits that were mentioned in the online CIH-related content included in this review. Aside from pain reduction and psychological benefits, a wide variety of benefits including sleep, fatigue, energy, stress, immune function, and wellness were also mentioned. These findings are consistent with previous review literature. Extant studies have reported significant associations between CAM use and health factors including arthritis, anxiety or depression, cancer, diabetes, chronic conditions, psychological health, as well as self-rated general health [71]. Secondary analysis of National Health Interview Survey data has also illustrated the importance of wellness and of CAM as part of a self-management style, to CAM users [15].

Limitations

This study had several limitations. First, our study was limited in the number of websites that was reviewed. Our search strategy and website inclusion criteria were similar to those that have been used in previous research on online health information quality; however, when an individual is searching for CIH information, it is likely that they will perform multiple searches, and thus be exposed to websites that were not included in our review. Second, the rankings provided by search engines are generated by "commercial" algorithms and are not necessarily consistent over time or place [72]. Third, in this study, we employed checklists for which there was a limited set of choices (0, 1, 2 or 0, 1). It is possible that if we had employed a more granular set of scales, then we would have seen more variability among websites.

Last, in this study, we selected 5 types of CIH to provide a richer and more comprehensive assessment of the diversity of online health information quality in CIH. Though we endeavored to select a diverse set of modalities, there is considerable diversity both within and across CIH domains, and thus our conclusions are limited insofar as the domains may be representative. CIH modalities vary in terms of many characteristics, including their degree of acceptance by the public, the ways in which modalities are utilized (eg, through the services of a provider, class or product), the degree to which provider associations exist and provide support to practitioners, and the extent to which evidence-based information is available. Considering the quality of online health information in the context of these characteristics could potentially be of considerable interest, but was beyond the scope of this study.

Conclusion

In this study, we selected multiple domains of CIH and characterized them through multiple dimensions (1) type and quality, (2) message, structural, and presentation characteristics; and (3) harms, benefits and website purposes. To our knowledge, this is the first study to perform a multidimensional assessment of websites in multiple CIH domains. In considering CIH, it is important to realize that the term does not represent a single tradition. There are differences in modalities resulting from a myriad of factors including the history, conceptual foundations, level of recognition, service delivery models, and perceptions in the media. This review showed that while there are similarities among websites of different CIH domains, there are also differences. Professional associations and regulatory bodies for these different domains might use this study to develop guidelines that they could provide to practitioners. There are also characteristics on which websites tended to be split in terms of presence and prevalence, such as the provision of quotations, testimonials, external links, and terminology. While the presence or absence of such characteristics may not necessarily be a positive or a negative characteristic, the findings of this study could be helpful for practitioners to consider how to present information to their clients and patients, as well as to consumers, to assist them in evaluation of the content that they encounter. As such, this article might be used by a diverse audience for a variety of purposes (eg, by professional associations interested in developing guidelines for website development), CIH practitioners who are interested to learn about others in similar and different clinical care modalities, practitioners of allopathic medicine, and patients. A general understanding of some of the variability of website characteristics could help both health care providers and patients consider the websites that they encounter with greater discernment. Lastly, this article can assist researchers who are interested in understanding the content and manner of presentation of information that lay health consumers are exposed to about CIH, as well as to identify what we do not yet understand about the presentation of this information.

XSL•FO

RenderX

Authors' Contributions

ATC conceptualized the study and wrote the bulk of the text, LTS and RWB performed the quantitative content analysis. RWB, ATC, and AP performed the statistical analysis. LTS contributed to parts of the text. All authors provided suggestions and feedback on the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Interrater reliability.

[PDF File (Adobe PDF File), 21KB - ijmr_v7i2e14_app1.pdf]

Multimedia Appendix 2

Chi-square tests of independence for website characteristics.

[PDF File (Adobe PDF File), 61KB - ijmr_v7i2e14_app2.pdf]

Multimedia Appendix 3

Pairwise comparisons of quality assessments.

[PDF File (Adobe PDF File), 47KB - ijmr_v7i2e14_app3.pdf]

References

1. Barnes PM, Bloom B, Nahin RL. Complementary and alternative medicine use among adults and children: United States, 2007. Natl Health Stat Report 2008 Dec 10(12):1-23 [FREE Full text] [Medline: 19361005]
2. Harris PE, Cooper KL, Relton C, Thomas KJ. Prevalence of complementary and alternative medicine (CAM) use by the general population: a systematic review and update. Int J Clin Pract 2012 Oct;66(10):924-939. [doi: 10.1111/j.1742-1241.2012.02945.x] [Medline: 22994327]
3. Complementary, Alternative, or Integrative Health: What's In a Name?. 2016 Jun. NCCIH URL: https://nccih.nih.gov/health/integrative-health [accessed 2018-09-05] [WebCite Cache ID 72D9wQJi1]
4. Clarke TC, Black LI, Stussman BJ, Barnes PM, Nahin RL. Trends in the use of complementary health approaches among adults: United States, 2002-2012. Natl Health Stat Report 2015 Feb 10(79):1-16 [FREE Full text] [Medline: 25671660]
5. Judson PL, Abdallah R, Xiong Y, Ebbert J, Lancaster JM. Complementary and Alternative Medicine Use in Individuals Presenting for Care at a Comprehensive Cancer Center. Integr Cancer Ther 2017 Mar;16(1):96-103. [doi: 10.1177/1534735416660384] [Medline: 27444312]
6. Ghildayal N, Johnson PJ, Evans RL, Kreitzer MJ. Complementary and Alternative Medicine Use in the US Adult Low Back Pain Population. Glob Adv Health Med 2016 Jan;5(1):69-78 [FREE Full text] [doi: 10.7453/gahmj.2015.104] [Medline: 26937316]
7. Hung A, Kang N, Bollom A, Wolf JL, Lembo A. Complementary and Alternative Medicine Use Is Prevalent Among Patients with Gastrointestinal Diseases. Dig Dis Sci 2015 Jul;60(7):1883-1888. [doi: 10.1007/s10620-014-3498-3] [Medline: 25556585]
8. Grant SJ, Bin YS, Kiat H, Chang DH. The use of complementary and alternative medicine by people with cardiovascular disease: a systematic review. BMC Public Health 2012 Apr 26;12:299 [FREE Full text] [doi: 10.1186/1471-2458-12-299] [Medline: 22536991]
9. Fox S, Jones S. The Social Life of Health Information. 2009 Jun 11. URL: http://www.pewinternet.org/2009/06/11/the-social-life-of-health-information/ [accessed 2018-09-05] [WebCite Cache ID 72DAkZGur]
10. Mazzocut M, Truccolo I, Antonini M, Rinaldi F, Omero P, Ferrarin E, et al. Web Conversations About Complementary and Alternative Medicines and Cancer: Content and Sentiment Analysis. J Med Internet Res 2016 Dec 16;18(6):e120 [FREE Full text] [doi: 10.2196/jmir.5521] [Medline: 27311444]
11. Astin JA. Why patients use alternative medicine: results of a national study. JAMA 1998 May 20;279(19):1548-1553. [Medline: 9605899]
12. Bishop FL, Yardley L, Lewith GT. A systematic review of beliefs involved in the use of complementary and alternative medicine. J Health Psychol 2007 Nov;12(6):851-867. [doi: 10.1177/1359105307082447] [Medline: 17956965]
13. O'Callaghan F, Jordan N. Postmodern values, attitudes and the use of complementary medicine. Complementary Therapies in Medicine 2003 Mar;11(1):28-32. [doi: 10.1016/S0965-2299(02)00109-7]
14. McFadden KL, Hernández TD, Ito TA. Attitudes toward complementary and alternative medicine influence its use. Explore (NY) 2010;6(6):380-388 [FREE Full text] [doi: 10.1016/j.explore.2010.08.004] [Medline: 21040887]

XSL·FO
RenderX

15. Upchurch DM, Rainisch BW. The importance of wellness among users of complementary and alternative medicine: findings from the 2007 National Health Interview Survey. BMC Complement Altern Med 2015 Oct 15;15:362 [FREE Full text] [doi: 10.1186/s12906-015-0886-y] [Medline: 26467652]

16. Salamonsen A. Use of complementary and alternative medicine in patients with cancer or multiple sclerosis: possible public health implications. Eur J Public Health 2016 Apr;26(2):225-229. [doi: 10.1093/eurpub/ckv184] [Medline: 26450914]

17. Bahall M. Prevalence, patterns, and perceived value of complementary and alternative medicine among cancer patients: a cross-sectional, descriptive study. BMC Complement Altern Med 2017 Jun 30;17(1):345 [FREE Full text] [doi: 10.1186/s12906-017-1853-6] [Medline: 28666435]

18. van der Schee E, Groenewegen PP. Determinants of public trust in complementary and alternative medicine. BMC Public Health 2010 Mar 12;10:128 [FREE Full text] [doi: 10.1186/1471-2458-10-128] [Medline: 20226015]

19. Frawley J, Adams J, Broom A, Steel A, Gallois C, Sibbritt D. Majority of women are influenced by nonprofessional information sources when deciding to consult a complementary and alternative medicine practitioner during pregnancy. J Altern Complement Med 2014 Jul;20(7):571-577. [doi: 10.1089/acm.2014.0028] [Medline: 24918451]

20. Khan MU, Jamshed SQ, Ahmad A, Bidin MA, Siddiqui MJ, Al-Shami AK. Use of Complementary and Alternative Medicine among Osteoarthritic Patients: A Review. J Clin Diagn Res 2016 Feb;10(2):JE01-JE06 [FREE Full text] [doi: 10.7860/JCDR/2016/15211.7169] [Medline: 27042482]

21. Ohlén J, Balneaves LG, Bottorff JL, Brazier ASA. The influence of significant others in complementary and alternative medicine decisions by cancer patients. Soc Sci Med 2006 Sep;63(6):1625-1636. [doi: 10.1016/j.socscimed.2006.03.050] [Medline: 16725245]

22. Goldman AW, Cornwell B. Social network bridging potential and the use of complementary and alternative medicine in later life. Soc Sci Med 2015 Sep;140:69-80 [FREE Full text] [doi: 10.1016/j.socscimed.2015.07.003] [Medline: 26207353]

23. Derose KP, Varda DM. Social capital and health care access: a systematic review. Med Care Res Rev 2009 Jun;66(3):272-306. [FREE Full text] [doi: 10.1177/1077558708330428] [Medline: 19174538]

24. Laporte A, Nauenberg E, Shen L. Aging, social capital, and health care utilization in Canada. Health Econ Policy Law 2008 Oct;3(Pt 4):393-411. [doi: 10.1017/S1744133108004568] [Medline: 18793479]

25. Umberson D, Montez JK. Social relationships and health: a flashpoint for health policy. J Health Soc Behav 2010;51 Suppl:S54-S66 [FREE Full text] [doi: 10.1177/0022146510383501] [Medline: 20943583]

26. Murthy V, Adams J, Broom A, Kirby E, Refshauge KM, Sibbritt D. The influence of communication and information sources upon decision-making around complementary and alternative medicine use for back pain among Australian women aged 60-65 years. Health Soc Care Community 2017 Jan;25(1):114-122. [doi: 10.1111/hsc.12275] [Medline: 26443878]

27. Palinkas LA, Kabongo ML, San Diego Unified Practice Research in Family Medicine Network. The use of complementary and alternative medicine by primary care patients. A SURF*NET study. J Fam Pract 2000 Dec;49(12):1121-1130. [Medline: 11132062]

28. Vincent C, Furnham A. Why do patients turn to complementary medicine? An empirical study. Br J Clin Psychol 1996 Feb;35 (Pt 1):37-48. [Medline: 8673034]

29. Olsen SA. A review of complementary and alternative medicine (CAM) by people with multiple sclerosis. Occup Ther Int 2009 Mar;16(1):57-70. [doi: 10.1002/oti.266] [Medline: 19222053]

30. Bonevski B, Wilson A, Henry D. An analysis of news media coverage of complementary and alternative medicine. PLoS One 2008 Jun 11;3(6):e2406 [FREE Full text] [doi: 10.1371/journal.pone.0002406] [Medline: 18545688]

31. Weeks LC, Strudsholm T. A scoping review of research on complementary and alternative medicine (CAM) and the mass media: looking back, moving forward. BMC Complement Altern Med 2008 Jul 19;8:43 [FREE Full text] [doi: 10.1186/1472-6882-8-43] [Medline: 18638413]

32. Price A, Grann VR. Portrayal of complementary and alternative medicine for cancer by top online news sites. J Altern Complement Med 2012 May;18(5):487-493. [doi: 10.1089/acm.2011.0110] [Medline: 22551076]

33. Turner L, Singh K, Garritty C, Tsertsvadze A, Manheimer E, Wieland LS, et al. An evaluation of the completeness of safety reporting in reports of complementary and alternative medicine trials. BMC Complement Altern Med 2011 Aug 22;11:67 [FREE Full text] [doi: 10.1186/1472-6882-11-67] [Medline: 21859470]

34. Brauer JA, El SA, Metz JM, Mao JJ. Complementary and alternative medicine and supportive care at leading cancer centers: a systematic analysis of websites. J Altern Complement Med 2010 Feb;16(2):183-186 [FREE Full text] [doi: 10.1089/acm.2009.0354] [Medline: 20180691]

35. Schmidt K, Ernst E. Assessing websites on complementary and alternative medicine for cancer. Ann Oncol 2004 May;15(5):733-742. [Medline: 15111340]

36. Walji M, Sagaram S, Sagaram D, Meric-Bernstam F, Johnson C, Mirza N, et al. Efficacy of quality criteria to identify potentially harmful information: a cross-sectional survey of complementary and alternative medicine web sites. J Med Internet Res 2004 Jun 29;6(2):e21 [FREE Full text] [doi: 10.2196/jmir.6.2.e21] [Medline: 15249270]

37. Martin-Facklam M, Kostrzewa M, Schubert F, Gasse C, Haefeli WE. Quality markers of drug information on the Internet: an evaluation of sites about St. John's wort. Am J Med 2002 Dec 15;113(9):740-745. [Medline: 12517364]

38. Koo M. Complementary and alternative medicine on wikipedia: opportunities for improvement. Evid Based Complement Alternat Med 2014;2014:105186 [FREE Full text] [doi: 10.1155/2014/105186] [Medline: 24864148]

XSL•FO
RenderX

39.   Sillence E, Briggs P, Harris P, Fishwick L. A framework for understanding trust factors in web-based health advice. International Journal of Human-Computer Studies 2006 Aug;64(8):697-713. [doi: 10.1016/j.ijhcs.2006.02.007]

40.   Sandvik H. Health information and interaction on the internet: a survey of female urinary incontinence. BMJ 1999 Jul 3;319(7201):29-32 [FREE Full text] [Medline: 10390457]

41.   Flanagin AJ, Metzger MJ. The role of site features, user attributes, and information verification behaviors on the perceived credibility of web-based information. New Media & Society 2007 Apr;9(2):319-342. [doi: 10.1177/1461444807075015]

42.   Fogg BJ, Soohoo C, Danielson DR, Marable L, Stanford J, Tauber ER. How do users evaluate the credibility of Web sites?: a study with over 2,500 participants. : ACM; 2003 Presented at: Proceedings of the 2003 Conference on Designing for User Experiences (DUX '03); June 06-07, 2003; San Francisco, CA URL: http://dl.acm.org/citation.cfm?id=997097 [doi: 10.1145/997078.997097]

43.   Rains SA, Karmikel CD. Health information-seeking and perceptions of website credibility: Examining Web-use orientation, message characteristics, and structural features of websites. Computers in Human Behavior 2009 Mar;25(2):544-553. [doi: 10.1016/j.chb.2008.11.005]

44.   Jansen BJ, Spink A. An analysis of documents viewing patterns of Web search engine users. Web Mining: Applications and Techniques 2004:339-354.

45.   Kelly D, Azzopardi L. How many results per page?: A Study of SERP Size, Search Behavior and User Experience. : ACM Press; 2015 Presented at: Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '15); August 09-13, 2015; Santiago, Chile p. 183-192. [doi: 10.1145/2766462.2767732]

46.   Riffe D, Lacy S, Fico F. Analyzing Media Messages: Using Quantitative Content Analysis in Research. New York: Routledge; 2014.

47.   Cohen J. A Coefficient of Agreement for Nominal Scales. Educational and Psychological Measurement 1960 Apr 01;20(1):37-46. [doi: 10.1177/001316446002000104]

48.   Breckons M, Jones R, Morris J, Richardson J. What do evaluation instruments tell us about the quality of complementary medicine information on the internet? J Med Internet Res 2008 Jan 22;10(1):e3 [FREE Full text] [doi: 10.2196/jmir.961] [Medline: 18244894]

49.   Pérez-López FR, Pérez RGR. Assessing the content and quality of information on the treatment of postmenopausal osteoporosis on the World Wide Web. Gynecol Endocrinol 2006 Dec;22(12):669-675. [doi: 10.1080/09513590601012603] [Medline: 17162708]

50.   Pérez-López FR. An evaluation of the contents and quality of menopause information on the World Wide Web. Maturitas 2004 Dec 10;49(4):276-282. [doi: 10.1016/j.maturitas.2004.07.006] [Medline: 15531123]

51.   Keshet Y. Network gatekeeping: complementary medicine information on the websites of medical institutions. Health (London) 2012 Mar;16(2):151-168. [doi: 10.1177/1363459311403950] [Medline: 21540253]

52.   McDonald JH. Handbook of biological statistics. Baltimore, MD: Sparky House Publishing; 2014.

53.   Agresti A. Categorical Data Analysis. Hoboken, New Jersey: John Wiley & Sons, Incorporated; 2002.

54.   Sharpe D. Your chi-square test is statistically significant: Now what? Pract Assess Res Eval 2015 Apr;20(8):1-10.

55.   Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B Methodol 1995;57(1):289-300.

56.   Chang KH, Brodie R, Choong MA, Sweeney KJ, Kerin MJ. Complementary and alternative medicine use in oncology: a questionnaire survey of patients and health care professionals. BMC Cancer 2011 May 24;11:196 [FREE Full text] [doi: 10.1186/1471-2407-11-196] [Medline: 21609461]

57.   British Acupuncture Council. URL: https://www.acupuncture.org.uk/ [accessed 2018-09-06] [WebCite Cache ID 72Dtsxy0d]

58.   State of California Acupuncture Board. URL: https://www.acupuncture.ca.gov/ [accessed 2018-09-06] [WebCite Cache ID 72Dtz6KNr]

59.   British Homeopathic Association. URL: https://www.britishhomeopathic.org/ [accessed 2018-09-06] [WebCite Cache ID 72Du2Zv4q]

60.   International Association of Reiki Professionals. URL: https://iarp.org/ [accessed 2018-09-06] [WebCite Cache ID 72DuUbmmJ]

61.   Cugelman B, Thelwall M, Dawes P. Website Credibility, Active Trust and Behavioural Intent. In: Oinas-Kukkonen H, Hasle P, Harjumaa M, Segerståhl K, Øhrstrøm P, editors. Persuasive Technology. Persuasive 2008. Lecture Notes in Computer Science, Vol 5033. Berlin: Springer-Verlag; 2008.

62.   Lorenc A, Ilan-Clarke Y, Robinson N, Blair M. How parents choose to use CAM: a systematic review of theoretical models. BMC Complement Altern Med 2009 Apr 22;9:9 [FREE Full text] [doi: 10.1186/1472-6882-9-9] [Medline: 19386106]

63.   Petty RE, Briñol P, Priester JR. Mass media attitude change. In: Bryant J, Oliver MB. editors. Media Eff Adv Theory Res 3rd ed New York: Routledge; 2009:125-164.

64.   Mankoff J, Kuksenok K, Kiesler S, Rode JA, Waldman K. Competing online viewpoints and models of chronic illness. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11).: ACM; 2011 Presented at: Proc SIGCHI Conf Hum Factors Comput Syst; 2011; Vancouver, BC, Canada URL: http://dl.acm.org/citation.cfm?id=1979027 [doi: 10.1145/1978942.1979027]

65.   Chen AT. The Relationship Between Health Management and Information Behavior Over Time: A Study of the Illness Journeys of People Living With Fibromyalgia. J Med Internet Res 2016 Oct 25;18(10):e269 [FREE Full text] [doi: 10.2196/jmir.5309] [Medline: 27780794]
66.   Robinson A, McGrail MR. Disclosure of CAM use to medical practitioners: a review of qualitative and quantitative studies. Complement Ther Med 2004;12(2-3):90-98. [doi: 10.1016/j.ctim.2004.09.006] [Medline: 15561518]
67.   Laiyemo MA, Nunlee-Bland G, Lombardo FA, Adams RG, Laiyemo AO. Characteristics and health perceptions of complementary and alternative medicine users in the United States. Am J Med Sci 2015 Feb;349(2):140-144 [FREE Full text] [doi: 10.1097/MAJ.000000000000363] [Medline: 25379625]
68.   Artus M, Croft P, Lewis M. The use of CAM and conventional treatments among primary care consulters with chronic musculoskeletal pain. BMC Fam Pract 2007 May 04;8:26 [FREE Full text] [doi: 10.1186/1471-2296-8-26] [Medline: 17480212]
69.   Azer SA, AlOlayan TI, AlGhamdi MA, AlSanea MA. Inflammatory bowel disease: An evaluation of health information on the internet. World J Gastroenterol 2017 Mar 07;23(9):1676-1696 [FREE Full text] [doi: 10.3748/wjg.v23.i9.1676] [Medline: 28321169]
70.   Kaicker J, Debono VB, Dang W, Buckley N, Thabane L. Assessment of the quality and variability of health information on chronic pain websites using the DISCERN instrument. BMC Med 2010 Oct 12;8:59 [FREE Full text] [doi: 10.1186/1741-7015-8-59] [Medline: 20939875]
71.   Bishop FL, Lewith GT. Who Uses CAM? A Narrative Review of Demographic Characteristics and Health Factors Associated with CAM Use. Evid Based Complement Alternat Med 2010 Mar;7(1):11-28 [FREE Full text] [doi: 10.1093/ecam/nen023] [Medline: 18955327]
72.   Kilgarriff A. Googleology is Bad Science. Computational Linguistics 2007 Mar;33(1):147-151. [doi: 10.1162/coli.2007.33.1.147]

Abbreviations

CAM: complementary and alternative medicine
CIH: complementary and integrative health
NCCIH: National Center for Complementary and Integrative Health
SERP: second search engine results page

XSL•FO
RenderX

Original Paper

Investigating the Role of Communication for Information Seekers' Trust-Related Evaluations of Health Videos on the Web: Content Analysis, Survey Data, and Experiment

Maria Zimmermann[1], Dr phil; Regina Jucks[1], Dipl-Psych, Dr phil

Department of Psychology and Sport Science, Institute for Psychology in Education, Münster, Germany

Corresponding Author:
Maria Zimmermann, Dr phil
Department of Psychology and Sport Science
Institute for Psychology in Education
Fliedner Straße 21
Münster,
Germany
Phone: 49 251 8339482
Fax: 49 251 8331399
Email: maria.zimmermann@uni-muenster.de

Abstract

Background: According to the language expectancy theory and the communication accommodation theory, health information seekers' trust evaluations of Web-based videos are determined by interplays between content and seekers' expectations on vloggers' appropriate language use in specific contexts of Web-based communication.

Objectives: Two investigations focused on differences both between vloggers' language styles and between users' general trust in specific Web-based platforms to investigate how the context of Web-based communication can be characterized (research question, RQ1). Thereafter, we investigated whether information uncertainty, vloggers' language style, and context of Web-based communication affect seekers' trust evaluations of videos (RQ2).

Methods: With a content analysis of 36 health videos from YouTube and Vimeo, we examined the extent of trust-related linguistic characteristics (ie, first-person and second-person pronouns). Additionally, we surveyed participants (n=151) on their trust in YouTube and Moodle (academic Web-based platform; RQ1). In an experiment, further participants (n=124) watched a video about nutrition myths and were asked to evaluate the information credibility, vloggers' trustworthiness, and accommodation of language by vloggers (RQ2). Following a $3 \times 2 \times 2$ mixed design, vloggers' explanations contained unambiguous (confirming or disconfirming) or ambiguous (neither confirming nor disconfirming) evidence on the myths (within factor). Furthermore, vloggers used YouTube-typical language (many first-person pronouns) or formal language (no first-person pronouns), and videos were presented on YouTube or Moodle (between factors).

Results: The content analysis revealed that videos on YouTube contained more first-person pronouns than on Vimeo ($F_{1,35}=4.64$; $P=.04$; $\eta^2=0.12$), but no more second-person pronouns ($F_{1,35}=1.23$; $P=.23$). Furthermore, when asked about their trust in YouTube or Moodle, participants trusted YouTube more than Moodle ($t_{150}\leq-9.63$; all $P \leq.001$). In the experiment, participants evaluated information to be more credible when information contained unambiguous rather than ambiguous evidence ($F_{2,116}=9.109$; $P<.001$; $\eta_p^2=0.14$). Unexpectedly, information credibility did not depend on vloggers' language style or the video platform ($F_{1,117}\leq2.40$; $P \geq.06$). Likewise, video's platform did not affect participants' evaluations of vloggers' trustworthiness ($F_{1,117}<0.18$; $P>.34$). However, participants judged vloggers who used a YouTube-typical language as being more benevolent, and their language use as being more appropriate in both video platforms ($F_{1,117}\geq3.41$; $P \leq.03$; $\eta_p^2\geq0.028$). Moreover, participants rated the YouTube-typical (vs formal) language as more appropriate for Moodle, but they did not rate one or the other language style as more appropriate for YouTube ($F_{1,117}=5.40$; $P=.01$; $\eta_p^2=0.04$).

Conclusions: This study shows that among specific Web-based contexts, users' typical language use can differ, as can their trust-related evaluations. In addition, health information seekers seem to be affected by providers' language styles in ways that depend on the Web-based communication context. Accordingly, further investigations that would identify concrete interplays

between language style and communication context might help providers to understand whether additional information would help or hurt seekers' ability to accurately evaluate information.

Introduction

Trust in Web-Based Nutrition Information

At least because of Popeye whose arms triple when he eats spinach and gains enormous power, the folk wisdom about spinach and its immense amount of iron has become anchored in the minds of the people. Even the information about the malpositioned decimal point that caused the falsely perceived iron content of spinach—and the followed scientific discussion about citation errors that have spread this false assumption—have not changed much [1].

Whenever people have questions about nutrition, the easiest way they can get information is on the internet. Using the internet has become a widespread way to inform people about nutrition [2,3], and how people process and evaluate Web-based health information affects their medical decisions as well as their health [4-7]. Although it can be challenging for people to get accurate information on the internet [2,4,8], the internet concurrently also offers the advantages of interactivity, anonymity, and a low threshold for getting information.

Regarding getting information from Web-based videos, people reported that they watch Web-based videos not only to gain knowledge [9] but also to support their individual learning [10,11]. However, viewing videos from video platforms on the Web may put (lay)people at risk of getting inaccurate, incomplete, or inadequate information not only because the content of Web-based videos is rarely reviewed and, therefore, the quality varies vastly [12,13] but also because information seekers have difficulty evaluating whether the obtained information is true [7]. In this sense, determining the objective accuracy of any health information (video or otherwise) is challenging as even medical evidence rarely attains absolute certainty and can always be proved wrong by supplemental evidence [7,14].

Thus, information seekers' risk of encountering inaccurate information on the Web comes from both the proportion of information that is likely inaccurate and the inability of a (lay)person to filter inaccurate information [8]. Consequently, information seekers need to rely on qualified information providers, as their own expertise about the topics of interest is likely incomplete, and their resources for evaluating the content according to academic criteria are limited (eg, evaluating whether the evidence given to support a claim is based on plausible scientific methods [4,15]). Instead, they need to identify criteria that indicate whom to rely on [7]. Therefore, in addition to or instead of understanding the academic criteria that can be used to assess the actual quality of information, it is also critically important to understand what kind of criteria

information seekers use to determine that they can trust someone to provide adequate information [7,8,16,17].

Although various phenomena related to trust are mirrored in diverse research fields—leading also to many definitions of trust [18-21]—we consider the common principles of trust when examining the relationship between information seekers and providers. As such, relationships between information seekers (trustors) and providers (trustees or trust objects) [22-25] entail an actual or perceived imbalance in each party's extent of knowledge. In turn, trusting the information provider requires that the information seeker willingly accepts that he or she can never really know if the information is adequate or whether the provider indeed has more knowledge (uncertainty of trustors [21-23,25,26]). Thus, health information seekers' risk being misinformed if the information provider offers inaccurate, incomplete, or inadequate information. In this sense, several characteristics of trust objects (such as the provider's trustworthiness or attractiveness) are considered antecedents of trust. Moreover, (dis)trust is considered to be either an outcome from, or a formative, or a conditional aspect of a trusting or a distrusting relationship [23,25,27].

So far, the research on seekers' evaluation of Web-based information has shown that several factors, such as the information source, the content of the information, and the type of media, are important when people evaluate the credibility of Web-based information in general [28-32]. Regarding health information specifically, the kind of criteria that people use to judge the credibility of Web-based health information is often related to the characteristics of the Web-based information and the provider of the Web-based information [16,22]. Although people's judgments of a provider's trustworthiness correspond to judgments about the credibility of information and vice versa [33], single factors, such as the provider's language and the content of information, seem to influence information seekers' evaluations of the information and the provider in different ways [34,35]. Therefore, in the rest of the study, we consider that people's willingness to rely on both the characteristics of the provided information and those of the information provider are concepts (hereinafter called the *credibility of information* and the *trustworthiness of providers*) that, although they depend on one another, can be influenced differently by individual factors.

The main aim of this study was to extend the research on how health information seekers evaluate information and information providers. We consider the relationship between health information seekers and providers to be mainly determined communicative patterns. Thus, in the following sections, we outline how the information providers' language style and the context of the Web-based communication can be viewed as factors that information seekers analyze to determine whom and which information to rely on.

The Communication Between Health Information Seekers and Providers

According to language expectancy theory (LET) [36] and communication accommodation theory (CAT) [37], gaining knowledge when seeking health information on the Web is not only determined by the content of the information (ie, what is said) but also by who is communicating (ie, who is saying it), their manner of communication (ie, how are they saying it), and the context of communication (ie, *where* takes the communication place). In particular, LET predicts people's attitude change (eg, imparting of knowledge) on any violations of their language expectancies that represent people's expectations on interlocutors' appropriate language use in a given context. Accordingly, information seekers should evaluate providers positively if providers use language styles that fit with cultural values and situational norms (ie, no violations). Similarly, information seekers should evaluate providers positively if providers use language styles more favorably than expected in a situation (ie, positive violation, eg, a doctor uses unexpectedly easy everyday terms as the doctor was used to use difficult medical jargon at previous appointments). Contrary, they should evaluate providers negatively if providers use language that conflicts with cultural values and situational norms (negative violations). Similar to what LET predicts, the CAT also predicts that information providers should be evaluated more positively if they use appropriate (ie, accommodative) language in a specific context. According to CAT, the interlocutor or audience is an important aspect that determines the contextual norms of communication.

In sum, both theories on interpersonal communication take into account the context of communication. In this sense, LET and CAT consider that the rules and norms embedded in the context of communication influence how people communicate, how people evaluate their interlocutors, and how people evaluate the appropriateness of their interlocutors' language use given a certain context; hence, these factors influence the communicative success.

Health Information Providers' Language Styles

On the Web, the language style used by the information provider offers especially salient information about the provider [15,16,38,39] and, therefore, is likely to influence an information seeker's perceived trustworthiness of the provider [5,15,38-43]. In this sense, an information seeker might perceive a specific language style used by the provider and then use it to judge some characteristics of the provider such as their competence or benevolence. In general, one's language style can be characterized by numerous linguistic aspects (eg, the use of technical terms, self-references, or hedges), and a health information provider's language style might impact a seeker's judgments about their credibility and trustworthiness [34,35,40,41].

Information providers' use of personal references (ie, referring to oneself or someone else) is closely associated with how people perceive providers' expertise [40,41] and how much people learn from them [44], and hence, such references are of special interest for seeking health information. Linguistically, as self-references (eg, first-person pronouns) are related to the extent of a speaker's self-disclosure, they might offer information about providers' trustworthiness in 2 ways: on the one hand, self-disclosure is considered important for establishing trust in Web-based communication, as it signals a willingness to open up and, hence, might promote the reciprocal exchange of information [38]; accordingly, high number of self-references can lead to higher credibility judgments [45]. On the other hand, it seems crucial that people share just the right amount of personal information, as sharing too much personal information and using too many first-person pronouns can also be perceived as unprofessional [41]; accordingly, high number of self-references are also found to negatively influence the trustworthiness of information providers on the Web [40,46]. Overall, the evidence on how self-disclosure affects trust-related evaluations is sparse and conflicting. Similarly, research has revealed conflicting evidence on how providers' use of technical language (also associated with providers' trustworthiness [41]) affects credibility judgments, as using high amounts of technical language has led to both higher [41,47] and lower credibility judgments [48,49]. These conflicting findings on the individual impact of providers' language styles on trust-related evaluations might strengthen the assumptions of LET and CAT, which state that people's evaluations of providers are not only influenced by the language style itself but by a complex interaction among the content, the interlocutors, the provider's language use, and the context of communication.

The Context of Web-Based Communication

Context is considered crucial in several approaches on people's evaluations of Web-based information [5,32,50-56]. Although context is always assumed to influence people's evaluations, what constitutes this context has been defined, however, with different levels of concreteness and using different aspects [56]. For instance, context has previously been described in terms of how long people spend searching for Web-based information [53,57], the characteristics of information seekers (such as their self-disclosure, expectations, or experiences) [52,58], the interaction of several aspects [50,56], or differences between seeking information on the Web or in *naturalistic* (ie, offline) settings [59].

Although context has often been assumed to influence people's evaluations, it has been, so far, rarely investigated in research about credibility on the Web, which may be because of the challenges in defining and conceptualizing context for Web-based as well as offline information seeking [33,56]. However, a review of the theoretical frameworks used to assess Web-based information emphasizes that to operationalize credibility, considerations, and conceptualizations of different Web-based contexts need to be outlined, as we must investigate these contexts to understand people's credibility judgments better [33]. Furthermore, Web-based health information seekers risk misjudging even accurate information if they do not recognize the context in which the information is supposed to be interpreted (also described as *context deficit* [5,60]).

By combining the concept of context in LET and CAT with approaches that have been used to study the impact of online media and its affordances (relating to its features) on people's evaluation of Web-based information [32,61], we consider the

XSL•FO

RenderX

context of Web-based communication to cover the norms and rules about users' typical use of online media as well as users' expectations of typical media use. Expectations of typical media use relate to the affordances that online media offer (such as navigability or recordability), where *affordances* refer not only to the simple features but also to the dynamic relationships among those features, the cues (eg, *button* for record) that features offer users, and how users use those cues [61]. Importantly, these affordances should be considered as part of the Web-based context, as research has considered several affordances of online media that might affect trust-related evaluations of Web-based information and information providers [30,32,35,62].

By considering the concept of online media affordances, we derived the following definition for the Web-based communication context.

The Web-based communication context is determined by certain norms. These norms emerge from and shape continuous dynamic relationships among (1) affordances of online media, (2) cues accompanying these affordances, and (3) how users use these cues [63].

This definition complements the concept of affordances by adding the concept of norms, which includes users' typical use and users' expectations on typical use. Norms or rules should constitute and emerge from users' (typical) use of online media—irrespective of explicitly named rules of the media. For instance, if specific affordances of online media create norms on how to communicate typically, these norms should also entail users' expectations about an appropriate way to communicate in this media [64]. In this sense, the community of online media (ie, users) is crucial for establishing norms and, hence, is an important aspect of the context [37]. Thus, the definition of contexts closely refers to research on norms of online communities: for instance, group member's awareness of norms that were caused by online communities' characteristics (ie, size and extent of use, lack of nonverbal cues, anonymity, warranting, communication record, and community type [65]), norms of online communities that are caused by sociotechnical interactions with characteristics of the Web-based environment [66], or sharing norms as interactions within virtual communities [67]. In this sense, it is likely that online communities may also evoke norms for community-specific language styles [64].

Previous Research on the Impact of Contexts of Web-Based Communication

With respect to rather typical language styles for specific contexts of Web-based communication, research has identified differences in linguistics characteristics between Facebook, Twitter, and YouTube, where the affordances of the media platform cause certain differences such as the length of the texts: Facebook posts can contain unlimited numbers of characters, whereas Twitter posts (tweets) are limited to 140 characters [68]. Similarly, YouTube's affordances of high anonymity and heterogeneity among users may cause more foul language on YouTube than on Facebook [69].

With respect to how the context of Web-based communication affects trust-related evaluations, previous studies have shown that users are affected by different types of online media itself (ie., websites, blogs, bulletin boards, and internet) [70] as well as by different affordances and cues related to a type of online media [71,72]. With respect to the community as an important aspect of Web-based contexts, information seekers evaluated medical experts as being more trustworthy when they provided information in online forums for medical professionals rather than in online forums for lay persons; hence, seekers evaluated the trustworthiness of medical experts differently depending on the specific online community [34].

Moreover, research also indicates that the appropriateness of a health information provider's language style is crucial not only when health information is sought offline [73,74] but also when it is sought on the Web [34,75-79]. In this sense, when providers use language that is accommodative toward the (online) audience, this is considered appropriate, as it takes into account the audience's communication preferences and, in turn, enhances their satisfaction with the communication [37,74].

As users of online media are considered to show and expect the norms of a particular (Web-based) context of communication [37,65], it is interesting that some of the research done so far, although there are not many studies, have identified differences in users' language styles for different online media [68,69]. Furthermore, 1 study also found that using a particular language style (ie, the number of self-references in messages) led users to evaluate a provider's trustworthiness differently depending on whether the language style was used on Facebook or Twitter [62]. Thus, participants' evaluations were seemingly impacted by the provider's language style as well as by the context of the Web-based communication. Overall, research indicates that users of different Web-based platforms communicate in different ways, and they evaluate information and providers' language styles differently depending on the context of Web-based communication.

Rationale

Seeking health information on the Web is a common way to obtain knowledge about health-related information and can influence personal health decisions [6]. In this sense, health information seekers need to use criteria that indicate whom they can trust to provide accurate, complete, and adequate information [7]. As such, it is important to identify these criteria and investigate their impact on trust-related evaluations. As seeking health information on the Web involves communication between information seekers and providers, trust-related evaluations depend not only on the content of the information but also on how seekers evaluate the providers and their manner of communicating and also on the context in which the communication takes place. Moreover, information seekers' evaluations also are shaped by their expectations about appropriate language use, given a specific context [36,37]. According to LET and CAT, health information providers' language styles and the context of Web-based communication should, therefore, impact health information seekers' evaluations of providers and the information in 2 ways, individually and reciprocally. However, research considering both aspects is sparse.

As it is difficult to conceptualize the context of Web-based communication, we aimed to identify differences between specific Web-based contexts that are associated with users' typical use [36,37,61], as this would not only extend previous findings on Web-based context-specific norms but it would also help to characterize specific contexts more concretely. For this purpose, we investigated differences in users' language styles (first investigation) and users' trust evaluations (second investigation) regarding specific contexts of Web-based communication that are relevant when watching Web-based health videos. As such, the Web-based communication context is defined by norms and rules about the online media affordances, cues accompanying affordances, and how users use these cues. Accordingly, diverse Web-based platforms may adequately represent and operationalize different contexts of Web-based communication, as users' use of Web-based platforms and these platforms' affordances should constitute the norms corresponding to users' typical use of the Web-based platforms.

> *Research question (RQ) 1: Are there differences in users' language styles and users' trust-related evaluations regarding specific contexts of Web-based communication?*

Building on the findings of this study's first and second investigations, which address RQ1, this study's third investigation examined experimentally whether health information seekers' trust-related evaluations of a Web-based video about nutrition myths were affected by the content of information, the language style of the information provider, and the context of the Web-based communication.

> *RQ2: Here, we consider an interaction between the language style of health information providers and the context of Web-based communication: are health information seekers' willing to rely on information providers and the information they provide impacted by the information itself, the information provider's language style, the appropriateness of this language styles given the specific context of Web-based communication, and the context of Web-based communication?*

The aim of this study's first 2 investigations was to gain empirical insights into certain contexts of Web-based communication. We, therefore, explored whether people communicate and trust differently depending on the Web-based context—namely, the Web-based platform. Accordingly, we first investigated whether speakers of Web-based health videos on Vimeo or YouTube used different linguistic styles. As we further aimed to extend the research on how Web-based platforms individually impact users' trust evaluations [70], we, in a second study, investigated whether users differ in their willingness to trust YouTube and Moodle (a platform for universities to share learning material). The empirical findings from both investigations allowed us to combine the theoretical assumptions about the context of Web-based communication derived above with actual findings from specific Web-based platforms.

Background for Investigation 1: Vloggers' Language Styles on YouTube and Vimeo

Although YouTube is arguably the most popular social media platform to watch, share, and comment on Web-based videos (eg, based on monthly unique viewers [80]), we aimed to compare YouTube with a Web-based platform that is similar in its purpose and is also popular. Thus, we selected Vimeo, belonging to InterActiveCorp, which is also a social media platform to watch, share, and comment on Web-based videos; it has about 60 million registered users [81] and is used by about 240 million unique users per month. As such, Vimeo has significantly fewer users than YouTube, which is used by over a billion people [82]. Moreover, every day 5 billion videos are watched on YouTube, and every minute 300 hours of video material is added [83]. Together, YouTube and Vimeo were accessed most frequently by internet users to watch videos in the United States [84], with Vimeo, after YouTube, reaching the second highest extent of internet users in the United Kingdom [85]. Furthermore, YouTube and Vimeo differ with respect to their financing. Although the free version of YouTube—which is probably used by users most often—is mainly financed by advertising displayed before and during the videos, the ad-free platform Vimeo offers both free memberships and *pro* memberships, and the whole platform is financed exclusively through these pro memberships [81]. As such, Vimeo also offers services that can increase the quality of videos, and it further curates films to provide high-quality content.

Therefore, YouTube and Vimeo can be considered examples of Web-based platforms whose differences might lead to community-specific norms, which may result in community-specific language styles [36,37,64], but that are still comparable in terms of most users' purposes for use. So far, however, only a few studies have identified linguistic differences between Web-based platforms [68,69]. In 1 study, a content analysis of contributions on Web-based platforms revealed that nonprofit advocacy groups use different propaganda characteristics in their messages on Facebook, Twitter, and YouTube [86]. When posting on YouTube compared with other platforms, these organizations used more authority figures; conversely, when posting on Twitter compared with Facebook or YouTube, they more often reduced complex issues. Similarly, users seemed to use more foul language on YouTube compared with other platforms like Facebook, which might because YouTube has heterogeneous users and is highly anonymous [69].

We aimed to extend these findings by investigating whether information providers in Web-based health videos use different number of personal references on Vimeo or YouTube. We focused on first- and second-person pronouns as they represent the use of personalization in language styles in Web-based videos [13,87] and are crucial for trust-related evaluations [38,40,41].

Background for Investigation 2: Users' Trust in YouTube and Moodle

As we aimed to investigate differences in users' trust-related evaluations in different Web-based contexts of communication,

we investigated whether users' trust in Web-based platforms differs. In this sense, it is worthwhile to determine if users do trust in Web-based platforms to provide Web-based services because trust in these trust objects (services) might also affect trust in associated trust objects (the information on the platforms). Accordingly, any differences between users' trust in these platforms additionally allow one to derive hypotheses about whether such platforms have individual effects on information seekers' evaluations of the information provider and the information (this study's third investigation). In this sense, differences between how users evaluate the same information (regarding whether they trust it) that occurs on different Web-based platforms [70] might be caused by users' general trust in those platforms [30,88].

