Published in final edited form as:

Int J Med Inform. 2009 November ; 78(11): 772–777. doi:10.1016/j.ijmedinf.2009.07.003.

Comparing Clinician Knowledge and Online Information Regarding Alli (Orlistat)

Stuart Nelson¹, Kevin O. Hwang, MD MPH², and Elmer V. Bernstam, MD MSE^{2,3}

¹Rice University, Houston, Texas, USA

²The University of Texas Health Science Center at Houston, Medical School Department of Internal Medicine, Section of General Internal Medicine, Houston, Texas, USA

³The University of Texas Health Science Center at Houston, School of Health Information Sciences, Houston, Texas, USA

Abstract

Background: Many consumers join online communities focused on health. Online forums are a popular medium for the exchange of health information between consumers, so it is important to determine the accuracy and completeness of information posted to online forums.

Objective: We compare the accuracy and completeness of information regarding the FDA-approved over-the counter weight-loss drug Alli (Orlistat) from forums and from clinicians.

Methods: We identified Alli-related questions posted on online forums and then posed the questions to 11 primary care providers. We then compared the clinicians' answers to the answers given on the forums. A panel of blinded experts evaluated the accuracy and completeness of the answers on a scale of 0 - 4. Another panel of blinded experts categorized questions as being best answered based on clinical experience versus review of the literature.

© 2009 Elsevier Ireland Ltd. All rights reserved.

Corresponding author Elmer Bernstam, MD Associate Professor School of Health Information Sciences University of Texas Health Science Center at Houston 7000 Fannin, Suite 600 Houston, TX 77030 USA Phone: 713 500 3927 Fax: 713 500 3929
elmer.v.bernstam@uth.tmc.edu.

Authors' contributions

Mr. Nelson, Dr. Hwang and Dr. Bernstam all participated in every phase of the work described in this manuscript. Each co-author participated in data collection, data analysis, and manuscript preparation.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conflicts of interest

None.

Summary table

What is known:

- Health care consumers are increasingly turning to the Internet for health information
- Health care consumers and professionals are concerned about the quality and accuracy of health information online
- However, it is not known how online information compares to information that a consumer might obtain from a clinician

What this study adds:

- Unfortunately, both online answers and physician answers are sub-optimal
- Answers to common questions about Alli, an over-the-counter weight loss drug, on an online forum are similar in completeness and accuracy to answers to the same questions provided by physicians

Results: The accuracy and completeness of responses was slightly better than forum responses, but there was no significant difference (2.3 vs 2.1, $p = 0.5$). Only one forum answer contained information that could potentially cause harm if the advice was followed.

Conclusions: Forum answers were comparable to clinicians' answers with respect to accuracy and completeness, but answers from both sources were unsatisfactory.

Keywords

orlistat; obesity; medical education; computer; Internet; medical informatics

Introduction

The Internet has supplanted clinicians as the primary source of health information for the American public [1]. This revolution in information-seeking behavior compels us to compare health information on the Internet to information offered by clinicians. Despite widespread concern regarding the ill effects of online information, there is currently little objective evidence of harm from online health information in the published literature [2]. This lack of objective evidence may be due to an actual lack of harm, or may be due to a lack of documentation (i.e., harm occurs, but is not documented in the published literature) [2].

Weight loss is a common search topic for online health seekers [3] and there are many large online communities (forums) focused on overweight and weight loss. For example, SparkPeople (<http://www.sparkpeople.com>) claims that over five million people have joined their online community [4]. Members of Internet weight loss communities ask each other for advice on many aspects of weight loss, including medications. While we have previously examined the quality of weight loss medication information exchanged on Internet weight loss forums [5], little is known about how the information on forums compares to information a patient might receive from a clinician on the same topic.

Health care consumers require a variety of information. Some information can be found in the biomedical literature, such as the average weight loss observed in clinical trials of Alli. Other information is better obtained from experienced users. For example, where should one buy Alli? Of course, the distinction between literature-based and experience-based knowledge may be subjective and some topics cannot be categorized.

