

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE ODONTOLOGIA DE BAURU

PATRICIA ESTEFANÍA AYALA AGUIRRE

**Quantitative and qualitative study of the structured and
unstructured digital information regarding early childhood caries**

**Estudo quantitativo e qualitativo das informações digitais
estruturadas e não-estruturadas sobre cárie na primeira infância**

BAURU

2022

PATRICIA ESTEFANIA AYALA AGUIRRE

Quantitative and qualitative study of the structured and unstructured digital information regarding early childhood caries

Estudo quantitativo e qualitativo das informações digitais estruturadas e não-estruturadas sobre cárie na primeira infância

Tese constituída por artigo apresentada à Faculdade de Odontologia de Bauru da Universidade de São Paulo para obtenção do título de Doutor em Ciências no Programa de Ciências Odontológicas aplicadas, na área de concentração Odontopediatria.

Orientador: Prof. Dr. Thiago Cruvinel da Silva

BAURU

2022

Ayala Aguirre, Patricia Estefanía
Quantitative and qualitative study of the
structured and unstructured digital information
regarding early childhood caries -- Bauru, 2022.
120p. : il. ; 31 cm.

Tese (doutorado) -- Faculdade de Odontologia
de Bauru, Universidade de São Paulo, 2022.

Orientador: Prof. Dr. Thiago Cruvinel da Silva

Autorizo, exclusivamente para fins acadêmicos e científicos, a
reprodução total ou parcial desta dissertação/tese, por processos
fotocopiadores e outros meios eletrônicos.

Assinatura:

Data:

Comitê de Ética da FOB-USP
Protocolo nº:
Data:

Universidade de São Paulo
Faculdade de Odontologia de Bauru
Assistência Técnica Acadêmica
Serviço de Pós-Graduação



FOLHA DE APROVAÇÃO


Tese apresentada e defendida por
PATRICIA ESTEFANIA AYALA AGUIRRE
e aprovada pela Comissão Julgadora
em 13 de maio de 2022.





Prof.^a Dr.^a **DANIELA PRÓCIDA RAGGIO**
FO-USP





Prof. Dr. **SAUL MARTINS DE PAIVA**
UFMG

Prof.^a Dr.^a **MARIA APARECIDA DE ANDRADE MOREIRA MACHADO**
FOB-USP

Prof. Dr. **THIAGO CRUVINEL DA SILVA**
Presidente da Banca
FOB - USP


Prof. Dr. Marco Antonio Hungaro Duarte
Presidente da Comissão de Pós-Graduação
FOB-USP

 Al. Dr. Octávio Pinheiro Brisolla, 9-75 | Bauru-SP | CEP 17012-901 | C.P. 73
 <https://posgraduacao.fob.usp.br>
 14 | 3235-8223 / 3226-6097 / 3226-6096
 posgrad@fob.usp.br

 [posgraduacaofobusp](https://www.facebook.com/posgraduacaofobusp)
 [@posgradfobusp](https://www.instagram.com/posgradfobusp)
 [fobuspoficial](https://www.youtube.com/fobuspoficial)
 [@Fobpos](https://twitter.com/Fobpos)

ERRATA

DEDICATÓRIA

Aos meus queridos pais Luis e Patricia, pelo amor e dedicação e por nunca medirem esforços para a minha formação, com o apoio e o incentivo constantes na busca pelos meus objetivos.

Los amo.

AGRADECIMENTOS

A **Deus**, que rege, guia e guarda todos os meus caminhos. Coloco sempre tudo nas tuas mãos, Senhor. Obrigada por todas as bênçãos diárias. Hoje não estaria aqui se não fosse por tua vontade.

Minha família, **meus pais**, meu irmão Fernando, minha irmã Gabriela, a minha querida sobrinha Doménica e meu cunhado Geovanny pela compreensão e incentivo. Amo vocês.

Ao meu prezado orientador **Prof. Dr. Thiago Cruvinel da Silva**, acredito que minhas palavras nunca conseguirão expressar a profunda gratidão e admiração que sinto pelo senhor. Agradeço a Deus todos os dias por ter me colocado no seu caminho, pela sua generosidade e presença constante, por oferecer uma oportunidade de trabalhar e deixar aprender com você e, em especial, por ter me apoiado e ser um amigo nos momentos mais difíceis nesses últimos anos. Deus abençoe você sempre.

À **Faculdade de Odontologia de Bauru - FOB/USP** e à Comissão de Pós-Graduação desta instituição.

Aos professores da Disciplina de Odontopediatria, **Prof.^a Dr.^a Maria Aparecida Andrade Moreira Machado**, **Prof.^a Dr.^a Thaís Marchini de Oliveira** e **Prof. Dr. Natalino Lourenço Neto** pelo conhecimento, experiências e confiança transmitidos à nossa turma.

Um agradecimento muito especial para a querida **Prof.^a Dr.^a Daniela Rios**. Você é um exemplo de uma pessoa trabalhadora e lutadora. Aprendi muito em todos os momentos que trabalhamos juntas. Estou muito agradecida pela confiança que depositou em mim.

Aos funcionários do Departamento de Odontopediatria **Andréia**, **Lílian** e **Estela**. Agradeço de coração pela paciência, disponibilidade, boa vontade e sorrisos compartilhados.

À **Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP)**, pela concessão da bolsa de Doutorado (processo nº. 2018/02563-6) essencial para o meu aprimoramento profissional e pessoal.

O presente trabalho também recebeu apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001”.

Às minhas queridas amigas do Doutorado, **Dani, Natalia, Mariel, Mayara, Eloá, Bianca Z.**, e a **Prof.^a Dr.^a Paulinha**, pelas experiências vividas e trocadas.

Aos meus companheiros de pesquisa, **Matheus e Olívia**, pela ajuda constante nos nossos trabalhos. Muito obrigada!

À minha querida amiga e irmã **Anna Paola**. Estou muito agradecida pelo companheirismo e carinho nos momentos mais difíceis.

Aos melhores amigos que o Brasil me deu, **Mateus e Arón**. A saudade de vocês dói todos os dias.

E finalmente, à minha querida família de Bauru, **Bárbara, Fernanda, Lucas, Bia, Renan e Felipe**. Todos os dias agradeço a Deus por tê-los na minha vida. Todos os conselhos, risadas, aventuras, tristezas, absolutamente toda a vivência me ajudou muitíssimo no meu caminhar durante esses nove anos em Bauru. Amo vocês.

“Es una locura no ser lo que se es, con la mayor plenitud posible”.

Santa Maria Eugenia Milleret de Jesús

RESUMO

A cárie na primeira infância (CPI) é uma condição bucal que afeta 532 milhões de crianças em todo o mundo, representando um grande desafio econômico e de saúde pública. O presente estudo teve como objetivo (1) analisar o volume e o perfil das pesquisas em consultas na Internet sobre CPI realizadas por usuários do Google de diferentes países, (2) avaliar a legibilidade e a qualidade das informações relacionadas à CPI de websites em inglês, espanhol e português brasileiro, (3) avaliar a completude das informações sobre CPI encontradas em vídeos do Youtube nos mesmos três idiomas e (4) determinar os sentimentos e interesses de usuários do Twitter relacionados a informações verdadeiras ou falsas/incorrectas sobre cárie dentária nos mesmos três idiomas. A variação mensal do Volume Relativo de Buscas (RSV) para consultas relacionadas à CPI foi determinada pelo API do Google Trends para cada país. A associação de RSV com anos vividos com incapacidade (YLDs) para cárie dentária não tratada em dentes decíduos e penetração da Internet foram testadas por modelos de regressão linear. Os tópicos mais relevantes pesquisados pelos usuários do Google foram avaliados qualitativamente. Para a análise da qualidade da informação, os websites foram recuperados pelo Google, Yahoo! e Bing de acordo com a participação de mercado de cada ferramenta nos EUA, México e Brasil. O questionário DISCERN, critérios de referência JAMA e fórmulas de legibilidade específicas para cada língua foram utilizados por dois investigadores independentes para avaliar a qualidade e a legibilidade dos websites. A completude das informações dos vídeos do Youtube foi avaliada por uma lista predefinida de 17 tópicos relacionados a informações educativas sobre CPI. Categorias dicotomizadas dos vídeos de acordo com a mediana da pontuação de completude e a natureza de sua autoria foram comparadas em relação às características dos vídeos e aos parâmetros de interação com a audiência do YouTube. Para avaliação do Twitter, foi construída uma estratégia de busca específica de acordo com cada idioma para coletar pelo menos 1.000 tweets para cada país (EUA, México e Brasil). Dois examinadores classificaram os tweets em seu idioma original de acordo com o sentimento (positivo, neutro ou negativo), tópicos relacionados à CPI e presença de informação falsa/incorrecta com base em evidências científicas atualizadas sobre a cárie dentária. Os resultados indicaram que o interesse dos usuários do Google por

informações sobre CPI ainda é baixo, porém está incrementando ao longo do tempo. Os usuários interessados pela doença podem se deparar com informações de baixa qualidade em websites e vídeos do Youtube. Além disso, os tweets relacionados à cárie dentária contendo informações falsas/incorretas estão associados a sentimentos positivos e frequentemente relacionados a conceitos errados sobre o consumo de açúcar e o desenvolvimento da doença. Em conclusão, a análise de Big Data é uma estratégia importante para identificar os interesses e preocupações dos usuários da Internet em relação às informações relacionadas à CPI de populações específicas. Além disso, o conhecimento sobre o que os pais e cuidadores estão expostos ao buscar por informações sobre a CPI na Internet é fundamental para melhorar o manejo da doença. Portanto, os profissionais da Odontologia devem estar atentos às necessidades de seus pacientes, produzindo materiais digitais de boa qualidade e promovendo o aconselhamento informativo durante as consultas clínicas.

Palavras-chave: Cárie na primeira infância. Internet. eHealth

ABSTRACT

Quantitative and qualitative study of the structured and unstructured digital information regarding early childhood caries

Early childhood caries (ECC) is an oral condition that affects 532 million children worldwide, representing a major economic and public health challenge. The present study aimed to (1) analyze the volume and profile of searches related to ECC performed by Google users from different countries, (2) evaluate the readability and quality of information related to ECC from websites in English, Spanish and Brazilian Portuguese, (3) to assess the completeness of ECC information found in Youtube videos in the same three languages, and (4) to determine the sentiment and interests of Twitter users regarding true or false/incorrect information about dental caries in the same three languages. The monthly variations in Relative Search Volume (RSV) for ECC-related queries was determined by the Google Trends API for each country. The association of RSV with years lived with disability (YLDs) for untreated dental caries in primary teeth and Internet penetration were tested by linear regression models. The most relevant topics searched by Google users were qualitatively evaluated. For the analysis of the quality of the information, the websites were retrieved by Google, Yahoo! and Bing according to the market share of each engine in the US, Mexico and Brazil. The DISCERN questionnaire, JAMA benchmarks and language-specific readability formulas were used by two independent researchers to assess the quality and readability of the websites. Completeness of information from Youtube videos was assessed by a predefined list of 17 topics related to educational information on ECC. Dichotomized categories of videos according to the median completeness score and the nature of their authorship were compared in relation to the characteristics of the videos and the parameters of interaction with the YouTube audience. To evaluate Twitter, a specific search strategy was built according to each language to collect at least 1,000 tweets for each country (USA, Mexico and Brazil). Two examiners rated tweets in their original language according to sentiment (positive, neutral, or negative), ECC-related themes, and presence of false/incorrect information based on up-to-date scientific evidence on dental caries. The results indicated that the interest of Google users for ECC information is still low, but it is increasing over time. Users interested in the disease may come across low-quality information on websites and YouTube videos. Furthermore, tweets related to dental caries containing false/incorrect

information are associated with positive sentiment and are often related to misinformation about sugar consumption and the development of the disease. In conclusion, Big Data analysis is an important strategy to identify Internet users' interests and concerns regarding ECC-related information from specific populations. In addition, understanding what parents and caregivers are exposed to when searching for information about ECC on the Internet is critical to improving disease management. Therefore, dental professionals must be attentive to the needs of their patients, producing good quality digital materials and promoting informative counseling during clinical consultations.

Keywords: Early childhood caries. Internet. eHealth. Misinformation.

LIST OF ILLUSTRATIONS

- ARTICLE 1

- Figure 1** - Predictive charts of RSV values for ECC-related searches performed in Brazil, France, United Kingdom, Italy, Mexico, Spain, and United States. The curves of observed values (blue lines), fit and forecast values (purple lines), and upper and lower bound of confidence intervals (green lines) are depicted from January 2004 through December 2020. RSV values presented after December 2020 (black line) represent 12- month forecasts. RSV: Relative Search Volume. 42
- Figure 2** - Heat map of monthly variation of the predictive generalized additive model values for ECC-related interest in all countries..... 43

- ARTICLE 2

- Figure 1** - Flowchart depicting the systematic selection of ECC-related websites by language..... 62

- ARTICLE 3

- Figure 1** - Flowchart depicting the systematic selection of Youtube videos containing ECC-related information, according to each language..... 80

- ARTICLE 4

- Figure 1** - Flowchart showing the synthesis of the methodological approach.... 104
- Figure 2** - Word cloud of the most frequent words found in dental caries-related tweets from three countries. 105
-
-

LIST OF TABLES

- ARTICLE 1

- Table 1** - Search strategies developed for each country 38
- Table 2** - Comparison of medians of RSV for ECC-related information between 2 time periods by country 39
- Table 3** - Regression models for the association of RSV for ECC-related information with YLDs and Internet penetration in different countries, regarding total population, females, and males. 40
- Table 4** - ECC-related topics and their respective RSV values for each country 41

- ARTICLE 2

- Table 1** - Descriptive statistics of quality and readability scores according to the languages of websites 60
- Table 2** - The comparison of means (SD) of quality and readability scores between dichotomised categories of websites published in distinct languages. Different superscript letters represent significant statistical differences between dichotomised categories regarding the same quality and readability measure (Mann-Whitney U test, $P < 0.05$) 61

- ARTICLE 3

- Table 1** - Descriptive statistics of videos regarding their completeness scores and characteristics by language 77
- Table 2** - Distribution of frequency of videos containing specific completeness score items by language 78
-
-

Table 3 - Distribution of the frequency of videos according to dichotomized categories of completeness score and authorship regarding video characteristics and languages. Different superscript letters represent significant statistical differences between dichotomized categories in each language (Mann-Whitney U test, $P < 0.05$)..... 79

- ARTICLE 4

Table 1 - Search strategies for Twitter API according to each language 98

Table 2 - Distribution of different variables according to different countries, Distinct letters represent significant statistical differences of distributions between countries (Chi-square test)..... 99

Table 3 - Descriptive statistics of interaction metrics according to different countries. Distinct letters represent significant statistical differences between countries (Kruskal-Wallis and Dunn's tests)..... 100

Table 4 - Distribution of human- and AI-based sentiment analyses according to different countries. Accuracies, sensitivities, and specificities of VADER and TextBlob are demonstrated in relation to human-based analysis. Distinct letters represent significant statistical differences of distributions between countries (Chi-square test)..... 101

Table 5 - Distribution of sentiment, misinformation, interaction, languages, sex, and accounts in relation to dichotomized themes. Distinct letters represent significant statistical differences within each variable (Chi-square test) 102

Table 6 - Multiple logistic regression model showing the association between misinformation with different factors 103

LIST OF ABBREVIATIONS

ECC	Early Childhood Caries
RSV	Relative Search Volume
YLD	Years Lived with Disability
BIC	Bayesian information criteria
HISB	Health Information Seeking Behavior
ICT	Information and Communication Technologies
SM	Social Media
HINTS	National Health Information Trends Survey

TABLE OF CONTENTS

1	INTRODUCTION	19
2	ARTICLES	22
2.1	ARTICLE 1 – Digital surveillance: monitoring the activity of Internet users searching for information related to early childhood caries (published in the <i>Health Informatics Journal</i>).....	23
2.2	ARTICLE 2 – Online quality and readability assessment of Early childhood caries information available on Websites from distinct countries: A cross-sectional study (published in the <i>European Journal of Pediatric Dentistry</i>)....	44
2.3	ARTICLE 3 – A cross-sectional completeness assessment of Youtube videos’ information regarding early childhood caries (formatted according to the guidelines of the <i>Journal of Medical Internet Research</i>).....	63
2.4	ARTICLE 4 – Tweeting about dental caries: an ecological digital study (formatted according to the guidelines of <i>PloS One</i>).....	81
3	DISCUSSION.....	108
4	CONCLUSIONS	113
	REFERENCES	116

1 INTRODUCTION

1 INTRODUCTION

Dental caries still represents a major public health challenge [1-3] affecting about 532 million children in the primary dentition stage [3] with a prevalence of 48% [95% CI 42, 52] resulting from 29 countries with available information of [4]. The presence of one or more decayed primary teeth (cavitated or not), lost and restored due to the progression of demineralization by caries in children under 6 years old characterizes early childhood caries (ECC) [5], also known as “early childhood decay”, “bottle decay” or “bottle-feeding syndrome” [6]. The importance of preventing ECC has been widely discussed in the literature [4, 5, 7-9]. The identification of individual risk factors, parental counseling and health promotion have played an important role in the disease prevention [5], avoiding possible damage to the quality of life of children and their families caused by advanced stages of the illness, such as alterations in mastication and speech, pain, psychological problems, and negative effects on the child's weight and growth [10, 11]. In addition, children diagnosed with ECC also have higher probabilities of developing carious lesions in the permanent dentition [6].

In recent years, the Internet has become the most popular means for obtaining health-related information [12], due to its rapid dissemination [13] associated with changes in the structure of health care, from the biomedical model to person-centered care [14]. As of January 2022, there are 298.8 million Internet users in United States alone, with a increase of 11 million between pandemic years, spending 7h and 11 minute daily in average, with a total of 17.3 billion Google search visits in December 2020 [15]. In Brazil, 85% of Internet users accessed the web every day and 47% search for content related to health or health services [16]. In addition, approximately 9 out 10 parents do search for health information about their children [17]. The analysis of Big Data produced by the activity of Internet users could contribute to the recognition of the dental needs of different groups to guide the planning and implementation of health actions [18]. The strategies would be based on data related to Behavioral Medicine/Dentistry, a multidisciplinary area that combines knowledge about the humanities and health for a better understanding of health and disease processes in different communities [19, 20]. Health concerns related to ECC could encourage parents and caregivers of children affected by the disease to seek additional

information on the Internet [21], similarly to what is increasingly happening among adults from different countries [22]. However, this behavior could pose risks to the target population if the information on ECC was also considered of low quality, as previously demonstrated to dental caries-related information available on Brazilian websites [23]. This concern becomes even more important when the increase in the Internet penetration in the poorest populations of developing countries is regarded, because the disease is more prevalent among individuals of lower socioeconomic levels [24]. Social networks have also been widely used by people to express their concerns and doubts or seek for advice about a specific condition [25], which generates a considerable amount of information that can be analyzed to understand the specific oral needs of a country or region [18]. However, to our knowledge, aspects of structured and unstructured ECC-related digital information regarding the volume of searches and shares in near-real time, and the quality of information available on websites and videos have not yet been determined. Additionally, the impact of the disease on social media users has not yet been analyzed and compared with the impact on parents and/or caregivers of children seeking pediatric dental care. According to the agenda proposed by the American Academy of Pediatric Dentistry, research investigating the use of technology for monitoring health behaviors has fundamental importance to reduce disparities related to oral health among different populations [26].

Therefore, this study aimed to analyze and characterize different types of ECC-related information found in texts, videos, posts, and opinions available on the Internet.

2 ARTICLES

2 ARTICLES

This Thesis is presented as a compilation of 4 articles written according to specific instructions and guidelines of journals, as follows:

- ARTICLE 1 – Digital surveillance: monitoring the activity of Internet users searching for information related to early childhood caries (published in the *Health Informatics Journal*).
 - ARTICLE 2 – Online quality and readability assessment of Early childhood caries information available on Websites from distinct countries: A cross-sectional study (published in the *European Journal of Pediatric Dentistry*).
 - ARTICLE 3 – A cross-sectional completeness assessment of Youtube videos' information regarding early childhood caries (formatted according to the guidelines of the *Journal of Medical Internet Research*).
 - ARTICLE 4 – Tweeting about dental caries: an ecological digital study (formatted according to the guidelines of *PloS One*).
-
-

2.1 ARTICLE 1 - Digital surveillance: monitoring the activity of Internet users searching for information related to early childhood caries (published in the *Health Informatics Journal*).

Abstract

This study evaluated the behavior of Internet users from seven countries through the analysis of the monthly variation of the Relative Search Volume (RSV) for queries related to early childhood caries (ECC), retrieved by Google Trends application programming interface between January 2004 to December 2020. The association of RSV with years lived with disability (YLDs) for untreated dental caries in deciduous teeth and Internet penetration were tested by regression models. Forecasting models were developed to predict the interests until December 2021. Relevant topics linked to searches were assessed qualitatively. Increasing levels of interests were observed in most countries, regarding the comparison of periods 2004–2012 and 2013–2021, without influence of seasonality. These results Internet were associated with YLDs only in France and with the Internet penetration in Mexico and France. Searches were mostly related to the definition, risk factors, and preventive care of ECC. The interests of Internet users in ECC-related information increased in the last years in all seven countries, although it was maintained low over time. It might indicate a lack of awareness about the negative consequences of the disease in the early stages of life. Specific policies should be developed toward the dissemination of oral health information and the prevention of the disease.

Keywords: dental caries, deciduous tooth, health behavior, information seeking behavior, eHealth

Introduction

Early childhood caries (ECC) is defined as a biofilm-mediated, sugar-driven, multifactorial dynamic disease that results in the phasic demineralization and remineralization of dental hard tissues of primary teeth of a child under 6 years old [1]. It affects 532 million infants, being 48.1% of cases concentrated in low-middle-income countries, portraying a challenging scenario for the public healthcare system worldwide [2].

Since the perception of risk of a disease is a significant determinant of health information seeking behavior, hypothetically, the burden caused by ECC could motivate parents and caregivers to engage into online searches, especially considering the context of the almost ubiquitous availability of digital sources in daily lives [3]. Indeed, it has been reported that nine out of 10 parents or caregivers are interested in health information regarding their children's conditions, employing Google Search as the first choice for their consultations [4]. Although this outcome indicates a positive trend of the development of a sense of independence toward self-care and health experiences, the low level of quality of information [5] and common difficulties of laypeople in retrieving, analyzing, and using digital contents could not have a beneficial effect for aiding parents or caregivers in improving children's health status [6,7]. Actually, it could have significant negative repercussions on the ability of people in making adequate treatment choices, affecting their confidence during clinical conversation and their relationship with professionals [4].

According to the American Academy of Pediatric Dentistry, research focused on the use of digital technology for monitoring people's health behaviors is fundamental to the better understanding of community needs, contributing to the reduction of oral health disparities observed in distinct populations [8,9]. Recently, the analysis of Big Data

related to the activity of Internet users have already been proved to be suitable for elucidating oral health experiences and specific interests of populations [10–13].

In this sense, the present study aimed to assess the levels of interest of Google user's regarding ECC-related information in different countries, regarding possible associations with the burden of the disease and the Internet penetration.

Methods

This report followed the STrengthening the Reporting of OBservational Studies in Epidemiology checklist for cross-sectional studies.¹⁴

Study design

Seven countries between January 2004 and December 2020. The monthly variation of Relative Search Volume (RSV) for queries linked to ECC was collected using Google Trends application programming interface (API), from search strategies developed by the combination of specific terms in the native language of each country. To elucidate a possible influence of the burden of the disease and the availability of the Internet on the health information seeking behavior, years lived with disability (YLDs) for untreated dental caries in deciduous teeth in children under 5 years old and Internet penetration were collected for each country over time. Data analysis was conducted to assess (a) search volume trends; (b) seasonality; (c) the association between RSV, YLDs, and Internet penetration; (d) forecasting models; and (e) the main Google users' interests in ECC.

