A reliability and content analysis of Italian language anorexia nervosa-related websites

Nicola Luigi Bragazzi1,4
Giulia Prasso5,6
Tania Simona Re3,4
Riccardo Zerbetto3
Giovanni Del Puente2

1School of Public Health, Department of Health Sciences (DISSAL), University of Genoa, Genoa, Italy;
2Department of Neuroscience, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health (DINOGMI), Section of Psychiatry, University of Genoa, Genoa, Italy;
3Centro Studi Terapia Della Gestalt (CSTG), Milan, Italy;
4UNESCO Chair “Anthropology of Health-Biosphere and Healing System”, University of Genova, Genoa, Italy;
5Department of Psychology (DISFOR), University of Genova, Genoa, Italy;
6School of Psychotherapy “Associazione Italiana di Psicoterapia Psicoanalitica dell’Infanzia dell’Adolescenza e della Famiglia” (AIIPPI), Milan, Italy

Objective: Internet represents a major source of information related to health issues, increasingly used by providers. Indeed, there are numerous websites concerning eating and weight disorders, such as anorexia nervosa. The aim of the current investigation was to systematically perform a reliability and content analysis of Italian language anorexia nervosa-related websites.

Methods: Reliability of the anorexia nervosa website pages was assessed using the Health on the Net Foundation Code of Conduct Standards (HonCode®) certification mark. A comprehensive content analysis (thematic cluster analysis, correspondence analysis of emerging themes and topics, density and degree-centrality analysis, and co-word analysis) was performed using the commercial software T-Lab.

Results: 402 unique website pages were included and analyzed in the current study. Statistically significant differences related to the accomplishment or less to the items of the HonCode® certificate mark were found for all items, except for confidentiality and webmaster/additional contact. From the thematic cluster analysis, four clusters were identified: namely, treatment (26.6%), vomiting (24.7%), weight (24.4%) and onset-age (24.3%). Treatment (in particular, cognitive-behavioral psychotherapy) was the most represented theme (with a frequency of 97.76%, recurring in 393 websites). Pro-ana content characterized 147 website pages (with a frequency of 36.57%).

Conclusion: The quality of Italian language anorexia nervosa-related websites was rather moderate-poor, being generally inconsistent with the principles of the HonCode® certification mark. Therefore, physicians and health authorities should be aware of such findings to provide their patients with proper answers and education.

Keywords: anorexia nervosa, internet, content analysis, qualitative analysis

Introduction

Eating disorders affect approximately 1% of the total population.1 Such prevalence rate raises concerns due to the lack of definitive therapeutic options, even though cognitive-behavioral therapy (CBT) represents the gold standard treatment for bulimia nervosa and binge-eating disorder, whereas family-based treatment (FBT) may be more indicated for adolescents and young adults with anorexia or bulimia nervosa.

The complexity of the disease etiology and the lack of favorable results of pharmacological and psychological treatments utilized may lead patients suffering from eating disorders to seek alternative options through available resources, for instance, the Internet.2

Nowadays, the Internet represents, indeed, a major source of information related to health issues, increasingly used both by providers and clients.3 In particular, numerous
websites concerning eating and weight disorders, such as anorexia nervosa, exist. Lladó et al performed an extensive analysis of websites, analyzing the top 20 results for each search according to their quality and ranking positions. More than a million entries were found with blog contents being the most shared among all the analyzed pages. Authors observed a high rate of contacts and interactions between the different anorexia nervosa-related communities scattered within the entire blogosphere.

However, health information available on the Internet remains unregulated and extremely varies in terms of quality and accuracy, in that, in most cases, it is not written or revised/approved by health professionals. As such, because of the lack of any editorial review policy, websites risk to disseminate and divulge potentially misleading information. Currently, very few studies have assessed such topics. In the existing scholarly literature, Hernández-Morante et al found that the quality of anorexia nervosa-related websites was moderate, and the information quality was rather fairly poor. Some websites, called pro-ana sites, are of particular concern in that they support and promote behaviors and practices that can lead to the development of eating disorders, encouraging a positive view on illness, with some web-pages even denying that anorexia or bulimia nervosa are disorders, being rather defined as lifestyle choices.

On the other hand, high-quality websites manually curated by physicians and workers in the field of public health can offer valuable support for patients with anorexia nervosa, as well as for patients suffering from other eating and weight disorders, making different health care services more accessible, providing web-based treatments and psycho-educational interventions or enabling to create self-help virtual groups.

As such, patients could visit a range of websites that offer different kinds of information of different levels of scientific evidence, from personal accounts and blogs to academic websites including sites curated by the WHO, the Centers for Disease Control and Prevention or scholarly repositories and bibliographic thesauri like PubMed. Therefore, while it is empowering for individuals to access this important source of health-related information, which provides more assistance in the shared decision-making process between the physician and the patients, the quality of health information is highly variable. Thus, in most instances, individuals could be faced with the challenge of understanding which website offers high quality, evidence-based information.