In this study, we compared the accuracy and completeness of responses to questions posted to online health forums with answers to the same questions provided by clinicians. We focused on information related to Alli, the only over-the-counter weight loss medication approved by the U.S. Food and Drug Administration. Since it is available without a prescription, consumers may or may not consult a clinician before taking Alli. Thus, they may obtain information online (such as from a forum), from a clinician, or both. We hypothesized that forum information on Alli accurate and complete. Further, we hypothesized that forum information is complementary to clinician information in the sense that knowledge best gained from experience would be available on forums while knowledge best learned from the scientific literature would be provided by clinicians. We based these hypotheses on our clinical experience and informal review of online information.

Background

Multiple studies addressed various aspects of online health information. However, the lack of precise, shared definitions for concepts in this field makes it difficult to directly compare study results [6]. Investigators have used a variety of definitions for accuracy of clinical information [7-12]. For example, some studies defined accuracy as concordance with a particular gold

standard (e.g., a clinical practice guideline, or something that the authors develop) [12], while others asked one or more experts to use their best judgment to rate the accuracy of information [7]. Thus, it is not surprising that estimates of accuracy vary widely [6].

Similarly, quality has been defined in a variety of ways including accuracy, completeness or concordance with some rating instrument (e.g., HONcode [13]) [14,15]. In this study, we use the term quality to refer to a combination of accuracy and completeness.

Recently, user-generated content, sometimes called “Web 2.0” [16] has become increasingly common. User-generated content can be found on forums, social networking websites, microblogs (Twitter), user-submitted video websites (YouTube), virtual worlds (e.g., Second Life) and others [17]. In contrast to traditional printed or Web content created by a single author, user-generated content is created by users and may even be “self-correcting” [5,18]. Answers to questions posted to online forums appear to be generally accurate [18,19]. However, we are not aware of any study that directly compares answers posted to online forums and information provided by clinicians in response to the same questions.

Methods

This study was conducted in four steps. First, we identified questions posted on online health forums. Second, we posed these questions to clinicians. Third, a panel of three clinical experts evaluated answers from each source. The panel was blinded to the study hypotheses as well as the source of the information (i.e., forums vs. clinicians). Finally, a separate panel of two clinicians unaware of the relevant hypotheses categorized questions into two categories: ones best answered based on clinical experience versus review of the scientific literature. The study was approved by the Committee for the Protection of Human Subjects at UT-Houston; our institutional review board.

On October 18, 2008 we searched Google (<http://www.google.com>) for the string “Alli forum.” We used Google because it is the most heavily used search engine [20]. “Alli forum,” is a general query that is short [21,22] and uses words that directly describe what users are trying to find. Each forum contained multiple threads, and the titles of each thread could be viewed without viewing the postings themselves. We focused on questions contained within the titles of the threads. We selected questions in reverse chronological order using original thread titles (i.e., most recent first). We designed this methodology to avoid the possibility of unintentional bias as we selected questions (e.g., select questions only where forum answers were accurate and complete). We included questions that:

1. Pertained directly to Alli, or the effects of Alli.
2. Were sufficiently focused on a single aspect of Alli that could be answered objectively, and is likely to be asked in a clinical environment. For example, “Do you like Alli?” was not a valid question because there is no objective answer.

The top four non-sponsored forums returned by Google were reviewed for questions. A total of 16 unique questions met the above criteria. Reviewers may (consciously or unconsciously) judge the answer to be worse if they encounter grammar or spelling errors. Therefore, we fixed such errors in forum text. Information that could be used to identify individuals such as names and addresses was also eliminated.

Forum answers to the questions were collected from the threads. There was no limit to the number of answers that were taken. To ensure that the grading panel remained blind to the source of the answers, postings were omitted if they:

1. Contained a question, rather than an answer to the original question.
2. Strayed off topic or were commenting subjectively. For example, answers that insulted another poster or were nonsensical were eliminated.

Answers were stored in the research database.