Ethics

Since federal regulations consider that research using publicly available data does not involve human subjects, this study did not require institutional review board approval from the Council of Ethics in Human Research of the Bauru School of Dentistry, University of Sao Paulo.

Search volume trends and data collection

Google Trends is a free online tool that calculates the weekly or monthly variation of the RSV values for queries performed in Google Search. These values are normalized, scaled from 0 to 100, with RSV = 100 representing the maximum volume rate in relation to total searches performed in a specific time range and location. In this study, we collected the metadata through the Google Trends API, using the Google Colaboratory (Colab) notebook (<https://colab.research.google.com/> platform). The API was programmed in Python 3, applying the PyTrends and Pandas libraries.

First, we used the topic-Early Childhood Caries to identify all countries with sufficient volume of data for analysis on Google Trends. Then, different search strategies were developed for each of the included countries by the analysis of the relevance of ECC-related terms in their respective native language, excluding terms with insufficient volume of data (Table 1). Finally, the specific search strategy was used in Google Trends API for data collection. In addition, a list of the most relevant topics linked to searches was also collected to determine the main people's interests.

Years lived with disability

The Global Burden Disease (GBD) study analyzed the burden estimates of 291 diseases and injuries including untreated caries in deciduous teeth. The metric used

for analysis is disability-adjusted life years (DALYs) that represents the sum of estimates of years of life lost due to premature mortality and YLDs. Nevertheless, as untreated caries would rarely lead to direct death, one DALY is represented by YLDs for all oral conditions.² The variation of YLDs for untreated caries in deciduous teeth of children under 5 years old between 2004 and 2019 was retrieved from the Institute for Health Metrics and Evaluation (<http://vizhub.healthdata.org/gbd-compare/>) database, regarding both genders and selected countries.

Internet penetration

The percentage of people with access to Internet of each country was obtained in The World Bank database (<https://data.worldbank.org/indicator/IT.NET.USER.ZS>) for the available time between 2004 and 2019.

Data analysis

Collected data was analyzed using the Statistical Package for Social Sciences (SPSS v. 25.0; Chicago, IL, USA), considering the following aspects [10,11]:

1. Forecasting models and search volume trends: 12-month predictive models were developed for RSV values from collected data. The best-fitted models were provided by the Expert Modeler based on the lowest values of normalized Bayesian information criteria (normalized BIC). The trends of time series were heuristically evaluated by the observation of curves of the collected values. Also, the generalized additive model (GAM) was applied to detect significant variations of RSV values between 2004 and 2020. Finally, the activity of Google users in the periods 2004–2012 and 2013–2021 were compared using the Mann–Whitney U test.

2. Association of RSV with YLDs and Internet penetration: the annual means of RSV were calculated between the years 2004 and 2019 (period with available data for the three indicators). Linear regression models were conducted to analyze the association of RSV with YLDs and Internet penetration of different countries. The Durbin–Watson (DW) test was applied to detect the influence of autocorrelation over time. According to significance tables, the critical values of the DW test for one predictor (YLDs or Internet penetration) and a sample size of 16 (2004–2019), DW values <0.844 indicate a significant influence of autocorrelation, values between 0.844 and 1.086 are inconclusive, and values >1.086 indicate no influence of autocorrelation on time series. The Ljung–Box test was used to detect a significant autocorrelation on time series with inconclusive DW values. Linear regressions serially correlated over time were adjusted by Cochrane–Orcutt estimation since all models satisfied the following assumptions: their residuals were found to be a stationary first-order autoregressive structure with the errors being white noise [15]. When no influence of autocorrelation was observed, the association was evaluated by linear regression models. The variable time was used as a confounding factor in all cases.

3. Seasonality: the influence of seasonality on time series was evaluated by the GAM. It consisted in a previous detrending of each long-term curve by its lag-1 difference and subsequently the application of two distinct generalized linear models on these differences, to evaluate separately the effect of monthly and quarterly seasonality on time series.

4. Main topics: the most popular ECC-related topics were defined for each country for subsequent qualitative analysis.

For all analyses, p values $<.05$ were considered significant.

Results

Among all geographic regions available in Google Trends, only seven countries had sufficient volume of data for analysis: Brazil (BRA), France (FRA), United Kingdom (GBR), Italy (ITA), Mexico (MEX), Spain (SPA), and United States (USA).

Forecasting models and search volume trends

Figure 1 demonstrates the trends of the variation of the interests in ECC-related information for each country over time. In most countries, the heuristic analysis of curves indicates a slightly increasing trend of RSV values, being more evident in FRA and GBR. Nevertheless, GAM analysis confirmed this trend only in FRA by the observation of a significant increase of RSV between 2004 and 2020. The fit statistics for predictive models for each country are presented as a supplemental file. When comparing the periods 2004–2012 and 2013–2021 (including 12-month forecasts), a significant increase in the interests in ECC-related information was observed for all countries, except for the USA (Table 2).

Association of Relative Search Volume with years lived with disability and Internet penetration

RSV and YLDs were significantly and negatively associated only in France. Also, RSV was positively and negatively associated with Internet penetration in Mexico and France, respectively (Table 3).

Seasonality

Figure 2 depicts the heat map that represents the monthly variation of the predictive GAM values for ECC-related interests in all countries. It was noted a significant monthly

seasonality only in the USA, with the highest and lowest levels of interest observed in February and November, respectively.

Main topics

The lists of most popular topics linked to queries of each country are shown in Table 4. Tooth decay (syndrome) was the first related topic found in all countries, except for ITA (tooth). Topics related to the definition of the disease (tooth, infant, child, dentistry, and deciduous teeth), risk factors (breastfeeding, milk, baby bottle, and sleep), and preventive care (preventive healthcare, fluoride, and brushing) were commonly detected among countries.

Discussion

In general, these findings indicate increasing trends in the interest of people in ECC-related digital information, although the unequivocal identification of low search activity of Google users over time. The main queries were usually linked to the definition, risk factors, and preventive care of the disease. This information seeking behavior was not influenced by the effects of quarterly and monthly seasonality, and also it was not associated with the burden of untreated dental caries in children under 5 years old and the Internet penetration in most countries.

There are some possible explanations for these outcomes. The small search volume associated with the large percentage of queries on the definition of the disease could indicate a low level of population awareness on the impact of ECC [16]. It is quite common the observance of negligent parental behaviors in relation to the maintenance of a healthy primary dentition by the misperception of irrelevance derived from the temporary nature of deciduous teeth [17]. Consequently, parents and caregivers

usually seek treatment for babies and toddlers only after the recognition of symptoms as dental pain, which appear only in specific situations as the final stage of dental caries [18,19]. In this sense, the hypothesis of a significant positive association between RSV (digital activity/interest) and YLDs for untreated caries (burden of disease determined as dental pain) seemed plausible; however, only in France, those indicators were associated negatively. It might represent a greater engagement of people in consuming information for the maintenance of good oral health conditions, which can contribute to the low burden of the disease observed over time, and vice versa. The slightly increase in the interest in ECC-related topics observed in the last years could be associated with the growth of initiatives to the management of dental caries in early stages of life [20,21]. On the other hand, the activity of parents and caregivers to seek information might be reduced to their perception of unsatisfactory digital contents [3]. Indeed, the quality of Websites from distinct countries containing ECC-related information was considered invariably very poor in our recent publication [5].

Additionally, the availability of the Web could influence the increment of health information seeking behaviors over the years [3], as supported by the positive association with RSV found in Mexico, the country with the highest raise of Internet penetration since 2004 (397.2%) [22]. The digital divide still represents a great challenge worldwide, widening the disparities in healthcare [23]. Socioeconomic-related barriers as insufficient technology skills, health literacy, and self-efficacy could prevent people with health information needs to make online searches, even when they have access to the Internet [6]. Furthermore, YLDs for untreated caries in deciduous teeth in children under 5 years old are notably lower in high-income countries (0.02 million YLDs) in comparison to lower middle-income ones (0.06 million

YLDs) according to the GBD study [2]; that disinterest in relation to their demands. Then, dental professionals must be prepared to guide parents and caregivers in searching online health information toward good quality digital content, addressing parental concerns and doubts regarding their children's conditions, through the implementation of personalized strategies for the management of the disease [25–27]. This attitude is able to create a great opportunity to develop health literacy and construct an evidence-based parental criticism, aiding in the combat of pernicious effects of digital misinformation and disinformation [28].

Two inherent limitations of infodemiological studies must be considered in the interpretation of these outcomes. First, the present analysis was based on data retrieved and compiled through secret algorithms of Google, which makes difficult the analysis of terms according to inclusion and exclusion criteria; however, ECC-related topics depicted in Table 4 clearly support the quality and soundness of these methods and analyses. Second, it is not possible to define a profile of Internet users who seek for ECC-related digital information using this approach, because RSV is determined from the registration of IPs (Internet protocols) linked to specific computers.

In conclusion, the interests of Google users in ECC-related information is increasing slowly in all seven countries over the years, and it is not being associated with the burden of untreated dental caries in children under 5 years old and the Internet penetration in most countries, except for France and Mexico. It might indicate a lack of awareness about the negative consequences of the disease in the early stages of life, and also the parental distrust on the quality and the usefulness of online information. In this sense, specific policies should be developed toward the dissemination of oral health information and the prevention of the disease. Additionally, dental professionals

must be aware of their patient's needs, producing good quality digital materials and promoting the in-formational counseling during clinical appointments.

Author contributions

Conceived and designed the study: PEAA and TC; acquired the data: PEAA and ML; analyzed and interpreted the data: PEAA and TC; and drafted and critically revised the manuscript: PEEA, APS, ML, and TC.

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 research was funded by the São Paulo Research Foundation (#2018/02563-6). The funder had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

1. Machiulskiene V, Campus G, Carvalho JC, et al. Terminology of dental caries and dental caries management: consensus report of a workshop organized by ORCA and cariology research group of IADR. *Caries Res* 2020; 54: 7–14.
 2. GBD Oral Disorders Collaborators, Bernabe E, Marcenes W, et al. Global, regional, and national levels and trends in burden of oral conditions from 1990 to 2017: a systematic analysis for the global burden of disease 2017 study. *J Dent Res* 2020; 99: 362–373.
 3. Wang X, Shi J and Kong H. Online health information seeking: a review and meta-analysis. *Health Commun* 2021; 36: 1163–1175.
 4. Kubb C and Foran HM. Online health information seeking by parents for their children: systematic review and agenda for further research. *J Med Internet Res* 2020; 22: e19985.
 5. Aguirre PEA, Aníbal I, Strieder AP, et al. Online quality and readability assessment of early childhood caries information available on websites from distinct countries: a cross-sectional study. *Eur J Paediatr Dent* 2022 (in press).
 6. Jiang S and Liu PL. Digital divide and Internet health information seeking among cancer survivors: a trend analysis from 2011 to 2017. *Psychooncology* 2019; 29: 61–67.
 7. Messadi D, Macek M, Markovic D, et al. Oral health literacy, preventive behavior measures, and chronic medical conditions. *JDR Clin Trans Res* 2018; 3: 288–301.
 8. AAPD: American Academy of Pediatric Dentistry. *The Reference Manual of Pediatric Dentistry* 2020. Chicago, IL: American Academy of Pediatric Dentistry, 2020.
 9. AAPD: American Academy of Pediatric Dentistry. *American Academy of Pediatric Dentistry Research Agenda*. *Pediatr Dentistry* 2017; 38: 10.
-

10. Lotto M, Aguirre PEA, Rios D, et al. Analysis of the interests of Google users on toothache information. *PLoS One* 2017; 12: e0186059.
 11. Aguirre PE, Coelho M, Oliveira T, et al. What can Google inform us about people's interests regarding dental caries in different populations? *Caries Res* 2018; 52: 177–188.
 12. Aguirre PEA, Strieder AP, Lotto M, et al. Are the internet users concerned about molar incisor hypo- mineralization? An infoveillance study. *Int J Paediatr Dent* 2020; 30: 27–34.
 13. Strieder AP, Aguirre PEA, Lotto M, et al. Digital behavior surveillance for monitoring the interests of Google users in amber necklace in different countries. *Int J Paediatr Dent* 2019; 29: 603–614.
 14. Vandembroucke JP, von Elm E, Altman DG, et al. Strengthening the reporting of observational studies in epidemiology (STROBE): explanation and elaboration. *Ann Intern Med* 2007; 16147(8): W163–W194.
 15. Wooldridge JM. Serial correlation and heteroskedasticity in time series regression. In: *Introductory econometrics: a modern approach*. Mason, USA: Southern-West Cengage Learning, 2013, pp. 415–417.
 16. Broadbent J, Zeng J, Foster Page L, et al. Oral health–related beliefs, behaviors, and outcomes through the life course. *J Dent Res* 2016; 95: 808–813.
 17. Al-Batayneh OB, Al-Khateeb HO, Ibrahim WM, et al. Parental knowledge and acceptance of different treatment options for primary teeth provided by dental practitioners. *Front Public Health* 2019; 7: 322.
-
-

18. Sanguida A, Vinothini V, Prathima GS, et al. Age and reasons for first dental visit and knowledge and attitude of parents toward dental procedures for Puducherry children aged 0–9 years. *J Pharm Bioallied Sci* 2019; 11: S413.
 19. Ramakrishnan M, Banu S, Ningthoujam S, et al. Evaluation of knowledge and attitude of parents about the importance of maintaining primary dentition-A cross-sectional study. *J Fam Med Prim Care* 2019; 8: 414.
 20. Schwendicke F, Splieth C, Breschi L, et al. When to intervene in the caries process? An expert Delphi consensus statement. *Clin Oral Investig* 2019; 23: 3691–3703.
 21. Pitts N and Zero D. White paper on dental caries prevention and management. In: Goldberg M (ed) *Understanding dental caries - from pathogenesis to prevention and therapy*. Geneva, Switzerland: FDI World Dental Federation, 2016, pp. 3–9.
 22. The World Bank. Individuals using the internet (% of population) - United Kingdom, Mexico, Brazil. <https://data.worldbank.org/indicator/IT.NET.USER.ZS?locations=GB-MX-BR> (2019, accessed April 2021).
 23. Arighi A, Fumagalli GG, Carandini T, et al. Facing the digital divide into a dementia clinic during COVID- 19 pandemic: caregiver age matters. *J Neurol Sci* 2021; 42: 1247–1251.
 24. Hatef E, Ma X, Shaikh Y, et al. Internet access, social risk factors, and web-based social support seeking behavior: assessing correlates of the “digital divide” across neighborhoods in the State of Maryland. *J Med Syst* 2021; 45: 1–14.
 25. McGrath C.. Behavioral sciences in the promotion of oral health. *J Dent Res* 2019; 98: 1418–1424.
-
-

26. Lotto M, Strieder AP, Aguirre PEA, et al. Parental-oriented educational mobile messages to aid in the control of early childhood caries in low socioeconomic children: a randomized controlled trial. *J Dent* 2020; 101: 103456.
 27. Colvara BC, Faustino-Silva DD, Meyer E, et al. Motivational interviewing for preventing early childhood caries: a systematic review and meta-analysis. *Community Dent Oral Epidemiol* 2021; 49: 10–16.
 28. Wang Y, McKee M, Torbica A, et al. Systematic literature review on the spread of health-related misinformation on social media. *Soc Sci Med* 2019; 240: 112552.
-

Table 1. Search strategies developed for each country

Country	Search strategy (API)
BRA	'cárie em bebê' + 'carie em bebe' + 'cárie de mamadeira' + 'carie de mamadeira'
FRA	'carie dent de lait' + 'carie du biberon' + 'carie bebe'
GBR	'baby tooth decay' + 'toddler tooth decay'
ITA	'carie denti da latte'
MEX	'caries infantil' + 'caries rampante' + 'caries de biberon' + 'caries en dientes de leche'
SPA	'caries infantil' + 'caries rampante' + 'caries en dientes de leche'
USA	'early childhood caries' + 'bottle caries' + 'baby tooth decay' + 'baby bottle tooth decay'

Table 2. Comparison of medians of RSV for ECC-related information between 2 time periods by country

	2004-2012		2013-2021		<i>p</i>	Trends
	Median	Interquartile Range	Median	Interquartile Range		
BRA	3.00	11.00	9.00	5.00	<0.001	increase*
FRA	0	0	28.50	33.00	<0.001	increase*
GBR	0	0	14.00	13.00	<0.001	increase*
ITA	0	0	16.00	18.00	<0.001	increase*
MEX	0	9.00	5.00	4.00	0.004	increase*
SPA	0	20.00	16.50	17.00	<0.001	increase*
USA	22.00	28.00	22.50	10.00	0.788	increase

Table 3. Regression models for the association of RSV for ECC-related information with YLDs and Internet penetration in different countries, regarding total population, females, and males.

		BRA	FRA	GBR	ITA	MEX	SPA	USA	
RSV and YLDs	TOTAL	R2	0.038	0.959	0.004	0.582	0.054	0.172	0.395
		Constant	1586.89	-6518.64	-143.42	-2663.00	683.95	-1570.11	1949.90
		<i>B</i>	1.783	-9.405	0.021	0.057	-0.462	-3.874	1.320
		<i>SE</i>	4.929	2.940	3.572	0.725	4.784	5.466	1.097
		β	0.432	-0.300	0.002	0.016	-0.029	-0.321	0.654
		<i>p</i>	0.723	0.008*	0.996	0.938	0.925	0.491	0.250
	FEMALES	R2	0.037	0.951	0.004	0.646	0.051	0.164	0.399
		Constant	1463.82	-6442.86	-139.81	-3659.84	662.19	-1426.86	1976.67
		<i>B</i>	1.609	-8.287	0.043	-3.483	-0.232	-2.199	1.295
		<i>SE</i>	4.706	3.048	3.997	2.270	4.460	3.574	1.039
		β	0.386	-0.279	0.004	-0.378	-0.015	-0.256	0.671
		<i>p</i>	0.738	0.002*	0.992	0.149	0.959	0.549	0.235
	MALES	R2	0.040	0.963	0.004	0.029	0.062	0.19	0.389
		Constant	1734.29	-6539.89	-149.06	134.58	738.46	-1870.91	1912.54
		<i>B</i>	1.994	-10.113	-0.013	-2.211	-1.000	-8.678	1.329
<i>SE</i>		5.173	2.849	3.451	7.805	5.072	9.737	1.162	
β		0.488	-0.310	-0.001	-0.139	-0.059	-0.463	0.628	
<i>p</i>		0.706	0.005*	0.997	0.782	0.847	0.389	0.273	
RSV and Internet penetration	R2	0.116	0.943	0.683	0.594	0.583	0.149	0.434	
	Constant	-10054.19	-6626.60	-2380.76	-4059.08	7765.74	-3592.87	1814.65	
	<i>B</i>	-1.512	-0.281	-0.11	-0.254	0.826	-0.446	0.282	
	<i>SE</i>	1.33	0.119	0.254	0.404	0.3	1.237	0.18	
	β	-3.91	-0.291	-0.181	-0.405	3.151	-1.069	0.615	
	<i>p</i>	0.276	0.037	0.672	0.54	0.016	0.724	0.141	

The Cochrane-Orcutt estimation was applied to adjust linear models with residuals serially correlated over time. The asterisk indicates the significance of model. *B*, unstandardized coefficient; *SE*, standard error.

Table 4. ECC-related topics and their respective RSV values for each country

BRA		FRA		GBR		ITA	
Tooth decay (Syndrome)	100	Tooth decay (Syndrome)	100	Tooth decay (Syndrome)	100	Tooth (Topic)	100
Baby bottle (Topic)	87	Tooth (Topic)	76	Tooth (Topic)	92	Tooth decay (Syndrome)	100
Infant (Topic)	17	Deciduous teeth (Topic)	70	Infant (Topic)	55	Milk (Dairy product)	98
Child (Topic)	14	Infant (Topic)	27	Toddler (Topic)	45	Child (Topic)	17
Tooth (Topic)	12	Baby bottle (Topic)	13	Deciduous teeth (Topic)	24	Pain (Medical condition)	9
Mouth (Topic)	6	Dentist (Occupation)	8	Bottle (Beer container)	18	Dentistry (Branch of medicine)	5
Dentistry (Branch of medicine)	5	Milk (Dairy product)	3	Baby bottle (Topic)	18	Molar (Tooth)	5
Pediatric dentistry (Medical specialty)	4	Risk (Topic)	1	Child (Topic)	18	Medical treatment (Diagnostic center in Skala Oropou,Greece)	2
Deciduous teeth (Topic)	3	Syndrome (Topic)	1	Medical Treatment (Diagnostic center in Skala Oropou,Greece)	12	Dentist (Occupation)	2
Medical treatment (Diagnostic center in Skala Oropou, Greece)	2	Pain (Medical condition)	1	10- Cause (Topic)	10		
Brush (Topic)	2	Molar (Tooth)	1	Dentistry (Branch of medicine)	8		
Oral hygiene (Topic)	2	Cure (Topic)	1	Gums (Topic)	6		
Year (Topic)	2	Doctolib (Company)	1	Brush (Topic)	6		
Dentist (Occupation)	2	Medical treatment (Diagnostic center in Skala Oropou,Greece)	1	Milk (Dairy product)	6		
Childhood (Topic)	1	Wisdom (Topic)	1	Night (Topic)	4		
Antibiotics (Drug type)	1	Carrie (Film series)	1	Breastfeeding (Topic)	4		
Prevalence (Topic)	1	Stain (Topic)	1	Molar (Tooth)	4		
Age (Topic)	1	Toothpaste (Topic)	1	Sleep (Topic)	4		
Pediatrics (Branch of medicine)	1	Carrie (1976 film)	1	Toothpaste (Topic)	4		
Pacifier (Topic)	1	Wisdom tooth (Topic)	1	Fluoride (Chemical compound)	2		
Hygiene (Topic)	1	Gums (Topic)	1	Tooth enamel (Topic)	2		
Disease (Topic)	1			Tooth brushing (Topic)	2		
				Pain (Medical condition)	2		
MEX		SPA		USA			
Tooth decay (Syndrome)	100	Tooth decay (Syndrome)	100	Tooth decay (Syndrome)	100		
Baby bottle (Topic)	35	Childhood (Topic)	36	Infant (Topic)	78		
Dentistry (Branch of medicine)	24	Tooth (Topic)	30	Tooth (Topic)	74		
Childhood (Topic)	14	Child (Topic)	22	Baby bottle (Topic)	41		
Child (Topic)	12	Deciduous teeth (Topic)	20	Early childhood caries (Topic)	40		
Human tooth (Topic)	9	Dentistry (Branch of medicine)	13	Bottle (Beer container)	38		
Deciduous teeth (Topic)	8	Human tooth (Topic)	13	Deciduous teeth (Topic)	21		
Tooth (Topic)	7	Medical Treatment (Diagnostic center in Skala Oropou, Greece)	7	Dentistry (Branch of medicine)	13		
Medical treatment (Diagnostic center in Skala Oropou,Greece)	7	Preventive healthcare (Topic)	6	Child (Topic)	13		
Pediatric dentistry (Medical specialty)	5	Helminthiasis (Disease)	6	Toddler (Topic)	9		
Childhood obesity (Topic)	3	Pediculosis (Topic)	4	Medical Treatment (Diagnostic center in Skala Oropou,Greece)	9		
Pulpotomy (Topic)	2	Prevalence (Topic)	3	Cause (Topic)	5		
Hygiene (Topic)	2	Lily of the valley (Plants)	3	Childhood (Topic)	5		
Dentition (Topic)	2	Incidence (Topic)	3	Breastfeeding (Topic)	5		
Mouth (Topic)	2	Oral hygiene (Topic)	3	Health (Topic)	4		
Health (Topic)	2	Risk (Topic)	1	Preventive healthcare (Topic)	4		
Dental consonant (Topic)	2	Security (Topic)	1	Dentist (Occupation)	3		
Oral hygiene (Topic)	2	Health (Topic)	1	Risk (Topic)	3		
Breastfeeding (Topic)	1	Dental restoration (Topic)	1	Pediatrics (Branch of medicine)	3		
Prevalence (Topic)	1	Head lice infestation (Disease)	1	Early childhood education (Topic)	3		
Preventive healthcare (Topic)	1	Definition (Topic)	1	Human mouth (Topic)	3		
Syndrome (Topic)	1			Gums (Topic)	2		
				Fluoride (Chemical compound)	2		
				Oral hygiene (Topic)	2		
				Signs and symptoms (Topic)	2		