Therefore, given the relationship between media exposure and eating and weight disorder symptom severity, and considering that the Internet provides easily accessible resources and treatment modalities for individuals who otherwise lack access to evidence-based care, it is of crucial importance to assess the content and the material present on the web. As such, the aim of the current investigation was to systematically perform a reliability and content analysis of Italian language anorexia nervosa-related websites. This would be beneficial both for patients and physicians, in that would offer them a snapshot of the information available online, helping to develop and implement ad hoc interventions for patients education and empowerment.

Materials and methods

Search strategy

Seven search engines (namely, Google, Yahoo, Virgilio, Microsoft Network or MSN, Ask, Ixquick, and Iasse), among the most popular in Italy, were selected for the current investigation. The terms “anorexia” and “anorexia nervosa” (in the Italian language, “anoressia” and “anorexia nervosa”) were used every time in individual searches for each search engine, selecting the first 100 results. Duplicate or not relevant website pages were not retained for the subsequent analyses. Further details related to the search strategy adopted are listed in Table 1.

Reliability assessment

Health on the Net Foundation Code of Conduct Standards (HonCode®) certification mark is based on eight principles,

<table>
<thead>
<tr>
<th>Search strategy item</th>
<th>Search strategy details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Search engines</td>
<td>Google, Yahoo, Virgilio, MSN, Ask, Ixquick, Iasse</td>
</tr>
<tr>
<td>Used keywords</td>
<td>Anoressia, anoressia nervosa (Italian, anorexia, anorexia nervosa)</td>
</tr>
<tr>
<td>Time filter</td>
<td>December 2014</td>
</tr>
<tr>
<td>Language filter</td>
<td>Italian</td>
</tr>
<tr>
<td>Inclusion criteria</td>
<td>Pertinent to the eating disorder “anorexia nervosa,” appearing among the first 100 results returned by the search engine</td>
</tr>
</tbody>
</table>
namely, authority (whether the website provides the qualifications and credential of authors: if, for example, they are medical professionals or not), complementarity (that is to say, a statement that maintains that information provided to readers, such as help, tips, advices or suggestions, is meant to support, not replace, patient–health care provider relationships), confidentiality (that is to say, how both personal and non-personal data of readers are handled), attribution (easily discoverable references to the source of information, such as the uniform resource locator or URL address, and information about the last update performed), justifiability (whether statements related to medical procedures or products are conveyed in terms of balanced, well-supported and well-referenced information), transparency of authorship (in terms of contact information, such as mail address or telephone number), financial disclosure (providing clear, exhaustive funding details) and advertising (information concerning advertising and editorial policy, and a clearly stated distinction between advertised and editorial material).

HonCode certification mark-based reliability assessment has been performed in the existing literature to evaluate the quality of otitis media-, silicosis-, orthopedics-, and cancer-related website pages, among others.

Content analysis

Content analysis was performed using T-Lab software, which exploits a text-driven automatic approach allowing meaningful patterns of words and themes to emerge. The content of the unique websites included in the current study was copied, manually pre-processed and computationally analyzed: thematic cluster analysis (which is based on the results obtained with the correspondence analysis and utilizes the K-means clustering technique, performed according to the MacQueen’s unsupervised learning algorithm), correspondence analysis of emerging topics and themes, density and degree-centrality analysis, and co-word analysis (when two or more keywords representing an emerging theme/topic occur on the same website page, they have essential relationships, and the more co-occurrence between two keywords, the closer their relationship is) were carried out. Each thematic cluster contained a set of elementary contexts/structures (namely, sentences, short texts, and paragraphs) characterized by the same linguistic/lexical patterns and keywords. Clusters were described through lexical units (that is to say, words, lemmas, or categories) and the variables/features most characteristic of the context units they were composed of.

For the cluster analysis, the following objective function was used:

$$J = \sum_{j=1}^{k} \sum_{i=1}^{n} \left| x_i^{(j)} - c_j^2 \right|$$

where $x_i^{(j)}$ is a given data point/object and $c_j$ is the cluster center. The MacQueen’s algorithm comprises four steps: namely, placing $k$ points within the space represented by the data points/objects that are being clustered, as initial group centroids (step I); assigning each data point/object to the group with the closest centroid (step II) and recalculating iteratively the positions of the $k$ centroids, once the procedure of assigning the objects has been completed (step III). Steps II and III are repeated until the stability of centroids is achieved, in such a way to obtain a separation of the data points/objects into groups, to minimize and to compute the final metrics (step IV).