We decided to investigate users' trust in YouTube versus in a Web-based platform that obviously differs in its operators but still offers users the opportunity to watch Web-based videos. YouTube and Moodle represent good platforms to compare, as Moodle is a platform provider for educational institutions, and its main purpose is to share learning materials such as book chapters, course curricula, and videos; conversely, YouTube is the most popular social media platform to watch, share, and comment on Web-based videos.

Research on trust in Web-based platforms [89-92] considers trust similarly to how we have described it above, where research aims to both investigate trust as a formative aspect leading to trust-related actions regarding the trust object (such as using a Web-based platform [89,90]) as well as to investigate characteristics of Web-based platforms as antecedents of trust (such as trustworthiness or usability [91,92]). To investigate whether there are differences between both platforms in terms of people's trust in the platform, we assessed people's perception of each platform's characteristics and people's self-reported willingness to interact with the platform as well as people's self-reported familiarity with the platform when using for a certain purpose [86].

Background for Investigation 3: The Impacts of the Uncertainty of Information, Information Provider's Language Style, and the Context of Web-Based Communication on Health Information Seekers' Trust-Related Evaluations for a Web-Based Video About Nutrition Myths

According to LET and CAT, whether health information seekers rely on information providers and their provided information depends not only on the content of information but also on the providers' language styles, the contexts of Web-based communication, as well as on seekers' expectancies about appropriate language styles, given specific Web-based contexts (ie, the reciprocal impact of language styles and contexts of Web-based communication). In this sense, research indicates that information seekers make use of not only single factors related to information providers, the information itself, or the media through which information is transferred when evaluating Web-based information [29,33,34,46]. Thus, this study's third investigation intended to investigate how information seekers are affected by the integrative consideration of single factors at these levels. Hence, building on the findings of this study's first

2 investigations, the following experiment examined how the uncertainty of information, the language style of the provider, and the context of Web-based communication affect individually and reciprocally seekers' evaluations of information credibility and the provider's trustworthiness when watching a Web-based video about nutrition myths (RQ2). Therefore, a $3 \times 2 \times 2$ mixed-design Web-based experiment was conducted with the factors being uncertainty of information (uncertain, confirming the myth, and not confirming the myth), language style of the vlogger (YouTube-typical vs formal language style), and the context of Web-based communication (Moodle vs YouTube).

The Web-based platforms YouTube and Moodle were chosen as contexts of Web-based communication because they contain features that should determine differences between users' typical use of both platforms and should, therefore, impact health information seekers' trust-related evaluations. Moodle is a platform provider for educational institutions, and the main purpose of the use is to share learning materials, such as book chapters, course curricula, and videos. Accordingly, videos on Moodle are most often uploaded by academics who are providing learning material to their students. Conversely, videos on YouTube are not necessarily uploaded by academics because every user is allowed to upload content for several purposes. Hence, on Moodle, the homogenous community is made up of rather *educationally close* members compared with the heterogenous community on YouTube. As the community (ie, users) of both Web-based platforms differs, it is likely that these platforms represent contexts of Web-based communication with different norms [36,37,65]. Accordingly, the platforms should impact how users communicate on these Web-based platforms and how they evaluate the appropriateness of information providers' language use [64]. In addition, users on both platforms can watch Web-based videos, but users seem to trust in these platforms differently. In particular, people seem to be more familiar and more willingness to transact with YouTube compared with Moodle, whereas they do not ascribe higher ability, benevolence, nor integrity to YouTube compared with Moodle (this study's second investigation).

In the following, we summarize the theoretical and empirical approaches for each factor and the interplay among a vlogger's language style and the context of Web-based communication to derive hypotheses in terms of factors' individual and reciprocal impacts.

Uncertainty of Health Information on the Web

Research on the credibility of Web-based information indicates that the content of information affects people's evaluation [29,33,35]. Thereby, the uncertainty of information has been highlighted in several studies because often scientific information entails preliminary findings—which are further often discussed controversially, especially if findings' origins and implications are attributed to conflicting scientific trends [93]. On the Web, seeking health information can bring to light the uncertainty of information, as people not only become aware of conflicting opinions about symptoms' origins, diagnosis validity, and optimal treatments but they can also be confronted with numerous conflicting opinions [94,95]. By focusing on ambiguity as an antecedent of uncertainty, the ambiguous

representation of an issue has been shown to make people see evidence as risky and avoid making decisions, and these outcomes appear to depend on the origin of the ambiguity—for instance, whether the information is either conflicting or incomplete [96].

Accordingly, people's confusion has also been investigated for conflictual information in Web-based newspapers [97]. In that work, studies with more conflictual information about deep brain stimulation was associated with people expressing higher levels of uncertainty and commenting more negatively on these studies. Similarly, people judged an original message about the risk of red meat to be more credible if they received another message about its risks (a consistent message) compared with a message about its benefits (a conflicting message); interestingly, these conflicting and consistent messages did not influence the perceived trustworthiness of the information provider [98]. In addition, conflicting rather than consistent information about the risky use of a drug to treat high cholesterol made participants less likely to recommend the medication to a fictitious friend [95]. However, in a study where people received analogous information about both the risks and benefits of either an unfamiliar or a familiar food supplement (ie, conflicting information), people reflected more about these risks and benefits if they generally believe knowledge to be dynamic and complex and if they were given information about the less familiar (CQ10) supplement [99]. This, in turn, indicates that people's epistemic beliefs and their familiarity with the topic influence how they process ambiguous information. In general, people seem to heuristically use rules of thumb when receiving conflicting instead of consistent information, and therefore, they seemingly use different strategies to decide about their health depending on whether they are processing conflicting or consistent information [100].

Although these findings predominantly indicate that people seem to perceive consistent information to be more credible than conflicting information, people's judgments in this context are, for instance, also influenced by the following factors: the order of presented information, people's prior epistemic beliefs, and people's attitude toward the topic [101]; the complexity of information and the number of sources that provide the information [102]; and the overall trustworthiness of the information provider [103]. Hence, again, instead of considering only the uncertainty of information, it is more realistic to also consider characteristics regarding the provider of information.

As nutrition myths are widespread and collective conceptions about nutrition arise because they are reiterated over and over, people's familiarity with such (mis)conceptions causes them to believe that the (mis)information is correct—irrespective of the scientific origin of the information [1,104,105]. After considering the summarized evidence for how people evaluate the uncertainty of Web-based information, we have derived the following presumptions and accompanying hypotheses for how people evaluate Web-based videos about nutrition myths:

> *H1: When discussing a popular nutrition myth, a vlogger's scientific explanations that confirm, disconfirm, or neither confirm nor disconfirm the myth are expected to impact how people judge the credibility of information. People should evaluate a video's information that is consistent with the nutrition myth (confirming the myth) as being more credible compared with a video that gives conflicting information (disconfirming the myth) or that gives information that neither confirms nor disconfirms the nutrition myth. In addition, people should evaluate a video's information to be more credible when scientific explanations neither confirm nor disconfirm the myth compared with those explanations that disconfirm the nutrition myth.*

Language Style of Vloggers

According to LET and CAT, the success of seeking health information on the Web is determined by how health information seekers evaluate the providers' language style. Especially on the Web, the language style of an information provider is considered an outstanding cue to assess characteristics about the provider [15,16,38,39]. Indeed, research, so far, indicates that the provider's language style is a cue that impacts whether people perceive the information to be credible and the information provider to be trustworthy [34,35,40,41].

In this context, linguistic aspects of personal references (eg, first-person and second-person pronouns) are of special interest for seeking health information, as providers' use of personal references is closely associated with providers' expertise [35,40,41]. According to the findings in this study's first study, a typical characteristic of health videos on YouTube is the usage of first-person pronouns. Similarly, so-called conversational language is achieved by linguistic aspects of self-disclosure [106]: That is, changing third-person into first-person pronouns makes it seem as if the author (or vlogger) is talking to you and is expressing a personal experience. As a YouTube-typical language style involves self-disclosure by the communicator, it might be an important mechanism for establishing trust in Web-based communication [38,41]. However, it is unclear to what extent sharing personal information might promote a trustful relationship: it could be that sharing personal information entails the risk of reducing the perceived expertise of the provider, as experts are expected to conduct themselves in an objective and unbiased manner. Accordingly, fewer first-person pronouns in Web-based medical advice predicted a higher perceived expertise of advice givers and helped people to determine that the advice givers were experts rather than laypersons [41]. Similarly, mental health information in messages on Facebook and Twitter that included personal testimonials (compared with those without personal testimonials) led people to think more critically about the provider [46]. In addition, another study showed that high self-disclosure in Web-based advice negatively impacted the perceived benevolence of the provider (as one aspect of trustworthiness) [40]. In contrast to this, it could be particularly self-references that may signal the willingness to open up and, hence, may promote a trustworthy relationship, as this could lead to reciprocal exchange of information [38]. Accordingly, there is also some evidence that a high number of self-references led to higher trust-related judgments [45].

To sum up, the use of personal reference seems to affect people's judgments of a provider's trustworthiness. The assumptions on the impact of personal references are ambiguous, but in the majority, the use of first-person pronouns decreased people's perceived trustworthiness of providers, which could be caused by people's expectation on how experts should communicate. On the basis of these thoughts, we formulated the following hypothesis:

> *H2: A vlogger's use of YouTube-typical language (the use of first-person pronouns) in a Web-based video causes people to evaluate the video's information as being less credible and the video's vlogger as being less trustworthy compared with when vloggers use formal language (no first-person pronouns).*

Context of Web-Based Health Videos

As previously described, the context of Web-based communication should be taken into account in addition to the content of information and the language style of information providers when investigating health information seekers' trust-related evaluations of Web-based information and information providers. According to LET and CAT, the context of Web-based communication should impact not only how people communicate but also people's evaluations of the interlocutors—irrespective of one's used language style [36,37]. Hence, besides affecting users' language use [68,69] (see also this study's first investigation), different Web-based contexts might also individually influence people's judgments of information and providers. The context of Web-based communication is considered norms and rules about the online media affordances, cues accompanying affordances, and how and what for these cues are used by users. As such, diverse Web-based platforms may represent different contexts of Web-based communication. However, evidence on any individual impact of Web-based platforms on trust-related evaluations is sparse [70]. Furthermore, even within a Web-based platform, some features are associated with distinguishable norms. For instance, this would be the case for an online health forum platform where a panel of experts communicates, in 1 forum thread, mainly to an audience of medical professionals and, in another forum thread, to an audience primarily of laypersons. In this context, the experts who provided the same information to medical professionals were evaluated to be more trustworthy compared with those who provided the information to laypersons [34].

In case people judge the credibility of Web-based health information and information providers differently depending on where the information came from [34,70], the communication context might serve (heuristically) as a cue for health information seekers and, hence, influence how they evaluate Web-based information. Thus, in line with approaches focusing on the *halo effect* as a phenomenon leading information seekers to biased judgments [30,88], seekers' perceptions and evaluations of the Web-based platform might be transferred to the information found on this platform. Accordingly, a positive or negative impression of a Web-based platform would affect seekers' judgments of information and providers in the same

direction regardless of whether the platform is the source of information or rather acts as a mediator. Putting together, the contexts of Web-based communication are assumed to influence information seekers' trust-related evaluations [30,36,37,88]. Thus, people's impressions of Moodle and YouTube might likewise affect people's evaluations of the Web-based health videos being presented on these platforms. As people seem to trust in YouTube rather than in Moodle (this study's second investigation), the following hypothesis is derived.

> *H3: As people are more familiar and willingness to transact with YouTube compared with Moodle and this causes them to trust YouTube more than in Moodle, a Web-based video presented on YouTube causes health information seekers to evaluate the video's information to be more credible and the video's vlogger to be more trustworthy compared with when they evaluate the information and the vlogger of a video presented on Moodle.*

Language Style in Specific Contexts of Web-Based Communication

In addition to studying the main effects of the context of Web-based communication and the vlogger's language style in a Web-based video, we are also interested in whether both of these factors interact and, hence, lead people to evaluate the credibility of information and trustworthiness of vloggers differently because of a perceived appropriateness of language style, given a certain context of Web-based communication. Again, by referring to our considerations described in this study previously, the context of communication is expected to influence users' language use within a certain context and therefore should also influence other users' expectancies on either appropriate or not appropriate language use given this context. In turn, users' evaluations of the appropriateness of information providers' language use given a certain context, should also impact users' trust-related evaluations of information and information providers [5,36,37]. Thus, even the same provider's language style can cause differences in the quality of trust-related evaluations [76-78], as according to LET and CAT, the appropriateness of language styles given a certain context—respectively community—is also crucial for people's evaluations.

There are some studies indicating that information seekers' trust-related evaluations of the Web-based information and the information provider are influenced by whether people perceive the provider's language use as either appropriate or not appropriate given a certain Web-based context. For instance, when (lay)people were exposed to nutrition information that was provided in the forum intended for medical professionals, they perceived the information which had a high amount of medical technical jargon to be more credible compared with the same information that was provided in the forum intended for laypersons. Conversely, they perceived the information that was provided in the forum intended for laypersons, which had a low amount of medical technical jargon to be more credible compared with the same information that was provided in the forum intended for medical professionals [34]. Hence, people evaluated the information to be more credible when the language

was adapted to the intended audience. Similarly, in a different study, when information providers used either personalized messages (ie, messages with first-person pronouns and self-disclosure concerning provider's health) or depersonalized messages (ie, messages without first-person pronouns and without self-disclosure), people rated the providers' trustworthiness differently depending on whether the messages were presented on Twitter or Facebook. Although it is not clear whether 1 of the 2 language styles was expected to be more typical for either Twitter of Facebook, people rated providers who tweeted depersonalized messages on Twitter to be more competent than providers who posted with depersonalized messages on Facebook; conversely, people rated providers who posted personalized messages on Facebook to be more competent than providers who tweeted personalized messages on Twitter [62]. Nevertheless, the research on any reciprocal impact of information provider's (appropriate) language style and the context of Web-based communication is sparse.

By considering these thoughts, we assumed that people find it more appropriate if a vlogger's language style in a Web-based video matches people's expected rules and norms of its platform [36,37]. More simply, a YouTube-typical language style (such as the use of first-person pronouns: this study's first investigation) might be perceived as appropriate on YouTube. Conversely, as Moodle is a learning platform that includes materials about academic issues, low self-disclosure (such as the nonuse of first-person pronouns in formal language) might be seen as more professional and appropriate on Moodle [41]. Hence, we derived the following hypothesis on an interaction effect:

H4: A Web-based video presented on YouTube causes people to evaluate the information to be more credible, the vlogger to be more trustworthy, and the vlogger's language use to be more accommodative when the vlogger uses YouTube-typical language than formal language. Instead, a Web-based video presented on Moodle causes people to evaluate the information to be more credible, the vlogger to be more trustworthy, and the vlogger's language use to be more accommodative when the vlogger uses formal rather than YouTube-typical language. As it is assumed that people evaluate the credibility of information and trustworthiness of the information provider depending on their perception of how appropriately a vlogger's language accommodates the community of a certain Web-based platform, participants were also asked to rate how they perceived the accommodation of the vlogger's language use.

Methods

Investigation 1

Procedures

To identify characteristics of users' language styles that are typical for either Vimeo or YouTube, we systematically selected health videos from each platform based on the most relevant hits of keyword searches. On YouTube, searches were conducted

on July 26, 2016 (for *health*), and July 28, 2016 (for *diet*); on Vimeo, searches were conducted for both *health* and *diet* on August 1. For each search term, the first most relevant 50 results were collected (the option of filtering by the *most relevant* was selected). Videos were excluded if they were longer than 10 min as the duration of about half of the 332,382 investigated Web-based videos in connection to YouTube's traffic were between 3 and 5 min [107]. In addition, videos were excluded for the following reasons: if the video was an advertisement (eg, movie trailers, brands, and products; except for explanations by health institutions), if the video did not have spoken content, if more than 1 person was speaking, or if the spoken language was not English. None of the videos appeared under both search terms. After exclusion and deletion, 36 videos remained (25 videos from YouTube and 11 from Vimeo). Furthermore, the following characteristics were recorded: original rank of video, the presence of a protagonist, duration, declared or apparent topic, or aim of the video (Multimedia Appendix 1).

Analyses

The transcripts of videos were analyzed with Linguistic Inquiry and Word Count (LIWC2007) by Pennebaker Conglomerates, Inc [108], to identify differences in the number and quality of pronouns. LIWC2007 is a computer program that analyzes the absolute and relative word frequency of texts based on included dictionaries and predefined categories of words. Transcribed texts of videos were implemented to LIWC2007, and analyses were conducted to identify differences in the relative number of pronouns used in the transcripts of the YouTube and Vimeo videos. Therefore, we used the LIWC's dictionary, including all singular and plural forms of first-person, second-person, and third-person pronouns (included in dictionary but irrelevant for this investigation). We analyzed the number of first-person and second-person pronouns relative to the overall word frequency of a transcript and tested differences between the extent of pronoun use on both platforms by conducting a Welch analysis of variance that is relatively robust against unequally distributed variances in dependent variables among Web-based platforms.

Investigation 2

Procedures

We conducted a Web-based survey by using *Questback's EFS Survey* [109]. Through an automatic balanced randomization, participants were randomly assigned the order in which they answered questions about the 2 platforms, either first about YouTube and then about Moodle, or vice versa. A total of 75 participants first answered items concerning Moodle, whereas 76 participants first answered items concerning YouTube. In the survey, both platforms were briefly introduced to guarantee that participants were equally familiar with each of the platform's purposes. YouTube was introduced as follows: "The YouTube video platform is mainly used for viewing, providing, commenting and sharing videos, and it is the largest and most comprehensive Web-based video portal." Moodle was introduced as follows: "The learning platforms of Moodle serve educational institutions (eg, universities) mainly to provide learning opportunities. In the Web-based courses on Moodle (which often accompany in-person courses), lecturers and

learners can upload learning materials such as texts and learning videos."

Participants

People were invited to take part in the survey through a link on Web-based platforms run by several German universities. A total of 151 participants aged between 18 and 61 years (mean 30.21 [SD 9.78]) took the survey. Of 151, 88 participants stated that they are studying, 54 participants declared to be employed, and 148 participants declared German to be their first language. The participants' self-reported weekly computer usage was, on average, 28.91 hours (SD 16.44), and their average weekly internet usage was 24.35 hours (SD 13.31). Furthermore, participants rated the frequency (from 1: daily to 5: never) of watching videos on YouTube, in online media libraries, and on streaming services to be, on average, 2.34 (SD 0.93)—meaning *several times a month*.

Measures

Participants answered 14 items on a 7-point Likert scale from 1 (I strongly disagree) to 5 (I strongly agree) once for Moodle and once for YouTube. Items were adapted from the scale *Trust in Online Firms* [86]. On the basis of a study by Mayer, Davis, and Schoorman [25], 2 items for ability, integrity, and benevolence and 1 item for the platform's overall trustworthiness were used to assess the trustworthiness of the platform (eg, *Moodle or YouTube has the skills and expertise to provide videos in an expected manner* or *Moodle (or YouTube)makes good-faith efforts to address most customer concerns*). Moreover, people's *willingness to transact* was assessed using 3 items (eg, *I watch videos on Moodle or YouTube* or *I am likely to utilize the services provided by Moodle or YouTube)* [86], and people's perceived familiarity with the platform was assessed by using 4 items (eg, *I am familiar with searching for videos on Moodle or YouTube*) [86].

Investigation 3

Exclusion Criteria and Participants

An analysis of power (assuming 1–beta=.80; Cohen f^2=.0625) was used in advance to determine that a sample size of N=113 would be sufficient assuming a small effect size. Using a mailing list of a large German university, participants were invited to participate in a Web-based survey and received a €8 voucher as reimbursement. In total, 128 participants completed the survey, in which people who were interested in participating were automatically excluded by the survey programming in the following cases: if they stated that they are studying or had studied medicine, food chemistry, or nutrition science (as it was assumed that their high prior knowledge would influence credibility and trustworthiness judgments) and if they stated that they are studying psychology or chemistry and pharmacy (it was possible that they might know the psychology lecturer acting as the nutrition vlogger in the experiment video or they might suspect the chemistry course in the Moodle condition to be artificial created). Furthermore, survey access was automatically denied when using a mobile phone (thus, the screen size of used devices was controlled). All participants specified at the end of the survey that they did want to provide their data for research purposes. Furthermore, 2 participants

were excluded from data analysis because they stated that they knew the psychology lecturer acting as the vlogger. In addition, 2 other participants were excluded because their time to complete the survey took more than 1 standard deviation above the overall mean duration of all participants (mean 40.5 min [SD 130.55]).

Although 20 participants did not recognize the video platform, it is unclear whether these participants might have unconsciously perceived platforms. Nonetheless, we decided to include these participants' data in analysis. Therefore, 124 participants (72 female) aged 18 to 48 years (mean 22.65 [SD 3.45]) were included in data analysis, with 30 to 32 participants in each experimental condition. The mean duration of time to complete the survey was 26.58 min (SD 13.68). In addition, 119 participants stated that they were students, and 121 participants declared German to be their first language (the other 3 participants had been speaking German for at least 17 years). Every week, participants used the computer, on average, for 27.89 hours (SD 17.53) and the internet for 34.76 hours (SD 20.83). Furthermore, participants rated their frequency (from 1: daily to 5: never) of watching videos on YouTube, in online media libraries, on streaming services, and learning platforms to be, on average, 2.96 (SD 0.60)—meaning *several times a month*, and watching videos specifically on learning platforms was, on average, 3.91 (SD 1.13)—meaning *less than once a month*. Moreover, all participants reported that their prior knowledge and motivation to learn about the topic *health and nutrition* (from 1: I disagree to 5: I agree) was, on average, 3.41 (SD 0.83); therefore, they were rather interested and cognizant of the topic.

Design

Participants were randomly assigned to 1 of the 4 cells of the 2×2 between-subject factors, which were context of Web-based communication (YouTube vs Moodle) and language style of the vlogger (YouTube-typical vs formal language), to examine the impact of both factors on information credibility, the vlogger's trustworthiness, and the perceived language accommodation of vloggers. Furthermore, participants were confronted with all manifestations of the 1×3 within-subject factor, which was uncertainty of information, meaning that they saw videos containing 2 of the following explanations: evidence for the nutrition myth (confirming), evidence against the nutrition myth (disconfirming), or unclear evidence that indicated the scientific findings neither confirm nor disconfirm the nutrition myth.

For varying the Web-based communication contexts, we manipulated the video's platform using 2 aspects that introduced where the video came from. First, to ensure a plausible comparison between Moodle and YouTube, participants read an email by a nutritional science lecturer who linked the video with a reference either to YouTube or Moodle (Multimedia Appendix 2). Accordingly, this introductory material explained that the videos are intended to supplement a nutrition course. This introduction was given to prevent participants from being suspicious, which might have been the case if they were told to watch a video on Moodle without any prior explanation. After reading the email by the nutrition science lecturer, participants

saw the video embedded in previously created screenshots according to the 2 conditions of either Moodle (an academic learning platform that allows for the exchange of material to supplement an academic course called *learnweb* for the German university where participants were recruited) or YouTube (Multimedia Appendix 3). The language used in videos was realized by a vlogger who used either YouTube-typical or formal language.

Procedure

Participants completed a Web-based survey via the tool Questback's EFS Survey [109]. In the beginning, a test page ensured that participants' devices could properly display the videos. After answering questions regarding demographic and control variables, participants were randomly assigned to the experimental conditions via a balanced randomization. According to the experimental conditions, they first read a lecturer's email with references to a video on either Moodle or YouTube (Multimedia Appendix 2). To ensure the participants paid attention to this introduction, they were able to continue after at least 25 seconds. Hereupon, participants viewed the video about nutrition myths spoken either in YouTube-typical or formal language. Both videos included 2 explanations that either confirmed or disconfirmed or (through unclear evidence) neither confirmed nor disconfirmed the nutrition myths. Once again, according to the experimental condition, videos were embedded in screenshots showing either a Moodle or YouTube surface. Continuing was allowed after at least 5 min to ensure that participants did not skip the video. After watching the video for the first time, participants rated the vlogger's trustworthiness, the vlogger's accommodation, and the control variables regarding the video's relevance in relation to the topic of nutrition. Every video sequence for each of the nutrition myths was shown a second time, which enabled participants to assess the credibility of information and control variables regarding feeling of knowing (3 items) for each of the nutrition myths individually. At the end, by answering open questions, participants were asked to assess the language of the vlogger, to remember the platform from which the video stemmed (manipulation check), and to state their generally used criteria when deciding whether to watch videos on a specific platform (explorative).

Experimental Materials

Videos included explanations of scientific findings with respect to 6 typical nutrition myths. All nutrition myths (in each video presented in this order: *coffee and dementia, cola and pretzel sticks, too many diet beverages leading to diabetes, the healthiness of low salt diet, harmfulness of too many eggs, healthy nutrition and cancer*) resulted from Web-based searches for frequent and typical nutrition myths in online forums. Although all explanations reflect the current scientific findings about related topics [110-116], explanations summarized the scientific evidence as if it speaks for, against, or neither for nor against the nutrition myths.

In the following, the explanations for the myths about coffee and dementia and light beverages and diabetes are summarized, whereby both conclusions were not clear as the scientific evidence was summarized such that it speaks neither for nor

against the myths. The first vlogger's explanation related to the nutrition myth about *coffee and dementia* that claims coffee inhibits the risk of developing dementia. The vlogger explained that coffee, in fact, has all sorts of positive and negative physical effects, and that coffee contains the substance caffeine, which like other building blocks of our DNA has a stimulating effect on the brain. The vlogger further reported that there are long-term studies showing coffee consumption reduces the risk of developing Alzheimer disease by 16%, but that it is still unclear whether only caffeine or other ingredients and factors are responsible for that. Another vlogger's explanation related to the myth of *too many diet beverages leading to diabetes*. The vlogger said that it is well known that in the long run, people are more likely to get diabetes if they often drink beverages with artificial sweeteners—compared with people who drink such drinks rarely or never. The vlogger said further that it is uncertain whether the sweeteners are responsible for the increase of diabetes, and that people instead might prefer to drink light beverages when they have a tendency to be overweight.

The explanations for the myths about cola and pretzel sticks and harmfulness of too many eggs contained conclusions that disconfirm the myths, as the scientific evidence was summarized such that it speaks against the myths. The vlogger explained underlying scientific results related to the myth of *cola and pretzel sticks*, which are said to help stop diarrhea. She explained that cola consists mostly of sugar, which leads to liquid being removed from the body. Further she explained that cola contains a lot of caffeine, which has different effects: it makes you awake, but stimulates your kidneys, which leads to more potassium loss. She concluded that even the popular mixture *cola with pretzel sticks* does not change anything because if one eats salty pretzel sticks, they only contain saline and, therefore, are not able to remedy the potassium deficiency. Similarly, the vlogger concluded that the myth regarding the *harmfulness of too many eggs* as not scientifically confirmed. According to this myth, the vlogger said that the yolk of an egg is not only quite rich in fat but also contains a lot of cholesterol. She explained that a medium-sized egg of 60 g supplies 270 mg of the fat, but there is no correlation between the risk of cardiovascular disease and the consumption of eggs in any large-scale study. Hence, she explained, if people get a lot of cholesterol in the form of eating eggs, this does not necessarily mean that it is harmful—especially because cholesterol also has positive properties. She mentioned that this is different for diabetics because there are some findings that indicate more cardiovascular disease for diabetics who ate lot of eggs.

Finally, the vlogger's explanations related to the myths of healthy nutrition and cancer and the healthiness of low salt diet concluded that the myths can be confirmed from a scientific perspective. In terms of *healthy nutrition and cancer*, the vlogger explained that there are many indications that for people who eat in a balanced way, move enough, and have a normal body weight have a lower risk of developing cancer. She further explained that cardiovascular diseases, obesity, hypertension, and sugar disease are less frequent for people who eat healthy, and there are studies showing that diet plays an essential role in preventing colorectal and breast cancer. She mentioned that for other types of cancer, for instance, esophageal cancer, there

are currently only a few indications that may indicate a connection. Moreover, she said that what is considered as healthy diet varies according to the state of knowledge, and there are no broad-brush recipes, and that the latest recommendations focus not only on what should be eaten but also on how much. Hence, she concluded that the energy balance—of what you eat and how much of it—is just as crucial as eating whole grain products, legumes, vegetables, and fruit. She said that people should only consume energy-rich foods, sugary drinks, red meat, and salty foods in moderation. In terms of *the healthiness of a low salt diet*, the vlogger explained that there is a lot of evidence showing that a low-salt diet affects blood pressure, which is important because high blood pressure will lead to heart disease and enlarged vessels in the long term. She further explained that, therefore, the World Health Organization recommends people to reduce their daily intake of common salt to about 6 g. This is supposed to reduce systolic blood pressure (the upper number in a blood pressure reading) by 5 to 6 mm and reduce the diastolic blood pressure (the lower number in a blood pressure reading) by 1 to 3 millimeters; this should have the same effect as losing weight.

To construct the YouTube-typical version, we adopted criteria developed by Mayer et al [106] and added some typical characteristics of language style which we extracted from 37 transcribed YouTube and Vimeo videos—more first-person pronouns and self-references in YouTube videos (this study's first study). Thus, we formulated a second version of the vlogger's language style for each video that replaced the personal pronouns (eg, *you*) with more formal words such as *the* and also omitted typical YouTube characteristics (eg, *I*). In the YouTube-typical language condition, texts contained more words because of personalization, so the video lengths differ by 63 seconds. Both videos were created and hence are identical regarding content, the vlogger, and design (videos were also used in another experiment by the authors [35]). In this vein, the information providers in both videos were comparable, as a psychology lecturer was acting the information provider in the same way. Furthermore, we did not inform participants about vlogger's expertise explicitly. This might be important to consider, as people also judge the trustworthiness and the credibility of information based on the information providers' expertise or other relevant characteristics [117].

To manipulate the platform of video, screenshots were produced showing either a Moodle or a YouTube surface. Therefore, frames of real YouTube videos were used, underlying titles were created, and the video about nutrition science was embedded. The Moodle screenshot was designed by creating an academic course about nutrition science where the embedded video about nutrition myths was uploaded (Multimedia Appendix 3).

Dependent Measures

Credibility of Information

Participants indicated on 1 item whether they agree with the given information, and they judged the information credibility on 5 additional items adopted by a measurement of trust in journalism [118]. Overall, these 6 items (5-point Likert scale, with 1: I strongly disagree to 5: I strongly agree) yielded internal

consistencies for each nutrition myth between Cronbach alpha=.83 and Cronbach alpha=.86.

Trustworthiness of the Vlogger

Vlogger's epistemic trustworthiness was assessed with the Muenster Epistemic Trustworthiness Inventory (METI) [16]. METI is composed of 3 subscales: *Expertise* reflects people's perception of an expert as truly knowledgeable, intelligent, and highly trained in her domain (6 items). *Integrity* reflects people's perception of an expert's good character, her values, and her as a person, which is acting in line with principles (4 items). *Benevolence* reflects people's perception of an expert's orientation toward others or society and represents whether an expert acts in accordance with the interest of others (4 items). Participants rated these items on 7-point semantic differentials (eg, 1: competent to 7: incompetent). Internal consistencies yielded Cronbach alpha=.91 for the 6 competence-related items, Cronbach alpha=.86 for the 4 integrity-related items, Cronbach alpha=.85 for the 4 benevolence-related items, and Cronbach alpha=.93 for the aggregated score of all 14 items.

Perceived Accommodation of the Vlogger's Language Style

Participants assessed how they perceived the vlogger's accommodation with an adaption of the Recipient Orientation Scale (ROS) [119]. *Audience design* reflects how people perceive the willingness of a vlogger to adapt to the audience (eg, *the person can imagine how it is to know little about this topic*). *Evaluation* reflects how people perceive the motivation of an expert to explain herself (eg, *the person cares about mediating her expertise*). *Subjective comprehension* reflects people's self-reported understanding about the topic (eg, *I understood the content*). Overall, internal consistencies yielded Cronbach alpha=.83 for the 9 items related to audience design, Cronbach alpha=.82 for the 4 items related to evaluation, Cronbach alpha=.77 for 2 items related to subjective comprehension, and Cronbach alpha=.87 for the aggregated score of all 15 items. For all items, a 5-point Likert scale from 1 (I strongly disagree) to 5 (I strongly agree) was used.

Control Variables

Participants were asked to report their prior knowledge and motivation to learn about the topic (4 items; eg, *I am familiar with the topic Health and Nutrition*) before they were randomly assigned to the experimental conditions. After watching the video 1 time, we assessed how participants perceived the relevance of videos in terms of the topic of nutrition (3 items; eg, *the video is relevant to a video about Health and Nutrition*). In addition, we assessed subjective familiarity, subjective complexity, and interest (3 items *feeling of knowing*) for each of the myths. For all items, a 5-point Likert scale from 1 (I strongly disagree) to 5 (I strongly agree) was used. Finally, a manipulation check asked *How do you evaluate the language of the person?* and *On what platform was the video?*

Analyses

We preliminarily analyzed differences between groups of our 4 conditions regarding any expected control variables. A multivariate variance analysis revealed no differences regarding the demographic variables of age, gender, and frequency of using the internet and a computer and the frequency of watching

Web-based videos, participants' prior knowledge, and their *feeling of knowing* in terms of the nutrition myths (all $F_{3,119} \leq 1.62$; $P \geq .19$; $\eta_p^2 \leq 0.04$). However, considering the significant Pearson correlations between age and trustworthiness, age and vlogger's accommodation (all $r \leq |.231|$; $P \geq .01$) between perceived relevance of the videos and all dependent variables (all from $r=|.197|$ to $r=|.533|$; $P \leq .03$) and between feeling of knowing and all dependent variables (all from $r=|.184|$ to $r=|.420|$; $P \leq .02$) these control variables should be included in the main analysis, as those variables might explain variance of the dependent variables. Pearson correlations between all other control variables and dependent variables were not significant (all from $r=|.001|$ to $r=|.165|$; $P \geq .07$). Taking these results together with the results of the multivariate variance analysis, which revealed no differences between experimental conditions regarding any control variables, none of these uncorrelated variables were included as control variables in the main analysis. Correlations for dependent and control variables for the third investigation are provided in Multimedia Appendix 4.

In addition, we analyzed whether the frequencies of participants who stated that they did not remember the video platform differed between conditions. Therefore, a chi-square test was used to compare the video platform and the manipulation check (recognized platform). Cell frequencies were between 2 and 58. Results showed a significant difference in the frequency between the video platform and the manipulation check ($\chi^2_1 = 14.1$; $P < .001$; $\phi = .337$). Of the 20 participants who stated that they did not remember the platform, 18 participants were in the condition that saw the video on the Moodle platform.

Moreover, we asked for the criteria participants used when deciding whether to watch videos on a specific platform, and we categorized their replies into aspects regarding (1) technical features of the platform, (2) quality features of the platform, (3) familiarity with the platform, (4) trust in the platform for a specific purpose, (5) others' recommendations of the platform, (6) the protagonist in the video, (7) the scientific nature of content, (8) the interest in content, (9) technical features of the video, (10) none, and (11) not specified. A chi-square test was used to compare participants' assignment to the video platform condition and the frequencies of these categories. Cell frequencies were between 2 and 43. Results showed no significant differences in frequencies between the video platform they were assigned to and these categories ($\chi^2_{10} = 6.7$; $P = .75$; $\phi = .010$). Hence, the number of participants who stated that they use criteria belonging to the above-mentioned categories was not different between the video platform conditions of Moodle and YouTube.

Setting the global alpha level of .05, we performed a multivariate variance analysis with the video platform and the language style as between-subject factors and the trustworthiness and the vloggers' accommodation as the dependent measures. Furthermore, another variance analysis with repeated measures was conducted in terms of the dependent measure credibility, again including the video platform and language style as the between-subject factors and the uncertainty of information as the within-subject factor. Myths were composited in terms of their overall conclusion and, therefore, in terms of whether the explanations and underlying scientific findings were (1) not clear (2) conflicting with, or (3) consistent with the myth. Hence, for each participant, 3 averaged values of credibility ratings were analyzed as within-subject measures because participants were asked to rate each of the myths on the same items each time they watched the video. As we conducted 2 analysis for the testing of hypotheses in 1 sample, we adjusted local alpha levels and set the local alpha level[1] of .025 and the local alpha level[2] of .05, according to the Bonferroni-Holm adjustment. All tests were 1-sided.

Results

Investigation 1

Results showed that the duration of videos from YouTube (mean 339.6 seconds [SD 134.32]) was longer than from Vimeo (mean 166.18 seconds [SD 116.33]; $F_{1,22} = 15.4$; $P = .001$). In addition, the total number of spoken words in videos from YouTube (mean 981.24 [SD 547.61]) was higher than the total number of words in videos from Vimeo (mean 320.18 [SD 175.82]; $F_{1,32} = 29.56$; $P < .001$). It is plausible that the higher word count from YouTube videos might be related to longer durations of the YouTube videos.

On average, videos from YouTube contained 50.3 (SD 59.25) first-person singular pronouns (me, I, and my) of 981.24 words. On average, videos from Vimeo contained 6.7 (SD 10.01) first-person singular pronouns of 320.18 words. The analysis of variance revealed that the relative number of first-person singular pronouns on YouTube (mean 4.4% [SD 3.94]) was higher than on Vimeo (mean 1.47% [SD 2.12]; $F_{1,32.35} = 8.17$; $P = .007$). On average, videos from YouTube contained 34.84 (SD 26.78) second-person singular pronouns (you, your, and thou) of 981.24 words; on average, videos from Vimeo contained 11.73 (SD 6.77) second-person singular pronouns of 320.18 words. The analysis of variance revealed no differences regarding the use of second-person pronouns in YouTube (mean 3.58% [SD 2.03]) and Vimeo (mean 4.41% [SD 2.95]; $F_{1,14.36} = .714$; $P = .41$).

Investigation 2

Paired-sample *t* tests revealed differences for people's trust in Web-based platforms (score of all 14 items) and for people's willingness to transact and people's familiarity with the Web-based platform; results showed lower trust in Moodle (mean 4.28 [SD 1.41]) than in YouTube (mean 5.25 [SD .84]), less willingness to transact with Moodle (mean 3.81 [SD 1.85]) than with YouTube (mean 5.91 [SD 1.02]), and less familiarity with Moodle (mean 3.82 [SD 1.81]) than with YouTube (mean 5.84 [SD .87]; all $t_{150} \leq -9.63$; all $P \leq .001$). There was no difference for the items related to trustworthiness (ability, benevolence, integrity, and overall trustworthiness) of Moodle (mean 4.74 [SD 1.23]) compared with YouTube (mean 4.63 [SD 1.18]; $t_{150} = 1.06$; $P = .29$), meaning that, overall, participants reported that they trust in YouTube more than in Moodle, although their answers regarding the items called *trustworthiness* did not differ between Moodle and YouTube. In particular, participants' familiarity and willingness to transact with

YouTube were higher compared with Moodle. A post-hoc analysis was conducted to identify whether participants' self-reported frequency of watching Web-based videos explains additional variance of the dependent variables. This analysis, indeed, reveals that participants' self-reported frequency of watching Web-based videos on YouTube (1 item) had significantly impacted participants' trust in Moodle and YouTube (both $F_{1,149} \geq 4.75$; $P \leq .03$; $\eta_p^2 \geq 0.03$). Furthermore, participants' frequency of watching Web-based videos in online media libraries (1 item) had significantly impacted participants' trust in Moodle ($F_{1,149}=10.8$; $P=.001$; $\eta_p^2=0.07$), but not participants' trust in YouTube ($F_{1,149}=.395$; $P=.53$). Finally, participants' frequency of watching Web-based videos on streaming services (1 item) had significantly impacted participants' trust in YouTube ($F_{1,149}=13.86$; $P=.001$; $\eta_p^2=0.09$), but not participants' trust in Moodle ($F_{1,149}=.353$; $P=.06$; $\eta_p^2=0.023$).

Investigation 3

Credibility of Information

The variance analysis with repeated measures yielded no between-subject main effect of video platform ($F_{1,117}=2.40$; $P=.06$; $\eta_p^2=0.02$), and language style ($F_{1,117}=.12$; $P=.37$; $\eta_p^2=0.001$), for participants' credibility judgments. Furthermore, there was no significant interaction of language style and video platform ($F_{1,117}=2.06$; $P=.08$; $\eta_p^2=0.02$). Furthermore, the analysis yielded a significant effect for uncertainty of information ($F_{2,116}=9.109$; $P<.001$; $\eta_p^2=0.136$). Post hoc group comparisons showed that explanations with unclear findings (*coffee and dementia* and *too many diet beverages leading to diabetes*) were judged to be significantly less credible than explanations conflicting with the myth (*cola and pretzel sticks* and *harmfulness of too many eggs*; $d=|1.9|$, SE 0.05; 95% CI 0.08-0.30; $P<.001$). Moreover, these explanations with unclear findings were judged to be significantly less credible than explanations confirming the myth (*healthy nutrition and cancer* and *the healthiness of low salt diet*; $d=|1.4|$, SE 0.05; 95% CI 0.25-0.03; $P=.004$). Furthermore, comparisons showed that those explanations conflicting with the myth (*cola and pretzel sticks* and *harmfulness of too many eggs*) and those confirming the myth (*healthy nutrition and cancer* and *the healthiness of low-salt diet*) did not lead to significantly different credibility ratings ($d=|0.05|$, SE 0.04, 95% CI 0.06-0.16; $P=.77$). Figure 1 shows credibility judgments for the uncertainty of information.

Trustworthiness of the Vlogger

The multivariate analysis revealed no main effects of video platform for all trustworthiness measures (all $F_{1,117} \leq .18$; $P \geq .34$; $\eta_p^2 \leq 0.002$). Furthermore, there were no main effects of language style for the METI score and the subscales competence and integrity (all $F_{1,117} \leq 0.86$; $P \geq .18$; $\eta_p^2 \leq 0.007$). However, there was a main effect of language style for the subscale benevolence ($F_{1,117}=3.41$; $P=.03$; $\eta_p^2=0.028$). That is, YouTube-typical language led to higher benevolence ratings (mean 2.63 [SD 0.93]) than formal language (mean 2.94 [SD 0.91]). Moreover, the multivariate analysis yielded no significant interactions of language style and video platform for all trustworthiness measures (all $F_{1,117} \leq 1.18$; $P \geq .14$; $\eta_p^2 \leq 0.01$).
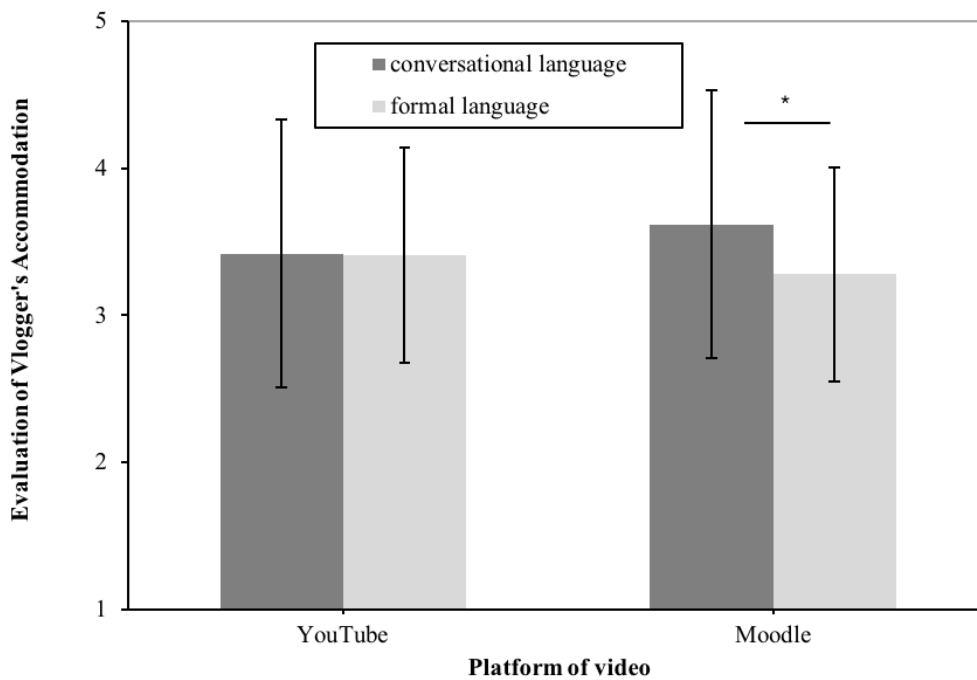
Perceived Language Accommodation by Vlogger

The multivariate analysis revealed no main effects of video platform for the perceived accommodation of vlogger's language (all $F_{1,117} \leq 1.29$; $P \geq .13$; $\eta_p^2 \leq 0.011$). However, analysis yielded a main effect of language style for the overall score of the recipients orientation scale ($F_{1,117}=8.41$; $P=.002$; $\eta_p^2=0.07$), and its subscale audience design ($F_{1,117}=10.59$; $P<.001$; $\eta_p^2=0.08$), with more ascribed vlogger's accommodation when she used YouTube-typical language compared with formal language. The analysis revealed no significant main effect of language style on the subscales evaluation and subjective comprehension (both $F_{1,117} \leq 1.31$; $P \geq .13$; $\eta_p^2 \leq 0.011$). Furthermore, analysis yielded a disordinal significant interaction effect of platform of video and language style for the subscale evaluation ($F_{1,117}=5.40$; $P=.01$; $\eta_p^2=0.04$). That is, more ascribed accommodation in Moodle for the vlogger who used YouTube-typical language compared with formal language ($d=|.43|$, SE 0.17; $P=.04$, Bonferroni post-hoc analysis). Conversely, in YouTube, the vlogger's accommodation for either formal or YouTube-typical language was not judged differently ($d=|.15|$, SE 0.18; $P=.50$). Similarly, none of the other pairwise comparisons of experimental conditions yielded significant differences of means (all $d \leq |.34|$, SE 0.18; $P \geq .17$). Figure 2 illustrates this interaction for the factors video platform and language style on the subscale evaluation.