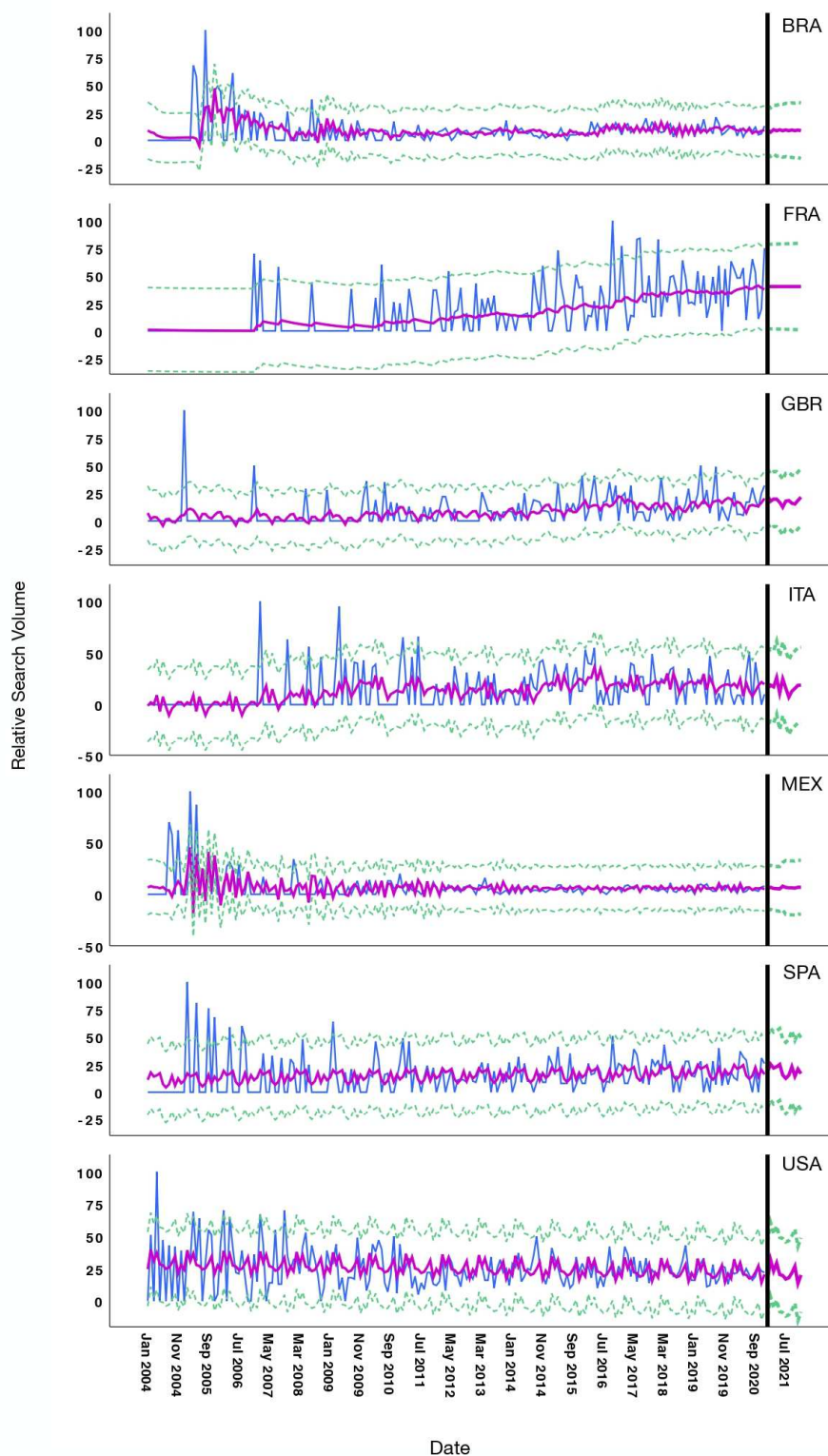


Figure 1. Predictive charts of RSV values for ECC-related searches performed in Brazil, France, United Kingdom, Italy, Mexico, Spain, and United States. The curves of observed values (blue lines), fit and forecast values (purple lines), and upper and lower bound of confidence intervals (green lines) are depicted from January 2004 through December 2020. RSV values presented after December 2020 (black line) represent 12- month forecasts. RSV: Relative Search Volume.



Figure 2. Heat map of monthly variation of the predictive generalized additive model values for ECC-related interest in all countries.

2.2 ARTICLE 2 - Online quality and readability assessment of Early childhood caries information available on Websites from distinct countries: A cross-sectional study

Abstract

Aim To assess the quality and readability of ECC-related Web information available in English, Spanish, and Brazilian Portuguese language.

Methods This study assessed the quality and readability of information related to ECC in three different languages found on the most popular Internet search engines worldwide. Websites were retrieved from different search engines using specific strategies. DISCERN questionnaire, JAMA benchmark criteria, and language-based readability formulas were used by two independent investigators to evaluate the quality and readability of websites. Also, contents were categorised according to aetiology, prevention, and treatment of ECC. The statistical analysis was performed using Spearman's rank correlation coefficient, hierarchical clustering analysis by Ward's minimum variance method, and Mann-Whitney U test. P values < 0.05 were considered significant.

Results Digital contents of 177 websites were considered of easy readability and low quality, without differences between websites produced by health- or non-health-related authors. Also, websites with ≥ 3 categories of information, or those that recommended dental visits or fluoride toothpastes presented higher DISCERN scores than their counterparts.

Conclusion ECC-related digital contents were considered simple, accessible and of poor quality, independently of their language and authorship. These findings indicate the importance of professional counseling to empower parents in selecting and consuming adequate information towards the improvement of children's oral health

KEYWORDS: Early Childhood Caries; Prevention; Health behaviour; Quality of information; Internet; eHealth.

Introduction

Early childhood caries (ECC) is a chronic disease defined as “the presence of one or more decayed (non-cavitated or cavitated lesions), missing (due to caries), or filled tooth surfaces in any primary tooth in a child under the age of six” [Pitts et al., 2019]. It impacts children’s and families’ quality of life due to its negative consequences, such as chronic pain, infections, sleeping disorders, loss of school days, and tooth extraction [Pitts et al., 2019; GBD Oral Disorders Collaborators et al., 2020]. In theory, these specific problems and circumstances lead parents and caregivers to recognize the need for a particular information, e.g., recommendations for the management of ECC [Yardi et al., 2018].

Although physicians are still the number one source of health information [Ipsos, 2018], the sense of self-efficacy in relation to selfcare [Bandura, 1988], next to the Internet’s pervasiveness, could trigger parents and caregivers’ seeking behaviour for online oral health consultations as means to fulfill their needs for information by themselves. Nevertheless, the access to digital health-related mis- or dis-information could hamper the primary prevention of ECC, because of their respective unintentional or deliberated incompleteness, uncertainty, or ambiguity natures [Stahl, 2006], in combination with the challenges represented for socially marginalized young parents in assessing the quality of information [Greyson, 2018]. Additionally, low-quality information could negatively impact on the person-professional relationship, especially among individuals who ignore the expertise of dentists for conflicting with their own beliefs confirmed with what was found online [Tan and Goonawardene, 2017; Glick, 2017].

Based on the aforementioned reasons, the aim of this study was to analyse the quality and readability of ECC-related information available on the Internet.

Methods

Design

The present study assessed the quality and readability of information related to ECC in three different languages (English, Spanish and Brazilian Portuguese) found on the most popular Internet search engines worldwide. Two instruments were used to analyse the quality of information, the DISCERN questionnaire and the Journal of American Medical Association (JAMA) benchmark criteria. For the readability analysis, specific metrics were used for each language. Websites were categorized regarding to their content and dichotomized according to the nature of their authorship, recommendation of fluoride toothpaste usage and dental visits.

Since federal regulations consider that research using publicly available data does not involve human subjects, this protocol did not require institutional review board approval from the Council of Ethics in Human Research of the Bauru School of Dentistry, University of São Paulo (Brazil). The present study followed the STrengthening the Reporting of OBservational Studies in Epidemiology (STROBE) checklist for cross-sectional studies.

Search strategy and data collection

First, all cookies and history browser were cleared of the computer used for research prior data collection. Then, country and language were selected on the browser's advanced section for each language studied: United States for English, Mexico for Spanish, and Brazil for Portuguese; otherwise, results could be influenced by the automated set up region. Three Internet search engines were used to retrieve the websites according to the market share by country [Statcounter, 2019; Statcounter, 2019b; Statcounter, 2019a]. A search strategy was constructed by the association of

the most commonly used terms related to ECC. For that, a list of terms in each language was tested in the Google Trends tool to convey those with the most volume of interest. Then, several strategies associating different terms were consulted until the development of a final strategy based on the higher number of results retrieved. Final data collection was performed on February 7th, 2019 using the strategies [“infant tooth decay” + “nursing bottle tooth decay”] for retrieving English websites, [“caries de bebe”] for retrieving Spanish websites, and [“carie em bebê” + “cárie de mamadeira” + “carie de mamadeira”] for retrieving Brazilian Portuguese websites. Websites with access difficulties (link errors), academic information (scientific papers), video links, journal news, duplicates, and nonspecific sources (Q&As, propagandas, pictures, etc.) were excluded. Before the assessment, all links were registered on archive.today (<http://archivetoday.com/>), an online archiving system that allows web material to remain unaltered for future evaluation.

Categorisation of websites

All websites were dichotomized according to the nature of the authorship in health- (dental professionals, clinics, universities and health companies) or non-health-related authors (blogs and others), the diversity of contents regarding aetiology, prevention, and treatment of ECC (higher = detection of 3 issues; lower = detection of 2 or less issues), the recommendation of fluoride toothpaste usage (yes/no), and the recommendation of dental visits (yes/no).

Quality assessment

Two instruments were used for the quality assessment. The JAMA benchmark consists of the identification of four concepts on the website: authorship (author’s name, affiliations, and credentials), attribution (effective references of content), currentness

(presence of dates of posts and updates of information), and disclosure (the statement of any potential conflicts of interest) of websites, granting 1 point for each criterion fulfilled, with a total score of 0 to 4) [Silberg et al., 1997].

The DISCERN instrument is a standardized questionnaire consisting of 16 Likert-scale questions, where 1 represents not meeting the criteria and 5 completely meets the criteria, divided into three sections: the first section describes the reliability of the content (questions 1 to 8), the second section analyses the information regarded treatment options (questions 9 to 15), and the third section represents an overall assessment of the website (question 16). Only the results of the first and second sections are commonly used to determine the quality of information of written materials, categorised in very poor (scores between 15 and 26), poor (27 to 38), fair (39 to 50), good (51 to 62), and excellent (63 to 75) [Charnock et al., 1999].

Two examiners independently assessed all websites (PEAA and IA). Regarding both instruments, when a divergent judgment was observed between the examiners, the website was re-assessed to the achievement of a consensus score.

Readability

The readability measures indicate the reading difficulty of a text based on specific metrics, such as average sentence length (ASL) and average number of syllables per word (ASW). Those metrics were collected from the online tool Readable.io (Readable.io, Bolney, England) through the Uniform Resource Locator (URL) of the website or the direct input of the texts in the platform.

Specific metrics for each language were used in this study. The Flesch Reading Ease (FRE) and the Flesch-Kincaid Grade Level (FKGL) are common metrics used to evaluate the grade of difficulty of English texts [Flesch, 1948; Kincaid et al., 1975]. The

formulas predict the level of education that a person requires to understand a specific text: $FRE = 206.835 - (1.015 \times ASL) - (84.6 \times ASW)$. A text scored between 90 and 100 is considered easily understandable by an average 5th grader. Scores between 60 and 70 indicate a text easily understandable by 8th and 9th graders, while scores between 0 and 30 are related to difficult texts, understandable only by graduates. The FKGL scores are equivalent to the US grade level of education that the reader needs to be able to comprehend written contents: $FKGL = 0.39$ (average number of words per sentence) + 11.8 (average number of syllables per word) - 15.59. The Fernández-Huerta Readability Formula (FHRF) is a modified version of the Flesch Reading Ease formula for Spanish texts in which 0 is the score for the greatest difficulty and 100 is the score for the easiest reading, similarly to FRE (Fernandez Huerta, 1959). Its formula is expressed as $FHRF = 206.84 - (0.6 \times \text{number of syllables per 100 words}) - (1.02 \times \text{number of sentences per 100 words})$. Brazilian Portuguese contents were evaluated by the Brazilian Portuguese version of the Flesch Reading Ease (FRE-BP): $FRE-BP = 248.835 - (84.6 \times \text{syllables per word}) - (1.015 \times \text{words per sentence})$. The reading difficulty of written materials is indicated by the scores in very easy (75-100), easy (50-75), difficult (25-50), and very difficult (0-25) [Martins et al., 1996].

Data analysis

All data were analyzed with the Statistical Package for Social Science (v 25.0; SPSS, Chicago, USA). The internal consistency of DISCERN was determined by Cronbach alpha. The interexaminer reliability was determined by the intraclass correlation coefficient (ICC) for absolute concordance to DISCERN and JAMA scores. The correlations between distinct measures were demonstrated by the Spearman rank correlation coefficients. The websites of distinct dichotomized natures were compared by Mann-Whitney U test. The quality measures of different countries were compared

by Kruskal Wallis and Dunn's post-hoc tests. P values < 0.05 were considered significant for all analyses.

Results

A total of 297 websites were collected among the three languages, excluding 120 links due to link errors (n = 35), academic source (n = 23), video links (n = 6), news (n = 19), duplicates (n = 10) and nonspecific sources (n = 27). Finally, 177 links met the criteria for evaluation (English n = 64, Spanish n = 41, and Brazilian Portuguese n = 72). Details of exclusion criteria by language are shown on Figure 1.

DISCERN and JAMA benchmark presented good internal consistencies (Cronbach's alpha= 0.729 and 0.677, respectively). The inter-examiner reliability was excellent for both instruments (DISCERN ICC=0.81, 95% CI: 0.72-0.87; JAMA benchmark ICC=0.80, 95% CI: 0.74-0.85).

In an overall analysis, the mean of DISCERN indicated a very poor quality of information for the three languages (24.79±4.66) in concordance with the JAMA benchmark (1.38±0.96). In total, 75.10% of the websites were classified as very poor quality of information, 22.60% were classified as poor quality of information and only 2.30% as fair quality of information. There were no websites classified as either good or excellent quality of information. Also, the sum of partial DISCERN scores (sections 1 and 2) were fair correlated with the scores of section 3 ($\rho=0.55$, $P<0.001$). Regarding to JAMA criteria, only one website published in Portuguese presented the maximum score (JAMA = 4). Ninety-four websites (53.11%) showed information related to the authorship, affiliations or credentials, 143 websites (80.79%) did not reference the content, 97 websites (54.80%) dated the information posted, and only 19 websites

(10.73%) stated any potential conflict of interest. DISCERN and JAMA benchmark were weakly positive correlated ($\rho=0.25$, $P<0.001$).

Table 1 depicts the descriptive statistics of quality and readability scores according to distinct languages. The mean of FRE (63.65 ± 7.44) and FKGL (7.85 ± 1.61) for English websites indicated accessible and plain texts, easily understood by 7th – 8th graders. Only 2 websites were measured as difficult to read. The mean of Fernandez-Huerta formula (67.37 ± 11) showed that Spanish websites were easy to read. Only 5 websites were considered difficult and 2 were considered very easy to read. Finally, the FRE-BP demonstrated that Brazilian Portuguese websites were accessible and easy to understand (58.83 ± 8.47). Overall, quality and readability scores were not correlated significantly, except to Fernández- Huerta that showed weak negative correlations with DISCERN (S1 + S2) ($\rho = -0.430$, $P<0.001$) and JAMA scores ($\rho = -0.402$, $P<0.001$).

A total of 49.20% websites were related to health authors, 24.90% recommended fluoride toothpaste usage, and 71.80% indicated dental visits. Regarding specific languages, most websites published in Spanish (68.30%) and Brazilian Portuguese (59.70%) were developed by non-health authors, whereas 70.30% of websites published in English were developed by health-related authors. The DISCERN scores were significantly higher among websites that recommended fluoride toothpaste usage and dental visits in comparison to their counterparts. On the other hand, the authorship of websites did not significantly influence the quality of information (Table 2).

Most websites referred to etiology (94.40%) and prevention of ECC (96.60%), and only 17.50% described treatment options for the disease. DISCERN scores were significantly higher in websites of higher diversity contents in comparison with their counterparts, independently of country/language (Table 2).

Discussion and conclusion

To the best of our knowledge, this is the first study that evaluated the quality of ECC-related information in different languages. Our findings showed a predominance of accessible, understandable, and low-quality content on the Internet, independently of the country and language, although the quality of information and their readability were only slightly negatively correlated in Spanish websites. Interestingly, the nature of the authorship did not influence the results, i.e., the quality of information produced even by health-related authors was also very poor. On the other hand, contents with higher diversity presented a better-quality level of information in all languages. In general, websites recommending the use of fluoride dentifrices and dental visits were better qualified than their counterparts. These results presented a good interexaminer reliability and internal consistency as proven with a previous study [Aguirre et al., 2017].

The almost ubiquitous availability of the Internet permitted the easy access to information, defeating spatial and temporal barriers. Not surprisingly, online search engines represent the second most used source for health information among adults worldwide [Ipsos, 2018]. Parents and caregivers who refer to websites for information regarding their children's health could rely on the content they found [Kubb and Foran, 2020], especially if it confirms something they already believe, read or heard before [Meppelink et al., 2019]. Hence, the lack of quality control of Web information could jeopardise user's health condition based on poor-quality contents [Kubb and Foran, 2020]. In this sense, a challenge sets when the internet is filled with blogs consisting in experiences, health beliefs and opinions regarding ECC exposing their readers to misinformation, e.g. "...avoid fluoride toothpastes because they are toxic".

Since the production of good quality information is a complex and arduous task, which involves the need of skilled specialists, trained to attain target audiences and produce accessible contents according to specific guidelines and instruments, the poor quality of information authored by health professionals might indicate their technical incapacity combined with a primary interest in producing oral health information for advertisements and commercial profit, linked to a possible marketing sense of obligation in developing a website for describing treatments to their patients [Mourad et al., 2020]. It can be confirmed by a great volume of content replication found among diverse websites. Indeed, professionals seemed to be unconscious about their responsibility in producing good quality educational information to stimulate healthy behaviours; instead, they are more concerned to show their current practice describing procedures not necessarily committed to the evidence-based dentistry. In addition, the better qualification of more diverse information emphasizes the importance of previous knowledge and professional planning to achieve contents of adequate educational performance [Kulasegarah et al., 2018].

Although dental caries is the most prevalent disease in the world (GBD Oral Disorders Collaborators et al., 2020), its impact in children could still be disregarded. For instance, oral health education during pregnancy [Paglia, 2019], the need of early dental visits, the importance of the maintenance of primary teeth for adequate oral and general health conditions are probably overlooked by parents [Colombo et al., 2019]. In this scenario, most websites described information related to natural breastfeeding, formulas and dental hygiene as formal “protocols” or steps that should be followed by parents. Hypothetically, strict non-individualized recommendations could dissuade parents and caregivers towards good oral health habits for not fitting in their own routines and lifestyles [Ferrazzano et al., 2019; Lotto et al., 2020].

The consumption of digital contents empowers people to discuss diverse aspects of their own health during the shared decision-making, causing possible negative effects on the relationship with professionals when based on beliefs and misinformation [Kubb and Foran, 2020]. In this scenario, when health professionals are welcoming towards their patients' Internet retrievals, patients are more positive about their relationship, otherwise they seemed challenged facing a negative perception, especially if they have concerns about the quality of the information gathered [Tan and Goonawardene, 2017]. Acknowledging the vast volume of oral health misinformation online and, moreover, the challenges that represent for parents to appraise good quality of information [Greyson, 2018], health professionals have the responsibility to guide their patients toward clear and accurate information, encouraging them to discuss their findings, especially when conditions such as ECC are behavior dependent [Pitts et al., 2019]. To improve health literacy levels, the greater and frequent involvement with health information should be stimulated, enhancing their proper capability in dealing with contents. Therefore, online search techniques as clicking restraint and lateral reading should be instructed to support consumers in the analysis of the veracity of information [Wineburg and McGrew, 2019]. It consists in analyzing thoroughly the links depicted before selecting a website, scanning a webpage without diving completely in its content and opening new tabs laterally regarding the same subject to compare results. Also, triangulation is a method used by researchers to cross-check information consulting different perspectives and sources [Barnes and Vidgen, 2006], including the use of escalation of authorities to verify any information, e.g. hearing information about parenthood, consulting first with the Internet, discussing it within their inner circle, to finally verify it with their health care provider [Greyson, 2018]. Altogether, these techniques emphasize the need to consult different sources to confirm an information

of interest. In addition, the production of good quality educational contents, although a challenging task, is critical to tackle misinformation. Important aspects must be considered before the production of health content, e.g., the disclosure of objectives, authorship, references that could be listed under de main text [Winker et al., 2000], and dates of first posting and updates shown to avoid possible hazards for health information consumers when reading information out of date (National Library of Medicine-NIH, 2018).

The present study has some limitations. First, although the total score could have been influenced by the section 2 of the DISCERN questionnaire, when a website did not provide information regarding treatment options, the independent analysis of sections 1 and 3 also corroborated with the same classification of poor quality of information. Second, although other populations are also native speakers of these three languages, the core of quality of ECC-related information presented here must be regarded specifically by selected countries, since other websites could be available for distinct geographical areas. Third, both assessment instruments were developed to assess the quality of only textual contents, i.e., graphic elements, figures and pictures were disregarded in this evaluation. For the same reason, 7.72% links related to scientific papers were also excluded from this analysis; however, it is presumed that lay readers would have a limited interested in formal contents, especially due to their difficult in understanding technical language and terms.

In conclusion, the information related to ECC published in websites of Brazil, Mexico, and USA were considered of poor quality, presenting accessible and simple contents for most population. The awareness on what parents and caregivers are exposed to when searching for ECC online could be a relevant approach to improve the management of the disease, considering that the information influences the parental

decision-making process decisively. Therefore, paediatric dentists should receive sufficient technical formation and stimulation to produce Web contents directly to parents and children, focused on the spread of information related to good oral health habits for the prevention and control of ECC.

