

# The process of developing a content analysis study to evaluate the quality of breastfeeding information on the Internet-based media

Methodological Innovations  
May-August 2019: 1–6  
© The Author(s) 2019  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/2059799119863286  
journals.sagepub.com/home/mio

Juliana Cristina dos Santos Monteiro<sup>1</sup> and Solina Richter<sup>2</sup>

## Abstract

The Internet offers a powerful network of information on breastfeeding that is used by doctors, patients, and scientists. The objective of this study is to describe the process of development of a data extraction tool to evaluate the content and quality of breastfeeding information on the Internet. Using a descriptive study method, we examined Internet pages to determine which variables needed to be measured in order to develop the data extraction tool. A purposive sampling of websites was selected to pilot test this tool. The developed data extraction tool has a descriptive structure to characterize websites and text pages. Using the developed tool, we can assess whether the information on text pages is supportive of breastfeeding and whether other strategies that protect breastfeeding are followed. The developed data extraction tool is a useful instrument that can assist researchers in evaluating the quality of information posted on the Internet related to breastfeeding.

## Keywords

Breastfeeding, communications media, content analysis, Internet, maternal health

## Introduction

Breastfeeding is universally recognized as the optimal method of infant feeding as it improves children's health and mothers' health. It contributes to the human capital development in high-, middle-, and low-income countries (Rollins et al., 2016; Victora et al., 2016). The World Health Organization (WHO) recommends that all babies should be exclusively breastfed until 6 months of age. After this period, breastfeeding should be complemented with solid foods up to 2 years of age or over (WHO and United Nations Children's Fund, 2009).

However, in spite of scientific evidence regarding the benefits as well as supportive guidelines, such as the International Code of Marketing of Breast-Milk Substitutes (WHO, 2017), breastfeeding practices are seriously impaired for several reasons. These reasons include the devaluing of breastfeeding by health professionals; inappropriate hospital practices, particularly in the postpartum period; pressure from the commercial milk and food product industries to use their products; changes in women's lifestyles; and the lack of

supportive advice and information (Dykes, 2011). The introduction of supplementary milk products and food are frequently too early and children and their mothers often do not benefit fully from breastfeeding (Laantera et al., 2011). Furthermore, it is important to highlight that breastfeeding is not only biologically conditioned but also socially, culturally, and individually constructed. In this sense, breastfeeding practices adapt according to the cultural and contextual circumstances of each population and country and also follow different norms in each region (Cattaneo, 2012; Rollins et al., 2016).

<sup>1</sup>University of São Paulo at Ribeirão Preto College of Nursing, Ribeirão Preto, Brazil

<sup>2</sup>Faculty of Nursing, University of Alberta, Edmonton, AB, Canada

## Corresponding author:

Juliana Cristina dos Santos Monteiro, University of São Paulo at Ribeirão Preto College of Nursing, Avenida Bandeirantes, 3900, SP 14040-902 Ribeirão Preto, Brazil.

Email: jumonte@eerp.usp.br



Nowadays, with the introduction of the Internet, a new medium has developed for women to seek information about and support for breastfeeding. A new generation of mothers live “online,” continuously connected to their computers, tablets, and smartphones. They often use more than one device at the same time (Wolynn, 2012). The World Wide Web (WWW) is a “global and decentralized network of hyperlinked multimedia objects” that can be accessed by the Internet (Weare and Lin, 2000, p. 272). The Internet is a powerful network of information and is increasingly used by doctors, patients, and scientists for health care information (Hughes et al., 2008). While a variety of sites and Internet pages offer breastfeeding information, it is unclear how mothers use the Internet to obtain information on breastfeeding their infants (Giglia and Binns, 2014). In contrast to other themes related to motherhood, the use of the Internet to support breastfeeding is a relatively new health education intervention method in an area that traditionally has been conducted face-to-face and one-on-one (Giglia and Binns, 2014).