Table 1 shows descriptive values for the factors' *platform of video* and *language style* for the dependent variables, credibility of information, trustworthiness, and vlogger's language accommodation. In addition, Table 2 shows values for the multivariate analysis of variance including the control variables age, relevance of video, and feeling of knowing and the factors' platform of video and language style for the dependent variables' trustworthiness and vlogger's language accommodation.

Figure 1. Participants' credibility judgments in terms of the uncertainty of information in the experiment (within subjects). The following covariates were included in the analysis due to the following values: age=22.65, relevance of video=3.36, feeling of knowing=3.55.
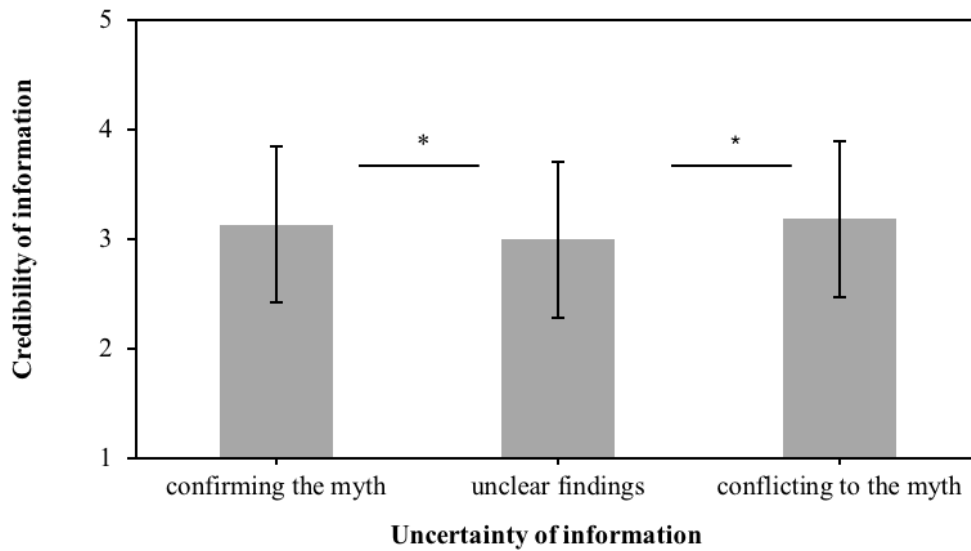
Figure 2. Participants' evaluations on the Evaluation subscale (ROS Evaluation) for the factors video platform and language style in the experiment. The following covariates were included in the analysis due to the values: age=22.65, relevance of video=3.36, feeling of knowing=3.55.

Table 1. Descriptive values for the factors' platform of video and language style of vlogger for the dependent variables trustworthiness (eg, 1: competent to 7: incompetent), credibility of information, and vlogger's language accommodation (1: I strongly disagree to 5: I strongly agree).

Dependent variables and platform of video	YouTube-typical language, mean (SD)	Formal language, mean (SD)	Total, mean (SD)
Trustworthiness of vlogger			
METI[a] score			
Moodle	2.80 (1.01)	2.77 (0.77)	2.79 (0.89)
YouTube	2.78 (0.83)	2.97 (0.79)	2.88 (0.81)
Total	2.79 (0.92)	2.87 (0.78)	2.83 (0.85)
METI competence			
Moodle	3.05 (1.21)	2.94 (1.07)	2.99 (1.13)
YouTube	3.12 (1.00)	3.01 (.98)	3.06 (0.98)
Total	3.08 (1.11)	2.97 (1.02)	3.03 (1.06)
METI integrity			
Moodle	2.59 (1.07)	2.49 (0.73)	2.54 (0.91)
YouTube	2.42 (0.91)	2.83 (0.73)	2.63 (0.84)
Total	2.51 (0.99)	2.65 (0.75)	2.58 (0.88)
METI benevolence			
Moodle	2.64 (1.02)	2.80 (0.80)	2.72 (0.91)
YouTube	2.61 (0.85)	3.08 (1.02)	2.84 (0.96)
Total	2.63 (0.93)	2.94 (0.91)	2.78 (0.93)
Vlogger's language accommodation			
ROS[b] score			
Moodle	3.84 (0.58)	3.53 (0.54)	3.69 (0.58)
YouTube	3.81 (0.50)	3.61 (0.58)	3.71 (0.55)
Total	3.83 (0.54)	3.57 (0.56)	3.70 (0.56)
ROS audience design			
Moodle	3.91 (0.61)	3.53 (0.62)	3.72 (0.64)
YouTube	3.94 (0.45)	3.67 (0.67)	3.81 (0.58)
Total	3.92 (0.54)	3.60 (0.64)	3.76 (0.61)
ROS evaluation			
Moodle	3.62 (0.94)	3.28 (0.79)	3.45 (0.88)
YouTube	3.42 (0.88)	3.41 (0.67)	3.41 (0.78)
Total	3.52 (0.91)	3.34 (0.73)	3.43 (0.83)
ROS subjective comprehension			
Moodle	3.98 (0.79)	4.02 (0.69)	4.99 (0.73)
YouTube	4.02 (0.79)	3.73 (0.76)	3.88 (0.78)
Total	4.00 (0.78)	3.88 (0.73)	3.94 (0.76)
Credibility of information			
Unclear			
Moodle	3.21 (0.83)	3.12 (0.61)	3.16 (0.73)
YouTube	3.25 (0.69)	3.17 (0.67)	3.21 (0.68)
Total	3.23 (0.76)	3.14 (0.64)	3.18 (0.70)
Conflicting			
Moodle	3.08 (0.84)	2.94 (0.64)	3.01 (0.74)

Dependent variables and platform of video	YouTube-typical language, mean (SD)	Formal language, mean (SD)	Total, mean (SD)
YouTube	2.89 (0.67)	3.07 (0.67)	2.98 (0.67)
Total	2.99 (0.76)	3.00 (0.65)	3.00 (0.71)
Confirming			
Moodle	3.07 (0.79)	3.05 (0.76)	3.06 (0.77)
YouTube	3.16 (0.64)	3.27 (0.66)	3.21 (0.65)
Total	3.11 (0.72)	3.15 (0.72)	3.13 (0.71)

[a]METI: Muenster Epistemic Trustworthiness Inventory.

[b]ROS: Recipient Orientation Scale.

Table 2. Multivariate analysis of variance including the factors' platform of video and language style of vlogger and the control variables age, relevance of video, and feeling of knowing for the dependent variables' trustworthiness and vlogger's language accommodation.

Source of variance	df	F	P value	η_p^2
Age				
METI[a] score	1	4.631	.02	0.038
METI competence	1	7.269	.004	0.058
METI integrity	1	2.281	.07	0.019
METI benevolence	1	0.616	.22	0.005
ROS[b] score	1	8.807	.002	0.07
ROS audience design	1	5.722	.009	0.047
ROS evaluation	1	10.312	.001	0.081
ROS subjective comprehension	1	0.011	.46	0
Relevance of video				
METI score	1	22.332	<.001	0.16
METI competence	1	25.078	<.001	0.177
METI integrity	1	15.618	<.001	0.118
METI benevolence	1	6.627	.001	0.054
ROS score	1	15.748	<.001	0.119
ROS audience design	1	6.216	.007	0.05
ROS evaluation	1	23.785	<.001	0.169
ROS subjective comprehension	1	1.72	.1	0.014
Feeling of knowing				
METI score	1	6.694	.001	0.054
METI competence	1	5.175	.01	0.042
METI integrity	1	5.962	.008	0.048
METI benevolence	1	3.559	.03	0.03
ROS score	1	21.529	<.001	0.155
ROS audience design	1	16.767	<.001	0.125
ROS evaluation	1	11.107	.001	0.087
ROS subjective comprehension	1	4.627	.02	0.038
Platform of video				
METI score	1	0	.5	0
METI competence	1	0.1	.38	0.001
METI integrity	1	0.003	.48	0
METI benevolence	1	0.182	.34	0.002
ROS score	1	0.583	.23	0.005
ROS audience design	1	1.289	.13	0.011
ROS evaluation	1	0.158	.35	0.001
ROS subjective comprehension	1	0.699	.2	0.006
Language style				
METI score	1	0.157	.35	0.001
METI competence	1	0.861	.17	0.007
METI integrity	1	0.721	.2	0.006
METI benevolence	1	3.412	.03	0.028

Source of variance	df	F	P value	η_p^2
ROS score	1	8.409	.002	0.067
ROS audience design	1	10.587	<.001	0.083
ROS evaluation	1	1.308	.13	0.011
ROS subjective comprehension	1	0.749	.19	0.006
Platform of video X language style				
METI score	1	0	.49	0
METI competence	1	0.83	.18	0.007
METI integrity	1	1.177	.14	0.01
METI benevolence	1	0.252	.31	0.002
ROS score	1	2.931	.05	0.024
ROS audience design	1	1.962	.08	0.016
ROS evaluation	1	5.404	.01	0.044
ROS subjective comprehension	1	0.769	.19	0.007

[a]METI: Muenster Epistemic Trustworthiness Inventory.

[b]ROS: Recipient Orientation Scale.

Discussion

Investigation 1

In summary, the findings of the content analysis suggest that it might be more typical for the most relevant health videos on YouTube to use first-person pronouns compared with the most relevant health videos on Vimeo because the relative proportion of first-person pronouns used in each investigated video about *health* and *diet* was higher on YouTube than on Vimeo. On the other hand, there were no differences in the number of second-person pronouns. These findings supplement findings that suggest specific Web-based platforms their own typical language styles [68,69]. Although both platforms do not explicitly declare rules for how people should use language (despite vulgar language considering netiquette), people seem to use different words. Depending on implicit rules and norms of YouTube and Vimeo—which might result from the specific affordances of both platforms—users of both platforms might have evolved a typical communication behavior. According to LET and CAT, a specific language style might fit the purpose of presenting videos on a particular platform, address a specific audience more appropriately, or is perceived by other users as more appropriate. It would be of further interest to investigate not only whether users use different language styles on platforms beyond those investigated here (YouTube and Vimeo) but also whether they use different language styles in videos with different content, beyond those related to health and diet, which might represent a subcontext (or a subcommunity) of each Web-based platform. Moreover, the language styles might be a different change for videos that have comparatively lower rankings than the ones investigated here, which are more highly ranked. As highly ranked videos represent videos recommended by platforms (through opaque criteria), these videos are indeed likely to represent platforms' *typical* videos but do not represent all available Web-based videos.

Investigation 2

In summary, findings indicate differences in users' general trust in investigated Web-based platforms. Participants seemed to trust in YouTube more than in Moodle to offer a Web-based video platform, as notably, participants' familiarity and willingness to transact with YouTube were higher than with Moodle. In more detail, it is of special interest that participants' trust-related evaluations in terms of the platform's expertise, benevolence, and integrity did not differ for YouTube and Moodle although participants' reported familiarity and willingness to transact differed for YouTube and Moodle. Accordingly, it seems worthwhile for further research to consider users' trust attitudes and users' trust actions toward Web-based platforms, as trust actions may not necessarily be caused by users' actual trust attitudes. Instead, users may use a Web-based platform such as YouTube because it is popular although users may not necessarily believe this platform to be more competent, benevolent, or to have more integrity than other platforms [27,120].

Investigation 3

Main Findings

Participants judged the credibility of information differently depending on whether the scientific explanations concluded that the associated nutrition myth was either confirmed or disconfirmed or neither confirmed nor disconfirmed by scientific evidence (H1). Interestingly, the credibility of information in videos was not judged depending on the language style of vloggers (H2) nor depending on the context of Web-based communication (H3). Hence, people seem not to consider the language style of vloggers or the context of communication when evaluating a video's information. Although the language style of the vlogger impacted participants' judgments of the vlogger's benevolence (H2; ie, participants judged vloggers who used YouTube-typical language to be more benevolent than vloggers who used formal language), the video's platform

XSL•FO
RenderX

individually had no impact on participants' judgments of the trustworthiness of the vlogger (H3). Although, contrary to the assumption that the appropriateness of a vlogger's language style given a certain context of communication (ie, platform of video) should lead people to rate the vlogger's credibility and trustworthiness higher, participants' trust-related evaluations of the information and the providers were not affected by an interaction between the vlogger's language style and the context of communication (H4). However, participants judged the vlogger's YouTube-typical language style to be more accommodative (ie, appropriate, as adapted toward the addressed audience) regardless of the Web-based platform on which the video was presented. In addition, they ascribed more language accommodation in Moodle for the vlogger who used YouTube-typical language compared with formal language, whereas their ascription of language accommodation did not differ for formal and YouTube-typical language in YouTube. Thus, participants were reciprocally impacted by the context of Web-based communication and the provider's language style, as they ascribed language appropriateness differently between YouTube and Moodle (H4).

To sum up, because participants did not judge the information or the vloggers on YouTube to be more reliable than on Moodle, the expected differences in participants' trust in these platforms did not carry over into their judgments on reliability of the information and the vlogger (H3). Similarly, participants did not judge the information or the vloggers to be more reliable when those vloggers used formal versus YouTube-typical language (H2). Instead, participants judged vloggers to be more benevolent and their language use to be more accommodative toward the audience if the vloggers used YouTube-typical language, which might in turn indicate that people do not expect to identify whether a vlogger is or is not a competent expert according to their use of formal language alone. Perhaps, people perceive the use of YouTube-typical language in Web-based videos as being a common choice for Web-based videos in general. That would also explain why participants did not judge information or vloggers to be more reliable if the video contained a rather appropriate (ie, YouTube-typical language on YouTube and formal language on Moodle) compared with a rather inappropriate language style (ie, YouTube-typical language on Moodle and formal language on YouTube) but instead ascribed more accommodation for the vloggers who used YouTube-typical language compared with formal language on Moodle. At the same time, because participants did not ascribe the accommodation of the vlogger's YouTube-typical and formal language use differently on YouTube, the use of YouTube-typical language is, perhaps, universal for Web-based videos. Together, these findings indicate rather complex interdependencies for how a vlogger's language use and its *appropriateness* given a certain Web-based context affects people's trust-related evaluations.

Finally, participants judged unclear explanations containing scientific evidence neither or nor against the nutrition myths (*coffee and dementia* and *too many diet beverages leading to diabetes*) to be less credible than those explanations concluding that the scientific evidence speaks either against (*cola and pretzel sticks* and *harmfulness of too many eggs*) or for (*healthy nutrition and cancer* and *the healthiness of low-salt diet*) the myth; this partially confirms the derived hypothesis (H1). Surprisingly, participants judged the credibility of explanations similarly when the explanations confirmed a myth and when they disconfirmed a myth. As people heuristically tend to make associations between how much effort they believe they put into understanding the given information and how they evaluate the information's credibility [30,32], participants might have perceived that both forms of explanations that gave clear scientific evidence (either for or against the common myth) were credible because participants did not have to put much effort into processing the information. Although it is likely that ambiguous information is rather complex and that people, therefore, may use more effort to process this information, it is also interesting to reflect on potential differences for how confirming and disconfirming information is assessed.

Limitations and Implications

In addition to the findings summarized above, in the following, the limitations of this study's third investigation will be emphasized. Furthermore, by taking into account the related research on health information seekers' trust-related evaluations, the results and limitations of this study's third investigation still emphasize the need for considering a complex interplay between single factors (ie, language style, information content, and the context of Web-based communication) when investigating seekers' tendencies to rely on Web-based information and providers.

Against the assumption that different Web-based contexts would influence people's judgments on the credibility of information and the trustworthiness of vloggers [34,70], videos were not judged differently depending on whether they were presented on YouTube or Moodle. In retrospect, there are factors lying in the manipulation of the experiment that might have influenced these findings. As videos on both platforms needed to be introduced by a lecturer to guarantee the valid and plausible comparison of a Web-based video platform and an academic Web-based platform, participants might have evaluated the video's reliability by also considering the lecturer as a gatekeeper. Thereby, the presence of the lecturer might have caused participants to perceive the content as filtered through a professional gatekeeper with high expertise [30]. Therefore, potential differences between Moodle and YouTube related to the vlogger's trustworthiness may have remained undetected because the gatekeeper may have caused the ratings to be relatively high to start out with (ceiling effect).

Furthermore, static screenshots of both platforms might have weakened the external validity, as participants were not able to interact with the platform's usual features. As such, research that assumes a platform's norms and rules will have an impact on people's trust-related evaluations of the platform's content might also need to consider both limitations of this study. Moreover, although the investigated platforms are limited to their specific features, which may represent plausible differences in platforms' communities and people's trust in Moodle and YouTube, people, in particular, may also differ in their frequency of using both platforms. Accordingly, the manipulation check showed that fewer participants remembered

the screenshots of Moodle than of YouTube. This, in turn, may indicate differences in people's information processing between both platforms; it is unclear whether participants may have unconsciously recognized the screenshots. However, further research might also focus on platforms that users are equally familiar with, so that researchers can control for potential differences in information processing when investigating whether the Web-based platforms influence how people evaluate the reliability of a platform's content.

According to users' trust in both platforms, another limitation lies in the comparison between participants of this study's second and third investigations and, therefore, in the causal explanation for the derived third hypothesis. Although plausible, it is still unclear whether participants in the third investigation would have rated their trust in Moodle and YouTube similar to those participants in the second investigation. As participants in the third investigation had to be unaware of this study's aim (ie, to investigate a potential impact of 2 Web-based platforms that would cause differences in participants' reliability judgments), they had to be recruited independently from those recruited in the second investigation. Hence, even if it is likely that participants in the second investigation were equally representative as those in the third investigation, it is presumable that participants in the third investigation were at least more conversant with the specific type of Moodle called learnweb. As participants in the third investigation were recruited using a mailing list from the same university that provides this specific form of Moodle, the presentation of a screenshot showing this specific Moodle surface may have triggered a higher perceived familiarity with (and a higher trust in) Moodle compared with an introduction and written explanation about Moodle's purpose and usage (presented within the second investigation). Accordingly, further research that is still aims to investigate if the context of Web-based communication affects information seekers' trust-related evaluations might not only focus on different platforms with specific rules and norms but should also consider people's experiences and habits in using these platforms. In this vein, different contexts of Web-based communication might not only be determined by different Web-based platforms but also within the same Web-based platform, as, for instance, subgenres within YouTube might evolve specific norms or rules [35].

In line with research on the credibility of information, participants in this investigation assessed the credibility of information depending on its uncertainty. In fact, it is interesting that the content was judged depending on whether it provided a strong conclusion about a common nutrition myth or whether it prevented a strong conclusion about the myth. Accordingly, the uncertainty of information may not only concern scientifically controversial and ambiguous findings but it may also entail the conflict between people's conceptions of well-known (ie, allegedly to be true) information and scientific evidence for or against this information. Thereby, giving strong evidence either for or against a nutrition myth was judged as being more credible than giving ambiguous scientific findings that are unable to confirm or disconfirm the myth. Although research about the uncertainty of information indicates consistent information to be more credible than conflicting information,

because of people's tendency to evaluate the conflicting information in a way that is beneficial to their existing knowledge [51,95,98,100], the conflicting information in this study may reflect conflicts with rather low relevance. Although it is likely that well-known nutrition myths lead people to believe in these myths because of their familiarity and wide-spread nature [105], they still represent information that is less relevant to people's social identity, needs, or interests [7]. Hence, participants in this study may not have been affected in a biased way to integrate the conflicting information into their existing knowledge, as busting a common nutrition myth may not be particularly relevant to a participant's identity. Hence, it seems fruitful for further investigations on people's credibility judgments to focus not only on provided information that just disconfirms or confirms existing information but also on the comparison between information that conflicts with people's beliefs in various ways (relevance, scientific evidence, myths, beliefs, etc). However, taking into account people's existing (mis)conceptions could help educate people, as they may judge information as more credible when they are given clear evidence for or against their existing (mis)conceptions; in addition, this new clear information may be easier for them to integrate into their existing knowledge structure [105].

Summarizing Discussion

Principal Findings

This study's investigations aimed to supplement research on people's evaluations of information and information providers on the Web. As it can be challenging for (lay)people to evaluate information found on the internet based on academic criteria that indicate whether information is correct, complete, and appropriate, it is important to investigate what cues people use instead of academic criteria to evaluate the reliability of information; this is important not only for understanding people's trust-related judgments but also for understanding how people process cues and make decisions. In this context, the main purpose of this study was to investigate whether the context of Web-based communication has an impact on how people evaluate the credibility of information and the information provider's trustworthiness.

In line with the idea that the context of Web-based communication—along with its rules and norms—determines users' language use and health information seekers' proper understanding of information [5,36,37], this study's first 2 investigations aimed to characterize the Web-based context of communication more concretely (RQ1). Therefore, the relationship between individuals and their expectations of a platform's rules and norms [61] was investigated by focusing on users' typical language use on YouTube and Vimeo and on users' trust in YouTube and Moodle. In line with previous research on different language styles for specific Web-based platforms [68,69], health videos on YouTube seemed to use more first-person pronouns than health videos on Vimeo. It appears plausible that affordances of Web-based platforms constitute users' typical language use as well as their expectations about how others should use language within these Web-based platforms [64]. Furthermore, participants in the second investigation trusted more in YouTube than in Moodle.

These differences between users' trust in Web-based platforms may likewise cause differences in users' trust-related evaluations for the very same Web-based information or provider given on different Web-based platforms [34,70]. Although it is challenging to identify implicit norms and rules of platforms because of a constantly changing Web-based platform environment, findings of the content analysis study and the survey data support the assumptions that users use different language styles on certain Web-based platforms, and they also have different levels of trust for different Web-based platforms.

According to LET and CAT, the context of Web-based communication not only determines users' typical language use but it may also affect health information seekers' proper understanding of information, as seekers may make judgments based on the interplay of the information content, the provider's language style, and the context of Web-based communication. Building on the first 2 investigations, in a Web-based experiment, we investigated whether health information seekers' trust-related evaluations of a Web-based video about nutrition myths were influenced not only by the content of the information but also by the providers' language style and the context of Web-based communication, where the context was the Web-based platform, either Moodle or YouTube (RQ2). Moreover, as research indicates that people's willingness to rely on information and information providers on the Web is influenced in a complex manner [30,33], we experimentally investigated any individual and reciprocal impact of the following factors: the uncertainty of the information being presented in the video, the language use of the vlogger presenting the information, and the Web-based context of communication. Hence, specific aspects of the message, the source, and the media were investigated, as they are expected to influence not only the communication but also the evaluation of communication and, in turn, the communication success [36,37].

Accordingly, participants' trust-related evaluations in the experiment study indicate rather complex interdependencies between health information seekers' evaluations of the vlogger's language style and the Web-based video platform. Hence, both the language style of the provider and the context of communication seemed to have an individual and reciprocal impact on seekers' trust-related evaluations, but not in the expected manner. That is, although participants judged the credibility of information in videos equally regardless of the platform and the vlogger's language use, for the videos presented on Moodle, participants thought that vloggers who used a YouTube-typical language were more accommodating toward the audience than vloggers who used formal language. Moreover, when participants assessed the trustworthiness of the vlogger, their judgments of her benevolence were affected by her language style; conversely, when participants assessed the credibility of information, their judgments were influenced by the uncertainty of information.

Limitations and Implications

Of note, all investigations in this study only emphasize the topic of health and nutrition. It is important to understand how people judge nutrition information found on the Web, as using the internet to seek out information about nutrition continues to increase [2,3], and people's decisions based on this information may influence people's health [6]. Future research could also focus on health information beyond explanations about common nutrition myths, as information seekers' decisions about whom and which information to rely on could be biased in many ways. In this vein, a strong bias might result if the Web-based information conflicts with seekers' motivational and emotional constitutions, such as might be the case when smokers hear health information about smoking [7]. Accordingly, information about nutrition myths might be less prone to be processed in a biased manner, as it could be less conflicting to seekers' own attitudes. As we found that people evaluate the credibility of nutrition information differently based on how uncertain it is, namely that they find nutrition information more credible when the provided scientific evidence speaks either strongly for or against the common nutrition myth rather than when the evidence is inconclusive, providing nutrition information on the Web should take into account people's existing knowledge about this information. Similarly, further research might investigate whether additional information might help people to more accurately judge provided information that relies on uncertain scientific evidence.

Considering the integrative investigation of factors, including the language style of the provider, the content of information, and the context of Web-based communication, all of this study's investigations should be treated as an approach to conceptualize the context of Web-based communication and determine its importance for seeking health information on the Web. In line with communication theories [36,37], communication and its success are influenced by the context in which it takes place. In addition, especially health information seekers risk to misinterpret the correctness of the information, if they neglect the context of Web-based communication [5]. However, conceptualizing and operationalizing the context of Web-based communication have many challenges [56], and by considering the accompanying rules and norms of Web-based platforms that are constituted by a platform's affordances and users' use of these affordances [61], the way context is operationalized in this study fails to address other aspects that might be considered context [57]. Similarly, research to date has identified that there are several aspects important for seeking health information on the Web, such as seekers' (epistemic) beliefs [30,95]. Thus, more aspects should be considered in future research in addition to the aspects investigated in this study. Furthermore, the language style of providers and the context of Web-based communication in this study were operationalized by specific aspects that do not capture all relevant aspects of language styles (eg, the use of technical jargon) and contexts of Web-based communication (eg, subgenres on YouTube). Furthermore, the Web-based environment is constantly evolving over time. Hence, seeking health information on the Web is also continually changing. In the future, seeking information on the Web might be complemented by completely new affordances of online media (eg, further developments of augmented reality). Hence, a generalization of the concrete context of Web-based communication investigated in this study is limited to some extent. Although defining context does somewhat limit the external validity, it does, however, take into account what might

be the most common aspects of evaluating Web-based health information within specific Web-based environments. In this vein, this approach might capture health information seekers' actual experiences in an even more valid way. It seems fruitful to consider additional Web-based platforms with different aspects that may constitute differences between users' typical use of Web-based platforms when investigating Web-based health information seekers' trust-related evaluations. As the definition of context refers to norms that are determined by affordances of online media and that entail users' expectations about how to use cues for the specific media, further research might particularly highlight users' expectations about (other) users' media habits. In this vein, research on expectations will face some challenges (eg, pre-expectations have most often been induced instead of assessed directly, as this would prime their actual expectancies [121]). Accordingly, it might be useful for research to focus on users who are familiar with online media, such as frequent users. Unfortunately, the participants in this study's experiment were rather infrequent users, and YouTube and Moodle were not used with the same frequency by participants. Hence, further research could investigate frequent users of Web-based platforms to identify if user expectations regarding the use of media affordances affect how users evaluate information and providers. In this context, it also seems valuable to explicitly ask people what they expect when using a typical platform.

Conclusions

Often, research on how health information seekers evaluate information providers' trustworthiness and information credibility on the Web focuses on single aspects of either the provider or the information. Much less research has addressed how these factors are affected by the context of communication. By focusing on the Web-based platform as an entity that enables one to conceptualize and operationalize the context of Web-based communication, it seems fruitful to investigate the impact of this communication context on people's reliability judgments, as the context of Web-based communication is expected not only to constitute the communication itself but also to influence the evaluation of communication and its communicative success. A future challenge, then, will be not only to specify the context of Web-based communication by identifying people's expectations and uses derived from a platform's affordances but also to investigate in a valid way the individual and reciprocal impact of individual Web-based factors on people's willingness to rely on information and information providers. As the way people process Web-based information seems to be influenced in a complex manner, understanding how people rely on nutrition information has to consider various Web-based aspects and whether they have any individual and integrative impact on people's evaluations.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Description of videos for YouTube and Vimeo video content analysis.

[PDF File (Adobe PDF File), 45KB - ijmr_v7i2e10282_app1.pdf]

Multimedia Appendix 2

Introduction to inform about where the videos came from according to the experimental conditions.

[PDF File (Adobe PDF File), 27KB - ijmr_v7i2e10282_app2.pdf]

Multimedia Appendix 3

Screenshots of experimental videos being embedded within a screenshot that showed the interface of either Moodle or YouTube.

[PDF File (Adobe PDF File), 147KB - ijmr_v7i2e10282_app3.pdf]

XSL•FO

RenderX

Multimedia Appendix 4

Correlations for the dependent and control variables (experiment).

[PDF File (Adobe PDF File), 111KB - ijmr_v7i2e10282_app4.pdf]

References

1. Rekdal OB. Academic urban legends. Soc Stud Sci 2014 Aug;44(4):638-654 [FREE Full text] [doi: 10.1177/0306312714535679] [Medline: 25272616]
2. Cash T, Desbrow B, Leveritt M, Ball L. Utilization and preference of nutrition information sources in Australia. Health Expect 2015 Dec;18(6):2288-2295 [FREE Full text] [doi: 10.1111/hex.12198] [Medline: 24798108]
3. Pollard CM, Pulker CE, Meng X, Kerr DA, Scott JA. Who uses the Internet as a source of nutrition and dietary information? An Australian population perspective. J Med Internet Res 2015 Aug 26;17(8):e209 [FREE Full text] [doi: 10.2196/jmir.4548] [Medline: 26310192]
4. Eysenbach G. The impact of the internet on cancer outcomes. CA Cancer J Clin 2003;53(6):356-371. [Medline: 15224975]
5. Eysenbach G. Credibility of Health Information and Digital Media: New Perspectives and Implications for Youth. In: Metzger MJ, Flanagin AJ, editors. Digital Media, Youth, and Credibility. Cambridge, MA: The MIT Press; 2008.
6. Chen YY, Li CM, Liang JC, Tsai CC. Health information obtained from the internet and changes in medical decision making: questionnaire development and cross-sectional survey. J Med Internet Res 2018 Feb 12;20(2):e47 [FREE Full text] [doi: 10.2196/jmir.9370] [Medline: 29434017]
7. Bromme R, Goldman SR. The public's bounded understanding of science. Educ Psychol 2014 Jun 5;49(2):59-69. [doi: 10.1080/00461520.2014.921572]
8. Eysenbach G, Powell J, Kuss O, Sa ER. Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. J Am Med Assoc 2002;287(20):2691-2700. [Medline: 12020305]
9. Lagger C, Lux M, Marques O. What makes people watch online videos: an exploratory study. Comput Entertain 2017 Apr 4;15(2):1-31. [doi: 10.1145/3034706]
10. Bauer J. Learner Perceptions of Using YouTube for Online Instruction. In: Proceedings of E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education. 2009 Presented at: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education; Oct 26, 2009; Vancouver, Canada.
11. Taslibeyaz E, Aydemir M, Karaman S. An analysis of research trends in articles on video usage in medical education. Educ Inf Technol 2016 Jan 14;22(3):873-881. [doi: 10.1007/s10639-015-9461-x]
12. Desai T, Shariff A, Dhingra V, Minhas D, Eure M, Kats M. Is content really king? An objective analysis of the public's response to medical videos on YouTube. PLoS One 2013;8(12):e82469 [FREE Full text] [doi: 10.1371/journal.pone.0082469] [Medline: 24367517]
13. Ho M, Stothers L, Lazare D, Tsang B, Macnab A. Evaluation of educational content of YouTube videos relating to neurogenic bladder and intermittent catheterization. Can Urol Assoc J 2015;9(9-10):320-354 [FREE Full text] [doi: 10.5489/cuaj.2955] [Medline: 26644803]
14. Upshur RE. Seven characteristics of medical evidence. J Eval Clin Pract 2000 May;6(2):93-97. [Medline: 10970003]
15. Bromme R, Jucks R. Discourse and expertise: The challenge of mutual understanding between experts and laypeople. In: Schober MF, Rapp DN, Britt MA, editors. The Routledge Handbook of Discourse Processes. 2nd edition. London: Routledge; 2017.
16. Hendriks F, Kienhues D, Bromme R. Measuring laypeople's trust in experts in a digital age: the Muenster Epistemic Trustworthiness Inventory (METI). PLoS One 2015;10(10):e0139309 [FREE Full text] [doi: 10.1371/journal.pone.0139309] [Medline: 26474078]
17. Stvilia B, Mon L, Yi YJ. A model for online consumer health information quality. J Am Soc Inf Sci 2009 Sep;60(9):1781-1791. [doi: 10.1002/asi.21115]
18. Li PP. A tentative typology of context for trust research and beyond. J Trust Res 2014 Oct 13;4(2):83-89. [doi: 10.1080/21515581.2014.966489]
19. Lyon F, Möllering G, Saunders M. Introduction. Researching trust: the ongoing challenge of matching objectives and methods. In: Lyon F, Möllering G, Saunders MN, editors. Handbook of Research Methods on Trust. Cheltenham: Edward Elgar Publishing; 2015:1-24.
20. Rousseau DM, Sitkin SB, Burt RS, Camerer C. Not so different after all: a cross-discipline view of trust. Acad Manage Rev 1998 Jul 1;23(3):393-404. [doi: 10.5465/Amr.1998.926617]
21. Mcknight DH, Carter M, Thatcher JB, Clay PF. Trust in a specific technology. ACM Trans Manage Inf Syst 2011 Jun 1;2(2):1-25. [doi: 10.1145/1985347.1985353]
22. Kelton K, Fleischmann K, Wallace W. Trust in digital information. J Am Soc Inf Sci 2008 Feb 1;59(3):363-374. [doi: 10.1002/asi.v59:3]
23. Hall MA, Dugan E, Zheng B, Mishra AK. Trust in physicians and medical institutions: what is it, can it be measured, and does it matter? Milbank Q 2001;79(4):613-39, v [FREE Full text] [Medline: 11789119]
24. Hardin R. Trust and Trustworthiness. New York, NY: Russell Sage Foundation; 2004.

25.   Mayer RC, Davis JH, Schoorman FD. An integrative model of organizational trust. Acad Manage Rev 1995 Jul 1;20(3):709-734. [doi: 10.5465/AMR.1995.9508080335]

26.   Friedman B, Khan PH, Howe DC. Trust online. Commun ACM 2000;43(12):34-40. [doi: 10.1145/355112.355120]

27.   Pearce WB. Trust in interpersonal communication. Commun Monogr 1974 Aug;41(3):236-244. [doi: 10.1080/03637757409375842]

28.   Fogg BJ, Soohoo C, Danielson DR, Marable L, Stanford J, Tauber ER. How do users evaluate the credibility of web sites? A study with over 2,500 participants. In: DUX '03 Proceedings of the 2003 conference on Designing for user experiences. 2003 Presented at: Proceedings of the conference on Designing for user experiences; June, 6-7, 2003; San Francisco, CA p. 1-15. [doi: 10.1145/997078.997097]

29.   Metzger MJ, Flanagin AJ, Eyal K, Lemus DR, Mccann RM. Credibility for the 21st century: integrating perspectives on source, message, and media credibility in the contemporary media environment. Ann Int Commun Assoc 2003;27(1):293-335. [doi: 10.1080/23808985.2003.11679029]

30.   Metzger MJ, Flanagin AJ. Credibility and trust of information in online environments: the use of cognitive heuristics. J Pragmat 2013 Dec;59:210-220. [doi: 10.1016/j.pragma.2013.07.012]

31.   Rieh SY, Danielson DR. Credibility: a multidisciplinary framework. Ann Rev Info Sci Tech 2008 Oct 24;41(1):307-364. [doi: 10.1002/aris.2007.1440410114]

32.   Sundar SS. The MAIN Model: A heuristic approach to understanding technology effects on credibility. In: Metzger MJ, Flanagin AJ, editors. Digital media, youth, and credibility. Cambridge: The MIT Press; 2008:72-100.

33.   Choi W, Stvilia B. Web credibility assessment: conceptualization, operationalization, variability, and models. J Assoc Inf Sci Technol 2015 May 13;66(12):2399-2414. [doi: 10.1002/asi.23543]

34.   Zimmermann M, Jucks R. How experts' use of medical technical jargon in different types of online health forums affects perceived information credibility: randomized experiment with laypersons. J Med Internet Res 2018 Jan 23;20(1):e30 [FREE Full text] [doi: 10.2196/jmir.8346] [Medline: 29362212]

35.   Zimmermann M, Jucks R. With a view to the side: YouTube's sidebar and YouTuber's linguistic style as hints for trust-related evaluations. Int J Human Comput Interact 2018:1-13. [doi: 10.1080/10447318.2018.1519165]

36.   Burgoon JK, Dillman LA, Stern L. Interpersonal Adaptation: Dyadic Interaction Patterns. Cambridge: Cambridge University Press; 2007.

37.   Giles H. Communication Accommodation Theory: Negotiating Personal Relationships And Social Identities Across Contexts. Cambridge: Cambridge University Press; 2016.

38.   Jucks R, Linnemann GA, Thon FM, Zimmermann M. Trust the words. Insights into the role of language in trust building in a digitalized world. In: Blöbaum B, editor. Trust and Communication in a Digitized World. Switzerland: Springer, Cham; 2016:225-237.

39.   Liao W, Bazarova NN, Yuan YC. Expertise judgment and communication accommodation in linguistic styles in computer-mediated and face-to-face groups. Commun Res 2016 Feb 5;45(8):1122-1145. [doi: 10.1177/0093650215626974]

40.   Thon FM, Jucks R. Regulating privacy in interpersonal online communication: the role of self-disclosure. Stud Commun Sci 2014;14(1):3-11. [doi: 10.1016/j.scoms.2014.03.012]

41.   Toma CL, D'Angelo JD. Tell-Tale words: linguistic cues used to infer the expertise of online medical advice. J Lang Soc Psychol 2014 Oct 16;34(1):25-45. [doi: 10.1177/0261927X14554484]

42.   Elsbach KD, Elofson G. How the packaging of decision explanations affects perceptions of trustworthiness. Acad Manag J 2000 Feb 1;43(1):80-89. [doi: 10.2307/1556387]

43.   Sillence E. Having faith in the online voice. Exploring contemporary issues of trust, language and advice in the context of e-health. Linguistik online 2017;87(8):107-125. [doi: 10.13092/lo.87.4175]

44.   Bol N, van Weert JC, de Haes HC, Loos EF, Smets EM. The effect of modality and narration style on recall of online health information: results from a web-based experiment. J Med Internet Res 2015 Apr 24;17(4):e104 [FREE Full text] [doi: 10.2196/jmir.4164] [Medline: 25910416]

45.   Ma X, Hancock JT, Mingjie KL, Naaman M. Self-disclosure and perceived trustworthiness of Airbnb host profiles. In: CSCW '17 Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing. 2017 Presented at: 17th ACM Conference on Computer Supported Cooperative WorkSocial Computing; - Mar 01; Portland, OR, USA; 2017 Feb 25 - Mar 01; Portland, OR. [doi: 10.1145/2998181.2998269]

46.   Quintero Johnson JM, Yilmaz G, Najarian K. Optimizing the presentation of mental health information in social media: the effects of health testimonials and platform on source perceptions, message processing, and health outcomes. Health Commun 2017 Dec;32(9):1121-1132. [doi: 10.1080/10410236.2016.1214218] [Medline: 27573748]

47.   Thomm E, Bromme R. "It should at least seem scientific!" Textual features of "scientificness" and their impact on lay assessments of online information. Sci Ed 2011 Nov 7;96(2):187-211. [doi: 10.1002/sce.20480]

48.   Thon FM, Jucks R. Believing in expertise: how authors' credentials and language use influence the credibility of online health information. Health Commun 2017 Jul;32(7):828-836. [doi: 10.1080/10410236.2016.1172296] [Medline: 27466693]

49.   Scharrer L, Stadtler M, Bromme R. You'd better ask an expert: mitigating the comprehensibility effect on laypeople's decisions about science-based knowledge claims. Appl Cognit Psychol 2014 Mar 12;28(4):465-471. [doi: 10.1002/acp.3018]

50.   Hilligoss B, Rieh SY. Developing a unifying framework of credibility assessment: construct, heuristics, and interaction in context. Inform Process Manag 2008 Jul;44(4):1467-1484. [doi: 10.1016/j.ipm.2007.10.001]

51.   Metzger MJ, Flanagin AJ, Medders RB. Social and heuristic approaches to credibility evaluation online. J Commun 2010;60(3):413-439. [doi: 10.1111/j.1460-2466.2010.01488.x]

52.   Fogg BJ. Persuasive technology: using computers to change what we think and do. San Francisco, CA: Morgan Kaufmann Publishers; 2002.

53.   Hargittai E, Fullerton L, Menchen-Trevino E, Thomas KY. Trust online: young adults' evaluation of Web content. Int J Commun 2010;4 [FREE Full text]

54.   Meola M. Chucking the checklist: a contextual approach to teaching undergraduates web-site evaluation. Portal-Libr Acad 2004;4(3):331-344. [doi: 10.1353/pla.2004.0055]

55.   Yeo G. Trust and context in cyberspace. Archives and Records 2013 Oct 2;34(2):214-234. [doi: 10.1080/23257962.2013.825207]

56.   Kelly D. Measuring online information seeking context, Part 1: background and method. J Am Soc Inf Sci Technol 2006 Nov;57(13):1729-1739. [doi: 10.1002/asi.20483]

57.   Savolainen R. Time as a context of information seeking. Libr Inform Sci Res 2006 Mar;28(1):110-127. [doi: 10.1016/j.lisr.2005.11.001]

58.   Zhang L, Peng TQ, Zhang YP, Wang XH, Zhu JJ. Content or context: which matters more in information processing on microblogging sites. Comput Human Behav 2014 Feb;31:242-249. [doi: 10.1016/j.chb.2013.10.031]

59.   Ramirez A, Wang Z. When online meets offline: an Expectancy Violations Theory perspective on modality switching. J Commun 2008 Mar;58(1):20-39. [doi: 10.1111/j.1460-2466.2007.00372.x]

60.   Eysenbach G, Diepgen TL. Towards quality management of medical information on the internet: evaluation, labelling, and filtering of information. Br Med J 1998 Nov 28;317(7171):1496-1500 [FREE Full text] [Medline: 9831581]

61.   Evans SK, Pearce K, Vitak J, Treem JW. Explicating affordances: a conceptual framework for understanding affordances in communication research. J Comput Mediat Comm 2016 Dec 26;22(1):35-52. [doi: 10.1111/jcc4.12180]

62.   Yilmaz G, Quintero Johnson JM. Tweeting facts, Facebooking lives: the influence of language use and modality on online source credibility. Commun Res Report 2016 Apr 13;33(2):137-144. [doi: 10.1080/08824096.2016.1155047]

63.   Zimmermann M. Trust in health information online. On the role of language style and context of online communication. In: Published Doctoral-Thesis. Münster, Germany: University Münster; 2018.