Acknowledgements

This research was funded by the São Paulo Research Foundation (#2018/02563-6). The funder had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

- Aguirre PEA, Coelho MM, Rios D, Machado MAAM, Cruvinel AFP, Cruvinel T. Evaluating the dental caries-related information on Brazilian websites: qualitative study. *J Med Internet Res* 2017; 19, e415.
- Bandura A. Organisational Applications of Social Cognitive Theory. *Aust J Manag* 1988; 13, 275-302.
- Barnes SJ, Vidgen RT. Data triangulation and web quality metrics: A case study in e-government. *Inf Manag* 2006; 43, 767-777.
- Charnock D, Shepperd S, Needham G, Gann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. *J Epidemiol Community Health* 1999; 53, 105-111.
- Colombo S, Gallus S, Beretta M, Lugo A, Scaglioni S, Colombo P, Paglia M, Gatto R, Marzo G, Caruso S. Prevalence and determinants of early childhood caries in Italy. *Eur J Paediatr Dent* 2019; 20, 267-273.
- Fernandez Huerta J. Simple readability measures. *Consigna* 1959; 214, 29-32.
- Ferrazzano G, Sangianantoni S, Mitrano R, Ingenito A, Alcidi B, Cantile T. Assessing changes in oral health-related quality of life and body growth in 3-5 years old children following dental treatment under general anaesthesia due to severe dental caries. *Eur J Paediatr Dent* 2019; 20, 214-218.
- Flesch R. A new readability yardstick. *J Appl Psychol* 1948; 32, 221.
- GBD Oral Disorders Collaborators, Bernabe E, Marcenes W, Hernandez C, Bailey J, Abreu L, Alipour V, Amini S, Arabloo J, Arefi Z. Global, regional, and national levels and trends in burden of oral conditions from 1990 to 2017: a systematic analysis for the global burden of disease 2017 study. *J Dent Res* 2020; 99, 362-373.
- Glick M. Believing is seeing: Confirmation bias. *J Am Dent Assoc* 2017; 148, 131-132.
- Greyson D. Information triangulation: A complex and agentic everyday information practice. *J Assoc Inf Sci Technol*. 2018; 69, 869-878.
- Ipsos 2018. Global views on healthcare: Percentage of adults worldwide who used select sources for health information as of 2018; 84 [Online]. Available: <https://www.ipsos.com/sites/default/files/ct/news/documents/2018-07/global-views-healthcare-report-2018-07-24.pdf> [Accessed May 2020].
- Kincaid JP, Fishburne Jr RP, Rogers RL, Chissom BS. Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel. Institute for Simulation and Training 1975.
- Kubb C, Foran HM. Online Health Information Seeking by Parents for Their Children: Systematic Review and Agenda for Further Research. *J Med Internet Res* 2020; 22, e19985.
-

Kulasegarah J, McGregor K, Mahadevan M. Quality of information on the Internet—has a decade made a difference? *Ir J Med Sci* 2018; 187, 873-876.

Lotto M, Strieder AP, Ayala Aguirre PE, Andrade Moreira Machado MA, Rios D, Cruvinel A, Cruvinel T. Parental perspectives on early childhood caries: a qualitative study. *Int J Paediatr Dent* 2020; 30(4):451-8.

Martins TB, Ghiraldelo CM, Nunes MdGV, Oliveira Junior ONd. Readability formulas applied to textbooks in brazilian portuguese. *Icm-sc-Usp* 1996. Meppelink CS, Smit EG, Franssen ML, Diviani N. “I was right about vaccination”: Confirmation bias and health literacy in online health information seeking. *J Health Commun* 2019; 24, 129-140.

Mourad M, Brendebach J, Alkilzy M, Splieth C. Choosing a paediatric dentist: Factors influencing parents’ decision. *Eur J Paediatr Dent* 2020; 21, 74-79.

National Library of Medicine (NIH). 2018. Evaluating Internet Health Information: A Tutorial [Online]. Available: <https://medlineplus.gov/webeval/webeval.html> [Accessed May 2020].

Paglia L. Oral prevention starts with the mother. *Eur J Paediatr Dent* 2019; 20, 173-173.

Pitts NB, Baez RJ, Diaz-Guillory C, Donly KJ, Alberto FC, McGrath C, Phantumvanit P, Seow WK, Sharkov N, Songpaisan Y. Early Childhood Caries: IAPD Bangkok Declaration. *J Dent Child (Chic)* 2019; 86, 72. Silberg WM, Lundberg GD, Musacchio RA. Assessing, controlling, and assuring the quality of medical information on the Internet: Caveant lector et viewer—Let the reader and viewer beware. *JAMA* 1997; 277, 1244-1245.

Stahl BC. On the Difference or Equality of Information, Misinformation, and Disinformation: A Critical Research Perspective. *Inf Sci* 2006; 9. Statcounter. 2019. Search Engine Market Share Brazil [Online]. Available: <http://gs.statcounter.com/search-engine-market-share/all/brazil> [Accessed February 2019].

Statcounter. 2019 -a. Search Engine Market Share Mexico [Online]. Available: <http://gs.statcounter.com/search-engine-market-share/all/mexico> [Accessed February 2019].

Statcounter. 2019 -b. Search Engine Market Share United States Of America [Online]. Available: <http://gs.statcounter.com/search-engine-market-share/all/united-states-of-america> [Accessed February 2019]. Tan SS-L, Goonawardene N. Internet health information seeking and the patient-physician relationship: a systematic review. *J Med Internet Res* 2017; 19, e9.

Wineburg S, McGrew S. Lateral reading and the nature of expertise: Reading less and learning more when evaluating digital information. *Teach Coll Rec* 2019; 121, 1-40.

Winker MA, Flanagan A, Chi-Lum B, White J, Andrews K, Kennett

RL, DeAngelis CD, Musacchio RA. Guidelines for medical and health information sites on the internet: principles governing AMA web sites. *JAMA* 2000; 283, 1600-1606.

Yardi S, Caldwell PH, Barnes EH, Scott KM. Determining parents' patterns of behaviour when searching for online information on their child's health. *J Paediatr Child Health* 2018; 54, 1246-1254.

Table 1. Descriptive statistics of quality and readability scores according to the languages of websites

	Outcomes	S1*	S2*	S3*	DISCERN (S1+S2)	JAMA†	FKGL‡	FKRE§	FH¶	FRE-BP#
Overall	Mean (SD)	15.94 (2.98)	8.84 (2.80)	1.93 (0.61)	24.79 (4.66)	1.38 (0.96)	-	-	-	-
	Median	16	8	2	24	1	-	-	-	-
	Minimum	9	3	1	16	0	-	-	-	-
	Maximum	27	21	4	44	4	-	-	-	-
USA (English)	Mean (SD)	16.05 (2.94) ^a	9.25 (3.04) ^a	1.92 (0.72) ^a	25.30 (5.27) ^a	1.09 (1.05) ^a	7.85 (1.61)	63.65 (7.44)	-	-
	Median	15.50	8.50	2	24	1	7.8	65.45	-	-
	Minimum	12	7	1	19	0	4.60	37.60	-	-
	Maximum	27	21	4	44	3	11.30	78.30	-	-
Mexico (Spanish)	Mean (SD)	15.61 (2.87) ^a	9.22 (3.59) ^a	1.85 (0.57) ^a	24.83 (5.05) ^a	1.46 (0.87) ^{a,b}	-	-	67.37 (11)	-
	Median	15	8	2	24	2	-	-	67.29	-
	Minimum	9	3	1	16	0	-	-	43.34	-
	Maximum	23	21	3	40	3	-	-	94.98	-
Brazil (Portuguese)	Mean (SD)	16.04 (3.11) ^a	8.26 (1.86) ^a	1.97 (0.53) ^a	24.31 (3.80) ^a	1.58 (0.88) ^b	-	-	-	58.83 (8.47)
	Median	16	7	2	23.50	2	-	-	-	59.32
	Minimum	9	7	1	18	0	-	-	-	33.24
	Maximum	27	15	3	36	4	-	-	-	75.20

*S1, S2, and S3: 3 different sections of DISCERN.

†The Journal of American Medical Association.

‡Flesch Kincaid Grade Level.

§Flesch Kincaid Reading Ease.

¶Fernandez-Huerta.

#Flesch Reading Ease adapted to Brazilian Portuguese.

Table 2. The comparison of means (SD) of quality and readability scores between dichotomised categories of websites published in distinct languages. Different superscript letters represent significant statistical differences between dichotomised categories regarding the same quality and readability measure (Mann-Whitney U test, $P < 0.05$).

		S1	S2	S3	DISCERN (S1+S2)	JAMA	FKGL	FRE	FH	FRE-BP		
Overall	Authorship	Health	15.74 (0.27) ^a	8.87 (0.24) ^a	1.90 (0.07) ^a	24.61 (0.42) ^a	1.33 (0.10) ^a	-	-	-	-	
		Non-health	16.14 (0.35) ^a	8.81 (0.35) ^a	1.96 (0.06) ^a	24.96 (0.56) ^a	1.42 (0.10) ^a	-	-	-	-	
	Dentist	Yes	16.27 (0.28) ^a	9.17 (0.27) ^a	1.96 (0.06) ^a	25.44 (0.45) ^a	1.43 (0.08) ^a	-	-	-	-	
		No	15.12 (0.32) ^b	8.00 (0.28) ^b	1.84 (0.08) ^a	23.12 (0.41) ^b	1.24 (0.15) ^a	-	-	-	-	
	Fluoride	Yes	17.39 (0.48) ^a	9.34 (0.48) ^a	2.11 (0.10) ^a	26.73 (0.70) ^a	1.70 (0.14) ^a	-	-	-	-	
		No	15.47 (0.24) ^b	8.68 (0.23) ^b	1.86 (0.05) ^b	24.14 (0.39) ^b	1.27 (0.08) ^b	-	-	-	-	
English	Authorship	Health	15.36 (0.33) ^a	8.73 (0.31) ^a	2.49 (0.08) ^a	25.89 (0.59) ^a	1.02 (0.13) ^a	7.91 (0.23) ^a	64.11 (0.83) ^a	-	-	
		Non-health	17.68 (0.87) ^b	10.47 (1.02) ^a	2.84 (0.16) ^a	30.05 (1.78) ^a	1.26 (0.31) ^a	7.72 (0.41) ^a	62.57 (2.47) ^a	-	-	
	Clusters	1	17.17 (4.04) ^a	12.50 (4.95) ^b	2.75 (0.75) ^a	32.00 (8.50) ^b	1.50 (0.09) ^a	8.37 (1.56) ^a	58.83 (9.95) ^a	-	-	
		2	15.79 (2.60) ^a	8.50 (1.76) ^a	2.56 (0.57) ^a	26.00 (4.08) ^a	1.00 (0.03) ^a	7.74 (1.61) ^a	64.77 (6.35) ^a	-	-	
	Dentist	Yes	16.52 (0.44) ^a	9.58 (0.47) ^a	2.67 (0.09) ^a	28.02 (0.86) ^a	1.19 (0.15) ^a	7.75 (0.25) ^a	64.33 (1.07) ^a	-	-	
		No	14.63 (0.50) ^b	8.25 (0.51) ^a	2.38 (0.13) ^a	24.44 (0.88) ^b	0.81 (0.28) ^a	8.08 (0.31) ^a	61.63 (1.86) ^a	-	-	
	Fluoride	Yes	18.62 (0.80) ^a	10.54 (0.96) ^a	3.08 (0.18) ^a	31.31 (1.30) ^a	1.77 (0.30) ^a	7.20 (0.47) ^a	65.22 (2.74) ^a	-	-	
		No	15.39 (0.36) ^b	8.92 (0.40) ^b	2.47 (0.08) ^b	26.06 (0.75) ^b	0.92 (0.14) ^b	8.02 (0.22) ^a	63.25 (0.95) ^a	-	-	
	Spanish	Authorship	Health	15.31 (0.76) ^a	9.92 (0.82) ^a	2.54 (0.18) ^a	25.23 (1.27) ^a	1.54 (0.24) ^a	-	-	69.20 (3.40) ^a	-
			Non-health	15.75 (0.56) ^a	8.89 (0.73) ^a	2.36 (0.12) ^a	24.64 (1.01) ^a	1.43 (0.17) ^a	-	-	66.52 (1.98) ^a	-
Clusters		1	15.88 (0.46) ^a	9.59 (0.67) ^a	2.50 (0.10) ^a	25.47 (0.93) ^a	1.44 (0.16) ^a	-	-	65.87 (1.95) ^a	-	
		2	14.67 (1.23) ^a	7.89 (0.79) ^a	2.11 (0.26) ^a	22.56 (1.18) ^a	1.56 (0.29) ^a	-	-	72.70 (3.19) ^a	-	
Dentist		Yes	15.78 (0.49) ^a	9.41 (0.66) ^a	2.41 (0.12) ^a	25.19 (0.93) ^a	1.50 (0.15) ^a	-	-	65.12 (1.73) ^a	-	
		No	15.00 (1.09) ^a	8.56 (1.00) ^a	2.44 (0.18) ^a	23.56 (1.46) ^a	1.33 (0.33) ^a	-	-	75.34 (3.98) ^b	-	
Fluoride		Yes	17.25 (0.52) ^a	9.58 (1.26) ^a	2.75 (0.18) ^a	26.83 (1.24) ^a	1.33 (0.26) ^a	-	-	64.04 (2.40) ^a	-	
		No	14.93 (0.55) ^b	9.07 (0.61) ^a	2.28 (0.11) ^a	24.00 (0.96) ^b	1.52 (0.16) ^a	-	-	68.75 (2.19) ^a	-	
Portuguese		Authorship	Health	16.52 (0.46) ^a	8.62 (0.38) ^a	2.03 (0.12) ^a	25.14 (0.75) ^a	1.72 (0.19) ^a	-	-	-	58.04 (1.60) ^a
			Non-health	15.72 (0.49) ^a	8.02 (0.26) ^a	1.93 (0.07) ^a	23.74 (0.54) ^a	1.49 (0.12) ^a	-	-	-	59.37 (1.28) ^a
	Clusters	1	17.08 (1.10) ^a	11.00 (0.62) ^b	2.23 (0.17) ^a	28.08 (1.38) ^b	1.69 (0.29) ^a	-	-	-	56.55 (2.46) ^a	
		2	15.81 (0.37) ^a	7.66 (0.14) ^a	1.92 (0.65) ^a	23.47 (0.38) ^a	1.56 (0.11) ^a	-	-	-	59.34 (1.09) ^a	
	Dentist	Yes	16.31 (0.51) ^a	8.56 (0.29) ^a	2.02 (0.08) ^a	24.88 (0.61) ^a	1.60 (0.13) ^a	-	-	-	58.95 (1.17) ^a	
		No	15.50 (0.42) ^a	7.67 (0.29) ^b	1.88 (0.11) ^a	23.17 (0.48) ^a	1.54 (0.19) ^a	-	-	-	58.62 (1.90) ^a	
	Fluoride	Yes	16.63 (0.88) ^a	8.37 (0.34) ^a	2.00 (0.13) ^a	25.00 (0.90) ^a	1.89 (0.19) ^a	-	-	-	56.44 (1.55) ^a	
		No	15.83 (0.39) ^a	8.23 (0.23) ^a	1.96 (0.07) ^a	24.06 (0.52) ^a	1.47 (0.12) ^a	-	-	-	59.69 (1.22) ^a	

^aS1, S2, and S3: 3 different sections of DISCERN.

^bJAMA: the Journal of American Medical Association.

^cFKGL: Flesch Kincaid Grade Level

^dFRE: Flesch Reading Ease

^eFH: Fernandez Huerta readability metric

^fFRE-BP: Flesch Reading Ease adapted to Brazilian Portuguese

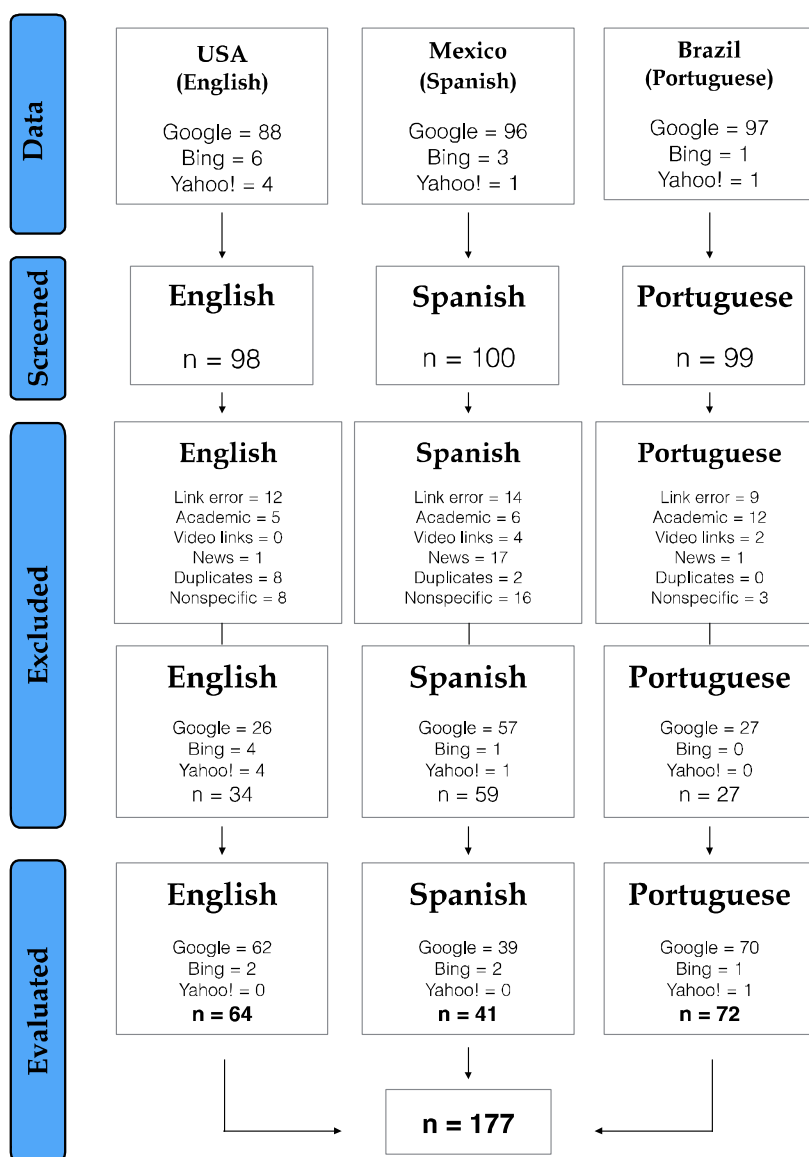


Figure 1. Flowchart depicting the systematic selection of ECC-related websites by language.

2.3 ARTICLE 3 - A cross-sectional completeness assessment of Youtube videos' information regarding early childhood caries

Abstract

Purpose: The present study aimed to assess the completeness of videos regarding early childhood caries (ECC) available on Youtube in 3 different languages.

Design: Search strategies were developed in English, Spanish and Portuguese for the collection of the first 60 videos of each language depicted on Youtube. The completeness of videos was assessed by a 17-topic checklist regarding ECC information. Videos were dichotomized according to the median of the completeness score and the nature of their authorship (health and non-health authors) to compare groups. The statistical analysis was performed using Spearman rank correlation coefficient and Mann-Whitney U test. $P < 0.05$ values were considered significant.

Results: One hundred and twenty videos met the inclusion criteria. Overall, information related to the etiology and prevention of ECC was incomplete (median = 5, Q1-Q3 = 3-7). The completeness score was not correlated to other characteristics of videos, whereas the interaction metrics as views, likes, dislikes and viewing rates were significantly correlated between them. It was found a predominance of videos produced by health authors, although channels of non-health authors presented higher number of subscribers. Interestingly, the frequent snacking of sugary foods and beverages was the most popular topic mentioned in videos.

Conclusions: In conclusion, Youtube videos containing information about etiology and prevention of ECC presented incomplete information in all 3 languages. In this sense, dental practitioners should embrace the benefits of social media to enhance health literacy skills of laypeople toward the correct identification of good quality information on the Internet.

Keywords: Dental caries, Deciduous tooth, Early childhood caries, YouTube, Social media

Introduction

Early childhood caries affects 532 million children around the world [1], representing a major public health and economic challenge [2, 3]. Considering its persistent increase, especially in low-income countries [1], the International Association of Paediatric Dentistry (IAPD) emphasizes the importance to improve the health literacy of parents, caregivers, and health care providers as a primary need for reducing the burden of the disease [4]. In this sense, through health education strategies contributes to strengthen health literacy levels and reduce health disparities [5-7], especially for the prevention of chronic diseases as ECC, avoiding the maintenance of risk behaviors, presence of untreated caries and dental pain [8].

At the beginning of 2021, social media had over 4.33 billion users that accessed contents massively via smartphones worldwide [9]. This almost ubiquitous availability stimulates users to exploit features for their daily needs, resulting in higher probabilities of seeking for health information [9-11], a behavior determined by a set of social and motivational factors that has notably increased over the years, such as social support, self-care and the interest to improve the communication with their physician [12-14]. Within this context, Youtube is an Internet-based video sharing platform commonly consulted by parents and caregivers interested in children-related health information videos [10], i.e., this platform can influence the adoption of healthy behaviors mainly among low health literacy populations [15, 16]. The type of content, the completeness of the information, the duration, and the accessibility of information presented are determinant aspects to the engagement of users with the content. On the other hand, the vast amount of misinformation found in social media can hamper the outcomes of this autonomous education process [17].

Taking into consideration the aforementioned evidence, the present study aimed to assess the completeness of the information regarding ECC available on Youtube videos produced in 3 different languages.

Methods

Search strategy and data collection

This cross-sectional study collected ECC-linked data from Youtube in November 2019. Prior data collection, specific country and language was set with a paid Virtual Private Network (VPN) (<https://www.expressvpn.com/what-is-vpn>) to avoid automated regional results. Then, a search strategy was developed by the association of relevant terms in 3 different languages, as follows: [*'early childhood caries' + 'bottle caries' + 'baby tooth decay' + 'baby bottle tooth decay'*] for American English, [*'caries infantil' + 'caries rampante' + 'caries de biberon' + 'caries en dientes de leche' + 'caries de la primera infancia'*] for Mexican Spanish, and [*'cárie em bebê' + 'carie em bebe' + 'cárie de mamadeira' + 'carie de mamadeira' + 'carie da primeira infância'*] for Brazilian Portuguese. Two independent researchers (PEAA and IA) screened videos regarding the exclusion criteria: a) videos showing clinical procedures, b) duplicates, c) blogs unrelated to the issue, d) videos produced in other languages, e) animated videos, and f) brushing technique content without ECC information.

Video characteristics

Quantitative data of subscribers, number of views, likes, dislikes, number of days since upload, and duration were collected for each video. The interaction index $\{[(\text{number of likes} - \text{number of dislikes})/\text{number of views}] * 100\}$ and the viewing rate $[(\text{number of views}/\text{number of days since upload}) * 100]$ were calculated.

Completeness of information

A previous study analyzed the usefulness of Youtube ECC videos for parental education, based on the AAPD policy on early childhood caries [18]. It consists in a scale of 13 items, being 1 point granted for the identification of each item in the videos. For this analysis, we decided to expand the scale to 17 topics, to qualify the completeness of the information regarding the etiology and prevention of the disease, according to the IAPD Bangkok Declaration [4]. The content items of the videos were also assessed by 2 investigators (PEAA and IA) independently. The items were grouped according to the presence of representative images, ECC effects (general health effects, local health effects, and effects on permanent teeth), importance of primary teeth, etiology (nighttime bottle-feeding, prolonged breastfeeding, and frequent sugary foods and beverages snacking), and prevention of the disease (brushing, use of fluoridated toothpaste, dietary advice for sugar-containing beverages, dietary advice on the avoidance of nighttime bottle-feeding, dietary advice on the avoidance on-demand breastfeeding after eruption of the first primary tooth, dietary advice on the avoidance of frequent snacking on sugar-containing foods and beverages, use of topical fluoride, recommendations of early dental visits, and recommendations on the decreasing transmission of microorganisms).