The benefits of using Internet-based media for seeking health information are many, but there are also some limitations regarding the quality of the information and the identification, reliability, and privacy of the authors. Potential health risks can arise from incorrect or unreliable information or advice (Moorhead et al., 2013). It is important to evaluate the information available on the Internet and to teach mothers how to judge the quality of the information. The use of the Internet as a health information source for patients is not commonly questioned or attended to by nurses, midwives, and other workers that are frontline health care providers. Indeed, it is important to consider the role of health care providers in breastfeeding support since it is culturally specific and implemented by different professionals. Therefore, they need to be able to evaluate the content and quality of health information on the Internet.

This article presents the process of conducting a content analysis study and the development of a data extraction tool that was used to evaluate the content of breastfeeding information obtained from the Internet. Neuendorf (2002) defines content analysis as “a summarizing, quantitative analysis of messages that relies on the scientific method (including attention to objectivity-intersubjectivity, a priori design, reliability, validity, generalizability, replicability, and hypothesis testing)” (p. 10).

Content analysis recognizes that media texts have different meanings to different readers; thus, this research technique tries to define the meaning of texts that are aimed at the target population (Macnamara, 2005). In this sense, applying content analysis to Internet-based media is a challenge, since the rapid growth and change of the web can complicate the definition of the research question, the decisions on the unit of analysis, and the sampling, coding, and training of coders. Nonetheless, this well-developed research technique can be adapted to a dynamic communication environment such as the web (McMillan, 2009).

In this study, the terms “site,” website, and main site are used to refer to a set of text pages on the Internet. “Text page” is defined as hypertexts accessible on the Internet (hypertext being a text aggregated from other information in the form of blocks of texts, words, images, or sounds, which are accessed through specific references, termed hyperlinks, or simply links on the Internet) (Site, 2014).

## Methodology

### Design

This is a descriptive study about the process of conducting a content analysis and the development of a data extraction tool. We will supplement the description with how it was applied in our study, which focused on conducting a content analysis and quality check of the information provided on the Internet related to breastfeeding. As this research was a partnership between Brazilian and Canadian researchers, we considered the Brazilian and Canadian Internet context. In Brazil, an increase in breastfeeding rates over the last decades is achieved due to the government investments and civil society participation (Rollins et al., 2016). According to the latest national survey in 2009, exclusive breastfeeding for 6 months or less was 41% (Ministério da Saúde, Brazil, 2009). In Canada, where the majority of women who had birth assistance in the hospital had support to initiate breastfeeding, exclusive breastfeeding for 6 months or more improved from 17% in 2003 to 26% in 2011–2012 (Gionet, 2013). It was relevant to compare breastfeeding information between a high- and a low-middle-income country.

### Ethical considerations

According to Resolution 466/12 of the Brazilian National Health Council, research that involves only public data and does not identify survey participants does not require approval by the Ethics Committee for Research. We followed that resolution since the first author is a Brazilian researcher. However, in this case, although ethical approval was not required, the discussion of our findings was respectful and considered ethical principles.

### Steps in conducting content analysis

The study to evaluate the content and quality of breastfeeding information on the Internet followed the steps suggested by McMillan (2009) to perform a content analysis of information available on the WWW. McMillan proposed the following steps: (1) formulate a research question, (2) select a sample, (3) define categories or the context units and develop the data extraction tool, (4) train coders and evaluate their reliability, and, finally, (5) analyze and interpret data.

## Formulating the research question

The research question for this study was “What is the content and quality of information regarding breastfeeding that is made available on Brazilian and Canadian Internet pages and accessed by laypersons, in this case breastfeeding mothers, using the Google search mechanism?”

## Selecting a sample

In content analysis, it is important to decide on the units of sampling. It can be, for example, time periods, dates, search engines, or television channels (Neuendorf, 2002). We used purposive sampling to identify sites on the Internet to pilot test the data extraction tool using the Google search engine. This technique allows the identification and storing of Internet pages to be used for analysis in greater depth (Ferreira et al., 2013; Hargrave et al., 2006; Wang et al., 2012). Sample selection was performed on the Brazilian and the Canadian pages of the Google search mechanism using the keywords *amamentação* (breastfeeding) (Brazilian page of Google) or *breastfeeding* (Canadian page of Google). The first five results in the Portuguese language and first 10 results in the English language were stored using the Uniform Resource Locator (URL), which is the address of a resource, such as a file, available on the Internet. Each URL was copied and saved to ensure the same analysis by the various coders.