Next, a pencil-and-paper “quiz” consisting of three questions randomly chosen from the pool of 16 forum questions was administered to 11 clinicians. To ensure that every question was answered, we created six versions of the quiz, and each contained a group of three questions. Fourteen questions were answered by two clinicians each, and two questions were answered by three clinicians each. This difference was due to a discrepancy between number of physicians quizzed and number of questions available. All clinician subjects were primary care providers and members of the teaching faculty in the UT-Houston division of general internal medicine. One was a nurse practitioner and the others were board-certified general internists; all were active primary care providers. The clinicians were given 10 minutes to write their answers to the questions. We chose 10 minutes because previous literature reports that a clinician is likely to have roughly three minutes to address each question that is posed [23]. Before taking the quiz, the clinicians were only told that the study pertained to Alli (Orlistat). The only instructions they were given was to answer the provided questions as best they could, as if they were in a clinical environment. The clinicians were not paid, and were quizzed during a routine faculty meeting. During the quiz the clinicians were not allowed to confer with colleagues, access the Internet, or use any source of information that may help them to answer the questions. The quizzes were then collected and stored in the research database. Even though they wrote their answers, which takes longer than giving verbal answers (as they would in a clinic environment), all clinicians finished within the allowed time.

A panel composed of one general internist, one cardiologist and one endocrinologist was assembled to compare the answers posted on the forums to the answers given by the clinicians. All panel members had extensive clinical experience in treating obesity and are attending clinicians. The task of the panel was to grade the answers for each individual question. Every forum answer that made a direct attempt to answer the posted question was considered a part to that question's answer. The collective forum answers were compared with the collective clinician answers for that particular question. Individual answer postings such as “I don't know” were excluded from the forum answer block. Panelists were asked to evaluate answers in groups (forum or clinician). Forum postings and individual clinician responses were not graded individually.

It was impractical to present the forum answers and clinician answers identically. For example, forum answers were longer than clinician answers. There is no objective way to expand the clinician answers or summarize the forum answers without changing their content. Instead, we masked the study hypotheses from the review panel.

Each panel member was provided with a briefing page describing their task, the questions and answers that they were to grade, a grading guideline, the package insert for Alli, information regarding the safety and efficacy of Alli, and a document with answer guidelines collected from the literature. Panel members worked independently, each grading one third of the collected data. The panel members were told only that they were comparing two different sources. Thus, the panel members were blind to the source of questions, source of the answers and the study hypotheses.

It was made clear to the panel that content (facts), and not verbiage, was to be used in evaluating the answers. Questions were presented with both sets of answers (forum and clinician), but a random number generator was used to determine which answer group was presented first. This

was done to decrease the probability that panelists would identify trends. The task was to grade all answers on a scale from 0-4 using the scale outlined in Table 1.

Grades were simply listed and averaged for each answer grouping. Panel members were also asked to identify any harmful information contained within the answers. In the event of potentially harmful information being found, graders were instructed to identify the harmful answer, and explain why they thought it was harmful. We defined “potentially harmful information” to be information that could cause harm if acted upon by a health care consumer. A two-tailed, paired T-test was used to compare mean scores.

Finally, two additional clinicians who did not otherwise participate in the data collection were asked to categorize the 16 questions listed in Table 2 into one of three categories: 1) best answered using knowledge about science (i.e., answer is likely to be found in the literature), 2) best answered using knowledge based on experience (i.e., someone who has actually prescribed or taken Alli would be best able to answer the question) or 3) can't categorize into 1 or 2. We hypothesized that questions that relied more on personal experience would be answered better by forum participants and that questions that relied more on scientific knowledge would be best answered by clinicians. The clinicians who categorized questions were not aware of the hypothesis. Within both categories of questions (knowledge and experience), we used two-tailed Student's t-tests to compare the response quality ratings for forums vs. clinicians.

Results

Table 2 shows the complete set of ratings for each answer group (clinician and forum) found for each question. The mean clinicians' response quality rating was not significantly different than the mean forum response rating (N=16 questions; 2.3 $\pm$ 1.14 vs. 2.1 $\pm$ 1.01, $p = 0.51$). The only potentially harmful content found was a partial response to the question regarding irritable bowel syndrome, in which a forum poster stated, “I would recommend using what's called ‘colon cleanse’ [...]” There were three answers from doctors that indicated that they held little or no knowledge on the subject. Such answers were given a “1” rating. Panelists were asked to evaluate answers in groups (forum or clinician). Forum postings and individual clinician responses were not graded individually.