Video categorization

Videos were dichotomized based on the median of the completeness score and according to the nature of their authorship in health authors (health care professionals and educational/professional organizations) and non-health authors (TV channels/news agencies, healthcare accounts, and laypeople blogs).

Data analysis

Statistical analysis was performed with the Statistical Package for Social Science (version 25.0; SPSS, Chicago, USA). The interrater reliability of the completeness score provided by the independent examiners was assessed by intraclass correlation coefficient (ICC) for the absolute concordance. The significant differences between the dichotomized groups of the natures of authorship and completeness score were observed by Mann-Whitney U test. Correlations between distinct measures were demonstrated by the Spearman's correlation coefficients. *P* values <0.05 were considered significant for all analyses.

Results

An excellent inter-examiner concordance level was observed for the overall analyses of videos (ICC=0.86, 95% CI 0.43-0.94), ranging from 0.85 to 0.87. A total of 120 video links were included in this study (American English n=40, Mexican Spanish n=37, and Brazilian Portuguese n=43). Detailed exclusion criteria are described in Figure 1.

Video characteristics

The videos were collected from 2005 to 2019. The median number of views was 3,660 (Q1-Q3 = 495 - 18,705.25), with a total of 5,329,943 views regarding the 3 languages. The mean length of videos was 4 min and 33 s (ranging from 34 s to 12 min and 59 s). The median of likes and dislikes were 17 (Q1-Q3 = 2.25 - 89.50) and 2 (Q1-Q = 0 - 12), respectively. The mean completeness score was 5.22 (\pm 3.15), with a median of 5.00 (Q1-Q3= 3 - 7). An American channel presented the higher number of subscribers (1,650,000). The interaction index was positive and low, with a mean of 1.20% (SD =

± 2.34 ; median = 0.34%). In average, the videos were watched 2,032.17 times daily. The video with the highest viewing rate was found among Mexican links (61,567.73). Table 1 depicts the descriptive statistics of the videos according to distinct languages. Videos produced in Spanish presented significantly higher numbers of subscribers ($P = 0.011$) and views ($P = 0.041$) in comparison to videos in Portuguese. Additionally, videos produced in Portuguese had significantly higher duration ($P = 0.018$) and lower number of days since upload ($P = 0.003$) than English videos. Finally, videos in Portuguese had a significant higher interaction index than the other 2 languages ($P < 0.001$). No statistical differences were shown for likes, dislikes, viewing rates and completeness scores among the 3 languages.

Completeness of information

Overall, the most popular issues were related to frequent sugary foods and beverages snacking (60%), toothbrushing (59.20%), nighttime bottle-feeding (54.20%), recommendations of early dental visits (45.80%), and the presence of representative images (45%). The less popular information supported the avoidance of on-demand breastfeeding after eruption of the first primary tooth (7.50%), and prolonged breastfeeding (10%), as well as to recommend the application of topical fluoride (11.70%) and the use of fluoridated toothpaste (13.30%) (Table 2). Only one video published in English presented the maximum score of 17, with a duration of 10 min and 11 s. Fifty-seven videos scored between 1 and 4, and 63 videos scored between 5 and 17. No statistical differences were detected regarding the comparison of dichotomized groups of completeness score, except for the proper score ($P < 0.001$) (Table 3).

Misinformation

Few videos mentioned the use of homemade strategies to prevent dental caries, or presented images of cavities only, associating the disease exclusively with severe stages of its progression. Also, images of children using incorrect training cups (sippy cups) were identified. Nevertheless, most misinformation detected was linked to the use of fluoridated toothpaste, e.g., fluoride concentration and toxicity, and alternative products for toothbrushing as coconut oil; however, the main argument for the use of fluoride-free toothpaste was the risk of fluoride intake by children. Unfortunately, this misinformation was presented by a pediatric dentist on a popular morning TV show, probably attained by a greater audience than the Youtube channel.

Authorship

Overall, videos were predominantly authored by health care professionals or educational/professional organizations (61.70%); however, this pattern was not observed among Spanish videos (45.90%). The channels of non-health authors had significantly more subscribers ($P = 0.001$), except in Brazil. In addition, non-health videos in Spanish also presented more views ($P = 0.014$), likes ($P = 0.045$), dislikes ($P = 0.045$), and higher viewing rates ($P = 0.036$). The mean time since the upload of Brazilian Portuguese videos produced by health authors was significantly shorter than those produced by non-health authors ($P = 0.028$) (Table 3).

Correlations

In general, the duration of the videos presented a weak positive correlation with interaction rate ($r_s = 0.23$, $P = 0.011$). The number of days since the upload of videos had a positive weak correlation with views ($r_s = 0.23$, $P = 0.013$) and a good negative

correlation with interaction rate ($r_s = -0.62$, $P < 0.001$). Weak significant positive correlation was observed between subscribers and views ($r_s = 0.36$, $P < 0.001$), likes ($r_s = 0.36$, $P < 0.001$), dislikes ($r_s = 0.35$, $P < 0.001$), and viewing rate ($r_s = 0.39$, $P < 0.001$). Views had a strong positive correlation with likes ($r_s = 0.80$, $P < 0.001$), dislikes ($r_s = 0.86$, $P < 0.001$), and viewing rate ($r_s = 0.97$, $P < 0.001$). Likes had a strong correlation with dislikes ($r_s = 0.87$, $P < 0.001$) and viewing rate ($r_s = 0.85$, $P < 0.001$), and a weak positive correlation with interaction index ($r_s = 0.38$, $P < 0.001$). Finally, dislikes had a strong correlation with viewing rate ($r_s = 0.87$, $P < 0.001$) (data not shown).

Discussion

These findings showed that information related to etiology and prevention of ECC found in YouTube videos is usually incomplete, being not correlated to interaction metrics. While videos developed by health authors were predominant, videos developed by non-health authors were more popular with higher number of subscribers. Interestingly, likes and dislikes were strongly correlated, whereas likes and interaction rates were weakly correlated. Also, the interaction rate of Brazilian Portuguese videos was significantly higher than those of other languages. As one of the main causes and preventive strategies of the disease were frequently cited the snacking of sugary foods and beverages and toothbrushing, respectively. Distinctly, the presentation of prolonged breastfeeding as a risk factor of the disease was uncommon, not being mentioned in any of the Spanish videos.

Regardless the consistent global digital divide [19], no significant differences were observed between view numbers of videos from developed and developing countries; however, videos produced in Spanish were more viewed than those produced in

Brazilian Portuguese. Considering that internet mobile penetration has increased notably in those countries over time [19], and parents and caregivers report the use of mobile phones for health information [20, 21] this seeking behavior tend to continue increasing with the expansion of internet access [11], with the self-perception of credibility of online information, and the sense of capacitation in obtaining, interpreting and using available Web content, especially in developing countries [22], which could justify higher views and interaction rate observed in Mexican and Brazilian Portuguese videos, respectively.

The scale of completeness of information has been previously called as “usefulness of video content” synthesized by the “utility score” [23-25]. In our opinion, the concept of utility of health information is quite complex, not being determined simply by the presence or absence of a topic found in a checklist. Contrarily, an informational video can be planned and developed to present a specific aspect of a condition; hence, its aims, intelligibility and correctness also must be considered to define usefulness appropriately. Although these findings showed a similar duration of videos scored <5 or ≥ 5 , only one video with 10 min and 11 s achieved the maximum score 17 (<https://www.youtube.com/watch?v=ZmacipkzRY>). Regarding the overall mean of duration of videos was 4 min and 33, the completeness of information still seems to depend on longer time for its presentation, which can be confirmed by the increase of duration of videos from 277 s to 376 s regarding percentiles 75th and 95th for completeness score. It could discourage the audience to watch the whole content of videos, which might explain the weak positive correlation between the interaction rate and duration found in these results. From the latter situation emerges a dilemma: the same piece of information could be simultaneously complete by the technical perspective and incomplete by the consumption perspective. For instance, 2 different

educational videos produced by the American Dental Association (ADA) about ECC scored 13 (<https://www.youtube.com/watch?v=SYGYTU7VP3I>) and 8 (<https://www.youtube.com/watch?v=SlwiXRmILyA>). Nevertheless, we believe that is speculative stating those videos as ineffective educational sources for parents and caregivers, regarding only the lack of some aspects of the disease considered vital for all consumers. Overall, videos produced by health professionals were shorter and presented lower completeness scores than other videos, although no significant differences were observed. In this sense, titles or description of videos should state clearly the aim of the information transmitted, as highlighted by quality assessment instruments for health information [26, 27]. The production of good quality educational videos requires effort and planning, which emphasizes the importance of dental associations and health organizations in producing information and advising dentists to disseminate accurate information [28].

In our study, we observed a strong correlation between views and likes. The engagement of social media users can be explained by different behavioral theories; e.g., the motivated seeking health information behavior is a relevant predictor for likes [29]. It is expected that parents and caregivers feel the need to seek information about their children's health [30, 31], turning to social media to respond their doubts [12]. For instance, Youtube is considered a helpful source for breastfeeding information [12, 32]. In our results, prolonged breastfeeding as a causing factor for the development of ECC were the least topic mentioned in all videos. Therefore, the knowledge about parental backgrounds and demands is fundamental for maximizing the use of social media as health communication tools in direction to significant education strategies.

The almost ubiquitous access of Internet and the freedom of speech empower social media users to engage in digital communities by sharing their experiences, beliefs or

asking for advice regarding a disease [33]; however, available content could mislead consumers and hamper health outcomes [34]. For instance, we usually identified misinformation such as home recipes for the management of dental caries, use of images showing the incorrect amount of toothpaste, and recommendations of fluoride-free toothpastes.

Although the present study assessed the completeness of correct information by the identification of topics from a predefined list containing different aspects of the etiology and prevention of ECC, we did not assess the quality of information using specific instruments, or still quantify the number of misinformation detected in those videos; nevertheless, a video was granted with 1 point only if a specific aspect was presented according to evidence provided by AAPD and IAPD guidelines [4, 35]. Also, subjects commonly associated with misinformation were identified during the analysis. In addition, these results must be interpreted with caution since they represent the information uploaded in a specific timeframe in 3 different languages and regions.

In conclusion, according to the criteria of this study, Youtube videos that presented information about etiology and prevention of ECC in English, Spanish and Brazilian Portuguese invariably focused on partial aspects of the disease. To combat misinformation and establish a trustful digital environment for population, dental teams are the main responsible for counseling people to obtain and consume good quality information towards the improvement of their own health literacy skills.

Acknowledgements

This research was funded by the São Paulo Research Foundation (#2018/02563-6). The funder had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

1. GBD Oral Disorders Collaborators, Bernabe E, Marcenes W, Hernandez C, Bailey J, Abreu L, et al. Global, regional, and national levels and trends in burden of oral conditions from 1990 to 2017: a systematic analysis for the global burden of disease 2017 study. *J Dent Res.* 2020;99(4):362-73. doi: 10.1177/0022034520908533.
 2. Kassebaum N, Bernabé E, Dahiya M, Bhandari B, Murray C, Marcenes W. Global burden of untreated caries: a systematic review and metaregression. *Journal of dental research.* 2015;94(5):650-8. doi: 10.1177/0022034515573272.
 3. Bernabé E, Marcenes W. Can minimal intervention dentistry help in tackling the global burden of untreated dental caries? *British Dental Journal.* 2020;229(7):487-91. doi: 10.1038/s41415-020-2155-9.
 4. Pitts NB, Baez RJ, Diaz-Guillory C, Donly KJ, Alberto FC, McGrath C, et al. Early Childhood Caries: IAPD Bangkok Declaration. *J Dent Child (Chic).* 2019;86(2):72. doi: 10.1111/ipd.12490.
 5. World Health Organization. Regional Office for the Eastern Mediterranean. Health education: theoretical concepts, effective strategies and core competencies-A foundation document to guide capacity development of health educators. EMRO Publications, 2012 978-92-9021-828-9. Available from: <https://apps.who.int/iris/handle/10665/119953>.
 6. Nutbeam D, Lloyd JE. Understanding and responding to health literacy as a social determinant of health. *Annual Review of Public Health.* 2021;42. doi: 10.1146/annurev-publhealth-090419-102529
 7. Nutbeam D, Muscat DM. Health Promotion Glossary 2021. *Health Promot Int.* 2021. doi: 10.1093/heapro/daaa157.
 8. Batista MJ, Lawrence HP, de Sousa MdLR. Oral health literacy and oral health outcomes in an adult population in Brazil. *BMC Public Health.* 2018;18(1):60. doi: 10.1186/s12889-017-4443-0.
 9. Hootsuite. Digital 2021. Global overview report. 2021. Available from: https://hootsuite.widen.net/s/zcdrtxwczn/digital2021_globalreport_en.
 10. Pretorius K, Johnson KE, Rew L. An integrative review: understanding parental use of social media to influence infant and child health. *Matern Child Health J.* 2019;23(10):1360-70. doi: 10.1007/s10995-019-02781-w.
 11. Jiang S, Liu PL. Digital divide and Internet health information seeking among cancer survivors: a trend analysis from 2011 to 2017. *Psychooncology.* 2020;29(1):61-7. doi: 10.1002/pon.5247.
 12. Pretorius K, Johnson KE, Rew L. An integrative review: understanding parental use of social media to influence infant and child health. *Matern Child Health J.* 2019;23(10):1360-70. doi: 10.1007/s10995-019-02781-w.
-
-

13. 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;92(3):426-31. doi: 10.1016/j.pec.2013.06.020.
 14. Laranjo L, Arguel A, Neves AL, Gallagher AM, Kaplan R, Mortimer N, et al. The influence of social networking sites on health behavior change: a systematic review and meta-analysis. *J Am Med Inform Assoc.* 2015;22(1):243-56. doi: 10.1136/amiajnl-2014-002841.
 15. Drozd B, Couvillon E, Suarez A. Medical YouTube videos and methods of evaluation: literature review. *JMIR Med Educ.* 2018;4(1):e3. doi: 10.2196/mededu.8527.
 16. Roberts M, Callahan L, O'Leary C. Social media: A path to health literacy. *Inf Serv Use.* 2017;37(2):177-87.
 17. Suarez-Lledo V, Alvarez-Galvez J. Prevalence of health misinformation on social media: systematic review. *J Med Internet Res.* 2021;23(1):e17187. doi: 10.2196/17187.
 18. ElKarmi R, Hassona Y, Taimeh D, Scully C. YouTube as a source for parents' education on early childhood caries. *Int J Paediatr Dent.* 2017;27(6):437-43. doi: 10.1111/ipd.12277.
 19. Zavazava C, Skhirtladze R, Eriksson F, Magpantay E, Montenegro L, Pokorna D, et al. Measuring the Information Society Report. 2018. Available from: <https://www.itu.int/pub/D-IND-ICTOI>.
 20. Haluza D, Böhm I. Mobile and online health information: Exploring digital media use among Austrian parents. *International Journal of Environmental Research and Public Health.* 2020;17(17):6053. doi: 10.3390/ijerph17176053.
 21. Casilang CG, Stonbraker S, Japa I, Halpern M, Messina L, Steenhoff AP, et al. Perceptions and Attitudes Toward Mobile Health in Development of an Exclusive Breastfeeding Tool: Focus Group Study With Caregivers and Health Promoters in the Dominican Republic. *JMIR Pediatrics and Parenting.* 2020;3(2):e20312. doi: 10.2196/20312.
 22. Wang X, Shi J, Kong H. Online health information seeking: A review and meta-analysis. *Health Commun.* 2020:1-13. doi: 10.1080/10410236.2020.1748829.
 23. Hassona Y, Taimeh D, Marahleh A, Scully C. YouTube as a source of information on mouth (oral) cancer. *Oral Dis.* 2016;22(3):202-8. doi: 10.1111/odi.12434.
 24. Fortuna G, Schiavo JH, Aria M, Mignogna MD, Klasser GD. The usefulness of YouTube™ videos as a source of information on burning mouth syndrome. *J Oral Rehabil.* 2019;46(7):657-65. doi: 10.1111/joor.12796.
 25. Hamdan AA, Shaqman M, Abu Karaky A, Hassona Y, Bouchard P. Medical reliability of a video - sharing website: The gingival recession model. *Eur J Dent Educ.* 2019;23(2):175-83. doi: 10.1111/eje.12417.
-

26. Charnock D, Shepperd S, Needham G, Gann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. *J Epidemiol Community Health*. 1999;53(2):105-11. doi: 10.1136/jech.53.2.105.
27. National Library of Medicine (NIH). Evaluating Internet Health Information: A Tutorial. 2018 [cited 2021 May]; Available from: <https://medlineplus.gov/webeval/webeval.html>.
28. Dhar V. Social Media and Its Implications in Pediatric Dentistry. *Pediatr Dent*. 2018;40(3):174-6.
29. Khan ML. Social media engagement: What motivates user participation and consumption on YouTube? *Comput Human Behav*. 2017;66:236-47. doi: 10.1016/j.chb.2016.09.024.
30. Bernhardt JM, Felter EM. Online pediatric information seeking among mothers of young children: results from a qualitative study using focus groups. *J Med Internet Res*. 2004;6(1):e7. doi: 10.2196/jmir.6.1.e7.
31. Kubb C, Foran HM. Online Health Information Seeking by Parents for Their Children: Systematic Review and Agenda for Further Research. *J Med Internet Res*. 2020;22(8):e19985. doi: 10.2196/19985.
32. Henshaw EJ, Cooper MA, Jaramillo M, Lamp JM, Jones AL, Wood TL. "Trying to figure out if you're doing things right, and where to get the info": parents recall information and support needed during the first 6 weeks postpartum. *Matern Child Health J*. 2018;22(11):1668-75. doi: 10.1007/s10995-018-2565-3.
33. Chou W-YS, Oh A, Klein WM. Addressing health-related misinformation on social media. *JAMA*. 2018;320(23):2417-8. doi: 10.1001/jama.2018.16865.
34. Wang Y, McKee M, Torbica A, Stuckler D. Systematic literature review on the spread of health-related misinformation on social media. *Soc Sci Med*. 2019;240:112552. doi: 10.1016/j.socscimed.2019.112552.
35. AAPD: American Academy of Pediatric Dentistry. The American Academy of Pediatric Dentistry research agenda. 2020. Available from: <https://bit.ly/37nUQYr>.
-
-

Table 1. Descriptive statistics of videos regarding their completeness scores and characteristics by language

		Completeness score	Duration (m)	Days since upload	Subscribers	Views	Likes	Dislikes	Interaction index	Viewing rate
Overall	Mean	5.22	4:33	2096.83	101009.66	44416.19	153.97	24.67	1.2	2032.17
	Median	5	3:07	1857	6915	3660	17	2	0.34	225.75
	IQ range	4	268	0:00	40747	18210	87	12	0.91	808.09
	Min	1	0:34	587	0	40	0	0	-0.95	4.01
	Max	17	12:59	5147	1650000	1331710	3900	1100	13.43	61567.73
English	Mean	5.8	3:34	2547.25	122759.75	50793.55	44.53	11.42	0.25	1791.87
	Median	5.50 ^a	02:13 ^a	2266.50 ^a	10080.00 ^{a,b}	2716.50 ^{a,b}	10.00 ^a	2.50 ^a	0.20 ^a	104.45 ^a
	IQ range	4	2:33	2440	82575	16502	55	9	0.3	866.85
	Min	1	0:34	600	0	105	0	0	-0.95	4.59
	Max	17	12:41	5147	1650000	1093112	769	225	1.47	26739.53
Spanish	Mean	4.97	5:01	2108.57	173836.16	58416.68	227.51	44.59	0.77	2918.56
	Median	4.00 ^a	03:33 ^{a,b}	1957.00 ^{a,b}	25800.00 ^a	10191.00 ^a	27.00 ^a	5.00 ^a	0.29 ^a	431.51 ^a
	IQ range	4	4:42	1411	75205	28749	182	22	0.44	1916.25
	Min	1	1:07	687	0	41	0	0	-0.28	4.86
	Max	12	12:59	4113	1630000	1331710	3900	1100	6.31	61567.73
Portuguese	Mean	4.88	5:05	1667.74	18112.35	26436.84	192.49	19.84	2.45	1503.97
	Median	4.00 ^a	03:41 ^b	1510.00 ^b	2160.00 ^b	3164.00 ^b	31.00 ^a	2.00 ^a	1.32 ^b	235.24 ^a
	IQ range	4	4:51	1423	11199	8732	115	6	3.02	528.13
	Min	1	0:53	587	0	40	0	0	0	4.01
	Max	12	12:58	4499	254000	567899	3300	441	13.43	33017.38

Table 2. Distribution of frequency of videos containing specific completeness score items by language

Completeness of information	Overall	English	Spanish	Portuguese
	N (%)	N (%)	N (%)	N (%)
The presence of representative images	54 (45.00)	19 (47.50%)	16 (43.20%)	19 (44.20%)
Local health effects	41 (34.20)	12 (30.00%)	12 (32.40%)	17 (39.50%)
General health effects	21 (17.50)	7 (17.50%)	4 (10.80%)	10 (23.30%)
Effect on permanent teeth	29 (24.20)	12 (30.00%)	8 (21.60%)	9 (20.90%)
Importance of primary teeth	27 (22.50)	11 (27.50%)	9 (24.30%)	7 (16.30%)
Etiology (Nighttime bottle-feeding)	65 (54.20)	22 (55.00%)	20 (54.10%)	23 (53.50%)
Etiology (Prolonged breastfeeding)	12 (10.00)	5 (12.50%)	6 (16.20%)	1 (2.30%)
Etiology (Frequent sugary foods and beverages snacking)	72 (60.00)	30 (75.00%)	23 (62.20%)	19 (44.20%)
Prevention (Brushing)	71 (59.20)	22 (55.00%)	25 (67.60%)	24 (55.80%)
Prevention (Fluoridated toothpaste)	16 (13.30)	8 (20.00%)	5 (13.50%)	3 (7.00%)
Dietary advice (Sugar-containing beverages)	47 (39.20)	17 (42.50%)	11 (29.70%)	19 (44.20%)
Dietary advice (Avoid nighttime bottle-feeding)	36 (30.00)	16 (40.00%)	6 (16.20%)	14 (32.60%)
Dietary advice (Avoid on-demand breastfeeding after eruption of first primary tooth)	9 (7.50)	7 (17.50%)	0 (0.00%)	2 (4.70%)
Dietary advice (Avoid frequent snacking on sugar-containing foods and beverages)	30 (25.00)	12 (30.00%)	9 (24.30%)	9 (20.90%)
Topical fluoride (varnish)	14 (11.70)	7 (17.50%)	3 (8.10%)	4 (9.30%)
Early dental visits	55 (45.80)	16 (40.00%)	18 (48.60%)	21 (48.80%)
Decreasing transmission of cariogenic bacteria	27 (22.50)	9 (22.50%)	9 (24.30%)	9 (20.90%)

Table 3. Distribution of the frequency of videos according to dichotomized categories of completeness score and authorship regarding video characteristics and languages. Different superscript letters represent significant statistical differences between dichotomized categories in each language (Mann-Whitney U test, $P < 0.05$)