64.   Stanoevska-Slabeva K. Toward a community-oriented design of internet platforms. Int J Electron Commer 2002;6(3):71-95. [doi: 10.1080/10864415.2002.11044244]

65.   Nicholls SB, Rice RE. A dual-identity model of responses to deviance in online groups: integrating social identity theory and expectancy violations theory. Commun Theor 2017 Jan 11;27(3):243-268. [doi: 10.1111/comt.12113]

66.   Fichman P, Sanfilippo MR. The bad boys and girls of cyberspace: how gender and context impact perception of and reaction to trolling. Soc Sci Comput Rev 2014 Jun 3;33(2):163-180. [doi: 10.1177/0894439314533169]

67.   Chou CH, Wang YS, Tang TT. Exploring the determinants of knowledge adoption in virtual communities: a social influence perspective. Int J Inform Manag 2015 Jun;35(3):364-376. [doi: 10.1016/j.ijinfomgt.2015.02.001]

68.   Auger GA. Fostering democracy through social media: evaluating diametrically opposed nonprofit advocacy organizations' use of Facebook, Twitter, and YouTube. Public Relations Rev 2013 Nov;39(4):369-376. [doi: 10.1016/j.pubrev.2013.07.013]

69.   Maia RC, Rezende TA. Respect and disrespect in deliberation across the networked media environment: examining multiple paths of political talk. J Comput Mediat Commun 2016 Mar 22;21(2):121-139. [doi: 10.1111/jcc4.12155]

70.   Hu Y, Shyam Sundar S. Effects of online health sources on credibility and behavioral intentions. Commun Res 2009 Nov 25;37(1):105-132. [doi: 10.1177/0093650209351512]

71.   Chung CJ, Nam Y, Stefanone MA. Exploring online news credibility: the relative influence of traditional and technological Factors. J Comput Mediat Commun 2012;17(2):171-186. [doi: 10.1111/j.1083-6101.2011.01565.x]

72.   Winter S, Krämer NC. A question of credibility? Effects of source cues and recommendations on information selection on news sites and blogs. Eur J Commun 2014;39(4):435-456. [doi: 10.1515/commun-2014-0020]

73.   Bientzle M, Fissler T, Cress U, Kimmerle J. The impact of physicians' communication styles on evaluation of physicians and information processing: a randomized study with simulated video consultations on contraception with an intrauterine device. Health Expect 2017 Dec;20(5):845-851 [FREE Full text] [doi: 10.1111/hex.12521] [Medline: 27860037]

74.   Stevenson FA, Cox K, Britten N, Dundar Y. A systematic review of the research on communication between patients and health care professionals about medicines: the consequences for concordance. Health Expect 2004 Sep;7(3):235-245 [FREE Full text] [doi: 10.1111/j.1369-7625.2004.00281.x]

75.   Pollard CM, Howat PA, Pratt IS, Boushey CJ, Delp EJ, Kerr DA. Preferred tone of nutrition text messages for young adults: focus group testing. JMIR Mhealth Uhealth 2016 Jan 19;4(1):e1 [FREE Full text] [doi: 10.2196/mhealth.4764] [Medline: 26787115]

76.   Jucks R, Paus E, Bromme R. Patients' medical knowledge and health counseling: what kind of information helps to make communication patient-centered? Patient Educ Couns 2012 Aug;88(2):177-183. [doi: 10.1016/j.pec.2012.01.011] [Medline: 22365590]

XSL•FO

RenderX

77. Bientzle M, Griewatz J, Kimmerle J, Küppers J, Cress U, Lammerding-Koeppel M. Impact of scientific versus emotional wording of patient questions on doctor- patient communication in an internet forum: a randomized controlled experiment with medical students. J Med Internet Res 2015 Nov 25;17(11):e268 [FREE Full text] [doi: 10.2196/jmir.4597] [Medline: 26607233]

78. Jucks R, Bromme R. Choice of words in doctor-patient communication: an analysis of health-related internet sites. Health Commun 2007 Jun;21(3):267-277. [doi: 10.1080/10410230701307865] [Medline: 17567258]

79. Paus E, Jucks R. Common ground? How the encoding of specialist vocabulary affects peer-to-peer online discourse. Discourse Process 2012 Oct 8;49(7):565-598. [doi: 10.1080/0163853X.2012.711671]

80. Statista. 2018. Most popular online video properties worldwide in December 2015, ranked by unique viewers (in millions) URL: https://www.statista.com/statistics/256897/leading-online-video-properties-worldwide-ranked-by-unique-viewers/ [accessed 2018-09-22] [WebCite Cache ID 72cm0GmRb]

81. IAC. 2017. URL: http://iac.com/brand/vimeo [accessed 2017-12-19] [WebCite Cache ID 6vpOHrIHF]

82. YouTube. 2017. URL: https://www.youtube.com/yt/about/press/ [WebCite Cache ID 6vpPsvq78]

83. Statistic Brain. URL: https://www.statisticbrain.com/youtube-statistics/ [accessed 2017-12-19] [WebCite Cache ID 6vpR72nT0]

84. Statistica. 2017. Social media platforms accessed by internet users in the United States to watch videos as of December 2017 URL: https://www.statista.com/statistics/809938/social-media-platforms-watch-videos-internet-users-us/ [accessed 2018-09-22] [WebCite Cache ID 72cmgHkNa]

85. Statista. 2016. Leading video streaming online services ranked by reach in the United Kingdom (UK) as of August 2016 URL: https://www.statista.com/statistics/616944/leading-video-streaming-online-services-in-the-uk-by-reach/ [accessed 2018-09-22] [WebCite Cache ID 72cnSfksl]

86. Bhatttacherjee A. Individual trust in online firms: scale development and initial test. JMIS 2014 Dec 23;19(1):211-241. [doi: 10.1080/07421222.2002.11045715]

87. Biel JI, Tsiminaki V, Dines J, Gatia-Perez D. Hi YouTube! Personality Impressions and Verbal Content in Social Video. 2013 Presented at: Proceedings of the 15th ACM on International conference on multimodal interaction; 2013; Sydney, Australia p. 405-412. [doi: 10.1145/2522848.2522894]

88. Lucassen T, Schraagen JM. Propensity to trust and the influence of source and medium cues in credibility evaluation. J Inf Sci 2012 Oct 5;38(6):566-577. [doi: 10.1177/0165551512459921]

89. Bart Y, Shankar V, Sultan F, Urban GL. Are the drivers and role of online trust the same for all web sites and consumers? A large scale exploratory empirical study. J Mark 2005 Oct;69(4):133-152. [doi: 10.1509/jmkg.2005.69.4.133]

90. Chang SE, Liu AY, Shen WC. User trust in social networking services: a comparison of Facebook and LinkedIn. Comput Human Behav 2017 Apr;69:207-217. [doi: 10.1016/j.chb.2016.12.013]

91. Mou J, Shin DH, Cohen JF. Trust and risk in consumer acceptance of e-services. Electron Commer Res 2015 Nov 5;17(2):255-288. [doi: 10.1007/s10660-015-9205-4]

92. Mou J, Shin DH, Cohen JF. Tracing college students' acceptance of online health services. Int J Human–Comput Interact 2016 Oct 7;33(5):371-384. [doi: 10.1080/10447318.2016.1244941]

93. Han PK, Klein WM, Arora NK. Varieties of uncertainty in health care: a conceptual taxonomy. Med Decis Making 2011;31(6):828-838 [FREE Full text] [doi: 10.1177/0272989X10393976] [Medline: 22067431]

94. Fleisher L, Bass S, Ruzek SB, McKeown-Conn N. Relationships among Internet health information use, patient behavior and self efficacy in newly diagnosed cancer patients who contact the National Cancer Institute's NCI Atlantic Region Cancer Information Service (CIS). Proc AMIA Symp 2002:260-264 [FREE Full text] [Medline: 12463827]

95. Kienhues D, Stadtler M, Bromme R. Dealing with conflicting or consistent medical information on the web: when expert information breeds laypersons' doubts about experts. Learn Instr 2011 Apr;21(2):193-204. [doi: 10.1016/j.learninstruc.2010.02.004]

96. Smithson M. Conflict aversion: preference for ambiguity vs conflict in sources and evidence. Organ Behav Hum Decis Process 1999 Sep;79(3):179-198. [doi: 10.1006/obhd.1999.2844] [Medline: 10471360]

97. Flemming D, Feinkohl I, Cress U, Kimmerle J. User comments about research findings: how conflictual information in online science journalistic articles influences laypeople?s understanding of scientific tentativeness. Communications 2017;42(4):465-480. [doi: 10.1515/commun-2017-0037]

98. Regan Á, McConnon Á, Kuttschreuter M, Rutsaert P, Shan L, Pieniak Z, et al. The impact of communicating conflicting risk and benefit messages: an experimental study on red meat information. Food Qual Prefer 2014 Dec;38:107-114. [doi: 10.1016/j.foodqual.2014.05.019]

99. Rogers ZF, Gould SJ. How do you know that? The epistemology of consumer health decision making under conditions of risk-benefit conflict. Psychol Mark 2015 Mar 10;32(4):450-466. [doi: 10.1002/mar.20791]

100. Elstad E, Carpenter DM, Devellis RF, Blalock SJ. Patient decision making in the face of conflicting medication information. Int J Qual Stud Health Well-being 2012;7:1-11 [FREE Full text] [Medline: 22943889]

101. Fischer AR, Frewer LJ. Consumer familiarity with foods and the perception of risks and benefits. Food Qual Prefer 2009 Dec;20(8):576-585. [doi: 10.1016/j.foodqual.2009.06.008]

102. Carpenter DM, Geryk LL, Chen AT, Nagler RH, Dieckmann NF, Han PK. Conflicting health information: a critical research need. Health Expect 2016 Dec;19(6):1173-1182 [FREE Full text] [doi: 10.1111/hex.12438] [Medline: 26709206]

103. Dean M, Shepherd R. Effects of information from sources in conflict and in consensus on perceptions of genetically modified food. Food Qual Prefer 2007 Mar;18(2):460-469. [doi: 10.1016/j.foodqual.2006.05.004]

104. Flynn DJ, Nyhan B, Reifler J. The nature and origins of misconceptions: understanding false and unsupported beliefs about politics. Adv Polit Psychol 2017 Jan 26;38:127-150. [doi: 10.1111/pops.12394]

105. Lewandowsky S, Ecker UK, Seifert CM, Schwarz N, Cook J. Misinformation and its correction: continued influence and successful debiasing. Psychol Sci Public Interest 2012 Dec;13(3):106-131. [doi: 10.1177/1529100612451018] [Medline: 26173286]

106. Mayer RE, Fennell S, Farmer L, Campbell J. A personalization effect in multimedia learning: students learn better when words are in YouTube-typical style rather than formal style. J Educ Psychol 2004;96(2):389-395. [doi: 10.1037/0022-0663.96.2.389]

107. Gill P, Arlitt M, Zongpeng L, Mahanti A. Youtube traffic characterization: a view from the edge. YouTube traffic characterization; 2007 Presented at: Proceedings of the 7th ACM SIGCOMM conference on Internet measurement; Oct 23-26, 2007; San Diego, CA p. 15-28. [doi: 10.1145/1298306.1298310]

108. Pennebaker JW, Booth RJ, Francis ME. Linguistic Inquiry and Word Count: LIWC Computer software. Austin, TX: LIWC; 2007. URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.600.7227&rep=rep1&type=pdf

109. Unipark. Survey-Software URL: https://www.unipark.com/en/survey-software/ [accessed 2018-11-13] [WebCite Cache ID 73uaeGnmE]

110. de Groot H, Farhadi J. Ernährungswissenschaft (Nutrition Science). Haan-Gruiten: Europa-Lehrmittel; 2015.

111. Eckel RH. Eggs and beyond: is dietary cholesterol no longer important? Am J Clin Nutr 2015 Aug;102(2):235-236. [doi: 10.3945/ajcn.115.116905] [Medline: 26178728]

112. Eskelinen MH, Kivipelto M. Caffeine as a protective factor in dementia and Alzheimer's disease. J Alzheimers Dis 2010;20(Suppl 1):167-174. [doi: 10.3233/JAD-2010-1404] [Medline: 20182054]

113. Greenwood DC, Threapleton DE, Evans CE, Cleghorn CL, Nykjaer C, Woodhead C, et al. Association between sugar-sweetened and artificially sweetened soft drinks and type 2 diabetes: systematic review and dose-response meta-analysis of prospective studies. Br J Nutr 2014 Sep 14;112(5):725-734. [doi: 10.1017/S0007114514001329] [Medline: 24932880]

114. Key TJ, Schatzkin A, Willett WC, Allen NE, Spencer EA, Travis RC. Diet, nutrition and the prevention of cancer. Public Health Nutr 2004;7(1a):187-200. [doi: 10.1079/PHN2003588]

115. Shin JY, Xun P, Nakamura Y, He K. Egg consumption in relation to risk of cardiovascular disease and diabetes: a systematic review and meta-analysis. Am J Clin Nutr 2013 Jul;98(1):146-159 [FREE Full text] [doi: 10.3945/ajcn.112.051318] [Medline: 23676423]

116. World health organization. Guideline: Sodium Intake for Adults and Children. In: WHO Guidelines Approved by the Guidelines Review Committee. Geneva: World Health Organization; 2012.

117. Yang Q, Beatty M. A meta-analytic review of health information credibility: belief in physicians or belief in peers? Health Inf Manag 2016 Aug;45(2):80-89. [doi: 10.1177/1833358316639432] [Medline: 27105476]

118. Matthes J, Kohring M. Operationalisierung von Vertrauen in Journalismus. M&K 2003;51(1):5-23. [doi: 10.5771/1615-634x-2003-1-5]

119. Bromme R, Jucks R, Runde A. Barriers and biases in computer-mediated expert-layperson-communication. In: Bromme R, Hesse FW, Spada H, editors. Barriers and biases in computer-mediated knowledge communication. New York, NY: Springer US; 2005:98-118.

120. Gefen D. E-commerce: the role of familiarity and trust. Omega 2000 Dec;28(6):725-737. [doi: 10.1016/S0305-0483(00)00021-9]

121. Burgoon JK, Le Poire BA. Effects of communication expectancies, actual communication, and expectancy disconfirmation on evaluations of communicators and their communication behavior. Human Commun Res 1993 Sep;20(1):67-96. [doi: 10.1111/j.1468-2958.1993.tb00316.x]

Abbreviations

CAT: communication accommodation theory
LET: language expectancy theory
LIWC: Linguistic Inquiry and Word Count
METI: Muenster Epistemic Trustworthiness Inventory
ROS: Recipient Orientation Scale
RQ: research question

XSL•FO

RenderX

Original Paper

Arthritis-Related Support in a Social Media Group for Quilting Hobbyists: Qualitative Study

Norina Gasteiger[1], BHealthSc (Hons); Rebecca Grainger[2], PhD; Karen Day[1], PhD

[1]Health Systems, School of Population Health, The University of Auckland, Auckland, New Zealand
[2]Department of Medicine, University of Otago, Wellington, New Zealand

Corresponding Author:
Karen Day, PhD
Health Systems
School of Population Health
The University of Auckland
Private Bag 92019
Auckland, 1142
New Zealand
Phone: 64 0278201125
Email: k.day@auckland.ac.nz

Abstract

Background: People with arthritis are increasingly seeking support online, particularly for information about social role participation while experiencing symptoms of chronic arthritis. Social media enables peer-to-peer support on how serious leisure (eg, hobbies such as quilting) can be adapted to allow participation. Research is needed to understand what type of peer support is provided online and how this support occurs.

Objective: The aim of our study was to explore what kind of support is offered by fellow hobbyists (with or without arthritis) in response to requests for advice in a social media group.

Methods: Three vignettes were posted on a Facebook quilting group regarding arthritis-related symptoms or impairments that affect how people quilt. A Facebook Insights report was used to examine the groups' demographics. Responses to the vignettes were thematically analyzed.

Results: The members of the quilting Facebook group were mostly women (18,376/18,478, 99.45%), aged 55 to 64 years, and most were located in the United States. In response to the vignettes, the 22 participants predominantly offered emotional support and shared information. Participants shared their real-life experiences and creative means in adapting medical advice to their crafting. More than half (30/54, 56%) of the advice that was offered aligned with the OrthoInfo medical best practice guidelines relevant to the vignettes.

Conclusions: Serious leisure social media groups can be useful forums for sharing information about arthritis-related issues. People do respond to requests for support and information, although there is a difference between quilting support (eg, "I need a new iron, what should I buy?") and health support (eg, "I have arthritis, what scissors should I buy?"). People provide emotional support for life events on serious leisure social media platforms (eg, offering condolences when a person states that she is making a memory quilt), and this extends to health issues when group members reveal them.

Introduction

Background

Quilters take their craft very seriously. Serious leisure is defined by Stebbins [1] as a "systematic pursuit of an amateur, hobbyist, or volunteer activity sufficiently substantial, interesting, and fulfilling for the participant to find a (leisure) career there acquiring and expressing a combination of its special skills, knowledge, and experience." Quilting is more than a sewing craft. It is the coming together of like-minded crafters to build a community of people who reflect culture, remember history, create art, and build quilted products demonstrating a range of

skills from novice to expert. Quilters are hobbyists who find benefit and pleasure in what they do and are committed to the development of knowledge, skills, and experience associated to enable participation in their hobby either at a serious or casual level [2]. Similarly, people living with long-term health issues acquire extensive knowledge about their condition, solve problems, and become an expert at handling their health [3].

Arthritis and painful musculoskeletal conditions are common long-term health issues that are a leading cause of disability worldwide [4,5]. Inflammatory arthritis (such as rheumatoid arthritis and the spondyloarthropathies) and osteoarthritis are leading causes of pain and can impair participation [6]. Arthritis often occurs in a person's most productive age range in adulthood and affects function and participation in everyday activities and hobbies [7,8]. Evidence-based medicine has resulted in diagnostic and treatment guidelines to enable knowledge translation from research to practice to improve health outcomes. From a clinician's perspective, adherence to evidence-based medicine is fraught [9,10] with adherence issues such as unpleasant side effects of medications. Medication adherence for rheumatoid arthritis, for example, can be as low as 30%, and interventions have been designed to improve adherence [11]. Translation of guidelines and medical advice into practice by patients encounters difficulties due to inequities (education and socioeconomic), low health literacy, and compromised cognitive abilities due to the illness itself and side effects of the medication [9].

Social media provides opportunities for people with health issues to elicit and share support and advice to adapt their activities to accommodate their condition. People use a variety of platforms (eg, Facebook, Twitter, and YouTube) to find and exchange information about health issues, elicit support, and improve self-management of their health issues [12]. People also join disease-specific groups provided by clinicians or patients to learn about their health issues and elicit and/or provide support [13]. Many join communities such as PatientsLikeMe, where individuals with similar conditions can share, view, and discuss their data and health issues, providing peer-to-peer support [14]. The literature does not explore serious leisure platforms (eg, quilting groups) where people may share and exchange information and support about long-term conditions as part of their participation in their leisure activity.

Objective

Our research question was as follows: What kind of support is offered by fellow hobbyists (with or without arthritis) in response to requests for advice in a social media group?

Methods

Study Setting

A quilting Facebook group was chosen as the research site because older women are predominantly affected by arthritis. Quilters are often older women and increasingly use Facebook groups for quilting advice [15]. The Quilting in America 2017 survey estimates there are 7 to 10 million quilters in the United States, most are female, have a mean age of 63 years, and are educated (70% attended college or university) and affluent [15]. Compared with the 2014 survey, time spent online in quilting-related sites has increased from 2.5 to 7.9 hours weekly and Facebook users from 14% to 50% [15]. Textile art has been shown to enhance well-being and quality of life and can have therapeutic benefits [16,17]. Given the widespread engagement with quilting-related social media sites and possible health benefits, a quilting Facebook group is appropriate for this research.

Recruiting Participants

To identify a group for our study, we reviewed 100 quilting groups on Facebook, resulting from a search using "quilt" as a keyword. A total of 10 groups were false positive results (eg, blanket knitting) and were excluded. Groups with 15,000 members or less were excluded from consideration for our study, as we were looking for a large enough group to elicit conversations with a likelihood of at least 10 responses to each of our scenarios (as per Table 1).

A general quilting group was deemed most appropriate because of the broad mix of skills (that was not evident in many of the specialty groups, eg, rag quilting) and range of experience (from beginner to expert). This left 6 qualifying groups to approach. The groups were approached in order of membership size. The administrators of the first 2 groups considered the researchers' invitation but declined to participate. It was decided that a smaller group among the specialty groups might be more flexible. Group sizes ranged from 563 to 150,000 members. A relatively small group of 18,478 members, dynamic in its conversation (>10 posts per day), with a growing membership, and that had previously posted arthritis-related questions, was deemed appropriate. It was a closed group, did not allow selling or self-promotion, and its function was primarily as a quilting support group. The nature of the specialty was broad, leveraging general quilting skills, for example, piecing and long-arm or free form quilting plus specialty skills.

The main administrator was approached by a private message and, after consultation with her group members, granted permission for the research. The administrator posted a comment indicating that KD had approached her to conduct research in the group saying, "Before allowing this in the group, I wanted to check with all of you. Please comment on this post and tell me if you're OK with this or not." A total of 78 people responded to it. Most group members were supportive in their comments (61), and 28 people clicked the *like* button for the initial post (we did not match these *likes* with the comments). A total of 4 posters indicated that they had arthritis-related health issues that affected their ability to quilt. They later commented in response to the vignettes. Most of the *likes* in the discussion were by KD and the group administrator to acknowledge comments. Some gave reasons for their support (they have arthritis, could learn something, and were happy to help). After 19 posts, new posters started indicating *no* with comments about medical research not being appropriate for a quilting group, and others wanted to avoid discussion about their own health issues when doing their craft. Thereafter, the conversation became mixed.

Table 1. Types of quilting groups reviewed for recruitment.

Group type	Large groups[a] (N=89), n (%)	Small groups[b] (N=64), n (%)
Buying and selling	12 (14)	9 (14)
Beginners	4 (5)	2 (3)
On the basis of a specific person's work	6 (7)	4 (6)
Specialty	45 (51)	36 (56)
General quilting	22 (25)	13 (20)[c]

[a]Greater than 15,000 members.

[b]15,000 members or less.

[c]An additional 3 were limited to one country, and one was limited to a research project.

A total of 3 members conversed with those who said *no*, defending the research, for example, the members could scroll over what was not of interest. Moreover, 2 people reiterated the purpose of the group as a support group for quilting and that other topics were not to be discussed. Near the end of the conversation, 1 person suggested that the group administrator make her own decision and this was supported by 2 others. Others offered alternative solutions, for example, researcher sets up her own Facebook page, whereas others were concerned that if permission was granted, it would be a *slippery slope* or would *take away from sharing and quilting*. The administrator made her decision after 3 days and consented to the research.

The University of Auckland Human Participants Ethics Committee approved this study on September 12, 2017 (reference 019783). The administrator of the Facebook group signed a consent form on behalf of the group members. Group members were informed that they could participate in the research by responding to scenarios. An explanation was provided with each post informing members that responses were being collected for research. A link to a summary of the study and full participant information sheets was also included in each post [18]. One researcher was a member of the group (KD) and collected the data, which were sent to the other researchers for analysis and reflection (NG and RG). The discussion threads have since been deleted to ensure privacy of the participants, as quotes used in this study could lead to identification of research participants by new or existing members of the Facebook group.

Data Collection

Overall, 3 vignettes (case scenarios) were posted as threads in the Facebook group, to which members responded (Figure 1).

Figure 1. Screenshot of the vignettes as posted in the quilting group.
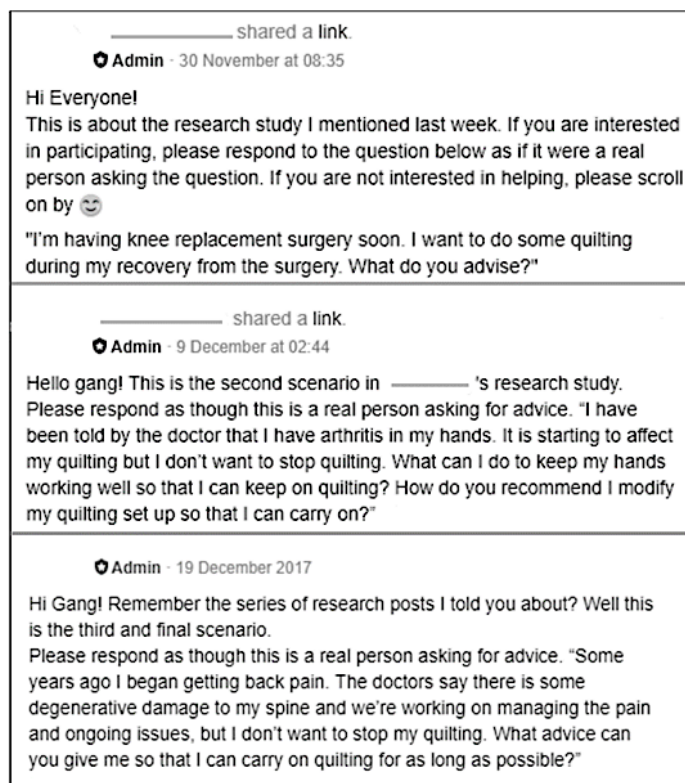
The rules of the quilting group stated the following: no selling or self-promotion is allowed, no politics, respect copyright, and be kind and supportive. The researchers designed vignettes to respect the primary role of the group, that is, quilting sharing and support. Vignettes were appropriate, as they enable sensitive qualitative data collection and are an effective tool for eliciting judgments and perceptions [19]. The vignettes alluded to arthritis-related issues that affect how people quilt and included a person undergoing knee joint arthroplasty (replacement) for knee osteoarthritis, a person with hand pain from hand osteoarthritis, and a person with chronic nonspecific lower back pain (Figure 1).

The vignettes were developed by a nurse (KD) and rheumatologist (RG). After no responses to the first vignette, it was removed, shortened, and simplified (response instructions were replaced by a link to KD's blog). The short version encouraged discussion. Vignettes were left on the Facebook page (November 30, 2017 and December 9 and 19, 2017) for approximately 1 week until no new responses were made. KD posted a summary of each discussion a week after the initial post to allow for further commentary or correction, which did not occur.

Data Analysis

Before the vignettes were posted, the administrator of the Facebook group extracted demographic data from the group using the *Page Insights* function. The resulting Excel spreadsheet was sent to the researchers for quantitative analysis of demographic characteristics, including summary statistics of age and gender. A map of the distribution of group members internationally (Figure 2) was created by NG using ArcGIS (Environmental Systems Research Institute, 2017, ArcGIS Desktop Help 10.5.1) [20].

Responses to each thread were screenshot, cut, and pasted and then transcribed onto an MS Word document for analysis and
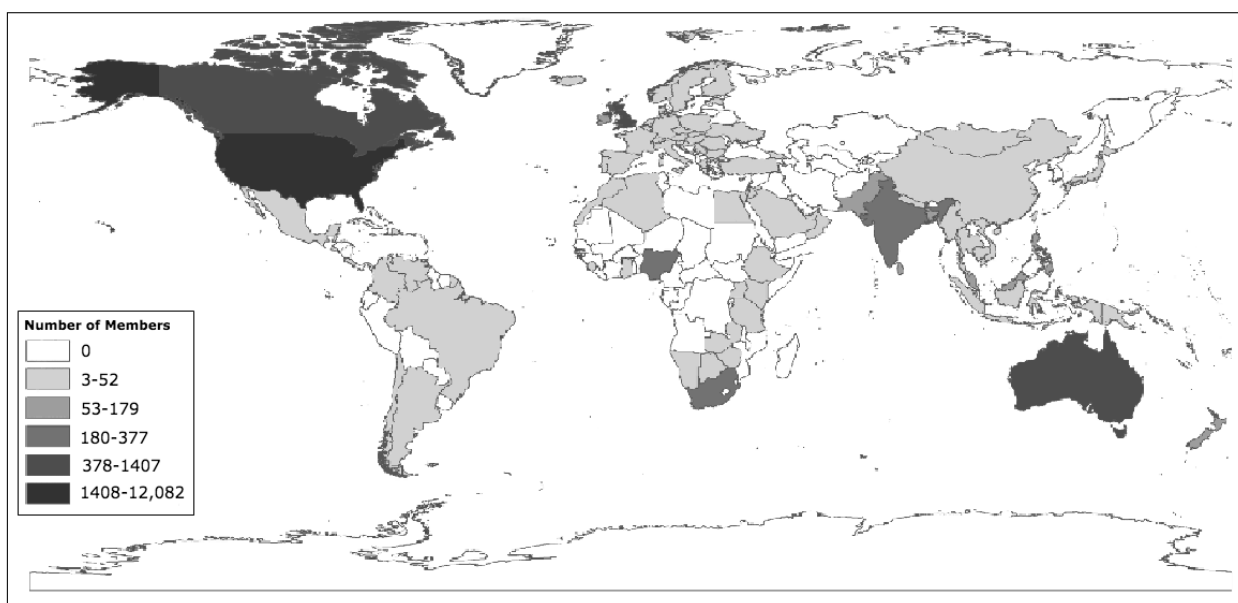
reflection using Williams' reflective diary guide [21] (see Textbox 1) to conduct an inductive thematic analysis [22]. The discussion threads and reflections were analyzed thematically. This consisted of coding the qualitative data and identifying themes that maintain the richness of the data and accurately represent and explain the phenomena [23]. Interactions between themes were examined and their relevance to the research aim was explained. All 3 researchers reflected on the data and made comments on the analysis.

The discussion thread data were compared with medical best practice for each vignette using the following steps:

- The peer advice from the posted comments was coded (eg, exercise or medication use) and grouped by codes in a table.
- Information for patients about recommended management and self-care from OrthoInfo [24] was summarized by RG (a rheumatologist) in MS Word. OrthoInfo is a website for lay people, developed and reviewed by members of the American Academy of Orthopaedic Surgeons (AAOS) and provides evidence-based information about treatment of, and self-care for, musculoskeletal conditions [24]. OrthoInfo was chosen because the AAOS is the world's largest professional association for musculoskeletal specialists, and the information has undergone peer review by an expert panel of 13-member editorial board [25].
- The self-care management options (ie, recommendations that could be implemented by an individual) proposed by OrthoInfo were identified from the summary.
- Coded peer advice that was congruent with self-care recommendations from OrthoInfo were matched in Table 2.
- Summary statistics about frequency of peer advice congruent (or not) with OrthoInfo recommendations were prepared.

Figure 2. Location of quilting closed Facebook group members around the world.

Textbox 1. Reflection tool questions.

At the end of the period covered by each vignette, note down and reflect on:

- The most interesting issue

- What worked well (eg, significant achievements)

- What did not work well

- The most puzzling or confusing issue

- The most unexpected issue

- Any risks and threats to the project

- Any opportunities for the project

- Any implications for the principles and purpose underpinning the project

- Differences between the plan and the action

- What was noticed but not addressed? Why? Would addressing that have improved the outcome?

- Any other comments or observations

Table 2. Comments and replies for each vignette.

Vignettes	Participants (n=22), n (%)	Comments (n=27), n (%)	Replies (n=11), n (%)
1	8 (36)	8 (30)	4 (36)
2	5 (23)	8 (30)	5 (46)
3	11 (50)	11 (40)	2 (18)

Results

The Context

Most quilting groups were international (of the 100 groups that we analyzed, only 3 were limited to 1 country). The Facebook group had 18,478 members when the first vignette was posted. Facebook Page Insights showed that the country with the most members was the United States of America, followed by Canada, the United Kingdom, and Australia. Members were spread all over the world, with high numbers in India, China, Europe, South Africa, and Nigeria and lowest density of members in South America, China, Europe, and parts of Africa (Figure 2). Thus, the group consisted of members from different countries (developed and developing) although membership was predominantly from the United States.

Almost all of the members from the quilting Facebook group identified as females (18,376/18,478, 99.44%; Figure 3), with a very small proportion identifying as males (92/18,478, 0.49%) or custom (10/18,478, 0.05%; Figure 4). The biggest proportion of female members (6433/18,478, 34.81%) was aged 55 to 64 years and for male members (21.6%) was 45 to 54 years. Those with custom genders were distributed more evenly across the age categories. None of the members who identified as male or custom were in the youngest age group, that is, 13 to 17 years.

Of the top 100 contributors to the quilting group, the posts ranged from 0 to 10 (87 members), 11 to 20 (6 members), and more than 20 (3 members). The top commenters (responses to posts) ranged from 13 to 99 (69 members), 100 to 199 (17), 200 to 299 (7), and 300 to 721 (4). Furthermore, 2 of the respondents

to our vignettes were among the top 100 contributors. While 1 person posted in vignette 1 (V1) and vignette 2 (V2), had 6 posts, and 518 comments before the research started, the other person posted in vignette 3 (V3), had no posts, and 270 comments before the research started. We were unable to reliably collect demographic data from the Facebook pages of individual participants because of most privacy settings blocking access to this kind of information. The Facebook Insights report did not supply demographics.

Discussions Relating to the Vignettes

A total of 22 people participated in the study by commenting or replying to comments in the vignettes (Table 1). There were 38 contributions (comments and replies) in total across the vignette discussions, with 71% (27/38) original posts and 29% (11/38) comments or responses. Participation was similar in all 3 case scenarios (V1, V2, and V3, n=14 each). Participants contributed an average of 1.7 times each. Each vignette was only posted once.

For 14 responses of V1, 2 participants responded to a comment by another participant or the researcher. There were 1 to 3 likes per comment—most of them were from the researcher and the group administrator to acknowledge the comments and encourage more comments. KD asked 8 questions in response to posts and received 1 reply. One participant also contributed to the consent conversation and was one of the top 100 contributors to the group. Of the 13 comments for V2, 1 participant posted 3 comments and responded to the researcher's question with 3 more comments. She also responded twice to comments by others.

Figure 3. Number of female members by age in quilting closed Facebook group.
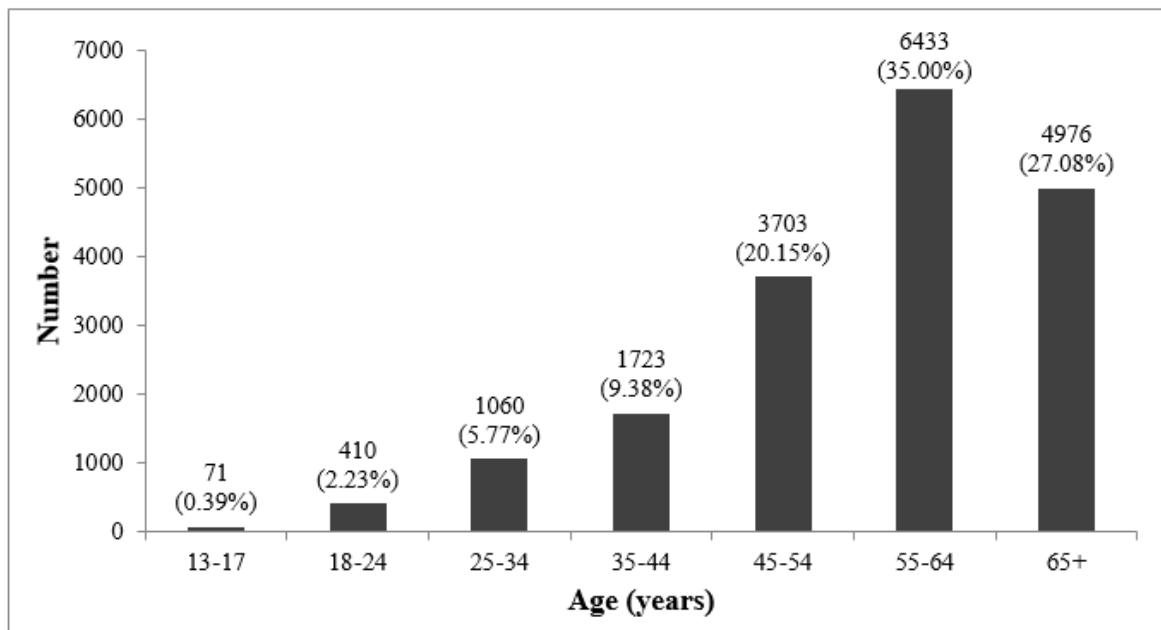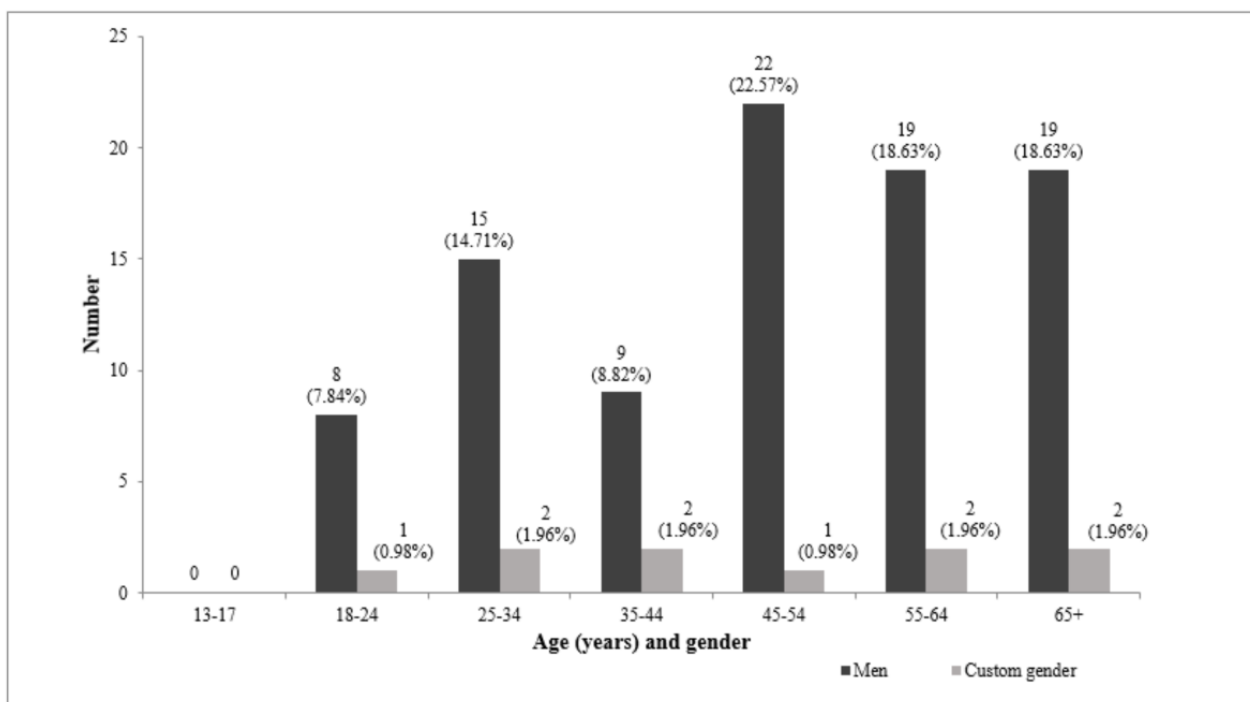

Figure 4. Number of male and custom (gender diverse) members by age in quilting closed Facebook group.

The *likes* were mostly given by the administrator but 2 were given by 1 of the participants plus someone who did not post a comment. Again, 1 participant also contributed to the consultation conversation and was 1 of the top 100 group contributors. Of 13 comments of V3, 1 was a question by the researcher to stimulate discussion (for which there was 1 reply by someone who had already posted) and 4 were responses to comments. No one posted more than once. The *likes* as for the other vignettes were given by the researcher and the administrator to encourage more participation. KD responded to 2 comments by asking questions, but there were no replies. In addition, 2 participants had contributed to the consultation

conversation and were also in the top 100 contributors. A total of 2 participants responded to both vignettes 1 and 2. Of those who participated in the consultation conversation, 2 responded to V1, and 1 responded to V3. No video links were shared. In addition, 2 website links were provided—1 in V2 for examples of spring-loaded scissors and 1 in V3 to a posture guide for computer use.

Emerging Themes

Emotional Support

Members offered emotional support in V1 and V2 through encouragement and reassurance. Members responding in V1

made encouraging comments such as wishing the hypothetical person good luck for their knee replacement surgery. Some members were encouraging about undergoing therapy by emphasizing that this was crucial to good recovery. Another member justified that she gave advice to encourage a person to continue quilting. She later reinforced the value of offering support:

> *I hope this advice encourages your "pretend" person to continue quilting...My pleasure if it helps someone else.*

Reassuring comments in V1 supported the decision for having knee replacement surgery and emphasized that recovery is only temporary and short term. Some members referenced activities that they were able to do again within a short time frame post surgery:

> *Was back at sewing machine for short periods 3 weeks post op.*

Comments that encouraged positive mindset were made in response to V1. These related to participants staying positive and hopeful throughout recovery and reassured that they were content during this time:

> *Keep cheerful, and remember you'll soon be able to continue normal activities...look forward to quilt shows for more ideas for even more quilts to make.*

Information Sharing

Participants shared a range of information on arthritis-related topics. Advice offered in V1 was specific to recovery, whereas that given in V2 and V3 was related to everyday life with arthritis. This included medical information, adapting quilting to their condition, and using assistive tools. Experience was mentioned to support a recommendation or to acknowledge a lack of knowledge. Those who had lived experience provided more in-depth and personal information and provided external links to information as well.

Medical information across the cases included additional treatments such as massage, exercises such as yoga or Pilates, continuing with quilting, dance, correct posture, pain and symptom management through the use of arthritis gloves, anti-inflammatory gel, heated massage pillows, and resting. Some recommendations, such as this advice from V3, were couched in humor:

> *Get up and boogie. LoL. Or limbering yoga moves.*

Members commenting in V1 emphasized the importance of trusting medical professionals and following instructions. They specifically recommended following medical instructions, asking questions, and attending physiotherapy appointments:

> *I fully followed all medical instructions, and had asked enough questions to know what to expect.*

Members commenting in V1 also valued preparation, as they assumed that exercise would be detrimental following a knee replacement surgery. They suggested resting, recovering, and staying off the stairs. As some of their craft rooms were downstairs, they recommended being prepared with kits,

handwork, and smaller quilts to work on. They also suggested having someone retrieve their gear when necessary:

> *If I needed something from my craft area I got someone else to fetch it for me until I was allowed on the stairs again. I suppose the best advice I could give would be to resolve yourself to limited activity for a short period of time and plan toward that end.*

Members' comments in all 3 cases emphasized the importance of taking breaks. While comments in V1 referred to prioritizing recovery, those in V2 related to resting during flare-ups until they passed and taking breaks to stretch hands. Comments by members in V3 acknowledged that they were unable to quilt on some days due to back pain and would watch people quilt instead:

> *Choose your battles. Some days I can deal with it and some days I can't. When I can't, I watch other people sew on YouTube :D.*

Responses to V3 also revealed the importance of posture to minimize back pain. Interestingly, members suggested applying good office ergonomics to their quilting environments. This included correct posture by using back braces or pillows to sit straight and adjusting their chairs. In addition, 1 member suggested tilting the sewing machine forward to alleviate stress on the back, shoulders, and neck. Another transferred posture advice from their employer to their sewing:

> *I worked at a fortune 500 company and posture is critical working at a station all day working on a computer. Along w/your breaks, make sure you have good posture at your sewing station. Sit up straight and make sure your chair adjusts for good posture.*

The use of assistive and labor-saving tools were recommended in V2, as hands often felt weak, tired, or sore. Participants suggested using branded spring-loaded scissors, ergonomic rotary cutters, and suction cups to hold down rulers. Members provided an external link to purchase the spring-loaded scissors:

> *Absolutely get spring-loaded scissors. I also use the smaller spring-loaded nippers for handwork/quilting-so much easier than regular scissors. The blades on the nippers are so pointy that I can use them to rip stitches out which is much easier on my hands.*

Experience with arthritis or lack thereof was referred to, either to reinforce or endorse information or to display a lack of certainty about advice. For V1, 4 of the 9 people participating in the conversation indicated they had had knee replacement surgery done. Of the 5 people conversing in V2, 1 stated that she had rheumatoid arthritis and another referred to her experience without saying if she had arthritis. Of the 12 people who responded to V3, 7 referred to their own experience but did not indicate if they have arthritis. Half of the participants either had arthritis or spoke from an experience that informed their responses. Members with lived experience also endorsed others' advice:

> *I had knee replacement 1 year ago yesterday. I agree with the ladies above me.*

Table 3. Alignment between best practice and advice given.