After individuating the main thematic clusters, the elementary contexts/structures of each cluster were used to extract the major narratives contained within each cluster. Emerging thematic nuclei/topics were co-mapped together using a multidimensional scaling approach, based on the Sammon mapping technique, utilizing the Sammon’s error or Sammon’s stress function:

$$E = \frac{1}{\sum_{i<j} d_{ij}} \sum_{i<j} \left( d_{ij}^* - d_{ij} \right)^2$$

where $d_{ij}^*$ is the distance between the $i^{th}$ and the $j^{th}$ object (theme/topic) within the original space and $d_{ij}$ is the distance between their projections. Summarizing, $E$ is a function that enables to interpret both the relationships between the different data points/objects and the dimensions that organize the space in which the given structures are represented.

T-Lab software is a validated, reliable tool, which has been used in the existing literature to perform content/lexicographic analyses of cancer- or environmental issues-related patient narrative and website pages, respectively.

The entire process was computer-assisted, with the involvement of psychologists, psychotherapists, and psychiatrists ensuring meaningful results.

Statistical analysis

To determine the websites quality, the frequency and percentage of the number of websites that met every item of the HonCode certification mark were calculated.
Exact chi-squared test was performed to analyze the statistical significance related to the presence or absence of the HonCode® certification mark.

The significant level was established at a $P<0.05$. All statistical analysis was performed through the commercial software “Statistical Package for Social Sciences” (SPSS Inc., Chicago, IL, USA) version 23.0 for Windows.

**Results**

From an initial list of 1400 websites, 402 unique website pages were included and analyzed in the current study (Figure 1). Statistically significant differences related to the accomplishment or less to the items of the HonCode® certification mark were found for all items, except for confidentiality and webmaster/additional contact (Table 2). Of the 8 procedural principles, 34.08% of the websites followed the authority principle, 86.07% followed the complementarity, 56.72% followed confidentiality, 15.67% complied to the justifiability concept, 29.35% to sources/date of information, 43.53% to attribution, 23.13% to financial disclosure, and 21.89% to advertising policy.

From the thematic cluster analysis, the MDS Sammon mapping of main themes/topics, the density and degree-centrality analysis, of the correspondence analysis of emerging themes/topics of all the anorexia nervosa-related website pages included in this study, four main clusters were identified: namely, treatment (26.6%), vomiting (24.7%), weight (24.4%) and onset-age (24.3%). The main topics are reported in Table 3 and pictorially shown in Figure 2: treatment, in particular CBT, was the most represented theme (with a frequency of 97.76%, recurring in 393 websites). Pro-ana content characterized 147 website pages (with a frequency of 36.57%).

**Discussion**

The present study aimed to assess Italian language websites on anorexia nervosa to determine their general as well as information-specific quality. Concerning the content, we found that the most represented topic was anorexia nervosa treatment (in particular, CBT), and furthermore, it was the first cluster in terms of weight.

This finding is perfectly in line with what was observed by other scholars. For instance, in Israel, Yom-Tov and coworkers investigated members of the main pro-ana website and found them significantly more interested in treatment, having wishes of procreation and reporting the highest goal weights. A minority exhibited interest in depression, self-harm and suicide, with a small percentage being severely malnourished. In Italy, Abbate Daga et al found 546,000, 212,000 and 39,100 websites devoted to anorexia nervosa treatment, anorexia nervosa psychotherapy, and anorexia nervosa pharmacotherapy, respectively. “Pro-anorexia,”

---

**Table 2** Compliance of the anorexia nervosa-related website pages included in the current study to the items of the HonCode certification mark and statistical significance

<table>
<thead>
<tr>
<th>HONcode item</th>
<th>Yes</th>
<th>No</th>
<th>Statistical significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authority</td>
<td>137 (34.08%)</td>
<td>265 (65.92%)</td>
<td>$P&lt;0.0001$</td>
</tr>
<tr>
<td>Complementarity</td>
<td>346 (86.07%)</td>
<td>56 (13.93%)</td>
<td>$P&lt;0.0001$</td>
</tr>
<tr>
<td>Confidentiality</td>
<td>228 (56.72%)</td>
<td>174 (43.28%)</td>
<td>$P&lt;0.0001$</td>
</tr>
<tr>
<td>Sources/date of information</td>
<td>118 (29.35%)</td>
<td>284 (70.65%)</td>
<td>ns</td>
</tr>
<tr>
<td>Justifiability</td>
<td>63 (15.67%)</td>
<td>339 (84.33%)</td>
<td>ns</td>
</tr>
<tr>
<td>Webmaster/additional contact</td>
<td>175 (43.53%)</td>
<td>227 (56.47%)</td>
<td>$P&lt;0.0001$</td>
</tr>
<tr>
<td>Sponsor</td>
<td>93 (23.13%)</td>
<td>309 (76.87%)</td>
<td>ns</td>
</tr>
<tr>
<td>Advertising policy/policy</td>
<td>88 (21.89%)</td>
<td>314 (78.11%)</td>
<td>$P&lt;0.0001$</td>
</tr>
</tbody>
</table>