		N (%)	Completeness score	Days since upload	Duration (s)	Subscribers	Views	Likes	Dislikes	Interaction index	Viewing rate	
Completeness	Overall	Scores <5	57 (47.50)	3 (2) ^a	1,797 (1924) ^a	191 (297) ^a	10,700 (39560) ^a	3,164 (15443) ^a	15 (73) ^a	2 (7) ^a	0.41 (0.99) ^a	207.40 (563.72) ^a
		Scores ≥5	63 (52.50)	7 (2) ^b	2,034 (1577) ^a	176 (250) ^a	3,730 (43099) ^a	4,379 (26057) ^a	20 (95) ^a	2 (18) ^a	0.27 (0.77) ^a	246.19 (1369.52) ^a
	English	Scores <5	16 (40.00)	3 (1) ^a	1,807.50 (2699) ^a	143 (529) ^a	21,782 (88743) ^a	1,167 (11330) ^a	10.50 (41) ^a	1 (6) ^a	0.31 (0.73) ^a	81.57 (502.59) ^a
		Scores ≥5	24 (60.00)	7 (3) ^b	2,719 (2146) ^a	122.50 (148) ^a	6,015 (57907) ^a	6,122 (32032) ^a	6.50 (66) ^a	3.50 (17) ^a	0.13 (0.21) ^a	165.93 (1176.85) ^a
	Spanish	Scores <5	19 (51.40)	3 (2) ^a	2,088 (1844) ^a	201 (253) ^a	30,200 (34100) ^a	10,191 (26286) ^a	27 (117) ^a	5 (13) ^a	0.31 (0.36) ^a	431.51 (1656.08) ^a
		Scores ≥5	18 (48.60)	6.5 (3) ^b	1,929.50 (1447) ^a	312 (417) ^a	6,920 (112205) ^a	9,179 (49129) ^a	18.50 (382) ^a	2 (41) ^a	0.22 (0.94) ^a	426.13 (3494.20) ^a
	Portuguese	Scores <5	22 (51.20)	3 (2) ^a	1,474 (1382) ^a	212.50 (537) ^a	1,132.50 (3773) ^a	2,315 (6166) ^a	29 (101) ^a	2 (6) ^a	1.33 (2.38) ^a	112.66 (537.01) ^a
		Scores ≥5	21 (48.80)	7 (3) ^b	1,510 (1560) ^a	221 (270) ^a	3,570 (32937) ^a	3,207 (9891) ^a	32 (118) ^a	2 (6) ^a	0.89 (3.17) ^a	256.36 (504.78) ^a
Authorship	Overall	Health	74 (61.70)	2 (4) ^a	1,809 (1674) ^a	157.50 (236) ^a	2,600 (21767) ^a	3,439 (14455) ^a	16 (75) ^a	2 (9) ^a	0.34 (1.22) ^a	225.75 (677.14) ^a
		Non-health	46 (38.30)	4 (4) ^a	1,991 (1474) ^a	207 (265) ^a	25,700 (114688) ^b	4,242 (26282) ^a	25.50 (111) ^a	2 (15) ^a	0.34 (0.71) ^a	226.80 (1220.06) ^a
	English	Health	28 (70.00)	5 (5) ^a	2,601.50 (2352) ^a	122.50 (141) ^a	2,625 (44797) ^a	3,132.50 (16142)	12.50 (55) ^a	4 (9) ^a	0.20 (0.28) ^a	250.04 (866.85) ^a
		Non-health	12 (30.00)	6 (5) ^a	1,800 (1876) ^a	149 (567) ^a	23,082 (770708) ^b	813.50 (15272) ^a	9 (69) ^a	1 (14) ^a	0.23 (0.50) ^a	68.32 (941.58) ^a
	Spanish	Health	17 (45.90)	6 (4) ^a	1,771 (1182) ^a	315 (453) ^a	3,040 (33274) ^a	3,593 (16119) ^a	10 (84) ^a	1 (9) ^a	0.24 (0.97) ^a	108.62 (1146.10) ^a
		Non-health	20 (54.10)	3.50 (3) ^b	2,350 (1447) ^a	195 (218) ^a	38,100 (161825) ^b	24,418 (38031) ^b	75.50 (324) ^b	12 (33) ^b	0.30 (0.27) ^a	762.68 (2470.65) ^b
	Portuguese	Health	29 (67.40)	4 (5) ^a	1,345 (1153) ^a	176 (279) ^a	2,160 (7719) ^a	3,980 (11730) ^a	47 (147) ^a	2 (11) ^a	1.33 (3.15) ^a	371.94 (616.86) ^a
		Non-health	14 (32.60)	5.50 (4) ^a	1,840.50 (1442) ^b	312 (330) ^a	1,922.50 (40850) ^a	1,464 (4945) ^a	18.50 (33) ^a	1.50 (2) ^a	0.91 (1.49) ^a	82.65 (245.90) ^a

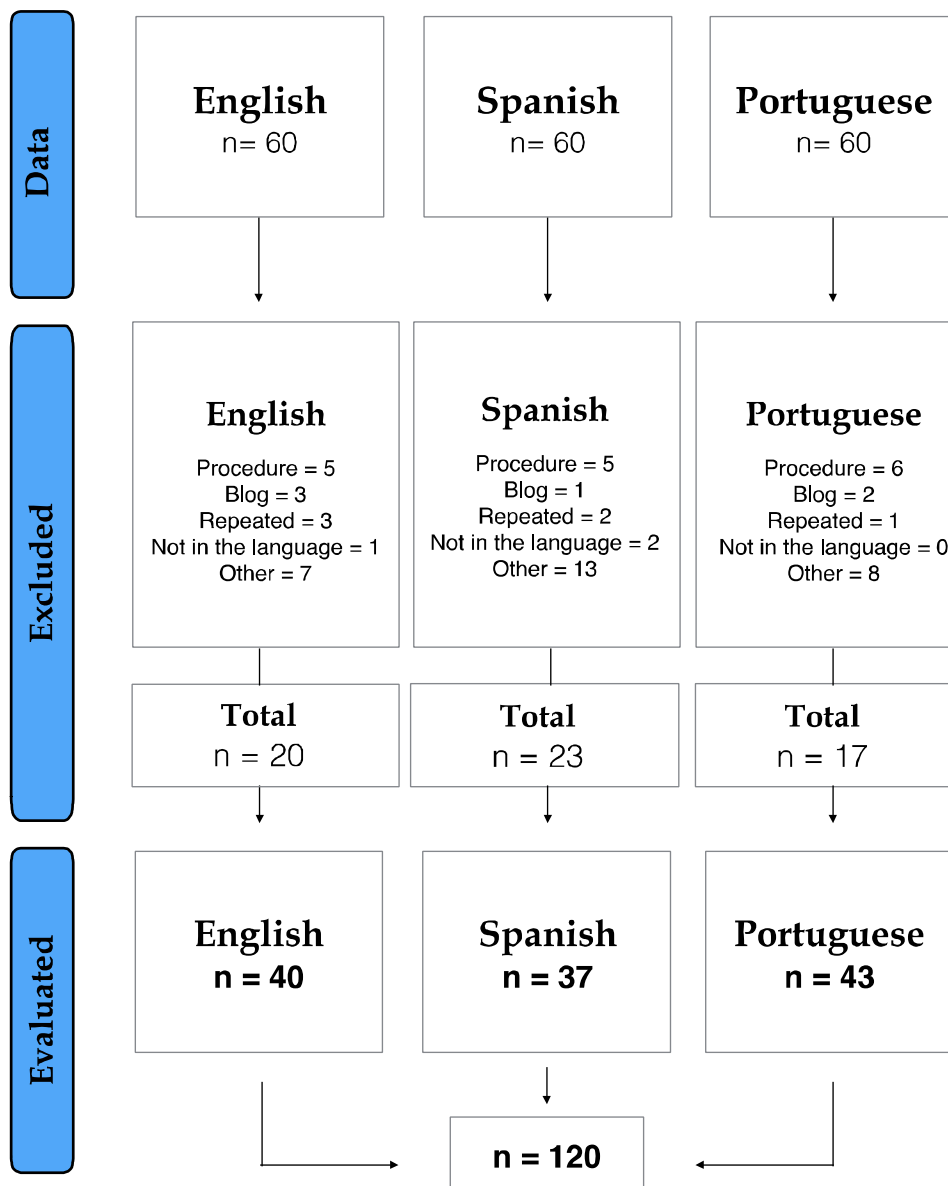


Figure 1. Flowchart depicting the systematic selection of Youtube videos containing ECC-related information, according to each language.

2.4 ARTICLE 4 - Tweeting about dental caries: an ecological digital study

Abstract

Twitter is a microblogging platform where active users communicate their ideas ranging vast topics, including oral health. In this sense, the analysis of this real-life content can contribute to the understanding of health conceptions and beliefs of specific populations. The present ecological digital study aimed to identify and characterize dental caries-related tweets from USA, Brazil, and Mexico. One thousand tweets were collected for each country using specific search strategies on Twitter API. After natural language processing, two independent investigators categorized tweets according to their sentiment (positive, neutral or negative), themes, and presence of misinformation (yes or no). Also, they compiled interaction metrics (retweets, replies, and likes) and information on authors' profiles (username, user's description, name, and followers). Additionally, sentiment analysis was performed using two artificial intelligence (AI) methods (VADER and TextBlob). Statistical analysis assessed the inter-examiner reliability (intraclass correlation coefficient), group comparisons (Chi-square, Kruskal-Wallis and Dunn's post-hoc tests), the accuracy of AI-based sentiment analysis (ROC curve), and the factors associated with interaction and misinformation (multiple binary logistic regression models) ($P < 0.05$). A good absolute concordance level between examiners were observed in all analysis of 2,669 tweets. Most tweets were related to the diagnosis of dental caries (23.0%) and negative sentiment (54.6%). It was observed significant higher percentages of positive tweets in the USA, and negative tweets and interaction in Brazil. Misinformation (18.40%) was associated with non-preventive themes (OR=19.90), positive sentiment (OR=7.35), and English language (OR=1.41), whereas tweets with interaction (38.80%) was not associated with any factor. The accuracy of VADER was higher than TextBlob to classify sentiments of these tweets. In conclusion, sentiment, interaction, themes and misinformation of tweets on dental caries were distinctly distributed between languages/countries, with predominance of interests on diagnosis and negative sentiment, differently from misinformation that was related to positive sentiment.

Keywords: Dental caries, Twitter, Social Media, eHealth, Misinformation

33 **Introduction**

34 Health information seeking behavior is increasingly on social media [1], where active
35 users consume and communicate ideas about diverse topics, such as perceptions of
36 oral health of themselves and their families [2]. Although this initiative is a desirable
37 mechanism of health participation and self-care [3, 4], these searches are frequently
38 motivated by emotional support, advice, and personal story report [5, 6], being
39 influenced by confirmation bias and cognitive dissonance [7, 8]. It means individuals
40 tend to accept messages aligned to their beliefs whereas refuse divergent ones,
41 independently the level of quality of information [7, 9]. Hence, the interaction of
42 consumers with content generates a myriad of novel assumptions, explanations, cause
43 and effect relationships, and interventions not supported by scientific evidence, but
44 which still find an eager audience among those with similar health values [10-12]. It
45 explains the faster propagation of misleading content in comparison to correct health
46 information on social media [13].

47 Twitter is a microblogging platform with approximately 436 million active users monthly
48 [14], with an audience predominantly composed by 25-34 year old people [15] and
49 males [16]. Usually, users post experiences of their routine life interacting with other
50 followers about diverse themes, such as politics, sports, entertainment real-time events
51 [17], and specific health topics as nutrition campaigns, smoking habits, influenza
52 symptoms [18], and COVID-19 and vaccination standpoints [19-21].

53 Dental caries is the most prevalent non-communicable oral disease worldwide,
54 although it is declining over the years [22]. This alarming situation impacts on people's
55 quality of life, hampering their oral health functions and daily activities [22, 23], which
56 might guide people to seek opinions and counseling, also sharing their knowledge
57 on social media [24-26]. All publications and interactions can be publicly and freely

58 accessed by researchers, aiming to understand the specific population needs on oral
59 conditions towards the development of new approaches for public health [27, 28] .
60 Previous studies demonstrated the interests of Internet users in information on dental
61 caries and its symptoms [29], also identifying the low quality of Web content to satisfy
62 population's demands [30-32]. Nevertheless, the Big Data provided by actions,
63 opinions and experiences posted by social media users in relation to the disease have
64 not yet been studied.

65 Based on the aforementioned evidence, this study aimed to identify and characterize
66 dental caries-related tweets from different countries, regarding interaction metrics,
67 authors' profiles, sentiment, themes, and the presence of misinformation.

68

69 **Methodology**

70 *Study design and ethics*

71 This ecological retrospective digital study analyzed tweets related to dental caries
72 shared by users from USA, Brazil and Mexico. One thousand tweets were collected for
73 each country using specific search strategies on the Twitter API. Two dental
74 investigators (PEEA and OSJ) were trained and calibrated to categorize 10% of tweets.
75 The calibration was confirmed by the intraclass correlation coefficient (ICC) for
76 absolute agreement, considering the analysis of 20% of tweets by language. After
77 natural language processing, tweets were categorized according to their sentiment
78 (positive, neutral or negative), themes, and presence of misinformation (yes or no). In
79 addition, investigators compiled information about interaction metrics (retweets,
80 replies, and likes) and authors' profiles (username, user's description, name, and
81 followers). Also, sentiment analysis was performed using two AI methods, VADER and
82 TextBlob. Statistical analysis assessed the inter-examiner reliability, group

83 comparisons, the accuracy of AI-based sentiment analysis, and the factors associated
84 with interaction and misinformation.

85 This study did not require institutional review board approval from the Council of Ethics
86 in Human Research of the *(track suppressed for blinded review)*.

87 *Search strategy and data collection*

88 Three different countries and their respective primary languages were selected for this
89 study: United States, English (USA); Brazil, Portuguese (BRA); and Mexico, Spanish
90 (MEX). They are listed among top 10 countries with most Twitter users [33]. Tweets
91 were scraped through Postman, an application recommended by the Twitter
92 Developers Platform to make requests to the Twitter API [34]. The Developers Platform
93 have enabled three collections for Twitter request. In this study, we used the Twitter
94 API v2. One hundred tweets of specific languages/countries were previously collected
95 to determine terms and hashtags to compose the final search strategies (Table 1).
96 According to the volume of tweets available for each country, different dates were
97 determined as USA (January 1, 2021), Brazil (January 1, 2019), and Mexico (January
98 1, 2017), resulting in 2,687, 1,914, and 1,073 tweets, respectively. The most recent
99 1,000 tweets per language were selected for analysis. All duplicates and non-dental-
100 caries-related content were excluded from analysis. Parameters of interaction metrics
101 (retweets, replies, and likes) and authors' profiles (author's id, name, user's number of
102 followers, and description) were determined for each tweet. The workflow of data
103 collection and analysis is summarized in Figure 1.

104

105 *Data preprocessing*

106 Raw data .json retrieved from Postman was converted to .csv format. Brazilian
107 Portuguese and Spanish content was translated into English with *Deep-Translator*, a

108 free and unlimited python package that supports multiple languages, integrating
109 different translators into one tool [35]. In this study, the translation was performed by
110 Google translator using the deep-translator library on Google Colaboratory notebook.
111 The number of “correct” translations to Brazilian Portuguese and Spanish were 84%
112 and 98.27%, respectively. After translation, duplicates or retweets were excluded
113 (filtered using the “RT” tag), resulting in a dataset that contained only original tweets
114 posted by individual users. Also, tweets containing information not related to dental
115 caries were excluded. The subsequent preprocessing steps of tweets depended on
116 specific analysis, as follows: removal of hashtags, mentions, quotations, and URLs
117 (sentiment analysis), with additional removal of punctuation marks, captions, emojis,
118 and stop words (word frequencies).

119

120 *Natural language processing (NLP)*

121 Natural language processing is a “*theoretically motivated range of computational*
122 *techniques for analyzing and representing naturally occurring texts at one or more*
123 *levels of linguistic analysis for the purpose of achieving human-like language*
124 *processing for a range of tasks or applications*”[36]. Those tasks or applications include
125 part-of-speech (POS) tagging, noun phrase extraction, sentiment analysis, etc. In this
126 study, NLP was performed to the analysis of sentiment and word frequencies related
127 to tweets about dental caries.

128 *Sentiment analysis*

129 We used three different approaches for sentiment analysis. First, the Valence Aware
130 Dictionary for sEntiment Reasoning (VADER) based on a wide lexicon to decode the
131 sentiment of each tweet into negative (≤ -0.05), neutral (0), and positive (≥ 0.05), using
132 a compound of values by measuring the intensity of words on a string [37]. Second,

133 we used TextBlob, also a lexical rule-based method analyzer that retrieves values
134 between -1 and +1, where -1 indicates an extremely negative sentiment and +1
135 extremely positive, whereas 0 is considered neutral [38]. Both sentiment analyzers
136 recognize punctuation marks, captions, emojis and jargons easily used in social media
137 content. VADER and TextBlob Python libraries on Google Colab were used for those
138 analysis. The third approach was human-based sentiment analysis, where two trained
139 and calibrated independent examiners analyzed all collected tweets into positive,
140 neutral, or negative sentiment (see details on calibration abovementioned).

141 *Word frequency*

142 To determine the frequency of words used in dental caries-related tweets, the content
143 was subjected to lowercasing and removal of stop words. After that, they were
144 submitted to tokenization, a process that identifies and separates each word or token
145 on a string. Finally, all tokens were lemmatized or reduced to its root form or a single
146 term. For this analysis, we used the *WordNetLemmatizer* from the NLTK library on
147 Google Colab notebook. The results were depicted on a word cloud.

148

149 *Determination of themes*

150 After consensus, the examiners categorized tweets into 12 topics that previously
151 emerged from the content analysis in the calibration process, as follows: 1 (no caries),
152 2 (toothache), 3 (dental fear), 4 (restorative procedures), 5 (prevention), 6 (diagnosis
153 of dental caries), 7 (political/social fact/comment), 8 (aesthetics/dental appearance
154 criticism), 9 (sugary foods, drinks and candies), 10 (other dental treatments), 11
155 (analogies/jokes), and 12 (dental professionals). The description and examples of
156 themes are provided as a supplemental file.

157

158 *Detection of misinformation*

159 Due to the unequivocal difficulties in recognizing the subjacent aims of Twitter users,
160 we considered the term *misinformation* as an umbrella concept that embraces false or
161 incorrect information published with or without the intention to cause harm [39]. All
162 information was analyzed by two independent dental investigators (PEEA and OSJ)
163 regarding scientific evidence on dental caries. Each tweet was categorized according
164 to the presence of misinformation (yes or no).

165

166 *Variables*

167 Interaction rate was calculated by the sum of the retweets, replies, and likes. Sexes of
168 authors were classified in male, female and unknown based on the information related
169 to their name, username, and description. To statistical analysis, categorical variables
170 were dichotomized as follows: presence of interaction (yes/no), sex (male-
171 unknown/female), misinformation (yes/no), language (English/non-English), sentiment
172 (positive-neutral/negative), themes (1, 5 and 9 = preventive/ others), and accounts
173 (health professionals/non-health professionals).

174

175 *Statistical analysis*

176 Data were analyzed with the Statistical Package for Social Science (version 25.0;
177 SPSS, Chicago, USA), as follows:

- 178 1. The intraclass correlation coefficient (ICC) was applied to assess the absolute
179 inter-examiner agreement for the analysis of human-based sentiment,
180 classification of themes, and identification of misinformation.
-

-
-
- 181 2. Kruskal-Wallis and post-hoc Dunn's tests were applied for the comparison of
182 countries in relation to interaction metrics, regarding the non-normal distribution
183 of data previously detected by the Shapiro-Wilk test.
- 184 3. The receiver operating characteristic (ROC) curve was employed to determine
185 the accuracies of AI-based sentiment analysis of VADER and TextBlob in
186 relation to the results of the human-based analysis.
- 187 4. Chi-square test was performed to evaluate differences in the distribution of
188 human-based sentiment, misinformation, and interaction regarding distinct
189 countries.
- 190 5. Multiple binary logistic regression models were conducted to determine the
191 association of dichotomized variables with interaction and misinformation,
192 regarding confounding factors. Only factors with Wald statistics with $p < 0.20$ in
193 prior univariate analyses were included in models.
- 194 6. For all analyses, p values < 0.05 were considered significant.

195

196 **Results**

197 A total of 2,669 tweets met the inclusion criteria, being 819 tweets from the USA, 923
198 from Brazil, and 927 from Mexico (Figure 1). An overall good absolute inter-examiner
199 agreement was observed for the sentiment analysis (USA ICC=0.82, BRA ICC=0.81,
200 and MEX ICC=0.74), classification of themes (USA ICC=0.84, BRA ICC=0.81, and
201 MEX ICC=0.65), and identification of misinformation (USA ICC=0.80, BRA ICC=0.90,
202 and MEX ICC=0.80). Also, 43.30% of the sample was identified as male, 43.60% as
203 female, and 13% as *unknown*. Only 6.10% of the accounts were classified as health
204 professionals, whereas 93.90% as non-health professionals (Table 2).

205

206 *Interaction*

207 Although differences were detected among the number of specific interaction metrics
208 as retweets, replies, and likes, no differences were observed in the interaction rates
209 between countries (Table 3). Also, regarding binary logistic regression analysis,
210 interaction was not associated with sex, gender, presence of misinformation, language,
211 sentiment, accounts, and themes.

212

213 *Sentiment*

214 In relation to human-based analysis, the accuracy of VADER (0.597) was higher than
215 TextBlob (0.565). Both tools presented a good sensitivity and a poor specificity (Table
216 4). Regarding the results of human-based sentiment analysis, most tweets were
217 negative (46.40%). Within countries, the percentages of positive and negative tweets
218 were significantly higher in USA (41,10%) and Brazil (62.80%), respectively (Table 2).

219

220 *Word frequency*

221 The word frequency is depicted on a word cloud (Figure 2). The most common word
222 found among tweets was 'cavity' with 1,600 mentions. Other frequent words detected
223 were 'teeth', 'tooth', 'caries', 'dentist', and 'decay'.

224 *Themes*

225 Overall, the 3 most mentioned themes were diagnosis of dental caries (19.60%),
226 prevention (15.20%), no caries (12,40%), and sugary foods, drinks and candies
227 (12.40%). Contrarily, the less mentioned themes were dental fear (1.70%), dental pain
228 (3.10%), and political/social fact/ comment (3.70%).

229 Table 5 presents the differences of distribution of sentiment, misinformation,
230 interaction, countries, sex, and accounts in relation to dichotomized themes. Tweets

231 containing preventive themes presented more negative sentiment, correct information,
232 and were more commonly authored by Brazilian and Mexican than tweets about other
233 themes.

234 Regarding specific countries, most tweets were about no caries (USA), diagnosis of
235 dental caries (Brazil), and prevention facts (Mexico). On the other hand, the lowest
236 number of tweets were about dental fear (USA and Mexico) and restorative procedures
237 (Brazil) (data not shown).

238

239 *Misinformation*

240 Out of 2,269 tweets, 491 (18.40%) were classified as containing misinformation and
241 they were significantly higher in USA (23.90% $p < 0.001$) than other countries (Table 2).

242 Tweets about sugary foods, drinks and candies were highly related to misinformation
243 (87.60%). Also, posts authored by unknown sex presented significantly more
244 misinformation (23.3%) than male (17.8%) and female (17.5%) ($p = 0.038$). Tweets
245 about toothache and dental fear were related to negative sentiment, whereas those
246 about no caries and sugary foods, drinks and candies were mostly positive and neutral.