Vignette and best practice	Times mentioned, n (%)
Vignette 1	
Knee replacement (n=22)	
Modify home or environment	9 (41)
Wound care and balanced diet	0 (0)
Resume normal activities	1 (5)
Exercise program	5 (23)
Use gait aid until advised to discontinue, graduated walking program	0 (0)
Other advice	7 (32)
Vignette 2	
Hand arthritis (n=9)	
Medication	1 (11)
Splinting	1 (11)
Other advice	7 (78)
Vignette 3	
Back pain (n=23)	
Physical therapy	3 (13)
Braces	2 (9)
Yoga or Pilates	3 (13)
Aerobic exercise	1 (4)
Healthy weight, avoid smoking, chiropractic, and care with lifting	0 (0)
Good posture	4 (17)
Other advice	10 (44)

Conversely, some people without arthritis or those who had not tried a treatment or tool acknowledged their lack of experience. For example, 1 member suggested trying arthritis gloves but then acknowledged that she has no lived experience with them:

> *I'm not much help.*

Medical Guideline Alignment

Of all of the self-care advice offered, just over half (30/54, 56%) aligned with the OrthoInfo recommendations for knee replacement surgery (V1: 15/22, 68%), hand arthritis (V2: 2/9, 22%), and back pain (V3: 13/23, 57%; see Table 3).

In several instances, participants appeared to offer advice on adapting both the environment and the quilt making in ways that aligned with the OrthoInfo medical guideline. Most of the information that did not align was valuable but was too diverse and quilting-specific. This was particularly evident in V1, where 4 comments referred to resting after a knee replacement surgery whereas the guideline suggests resuming normal activity. Moreover, 2 other comments included advice on pain management, which was not found in the OrthoInfo guideline.

With regard to quilting-specific advice, the guideline did not mention assistive tools for hand arthritis, yet in V2, most of the comments consisted of suggestions for adapting their quilting environment through the use of tools such as suction cups to hold rulers down or spring-loaded scissors. Other comments included resting and keeping fingers active.

Comments that did not align in V3 were diverse. These included varying positions, wearing supportive shoes, finding a comfortable chair or one that is on wheels, taking breaks, and using an orthopedic pillow. A lack of commentary on the best practice recommendation of *care with lifting* in V3 may have been explained by members having craft rooms dedicated to their quilting or assuming that others have craft rooms. These bespoke craft rooms may alleviate the need to carry heavy equipment.

Discussion

Principal Findings and Comparison With Prior Work

Participation Pattern

Low participation rates were noted, as only 22 people or 0.12% (22/18,478) of the Facebook group members participated in the study. A combination of factors may explain this. There was a low rate of posting for the top 100 contributors (from 0-29 posts) and commenting (4 people had made more than 300 comments). These low posting and commenting rates were reflected in the vignette conversations, where 4 of the top 100 contributors responded to the vignettes. As indicated in the conversation of the consent consultation, some members did not see the

relevance of the research to their quilting, and others felt that their quilting group was not the right place for research or medical discussions. Some people used their quilting and the group to avoid life issues while pursuing their hobby. There were more than 10 posts a day in this group, with top recent posts receiving 18 to 61 comments. It is possible that the vignettes were missed, as they were not pinned (prioritized) posts. As the frequency of posting and commenting was already low in the group, it is not surprising that only 22 people responded to the vignettes.

Could the vignettes themselves have contributed to low participation? They were constructed to mimic requests for assistance and advice in other Facebook groups that KD belongs to, including other quilting groups. Posts were pithy and extended a clear question for advice, assistance, or support. Ethics about social media research influenced the decision to use a hypothetical person requesting advice. Informed consent and transparency about data gathering are a prerequisite for research [26]. Consequently, the researcher who was a group member was positioned not only as a quilter but also as a researcher. Trust was important as exhibited by the response of the first group that was approached, and concern during the consent consultation that allowing research in the group would be the start of a slippery slope. To achieve informed consent, it was necessary to provide a link to the website (leading to participant information and consent statements) as part of the vignettes that were posted, resulting in an implied emphasis on the *pretend* person posting the vignette questions.

People using social media enjoy a degree of invisibility and anonymity, especially in closed discussion forums such as closed groups on Facebook. This is particularly valuable to people with long-term health issues who do not need to disclose much about themselves to gain support or find information [27]. They are able to seek information and emotional support, form relationships, affect behavior appropriate to the profile they are projecting, for example, as a quilter, and tell their story on their own terms. Social media becomes a level playing field for people with disabilities or health issues that affect their ability to pursue a hobby. It is possible that participants living with arthritis did not want to reveal their health issues to protect their quilting profile.

Emotional Support

Participants offered emotional support through reassuring and encouraging comments. Some participants justified their contributions by acknowledging that their encouragement is valuable and could ultimately benefit others. Receiving support (online or offline) has been linked to improved health outcomes [13]. The act of offering others support online can act as a vehicle for emotional support, for example, self-worth is enhanced by helping others [28]. Members indicated that recovery after knee replacement surgery is short term and encouraged a positive mindset. This exemplifies a form of advice that Mendelson [29] identified regarding continuity of life and events (despite diagnosis). Emotional support and, more specifically, advice on continuity of life are beneficial, as anxiety and depression are prevalent among people with rheumatoid arthritis [30]. It appears that acts of providing and receiving emotional support are valued by group members who responded to our vignettes.

Information Sharing and Medical Guideline Alignment

Is best practice medical advice accurately shared by quilters living with arthritis? About half of the advice offered aligned with medical best practice for self-care. Noncompliance for physiotherapy and physical exercise programs are common and may be as high as 70% [31,32]. People who assume exercising would not help or experienced barriers to exercising, for example, pain, were less compliant [31,32]. Some participants who responded to the surgical vignette indicated that although they had prepared themselves for convalescence, the blur between postsurgical pain and the need to rest may explain early avoidance of walking on stairs. As our research did not set out to measure offline behavioral changes or improvements in health outcomes, we are not able to draw conclusions about any effects of information sharing.

From a medical perspective, information shared online can be inaccurate, inappropriate, and misinformed [28,33,34]. The nonspecificity of the comments by our participants meant that they were open to interpretation by other members and did not consider potential safety impacts. For example, recommendations to tilt the sewing machine may have been creative or helpful, but there was no conversation about the possibility of the machine falling onto the user's lap or the development of an ergonomically tilted machine by a manufacturer. People have the opportunity to become more informed and empowered with access to health-related information online, but they also are at risk from the consequences of incorrect information because of the volume of shared advice or an inability to discern good from bad advice [35]. With an increase of some age groups using social media for health-related information [36,37], health professionals should be prepared to assist their patients in discerning the quality of information available online [35,38].

Some of the advice that did not align and carried minimal safety risk was still considered to be valuable. This was exemplified in the suggestion for using assistive tools, for example, spring-loaded scissors or suction pads on rulers, which were not explicitly included in the OrthoInfo guideline [24]. Best practice guidelines differ, as evidenced in other guidelines recommending the use of joint protection techniques and assistive tools for the nonpharmaceutical management of hand osteoarthritis [39]. Although some advice with minimal safety risk is useful, it is important to recognize that the advice offered online might raise safety concerns, as advice is open to interpretation and implementation by those who read it. Clinicians cannot be present in every social media platform, and concern for accuracy of information sharing can be mitigated by teaching clinicians about the value of serious leisure for well-being and how to help their patients to critique available support.

Limitations

This was an exploratory study. We acknowledge that there is a plethora of health-related social media groups and services. We were interested in how people integrate their health-related

information and support-seeking into their serious leisure activities, using quilting as context. The sample size was small, but the findings revealed that people do share health-related information and customize it to their serious leisure hobby such as quilting, for example, computer ergonomics applied to posture at a sewing machine to avoid back ache. The findings can be transferred to similar settings, for example, other social media platforms such as Twitter and should be further investigated on a larger scale and in more depth, at which point generalizability of findings and clearer representation can be explored.

As an exploratory study, we used only 1 Facebook group and chose not to compare, for example, Facebook with Twitter, Instagram, Pinterest, or Reddit. Now that it has been suggested that people who participate in serious leisure in the form of quilting do fairly and accurately share health-related information, further studies should be completed using different platforms, comparing different groups in the same platform, and comparing closed and public groups on Facebook.

Participants responded to hypothetical scenarios. KD had observed health-related (and indeed, arthritis-related) conversations in quilting groups on Facebook and crafted the scenarios to closely resemble authentic questions. The scenarios were overtly from a researcher to avoid concerns about deception, as group members are militant about remaining on topic, that is, talking only about quilting-related topics. The obvious presence of the researcher may have resulted in self-selection bias.

Bias could have been introduced by self-selection for participation. Editing of posts and comments (to come across as more helpful, polite, and informed, or to present a profile that pleased the researcher and other quilting group members)

could have introduced bias. The anonymity or crafting of a profile that is more positive than in real life, for example, not revealing health issues may have resulted in some people not participating in the research because they did not want to reveal their health issues. As our analysis of 100 Facebook quilting groups only revealed open pages that were linked to individual quilters (ie, not to actual groups or sets of followers that one might be more likely to find on Twitter or Instagram), it is possible that this research cannot be repeated outside a formal group if repeated on Facebook.

Conclusions

In our exploratory study, we set out to find out what information people share about specific health issues that affect a serious leisure activity such as quilting and how they elicit and provide support to one another to continue with their quilting. People do respond to requests for support and information, although there is a difference between quilting support (eg, I need a new iron, what should I buy?) and health support (eg, I have arthritis, what scissors should I buy?). People provide emotional support for life events in serious leisure social media platforms (eg, offering condolences when a person states that she is making a memory quilt), and this extends to health issues when group members reveal them.

Future research can be informed by our research. The next step is to create a survey that incorporates theories of peer-to-peer support, serious leisure and serious leisure flow, empowerment, and evidence-based health care to establish the utility of online serious leisure groups for supporting a health issue. Future research could explore the extent to which support provided online has the ability to adapt offline behaviors.

Authors' Contributions

KD and RG contributed to the study design and data collection. All authors contributed to the data analysis and interpretation. NG drafted the manuscript, which RG and KD edited and reviewed. KD responded to the reviewers' feedback. All authors read, reviewed, and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Stebbins R. Serious leisure and work. Sociol Compass 2009;3(5):764-774. [doi: 10.1111/j.1751-9020.2009.00233.x]
2. Stebbins R. Serious leisure: a conceptual statement. Pac Sociol Rev 1982;25(2):251-272. [doi: 10.2307/1388726]
3. Greenhalgh T. Patient and public involvement in chronic illness: beyond the expert patient. Br Med J 2009 Feb 17;338:b49. [Medline: 19223339]
4. Vos T, Flaxman S, Naghavi M, Lozano R, Michaud C, Ezzati M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012 Dec 15;380(9859):2163-2196. [doi: 10.1016/S0140-6736(12)61729-2] [Medline: 23245607]

XSL•FO

RenderX

5.   Cross M, Smith E, Hoy D, Nolte S, Ackerman I, Fransen M, et al. The global burden of hip and knee osteoarthritis: estimates from the global burden of disease 2010 study. Ann Rheum Dis 2014 Jul;73(7):1323-1330. [doi: 10.1136/annrheumdis-2013-204763] [Medline: 24553908]

6.   Al Maini M, Adelowo F, Al Saleh J, Al Weshahi Y, Burmester GR, Cutolo M, et al. The global challenges and opportunities in the practice of rheumatology: white paper by the World Forum on Rheumatic and Musculoskeletal Diseases. Clin Rheumatol 2015 May;34(5):819-829 [FREE Full text] [doi: 10.1007/s10067-014-2841-6] [Medline: 25501633]

7.   Kool MB, Geenen R. Loneliness in patients with rheumatic diseases: the significance of invalidation and lack of social support. J Psychol 2012;146(1-2):229-241. [doi: 10.1080/00223980.2011.606434] [Medline: 22303622]

8.   WHO. 2017. Chronic diseases and health promotion: chronic rheumatic conditions URL: http://www.who.int/chp/topics/rheumatic/en/ [accessed 2018-09-23] [WebCite Cache ID 72f4xBhi1]

9.   Greenhalgh T, Howick J, Maskrey N, Evidence Based Medicine Renaissance Group. Evidence based medicine: a movement in crisis? Br Med J 2014 Jun 13;348:g3725 [FREE Full text] [Medline: 24927763]

10. Hunt D, Koteyko N, Gunter B. UK policy on social networking sites and online health: from informed patient to informed consumer? Digit Health 2015;1:2055207615592513 [FREE Full text] [doi: 10.1177/2055207615592513] [Medline: 29942541]

11. Galo JS, Mehat P, Rai SK, Avina-Zubieta A, De Vera MA. What are the effects of medication adherence interventions in rheumatic diseases: a systematic review. Ann Rheum Dis 2016 Apr;75(4):667-673. [doi: 10.1136/annrheumdis-2014-206593] [Medline: 25667208]

12. Patel R, Chang T, Greysen SR, Chopra V. Social media use in chronic disease: a systematic review and novel taxonomy. Am J Med 2015 Dec;128(12):1335-1350. [doi: 10.1016/j.amjmed.2015.06.015] [Medline: 26159633]

13. Rains SA, Peterson EB, Wright KB. Communicating social support in computer-mediated contexts: a meta-analytic review of content analyses examining support messages shared online among individuals coping with illness. Commun Monogr 2015;82(4):403-430. [doi: 10.1080/03637751.2015.1019530]

14. Wicks P, Mack Thorley E, Simacek K, Curran C, Emmas C. Scaling PatientsLikeMe via a "Generalized Platform" for members with chronic illness: web-based survey study of benefits arising. J Med Internet Res 2018;20(5):e175. [Medline: 29735472]

15. The Quilting Company. FabShopNet. 2017. The Quilting in America 2017 survey URL: https://fabshopnet.com/wp-content/uploads/downloads/qia_summary.pdf [accessed 2018-09-24] [WebCite Cache ID 72f66XRlH]

16. Burt EL, Atkinson J. The relationship between quilting and wellbeing. J Public Health (Oxf) 2012 Mar;34(1):54-59. [doi: 10.1093/pubmed/fdr041] [Medline: 21646241]

17. Reynolds F. Textile art promoting well-being in long-term illness: some general and specific influences. J Occup Sci 2004;11(2):58-67. [doi: 10.1080/14427591.2004.9686532]

18. Day K. KDayInformatix. Using social media to elicit support for health issues URL: https://kdayinformatix.wordpress.com/research/using-social-media-to-elicit-support-for-health-issues/ [accessed 2018-09-23] [WebCite Cache ID 72f4bdz4N]

19. Barter C, Renold E. The use of vignettes in qualitative research. Soc Res Update 1999:1-6 [FREE Full text]

20. Esri. ArcGIS. ArcGIS Desktop Software (Version 10) URL: https://www.arcgis.com/index.html [accessed 2018-09-23] [WebCite Cache ID 72f5JNhUt]

21. Williams B, Harris B. Learning logs: structured journals that work for busy people. In: Sankaran S, editor. Effective Change Management Using Action Learning and Action Research: Concepts, Frameworks, Processes, Applications. Lismore, NSW: Southern Cross Press; 2001:97-119.

22. Thomas DR. A general inductive approach for analysing qualitative evaluation data. Am J Eval 2016 Jun 30;27(2):237-246. [doi: 10.1177/1098214005283748]

23. Boyatzis RE. Transforming Qualitative Information: Thematic Analysis and Code Development. London: Sage Publications; 1998.

24. American Academy of Orthopaedic Surgeons. 2018. OrthoInfo URL: https://orthoinfo.aaos.org/ [accessed 2018-09-23] [WebCite Cache ID 72f3GCPMx]

25. American Academy of Orthopedic Surgeons. OrthoInfo. 2017. OrthoInfo Editorial Policy URL: https://www.orthoinfo.org/en/about-orthoinfo/editorial-policy-article/ [accessed 2018-09-23] [WebCite Cache ID 72f3SgHck]

26. Golder S, Ahmed S, Norman G, Booth A. Attitudes toward the ethics of research using social media: a systematic review. J Med Internet Res 2017 Jun 06;19(6):e195 [FREE Full text] [doi: 10.2196/jmir.7082] [Medline: 28588006]

27. Allen C, Vassilev I, Kennedy A, Rogers A. Long-term condition self-management support in online communities: a meta-synthesis of qualitative papers. J Med Internet Res 2016 Mar 10;18(3):e61 [FREE Full text] [doi: 10.2196/jmir.5260] [Medline: 26965990]

28. Merolli M, Gray K, Martin-Sanchez F. Therapeutic affordances of social media: emergent themes from a global online survey of people with chronic pain. J Med Internet Res 2014 Dec 22;16(12):e284 [FREE Full text] [doi: 10.2196/jmir.3494] [Medline: 25533453]

29. Mendelson C. Gentle hugs: internet listservs as sources of support for women with lupus. ANS Adv Nurs Sci 2003;26(4):299-306. [Medline: 14674578]

30. Matcham F, Rayner L, Steer S, Hotopf M. The prevalence of depression in rheumatoid arthritis: a systematic review and meta-analysis. Rheumatology (Oxford) 2013 Dec;52(12):2136-2148 [FREE Full text] [doi: 10.1093/rheumatology/ket169] [Medline: 24003249]

31. Sluijs EM, Kok GJ, van der Zee J. Correlates of exercise compliance in physical therapy. Phys Ther 1993 Nov;73(11):771-782. [Medline: 8234458]

32. Campbell R, Evans M, Tucker M, Quilty B, Dieppe P, Donovan JL. Why don't patients do their exercises? Understanding non-compliance with physiotherapy in patients with osteoarthritis of the knee. J Epidemiol Community Health 2001 Feb;55(2):132-138 [FREE Full text] [Medline: 11154253]

33. Mazzoni D, Cicognani E. Sharing experiences and social support requests in an Internet forum for patients with systemic lupus erythematosus. J Health Psychol 2014 May;19(5):689-696. [doi: 10.1177/1359105313477674] [Medline: 23479300]

34. van Uden-Kraan CF, Drossaert CH, Taal E, Shaw BR, Seydel ER, van de Laar MA. Empowering processes and outcomes of participation in online support groups for patients with breast cancer, arthritis, or fibromyalgia. Qual Health Res 2008 Mar;18(3):405-417. [doi: 10.1177/1049732307313429] [Medline: 18235163]

35. Adams S, de Bont A. Information Rx: prescribing good consumerism and responsible citizenship. Health Care Anal 2007 Dec;15(4):273-290 [FREE Full text] [doi: 10.1007/s10728-007-0061-9] [Medline: 17972056]

36. Findahl O. Göteborg: Göteborgstryckeriet; 2012. Svenskarna och Internet URL: https://www.iis.se/docs/SOI2012.pdf [accessed 2018-09-24] [WebCite Cache ID 72ftOGQJ5]

37. Zickuhr K, Madden M. Pew Internet. Older adults and internet use 2012 URL: http://www.pewinternet.org/2012/06/06/older-adults-and-internet-use/ [accessed 2018-09-23] [WebCite Cache ID 72f4NxR94]

38. Diaz JA, Griffith RA, Ng JJ, Reinert SE, Friedmann PD, Moulton AW. Patients' use of the Internet for medical information. J Gen Intern Med 2002 Mar;17(3):180-185 [FREE Full text] [Medline: 11929503]

39. Hochberg MC, Altman RD, April KT, Benkhalti M, Guyatt G, McGowan J, American College of Rheumatology. American College of Rheumatology 2012 recommendations for the use of nonpharmacologic and pharmacologic therapies in osteoarthritis of the hand, hip, and knee. Arthritis Care Res (Hoboken) 2012 Apr;64(4):465-474. [Medline: 22563589]

Abbreviations

AAOS: American Academy of Orthopaedic Surgeons
V1: vignette 1
V2: vignette 2
V3: vignette 3

XSL•FO
RenderX

Original Paper

Calorie Estimation From Pictures of Food: Crowdsourcing Study

Jun Zhou[1], BE; Dane Bell[2], BA, MA, PhD; Sabrina Nusrat[3], BS, MS, PhD; Melanie Hingle[4], BS, MPH, PhD; Mihai Surdeanu[3], BS, MS, PhD; Stephen Kobourov[3], BS, MS, PhD

[1]Department of Computer Science, Columbia University, New York, NY, United States
[2]Department of Linguistics, University of Arizona, Tucson, AZ, United States
[3]Department of Computer Science, University of Arizona, Tucson, AZ, United States
[4]Department of Nutritional Sciences, University of Arizona, Tucson, AZ, United States

Corresponding Author:
Stephen Kobourov, BS, MS, PhD
Department of Computer Science
University of Arizona
Gould-Simpson 917
1040 East 4th Street
Tucson, AZ, 85721
United States
Phone: 1 520 621 4632
Email: kobourov@email.arizona.edu

Abstract

Background: Software designed to accurately estimate food calories from still images could help users and health professionals identify dietary patterns and food choices associated with health and health risks more effectively. However, calorie estimation from images is difficult, and no publicly available software can do so accurately while minimizing the burden associated with data collection and analysis.

Objective: The aim of this study was to determine the accuracy of crowdsourced annotations of calorie content in food images and to identify and quantify sources of bias and noise as a function of respondent characteristics and food qualities (eg, energy density).

Methods: We invited adult social media users to provide calorie estimates for 20 food images (for which ground truth calorie data were known) using a custom-built webpage that administers an online quiz. The images were selected to provide a range of food types and energy density. Participants optionally provided age range, gender, and their height and weight. In addition, 5 nutrition experts provided annotations for the same data to form a basis of comparison. We examined estimated accuracy on the basis of expertise, demographic data, and food qualities using linear mixed-effects models with participant and image index as random variables. We also analyzed the advantage of aggregating nonexpert estimates.

Results: A total of 2028 respondents agreed to participate in the study (males: 770/2028, 37.97%, mean body mass index: 27.5 kg/m^2). Average accuracy was 5 out of 20 correct guesses, where "correct" was defined as a number within 20% of the ground truth. Even a small crowd of 10 individuals achieved an accuracy of 7, exceeding the average individual and expert annotator's accuracy of 5. Women were more accurate than men ($P<.001$), and younger people were more accurate than older people ($P<.001$). The calorie content of energy-dense foods was overestimated ($P=.02$). Participants performed worse when images contained reference objects, such as credit cards, for scale ($P=.01$).

Conclusions: Our findings provide new information about how calories are estimated from food images, which can inform the design of related software and analyses.

XSL•FO
RenderX

Introduction

Background

Estimating calories in pictures of food is an important task, providing data to inform nutrition research and practice and helping individuals achieve optimal, balanced dietary intakes. Yet this task turns out to be difficult for both experts and nonexperts. We are using this study as an opportunity to enhance our understanding of whether and how calorie estimation works "in the wild," that is, in real-world scenarios. There are many applications of this understanding, ranging from improving the methodological rigor (and reducing the associated burden) of dietary assessment, a pervasive and unanswered question in nutrition science, to influencing the design of interventions focused on dietary behavior change.

The fact that individuals do not estimate calories well [1-4] has motivated the design of software apps to help individuals better estimate different aspects of dietary intake (eg, calories, energy density, nutrient density, and portions) using machine learning (ML) and by harnessing the "wisdom of the crowd." The latter phenomenon was first documented in a 1907 Nature paper [5] and has been successfully used in many domains, ranging from gene network inference [6] to computational problems [7]. Apps in this space remain quite difficult to use, requiring burdensome manual logging of what one eats, or, when ML is used to classify pictures of foods, explicit weight values to be entered manually. To a large extent, the identification of calorie content from images of food either through crowd sourcing or ML remains an open research question. This work is a necessary step toward the automated identification of calorie content from images of food.

Objectives

The aim of this study was to determine the accuracy of crowdsourced annotations of calorie content in food images and to identify and quantify sources of bias and noise as a function of respondent characteristics and food qualities (eg, energy density).

Methods

Procedure

The proposed task is essentially a combination of 2 tests individuals must engage in when estimating calories. The first test relates to the relative energy density of the food pictured, whereas the second test discerns the portion size. Thus, we contend that the ecological validity of our approach is high, despite the task's complexity. The study protocol described herein was reviewed by an institutional review board at the University of Arizona and met the criteria for exemption under 45 CFR 46.101(b).

We designed a simple online quiz administered by a custom-built webpage to measure the accuracy of calorie estimation in pictures of food, verify the existence of collective wisdom, and analyze data and find patterns and trends that can be useful in the design of calorie-tracking apps.

We posted the quiz to SampleSize [8], a subreddit (ie, a forum on reddit) dedicated to posting surveys and survey results. This choice was made on the basis of having a large, active user base that reflects the demographics likely to make large-scale food annotations for reasons of personal interest in self-quantification.

The quiz began with a short introduction: "We would like to see whether you have a good understanding about calories. We will show you several pictures of food and your task, should you choose to accept it, is to guess how many calories are in the food. We will not share any identifying information about you. All of the data is anonymous."

The quiz included 20 questions. Each question consisted of a picture of some food item (see Figure 1) and the prompt, "How many calories are in the food pictured here? (Type a number in the box between 50 and 800)." Implausible dietary data, from (un)intentional under-reporting or over-reporting, are a pervasive problem in nutrition research and can introduce bias or lead to erroneous interpretations of diet-weight or diet-disease relationships. A common way of handling this issue is to exclude extreme values after the fact based on the distribution of the data (eg, removing data more than 2 SDs from the mean) or by subjective assessment [9]. In contrast, we provided the upper and lower limits on the guesses, based on the ground truth data, to ease an already difficult task and thereby reduce the amount of data that would later be necessary to remove. The numbers also helped clarify that we were referring to kilocalories and helped reduce outliers. Neither the correct calorie amounts nor other participants' answers were visible to a participant during the estimation portion of the experiment, although it is possible that some might have read the reddit comments before participation, which revealed some calorie values. We decided to not add additional information to the pictures (eg, *does the sandwich contain mayonnaise?*) to keep the task closer to a realistic image annotation task.

Following the food-related questions, the participants were asked to provide their age group, gender, and body mass index (BMI). An option to calculate BMI via height and weight information was also available. We deliberately chose not to ask for additional demographic questions (eg, location, income, and education) to protect participant privacy. We reported the accuracy of the individual participant who just completed the quiz, as well as the average accuracy of all prior participants, using a breakdown showing the performance of each question.

Quiz Materials

We used 2 categories of food for the pictures: single-ingredient (eg, broccoli, cheese) and mixed-ingredients (eg, sandwich, pizza). There were 20 pictures of food items in total (Figure 1): 12 single-ingredient and 8 mixed. The food shown in the pictures ranged from 100 to 720 kcal. Importantly, we chose these food items according to the United States Department of Agriculture's (USDA) MyPlate model [10] that captures the building blocks for a healthy diet, and which includes 5 types of food (vegetables, fruits, protein, dairy, and grain), as well as mixed foods containing these ingredients. Our selection aimed to follow this model, to include realistic foods that appear in daily consumption, and to be concise so participants engage with the quiz.

The food portions selected are summarized in Table 1. The images were ordered so that each food type was maximally separated from other instances of its type, and the order was the same for each participant. We collected nutrition information about some food items from official restaurant websites. Although the calorie content of the foods pictured was not directly measured, US federal statute requires the published calorie values of restaurant food items to be within 20% of actual calorie value [11].

We chose not to inform the participants of the sources of the images, to reduce the potential that they would search the Web for "ground truth" data, for example, by going to the actual Burger King's website. Likewise, the participants were not explicitly told that some images were from fast food restaurants.

Patterns and Analysis

A total of 3 measures were relevant to our analysis:

1. Error, e, is estimated kilocalories ($\hat{c}$) minus ground truth kilocalories (c), $e = \hat{c} - c$, and percent error, η, is error as a percentage of the ground truth kilocalories, $\eta = e / c$, both of which are positive in overestimation and negative in underestimation. Because of the variation in the ground truth kilocalories of the foods, the latter is a more reliable indicator of the scale of response bias.

2. Absolute error, $|e|$, measures accuracy irrespective of the direction of estimation bias ($|e| = |\hat{c} - c|$).

3. Discrete accuracy, D, is the number of estimates that were within 20% of the true calorie value (out of 20 estimates); discrete accuracy was the measure reported to quiz participants:

Before this analysis, we removed participants who reported a BMI less than 15 or more than 50 kg/m^2 (which are unlikely to be correct), and participants who did not report their gender. In addition, we eliminated responses of less than 50 kcal or greater than 800 kcal and retained all the remaining ones.

We analyzed the results of the survey using linear mixed-effects modeling in R [21,22], allowing regression with random intercepts for both participants and foods simultaneously. The R^2 values are the proportion of the variance in the data that is described by the models' predicted values. For all analyses, a P value less than alpha=.05 was considered indicative of a statistically significant relation.

Figure 1. Untrained participants estimated the food calories in these 20 images.
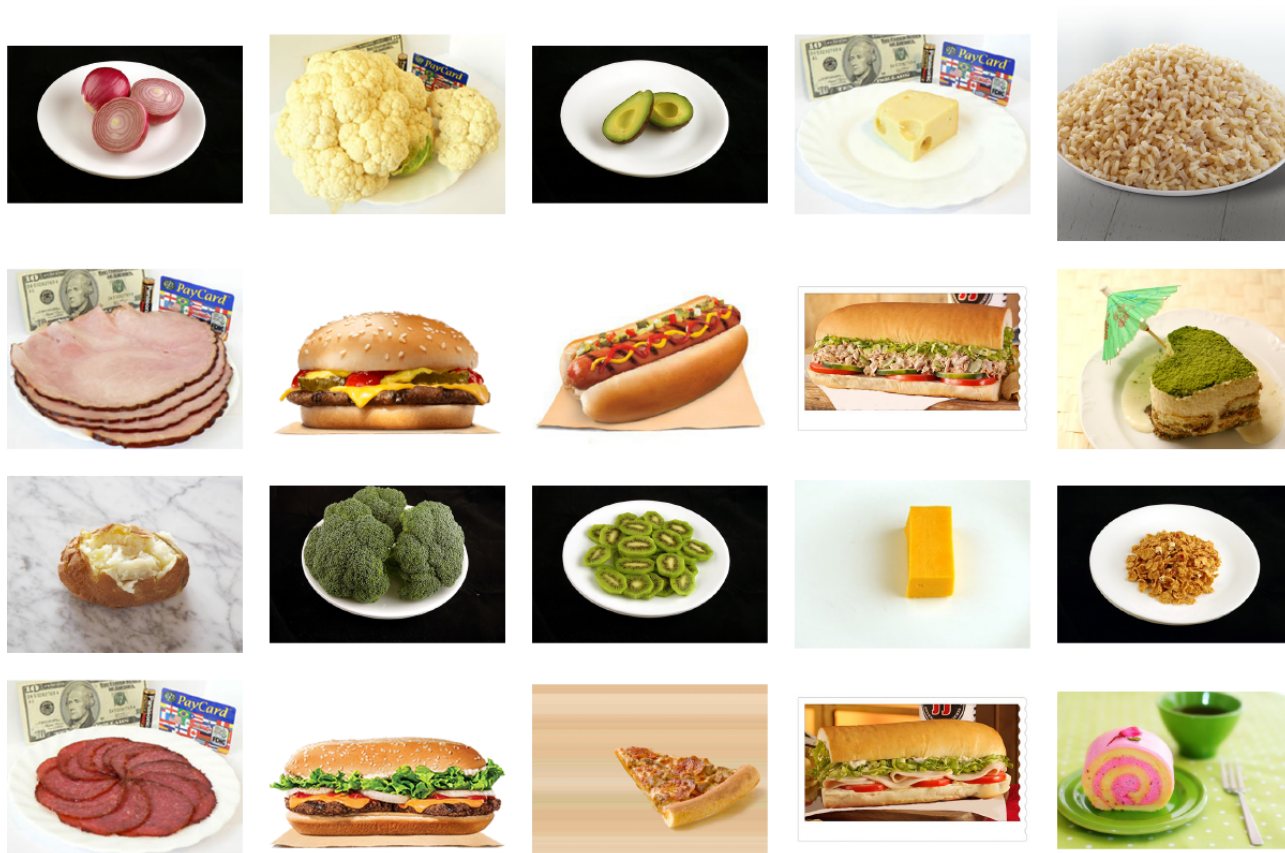
Table 1. Foods were chosen for the quiz to attain maximum coverage of food types encountered in daily life by likely participants. *Scaling* refers to the presence of reference objects, such as credit cards, which could indicate food volume.

Food	Type	Energy (kcal)	Mass (g)	Scaling?	Source
Cheddar cheese	Dairy	200	51	No	wiseGEEK [12]
Gouda cheese	Dairy	300	84	Yes	HealthAssist [13]
Avocado	Fruit	200	125	No	wiseGEEK [12]
Kiwi	Fruit	200	328	No	wiseGEEK [12]
Brown rice	Grain	420	297.7	No	Panda Express [14]
Cereal	Grain	200	55	No	wiseGEEK [12]
Ham	Meat	300	185.1	Yes	HealthAssist [13]
Salami	Meat	300	72.9	Yes	HealthAssist [13]
Red onion	Vegetable	200	475	No	wiseGEEK [12]
Potato	Vegetable	100	141.7	No	Food Network [15]
Broccoli	Vegetable	200	588	No	wiseGEEK [12]
Cauliflower	Vegetable	300	1200	Yes	HealthAssist [13]
Cheeseburger	Mixed	270	104	No	Burger King [16]
Hot dog	Mixed	310	123	No	Burger King [16]
Green tea cake	Mixed	136	40	No	Wit Co, Ltd [17]
Long cheeseburger	Mixed	590	213	No	Burger King [16]
Pepperoni and sausage pizza	Mixed	240	97	No	Papa John's Pizza [18]
Swiss roll	Mixed	251	96	No	Slism [19]
Tuna sandwich	Mixed	720	420	Yes	Jimmy John's [20]
Turkey sandwich	Mixed	510	254	Yes	Jimmy John's [20]

Results

In total, 2125 individuals participated in our reddit quiz. After removing 97 participants with missing or invalid demographic data, 2028 individuals were included in the analysis.

Participant Demographics

The demographics of the participants are summarized in Figures 2-5. Although we collected no location data, an earlier study, again recruiting from the SampleSize subreddit, found that 67.4% (421/625) of participants reported a location within the United States [23], a rate that is similar to the 64% reported in another voluntary survey with participants from across reddit [24]. We also have a higher percentage of female participants than the US average, and a larger fraction of people with BMI around 25 kg/m^2. It is possible that the participants in our quiz were more interested in this topic than the average person. However, in their self-selection, they are more demographically similar than the average person to likely crowdsourcing annotators for potential future app development.

Participant Feedback

The participants volunteered their BMI and other demographic information, and 18 participants left 31 comments on the reddit thread. Table 2 summarizes the types of feedback comments we received, as well as some examples.

The feedback from the participants demonstrates engagement, interest, and curiosity. This implies that such tasks could be legitimately gamified (applying game mechanics and game design techniques to engage and motivate people to achieve their goals). It also shows that unlike Mechanical Turk participants, the participants in our study were engaged and motivated by intrinsic interest.

Note that our work addresses some of the requests shown in Table 2. For example, we found no increased accuracy from the presence of reference objects for scale in the pictures.

XSL•FO

RenderX

Figure 2. The reported gender of respondents to our quiz is compared with data from National Health and Nutrition Examination survey (NHANES).
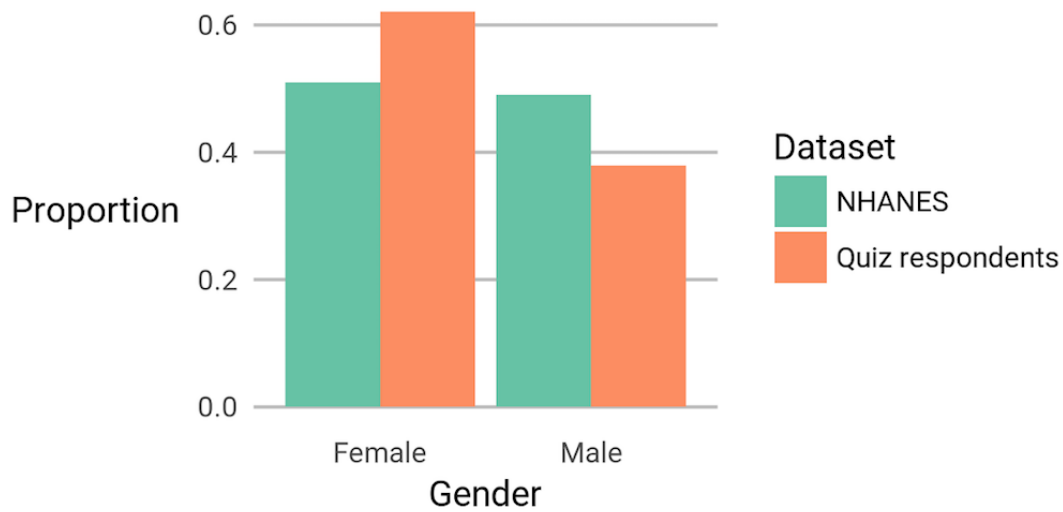
Figure 3. The age of respondents to our quiz is compared with data from National Health and Nutrition Examination survey (NHANES).
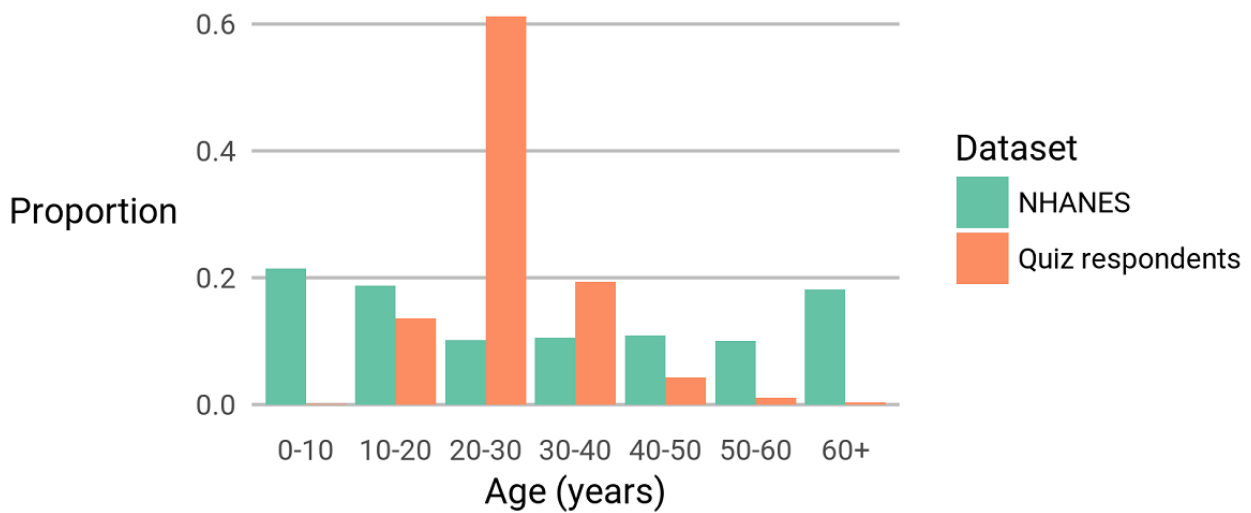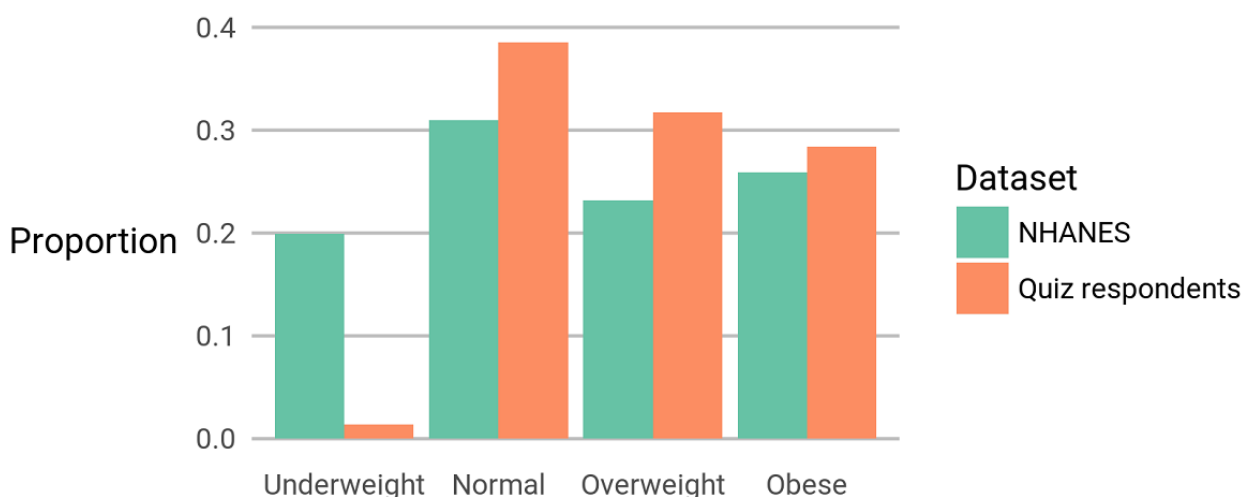

Figure 4. The body mass index (kg/m²) of respondents to our quiz is compared with data from National Health and Nutrition Examination survey (NHANES).
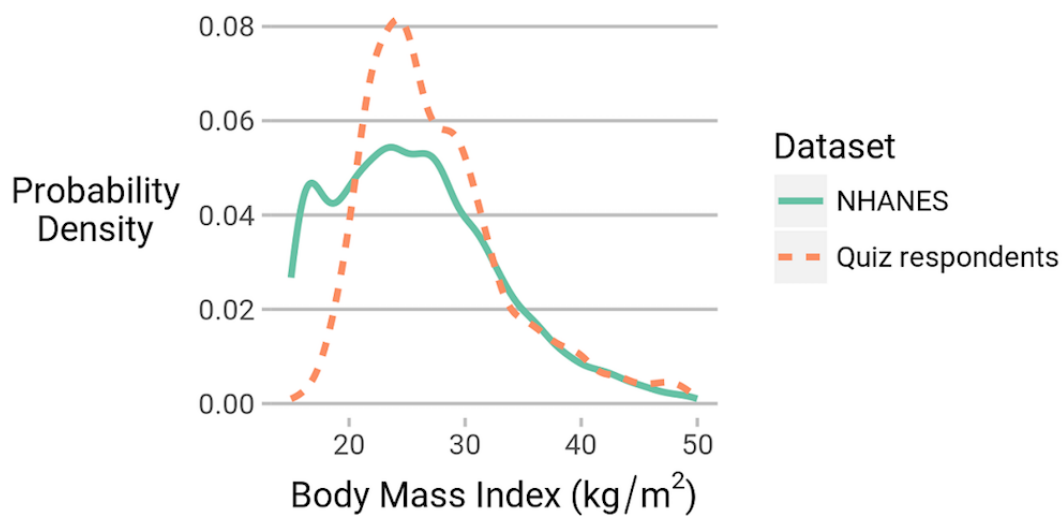
Figure 5. The body mass index (kg/m²) category of respondents to our quiz is compared with data from National Health and Nutrition Examination survey (NHANES).