The forum answers for every question were longer than the answers that clinicians provided. Doctors' answers were generally limited to one sentence each. Contrary to our hypothesis, “experience” questions were answered better by clinicians than forums, but there was no significant difference (N = 7 questions; mean 2.6 [SD = 0.8] vs. 2.4 [SD = 1.2], $p = .70$). Knowledge questions were also answered better by clinicians than forums but again, there was no significant difference (N = 9 questions 2.3 [SD = 1.2] vs. 1.7 [SD = 1.1], $p = .26$). Interestingly, both groups did better on “experience” questions compared to “knowledge” questions, but there were no significant differences. Clinicians scored higher than forums on eight questions and forums scored higher than clinicians on four questions. The scores were equivalent for the other four questions.

Discussion

We found that forum answers were approximately of the same quality as clinicians' answers to consumer questions regarding Alli. Unfortunately, the average answer score was low for both forums and clinicians. However, only one answer was found on the forums that contained information judged to be potentially harmful.

Because Alli is an over-the-counter drug, it is not entirely surprising that clinicians' knowledge was limited. A mean score of 2.3 suggests that the answers were judged to have “some accurate

information, and [were] mostly incomplete.” Generally, clinicians' answers were vague and imprecise. There were frequent answers that suggested a complete lack of knowledge on the subject (e.g., “I don't know the answer.”).

Our results suggest that information that patients can find regarding Alli on forums is similar in quality to that which they might receive from primary care providers. Some of the forum information was not repeated by clinicians, though we did not formally address this issue. Our hypothesis that forum answers were accurate and complete was not supported by our data. Clinicians and forums gave somewhat better answers for questions that were best answered based on experience rather than the literature.

We designed our study to minimize systematic bias. The panel that graded the answers remained blind to the study hypotheses and did not know the source of the answers they were evaluating. Nor did they know which two groups were being compared. Answer groups were presented randomly to avoid pattern recognition, and questions were randomized, so no panel member graded questions answered by a single clinician.

Two recent studies compared user-generated online information (sometimes referred to as Web 2.0) to more traditional sources of information such as encyclopaedias. Kortum, *et al* evaluated the effects of inaccurate information on the public's knowledge [24]. A group of 34 high school students were instructed to search for vaccine information online and then to answer questions about vaccines. Fifty-nine percent of participants thought that the Internet sites were accurate on the whole, even though over half of the links were inaccurate. About half of participants reported inaccurate statements about vaccines; 24 of 41 verifiable facts were false. Like our study, Kortum, *et al* demonstrated that online health information is not always accurate.

In another study, Clauson, *et al* compared an online open-source encyclopedia (Wikipedia, <http://www.wikipedia.org>) to the traditionally edited Medscape Drug Reference. [25]. The authors found that Wikipedia had a narrower scope, was less complete and had more errors than the traditionally-curated database. In contrast, we found that online forums performed poorly but comparably to clinicians.

This study was novel in two main ways. We analyzed information about a unique drug. Alli is the only FDA approved weight loss medication available over the counter. Since many consumers use weight loss medications, it is important to explore clinician and public knowledge of drugs in this class. A second novel aspect of our study is that it evaluated answers from two alternative information sources to a set of questions actually posed by online information-seekers.

Our study had several limitations. One limitation was that forum answers were generally much longer, while clinicians' answers rarely exceeded one sentence. This was partially addressed by blinding the panel to the hypotheses and to the sources of information. Also, answer presentation was randomized. However, it is possible that clinicians responded hastily, thereby under-representing their actual knowledge on the subject. In addition, some forum posters wrote a lot of extraneous (non-factual) information. A second limitation is that the clinicians were all from the same institution. The results may have been different if we conducted our study with a different group of clinicians or at another institution.