247 In addition, only 6.3% of misinformation items were tweeted by professional accounts,
248 although it was not statistically different from non-professional accounts (19.1% vs.
249 18.3%), regarding frequencies into the same category of accounts.

250 The multiple logistic regression model demonstrated a significant positive association
251 of misinformation with other themes (OR=19.90), positive sentiment (OR=7.37), and
252 the English language (OR = 1.41) (Table 6).

253

254

255

256 Discussion

257 To our knowledge, this is the first evidence on the sentiment, themes of interest and
258 detection of misinformation in dental caries-related tweets from three different
259 countries and languages. Most tweets were related to the diagnosis of dental caries
260 and negative sentiment, with the word 'cavity' receiving the highest number of
261 mentions. It was observed significant higher percentages of positive tweets in the USA,
262 and negative tweets and interaction in Brazil. Misinformation was associated with non-
263 preventive themes, positive sentiment, and English language, whereas tweets with
264 interaction were not associated with any factor. The accuracy of VADER was higher
265 than TextBlob to classify sentiments of these tweets.

266 Dental caries represents the most prevalent condition among all non-communicable
267 diseases (NCDs) assessed in the Global Burden of Disease (GBD) study [23],
268 presenting common risk factors to other NCDs [40, 41], such as excessive sugar
269 consumption, poor oral hygiene, and low salivary flow [42]. Interestingly, our results
270 indicates the common misinformation about the direct and inevitable development of
271 caries lesions after the consumption of sugary foods. In this direction, some tweets can
272 be highlighted as follows: (1) *"Ironically, I still like baking, but we can't have cake for
273 every meal. Well, technically we could, but we'd probably give ourselves diabetes
274 and/or cavities."*; (2) *"Man having had 6 cavities filled makes me not wanna eat candy
275 ever again 🤔👉"*; (3) *"No more candy...no more cavities"*. Hence, Twitter users
276 consider to be doomed to develop carious lesions when intaking free sugars
277 disregarding the importance of specific strategies as mechanical removal of biofilm and
278 fluoride usage to prevent and control the disease [41]. The direct association between
279 two situations that are not strictly connected is known as illusion of causality [10, 43].
280 It is important to understand how these illusory beliefs could hamper people's health

281 when information contradicts their convictions [8], and consequently could affect the
282 process of learning evidence-based information [43]. In fact, different strategies can
283 help patients in acquiring scientific knowledge [10], which remarks the role of dentists
284 in educating their patients during consultations, even when information is not
285 requested.

286 Nevertheless, misinformation items were also identified among tweets from dental
287 professionals, e.g. *“give your partner a cleaning and resins this 14th. This way the*
288 *bacteria of calculus and cavities will not spread...”*. As seen in other studies, dental
289 caries-related information created by health professionals is not complete and accurate
290 [30, 44]. The production of health content involves a complex professional
291 responsibility, and its impact seems to be overlooked when commercial interests
292 overlap the needs to educate and inform their patients and the society. On the other
293 hand, it is important to state the scientific gap between the production of scientific
294 information to reach clinical protocols [45, 46]. Interestingly, misinformation was not
295 detected among fluoride-related tweets distinctly of other studies that assessed the
296 quality of information of dental caries on the world wide web [44].

297 Twitter is a platform that has been studied for many years for health topic surveillance
298 [18], especially because its near real-time posts allows the identification of public health
299 issues [47]. Recognizing social media users' concerns regarding dental caries is an
300 important strategy to prevent and manage the disease, to promote the engagement of
301 patients into healthy oral behaviors. Additionally, sentiment analysis is also a
302 barometer of Internet population perceptions [47]. People is well aware of the burden
303 caused by the disease, i.e., the diagnosis of dental caries might be considered a
304 negative situation. In a previous study, it was confirmed the association of the burden
305 of the disease with the volume of queries made on a search engine [29]. On the other

306 hand, the closed association of a positive sentiment with misinformation items is a
307 distressing situation for the dental community. Information disorders are classified in
308 intentional and unintentional dissemination of inaccurate information. For the purpose
309 of this study, the term *misinformation* was employed without identifying the intentions
310 of information. We compared two sentiment analyzers with human-based analysis
311 resulting in lower accuracies; nevertheless, when eliminating neutral sentiment, the
312 specificity of VADER incremented to 0.654, although TextBlob did not followed the
313 same pattern. It might be explained by the hypothesis they could still present difficulties
314 when assessing dental caries jargon or identifying irony [47]. Although both automatic
315 sentiment dictionaries are widely used in the medical research, for example, “no caries”
316 was a common tweet among the three countries, and even though is a positive feeling
317 transmitted by the user, it was classified negatively for the weights of the detected word
318 ‘no’. In contrast, tweets about ‘no caries’ with happy emojis did not influence the
319 compound results for both analyzers. Future studies could develop a wider dictionary
320 for dental caries tweets assessment aiming the analysis of more than 3,000 tweets.
321 Moreover, these results also demonstrate the use of automatic translators are accurate
322 for the study of this disease, at least in two countries, allowing the opportunity of
323 collecting a larger data set for analysis.

324 The present results must be interpreted with caution. First, the specific population
325 studied was restricted to the social media users, i.e., the sample may not be
326 representative of the whole population of those regions, although in the United States,
327 Twitter users could be representative of the population of the country. In addition,
328 although a thorough search strategy was conducted for each country in its specific
329 language, a considerable volume of tweets could have been missed for the analysis.
330 Both limitations must consider that Twitter API retrieves only the 1% of the tweets

331 posted. Finally, misinformation dissemination was not analyzed, therefore, the impact
332 of the information disorder items on specific populations was not measured. In this
333 regard, we suggest the inclusion of this analysis in future studies.

334 In conclusion, the general sentiment of Twitter users from three different countries
335 regarding dental caries was negative. Also, misinformation items related to the disease
336 were identified on Twitter posts. Those items were related to the consumption of sugary
337 foods and the direct and inevitable development of carious lesions. Also,
338 misinformation items were associated to a general positive sentiment. The
339 identification of social media users concerned and interested in dental caries could be
340 a relevant approach for the development of specific prevention and management
341 strategies of the disease, focusing on the discussion of shared decision making for
342 treatment counseling. Educating patients in the etiology and development of NCDs
343 such as dental caries is imperative for the reduction of the burden of the disease in
344 specific populations.

References

1. Wang X, Shi J, Kong H. Online health information seeking: A review and meta-analysis. *Health Commun.* 2020:1-13. doi: 10.1080/10410236.2020.1748829.
 2. Bhatnagar S, Choubey N. Making sense of tweets using sentiment analysis on closely related topics. *Social Network Analysis and Mining.* 2021;11(1):1-11.
 3. Zhang X, Foo S, Majid S, Chang Y-K, Dumauual HTJ, Suri VR. Self-care and health-information-seeking behaviours of diabetic patients in Singapore. *Health Communication.* 2020;35(8):994-1003.
 4. Organization WH. WHO guideline on self-care interventions for health and well-being: World Health Organization; 2021.
 5. Alhusseini N, Banta JE, Oh J, Montgomery SB. Social Media Use for Health Purposes by Chronic Disease Patients in the United States. *Saudi J Med Med Sci.* 2021;9(1):51.
 6. Chen J, Wang Y. Social media use for health purposes: systematic review. *Journal of Medical Internet Research.* 2021;23(5):e17917.
 7. Meppelink CS, Smit EG, Franssen ML, Diviani N. "I was right about vaccination": Confirmation bias and health literacy in online health information seeking. *Journal of health communication.* 2019;24(2):129-40.
 8. Yarritu I, Matute H. Previous knowledge can induce an illusion of causality through actively biasing behavior. *Frontiers in Psychology.* 2015:389.
 9. Masaki S, Yusuke Y. Characterizing the Influence of Confirmation Bias on Web Search Behavior. *Frontiers.* 2021;12.
 10. Matute H, Blanco F, Yarritu I, Díaz-Lago M, Vadillo MA, Barberia I. Illusions of causality: how they bias our everyday thinking and how they could be reduced. *Frontiers in psychology.* 2015;6:888.
 11. Zhang X, Xu X, Cheng J, editors. *WeChatting for health: what motivates older adult engagement with health information.* Healthcare; 2021: Multidisciplinary Digital Publishing Institute.
 12. Bryan MA, Evans Y, Morishita C, Midamba N, Moreno M. Parental perceptions of the internet and social media as a source of pediatric health information. *Acad Pediatr.* 2020;20(1):31-8.
 13. Eysenbach G. How to fight an infodemic: the four pillars of infodemic management. *Journal of medical Internet research.* 2020;22(6):e21820.
 14. Statista Research Department. Most popular social networks worldwide as of January 2022, ranked by number of monthly active users: Statista; 2022 [cited 2022 March]. Available from: <https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/>.
 15. Statista Research Department. Distribution of Twitter users worldwide as of April 2021, by age group: Statista; 2021 [cited 2022 March]. Available from: <https://www.statista.com/statistics/283119/age-distribution-of-global-twitter-users/>.
 16. Statista Research Department. Distribution of Twitter users worldwide as of January 2022, by gender: Statista; 2022 [cited 2022 March]. Available from: <https://www.statista.com/statistics/828092/distribution-of-users-on-twitter-worldwide-gender/>.
 17. Statista Research Department. Twitter - Statistics & Facts: Statista; 2022 [cited 2022 March]. Available from: https://www.statista.com/topics/737/twitter/#topicHeader_wrapper.
-

-
18. Sinnenberg L, Buttenheim AM, Padrez K, Mancheno C, Ungar L, Merchant RM. Twitter as a tool for health research: a systematic review. *American journal of public health*. 2017;107(1):e1-e8.
 19. Eibensteiner F, Ritschl V, Nawaz FA, Fazel SS, Tsagkaris C, Kulnik ST, et al. People's willingness to vaccinate against COVID-19 despite their safety concerns: Twitter poll analysis. *Journal of medical Internet research*. 2021;23(4):e28973.
 20. Rosenberg H, Syed S, Rezaie S. The Twitter pandemic: The critical role of Twitter in the dissemination of medical information and misinformation during the COVID-19 pandemic. *Canadian journal of emergency medicine*. 2020;22(4):418-21.
 21. Al-Khalifa KS, Bakhurji E, Halawany HS, Alabdurubalnabi EM, Nasser WW, Shetty AC, et al. Pattern of dental needs and advice on Twitter during the COVID-19 pandemic in Saudi Arabia. *BMC oral health*. 2021;21(1):1-7.
 22. World Dental Federation. The Challenge of Oral Disease: A call for global action 2015 [cited 2019 May]. Available from: [www.fdiworldental.org/publications/oral-health-atlas/oral-health-atlas-\(2015\).aspx](http://www.fdiworldental.org/publications/oral-health-atlas/oral-health-atlas-(2015).aspx).
 23. GBD Oral Disorders Collaborators, Bernabe E, Marcenes W, Hernandez C, Bailey J, Abreu L, et al. Global, regional, and national levels and trends in burden of oral conditions from 1990 to 2017: a systematic analysis for the global burden of disease 2017 study. *J Dent Res*. 2020;99(4):362-73. doi: 10.1177/0022034520908533.
 24. Rzewuska M, Lamont TJ, Banister K, Gillies K, Goulao B, Locock L, et al. Twitter communication of the UK public on dental health and care during a COVID lockdown: "My kingdom for a dentist". *Community Dental Health*. 2021.
 25. Salim NA, Jubair F, Hassona YM, Izriqi S, Al-Fuqaha'a D. Esthetic Dentistry on Twitter: Benefits and Dangers. *International Journal of Dentistry*. 2021;2021.
 26. Grond SE, Kallies G, McCormick ME. Parental and provider perspectives on social media about ankyloglossia. *International Journal of Pediatric Otorhinolaryngology*. 2021;146:110741.
 27. Al-Khalifa KS, AlSheikh R, Alsaahafi YA, Alkhalifa A, Sadaf S, Al-Moumen SA, et al. Dental care in the Arab countries during the COVID-19 pandemic: an infodemiological study. *Risk Management and Healthcare Policy*. 2021;14:2153.
 28. Eliacik BK. Topical Fluoride Applications Related Posts Analysis on Twitter Using Natural Language Processing. *Oral Health & Preventive Dentistry*. 2021;19(1):457-64.
 29. Aguirre PE, Coelho M, Oliveira T, Rios D, Cruvinel AF, Cruvinel T. What can Google inform us about people's interests regarding dental caries in different populations? *Caries Res*. 2018;52(3):177-88.
 30. Aguirre PEA, Coelho MM, Rios D, Machado MAAM, Cruvinel AFP, Cruvinel T. Evaluating the dental caries-related information on Brazilian websites: qualitative study. *J Med Internet Res*. 2017;19(12):e415.
 31. Lotto M, Aguirre PEA, Neto NL, Cruvinel AF, Cruvinel T. Is the Quality of Toothache-Related Information Published in Brazilian Websites Adequate to Assist People in Seeking Dental Treatment? *Oral Health Prev Dent*. 2020;18(2):301-9.
 32. Ayala Aguirre PE, Aníbal I, Strieder A, Lotto M, AF PC, Rios D, et al. Online quality and readability assessment of Early childhood caries information available on websites from distinct countries: A cross-sectional study. *European Journal of Paediatric Dentistry*. 2022;23(1):15-20.
 33. Statista. Leading countries based on number of Twitter users as of January 2022 (in millions) 2022 [cited 2022 March]. Available from:
-

<https://www.statista.com/statistics/273172/twitter-accounts-with-the-most-followers-worldwide/>.

34. Postamn. Getting started with Postman 2020 [cited 2021 September]. Available from: <https://developer.twitter.com/en/docs/tutorials/postman-getting-started>.
 35. Nidhal Baccouri. Deep-Translator: MIT license; 2020 [cited 2021 September]. Available from: <https://deep-translator.readthedocs.io/en/latest/README.html#deep-translator>.
 36. Liddy ED. Natural Language Processing. In: Marcel Decker I, editor. In Encyclopedia of Library and Information Science. 2nd Edition ed. NY2001.
 37. Hutto C, Gilbert E, editors. Vader: A parsimonious rule-based model for sentiment analysis of social media text. Proceedings of the International AAAI Conference on Web and Social Media; 2014.
 38. Loria S. TextBlob: Simplified Text Processing. 2020 [cited 2021 September]. Available from: <https://textblob.readthedocs.io/en/dev/>.
 39. Wang Y, McKee M, Torbica A, Stuckler D. Systematic literature review on the spread of health-related misinformation on social media. Soc Sci Med. 2019;240:112552. doi: 10.1016/j.socscimed.2019.112552.
 40. Bernabé E, Marcenes W. Can minimal intervention dentistry help in tackling the global burden of untreated dental caries? British Dental Journal. 2020;229(7):487-91. doi: 10.1038/s41415-020-2155-9.
 41. Twetman S. Prevention of dental caries as a non-communicable disease. European journal of oral sciences. 2018;126:19-25.
 42. Pitts NB, Twetman S, Fisher J, Marsh PD. Understanding dental caries as a non-communicable disease. British dental journal. 2021;231(12):749-53.
 43. Yarritu I, Matute H, Luque D. The dark side of cognitive illusions: When an illusory belief interferes with the acquisition of evidence-based knowledge. British Journal of Psychology. 2015;106(4):597-608.
 44. Aguirre PEA, Aníbal I, Strieder AP, Lotto M, Rizzato VL, Cruvinel AFP, et al. (in press). Online quality and readability assessment of Early childhood caries information available on Websites from distinct countries: A cross-sectional study Eur J Paediatr Dent. 2021.
 45. Frantsve-Hawley J, Rindal DB. Translational research: bringing science to the provider through guideline implementation. Dental Clinics. 2019;63(1):129-44.
 46. Green LW, Ottoson JM, Garcia C, Hiatt RA. Diffusion theory and knowledge dissemination, utilization, and integration in public health. Annual review of public health. 2009;30:151-74.
 47. Ji X, Chun S, Wei Z, Geller J. Twitter sentiment classification for measuring public health concerns. Social Network Analysis and Mining. 2015;5(1):1-25.
-

Table 1. Search strategies for Twitter API according to each language

Language	Search strategy
English	("dental cavity" OR "cavities" OR "tooth decay" OR "caries" OR "dental caries" OR #dentalcaries OR #caries OR #cavities OR #cavity OR #toothdecay) lang:en place_country:US
Portuguese	("cárie dentária" OR "cárie dentaria" OR "carie dentária" OR "carie" OR "cárie" OR "dente podre" OR #cariedentaria OR #cáriedentária OR #cárie OR #cáriedentaria OR #cariedentária OR #dentepodre OR #carie) lang:pt place_country:BR
Spanish	("caries dental" OR "caries" OR "carie" OR "caries dentales" OR "dientes picados" OR "muelas picadas" OR "dientes cariados" OR #cariesdental OR #caries OR #carie OR #dientecariado OR #dientepicado OR #muelascariadas OR #muelapicada) lang:es place_country:MX

Table 2. Distribution of different variables according to different countries, Distinct letters represent significant statistical differences of distributions between countries (Chi-square test)

		Languages n(%)				
		Overall	English	Portuguese	Spanish	<i>p</i>
Sentiment	Negative	1238 (46.40%)	340 (41.50%) ^a	580 (62.80%) ^b	318 (34.30%) ^c	<0.001
	Neutral	696 (26.20%)	142 (17.30%) ^a	161 (17.40%) ^a	393 (42.40%) ^b	
	Positive	735 (27.50%)	337 (41.10%) ^a	182 (19.70%) ^b	216 (23.30%) ^b	
Misinformation	Yes	491 (18.40%)	196 (23.90%) ^a	132 (14.30%) ^b	163 (17.60%) ^b	<0.001
	No	2178 (81.60%)	623 (76.10%) ^a	791 (85.70%) ^b	764 (82.40%) ^b	
Interaction	Yes	1036 (38.80%)	491 (60.00%) ^a	604 (65.40%) ^b	538 (58.00%) ^a	0.003
	No	1633 (61.20%)	328 (40.00%) ^a	319 (34.60%) ^b	389 (42.00%) ^a	
Sex	Male	1159 (43.40%)	358 (30.90%) ^{a, b}	375 (32.40%) ^b	426 (36.80%) ^a	<0.001
	Female	1163 (43.60%)	334 (28.70%) ^a	472 (40.60%) ^b	357 (30.70%) ^a	
	Unknown	347 (13%)	127 (36.60%) ^a	76 (21.90%) ^b	144 (41.50%) ^a	
Account	Health professionals	162 (6.10%)	31 (19.10%) ^a	22 (13.60%) ^a	109 (67.30%) ^b	<0.001
	Non-Health professionals	2507 (93.90%)	788 (31.40%) ^a	901 (35.90%) ^a	818 (32.60%) ^b	

Table 3. Descriptive statistics of interaction metrics according to different countries. Distinct letters represent significant statistical differences between countries (Kruskal-Wallis and Dunn's tests)

		ReTweet	Reply	Likes	Followers	Interaction
English	Mean (SD)	0.16 (0.05) ^a	0.58 (0.06) ^a	3.69 (0.66) ^a	1,839.29 (232.03) ^a	4.43 (0.73) ^a
	Median (IR)	0 (0)	0 (1)	1 (3)	510 (1052)	1 (3)
	Minimum	0	0	0	1	0
	Maximum	37	22	481	125,613	522
Portuguese	Mean (SD)	0.14 (0.03) ^a	0.57 (0.04) ^b	2.78 (0.76) ^{a,b}	1,545.28 (220.27) ^a	3.49 (0,81) ^a
	Median (IR)	0 (0)	0 (1)	1 (1)	483 (856)	1 (2)
	Minimum	0	0	0	1	0
	Maximum	21	20	604	119,831	645
Spanish	Mean (SD)	0.67 (0.17) ^b	0.29 (0.02) ^c	2.54 (0.53) ^b	2,664.70 (454.63) ^b	3.50 (0.70) ^a
	Median (IQ)	0 (0)	0 (0)	0 (2)	390.50 (894)	1 (2)
	Minimum	0	0	0	1	0
	Maximum	114	8	347	289,605	445

Table 4. Distribution of human- and AI-based sentiment analyses according to different countries. Accuracies, sensitivities, and specificities of VADER and TextBlob are demonstrated in relation to human-based analysis. Distinct letters represent significant statistical differences of distributions between countries (Chi-square test)

		Human	VADER	TextBlob
Overall	Negative	1,238 (46.40%)	1,008 (37.80%)	548 (20.50%)
	Neutral	696 (26.10%)	519 (19.40%)	970 (36.30%)
	Positive	735 (27.50%)	1142 (42.80%)	1,151 (43.10%)
	Accuracy	-	0.597	0.565
	Sensitivity	-	0.712	0.855
	Specificity	-	0.481	0.275
USA	Negative	340 (41.50%) ^a	304 (37.10%) ^a	149 (18.20%) ^a
	Neutral	142 (17.30%) ^a	136 (16.60%) ^a	272 (33.20%) ^a
	Positive	337 (41.10%) ^a	379 (46.30%) ^a	398 (48.60%) ^a
	Accuracy	-	0.585	0.588
	Sensitivity	-	0.699	0.891
	Specificity	-	0.471	0.285
Brazil	Negative	580 (62.80%) ^b	394 (42.70%) ^b	229 (24.80%) ^b
	Neutral	161 (17.40%) ^a	183 (19.80%) ^{a,b}	328 (35.50%) ^{a,b}
	Positive	182 (19.70%) ^b	346 (37.50%) ^b	366 (39.70%) ^b
	Accuracy	-	0.626	0.551
	Sensitivity	-	0.732	0.816
	Specificity	-	0.521	0.286
Mexico	Negative	318 (34.30%) ^c	310 (33.40%) ^a	170 (18.30%) ^a
	Neutral	393 (42.40%) ^b	200 (21.60%) ^b	370 (39.90%) ^b
	Positive	216 (23.30%) ^b	417 (45%) ^a	387 (41.70%) ^b
	Accuracy	-	0.566	0.547
	Sensitivity	-	0.711	0.849
	Specificity	-	0.421	0.245

Table 5. Distribution of sentiment, misinformation, interaction, languages, sex, and accounts in relation to dichotomized themes. Distinct letters represent significant statistical differences within each variable (Chi-square test)

		Themes		
		Preventive	Others	Total
Sentiment	Negative	1027 (83.0%) ^a	211 (17.0%) ^b	1238 (100%)
	Neutral	271 (38.9%) ^a	425 (61.1%) ^b	696 (100%)
	Positive	304 (41.4%) ^a	431 (58.6%) ^b	735 (100%)
Misinformation	Yes	121 (24.6%) ^a	370 (75.4%) ^b	491 (100%)
	No	1481 (68.0%) ^a	697 (32.0%) ^b	2178 (100%)
Interaction	Yes	992 (60.7%) ^a	641 (39.3%) ^a	1633 (100%)
	No	610 (58.9%) ^a	426 (41.1%) ^a	1036 (100%)
Languages	English	408 (49.8%) ^a	411 (50.2%) ^a	819 (100%)
	Portuguese	645 (69.9%) ^b	278 (30.1%) ^b	923 (100%)
	Spanish	549 (59.2%) ^c	378 (40.8%) ^c	927 (100%)
Sex	Female	898 (59.6%) ^a	608 (40.4%) ^a	1506 (100%)
	Male/unknown	704 (60.5%) ^a	459 (39.5%) ^a	1163 (100%)
Account	Health professionals	95 (58.6%) ^a	67 (41.4%) ^a	162 (100%)
	Non-health professionals	1507 (60.1%) ^a	1000 (39.9%) ^a	2507 (100%)

Table 6 Multiple logistic regression model showing the association between misinformation with different factors