Table 2. Representative comments from the reddit post of the calorie estimation quiz.

Type	Example
Fun	• "That was fun! I think the folks in 'loseIt' [another subreddit] and on various MFP [MyFitnessPal] forums would enjoy taking this, too."
Surprise	• "I'm really really doubtful that burger is only 270 cal." • "[N]o way are two red onions 200 calories."
Units	• "[C]ountries other than the US use the actual unit of energy- Joules"
Scale	• "It would have been great to have a ruler next to the food." • "[I]f you show me a plate of rice, I can't guess how much rice are on the plate because I don't know how big the plate is."
Difficulty	• "Shoot, got 1 right out of 20 LOL. No wonder my BMI is 29." • "I dont know if there was mayo on [the submarine sandwiches] or not, which changes things a lot."

How Good Are People at Estimating Calories?

The participants' estimates had a mean absolute error ($|e|$) of 57.9% (136 kcal). In terms of discrete accuracy (D), the mean participant answered 5.15 questions correctly out of 20. Figure 6 shows the distribution of correct responses. Absolute error varied considerably by item from the most accurate item—a turkey sandwich, with a mean absolute error of 23.0% (39 kcal)—to the least—green tea cake, at 241.0% (327 kcal) absolute error. Figure 7 illustrates the variety of estimates and percent error (η) distributions for different items. Together, these facts show that human calorie estimates are both inaccurate overall and inconsistent in their inaccuracies.

Does the Wisdom of the Crowd Phenomenon Apply Here?

A consensus formed rapidly for each food, as shown in Figure 8 (see the dashed orange line in the figure), so that 10 responses gave a very good estimate of the next 1000 responses. In fact,

a bootstrap significance test shows that the average of 10 randomly selected participants' guesses is no more (or less) accurate than the average of those of 1000 random participants (P=.36). Moreover, the consensus responses had greater discrete accuracy (D) than that of the individual participants, achieving 7 correct responses out of 20, a 36% relative improvement over the 5.15 correct among individual participants. This result is consistent with previous studies demonstrating the wisdom of the crowd, in which the accuracy of consensus judgments exceeds that of individual judgments (see Comparison With Prior Work).

Another important observation is that although error was high for individual responses and individual foods, the bias in the errors was low overall across all questions, such that the median of the error across items and participants is 0 (when using crowdsourcing over 2028 participants). Although this result is not actionable in itself, as it is averaged across all questions, it does demonstrate the power of crowds to converge toward high-accuracy judgments.

Figure 6. A histogram of the number of correct estimates each participant made. See Patterns and Analysis for the definition of this measure, D.
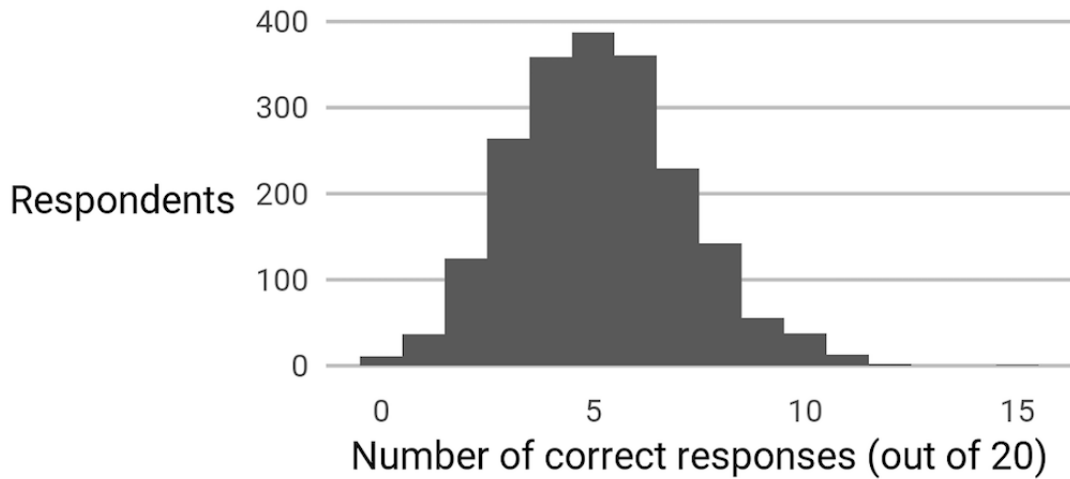

Figure 7. Calorie estimates and percent error (η) for each food item. For each food item, the violin plots represent the distribution of the calorie estimates by the participants and their percent error. The bottom and top of the boxes represent the first and third quartile, and the red band represents the mean of the calorie estimates, respectively. The green band represents the actual calorie value for each food item.
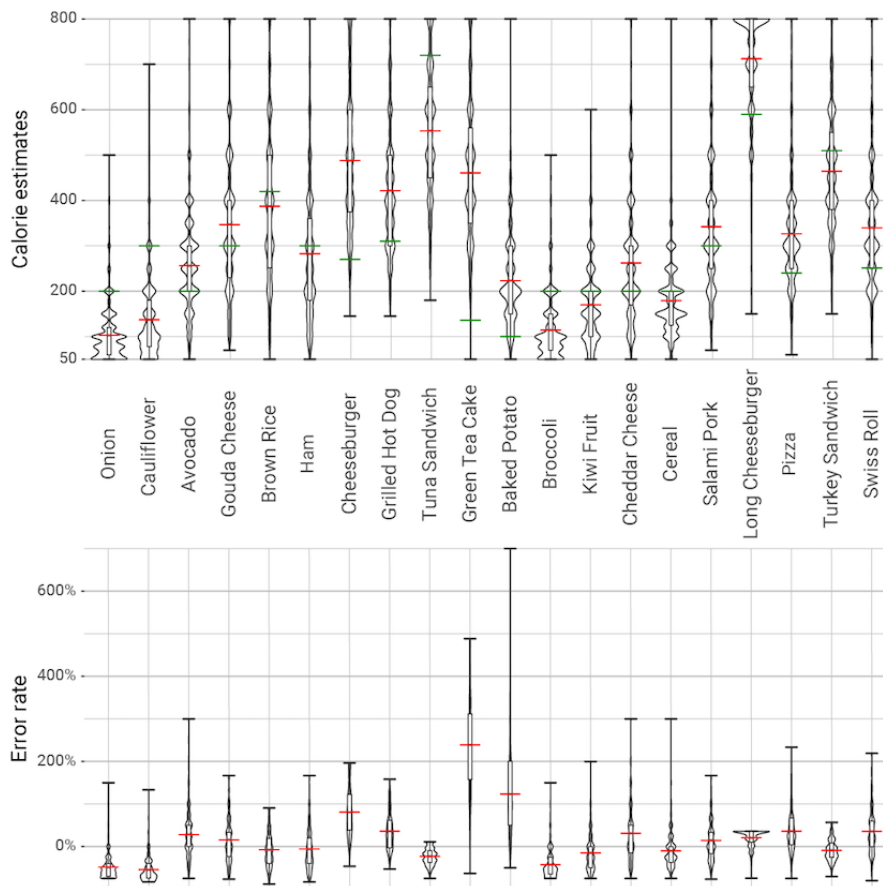
Figure 8. Mean estimates for each food as more participants are added show that a consensus forms rapidly. The dotted blue lines show the true calorie value for each food. The x-axis uses a logarithmic scale. The orange dashed line indicates the estimates of nonexperts. The green continuous line represents the estimates of nutrition science experts. Note that the range of acceptable calorie estimates was 50 to 800 calories for each food item.
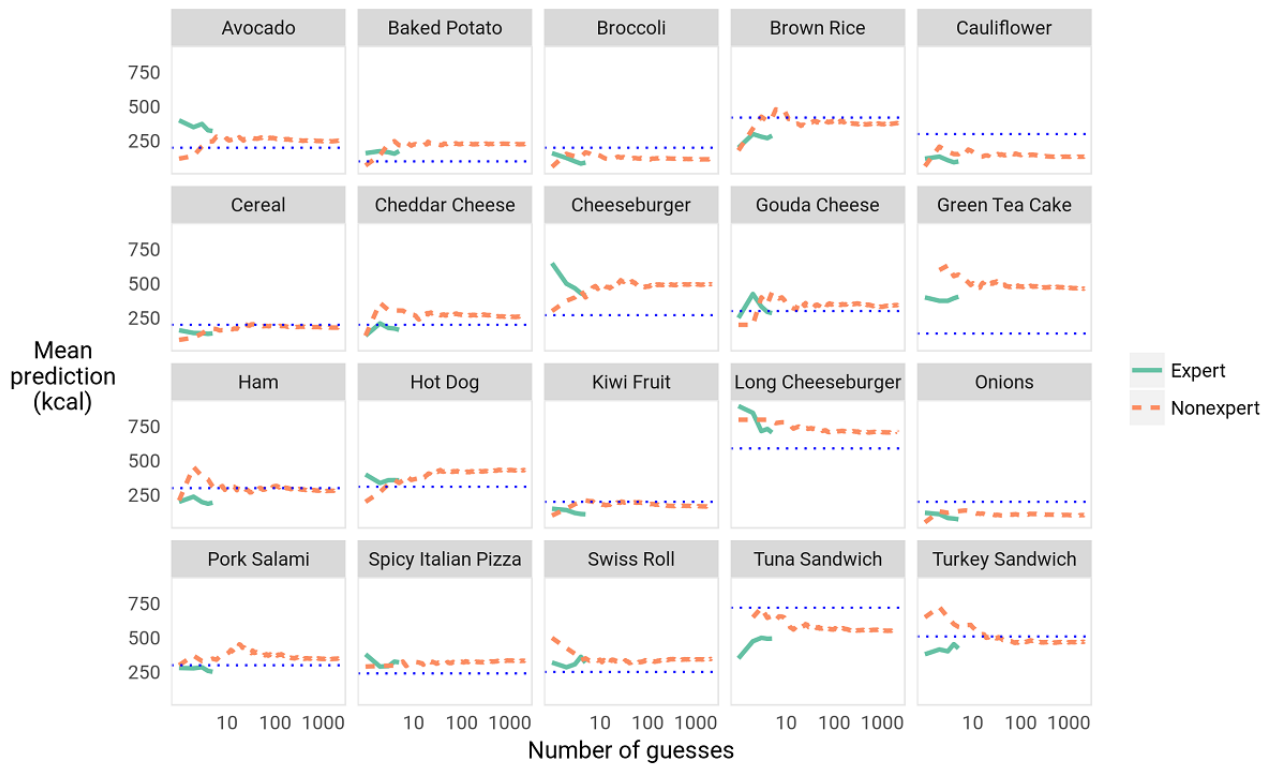
Figure 9. Participants underestimated the calorie content of calorie-sparse foods and overestimated that of calorie-rich foods.
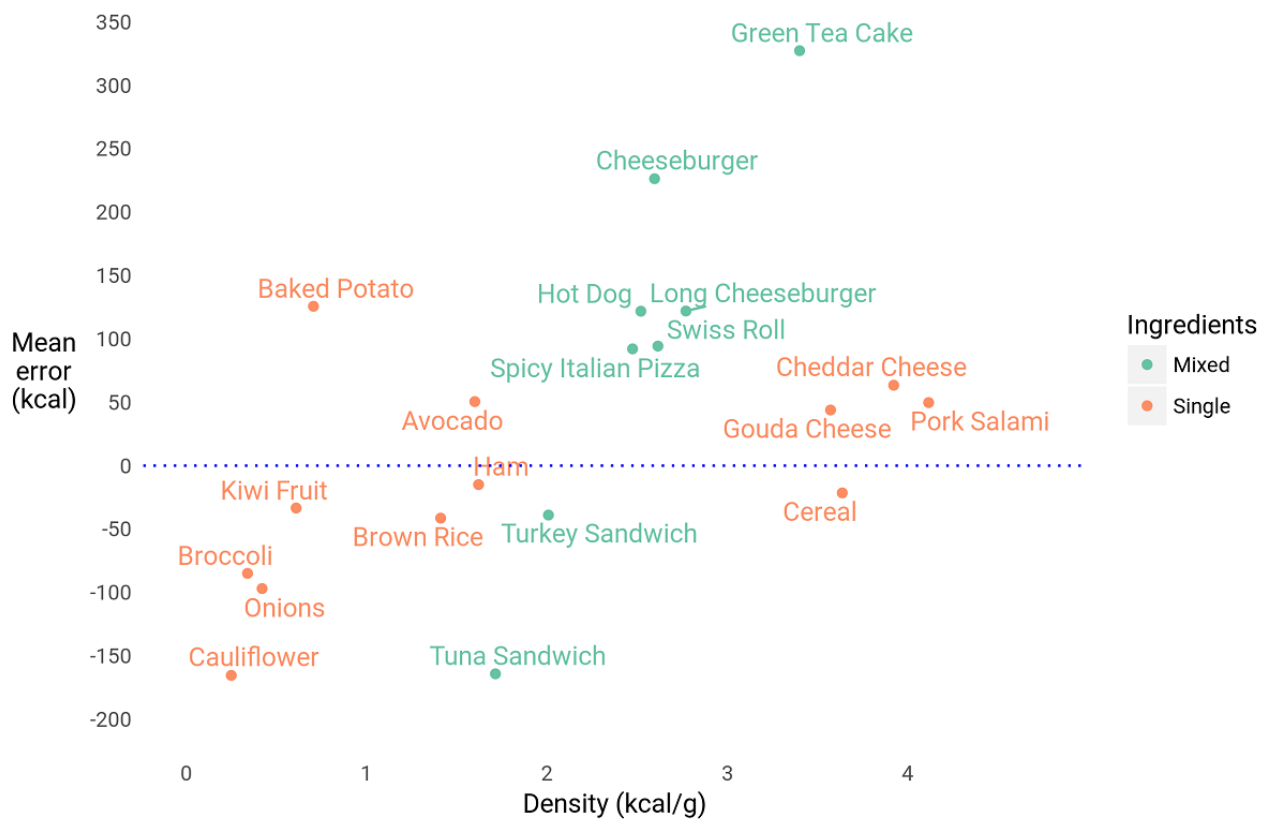
Do the Nutritional Experts Outperform the Crowd?

In addition to redditors, we solicited participation from 5 nutritional experts. We recruited a faculty on a voluntary basis from the Department of Nutritional Science at the University of Arizona and the School of Nutrition and Health Promotion at Arizona State University. Somewhat surprisingly, neither the absolute error of their responses nor their discrete accuracy was statistically different from those of the average nonexpert participant (P=.19). In fact, a small crowd of only 2 randomly selected nonexperts was required to outperform the highest performing expert, achieving an average absolute error ($|e|$) of 119.3 (52.3%) compared with the expert's 130.2 (55.3%). Expert performance is shown in comparison with nonexpert performance in Figure 8. This result is consistent with the hypothesis that the sources of error (eg, erroneous volume estimation due to a notion of typical portion size) apply equally to experts and nonexperts. Prior work in many domains of estimation has supported the notion that a relatively small group of nonexperts can estimate just as well as a single expert [25,26] (see also Comparison With Prior Work).

Does Having an Object for Scale in the Picture Help?

Several comments in the reddit thread expressed the hypothesis that pictures featuring a standard-sized reference object (such as a credit card) were easier to answer. The results showed that reference objects, far from aiding estimation, increased absolute error ($|e|$) by a mean 4.6 kcal (P=.01, R^2=.31). Our hypothesis is that participants used background knowledge about the typical size of foods to scale foods but were not able to profit from comparison against the reference objects. This is statistically significant evidence for the notion that scale information does not aid calorie estimation in digital images (compare [27]). However, it is important to note that this was a post hoc analysis only; the experiment was not designed to analyze this hypothesis. For example, we included objects that come in many different sizes (eg, forks) as reference objects, which may have confused the quiz takers. We leave a more careful evaluation of this particular observation as future work.

Does Energy Density of Foods Predict Estimation Error?

As shown in Figure 9, the caloric content of energy-dense foods was systematically overestimated, and that of energy-sparse foods underestimated, as measured by error (e, P=.02, R^2=.57). This bias is similar to one found by Almiron-Roig et al [28] in estimating in-person portion sizes and could reflect 2 nonexclusive sources. First, it could result from the perceived healthiness of the food items [29]. For example, broccoli is a

prototypically healthy food but is not devoid of calories; conversely, prototypically unhealthy foods such as cheeseburgers have often been "engineered" for low calories [30]. This explanation aligns with the results of Carels et al [1], who found that college students overestimated the caloric content of foods considered to be unhealthy while they underestimated the number of calories in healthy foods. Second, the bias could result from an assumption that the items would have a similar weight to one another, when in fact there was an inverse relationship between the energy density and weight of the items (Pearson correlation: ρ=−.70). We hypothesize that inelastic adjustment of portion size according to energy density could contribute to obesity.

Does Body Mass Index Predict Estimation Errors?

BMI itself does not predict accuracy or bias in these data, similar to Blake et al [31] and Chandon and Wansink [32]. Other studies show that overweight and obese individuals consistently under-report calorie intake to a greater degree than nonoverweight individuals [33,34]. However, BMI does significantly interact with energy density in predicting percent error (η, P=.002, R^2=.57), such that the higher a participant's BMI, the more they exaggerated the calorie content of calorie-rich foods. We hypothesize that overweight individuals are more sensitive to perceptions of food.

Do Gender and Age Predict Estimation Errors?

No biasing effect (toward underestimation, for example) was found, but absolute error ($|e|$) was greater for men than for women (P<.001, R^2=.31), similar to the portion judgment result by Almiron-Roig et al [28]. In addition, the absolute error was greater for older participants (P<.001, R^2=.31), but these effects did not interact. Figures 10 and 11 summarize these differences. We hypothesize that the primary reason for these differences is cultural, reflecting gender norms and the relatively recent cultural emphasis on calories as a measure of healthiness.

Do Estimation Errors Cluster by Food Type?

The over- and underestimation of errors for some foods correlate with those for others. For example, a participant who underestimates the calories in broccoli is likely to do so for cauliflower as well. Figure 12 shows an automatically generated map [35] illustrating these correlations, with clusters showing similar subnetworks. A larger map with more food items would be a strong basis for predicting human bias on clusters of food types (eg, vegetables).

Figure 10. The absolute error (|e|) of participants differs by gender. Box edges show the first and third quartiles and are split by the median. The boxes' whiskers extend to the farthest point within 1.5 times the interquartile range from the box ends. The notches denote the 95% CI of the median. The y-axis is on a square-root scale.
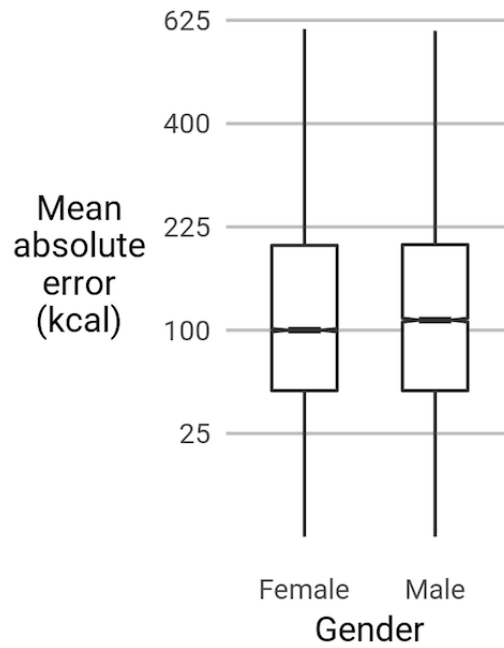

Figure 11. The absolute error (|e|) of participants differs by age. Box edges show the first and third quartiles and are split by the median. The boxes' whiskers extend to the farthest point within 1.5 times the interquartile range from the box ends. The notches denote the 95% CI of the median. The y-axis is on a square-root scale.
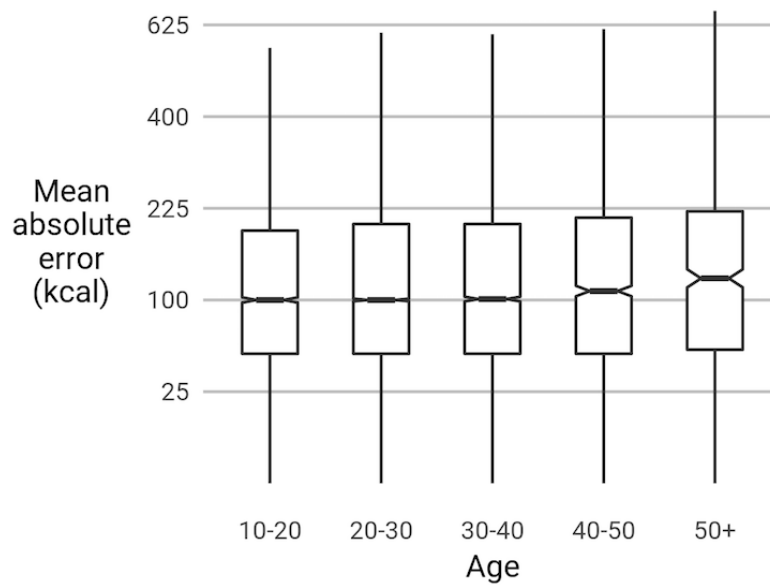
Figure 12. This map shows a network of food items in our survey based on correlation of estimation errors. Pairs of strongly correlated foods are connected by edges. The stronger the correlation, the closer the two (distances are inverse to correlation) are. Clusters show groups of similar subnetworks.
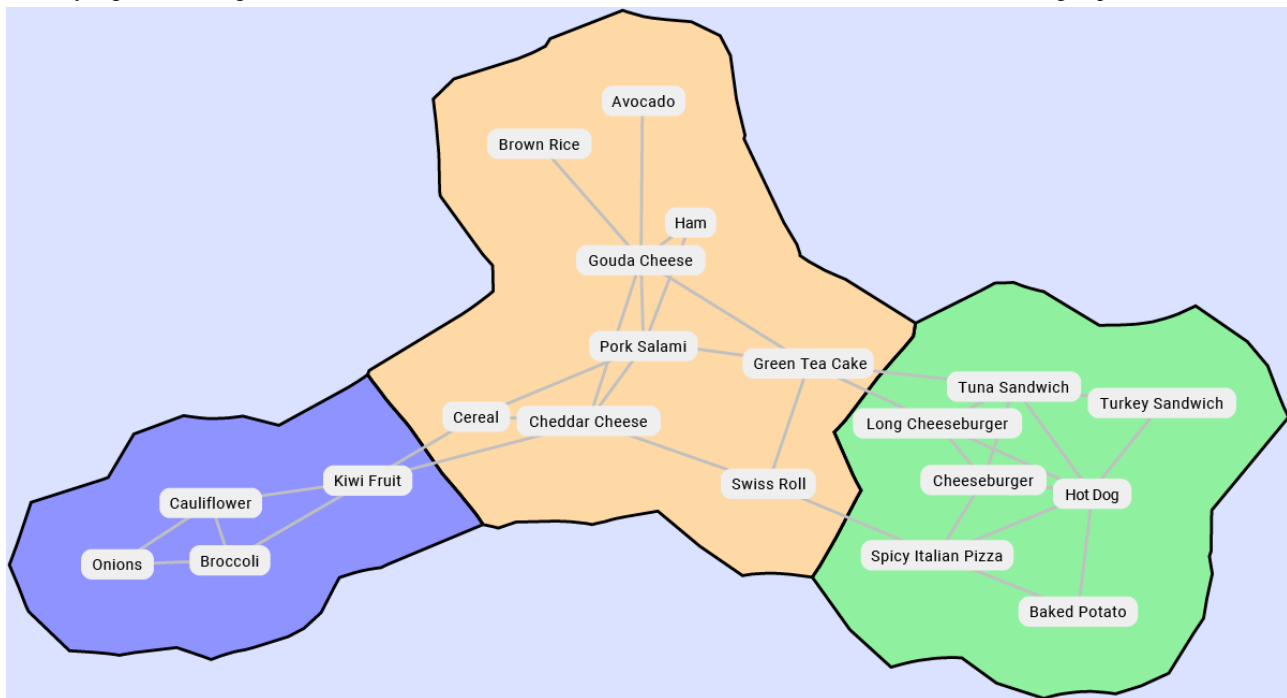
Discussion

Principal Findings

The above analysis identifies several patterns that are important for the design of calorie estimation apps.

First and foremost, our study demonstrates that individuals are poor judges of calorie content in images, and prior work has shown that they are poor judges of portion size in real-life situations (see Comparison With Prior Work). This suggests the utility of an ML approach to calorie estimation to facilitate meal planning. Keeping track of calories by describing foods and guessing quantities and values is a tedious and inaccurate strategy, yet it is the strategy most commonly used in apps today. Given that "a picture is worth a thousand words," our initial hypothesis was that using images (rather than descriptions of foods) should lead to better estimates. Our results, however, do not support this hypothesis: on average, participants performed poorly at estimating the amount of calories in pictures of food, answering 5.15 of 20 questions correctly on average. Our analysis indicates that participants in our dataset tended to exaggerate common dietary knowledge; they underestimated the number of calories in energy-sparse foods and overestimated them in energy-dense ones.

Our related work discussion (Comparison With Prior Work, below) highlights that estimating calories using ML remains an open research problem. However, our work suggests that such apps could take advantage of the wisdom of the crowd for estimation. We showed that the crowd performs better than experts, on average, even when the crowd is small. This suggests that this annotation could be implemented accurately and at low cost.

The results suggest that for apps that focus on calorie monitoring (including self-reporting), it might be a good idea to characterize

the users' demographic data (age, gender, and BMI) shown to influence the accuracy of calorie estimates either directly or when combined with other factors such as energy density.

We identified additional patterns that simplify the design and implementation of calorie-tracking apps. The first such pattern is that scale information does not improve estimation accuracy. The second is that estimation errors cluster by food types, which indicates that the app may extrapolate user patterns between foods in the same group.

It is important to note that the observations of this study are statistically significant and applicable to the population of interest to us (ie, individuals likely to participate in crowdsourced annotations). This population is considerably younger than the US population ($\chi^2_6=3362.5$, $P<.001$) and contains more women proportionally ($\chi^2_1=81.0$, $P<.001$). In future work, we aim to repeat this study for a larger population that matches known demographics to verify the validity of our analysis on such populations.

Comparison With Prior Work

Related work includes prior work in nutritional sciences, ML, image processing, and crowdsourcing. We review a small but representative subset below.

Nutrition and Diet

Bandini et al [36] and Schoeller et al [37] have reported that individuals tend to selectively under-report the energy intake when these data are manually logged. This seems to be especially true for overweight and obese individuals [33,34] and could be associated with a failure to accurately estimate portions, although Blake et al [31] and Chandon and Wansink [32] found that BMI does not correlate with the ability to estimate calories when this task is conducted in person. Portion

estimation of in-person food remains poor, whether in reference to images on computer screens or on printed images [38]. However, calorie estimation of large meals may be worse than that of small meals [39].

To monitor dietary intake more accurately, third-party automated food analysis systems have been proposed. Martin et al [40] used the remote food photography method (RFPM), which requires individuals to upload 3 pictures when having a meal: the plate of the foods selected by an individual, standard portions of known quantities of the foods, and the leftovers. These pictures are sent to trained dietitians who verify portions with participants and analyze these data using a standardized nutrient database. This approach relies on the judgment of trained nutrition professionals and argues for the validity of RFPM. Providing all 3 pictures for each meal is a challenge, as indicated by Williamson et al [41]. Beltran et al [42] tested the reliability of the eButton system, in which a camera worn on the chest records images continuously. The images are captured passively while the participant goes about their day, but such a system still requires experts to identify foods in the images and confirm them with participants. Similar to the RFPM employed by Martin et al [40], the eButton system requires valid pictures before and after each meal, camera placement at a certain angle, and proper lighting. Although promising, such systems are unlikely to scale to the millions of people who would like to accurately track their nutritional intake.

Machine Learning and Image Processing

Given the challenges of the systems described above, a system that can automatically measure calories in pictures of food would be in great demand. Image processing techniques can be used to recognize food in images, and ML can be used to estimate the calories in the food.

Menu-Match [43] uses a database of restaurants and Global Positioning System locations and attempts to guess what is in the picture, using image features such as color and scale-invariant feature transforms [44]. It has not been made available to the general public. Im2Calories [45] is built on the work of Menu-Match. A multi-label classifier is trained on a collection of images of food. The app locates the restaurant a user is dining in and, given an image from the user, the classifier (running on the user's phone) guesses which foods are present in the meal. Looking up the nutritional facts provided by the restaurant, using the resulting estimates, yields good results. Note, however, that Im2Calories has not been made available to the general public or even for research purposes.

Bettadapura et al [46] show that food recognition using location data improves accuracy. Such systems, however, are inherently limited to the restaurants whose menus are in the database. These also assume that menus do not change often and that the volume of food is the same from plate to plate. In reality, most meals are eaten either outside of restaurants or in restaurants whose menus are not included in some dataset. The "in the wild" problem is more natural but also more difficult.

The Web app Foodlog [47,48] divides food images into 300 blocks each and extracts discrete cosine transform coefficients and color histogram from each block. Using these data, Foodlog

classifies the food into 5 categories according to the USDA's My Pyramid system. Experimental results report 88% accuracy in the extraction of food and 73% accuracy in food balance estimation. The FoodCam system [49] segments the region of each food by GrabCut (an image segmentation approach based on iterative graph-cuts) [50], extracts image features of histogram of oriented gradients [51] and color patches with the Fisher Vector (an image representation obtained by pooling local image features) [52], and finally classifies it into 1 of 100 food categories using linear support vector machines.

With the exception of Im2Calories, the systems above achieve relatively good food recognition but without volume estimation. To estimate volume, Chae et al [53] minimize the false-segmented regions, smooth the segmentation boundaries of food, and reconstruct 3D primitive shapes from a single food image. He et al [54] estimate the weight of food given a single image using a shape template for regular-shaped foods and area-based weight estimation for irregularly shaped food. The Im2Calories system [45] estimates the distance of every pixel from the camera by using a convolutional neural net architecture, converts the depth map into a voxel representation, and estimates the volume of the food. Although such approaches are effective, there is no app for estimation of food volumes available to the general public.

Crowdsourcing

Crowdsourcing sometimes makes it possible to use multiple nonexpert judgments to approach the high quality of expert annotation [55]. Surowiecki [25] argues that in many instances, the average nonexpert estimates can even outperform a single expert. Watson has shown that the average of the individual judgments can be equal or superior to the judgment of the best individual within the group [26]. Moreover, the validity of judgments increases with more judges [56]. The strength of the wisdom of the crowd over ML is well understood and exploited in industry. For example, CardMunch (now a service of Evernote [57]) uses crowdsourcing with Amazon Mechanical Turk to convert pictures of business cards into digital contact information. Eloquent Labs [58] uses a mix of crowdsourcing with an artificial intelligence to implement a conversational assistant for customer service.

In the nutrition domain, Mamykina et al [59] show that crowdsourced ingredient annotations from food images are improved by expert annotation and by showing the annotators previous annotations of the images. The PlateMate [60] app leverages crowdsourcing to implement the first step in the RFPM. Rather than typing names of foods and estimating portions, users take photographs of their plates both at the beginning of the meal and at the end to accurately capture how much food was actually eaten. PlateMate uses annotations from nonexpert Amazon Mechanical Turk workers instead of expert dietitians to estimate the composition of foods in static images. PlateMate's results are as accurate as the experts. Similarly, the Im2Calories [45] project uses crowdsourcing to annotate all the food terms that apply to an image. Manually merging synonymous terms, they create the Food201-Multilabel dataset for training. Compared with the original Food101 classes, the new classes of Food201-Multilabel do better according to mean

average precision, as they often correspond to side dishes or small food items.

In sum, despite the abundance of interest in this and related topics, including calorie-tracking apps with manual entry, there exists no publicly available app that will accurately estimate calories from a single image. Likewise, although there are many studies of human bias in tracking calories and lack of skill in estimating portion sizes, no previous work establishes the accuracy and biases of crowdsourcing for calorie estimation, or what demographic factors might correlate with accuracy.

Learning from our study, we envision a very simple app, where the only action required from the user is to take a picture of her or his food. The estimation logic, driven by the wisdom of the crowd and ML, would be transparent to the user, that is, it would be triggered automatically when the camera is used. The logic includes (1) detecting if the picture is a picture of food using image classification [61,62] and (2) routing the image for crowd annotations (similar to CardMunch, which routes the task of processing images of business cards to the crowd). We hope that this simplicity will yield wide adoption, which, in turn, will lead to measurable effects in dietary choices.

Limitations

Participants were not directly informed that some images were of fast food and thus more likely to be subject to food engineering, for example, replacing sugar or using sweetness enhancers, or adding water or protein to enhance food properties and palatability [30,63]. The fact that this was not explicitly mentioned to the participants raises the possibility that participants might have considered these foods as "homemade," which may have influenced perceived energy density and calories. However, since a majority of hamburgers are eaten at restaurants rather than homemade, judgments about engineered foods are as or more relevant than home-cooked foods for both naturalistic and app-related purposes

Conclusions

We described a study measuring the ability of over 2000 individuals to estimate calories in 20 pictures of food chosen to capture the building blocks of a healthy diet [10]. We believe this study should be read as an analysis that drives the design of future food-related apps, with additional impacts on crowdsourcing strategies and the design of human-computer interfaces.

Our analysis confirms some earlier observations (eg, calorie estimation is a difficult task, even for the experts) and offers new insights:

1. Even a small crowd of 2 nonexperts achieves calorie estimation accuracy greater than that of the expert annotators. This suggests that semiautomated food labeling apps can be implemented at a low cost by harnessing the wisdom of the crowd, even when the crowd is small. Note that some prior approaches in this space, such as PlateMate [60], use crowdsourcing to provide calorie information to users. To the best of our knowledge, the crowdsourcing

method has never been tested as a source of data for algorithmic calorie estimation before.

2. We found new type-of-food effects, with energy-dense foods (such as hamburgers) being consistently overestimated and energy-sparse foods (such as broccoli) consistently underestimated. Future crowdsourcing (or ML) projects aiming to annotate food for calorie content will benefit from correction using these biases.

3. We found the absence of some expected correlations. For example, the presence of reference objects for scale does not improve accuracy but rather slightly decreases accuracy, and the BMI is not correlated with accuracy. These observations impact the design of interfaces for annotation apps, as well as data collection protocols.

All in all, this work suggests that calorie-estimation apps are needed and can be built at low cost (eg, using small annotator groups, and without the overhead of including reference objects in images, or controlling for the BMI of users).

Several interesting research questions remain. First, given the low calorie estimation accuracy (5 out of 20) and some clear patterns (underestimating "healthy" foods and overestimating "unhealthy" foods), it is natural to ask whether simple training with feedback can help improve accuracy for nonexperts. If so, how much training is required, what gains in accuracy can be obtained, and how much further can the crowd boost the results? Second, can we factor in biases (eg, age, gender) to obtain better crowdsourced prediction? Third, can better (more consistent) reference objects lead to improvements in accuracy? Fourth, assuming the baseline accuracy for "simple" foods (eg, fruits, vegetables, and sandwiches) can be improved with some of the ideas above, can we hope to tackle more difficult challenges, such as amorphous foods (porridge, mashed potatoes) and liquids (soups, smoothies) in which ingredients and volume are less obvious? Lastly, but perhaps most importantly, we aim to apply the knowledge gained from this study beyond the understanding of how (or how well) people estimate calories to include assessment of diet quality, which has become a dietary construct of interest in the past 5 years [64]. This change has occurred because dietary patterns and dietary quality (eg, increased nutrient density, nutrient diversity, and nutrient adequacy) have been strongly associated with health and disease outcomes. This information provides potentially more meaningful metrics than the number of calories (which says nothing about the quality or "healthiness" of the food) when providing participants or patients with feedback.

We believe this study should be read as an analysis that informs the design of future food-related apps (in particular, apps that feature calorie estimation), with additional potential impacts on crowdsourcing strategies and the design of human-computer interfaces. Our future goal is to provide estimates about judging calories from images for the purpose of mass annotation (eg, in support of a calorie-estimation app), which, in turn, is part of a larger system that analyzes text, images, and videos to estimate risk of diet-sensitive diseases such as type 2 diabetes mellitus [23,65].

Conflicts of Interest

DB and MS disclose a financial interest in lum.ai. This interest has been disclosed to the University of Arizona Institutional Review Committee and is being managed in accordance with its conflict of interest policies.

References

1. Carels RA, Konrad K, Harper J. Individual differences in food perceptions and calorie estimation: an examination of dieting status, weight, and gender. Appetite 2007 Sep;49(2):450-458. [doi: 10.1016/j.appet.2007.02.009] [Medline: 17428574]

2. Carels RA, Harper J, Konrad K. Qualitative perceptions and caloric estimations of healthy and unhealthy foods by behavioral weight loss participants. Appetite 2006 Mar;46(2):199-206. [doi: 10.1016/j.appet.2005.12.002] [Medline: 16466830]

3. Block JP, Condon SK, Kleinman K, Mullen J, Linakis S, Rifas-Shiman S, et al. Consumers' estimation of calorie content at fast food restaurants: cross sectional observational study. Br Med J 2013 May 23;346:f2907. [doi: 10.1136/bmj.f2907]

4. Brown RE, Canning KL, Fung M, Jiandani D, Riddell MC, Macpherson AK, et al. Calorie estimation in adults differing in body weight class and weight loss status. Med Sci Sports Exerc 2016 Mar;48(3):521-526. [doi: 10.1249/MSS.0000000000000796] [Medline: 26469988]

5. Galton F. Vox Populi. Nature 1907 Mar 7;75:450-451 [FREE Full text] [doi: 10.1038/075450a0]

6. Marbach D, Costello JC, Küffner R, Vega NM, Prill RJ, Camacho DM, The DREAM5 Consortium, et al. Wisdom of crowds for robust gene network inference. Nat Methods 2012 Jul 15;9(8):796-804. [doi: 10.1038/nmeth.2016] [Medline: 22796662]

7. Yi SK, Steyvers M, Lee MD, Dry MJ. The wisdom of the crowd in combinatorial problems. Cogn Sci 2012 Apr;36(3):452-470. [doi: 10.1111/j.1551-6709.2011.01223.x] [Medline: 22268680]

8. Reddit. 2017. /r/SampleSize: Where your opinions actually matter! URL: https://www.reddit.com/r/SampleSize/ [accessed 2017-09-22] [WebCite Cache ID 6tg5V8wTB]

9. Huang TT, Roberts SB, Howarth NC, McCrory MA. Effect of screening out implausible energy intake reports on relationships between diet and BMI. Obes Res 2005 Jul;13(7):1205-1217 [FREE Full text] [doi: 10.1038/oby.2005.143] [Medline: 16076990]

10. United States Department of Agriculture. Choosemyplate. 2017. MyPlate Model URL: https://www.choosemyplate.gov/MyPlate [WebCite Cache ID 6tbsR3lvJ]

11. Food and Drug Administration. Food labeling; nutrition labeling of standard menu items in restaurants and similar retail food establishments. Final rule. Fed Reg 2014;79(230):71155-71259 [FREE Full text]

12. wiseGEEK. 2017. What does 200 Calories Look Like? URL: http://www.wisegeek.com/what-does-200-calories-look-like.htm [WebCite Cache ID 6tbsWxbS1]

13. HealthAssist. 2017. 300 Calorie Food Picture Gallery URL: http://www.healthassist.net/food/300kcal/300.shtml [WebCite Cache ID 6tdAdUSkj]

14. Panda Express. 2017. Brown Steamed Rice URL: https://www.pandaexpress.com/menu/sides/brown-steamed-rice [WebCite Cache ID 6tbsukhFK]

15. Food Network. 2017. Foods with 100 Calories URL: http://www.foodnetwork.com/healthy/photos/foods-with-100-calories [WebCite Cache ID 6tbqxkJsp]

16. Burger King. 2017. URL: https://www.bk.com/ [WebCite Cache ID 6tbqVFj2e]

17. Wit Co, Ltd. 2017. Matcha calorie calculation of cake URL: http://www.asken.jp/calculate/meal/94371 [WebCite Cache ID 6tbs116BG]

18. Papa John's Pizza. 2017. URL: https://www.papajohns.com/ [WebCite Cache ID 6tbsSuImT]

19. Slism. 2017. Food nutrients can be seen calorie calculation at a glance URL: http://calorie.slism.jp/200528/ [WebCite Cache ID 6tbt9Z9Ym]

20. Jimmy John's. 2017. Jimmy John's Gourmet Sandwiches URL: https://www.jimmyjohns.com/ [WebCite Cache ID 6tbr3E3nW]

21. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. J Stat Softw 2015;67(1) [FREE Full text] [doi: 10.18637/jss.v067.i01]

22. Kuznetsova A, Brockhoff PB, Christensen R. lmerTest Package: tests in linear mixed effects models. J Stat Softw 2017;82(13):1-26. [doi: 10.18637/jss.v082.i13]

23. Bell D, Fried D, Huangfu L, Surdeanu M, Kobourov S. Towards using social media to identify individuals at risk for preventable chronic illness. 2016 Presented at: Tenth International Conference on Language Resources and Evaluation (LREC); May 23-28, 2016; Portorož, Slovenia p. 23-28 URL: http://www.lrec-conf.org/proceedings/lrec2016/summaries/30.html

24. Reddit. 2011. Who in the World is reddit? Results are in URL: https://redditblog.com/2011/09/12/who-in-the-world-is-reddit-results-are-in/ [WebCite Cache ID 6tbvF5LJE]

25. Surowiecki J. The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations. New York: Doubleday; 2004.

26. Watson GB. Do groups think more efficiently than individuals? J Abnorm Soc Psychol 1928;23(3):328-336. [doi: 10.1037/h0072661]

27. Hernández T, Wilder L, Kuehn D, Rubotzky K, Moser-Veillon P, Godwin S, et al. Portion size estimation and expectation of accuracy. J Food Compos Anal 2006 Aug;19:S14-S21. [doi: 10.1016/j.jfca.2006.02.010]

28. Almiron-Roig E, Solis-Trapala I, Dodd J, Jebb SA. Estimating food portions. Influence of unit number, meal type and energy density. Appetite 2013 Dec;71:95-103. [doi: 10.1016/j.appet.2013.07.012]

29. Faulkner GP, Pourshahidi LK, Wallace JM, Kerr MA, McCaffrey TA, Livingstone MBE. Perceived 'healthiness' of foods can influence consumers' estimations of energy density and appropriate portion size. Int J Obes Relat Metab Disord 2013 May 07;38(1):106-112. [doi: 10.1038/ijo.2013.69]

30. Pérez-Escamilla R, Obbagy JE, Altman JM, Essery EV, McGrane MM, Wong YP, et al. Dietary energy density and body weight in adults and children: a systematic review. J Acad Nutr Diet 2012;112(5):a. [Medline: 22480489]

31. Blake AJ, Guthrie HA, Smiciklas-Wright H. Accuracy of food portion estimation by overweight and normal-weight subjects. J Am Diet Assoc 1989 Jul;89(7):962-964. [Medline: 2745915]

32. Chandon P, Wansink B. Is obesity caused by calorie underestimation? A psychophysical model of meal size estimation. J Mark Res 2007 Feb;44(1):84-99. [doi: 10.1509/jmkr.44.1.84]

33. Bailey RL, Mitchell DC, Miller C, Smiciklas-Wright H. Assessing the effect of underreporting energy intake on dietary patterns and weight status. J Am Diet Assoc 2007 Jan;107(1):64-71. [doi: 10.1016/j.jada.2006.10.009]

34. Kretsch MJ, Fong AK, Green MW. Behavioral and body size correlates of energy intake underreporting by obese and normal-weight women. J Am Diet Assoc 1999 Mar;99(3):300-306. [doi: 10.1016/s0002-8223(99)00078-4]

35. Kobourov SG, Pupyrev S, Simonetto P. Visualizing graphs as maps with contiguous regions. 2014 Presented at: Eurographics Conference on Visualization (EuroVis); 2014; Swansea, Wales, UK URL: https://pdfs.semanticscholar.org/c3dd/bbb159ece5672c4488604fdbe1acd25fae59.pdf

36. Bandini LG, Schoeller DA, Cyr HN, Dietz WH. Validity of reported energy intake in obese and nonobese adolescents. Am J Clin Nutr 1990 Sep;52(3):421-425. [doi: 10.1093/ajcn/52.3.421] [Medline: 2393004]

37. Schoeller DA, Bandini LG, Dietz WH. Inaccuracies in self-reported intake identified by comparison with the doubly labelled water method. Can J Physiol Pharmacol 1990 Jul;68(7):941-949. [doi: 10.1139/y90-143]

38. Hernández T, Wilder L, Kuehn D, Rubotzky K, Moser-Veillon P, Godwin S, et al. Portion size estimation and expectation of accuracy. J Food Comp Anal 2006;19:S14-S21 [FREE Full text]

39. Wansink B, Chandon P. Meal size, not body size, explains errors in estimating the calorie content of meals. Ann Intern Med 2006 Sep 05;145(5):326. [doi: 10.7326/0003-4819-145-5-200609050-00005]

40. Martin CK, Han H, Coulon SM, Allen HR, Champagne CM, Anton SD. A novel method to remotely measure food intake of free-living individuals in real time: the remote food photography method. Br J Nutr 2009 Feb;101(3):446-456 [FREE Full text] [doi: 10.1017/S0007114508027438] [Medline: 18616837]

41. Williamson DA, Allen HR, Martin PD, Alfonso A, Gerald B, Hunt A. Digital photography: a new method for estimating food intake in cafeteria settings. Eat Weight Disord 2013 Jul 26;9(1):24-28. [doi: 10.1007/bf03325041]

42. Beltran A, Dadabhoy H, Chen T, Lin C, Jia W, Baranowski J, et al. Adapting the eButton to the abilities of children for diet assessment. 2016 Presented at: 10th International Conference on Methods and Techniques in Behavioral Research; May 25-27, 2016; Dublin, Ireland p. 25-27 URL: https://www.measuringbehavior.org/files/2016/MB2016_Proceedings.pdf

43. Beijbom O, Joshi N, Morris D, Saponas S, Khullar S. Menu-Match: Restaurant-Specific Food Logging from Images. 2015 Presented at: IEEE Winter Conference on Applications of Computer Vision; 2015; Waikoloa, HI, USA. [doi: 10.1109/wacv.2015.117]

44. Lowe D. Object recognition from local scale-invariant features. 1999 Presented at: Seventh IEEE International Conference on Computer Vision; 1999; Kerkyra, Greece. [doi: 10.1109/iccv.1999.790410]

45. Myers A, Johnston N, Rathod V, Korattikara A, Gorban A, Silberman N, et al. Im2Calories: Towards an Automated Mobile Vision Food Diary. 2015 Presented at: IEEE International Conference on Computer Vision (ICCV); 2015; Santiago, Chile. [doi: 10.1109/iccv.2015.146]

46. Bettadapura V, Thomaz E, Parnami A, Abowd G, Essa I. Leveraging Context to Support Automated Food Recognition in Restaurants. 2015 Presented at: IEEE Winter Conference on Applications of Computer Vision; January 5-9, 2015; Waikoloa, HI, USA. [doi: 10.1109/wacv.2015.83]

47. Kitamura K, Yamasaki T, Aizawa K. Food log by analyzing food images. 2008 Presented at: 16th ACM international conference on Multimedia - MM 08; 2008; Vancouver, BC, Canada. [doi: 10.1145/1459359.1459548]

48. Kitamura K, Yamasaki T, Aizawa K. FoodLog. 2009 Presented at: ACM multimedia workshop on Multimedia for cooking and eating activities - CEA 09; 2009; Beijing, China. [doi: 10.1145/1630995.1631001]

49. Kawano Y, Yanai K. FoodCam: A real-time food recognition system on a smartphone. Multimed Tools Appl 2014 Apr 12;74(14):5263-5287. [doi: 10.1007/s11042-014-2000-8]

50. Rother C, Kolmogorov V, Blake A. Grabcut: interactive foreground extraction using iterated graph cuts. ACM Trans Graph 2004;23(3):309-314. [doi: 10.1145/1015706.1015720]

51. Dalal N, Triggs B. Histograms of oriented gradients for human detection. 2005 Presented at: IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05); June 20-25, 2005; San Diego, CA, USA. [doi: 10.1109/CVPR.2005.177]

52. Csurka G, Perronnin F. Fisher vectors: beyond bag-of-visual-words image representations. In: Communications in Computer and Information Science Computer Vision, Imaging and Computer Graphics Theory and Applications. Berlin, Heidelberg: Springer; 2011:28-42.