In a content analysis of the web, it is necessary to define the time frame for analysis since Internet content can change quickly (McMillan, 2009). Data were selected and collected in one single day using the Google search mechanism. Furthermore, it is necessary to identify the context units for coding. In our study, the context unit defined was the Internet text pages. When the URL found on Google corresponded to the sites' main page and did not have text but only links, the information on breastfeeding was sought on the first link which had the word *amamentação* or *aleitamento materno* (sites in Portuguese) or *breastfeeding* or *breastfeeding advice* (sites in English). The Internet text pages were selected in accordance with eligibility criteria: to be available in Portuguese or English and with unrestricted access. The exclusion criteria were as follows: pages which were difficult to access; duplicate pages; those with information only in the form of files for downloading (such as presentations) or only audio and video links (such as YouTube); discussion and question-and-answer forums; pages marked as advertisement sites; and news clips about breastfeeding. Considering the magnitude and changing nature of the websites (McMillan, 2009), some information present in the main site was also collected so that the main sites and the creators of the site could be characterized.

## Defining categories or the context units

Development of the data extraction tool started in October and finished in December 2014. We went back to our research

question and examined web pages to determine which variables needed to be measured to answer it. As we continued to test the data extraction tool and variables to be collected, we added additional variables.

We considered content from selected text pages, identifying and noting the purpose of the main page and the essential purpose of the text page. We checked the quality of the information by comparing the content on the text pages with the recommendations for breastfeeding by WHO and the Brazilian and Canadian guidelines. Pictures and drawings within the text pages potentially were an issue since they could agree or conflict with official documents related to breastfeeding. In addition, we added to the tool questions about the advertisement of milk substitutes and other products for babies or mothers. A total of five text pages in Portuguese and 10 text pages in English were identified to be evaluated and pilot test the data extraction tool.

The first version of the tool raised questions related to the copyright and sponsorship of a website's main page, the author/creator of the website's main page, and the author/creator of the text page. These kinds of information were not originally required but were important to collect to identify and authenticate each text page and its main page. Issues related to copyright and sponsorship were added, and issues related to author/creator were clarified and rewritten.

## Training coders and evaluating the reliability

Following the steps suggested by McMillan (2009), we performed the process of training coders and evaluating their reliability. In our study, four coders were trained to evaluate and use the tool; two were native English-speaking Canadians and two were native Portuguese-speaking Brazilians. Canadian coders were an experienced midwife with extensive experience in obstetric care who is currently working as a faculty member at a Canadian university, Faculty of Nursing, and a nurse who is master's degree-prepared. Brazilians coders were an Obstetric Nurse who is currently working as a faculty member at a Brazilian university, Faculty of Nursing, and a nurse who was a PhD student. They are invited to be coders since they are experienced in supporting breastfeeding in their communities. The coders received training on the purpose of the tool and how to use it. To check the level of correspondence between coders, we followed the agreement criteria, which occurred when the coders agreed about the variable evaluated and the value assigned to it (Neuendorf, 2002). Following the training, data were collected.

Using the first version of the tool, coders raised a question about the use of scientific references in the text pages, and this question was added to the extraction tool. The extraction tool also made no reference to the presence and use of advertisements. The first evaluation of the tool and webpages surprised the Brazilian coders when they noticed a few websites containing product advertisements for babies and/or mothers. We then added questions to capture the use of advertisements.