**Abbreviation:** ns, not statistically significant.
“pro-ana sites,” “thinspiration” and “anorexicnation” resulted in 257,000, 18,600, 14,200 and 577 websites, respectively. Forty-seven of 100 randomly selected pro-ana websites were found to be often visited. Authors concluded that the Internet had a Ying-Yang nature, making, on the one hand, some topics such as treatments more available and accessible to the general public, but on the other hand, potentially encouraging psychopathological traits like asceticism, competition, purging behaviors and obsession for control.

Pro-ana virtual communities tend, indeed, to create strong ties and links on the Internet, exhibiting values such as support, solidarity and sense of belonging.25–28 Yom-Tov et al29 analyzed 242,710 anorexia nervosa-related photos shared via the photo-sharing site Flickr from 491 users and found that members of the pro-anorexia digital communities tended to interact with each other to a much higher degree than what could be expected from the distribution of contacts (only 59–72% of contacts versus 74–83% of comments made to members inside the community). Being actively engaged online (namely, commenting and posting observations and likes) elicits, in its turn, further digital activities. In our analysis, we found that pro-ana websites represented approximately one-third of the online material. This finding can be compared with that obtained by a research carried out in Italy by Bert and collaborators.30 They systematically retrieved 341 “pro-ana” Twitter accounts and analyzed them in terms of number of followers (mean followers 2360.9, range 5–32,700), tweets (mean tweets: 4351.2, range: 5–85,700), and biographical information of the users (97.9% girls, mean age 17.9 years, range 12–28 years). Authors concluded that pro-ana content on Twitter was rather popular and, as such, potentially dangerous.

Another topic highly represented in the websites analyzed in the present study was body dissatisfaction and diet/weight control. The media, indeed, tend to divulge images of abnormally thin models represented as the ideal.31

In conclusion, at the same time, the Internet represents an important and precious resource, as well as a source of risks.32 According to a survey conducted in 25 European countries,33 the Internet is increasingly used to surf for health-related content and the exposure of adolescents to pro-ana websites is rather high and relevant. Since different investigations have found a link between online behaviors and offline eating and weight disorder behaviors,34,35 stakeholders should urgently prioritize results of search engines not based on popularity ranking but on trustworthiness and quality.36–38

Our investigation suffers from some limitations: for example, it does not investigate some interesting features, like the ranking and rankability or readability of the website pages. On the other hand, it has a number of strengths, including the systematic search of anorexia nervosa-related website pages.

### Conclusion

The present systematic web searches-based study found that quality of Italian language anorexia nervosa-related websites was rather moderate-poor, being generally inconsistent with the principles of the HonCode® certification mark. The major themes/topics represented in the websites included in the analysis were treatment and management of anorexia nervosa. This has important, practical
implications for health care workers working with patients suffering from eating and weight disorders. They should be, indeed, aware of the importance of the Internet in disseminating anorexia nervosa-related information and should discuss this point with their patients, trying to educate and empower them, showing them how to recognize high-quality, reliable and evidence-based websites. Moreover, the information provided in the present investigation could be exploited to develop ad hoc online resources for patients suffering from eating and weight disorders, as well as novel web-based treatments and psycho-educational interventions.

However, based on the above-mentioned shortcomings and drawbacks, further research in the field is urgently needed.

Ethical standards
No informed consent was necessary for the realization of the study, since all data analyzed are publicly available and have been handled at aggregated level.

Disclosure
The authors report no conflicts of interest in this work.

References

Figure 2 Cluster analysis of all the anorexia nervosa-related website pages included in this study. Factor 1 (treatment), factor 2 (vomiting), factor 3 (onset-age), and factor 4 (weight).


35. Branley DB, Covey J. Is exposure to online content depicting risky behavior related to viewers’ own risky behavior offline? *Comput Hum Behav*. 2017;75:283–287. doi:10.1016/j.chb.2017.05.023


**Risk Management and Healthcare Policy**

**Publish your work in this journal**

Risk Management and Healthcare Policy is an international, peer-reviewed, open access journal focusing on all aspects of public health, policy, and preventative measures to promote good health. The journal welcomes submitted papers covering original research, basic science, clinical & epidemiological studies, reviews and evaluations, guidelines, expert opinion and commentary, case reports and extended reports. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit http://www.dovepress.com/testimonials.php to read real quotes from published authors.