Another limitation is the fact that answers were taken in groups, rather than individually. This was done to make clinician answers comparable in format to forum postings so that the review panel would be less likely to discern the source of information. A consumer accessing a forum may consider multiple answer posts, rather than just one. The problem with this approach, however, is that there were instances of conflicting information contained within one answer group. One member of the panel reported that it was sometimes difficult to give a single all-

inclusive rating to a group of conflicting statements. Our study also may have been under-powered to detect significant differences.

An additional limitation is that the 0-4 scale used for the study does not distinguish between completeness and accuracy of the answer groups. Answers may have received low scores due to factual error, lack of completeness, or both. However, the scale allowed us to address our main concern: “Is the answer a good one?”

Our findings, seen in the context of prior studies, generate questions about the relationship between information found online and information that clinicians impart to their patients. Our primary goal was to compare the quality of forum information to clinician information. We also tried to determine if forum “knowledge” is complimentary to clinician knowledge. We found no such complementary relationship, but we did not perform a content analysis. It is possible that such an analysis will reveal that forum information is actually complementary to clinician information for the same question. Another compelling issue is the degree of information overlap between the two sources. This would suggest which type of information each source is good at providing.

Evaluations of the quality of online information can guide public awareness and education campaigns. If a trend is found (e.g., most side effect information is incomplete and inaccurate, whereas most drug interaction information is deemed to be complete and accurate), then this would indicate the need for more public knowledge of (in this example) side effect information.

In summary, we found no difference between the quality of answers to questions pertaining to Alli displayed by online forums and that which clinicians can provide. In both cases, the information was not entirely accurate or complete. However, we only found one instance of potentially harmful forum information. Given the popularity of online forums focused on health topics, it is important to explore the utility of such forums as a complimentary information source for health care consumers.

Acknowledgements

Thomas R. Lux MD, Rocio A. Cordero MD, Heinrich Taegtmeier MD PhD, Funda Meric-Bernstam MD and Eric J. Thomas MD for their help with the study. Supported in part by the Center for Clinical and Translational Sciences at UT-Houston (NCRR grant 1UL1RR024148). The Center for Clinical and Translational Sciences and the NCRR had no role in the study design, in the collection, analysis and interpretation of data; in the writing of the manuscript; or in the decision to submit the manuscript for publication.

References

1. Hesse BW, et al. Trust and sources of health information: the impact of the Internet and its implications for health care providers: findings from the first Health Information National Trends Survey. *Arch Intern Med* 2005;165(22):2618–24. [PubMed: 16344419]
2. Crocco AG, Villasis-Keever M, Jadad AR. Analysis of cases of harm associated with use of health information on the Internet. *JAMA* 2002;287:2869–71. [PubMed: 12038937]
3. Fox, S. Health information online. Pew Internet and American Life Project; Washington DC: 2005.
4. Why is this free?. [cited 2009 March 8]; Available from: http://www.sparkpeople.com/why_is_this_free.asp.
5. Esquivel A, Meric-Bernstam F, Bernstam EV. Accuracy and self correction of information received from an internet breast cancer list: content analysis. *BMJ* 2006;332(7547):939–42. [PubMed: 16513686]
6. Eysenbach G, et al. Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. *JAMA* 2002;287(20):2691–700. [PubMed: 12020305]
7. Meric F, et al. Breast cancer on the world wide web: cross sectional survey of quality of information and popularity of websites. *BMJ* 2002;324(7337):577–81. [PubMed: 11884322]