- 1 Name of site
- 2 Date when data was collected
- 3 Date when information was posted **on text page**
- 4 Site is copyrighted
- 5 Breastfeeding is the site's main topic
- 6 If not, the site's main topic is
- 7 Country of origin (according to the site's domain)
- 8 Author/creator(s) of site's **main** page is mentioned
- 9 Author/creator(s) of site's **main** page
- 10 Author/creator(s) of **text** page is mentioned
- 11 Author/creator(s) of **text** page
- 12 Author/creator of the **site's** contact information
- 13 Author/creator of the **text page's** contact information
- 14 Site is sponsored
- 15 If yes, site is sponsored by
- 16 Target population of site
- 17 Resources used on **text page** for transmitting the specific information on breastfeeding
- 18 If there is an image in the **text page**, it is culturally suitable
- 19 Language of document
- 20 Authors present scientific references on **text page**
- 21 Text page is supportive of breastfeeding
- 22 Purpose/objective of **site** (If it is not stated, draw your own conclusions)
- 23 Main message/content/purpose of the **text page**
- 24 Unique characteristics of the site:
- 25 Does the **text page** conflict or agree with official documents related to breastfeeding?
- 26 When pictures and drawings are used, do they conflict or agree with official documents related to breastfeeding?
- 27 Is milk substitute advertisement included? Please explain.
- 28 Is an advertisement included for babies and/or mother products? (i.e., pacifiers and bottles for feeding, products for nipples or breasts) Please name and/or explain.

**Figure 1.** Data extraction tool.

We proceeded with a second round of testing of the data extraction tool, adding independent coders. The second review of the tool raised some controversial points about the characteristics of the site. Initially, coders had quickly glanced at the site and assessed and indicated its unique characteristics. Without a specific example, coders were confused; therefore, we decided to be more direct, stating some possible answers.

Another coder questioned the contact information on the website and text pages: the coders did not agree with what represented contact information. After some discussion, we agreed that addresses, email addresses, and phone numbers must be considered, as well as “contact us” pages, such as those with links to social media like Facebook, Twitter, and blogs, where readers could send messages to the author/creator of the main site or the text page. During the evaluation of the second version, Canadian coders questioned the use of images on text pages. To address this concern, one question related to images was inserted in the tool in order to know whether the picture or image was culturally suitable or not, in other words, to ascertain whether mothers felt represented in this image. After making all necessary adjustments, it was possible to reach the third and final version, which had the agreement of all coders.

Data were stored using a spreadsheet and separate columns were assigned to each coder. Thus, it was possible to compare the answers of each coder and assess their agreements. Discrepancies were discussed, and the tool adapted and reevaluated until a consensus was reached. In order to perform the data collection, the data extraction tool considered the variables depicted in Figure 1.

### *Analyzing and interpreting data*

Analyzing and interpreting data is the last step of content analysis and, according to McMillan's (2009) findings, is “descriptive in nature” (p. 64). We report our findings in a descriptive way and also compare it with the scientific evidence in order to identify whether breastfeeding information on the Internet follows the official recommendations.

## **Results and discussion**

Studies involving the media call attention to the use of new forms of technologies that can improve the quality of health of those who use these resources, as is the case with Internet-based media. According to Berry (2007), a substantial amount of health communication “takes place at a wider



public health, or mass communication, level” (p. 87). In this sense, the increase in Internet access for those seeking health information has advantages and disadvantages for health communication. While it facilitates access to up-to-date information, it can also increase confusion and anxiety in patients, mainly when they find conflicting information and advice. Furthermore, information obtained on Internet sites does not have the guarantee that it is accurate or reliable (Berry, 2007). However, creativity and careful adherence to the content analytic fundamentals can respond to the complexity of the huge amount of information found on the Internet (Neuendorf, 2002).

According to Weare and Lin (2000), in content analytic research, the structure of the analysis depends on the research question; however, “the outlines of the procedures required to develop valid and reliable measures and inferences are well established” (p. 273). Therefore, in our study, the content analysis considered the steps proposed by McMillan (2009) in order to achieve all requirements that are necessary when the evaluation of content on breastfeeding on the media of the Internet is performed.