53. Chae J, Woo I, Kim S, Maciejewski R, Zhu F, Delp EJ, et al. Volume estimation using food specific shape templates in mobile image-based dietary assessment. Proc SPIE Int Soc Opt Eng 2011 Feb 07;7873:78730K [FREE Full text] [doi: 10.1117/12.876669] [Medline: 22025936]

54. He Y, Xu C, Khanna N, Boushey CJ, Delp EJ. Food image analysis: Segmentation, identification and weight estimation. 2013 Presented at: IEEE International Conference on Multimedia and Expo (ICME); 2013; San Jose, CA, USA. [doi: 10.1109/icme.2013.6607548]

55. Brabham DC. Crowdsourcing as a model for problem solving: an introduction and cases. Convergence 2008;14(1):75-90. [doi: 10.1177/1354856507084420]

56. Gordon K. Group judgments in the field of lifted weights. J Exp Psychol 1924;7(5):398-400. [doi: 10.1037/h0074666]

57. Evernote. 2017. Get organized. Work smarter. Remember everything URL: http://evernote.com [accessed 2017-09-21] [WebCite Cache ID 6udUVuy9M]

58. Eloquent Labs. 2017. URL: https://www.eloquent.ai/ [WebCite Cache ID 6uhsMCtST]

59. Mamykina L, Smyth T, Dimond J, Gajos K. Learning from the crowd: Observational learning in crowdsourcing communities. 2016 Presented at: CHI Conference on Human Factors in Computing Systems - CHI 16; 2016; San Jose, CA, USA. [doi: 10.1145/2858036.2858560]

60. Noronha J, Hysen E, Zhang H, Gajos K. Platemate. 2011 Presented at: 24th Annual ACM symposium on User interface software and technology - UIST 11; 2011; Santa Barbara, CA, USA. [doi: 10.1145/2047196.2047198]

61. Kagaya H, Aizawa K. Highly accurate food/non-food image classification based on a deep convolutional neural network. 2015 Presented at: International Conference on Image Analysis and Processing; September 7, 2015; Genova, Italy p. 350-357. [doi: 10.1007/978-3-319-23222-5_43]

62. Kagaya H, Aizawa K, Ogawa M. Food Detection and Recognition Using Convolutional Neural Network. 2014 Presented at: Proceedings of the ACM International Conference on Multimedia - MM 14; 2014; Orlando, FL, USA. [doi: 10.1145/2647868.2654970]

63. Meister K, Doyle M. Obesity and Food Technology. American Council on Science and Health; 2009. URL: https://www.scribd.com/document/37170221/Obesity-and-Food-Technology [WebCite Cache ID 6tfsQIMLD]

64. Tapsell LC, Neale EP, Satija A, Hu FB. Foods, nutrients, and dietary patterns: interconnections and implications for dietary guidelines. Adv Nutr 2016 Dec;7(3):445-454 [FREE Full text] [doi: 10.3945/an.115.011718] [Medline: 27184272]

65. Rains SA, Hingle MD, Surdeanu M, Bell D, Kobourov S. A test of the risk perception attitude framework as a message tailoring strategy to promote diabetes screening. Health Commun 2018 Jan 26:1-8. [doi: 10.1080/10410236.2018.1431024] [Medline: 29373042]

Abbreviations

BMI: body mass index
ML: machine learning
RFPM: remote food photography method
USDA: United States Department of Agriculture

XSL•FO
RenderX

Extrahepatic Autoimmune Diseases are Prevalent in Autoimmune Hepatitis Patients and Their First-Degree Relatives: Survey Study

Rachel Fogel[1], BS; Megan Comerford[1], BS, MSPH; Prianka Chilukuri[1], MD; Eric Orman[1], MSc, MD; Naga Chalasani[1], MD; Craig Lammert[1], MD, MS

Division of Gastroenterology and Hepatology, Indiana University, Indianapolis, IN, United States

Corresponding Author:
Craig Lammert, MD, MS
Division of Gastroenterology and Hepatology
Indiana University
702 Rotary Circle
Indianapolis, IN,
United States
Phone: 1 3172746492
Email: clammert@iu.edu

Abstract

Background: Concurrent autoimmune illnesses contribute to increased medical burden and reduced quality of life in patients with autoimmune hepatitis (AIH). The frequency of coexisting autoimmune conditions among North American patients with AIH and their families remains incomplete. Challenges associated with disease capture in the electronic medical record, high study costs, and geographic spread of patients are formidable barriers to understanding the extent of concurrent autoimmune conditions in these groups.

Objective: This objective of this study was to examine the frequency of extrahepatic autoimmune diseases (EHAD) among AIH cases and healthy controls as well as their first-degree relatives using social networking sites (SNS).

Methods: We developed a 53-question survey detailing the history of autoimmune diseases. A survey link was posted at routine intervals within specific Web-based cohorts on SNS. Healthy controls, without self-reported autoimmune liver disease, were recruited from Amazon's Mechanical Turk. Continuous variables were summarized using medians and P values obtained with the Wilcoxon rank-sum test. Categorical variables were compared using the chi-square test.

Results: Compared with controls (n=1162), cases (n=306) were more likely to be older (median age: 49 vs 33 years), female (284/306, 92.81% vs 955/1162, 82.18%), and have an EHAD (128/306, 41.83% vs 218/1162, 18.76%; P=.001). The most frequent EHADs among cases were thyroid disease (49/306, 16.01%), Sjögren syndrome (27/306, 8.82%), Raynaud phenomenon (23/306, 7.52%), and psoriasis (22/306, 7.19%). Overall, 55.88% (171/306) of cases and 35.71% (1601/4484) of controls reported at least 1 first-degree relative (FDR) with a history of EHAD (P=.001). Cases had a significantly higher risk of EHAD than controls after the adjustment for age, sex, race, and body mass index: odds ratio 2.46 (95% CI 1.8-3.3); P=.001.

Conclusions: Patients with AIH report higher prevalence of coexistent EHAD than healthy controls, and their FDRs are also more likely to have autoimmune disorders.

Introduction

Autoimmune hepatitis (AIH) is characterized by T-cell-mediated inflammation of the liver and typical autoantibodies [1,2]. If left untreated, AIH can result in progressive liver disease, including cirrhosis and failure, often requiring liver transplantation [3]. Several epidemiological studies have yielded a wide array of incidence globally, with rates of 0.08 per 100,000 in Japan [4] to 43 per 100,000 among Alaska natives [5]. Regardless of the geographical location, AIH is known to affect women primarily and the age of onset appears to be bimodal [6].

The etiology of AIH remains unclear, but both environmental and genetic factors have been hypothesized despite few

supporting studies. Environmental contributions are supported by varying global AIH incidence rates [4-6], evidence of both drug-induced and viral-induced AIH [7-9], and disease risk associated with certain exposures [10]. Genetic risk associations have been observed predominantly at the human leukocyte antigen (*HLA*) locus [11] and in at least one non-*HLA* gene [12]. Beyond limited genome-wide data and prior candidate gene studies, descriptive European reports support an underlying autoimmune phenotype, as observation of extrahepatic autoimmune diseases (EHAD) occurring concurrently with AIH has been as high as 42% [6,13,14]. To date, only 2 European studies have examined the family history of EHAD in patients with AIH by collecting survey data and retrospectively reviewing patient records [6,14]. Furthermore, a case-control assessment of EHAD among patients with AIH and their first-degree relatives (FDR) remains incomplete.

Geographic barriers and lack of a nationwide medical record system have limited AIH investigation; however, the advent of Web-based social networking sites (SNS), such as Facebook and Twitter, have bridged this gap in research accessibility [15]. Web-based patient groups can now provide an integrative, synergistic system for collecting data and engaging patients in the participation and advancement of research [16,17]. This study aims to utilize a novel research method centered on SNS to examine the association between AIH and the presence of autoimmune diseases among patients with AIH and their FDRs.

Methods

Social Networking Sites

We implemented an AIH patient recruitment method using SNS for this study. We have previously described the advantages of these platforms to provide patients with AIH with access to health information, patient-directed support, and opportunities for research involvement [18]. The Autoimmune Hepatitis Research Network, a private Facebook group created and managed by a physician-led research team at Indiana University, was created in 2014 and currently hosts nearly 1800 members. Patients with AIH (cases) were recruited to participate through monthly electronic study advertisements over a 6-month study period (June 2015-January 2016) on the Autoimmune Hepatitis Research Network, as well as the Autoimmune Hepatitis Association public Facebook page. In addition, the study team posted a research invitation on associated Twitter accounts monthly (@craiglammertIU: 218 and and @AIH_Association: 367 followers, as of June 2017)] during the study period. Advertisements included disease background information, inclusion and exclusion information, and a direct survey link.

For participation, cases were required to be aged ≥18 years and have previously received a diagnosis of AIH from a medical doctor.

Acquisition of the Control Population

Controls without self-reported AIH were screened and recruited from Amazon's Mechanical Turk (MTurk), a crowdsourcing website for the completion of requester-directed tasks, which has been shown to approximate the demographics of American adults over the Web [19]. MTurk workers with high approval ratings have been shown to answer attention check questions correctly, thereby minimizing the selection bias [20]. Thus, to obtain high-quality data, participation was limited to high-reputation MTurk workers with a 95% approval rating or above. A survey link was posted to MTurk daily, and MTurk workers completed the survey in exchange for a small monetary reward. Controls were required to be aged ≥18 years, US residents, and report no history of autoimmune liver disease. The controls received US $0.25 compensation for their participation and were only able to complete the survey once. The Indiana University Institutional Review Board approved all study methods.

Questionnaire Development

An autoimmune disease questionnaire was created using Web-based survey development software and included key demographics, as well as 48 questions addressing personal and FDR medical history of other autoimmune diseases. The questionnaire was adapted from the Mayo Clinic Medical Questionnaire for Chronic Cholestatic Liver Diseases and took approximately 10 minutes to complete [21]. We assessed 12 common EHAD and 3 hepatic autoimmune diseases, including AIH, primary sclerosing cholangitis (PSC) and primary biliary cholangitis (PBC). A total of 2081 responses to the survey were collected over the 6-month study period. Incomplete surveys, duplicate entries, and self-reported liver disease in controls were removed. A total of 1468 participants were included in the study analysis, including 306 cases and 1162 controls (Figure 1).

Statistical Analysis

Continuous variables were summarized using medians and the 25th and 75th percentiles, and *P* values were obtained with the Wilcoxon rank-sum test. Categorical variables were compared using the chi-square test. Logistic regression was used to estimate the risk of concurrent EHAD and EHAD within FDR after adjusting for age, sex, race, and body mass index (BMI). Statistical analysis was completed using IBM SPSS Statistics software (version 2.0).

Figure 1. Electronic recruitment, exclusion, and inclusion of study participants. IRB: international review board, AHRN: Autoimmune Hepatitis Research Network, AIHA: Autoimmune Hepatitis Association.
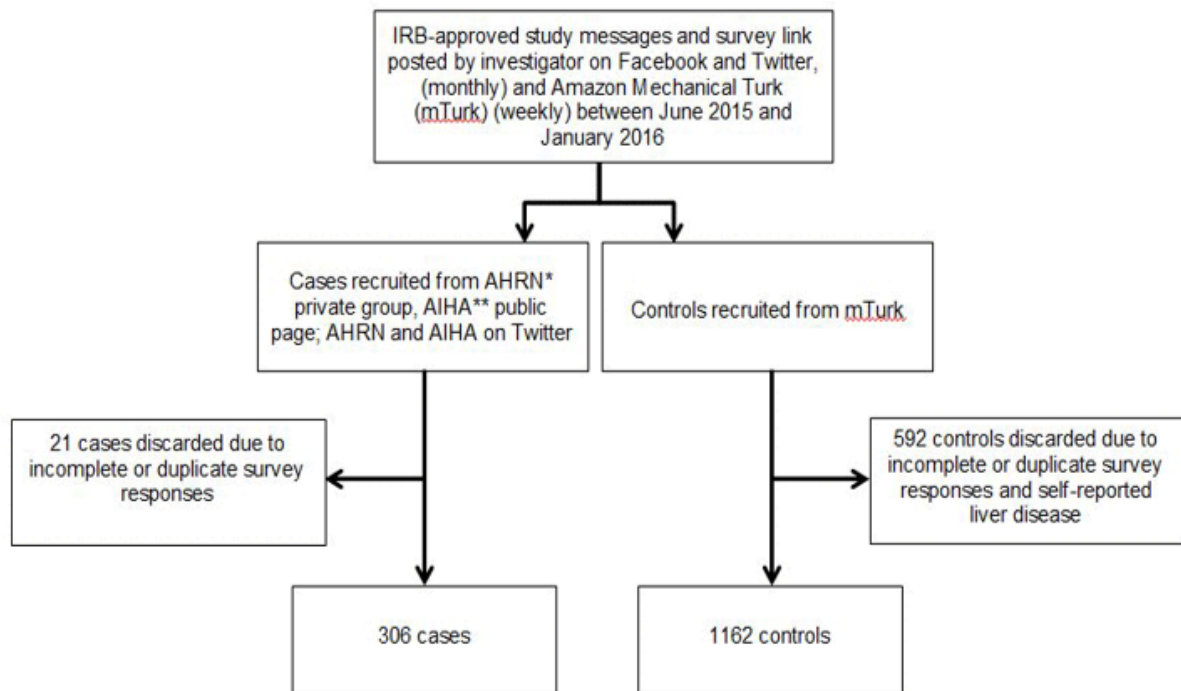
Results

Table 1 presents the demographic characteristics of cases and controls.

Compared with controls, cases were more likely to be female (284/306, 92.81% vs 955/1162, 82.18%), older (49 years vs 33 years, median), Caucasian (266/306, 86.93% vs 944/1162, 81.23%), and have a higher BMI (28.5 vs 25.8). At least 1 EHAD was reported in 41.83% (128/306) of cases and 18.76% (218/1162) of controls (*P*<.001; Table 2).

Autoimmune thyroid disease was the most common EHAD in cases (49/306, 16.01%) and controls (67, 1162, 5.77%). The next most prevalent diseases were Sjögren syndrome (27/306, 8.82%), Raynaud phenomenon (23/306, 7.52%), and psoriasis (22/306, 7.19%) among cases and psoriasis (60/1152, 5.16%), rheumatoid arthritis (46/1162, 4.00%), and type 1 diabetes mellitus (22/1162, 1.89%) among controls. All EHADs were markedly more frequent in cases than controls, except for Crohn disease and type 1 diabetes mellitus (Table 2). Both PBC and PSC were more frequently observed in cases (50/306, 16.34% and 46/1162, 4.00%, respectively) than in controls (2/306, 0.69% and 10/1162, 0.86%, respectively; *P*<.001). Among cases with the AIH-PBC overlap, the most frequent EHAD was autoimmune thyroid disease (12/50, 24%), Sjögren syndrome (8/50, 16%), and psoriasis (7/50, 14%). Cases with the AIH-PBC overlap were more likely than cases without the overlap to have Sjögren syndrome (8/50, 16% vs 15/256, 5.7%; *P*=.05) and

lupus (7/50, 14% vs 15/256, 5.9%; *P*=.02). Among cases with the AIH-PSC overlap, the most frequent EHAD was Crohn diseases (3/12, 25%), ulcerative colitis (2/12, 17%), and psoriasis (2/12, 17%). Cases with the AIH-PSC overlap were more likely to have Crohn disease (3/12, 25% vs 3/294, 1%; *P*=.001) and ulcerative colitis (2/12, 17% vs 12/294, 4.1%; *P*=.06) than cases without the overlap. A comparison of cases with and without EHAD revealed similar demographics. Female sex (93.81% vs 92.12% female), age at AIH diagnosis (47 years vs 42 years), race (88.32% vs 86.02% Caucasian) and BMI (29.2 vs 28.3) were similar between cases with EHAD and cases without EHAD, respectively.

Overall, 6080 FDR autoimmune disease histories were obtained, including 1596 FDRs in cases and 4484 in controls. Cases more frequently had an FDR with at least one autoimmune disease compared with controls (892/1596, 55.89% vs 1601/4484, 35.71%; *P*=.001; Table 2). Autoimmune thyroid disease was the most common reported EHAD in case FDRs (318/1596, 19.93%), whereas rheumatoid arthritis was the most common in controls (244/1596, 15.23%). FDR EHADs, such as autoimmune thyroid disease, celiac disease, ulcerative colitis, psoriasis, Raynaud phenomenon, Sjögren syndrome, and type 1 diabetes mellitus, were more prevalent in cases compared with controls (*P*=.001; Table 2). Autoimmune liver diseases were generally infrequent in FDRs of both groups. PSC was the only autoimmune liver disease seen more in FDRs of cases compared with controls (5/1596, 0.31% vs 0/4484, 0%; *P*=.001).

Table 1. Demographic characteristics of cases and healthy controls.

Characteristics	Cases (n=306)	Controls (n=1162)	P value
Female, n (%)	284 (92.81)	955 (82.18)	<.001
Age (years)			
At the time of study	49.0	33.0	<.001
At the time of diagnosis	44.0	N/A[a]	N/A
Caucasian, n (%)	266 (86.93)	944 (81.23)	.02
Weight (kg), median	76.0	72.6	.01
Body mass index, median	28.5	25.8	<.001
First-degree relatives, n (total=average #/participant)	1596 (5.2)	4484 (3.8)	N/A
Siblings, n	560	1417	N/A
Children, n	519	874	N/A

[a]N/A: not applicable.

Table 2. Concurrent autoimmune diseases among cases and controls as well as first-degree relatives.

Concurrent autoimmune diseases	Cases with condition (n=306), n (%)	Controls with condition (n=1162), n (%)	P value	Cases with an FDR[a] condition (n=306), n (%)	Controls with an FDR condition (n=1162), n (%)	P value
Any extrahepatic autoimmune disease	128 (41.83)	218 (18.76)	<.001	171 (55.88)	415 (35.71)	.001
Autoimmune hepatitis	306 (100.00)	0 (0.00)	NS[b]	14 (4.58)	26 (2.24)	NS
Autoimmune thyroid disease	49 (16.01)	67 (5.77)	<.001	61 (19.93)	89 (7.66)	.001
Celiac disease	16 (5.23)	15 (1.29)	<.001	19 (6.21)	33 (2.84)	.007
Crohn disease	7 (2.28)	13 (1.11)	NS	9 (2.94)	41 (3.53)	NS
Ulcerative colitis	12 (3.92)	12 (1.00)	<.001	25 (8.17)	46 (4.00)	.004
Lupus	14 (4.58)	14 (1.20)	<.001	15 (4.90)	33 (2.84)	NS
Multiple sclerosis	1 (0.33)	7 (0.60)	.008	8 (2.61)	33 (2.84)	NS
Primary sclerosing cholangitis	12 (3.92)	100 (0.86)	<.001	1 (0.31)	0 (0.00)	.001
Primary biliary cholangitis	50 (16.34)	80 (0.69)	<.001	5 (1.63)	13 (1.12)	NS
Psoriasis	22 (7.19)	60 (5.16)	.02	48 (15.69)	109 (9.38)	.002
Raynaud phenomenon	23 (7.52)	21 (1.81)	<.001	58 (18.95)	119 (10.24)	.001
Rheumatoid arthritis	19 (6.21)	46 (4.00)	<.001	56 (18.30)	177 (15.23)	NS
Scleroderma	3 (1.00)	30 (0.26)	.003	2 (0.65)	10 (0.86)	NS
Sjögren syndrome	27 (8.82)	50 (0.43)	<.001	30 (4.25)	53 (0.77)	.001
Type 1 diabetes mellitus	3 (1.00)	22 (1.89)	NS	30 (9.80)	53 (4.56)	.001

[a]FDR: first-degree relative.

[b]NS: not significant.

Table 3 shows further assessment of EHAD among specific FDRs. Compared with controls, there was a higher frequency of EHAD among mothers (95/295, 32.2% vs 274/1110, 24.7%; *P*=.01), siblings (82/600, 13.67% vs 120/1799, 6.67%; *P*=.001), and children (36/411, 8.76% vs 40/874, 4.58%; *P*=.004) of cases. In total, only 3.71% (14/306) cases and 2.14% (27/1162) controls reported an FDR with AIH (*P*=.05). Further depiction of AIH per FDR did not reveal any differences between cases

and controls, yet AIH within mothers appeared more likely among cases (7/295, 2.37% vs 10/1110, 0.90%; *P*=.06).

A logistic regression model was used to assess the risk of concurrent EHAD and EHAD within FDRs of all participants after adjusting for current age, sex, race, and BMI. After the adjustment, both the odds of EHAD (OR 2.46, 95% CI 1.8-3.3; *P*=.001) and EHAD among FDR (OR 2, 95% CI 1.5-2.7; *P*=.001) were significantly higher among cases than among controls.

Table 3. Proportion of autoimmune diseases among first-degree relatives of autoimmune hepatitis cases and healthy controls.

Prevalence of autoimmune disease	Cases	Controls	P value
Extrahepatic autoimmune disease, n (%)			
Mother[a]	95 (32.2)	274 (24.7)	.01
Father[b]	45 (15.3)	130 (11.7)	NS[c]
Siblings[d]	82 (13.67)	120 (6.67)	.001
Children[e]	36 (8.76)	40 (4.58)	.004
Autoimmune hepatitis, n (%)			
Mother	7 (2.37)	10 (0.90)	.06
Father	1 (0.34)	7 (0.63)	NS
Siblings	6 (1.00)	9 (0.50)	NS
Children	0 (0.00)	1 (0.11)	NS

[a]Cases: n=295; controls: n=1110.

[b]Cases: n=295; controls: n=1110.

[c]NS: not significant.

[d]Cases: n=600; controls: n=1799.

[e]Cases: n=411; controls: n=874.

Discussion

This first-ever SNS-supported assessment of EHAD within AIH cases and their FDRs revealed that EHAD is prevalent among both cases and their corresponding FDRs. Specifically, in this study, 41.83% (128/306) of cases and only 18.76% (218/1162) of controls reported an EHAD at the time of the survey completion. Furthermore, 55.88% (171/306) of cases and 35.71% (1601/4484) of controls reported at least one FDR with an EHAD. The assessment of AIH cases with and without EHAD did not demonstrate any demographic differences. The observation of increased EHAD in cases and their FDRs compared with controls remained statistically significant even after adjusting for current age, sex, race, and BMI.

The findings of our case-control study provide a similar estimate of concurrent EHAD in AIH compared with prior descriptive European reports that document the prevalence between 26% and 42% [8,13,14]. In our multivariate model, the odds of an EHAD in cases were almost 2.5 times than that of controls. A detailed comparison of individual autoimmune diseases within cases revealed similar increased EHAD frequencies as previous studies. For instance, we observed autoimmune thyroid disease reported in 16.01% (49/306) of cases, similar to that reported from a large study (9%) of concomitant autoimmune diseases in patients with AIH from the Netherlands [6]. Similarly, both this study and van Gerven et al's [6] study revealed inflammatory bowel disease, rheumatoid arthritis, and celiac disease occurring within 4%-10% of cases. In addition, we provide evidence that patients with overlap disorders (AIH-PBC and AIH-PSC) maintain a high prevalence of concurrent EHAD typically observed in PBC (Sjögren syndrome and psoriasis) and PSC (inflammatory bowel disease). It should be noted that our study was completed at a low cost and within 6 months by a single study team, while the Dutch study required 31 centers to complete. These findings are a testament to the application

of social media tools in medical research, further echoing the sentiment that large studies in rare disease can be completed efficiently and at very low cost.

The observation of increased EHAD within cases is not unexpected, particularly as reports within other autoimmune diseases have described an "autoimmune phenotype," in which autoimmune diseases frequently coexist in an individual. For example, studies in disorders, such as thyroid disease [22], celiac disease [23], and rheumatoid arthritis [24], have revealed high rates of other concurrent autoimmune diseases. This seemingly frequent observation of disease coexistence may be thematically rooted in the *HLA* region on chromosome 6. Variable alleles at this locus have been strongly associated with autoimmune disease risk, yet in some instances, other variants provide protection. The first and only AIH-specific genome-wide association study, to date, has confirmed the *HLA* locus at 6p21 as a region of heightened disease risk [12]. Assessments of other single-nucleotide polymorphisms associated with autoimmune disorders were interrogated in the AIH genome-wide association study, revealing trends of association with AIH even after adjusting for single-nucleotide polymorphisms from the HLA. The shared association of risk alleles within a spectrum of organ-specific autoimmune diseases exemplifies inherent pleiotropic effects of associated genes [25]. Future genome-wide studies will likely provide only a few disease-specific associations among autoimmune diseases, and even more; it remains unlikely that these specific risk loci will clarify the organ specificity. Careful investigation of the exposome, rare genetic variants, and epigenome within cases and their FDRs will provide a higher resolution of the disease risk.

Our interrogation of FDRs of patients with AIH remains the largest interrogation ever with >6000 individual histories, as well as the only case-control assessment. We found a high percentage of EHAD within FDR cases (171/306, 55.88%), compared with only 2 other published reports that examined

EHAD among FDR (42% in the Netherlands [6] and 16% in the United Kingdom [14]). In fact, the odds of having an FDR with EHAD in our cases was double than that of controls in an adjusted model. We argue that the observed frequency of EHAD in FDR cases in this study is accurate, as patients within the sampled social media groups commonly include more autoimmune disease awareness and familiarity with associated conditions [18]. We also believe case attestation to the diagnosis of AIH is reliable, as we have reported a high degree of concordance between reports of patients with AIH and their reviewed medical records within the same digital communities [26]. Furthermore, the often observed female sex and higher educational attainment in patients with AIH has been associated with the increased uniformity between patient reports and medical records [27]. The low rate of EHAD observed in FDRs in the United Kingdom study is likely artificially low because of the study methodology that was founded in the medical record review [14].

Despite a significant difference of any EHAD between case (171/306, 55.88%) and control (1601/4484, 35.71%) FDRs, close to 50% of the 12 assessed EHADs were no different between the FDR groups. Interestingly, the frequency of 2 autoimmune liver diseases, AIH and PBC, were also similar among case and control FDRs in this study. Further examination of autoimmune liver disease in FDRs within each respective disease may help elucidate distinct pathogenetic disparities or similarities among each (PBC, PSC, and AIH). For instance, Mantaka et al [28] and Jones et al [29] observed that between 6% and 10% of PBC cases have an FDR with PBC, respectively. The lower observed frequency of AIH within case FDRs in this study (3.71%) may represent important disease-specific differences in the genetic penetrance or susceptibility associated with the exposome. Moreover, Jones et al [29] found the highest

prevalence of FDR PBC among cases' daughters (2.3%). In this study, AIH was similarly prevalent in cases' mothers (2.37%) and siblings (1%), yet these frequencies were not markedly different from controls.

SNS are attractive research tools, and if implemented cautiously, may help reduce large gaps in rare disease research. There are >1.3 billion users on Facebook alone, and many patients use this as an application for medical information and health-related support from peers [30,31]. These applications are perfectly suited to transcend well-described issues in traditional research methods such as high study costs, coordination of multiple centers, and wide geographic patient distribution [18]. For example, this study cost US $667, was performed at a single center, and included patients from all over North America. SNS may not be suitable for all disease populations; however, in AIH, social media is an easy methodological choice, given a significant amount of demographic overlap.

In summary, this study utilized SNS as a low-cost, effective research method to examine the associations between AIH and other autoimmune diseases among patients and their FDR using a case-control study design. Our data maintain the previously observed autoimmune phenotype of patients with AIH and their family members, such that cases and case FDRs were almost twice as likely as controls and control FDRs to have been diagnosed with, at least, one EHAD. These findings lend further support to an inheritable genetic predisposition underscoring the etiology of AIH. Overall, the FDR prevalence of AIH was found to be quite low among cases, and, thus, broad FDR screening for AIH is not indicated. However, female FDRs of patients with AIH, mother-daughter pairs in particular, may be worth educating of associated symptoms, given a slightly higher risk of the disease development.

Acknowledgments

Conflicts of Interest

None declared.

References

1.   Krawitt EL. Autoimmune hepatitis. N Engl J Med 1996 Apr 04;334(14):897-903. [doi: 10.1056/NEJM199604043341406] [Medline: 8596574]

2.   Vergani D, Longhi MS, Bogdanos DP, Ma Y, Mieli-Vergani G. Autoimmune hepatitis. Semin Immunopathol 2009 Sep;31(3):421-435. [doi: 10.1007/s00281-009-0170-7] [Medline: 19533129]

3.   Lammert C, Loy VM, Oshima K, Gawrieh S. Management of Difficult Cases of Autoimmune Hepatitis. Curr Gastroenterol Rep 2016 Feb;18(2):9 [FREE Full text] [doi: 10.1007/s11894-015-0484-7] [Medline: 26780632]

4.   Toda G, Zeniya M, Watanabe F, Imawari M, Kiyosawa K, Nishioka M, et al. Present status of autoimmune hepatitis in Japan--correlating the characteristics with international criteria in an area with a high rate of HCV infection. Japanese National Study Group of Autoimmune Hepatitis. J Hepatol 1997 Jun;26(6):1207-1212. [Medline: 9210605]

5.   Hurlburt KJ, McMahon BJ, Deubner H, Hsu-Trawinski B, Williams JL, Kowdley KV. Prevalence of autoimmune liver disease in Alaska Natives. Am J Gastroenterol 2002 Sep;97(9):2402-2407. [doi: 10.1111/j.1572-0241.2002.06019.x] [Medline: 12358264]

6.   van Gerven NMF, Verwer BJ, Witte BI, van Erpecum KJ, van Buuren HR, Maijers I, Dutch Autoimmune hepatitis STUDY group. Epidemiology and clinical characteristics of autoimmune hepatitis in the Netherlands. Scand J Gastroenterol 2014 Oct;49(10):1245-1254. [doi: 10.3109/00365521.2014.946083] [Medline: 25123213]

7.   Björnsson E, Talwalkar J, Treeprasertsuk S, Kamath PS, Takahashi N, Sanderson S, et al. Drug-induced autoimmune hepatitis: clinical characteristics and prognosis. Hepatology 2010 Jun;51(6):2040-2048. [doi: 10.1002/hep.23588] [Medline: 20512992]

8.   van Gerven NMF, van der Eijk AA, Pas SD, Zaaijer HL, de Boer YS, Witte BI, Dutch Autoimmune Hepatitis Study Group. Seroprevalence of Hepatitis E Virus in Autoimmune Hepatitis Patients in the Netherlands. J Gastrointestin Liver Dis 2016 Mar;25(1):9-13 [FREE Full text] [Medline: 27014749]

9.   Czaja AJ. Transitioning from Idiopathic to Explainable Autoimmune Hepatitis. Dig Dis Sci 2015 Oct;60(10):2881-2900. [doi: 10.1007/s10620-015-3708-7] [Medline: 25999246]

10.   Ngu JH, Gearry RB, Frampton CM, Stedman CAM. Autoimmune hepatitis: the role of environmental risk factors: a population-based study. Hepatol Int 2013 Jul;7(3):869-875. [doi: 10.1007/s12072-013-9448-x] [Medline: 26201924]

11.   Doherty DG, Donaldson PT, Underhill JA, Farrant JM, Duthie A, Mieli-Vergani G, et al. Allelic sequence variation in the HLA class II genes and proteins in patients with autoimmune hepatitis. Hepatology 1994 Mar;19(3):609-615. [Medline: 8119685]

12.   de Boer YS, van Gerven NMF, Zwiers A, Verwer BJ, van Hoek B, van Erpecum KJ, Dutch Autoimmune Hepatitis Study Group, LifeLines Cohort Study, Study of Health in Pomerania. Genome-wide association study identifies variants associated with autoimmune hepatitis type 1. Gastroenterology 2014 Aug;147(2):443-52.e5. [doi: 10.1053/j.gastro.2014.04.022] [Medline: 24768677]

13.   Muratori P, Fabbri A, Lalanne C, Lenzi M, Muratori L. Autoimmune liver disease and concomitant extrahepatic autoimmune disease. Eur J Gastroenterol Hepatol 2015 Oct;27(10):1175-1179. [doi: 10.1097/MEG.0000000000000424] [Medline: 26148248]

14.   Wong G, Yeong T, Lawrence D, Yeoman AD, Verma S, Heneghan MA. Concurrent extrahepatic autoimmunity in autoimmune hepatitis: implications for diagnosis, clinical course and long-term outcomes. Liver Int 2017 Dec;37(3):449-457. [doi: 10.1111/liv.13236] [Medline: 27541063]

15.   Farmer AD, Bruckner HCEM, Cook MJ, Hearing SD. Social networking sites: a novel portal for communication. Postgrad Med J 2009 Sep;85(1007):455-459. [doi: 10.1136/pgmj.2008.074674] [Medline: 19734511]

16.   Schumacher KR, Stringer KA, Donohue JE, Yu S, Shaver A, Caruthers RL, et al. Social media methods for studying rare diseases. Pediatrics 2014 May;133(5):e1345-e1353 [FREE Full text] [doi: 10.1542/peds.2013-2966] [Medline: 24733869]

17.   Khatri C, Chapman SJ, Glasbey J, Kelly M, Nepogodiev D, Bhangu A, STARSurg Committee. Social media and internet driven study recruitment: evaluating a new model for promoting collaborator engagement and participation. PLoS One 2015 Mar;10(3):e0118899 [FREE Full text] [doi: 10.1371/journal.pone.0118899] [Medline: 25775005]

18.   Ganassi EE, Lyamin EA, Aptikaeva GF, Eidus LK. Statistical analysis of radiation damage to plant chromosomes in relation to the variability of experimental data. Sov Genet 1974 Mar 01;7(12):1533-1539. [Medline: 4827767]

19.   Jun. Running Experiments on Amazon Mechanical Turk Internet. Rochester, NY: Social Science Research Network; 2010. Report No. : ID 1626226 URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1626226 [accessed 2017-12-04] [WebCite Cache ID 6vRoL633Y]

20.   Peer E, Vosgerau J, Acquisti A. Reputation as a sufficient condition for data quality on Amazon Mechanical Turk. Behav Res Methods 2014 Dec;46(4):1023-1031. [doi: 10.3758/s13428-013-0434-y] [Medline: 24356996]

21.   Lammert C, Nguyen DL, Juran BD, Schlicht E, Larson JJ, Atkinson EJ, et al. Questionnaire based assessment of risk factors for primary biliary cirrhosis. Dig Liver Dis 2013 Jul;45(7):589-594 [FREE Full text] [doi: 10.1016/j.dld.2013.01.028] [Medline: 23490343]

22.   Weetman AP. Diseases associated with thyroid autoimmunity: explanations for the expanding spectrum. Clin Endocrinol (Oxf) 2011 Apr;74(4):411-418. [doi: 10.1111/j.1365-2265.2010.03855.x] [Medline: 20718761]

23.   Lauret E, Rodrigo L. Celiac disease and autoimmune-associated conditions. Biomed Res Int 2013;2013:127589 [FREE Full text] [doi: 10.1155/2013/127589] [Medline: 23984314]

24.   Somers EC, Thomas SL, Smeeth L, Hall AJ. Are individuals with an autoimmune disease at higher risk of a second autoimmune disorder? Am J Epidemiol 2009 Mar 15;169(6):749-755. [doi: 10.1093/aje/kwn408] [Medline: 19224981]

25.   Gulamhusein AF, Juran BD, Lazaridis KN. Genome-Wide Association Studies in Primary Biliary Cirrhosis. Semin Liver Dis 2015 Nov;35(4):392-401 [FREE Full text] [doi: 10.1055/s-0035-1567831] [Medline: 26676814]

26.   Comerford M, Fogel R, Bailey JR, Chilukuri P, Chalasani N, Lammert CS. Leveraging Social Networking Sites for an Autoimmune Hepatitis Genetic Repository: Pilot Study to Evaluate Feasibility. J Med Internet Res 2018 Jan 18;20(1):e14 [FREE Full text] [doi: 10.2196/jmir.7683] [Medline: 29348111]

27.   Okura Y, Urban LH, Mahoney DW, Jacobsen SJ, Rodeheffer RJ. Agreement between self-report questionnaires and medical record data was substantial for diabetes, hypertension, myocardial infarction and stroke but not for heart failure. J Clin Epidemiol 2004 Oct;57(10):1096-1103. [doi: 10.1016/j.jclinepi.2004.04.005] [Medline: 15528061]

28.   Hermanek P, Scheibe O, Spiessl B, Wagner G. [TNM classification of malignant tumors: the new 1987 edition]. Radiobiol Radiother (Berl) 1987;28(6):845-846. [Medline: 3444887]

XSL•FO

RenderX

29.   Jones DE, Watt FE, Metcalf JV, Bassendine MF, James OF. Familial primary biliary cirrhosis reassessed: a geographically-based population study. J Hepatol 1999 Mar;30(3):402-407. [Medline: 10190721]
30.   Cline RJ, Haynes KM. Consumer health information seeking on the Internet: the state of the art. Health Educ Res 2001 Dec;16(6):671-692 [FREE Full text] [Medline: 11780707]
31.   Sabanathan S, Hashimi H, Pradhan GN. Transhiatal oesophagectomy in the management of carcinoma of the thoracic oesophagus. J R Coll Surg Edinb 1988 Aug;33(4):192-195. [Medline: 3221337]

Abbreviations

AIH: autoimmune hepatitis
BMI: body mass index
EHAD: extrahepatic autoimmune diseases
FDR: first-degree relatives
*HLA***:** human leukocyte antigen
MTurk: Mechanical Turk
PBC: primary biliary cholangitis
PSC: primary sclerosing cholangitis
SNS: social networking sites

Original Paper

Hepatitis C Prevalence and Management Among Patients Receiving Opioid Substitution Treatment in General Practice in Ireland: Baseline Data from a Feasibility Study

Ross Murtagh[1]; Davina Swan[1], PhD; Eileen O'Connor[1], BSc, RGN, RCN; Geoff McCombe[1], PhD; John S Lambert[1,2], MD, PhD; Gordana Avramovic[1,2], MA; Walter Cullen[1], MD, MB, BCh, BAO

[1]University College Dublin, Dublin, Ireland
[2]Mater Misericordiae University Hospital, Dublin, Ireland

Corresponding Author:
Walter Cullen, MD, MB, BCh, BAO
University College Dublin
Belfield
Dublin, D14 YH57
Ireland
Phone: 353 1 4730893
Fax: 353 1 4544469
Email: walter.cullen@ucd.ie

Abstract

Background: Hepatitis C virus (HCV) infection is a major cause of chronic liver disease and death. Injection drug use is now one of the main routes of transmission of HCV in Ireland and globally with an estimated 80% new infections occurring among people who inject drugs (PWID).