8. Bateman, M.; Rittenberg, CN.; Gralla, RJ. Is the internet a reliable and useful resource for patients and oncology professionals: a randomized evaluation of breast cancer information; 34th Annual Meeting of the American Society of Clinical Oncology; Los Angeles, CA. 1998;
9. Biermann JS, et al. Evaluation of cancer information on the Internet. *Cancer* 1999;86(3):381–90. [PubMed: 10430244]
10. Sandvik H. Health information and interaction on the internet: a survey of female urinary incontinence. *BMJ* 1999;319(7201):29–32. [PubMed: 10390457]
11. Griffiths KM, Christensen H. Quality of web based information on treatment of depression: cross sectional survey. *BMJ* 2000;321(7275):1511–5. [PubMed: 11118181]
12. Fricke M, et al. Consumer health information on the Internet about carpal tunnel syndrome: indicators of accuracy. *Am J Med* 2005;118(2):168–74. [PubMed: 15694903]
13. HONcode Principles - Quality and trustworthy health information. [WWW] October 22, 2008 [cited 2009 July 9]; Available from: <http://www.hon.ch/HONcode/Conduct.html>
14. Breckons EV, et al. What do evaluation instruments tell us about the quality of complementary medicine information on the internet? *J Med Internet Res* 2008;10(1):e3. [PubMed: 18244894]
15. Bernstam EV, et al. Instruments to assess the quality of health information on the World Wide Web: what can our patients actually use? *Int J Med Inform* 2005;74(1):13–9. [PubMed: 15626632]
16. O'Reilly T. What is Web 2.0: Design patterns and business models for the next generation of software. 2005
17. Vance K, Howe W, Dellavalle RP. Social internet sites as a source of public health information. *Dermatol Clin* 2009;27(2):133–6. [PubMed: 19254656]
18. Hwang KO, et al. Quality of weight loss advice on internet forums. *Am J Med* 2007;120(7):604–9. [PubMed: 17602934]
19. Hoffman-Goetz L, Donelle L, Thomson MD. Clinical guidelines about diabetes and the accuracy of peer information in an unmoderated online health forum for retired persons. *Inform Health Soc Care* 2009;34(2):91–9. [PubMed: 19412842]
20. Sullivan D. Nielsen NetRatings Search Engine Ratings. 2006[cited 2009 July 9]; Available from: <http://searchenginewatch.com/2156451>
21. Jansen BJ, Spink A, Saracevic T. Real Life, Real Users, and Real Needs: A Study and Analysis of User Queries on the Web. *Information Processing and Management* 2000;36(2):207–227.
22. Silverstein, C., et al. SRC Technical Note. Digital Systems Research Center; Palo Alto, CA: 1998. Analysis of a very large AltaVista query log.
23. Tai-Seale, M.; McGuire, T. Time allocation in primary care with competing demands; *Economics of Population Health: Inaugural Conference of the American Society of Health Economist*; Madison, WI: 2006.
24. Kortum P, Edwards C, Richards-Kortum R. The impact of inaccurate Internet health information in a secondary school learning environment. *J Med Internet Res* 2008;10(2):e17. [PubMed: 18653441]
25. Clauson KA, et al. Scope, completeness, and accuracy of drug information in Wikipedia. *Ann Pharmacother* 2008;42(12):1814–21. [PubMed: 19017825]

Table 1

Grade	Description
1	Answer contains abounding false information and/or makes little attempt to fully resolve the question.
2	Answer contains some accurate information, and is mostly incomplete.
3	Answer provides mostly accurate information and is near-complete.
4	Answer contains all pertinent information and is completely accurate.
0	Answer is not related to the question.

Table 2

Question	Forum Response Cumulative Rating (1-4)	Doctor Response Cumulative Rating (1-4)	Harmful Content?
Can I take less than the recommended amount? **	2	2	No
Can I take a double dose? *	2	3	No
Is Alli only for the obese? **	3	4	No
Can you stop and start treatment freely? *	2	2	No
What are the major side-effects of Alli? **	3	2	No
What happens if you forget to take Alli? *	4	4	No
Does the body get used to Alli? **	2	1	No
How quickly does Alli work? **	No response	2	No
Will my fat come back if I discontinue use? *	4	2	No
Does Alli lower cholesterol? **	1	2	No
Do most people experience side effects? *	1.5	3	No
Does Alli block my vitamins from getting to my body? **	1	4	No
Can you combine Alli and phenteramine? **	0	1	No
Does Alli affect Irritable bowel syndrome? *	1	2	Forum answer, relating to colon cleansing
How long should I wait before eating after taking a pill? *	2	2	No
What are the interactions of blood pressure meds and Alli? ***	3	1	No
Averages (including 0)	2.1	2.3	-

* Question was categorized as being best answered using experience by at least one of two clinicians.

** Question was categorized as being best answered using scientific knowledge by both clinicians

*** Question could not be categorized by at least one of two clinicians.