The descriptive structure of the components added to the tool provides important information to characterize the websites and text pages, giving an overview of the most relevant sites that are found through the Google search mechanism. The quality of information provided on the Internet related to breastfeeding is evaluated in accordance with the recommendations based on scientific evidence for this practice. In this sense, using the developed tool, we can assess whether the information on text pages is supportive of breastfeeding and encouraging mothers to breastfeed as is recommended by WHO (WHO and United Nations Children’s Fund, 2009). In addition, the developed tool can assess whether other strategies that protect breastfeeding were followed. Data collected are compared with the International Code of Marketing of Breast-Milk Substitutes (WHO, 2017); the Brazilian Guidelines for the Marketing of Baby Food, Pacifiers and Bottles, which became a national law regarding the commercialization of foods for infants and young children and also related to nursery products (Brazil, 2006); or the guidelines for Breastfeeding Practices in Canada (Government of Canada, 2012). In order to protect and promote breastfeeding, the International Code of Marketing of Breast-Milk Substitutes was launched in 1981 as an instrument to regulate the marketing strategies of the industry providing breast milk substitutes and teats, pacifiers, and bottle-feeding. The code can be adapted for each country’s context. Brazil has implemented the code with a legislation that encompasses all provisions of the International Code; Canada has enacted legal measures that encompass a few provisions of the International Code (United Nations Children’s Fund, 2011).

Web pages and text pages are open access, and people have changed from being passive consumers to open and interactive creators of content (Hughes et al., 2008; Scanfield

et al., 2010). Thus, the web and text page creators are free and independent to give particular opinions or follow outdated references, which are often not supportive of breastfeeding. Open access also justifies the addition of questions related to the creators, copyright, and sponsorship, as such information is useful to identify lay, professional, or commercial website owners.

Furthermore, according to the International Code of Marketing of Breast-Milk Substitutes and the Brazilian Regulation, trade promotion about milk formulas, bottle-feeding, and pacifiers is forbidden in media, including merchandising and dissemination by electronic media and others (Brazil, 2006; WHO, 2017). Such regulation explains the questions added to the tool about advertisements for babies and mothers, whereby we can evaluate whether or not the text pages follow the recommendations.

In addition, it is significant to evaluate the cultural differences that exist between the countries because mothers who might have different ethnicities seek the text pages. Canadian coders were aware of these differences, since Canada is a major immigrant-receiving country (Ogilvie et al., 2012). Thus, the inclusion of a question to identify whether pictures in the text pages are culturally suitable was a positive change in the data extraction tool. Considering all the changes made to the tool, which had the final agreement of all coders, we assert that the developed data extraction tool is a useful instrument and can assist researchers in evaluating the quality of information posted on the Internet related to breastfeeding.

### Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by São Paulo Research Foundation—FAPESP (grant number 2014/14970-4).

### ORCID iD

Juliana Cristina dos Santos Monteiro  <https://orcid.org/0000-0001-6470-673X>

### References

- Berry D (2007) *Health Communication: Theory and Practice*. Maidenhead: Open University Press.
- Brazil (2006) *Law 11.265, January 3, 2006. Regulates the Marketing of Foods for Infants and Young Children and Also Correlates with Nursery Products*. Available at: [http://www.planalto.gov.br/ccivil\\_03/\\_Ato2004-2006/2006/Lei/L11265.htm](http://www.planalto.gov.br/ccivil_03/_Ato2004-2006/2006/Lei/L11265.htm) (accessed 12 March 2014).
- Cattaneo A (2012) Academy of breastfeeding medicine founder’s lecture 2011: Inequalities and inequities in breastfeeding: An international perspective. *Breastfeeding Medicine* 7(1): 3–9.