Objective: We aimed to examine whether patients receiving opioid substitution therapy in primary care practices in Ireland were receiving guideline-adherent care regarding HCV screening. Ireland has developed a model of care for delivering opioid substitution treatment in the primary care setting. We conducted this study given the shift of providing care for PWID from secondary to primary care settings, in light of current guidelines aimed at scaling up interventions to reduce chronic HCV infection and associated mortality.

Methods: We included baseline data from the Dublin site of the Heplink study, a feasibility study focusing on developing complex interventions to enhance community-based HCV treatment and improve the HCV care pathway between primary and secondary care. We recruited 14 opioid substitution treatment-prescribing general practices that employed the administration of opioid substitution therapy from the professional networks and databases of members of the research consortium. A standardized nonprobability sampling framework was used to identify 10 patients from each practice to participate in the study. Patients were eligible if aged ≥18 years, on opioid substitution treatment, and attending the practice for any reason during the recruitment period. The baseline data were collected from the clinical records of participating patients. We collected and analyzed data on demographic characteristics, care processes and outcomes regarding HCV and other blood-borne viruses, urinalysis test results, alcohol use disorders, chronic illness, and health service utilization. We examined whether patients received care concordant with guidelines related to HCV screening and care.

Results: The baseline data were collected from clinical records of 134 patients; 72.2% (96/134) were males; (mean age 43, SD 7.6; range 27-71 years); 94.8% (127/134) of patients had been tested for anti-HCV antibody in their lifetime; of those, 77.9% (99/127) tested positive. Then, 83.6% (112/134) of patients had received an HIV antibody test in their lifetime; of those, 6.3% (7/112) tested HIV positive. Moreover, 66.4% (89/134) of patients had been tested for hepatitis B virus in their lifetime and 8% (7/89) of those were positive. In the 12 months before the study, 30.6% (41/134) of patients were asked about their alcohol use by their general practitioner, 6.0% (8/134) received a brief intervention, and 2.2% (3/134) were referred to a specialist addiction or alcohol treatment service.

Conclusions: With general practice and primary care playing an increased role in HCV care, this study highlights the importance of prioritizing the development and evaluation of real-world clinical solutions that support patients from diagnosis to treatment completion.

XSL•FO
RenderX

Introduction

Background

Hepatitis C virus (HCV) infection is associated with considerable morbidity and public health burden globally. Worldwide, it is estimated that approximately 115 million people (1.6% of the world's population) have been infected with HCV, two-thirds of the infections being chronic [1]. The number of people infected with HCV infection in the European Union (EU) and European Economic Area is estimated at 5.6 million [2].

The major burden of HCV infection arises from its progression to chronic infection [3]. Approximately 74% of acutely infected patients progress to chronic infection [4] with 20% developing cirrhosis within 25 years and 25% of patients with cirrhosis developing hepatocellular carcinoma and decompensated liver disease [5,6]. In Europe, HCV is now a leading cause of cirrhosis and primary liver cancer [7].

People Who Inject Drugs and Hepatitis C Virus

Injection drug use is now one of the main routes of transmission of HCV globally, including Ireland, with an estimated 80% of new infections occurring among people who inject drugs (PWID) [8]. The number of HCV-positive cases among PWID is estimated globally at 10 million [9]. In Ireland, a study of anonymized data from the National Drug Treatment Reporting System in 1991-2014 estimated the total number of injectors as 16,382 up to 2014 with an estimated 56% chronically infected with HCV. After adjusting to account for injectors who had never shared injection equipment, it was estimated that 12,423 were infected with HCV with 9317 infections being chronic [10].

HCV infection is mostly asymptomatic in its early stages and the rate of progression is slow with manifest liver disease uncommon within the first 20 years of established infection [11]. As a result, without adequate screening measures, there is marked potential for an elevated burden of disease among an aging population of former or current PWID. Furthermore, injection drug use is associated with high levels of problem alcohol use as well as other factors associated with adverse outcomes [12,13]. In Ireland, a national cross-sectional study reported that 35% of patients attending general practitioners (in Ireland "general practitioner" is the term used for a primary care physician) for opioid substitution treatment also had problem alcohol use [14], whereas findings from a subsequent qualitative study highlighted the need for an educational intervention to address this problem in primary care [15]. Alcohol use disorders accelerate the progression to fibrosis and associated complications; therefore, it is important that alcohol use is addressed within HCV care [16,17]. Guidelines issued by the World Health Organization and the Health Services Executive in Ireland outline the best practices for HCV care, including addressing problem alcohol use [18,19].

In Europe, primary care is increasingly involved in providing continuing care, including opioid substitution treatment, for PWID. Ireland has developed a model of care for delivering opioid substitution treatment, which acknowledges the central role a general practitioner plays in delivering care [20]. Methadone is a long-acting synthetic opioid with documented safety, efficacy, and effectiveness prescribed for opioid substitution treatment [21]. To prescribe methadone in Ireland, general practitioners are subjected to clinical audit and must complete special training with general practitioners providing methadone treatment for ≥15 patients subject to more regular audit and advanced training. General practitioners who prescribe methadone for <15 patients are referred to as "level 1" general practitioners and those prescribing for ≥15 as "level 2" general practitioners [14]. Patients on opioid substitution treatment attend their general practitioner regularly (at least weekly), which offers the general practitioner a chance to monitor for drug and alcohol use and provide education around harm reduction strategies. However, general practitioners have highlighted challenges in providing opioid substitution treatment, including a fear of violence, alienating other patients, lack of knowledge, lack of community supports, and personal reasons such as a lack of empathy with addicts and discomfort with the topic of addiction [20].

Primary Care and Hepatitis C Virus

The important role of primary care in HCV screening and the provision of complex interventions tailored to the needs of PWID have been highlighted previously [22-25]. Although primary care can be a challenging place to deliver interventions given time and resource constraints, the mobility of new HCV diagnostic modalities, including liver stiffness measurement by transient elastography (FibroScan), means it can be transported to community sites to access patients' disease severity. Furthermore, utilizing an HCV liaison nurse in primary care settings can enhance HCV assessment, reduce patient-identified barriers, target therapy, and enable the triage of patients for more immediate care [26-28]. Trials of integrated models of care for HCV treatment (ie, incorporating treatment for mental health issues and substance abuse) in primary health care settings have shown promising results for improving outcomes among PWID [29-31].

The recent emergence of direct-acting antiviral (DAA) treatments has dramatically improved treatment outcomes for patients with HCV; these treatments have shown >90% efficacy in achieving the sustained virological response. The utilization of these new therapeutic agents, combined with improved testing, linkage to care, and adherence to treatment, shows promise for counteracting the expected rise in the future disease burden [32].

Irish and international strategies, plans, and guidelines have prioritized the provision of new pharmaceutical regimens for patients with the greatest clinical need initially, seeking to

XSL•FO
RenderX

balance the high cost of these drugs at present with a view for the wider implementation of these drug regimens in the future [33,34]. In addition, historically HCV screening has been problematic with blood tests and liver biopsy being the standard approach to assess the need for treatment causing many patients to avoid HCV care because of the perceived dangers associated with liver biopsy [35].

A cross-sectional survey of patients attending general practice (in Ireland "general practice" is the name given to the service provided by general practitioners) in the Eastern region of Ireland reported HCV testing rates of 34% and the HCV prevalence rate of 73% among those tested [36]; this low rate of screening for HCV in general practice has also been reported in other health care systems [37,38].

Aims

This study aims to describe the current management of HCV among patients on opioid substitution treatment attending general practice in Ireland in light of the current guidelines aimed at scaling up interventions to reduce chronic HCV infection and associated mortality [39,40]. The data presented in this paper are the baseline data from the Dublin site of the Heplink study [41], which is a feasibility study focusing specifically on developing complex interventions to enhance community-based HCV treatment and improve HCV care pathway in primary and secondary care. The Heplink study is one component of the Hepcare Europe project [41,42], an EU-supported service innovation project and feasibility study at 4 European sites (Dublin, London, Seville, and Bucharest) to develop, implement, and evaluate interventions to enhance the identification and treatment of HCV among PWID.

Methods

Study Design

The research examined the current HCV care practice using a nonprobability sample of patients attending general practices across North Dublin for opioid substitution treatment. Data were collected from clinical records as part of a feasibility study of a complex intervention to enhance HCV care. Outcomes will be discussed relating to blood-borne virus care and problem alcohol use.

The study sites were 14 opioid substitution treatment-prescribing general practices located in Dublin's north inner city. In Ireland, currently, there are 2 types of settings in which opioid substitution treatment is delivered in the community—specialist addiction clinics and general practice. All patients receiving opioid substitution treatment are registered on the Central Treatment List. "Level 1" general practitioners are responsible for the treatment of stabilized opiate-dependent persons referred from specialist addiction clinics or from "level 2" general practitioners. Practice as a "level 1" general practitioner requires the completion of a recognized training program conducted by the Irish College of General Practitioners and regular educational updates. General practitioners are audited by the Irish College of General Practitioners through Health Services Executive Audit Committee. "Level 1" general practitioners can treat up to a maximum of 15 patients. A "level 2" general practitioner

who has undergone additional training can initiate opioid substitution treatment and prescribe for a higher number of patients (up to a maximum of 35 patients or a maximum of 50 in a partnership with ≥2 doctors in their practice [43]). As of August 31, 2016, there were 9652 patients receiving treatment for opiate use in Ireland (excluding prisons), which included 4150 patients being treated by 350 general practitioners in the community [44].

We recruited 14 opioid substitution treatment-prescribing general practices in North Dublin from the professional networks and databases of the research team. Practices were eligible to participate if they were registered to prescribe opioid substitution treatment and were located within the Mater Misericordiae University Hospital (MMUH) catchment area. We wrote to all general practices (n=63) within the referral area of the MMUH Infectious Diseases Department and specified this inclusion criterion in our letter of invitation. We only invited expressions of interest to be returned if practices met this inclusion criterion. A standardized nonprobability sampling framework was used to identify 10 patients from each practice to participate in the study. Standardized nonprobability sampling is commonly used in feasibility studies in which samples are selected on the basis of the subjective judgment of the researcher rather than random selection (ie, probabilistic methods), which is the cornerstone of probability sampling techniques. Based on the recommendations for good practice in feasibility studies [45] and our previous feasibility studies among PWID [46,47] we estimated that 140 patients (attending 14 general practices) would be adequate to calculate the actual recruitment and retention rates (ie, feasibility) and provide data on the acceptability of study processes and outcome measures to inform a future definitive intervention trial. Patients were eligible to participate if they were aged at least 18 years, were on opioid substitution treatment, and attended the practice for any reason during the recruitment period. The researchers instructed participating general practitioners to recruit consecutively presenting patients who were eligible until they had attained a quota of 10.

General practitioners provided eligible patients with a verbal explanation of the study and a written information leaflet outlining the study's purpose, procedures, and how the findings would be utilized. Patients who were interested in participating were asked to sign a consent form, which was witnessed by a general practitioner or a member of the research team. Although the initial approach to participate was from a general practitioner, recruitment was facilitated by a member of the research team being "on site" (where feasible) to support the practice during the recruitment phase and answer any questions that the potential participants might have. Although practices were instructed to recruit 10 eligible patients, some practices recruited <10 because they had a smaller number of opioid substitution treatment patients, and some practices recruited >10 patients. Ethical approval for the study was received from the MMUH Research Ethics Committee.

Data Collection and Analysis

The clinical records of participating patients were examined by a member of the research team prior to the implementation of

XSL•FO

RenderX

the "Heplink" intervention. Baseline data were extracted on demographic characteristics, care processes and outcomes in relation to HCV and other blood-borne viruses (BBVs), urinalysis test results, problem alcohol use, chronic illness, and health service utilization. Variables were chosen to provide data on patients' present and historical HCV care as well as other BBVs (hepatitis B virus, HBV, and HIV). In addition, data were collected on service utilization (emergency department visits and accessing general practice out-of-office hours service) to identify high utilizers and access to care. The baseline data collection instrument was expanded midway through the study to collect additional information on BBV care from the 7 practices (n=60 patients) who participated in the study thereafter. It is our intention to collect this additional baseline information from the other 7 practices when we return to all practices to collect the 6-month postintervention follow-up data and the baseline information will be reported in our pre-post intervention data manuscript which is in preparation. Furthermore, means, frequencies, and percentages were calculated using IBM SPSS Statistics Software version 24.

Blood-Borne Virus Care

Lifetime and past 12-month data on HCV care were extracted from each clinical record, including the following: HCV antibody testing (yes or no) and status (positive or negative); whether a patient had been referred to a hepatology or infectious diseases specialist (yes or no); had attended a hepatology or infectious diseases specialist (yes or no); been assessed by Fibroscan (yes or no); Fibroscan scores (kPa); and initiated HCV treatment (yes or no). In addition, lifetime and past 12-month data in relation to other BBVs were also extracted, including the following: HIV antibody and hepatitis B surface antigen (HbsAg) or hepatitis B core antibody (anti-HBc) testing (yes or no) and status (positive or negative) and whether a patient had received any dose of HBV immunization (yes or no). Additional data on BBV care were extracted on a subset of 7 practices (n=60 patients), including the following: HCV RNA and antigen (Ag) testing (yes or no) and status (positive or negative); whether HCV treatment had been completed (yes or no); and whether a patient had received any dose of hepatitis A virus immunization (yes or no). Furthermore, lifetime and past 12-month data in relation to these variables were collected.

Problem Drug and Alcohol Use

Data extracted from clinical records included results of the last urine drug test, that is, whether positive or negative for metabolites of illicit drugs and non-prescribed benzodiazepines. In addition, clinical records were reviewed to determine whether a general practitioner, in the past 12 months, had screened or discussed alcohol use (yes or no); conducted a brief intervention (yes or no); and referred a patient to specialized treatment (yes or no).

Chronic Illness and Health Service Utilization

Data on the presence of chronic illnesses (yes or no) and whether a patient had any emergency department visits in the past month

(yes or no) or general practice out of hours visits in the past month (yes or no) were extracted.

Results

Sample Characteristics

In this study, 14 general practices and 135 patients were recruited. Although 7 practices were "level 1" prescribers, 7 practices were "level 2" prescribers. The baseline data were collected from the clinical records of 134 patients; of these, 71.6% (96/134) were males, and the mean age of the sample was 43 (range 27-71; SD 7.6) years. In addition, 37.3% (50/134) patients' most recent urine sample had tested positive for metabolites of nonprescribed drugs of abuse. Of note, 36.6% (49/134) had at least one chronic illness documented in their clinical record. Although 2 patients (1/134, 0.7%) had visited an emergency department in the past month. None of the patients had attended a general practice out of hours service in the past month.

Screening for Blood-Borne Viruses

In this study, 94.8% (127/134) patients had been tested for anti-HCV antibody in their lifetime; of those, 77.9% (99/127) tested positive. In the 12-month period before the study, 23.9% patients (32/134) had received an anti-HCV antibody test and 23 (23/32, 72%) of those tested positive.

Next, 112 of 134 (83.6%) patients had received an HIV antibody test in their lifetime; of those, 7 (6.3%) tested HIV positive. Furthermore, 34 (25.4%) patients had been tested for the HIV antibody in the previous 12 months and 1 (1/34, 3%) tested positive.

Furthermore, 66.4% (89/134) patients had been tested for HBV in their lifetime, and 7 (7/89, 8%) of those tested were positive; 22.4% (30/134) had been tested in the previous 12 months, with one (1/30, 3%) testing positive (Table 1).

Immunization Against Other Hepatotropic Viruses

Of 134 patients, 48.5% (65/134) had received at least one dose of HBV immunization in their lifetime, and 8.2% (11/134) had received at least one dose in the 12 months prior to the study. Additional data collected from 7 practices (n=60 patients) showed evidence that 28% (17/60) patients had received at least one dose of Hepatitis A virus immunization in their lifetime with 2% (1/60) receiving the vaccine in the past 12 months (Table 2).

Subsequent Care of Anti-Hepatitis C Virus Antibody-Positive Patients

Of 99 patients known to be anti-HCV antibody-positive, 17% (17/99) patients had undergone a Fibroscan in their lifetime, and 5% (5/99) had undergone a Fibroscan in the previous 12 months. Fibroscan scores were available for 16 patients; the mean score was 7.0 kPa (SD 3.4; range 0.1-16.9). In addition, 20% (20/99) patients had initiated HCV treatment in their lifetime, and 3% (3/99) had initiated treatment in the past 12 months.

Table 1. Blood-borne virus screening and infection status.

Outcome	Total (n=134), n (%)
HCV[a] Ab[b] test—lifetime	127 (94.8)
HCV Ab positive—lifetime	99 (73.9)
HCV Ab test—past year	32 (23.9)
HCV Ab positive—past year	23 (17.2)
HIV Ab test—lifetime	112 (83.6)
HIV Ab positive—lifetime	7 (5.2)
HIV Ab test—past year	34 (25.4)
HIV Ab positive—past year	1 (0.7)
Anti-HBc[c] or BsAg[d] test—lifetime	89 (66.4)
Anti-HBc or HBsAg positive—lifetime	7 (5.2)
Anti-HBc or HBsAg test—past year	30 (22.4)
Anti-HBc or HBsAg positive—past year	1 (0.7)

[a]HCV: hepatitis C virus.

[b]Ab: antibody.

[c]anti-HBc: hepatitis B core antibody.

[d]HBsAg: hepatitis B surface antigen.

Table 2. The hepatitis A and B immunization status.

Outcome	Total, n (%)
Any HBV[a] immunization dose lifetime[b]	65 (48.5)
Any HBV immunization dose past year[b]	11 (8.2)
Any HAV[c] immunization dose lifetime[d]	17 (28.3)
Any HAV immunization dose past year[d]	1 (1.7)

[a]HBV: hepatitis B virus.

[b]n=134.

[c]HAV: hepatitis A virus.

[d]n=60.

Table 3. The management of anti-HCV antibody-positive patients.

Outcome	Total, n (%)
Fibroscanned—lifetime[a]	17 (17)
Fibroscanned—past year[a]	5 (5)
Initiated HCV[b] treatment—lifetime[a]	20 (20)
Initiated HCV treatment past year[a]	3 (3)
HCV Ag[c] or RNA test—lifetime[d]	36 (76)
HCV Ag or RNA positive—lifetime[d]	22 (46)
HCV Ag or RNA test—past year[d]	5 (10)
HCV Ag or RNA positive—past year[d]	2 (4)

[a]n=99.

[b]HCV: hepatitis C virus.

[c]Ag: antigen.

[d]n=47.

XSL•FO
RenderX

Table 4. Screening and intervention for problem alcohol use in the past year.

Outcome	Total (n=134), n (%)
Alcohol screening past year	41 (30.6)
Alcohol brief intervention past year	8 (6.0)
Referred to specialist addiction or alcohol treatment service past year	3 (2.2)

Additional data collected from 7 (n=60 patients) of the 14 practices indicated that among 47 patients who tested HCV antibody positive in these practices, 77% (36/47) had received confirmatory HCV RNA or Ag testing in their lifetime with 47% (22/47) testing positive and 11% (5/47) had received RNA or Ag testing in the previous 12 months with 4% (2/47) testing positive (Table 3).

Alcohol Screening and Brief Intervention

Of 134 patients, 41 patients (30.6%) were asked about their alcohol use by their general practitioner in the 12-month period prior to the study; 6.0% (8/134) had received a brief intervention (ie, a structured discussion around alcohol harms), and 2.2% (3/134) had been referred to a specialist addiction or alcohol treatment service in the same 12-month period (Table 4).

Discussion

Summary

Our findings (Table 5) suggest that although the testing rates for HCV among PWIDs attending general practice for opioid substitution treatment have increased since 2003, access to further assessment and antiviral treatment remains a challenge. Although 94.8% (127/134) of patients had been tested for anti-HCV antibody in their lifetime (of whom 99/127, 77.9% tested positive), only 17.2% (23/134) had had a Fibroscan and 20.2% (27/134) had initiated antiviral therapy. The fact that only 20.2% of patients have commenced antiviral therapy reflects the current high cost of antiviral drugs in Ireland, which prohibits the number of patients who can access treatment. Because the cost of antiviral drugs is predicted to decrease considerably over the coming years, the number accessing antiviral therapy will increase. In addition, it is a concern that 37.3% (50/134) of patients' most recent urine sample had tested positive for metabolites of nonprescribed drugs of abuse; this highlights the need for general practitioners to engage in continued harm reduction education with this cohort of patients. The low immunization rates recorded (48.5%) may be attributed to immunization being carried out elsewhere and therefore not recorded in patients' general practice records. The low screening rates for alcohol use (41/134, 30.6%) are similar to control group rates in a recent study examining screening and brief intervention for alcohol in general practice in Ireland [47] and highlight the need for more general practitioner training.

Comparison With Existing Literature

In this study, 94.8% (127/134) of patients had been tested for anti-HCV antibody in their lifetime with 77.9% (99/127) testing positive; this shows an increase in the number of patients being tested for anti-HCV antibody compared with data collected in 2003 in a study examining HCV among opioid substitution

treatment patients in primary care in Ireland (69%; Table 5) [48]. This increase in testing for the anti-HCV antibody is to be welcomed and reflects the increase in HCV outreach programs and better education of general practitioners in HCV care since the introduction of DAA drugs. The percentage of patients testing positive for HCV is higher than that reported in previous studies on opioid substitution treatment patients in primary care in Ireland [14,48,49] (Table 5) and higher than that in an at-risk cohort from the HepCAT study in primary care in the United States [50]. Furthermore, the percentage of patients testing HCV positive is considerably higher than available data on the prevalence of HCV among injection drug users in Ireland (41.5%) from the European Monitoring Centre for Drugs and Drug Addiction [51].

In this study, 6.3% (7/112) of patients tested positive for HIV, and 8% (7/89) tested positive for HBV (anti-HBc or HBsAg); these rates are lower for HIV but higher than the HBV data reported by Cullen et al (HIV 10%; HBV 4%) [48]. Compared with a more recent study by Klimas et al among opioid substitution treatment patients in primary care in Ireland, the rate for HIV is similar, but that for HBV is higher (HIV 5.7%; HBV 2.7%; Table 5) [49].

In addition, 30.6% of patients had been asked about their alcohol use by their general practitioner in the 12-month period prior to the study, 6.0% had received an alcohol brief intervention, and 2.2% had been referred to a specialist addiction or alcohol treatment service. Although these data show an improvement in screening, brief intervention, and referral to treatment for alcohol compared with the baseline data from a recent alcohol intervention study among opioid substitution treatment patients in primary care in Ireland [49,52], alcohol screening and brief intervention should be systematically performed in this cohort [53].

Limitations

Limitations of this study include the use of a nonprobability sampling strategy; although this results in a lower level of generalizability of research findings and inability to calculate CIs and margins of error, it is an appropriate sampling strategy to use for populations of PWID [54] and when conducting a feasibility study in which lower sample sizes make probability sampling impractical [55]. In addition, potential bias and lack of generalizability may occur from general practitioners who are more motivated and enthusiastic about the issue under study being overrepresented among those recruited. Owing to their interest in the issue, general practitioners who self-selected to the study may be providing better HCV care to their patients than the wider general practitioner population, leading to higher screening rates being detected in the clinical records of participating patients.

Table 5. The blood-borne virus status among patients attending general practice in Ireland for opioid substitution therapy—comparison of 2017, 2013, 2003, and 1999.

Variable	1999	2003	2013	2017
Age (y)	28	32.2	40.9	43
Gender (male), n (%)	409 (72)	141 (72)	69 (65)	96 (72)
Tested for anti-HCV[a] (lifetime), n (%)	380 (67)[b]	151 (77)	104 (99)[c]	127 (95)
HCV positive, n (%)	276 (73)[d]	104 (69)	54 (51)	99 (78)
Tested for HIV (lifetime), n (%)	326 (57)	135 (69)	103 (98.2)[c]	112 (84)
HIV positive, n (%)	27 (8)	14 (10)	6 (5)	7 (6)
Tested for HBV[e] (anti-HBc[f] or HBsAg[g]), n (%)	316 (55)	118 (60)	88 (83)	89 (66)
HBV positive (anti-HBc or HBsAg), n (%)	43 (14)	5 (4)	3 (2)	7 (8)

[a]HCV: hepatitis C virus.

[b]113 of 380 was self-reported data.

[c]Self-reported data.

[d]75 of 276 was self-reported data.

[e]HBV: hepatitis B virus.

[f]HBc: hepatitis B core antibody.

[g]HBsAg, hepatitis B surface antigen.

Furthermore, although only data on patients who consented to a researcher having access to their clinical records were collected, consent bias is likely to be minimal, particularly given the high proportion of those asked who provided consent. In addition, owing to the complex needs of PWID, many will have previously attended a different general practitioner(s) during their lifetime. Therefore, it is unlikely that practice clinical records can accurately capture "lifetime" blood-borne screening. The accuracy of practice clinical records can often be problematic in primary care research because of data not being updated or recorded consistently. However, the rich information contained in patients' clinical records offers considerable potential for research purposes, and our research team aimed to minimize inaccuracies by asking the general practitioner to review the extracted data. Furthermore, data from clinical records do not capture important variables such as socioeconomic variables (income and education) and psychosocial variables (levels of social support), and a data instrument to collect these variables should be added to future research in this area. In addition, the "chronic illness" variable used was dichotomous and therefore hides a lot of information, and this should be addressed for future studies. A further limitation is the comparison data presented in Table 5. There were wide variations in sample sizes between the datasets with some data being self-reported. As such, the conclusions drawn from comparisons of these datasets must be treated with some caution.

Despite these sources of potential bias, the general practitioners and patients who participated in the study were comparable in their profile to other studies from Ireland [48]. Furthermore, this is the first study in Ireland to examine the feasibility of integrated HCV care among problem drug users attending primary care. It will provide positive public health implications with key data to enhance the scientific understanding of interventions that prevent risk behaviors, inform policy and service development, and contribute to the health and social gain locally and internationally. Furthermore, it is the intention of the research team to complement this urban-living population study with similar studies in rural-living populations in Ireland. If the intervention proves feasible, we also intend to scale up the study to other geographical areas recruiting larger sample sizes.

Conclusions

The advent of highly effective DAAs have made eradicating hepatitis C possible, but for this to occur, health care systems must address the complex and wide-ranging difficulties associated with effective HCV screening, assessment, and treatment in the community. With general practice and primary care playing an increased role in HCV care, this study suggests that the development and evaluation of real-world clinical solutions, which support patients from diagnosis to completing treatment, are a priority.

XSL·FO

RenderX

Authors' Contributions

RM developed the manuscript with other coauthors contributing specific components. WC is the principal investigator and conceived the study. RM and WC prepared the manuscript with a core group of authors.

Conflicts of Interest

None declared.

References

1. Gower E, Estes C, Blach S, Razavi-Shearer K, Razavi H. Global epidemiology and genotype distribution of the hepatitis C virus infection. J Hepatol 2014 Nov;61(1 Suppl):S45-S57 [FREE Full text] [doi: 10.1016/j.jhep.2014.07.027] [Medline: 25086286]

2. Bell J, Batey R, Farrell G, Crewe E, Cunningham A, Byth K. Hepatitis C virus in intravenous drug users. Med J Aust 1990 Sep 03;153(5):274-276. [Medline: 2118227]

3. Perz JF, Armstrong GL, Farrington LA, Hutin YJF, Bell BP. The contributions of hepatitis B virus and hepatitis C virus infections to cirrhosis and primary liver cancer worldwide. J Hepatol 2006 Oct;45(4):529-538. [doi: 10.1016/j.jhep.2006.05.013] [Medline: 16879891]

4. Micallef J, Kaldor J, Dore GJ. Spontaneous viral clearance following acute hepatitis C infection: a systematic review of longitudinal studies. J Viral Hepat 2006 Jan;13(1):34-41. [doi: 10.1111/j.1365-2893.2005.00651.x] [Medline: 16364080]

5. Reddy KR, Bruno S, Rossaro L, Ravendhran N, Pauly MP, Boparai N, et al. Predictors of Sustained Virologic Response Among Treatment-NaïVE Patients With Hepatitis C Virus Genotype 1 When Treated With Boceprevir (Boc) Plus Peginterferon Alfa-2B-Ribavirin (PR). Gastroenterology 2011 May;140(5):S-943. [doi: 10.1016/S0016-5085(11)63910-3]

6. Freeman A, Dore G, Law M, Thorpe M, Von Overbeck J, Lloyd A, et al. Estimating progression to cirrhosis in chronic hepatitis C virus infection. Hepatology 2001 Oct;34(4 Pt 1):809-816. [doi: 10.1053/jhep.2001.27831] [Medline: 11584380]

7. Blachier M, Leleu H, Peck-Radosavljevic M, Valla D, Roudot-Thoraval F. The burden of liver disease in Europe: a review of available epidemiological data. J Hepatol 2013 Mar;58(3):593-608 [FREE Full text] [doi: 10.1016/j.jhep.2012.12.005] [Medline: 23419824]

8. Lazarus J, Sperle I, Maticic M, Wiessing L. A systematic review of Hepatitis C virus treatment uptake among people who inject drugs in the European Region. BMC Infect Dis 2014;14 Suppl 6(6):S16 [FREE Full text] [doi: 10.1186/1471-2334-14-S6-S16] [Medline: 25252742]

9. Nelson PK, Mathers BM, Cowie B, Hagan H, Des JD, Horyniak D, et al. Global epidemiology of hepatitis B and hepatitis C in people who inject drugs: results of systematic reviews. Lancet 2011 Aug 13;378(9791):571-583 [FREE Full text] [doi: 10.1016/S0140-6736(11)61097-0] [Medline: 21802134]

10. Carew A, Murphy N, Long J, Hunter K, Lyons S, Walsh C, et al. Incidence of hepatitis C among people who inject drugs in Ireland. Hepatology Medicine Policy 2017 Jan 26;2(1). [doi: 10.1186/s41124-017-0024-1]

11. Seeff LB. Natural history of chronic hepatitis C. Hepatology 2002 Nov;36(5 Suppl 1):S35-S46. [doi: 10.1053/jhep.2002.36806] [Medline: 12407575]

12. Cornford C, Close H. The physical health of people who inject drugs: complexities, challenges, and continuity. Br J Gen Pract 2016 Jun;66(647):286-287 [FREE Full text] [doi: 10.3399/bjgp16X685333] [Medline: 27231287]

13. Onyeka I, Olubamwo O, Beynon C, Ronkainen K, Föhr J, Tiihonen J, et al. Factors associated with hospitalization for blood-borne viral infections among treatment-seeking illicit drug users. J Subst Abuse Treat 2015 Jun;53:71-77. [doi: 10.1016/j.jsat.2015.01.005] [Medline: 25736625]

14. Ryder N, Cullen W, Barry J, Bury G, Keenan E, Smyth BP. Prevalence of problem alcohol use among patients attending primary care for methadone treatment. BMC Fam Pract 2009 Jun 11;10:42 [FREE Full text] [doi: 10.1186/1471-2296-10-42] [Medline: 19519882]

15. Field C, Klimas J, Barry J, Bury G, Keenan E, Smyth B, et al. Problem alcohol use among problem drug users in primary care: a qualitative study of what patients think about screening and treatment. BMC Fam Pract 2013 Jul 13;14(1):98 [FREE Full text] [doi: 10.1186/1471-2296-14-98] [Medline: 23849081]

16. Tumurbaatar B, Tikhanovich I, Li Z, Ren J, Ralston R, Kuravi S, et al. Hepatitis C and alcohol exacerbate liver injury by suppression of FOXO3. Am J Pathol 2013 Dec;183(6):1803-1814 [FREE Full text] [doi: 10.1016/j.ajpath.2013.08.013] [Medline: 24225087]

17. Schwarzinger M, Baillot S, Yazdanpanah Y, Rehm J, Mallet V. Contribution of alcohol use disorders on the burden of chronic hepatitis C in France, 2008-2013: A nationwide retrospective cohort study. J Hepatol 2017 Sep;67(3):454-461 [FREE Full text] [doi: 10.1016/j.jhep.2017.03.031] [Medline: 28396173]

18. Guidelines for the screening, caretreatment of persons with hepatitis C infection.: World Health Organization; 2014. URL: http://www.who.int/hiv/pub/hepatitis/hepatitis-c-guidelines/en/

19. Hepatitis C Screening (NCEC National Clinical Guideline No 15 Summary). 2017. URL: http://health.gov.ie/national-patient-safety-office/ncec/national-clinical-guidelines/

20. Delargy I, O'Shea M, Van Hout MC, Collins C. General Practitioner perspectives on and attitudes toward the methadone treatment programme in Ireland. Heroin Addiction and Related Clinical Problems 2016;18(4):43-50 [FREE Full text]

21. Klimas J, Egan M, Tobin H, Coleman N, Bury G. Development and process evaluation of an educational intervention for overdose prevention and naloxone distribution by general practice trainees. BMC Med Educ 2015 Nov 20;15(2):206-286 [FREE Full text] [doi: 10.1186/s12909-015-0487-y] [Medline: 26590066]

22. Seidenberg A, Rosemann T, Senn O. Patients receiving opioid maintenance treatment in primary care: successful chronic hepatitis C care in a real world setting. BMC Infect Dis 2013 Jan 08;13:9 [FREE Full text] [doi: 10.1186/1471-2334-13-9] [Medline: 23298178]

23. Bechini A, Levi M, Falla A, Ahmad A, Veldhuijzen I, Tiscione E, et al. The role of the general practitioner in the screening and clinical management of chronic viral hepatitis in six EU countries. J Prev Med Hyg 2016;57(2):E51-E60 [FREE Full text] [Medline: 27582629]

24. Artenie AA, Bruneau J, Lévesque A, Wansuanganyi JMB. Role of primary care providers in hepatitis C prevention and care: one step away from evidence-based practice. Can Fam Physician 2014 Oct;60(10):881-2, e468 [FREE Full text] [Medline: 25316735]

25. Artenie AA, Roy E, Zang G, Jutras-Aswad D, Bamvita J, Puzhko S, et al. Hepatitis C Virus seroconversion among persons who inject drugs in relation to primary care physician visiting: The potential role of primary healthcare in a combined approach to Hepatitis C prevention. Int J Drug Policy 2015 Oct;26(10):970-975. [doi: 10.1016/j.drugpo.2015.04.013] [Medline: 25316735]

26. Arain A, De Sousa J, Corten K, Verrando R, Thijs H, Mathei C, et al. Pilot Study: Combining Formal and Peer Education with FibroScan to Increase HCV Screening and Treatment in Persons who use Drugs. J Subst Abuse Treat 2016 Dec;67:44-49. [doi: 10.1016/j.jsat.2016.04.001] [Medline: 27296661]

27. Marshall A, Cunningham E, Nielsen S, Aghemo A, Alho H, Backmund M, et al. Restrictions for reimbursement of interferon-free direct-acting antiviral drugs for HCV infection in Europe. The Lancet Gastroenterology & Hepatology 2018 Feb;3(2):125-133. [doi: 10.1016/S2468-1253(17)30284-4]

28. Grebely J, Morris MD, Rice TM, Bruneau J, Cox AL, Kim AY, et al. Cohort profile: the international collaboration of incident HIV and hepatitis C in injecting cohorts (InC3) study. Int J Epidemiol 2013 Dec;42(6):1649-1659 [FREE Full text] [doi: 10.1093/ije/dys167] [Medline: 23203695]

29. Ho S, Bräu N, Cheung R, Liu L, Sanchez C, Sklar M, et al. Integrated Care Increases Treatment and Improves Outcomes of Patients With Chronic Hepatitis C Virus Infection and Psychiatric Illness or Substance Abuse. Clin Gastroenterol Hepatol 2015 Nov;13(11):2005-14.e1. [doi: 10.1016/j.cgh.2015.02.022] [Medline: 25724704]

30. Mason K, Dodd Z, Sockalingam S, Altenberg J, Meaney C, Millson P, et al. Beyond viral response: A prospective evaluation of a community-based, multi-disciplinary, peer-driven model of HCV treatment and support. Int J Drug Policy 2015 Oct;26(10):1007-1013. [doi: 10.1016/j.drugpo.2015.04.012] [Medline: 26005037]

31. Groessl E, Sklar M, Cheung R, Bräu N, Ho SB. Increasing antiviral treatment through integrated hepatitis C care: a randomized multicenter trial. Contemp Clin Trials 2013 Jul;35(2):97-107. [doi: 10.1016/j.cct.2013.05.002] [Medline: 23669414]

32. Cousien A, Tran VC, Deuffic-Burban S, Jauffret-Roustide M, Dhersin J, Yazdanpanah Y. Hepatitis C treatment as prevention of viral transmission and liver-related morbidity in persons who inject drugs. Hepatology 2016 Apr;63(4):1090-1101. [doi: 10.1002/hep.28227] [Medline: 26390137]

33. Maticic M, Videcnik Zorman J, Gregorcic S, Schatz E, Lazarus JV. Are there national strategies, plans and guidelines for the treatment of hepatitis C in people who inject drugs? A survey of 33 European countries. BMC Infect Dis 2014;14 Suppl 6:S14 [FREE Full text] [doi: 10.1186/1471-2334-14-S6-S14] [Medline: 25252635]

34. European Association for the Study of the Liver. Electronic address: easloffice@easloffice.eu. EASL Recommendations on Treatment of Hepatitis C 2016. J Hepatol 2017 Dec;66(1):153-194. [doi: 10.1016/j.jhep.2016.09.001] [Medline: 27667367]

35. Akay S, Karasu Z, Noyan A, Pala S, Musoglu A, Ilter T, et al. Liver biopsy: is the pain for real or is it only the fear of it? Dig Dis Sci 2007 Feb;52(2):579-581. [doi: 10.1007/s10620-006-9493-6] [Medline: 17211704]

36. Cullen W, Bury G, Barry J, O'Kelly FD. Hepatitis C infection among drug users attending general practice. Ir J Med Sci 2003;172(3):123-127. [Medline: 14700114]

37. Nicklin DE, Schultz C, Brensinger CM, Wilson JP. Current care of hepatitis C-positive patients by primary care physicians in an integrated delivery system. J Am Board Fam Pract 1999;12(6):427-435. [Medline: 10612360]

38. Shehab T, Sonnad S, Jeffries J, Gunaratnum N, Lok AS. Current practice patterns of primary care physicians in the management of patients with hepatitis C. Hepatology 1999 Sep;30(3):794-800. [doi: 10.1002/hep.510300328] [Medline: 10462388]

39. World Health Organization. Global health sector strategy on viral hepatitis 2016-2021. 2016. URL: https://www.who.int/hepatitis/strategy2016-2021/ghss-hep/en/

40. Duffell E, Hedrich D, Mardh O, Mozalevskis A. Towards elimination of hepatitis B and C in European Union and European Economic Area countries: monitoring the World Health Organization's global health sector strategy core indicators and scaling up key interventions. Euro Surveill 2017 Mar 02;22(9) [FREE Full text] [doi: 10.2807/1560-7917.ES.2017.22.9.30476] [Medline: 28277217]

41. McCombe G, Swan D, O'Connor E, Avramovic G, Vickerman P, Ward Z, et al. Integrated Hepatitis C Care for People
     Who Inject Drugs (Heplink): Protocol for a Feasibility Study in Primary Care. JMIR Res Protoc 2018 Jun 04;7(6):e149
     [FREE Full text] [doi: 10.2196/resprot.9043] [Medline: 29866641]

42. Swan D, Cullen W, Macias J, Oprea C, Story A, Surey J, et al. Hepcare Europe - bridging the gap in the treatment of
     hepatitis C: study protocol. Expert Rev Gastroenterol Hepatol 2018 Mar;12(3):303-314. [doi:
     10.1080/17474124.2018.1424541] [Medline: 29300496]

43. Farrell M. Barry J: The introduction of the opioid treatment protocol. In. Dublin: Health Service Executive (HSE); 2010.

44. Health SPRA2. In. Dublin: Health Service Executive (HSE); 2016.

45. Arain M, Campbell MJ, Cooper CL, Lancaster GA. What is a pilot or feasibility study? A review of current practice and
     editorial policy. BMC Med Res Methodol 2010 Jul 16;10:67 [FREE Full text] [doi: 10.1186/1471-2288-10-67] [Medline:
     20637084]

46. Klimas J, Anderson R, Bourke M, Bury G, Field CA, Kaner E, et al. Psychosocial interventions for alcohol use among
     problem drug users: protocol for a feasibility study in primary care. JMIR Res Protoc 2013 Aug 02;2(2):e26 [FREE Full
     text] [doi: 10.2196/resprot.2678] [Medline: 23912883]

47. Henihan AM, McCombe G, Klimas J, Swan D, Leahy D, Anderson R, et al. Feasibility of alcohol screening among patients
     receiving opioid treatment in primary care. BMC Fam Pract 2016 Dec 05;17(1):153 [FREE Full text] [doi:
     10.1186/s12875-016-0548-2] [Medline: 27816057]

48. Cullen W, Stanley J, Langton D, Kelly Y, Bury G. Management of hepatitis C among drug users attending general practice
     in Ireland: baseline data from the Dublin area hepatitis C in general practice initiative. Eur J Gen Pract 2007;13(1):5-12.
     [doi: 10.1080/14017430601049365] [Medline: 17366287]

49. Klimas J, Marie HA, McCombe G, Swan D, Anderson R, Bury G, et al. Psychosocial Interventions for Problem Alcohol
     Use in Primary Care Settings (PINTA): Baseline Feasibility Data. J Dual Diagn 2015;11(2):97-106. [doi:
     10.1080/15504263.2015.1027630] [Medline: 25985200]

50. Southern W, Norton B, Steinman M, DeLuca J, Drainoni M, Smith B, et al. A Birth-cohort testing intervention identified
     hepatitis c virus infection among patients with few identified risks: a cross-sectional study. BMC Infect Dis 2015 Dec
     01;15:553 [FREE Full text] [doi: 10.1186/s12879-015-1283-3] [Medline: 26626449]

51. Ireland: Country Drug Report 2018. Lisbon: European Monitoring Centre for Drugs and Drug Addiction; 2018. URL: http://
     /www.emcdda.europa.eu/countries/drug-reports/2018/ireland_en [accessed 2018-10-26]

52. Klimas J, Anderson R, Bourke M, Bury G, Field CA, Kaner E, et al. Psychosocial interventions for alcohol use among
     problem drug users: protocol for a feasibility study in primary care. JMIR Res Protoc 2013 Aug 02;2(2):e26 [FREE Full
     text] [doi: 10.2196/resprot.2678] [Medline: 23912883]

53. Klimas J, Cullen W, Field C, Problem Alcohol/Drug Use Guideline Development Group. Problem alcohol use among
     problem drug users: development and content of clinical guidelines for general practice. Ir J Med Sci 2014 Mar;183(1):89-101.
     [doi: 10.1007/s11845-013-0982-2] [Medline: 23820987]

54. Thompson S, Collins LM. Adaptive sampling in research on risk-related behaviors. Drug Alcohol Depend 2002 Nov;68
     Suppl 1:S57-S67. [Medline: 12324175]

55. Keiding N, Louis TA. Perils and potentials of self-selected entry to epidemiological studies and surveys. Journal of the
     Royal Statistical Society Series A (Statistics in Society) 2016;179(2):319-376 [FREE Full text]

Abbreviations

Ag: antigen
BBV: blood-borne virus
DAA: direct-acting antiviral
EU: European Union
HAV: hepatitis A virus
HBc: hepatitis B core antibody
HBsAg: hepatitis B surface antigen
HBV: hepatitis B virus
HCV: hepatitis C virus
MMUH: Mater Misericordiae University Hospital
PWID: people who inject drugs