- Dykes F (2011) Twenty-five years of breast-feeding research in Midwifery. *Midwifery* 27(1): 8–14.
- Ferreira D, Carreira H, Silva S, et al. (2013) Assessment of the contents related to screening on Portuguese language websites providing information on breast and prostate cancer. *Cadernos de Saúde Pública* 29(11): 2163–2176.
- Giglia R and Binns C (2014) The effectiveness of the Internet in improving breastfeeding outcomes: A systematic review. *Journal of Human Lactation* 30(2): 156–160.
- Gionet L (2013) *Breastfeeding Trends in Canada* (Catalogue No. 82-624-X). Ottawa, ON, Canada: Statistics Canada.
- Government of Canada (2012) *Breastfeeding Practices in Canada: Overview*. Available at: <https://www.canada.ca/en/health-canada/services/food-nutrition/food-nutrition-surveillance/health-nutrition-surveys/canadian-community-health-survey-cchs/breastfeeding-practices-canada-overview-food-nutrition-surveillance-health-canada.html> (accessed 12 March 2014).
- Hargrave DR, Hargrave UA and Bouffet E (2006) Quality of health information on the Internet in pediatric neuro-oncology. *Neuro-Oncology* 8(2): 175–182.
- Hughes B, Joshi I and Wareham J (2008) Health 2.0 and medicine 2.0: Tensions and controversies in the field. *Journal of Medical Internet Research* 10: e23.
- Laanterä S, Pölkki T and Pietilä A-M (2011) A descriptive qualitative review of the barriers relating to breast-feeding counselling. *International Journal of Nursing Practice* 17(1): 72–84.
- McMillan SJ (2009) The challenge of applying content analysis to the World Wide Web. In: Krippendorff K and Bock MA (eds) *The Content Analysis Reader*. Thousand Oaks, CA: SAGE, pp. 60–67.
- Macnamara JR (2005) Media content analysis: Its uses; benefits and best practice methodology. *Asia Pacific Public Relations Journal* 6(1): 1–34.
- Ministério da Saúde, Brazil (2009) *II Pesquisa de Prevalência de Aleitamento Materno nas Capitais Brasileiras e Distrito Federal* [II Research of Breastfeeding Predominance in Brazilian Capitals and Federal District]. Brasília, Brazil: Ministério da Saúde.
- Moorhead SA, Hazlett DE, Harrison L, et al. (2013) A new dimension of health care: Systematic review of the uses, benefits, and limitations of social media for health communication. *Journal of Medical Internet Research* 15(4): e85.
- Neuendorf KA (2002) *The Content Analysis Guidebook*. Thousand Oaks, CA: SAGE.
- Ogilvie L, Higginbottom G, Burgess-Pinto E, et al. (2012) Fostering excellence: Development of a course to prepare graduate students for research on migration and health. *Nursing Inquiry* 20(3): 211–222.
- Rollins NC, Bhandari N, Hajeebhoy N, et al. (2016) Why invest, and what it will take to improve breastfeeding practices? *The Lancet* 387: 491–504.
- Scanfeld D, Scanfeld V and Larson E (2010) Dissemination of health information through social networks: Twitter and antibiotics. *American Journal of Infection Control* 38(3): 182–188.
- Site (2014) In Wikipedia, the free encyclopedia. Available at: <http://pt.wikipedia.org/wiki/Site> (accessed 18 June 2014).
- United Nations Children's Fund (2011) *National Implementation of the International Code Marketing of Breast-Milk Substitutes*. New York: United Nations Children's Fund.
- Victora CG, Bahl R, Barros AJD, et al. (2016) Breastfeeding in the 21st century: Epidemiology, mechanisms, and lifelong effect. *The Lancet* 387: 475–490.
- Wang L, Wang J, Wang M, et al. (2012) Using Internet search engines to obtain medical information: A comparative study. *Journal of Medical Internet Research* 14(3): e74.
- Weare C and Lin W-Y (2000) Content analysis of the World Wide Web: Opportunities and challenges. *Social Science Computer Review* 18(3): 272–292.
- Wolynn T (2012) Using social media to promote and support breastfeeding. *Breastfeeding Medicine* 7(5): 364–365.
- World Health Organization (2017) *The International Code of Marketing of Breast-Milk Substitutes: Frequently Asked Questions – 2017 Update*. Geneva: World Health Organization.
- World Health Organization and United Nations Children's Fund (2009) *Baby-Friendly Hospital Initiative: Revised, Updated and Expanded for Integrated Care: Section 1: Background and implementation*. Geneva: World Health Organization.

### Author biographies

Juliana Cristina dos Santos Monteiro is an Associate Professor at the University of São Paulo at Ribeirão Preto College of Nursing. The focus of her research is to support women during pregnancy, birth, postpartum period and early motherhood to acquire confidence and well-being.

Solina Richter is Professor and Associate Dean Global Health in the Faculty of Nursing, University of Alberta. Her research program focuses on the social determinants of health, and more specifically research to inform public policies and frontline practices that protect and promote the health of the low socioeconomic populations.