Factors	Misinformation				
	B [†]	S.E. ‡	Wald	p	OR [§] (95% CI)
Theme (Others)	2.99	0.16	346.08	<0.001*	19.90 (14.52-27.27)
Language (English)	0.34	0.12	8.58	0.003*	1.41 (1.12-1.78)
Sentiment (Positive)	1.99	0.15	168.08	<0.001*	7.35 (5.44-9.93)
Constant	-1.13	0.16	50.29	<0.001*	

†Unstandardized coefficient

‡Standard error

§Odds ratio

*indicates significant differences

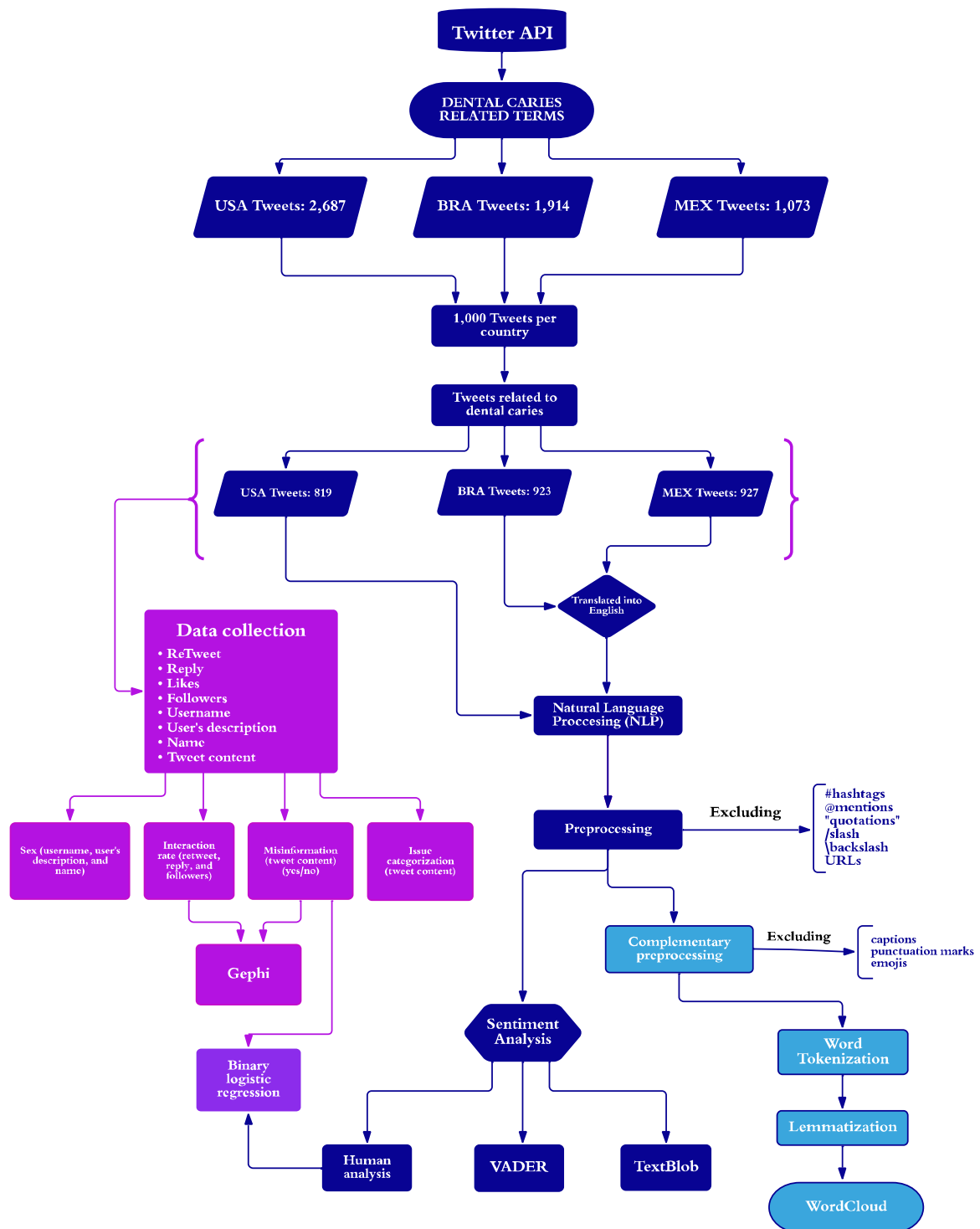


Figure 1. Flowchart showing the synthesis of the methodological approach.

Supplemental file. Description of the categories of themes of dental caries-related tweets in three different countries

Themes	Description	Example
1	About no caries experience	No cavities 🙄 I've been eating more sweets than usual lately so I was a tad paranoid lol
2	About toothache or dental painful symptoms	Tomorrow I have a dentist appointment...it's been 4 years since my last one... I know I have so many cavities and my impacted wisdom teeth have been causing me excruciating pain the last 4 days...I wish anxiety didn't let it get this far... 😊
3	About dental fear in diverse clinical situations	Went to the dentist for the first time in 4 years and I have 10 cavities. Pray for me y'all 😊 I have a deathly fear of the dentist, don't judge...
4	About restorative procedures	Had a couple cavities filled today. Idk if it was the novocaine shot or what but my sinuses went nuts and I basically got a black eye for like an hour.
5	About prevention facts or statements	Cavities can be prevented, we suggest brushing twice a day for 2 minutes. #NationalChildrensHealthMonth
6	About diagnosis of dental caries	She got 3 cavities 🙄 but I guess that's not really bad considering she's about to be 6 and haven't had any until just now.
7	About political or social facts and comments	Dental insurance is a joke. I pay for dental insurance and I still had to pay \$500 to fill my son's cavities. 😊 and let's not even talk about braces!!!
8	About aesthetics and dental appearance criticism	How can people not take care of they teeth...gross!! Dont kiss me if yo teeth are black full of cavities 🙄🙄🙄🙄🙄
9	About consumption of sugary foods, drinks, and candies	Butterfingers have the added benefit of pulling out existing fillings and giving you new cavities. A+
10	About other dental treatments distinct to restorative ones	JUST PULLED SIX TEETH! 🗄️🗄️🗄️🗄️🗄️🗄️💕 #Dentist #COVID19 #instagram Been waiting a year to take care of these painful teeth and cavities. Finally! 🗄️
11	About analogies or jokes	What's 'dental calculus'? $\lim_{x \rightarrow 0} x = \text{tooth decay?}$ dx/dy (d for dentist)?
12	About dental professional comments	Dentists are about to bank off of quarantine cavities

3 DISCUSSION

3 DISCUSSION

The present study compiled 4 infoveillance studies to determine the Internet and social media users' interest and concerns regarding ECC and dental caries, also regarding the quality of the information on the disease found online. Our findings showed the interests of Google users in ECC-related information is increasing; however, it was maintained low over time, which could indicate a lack of awareness about the negative consequences of the disease in the early stages of life. In addition, those concerned with the disease find low quality of information on Websites and incomplete content on YouTube videos. Interestingly, tweets containing misinformation about the disease was charged with positive sentiment, frequently associated with the sugar consumption and the development of the disease.

With the advent of the *information era*, a novel field of research called Infodemiology emerged to study “*the distribution and determinants of information in an electronic medium, specifically the Internet, or in a population, with the ultimate aim to inform public health and public policy*” [27]. Infoveillance uses the same methods as infodemiology but for surveillance purposes [27]. Since its intention is not to substitute traditional epidemiological surveys, its interest is directed towards elucidating “what and when, why, and who” seeks health information on the Internet.

What and When: types of research

Much of the information obtained on the Internet comes from the results of previous search performed in online search tools, which are stored in a database freely available for analysis. Google Trends is an online tool that presents the weekly and monthly variation of Relative Search Volume (RSV) values for searches performed on Google during a specific period of time. The platform also makes available a list of popular terms used to search the information of interest, allowing a quick, affordable and accurate analysis of data related to the consumption behavior of health information by netizens [28]. For instance, our results on ECC depicted an interest linked to the definition of the disease (tooth, infant, child, dentistry, and deciduous teeth), risk factors (breastfeeding, milk, baby bottle, and sleep), and preventive care (preventive healthcare, fluoride, and brushing). The activity of Internet users seeking dental

information [29] motivated the production of evidence of very poor quality and easy readability of Web content items related to the disease and its consequence [30]. In addition, Websites commonly recommended the avoidance of the consumption of “toxic fluoride toothpastes” [30]. These patterns were observed in three countries.

Why: health information needs

Theories of health information seeking behavior (HISB) developed conceptual models to identify factors that motivate the behavior [31] such as contextual factors as health status, primary health care, environmental information factors, and family members who have an illness or are at risk of a disease [32, 33]. On the other hand, personal factors are associated to demographics, socioeconomics, health history, genetics, education, culture, language, attitudes, intentions, behavior, and others [34]. The first encounter with the information could also determine the interest in the information itself, resulting in active or passive consumers of the information. Active consumers, according to their specific needs (contextual or personal), search personally for information of interest. Passive consumers, on the contrary, could have an incidental encounter with information, either through mass media (television, radio, magazines) or through Internet advertisements, posts, or even during the consultation with their health physician, who could provide the information without being requested. Other factors can also influence searches for health information, for example, machine-human relationship (the person's familiarity with browsing the internet or ease of using electronic devices), beliefs or convictions, organizational factors (norms, social aspects), and health literacy [35]. Interestingly, most people who search for health information on the Internet are users of a social network [36], usually recognizing their own influence by the information consumed [37]. Therefore, identifying why people prefer to seek health information on social media can assist in our planning for health promotion and disease prevention [37].

Who: e-patient.

Big data enables to study people's interests related to their queries placed on search engines; however, the identification of the person who performed the query is retained by Google Trends, respecting the privacy rights of its users. Nevertheless, surveys such as the National Health Information Trends Survey (HINTS) are able to

identify the health information consumer profile [38]. Initially, e-patients tended to be young adult women, with a higher level of education, with previous Internet browsing experience, with a high income and Internet at home. Also, people who have a specific disease [32] would be more likely to seek for information on the Internet [34, 39], although, a study negatively related health behaviors with active search on the Internet [40]. Due to the democratization of the internet, its access has been dissipating geographical barriers allowing it to different social status, which could indicate a progressive modification in the e-patient profile. For instance, 73% of the Brazilian population has access to the Internet [41], among which 47% access the Internet through their smartphones. A singular feature of e-patients is related to the time spent browsing the Internet, which makes them to develop a skill or familiarity with the search for different topics [34]. Interestingly, individuals who frequently access information seem to have better health outcomes in relation to their conditions [42-45].

Oral health professional role: Dentist v2.0

Dental caries in deciduous teeth is the 10th most prevalent disease in the world [3], affecting the quality of life of children and their parents. Although our results did not find an association between the burden of the disease and the volume of searches performed on a specific search tool, the HISB theories support that the presence of the disease in oneself or in a loved one could stimulate the needs for health information. Indeed, parents do search for information related to their children resorting even to SM to respond their parental doubts.

When information available is misleading, incorrect, or incomplete it could hamper the engagement with prevention techniques, adherence to treatments and the relationship with the dentist. Unfortunately, the dissemination of misleading information is not a new phenomenon. The continuous advances in technology and the use of digital platforms as their main source of information could result in an aggravation of this phenomenon [46]. Why would this situation represent a threat to people's health and to the public health system? Misleading information spreads easily, people review contents without verifying their veracity, under the influence of *confirmation bias*, where the user's own beliefs unconsciously trick the user into trust what he reads, especially if the information has been read multiple times. For instance, we found a positive association between misinformation items and a positive sentiment of the tweets. In

addition, *information overload* discourages people from searching for trustful sources, relying on the first information found. The proliferation of opinions that influence the lifestyle of users and, especially, the influence of celebrities on their followers [47], for example, the use of carbon toothpastes or non-fluoride toothpastes.

How to deal with this situation? We face ourselves with a challenging infodemic [48], and preventing the posting or circulation of misinformation items could be against people's right to freely post their ideas or believes; however, the focus should not be on social media alone, although there are non-profit institutions charged with disproving misinformation, but in the person seeking care and in the oral health professionals. Communication with our patient is essential and stimulating people to seek information is very important, especially when people who are not very motivated to seek information on health at some point will have found information on health (incidental encounter with information). Therefore, we must be prepared to have self-confidence to know how to discern between true or false information, as well as having the security and confidence that the information obtained is still doubtful and can be taken in consultation to be discussed with a health professional. As long as our role in the battle must also be focused on disseminating true information, based on scientific evidence, following the guidelines of quality assessment instruments. The production of good quality information is a complex and arduous task, which involves the need of skilled specialists, trained to attain target audiences, and produce accessible contents according to specific guidelines and instruments. The poor quality of information authored by health professionals might indicate their technical incapacity combined with a primary interest in producing oral health information for advertisements and commercial profit, linked to a possible marketing sense of obligation in developing a website for describing treatments to their patients [49].

In conclusion, the consumption of misleading information could deteriorate the person-dentist relationship; hence, the professional should be prepared to face this challenge effectively, advising and encouraging their patients to explore information on recommended websites, warning people about the risks of health home practices, and contributing with the production of good quality materials. These findings indicate the need of the development of special policies focused on the production and publication of web health information, encouraging dentists to guide their patients in searching specific oral health websites.

4 CONCLUSIONS

4 CONCLUSIONS

- The interest of Google users in ECC-related information is increasing slowly in all seven countries over the years, and it is not being associated with the burden of untreated dental caries in children under 5 years old and the Internet penetration in most countries, except for France and Mexico.
 - The information related to ECC published in websites of Brazil, Mexico, and USA were considered of poor quality, presenting accessible and simple contents for most population.
 - Youtube videos that presented information about etiology and prevention of ECC in English, Spanish and Brazilian Portuguese invariably focused on partial aspects of the disease.
 - The general sentiment of Twitter users from three different countries regarding dental caries was negative. Also, misinformation items related to the disease identified on Twitter posts were associated to a general positive sentiment and to the consumption of sugary foods and the direct and inevitable development of carious lesions.
 - Specific policies should be developed toward the dissemination of oral health information and the prevention of the disease.
 - Dental professionals must be aware of their patient's needs, producing good quality digital materials and promoting the informational counseling during clinical appointments.
 - The awareness on what parents and caregivers are exposed to when searching for ECC online could be a relevant approach to improve the management of the disease, considering that the information influences the parental decision-making process decisively.
 - Pediatric dentists should receive sufficient technical formation and stimulation to produce Web contents directly to parents and children,
-
-

focused on the spread of information related to good oral health habits for the prevention and control of ECC.

- To combat misinformation and establish a trustful digital environment for population, dental teams are the main responsible for counseling people to obtain and consume good quality information towards the improvement of their own health literacy skills.
 - Educating patients in the etiology and development of NCDs such as dental caries is imperative for the reduction of the burden of the disease in specific populations
-
-

REFERENCES

REFERENCES

1. Berkowitz RJ. Causes, treatment and prevention of early childhood caries: a microbiologic perspective. *Journal-Canadian Dental Association*. 2003;69(5):304-7.
 2. Wong PD, Birken CS, Parkin PC, Venu I, Chen Y, Schroth RJ, et al. Total breast-feeding duration and dental caries in healthy urban children. *Academic pediatrics*. 2017;17(3):310-5.
 3. GBD Oral Disorders Collaborators, Bernabe E, Marcenes W, Hernandez C, Bailey J, Abreu L, et al. Global, regional, and national levels and trends in burden of oral conditions from 1990 to 2017: a systematic analysis for the global burden of disease 2017 study. *J Dent Res*. 2020;99(4):362-73.
 4. Uribe SE, Innes N, Maldupa I. The Global Prevalence Of Early Childhood Caries: A Systematic Review with Meta-analysis Using the WHO Diagnostic Criteria. *Int J Paediatr Dent*. 2021.
 5. Pitts NB, Baez RJ, Diaz-Guillory C, Donly KJ, Alberto FC, McGrath C, et al. Early Childhood Caries: IAPD Bangkok Declaration. *J Dent Child (Chic)*. 2019;86(2):72.
 6. Jordan AR, Becker N, Jöhren H-P, Zimmer S. Early Childhood Caries and Caries Experience in Permanent Dentition: A 15-year Cohort Study. *Swiss dental journal*. 2016;126(2):114-9.
 7. Moynihan P, Tanner L, Holmes R, Hillier-Brown F, Mashayekhi A, Kelly S, et al. Systematic review of evidence pertaining to factors that modify risk of early childhood caries. *JDR Clin Trans Res*. 2019;4(3):202-16.
 8. Colvara BC, Faustino-Silva DD, Meyer E, Hugo FN, Celeste RK, Hilgert JB. Motivational interviewing for preventing early childhood caries: A systematic review and meta-analysis. *Community Dent Oral Epidemiol*. 2021;49(1):10-6.
 9. Edelstein BL, Ng MW. Chronic disease management strategies of early childhood caries: support from the medical and dental literature. *Pediatric dentistry*. 2015;37(3):281-7.
 10. Organization WH. Ending childhood dental caries: WHO implementation manual. 2019.
-
-

11. IAPD. Foundational Articles and Consensus Recommendations: Management of Early Childhood Caries 2020.
 12. Fox S, Duggan M. Health online 2013. *Health*. 2013;2013:1-55.
 13. Statista. Number of internet users worldwide from 2005 to 2021(in millions) 2021 [Available from: <https://www.statista.com/statistics/273018/number-of-internet-users-worldwide/>].
 14. Zhao J, Gao S, Wang J, Liu X, Hao Y. Differentiation between two healthcare concepts: person-centered and patient-centered care. *J Nurs*. 2016;2352:0132.
 15. Hootsuite - Simon Kemp. Digital data 2021: The United States of America. 2021.
 16. Statista. Internet usage in Brazil - statistics and facts. 2021.
 17. Kubb C, Foran HM. Online Health Information Seeking by Parents for Their Children: Systematic Review and Agenda for Further Research. *J Med Internet Res*. 2020;22(8):e19985.
 18. Cruvinel T, Ayala Aguirre PE, Lotto M, Marchini Oliveira T, Rios D, Pereira Cruvinel AF. Digital behavior surveillance: monitoring dental caries and toothache interests of Google users from developing countries. *Oral diseases*. 2019;25(1):339-47.
 19. Miller NE. Behavioral medicine: Symbiosis between laboratory and clinic. *Annual Review of Psychology*. 1983;34(1):1-32.
 20. Ayers JW, Althouse BM, Dredze M. Could behavioral medicine lead the web data revolution? *Jama*. 2014;311(14):1399-400.
 21. Ríordáin RN, McCreary C. Dental patients' use of the Internet. *British dental journal*. 2009;207(12):583.
 22. Aguirre PE, Coelho M, Oliveira T, Rios D, Cruvinel AF, Cruvinel T. What can Google inform us about people's interests regarding dental caries in different populations? *Caries research*. 2018;52(3):177-88.
 23. Aguirre PEA, Coelho MM, Rios D, Machado MAAM, Cruvinel AFP, Cruvinel T. Evaluating the dental caries-related information on Brazilian websites: qualitative study. *J Med Internet Res*. 2017;19(12):e415.
-

24. Mascarenhas AK. Who needs more than 1,000 ppm? The epidemiology of high-risk populations. *Caries Research*. 2016;50(Suppl. 1):1-8.
 25. Heavilin N, Gerbert B, Page J, Gibbs J. Public health surveillance of dental pain via Twitter. *Journal of dental research*. 2011;90(9):1047-51.
 26. AAPD: American Academy of Pediatric Dentistry. American Academy of Pediatric Dentistry Research Agenda. 2017;39:10.
 27. Eysenbach G. Infodemiology and infoveillance: tracking online health information and cyberbehavior for public health. *American journal of preventive medicine*. 2011;40(5):S154-S8.
 28. Carr LJ, Dunsiger SI. Search query data to monitor interest in behavior change: application for public health. 2012.
 29. Aguirre PEA, Lotto M, Strieder AP, Cruvinel T. Digital surveillance: Monitoring the activity of Internet users searching for information related to early childhood caries. *Health Informatics Journal*. 2022;28(1):14604582211073057.
 30. Ayala Aguirre PE, Aníbal I, Strieder A, Lotto M, AF PC, Rios D, et al. Online quality and readability assessment of Early childhood caries information available on websites from distinct countries: A cross-sectional study. *European Journal of Paediatric Dentistry*. 2022;23(1):15-20.
 31. Wang X, Shi J, Kong H. Online health information seeking: A review and meta-analysis. *Health Commun*. 2020:1-13.
 32. Longo DR, Schubert SL, Wright BA, LeMaster J, Williams CD, Clore JN. Health information seeking, receipt, and use in diabetes self-management. *The Annals of Family Medicine*. 2010;8(4):334-40.
 33. Longo DR. Understanding health information, communication, and information seeking of patients and consumers: a comprehensive and integrated model. *Health Expect*. 2005;8(3):189.
 34. Cantor D, Coa K, Crystal-Mansour S, Davis T, Dipko S, Sigman R. Health information national trends survey (HINTS) 2007. Rockville, MD: Westat. 2009.
 35. Bernhardt JM, Felter EM. Online pediatric information seeking among mothers of young children: results from a qualitative study using focus groups. *J Med Internet Res*. 2004;6(1):e7.
-

36. Ventola CL. Social media and health care professionals: benefits, risks, and best practices. *Pharmacy and therapeutics*. 2014;39(7):491.
 37. Park DY, Goering EM. The health-related uses and gratifications of YouTube: Motive, cognitive involvement, online activity, and sense of empowerment. *Journal of Consumer Health on the Internet*. 2016;20(1-2):52-70.
 38. Chung JE. Patient–provider discussion of online health information: results from the 2007 Health Information National Trends Survey (HINTS). *J Health Commun*. 2013;18(6):627-48.
 39. Zhang L, Qin Y, Li P. Media Complementarity and Health Information Acquisition: A Cross-sectional Analysis of the 2017 HINTS-China Survey. *Journal of Health Communication*. 2020:1-10.
 40. Jaafar NI, Ainin S, Yeong MW. Why bother about health? A study on the factors that influence health information seeking behaviour among Malaysian healthcare consumers. *International Journal of Medical Informatics*. 2017;104:38-44.
 41. Statista Digital Market Outlook. Internet user penetration in Brazil from 2017 to 2023 2019 [Available from: <https://www.statista.com/statistics/292757/brazil-internet-user-penetration/>].
 42. Lee K, Hoti K, Hughes JD, Emmerton LM. Consumer use of “Dr Google”: a survey on health information-seeking behaviors and navigational needs. *Journal of medical Internet research*. 2015;17(12).
 43. McCloud RF, Okechukwu CA, Sorensen G, Viswanath K. Beyond access: barriers to internet health information seeking among the urban poor. *Journal of the American Medical Informatics Association*. 2016:ocv204.
 44. Pluye P, Grad R, Repchinsky C, Jovaisas B, Johnson-Lafleur J, Carrier ME, et al. Four levels of outcomes of information-seeking: A mixed methods study in primary health care. *Journal of the American Society for Information Science and Technology*. 2013;64(1):108-25.
 45. Barry MJ, Edgman-Levitan S. Shared decision making—the pinnacle of patient-centered care. *New England Journal of Medicine*. 2012;366(9):780-1.
 46. Fernando Esteves & Gustavo Sampaio. *Viral: A Epidemia de Fake News e a Guerra da Desinformação*. Portugal: Desassossego; 2020.
-
-

47. da Silva MAD, Walmsley AD. Fake news and dental education. *Br Dent J.* 2019;226(6):397-9.
 48. Eysenbach G. How to fight an infodemic: the four pillars of infodemic management. *Journal of medical Internet research.* 2020;22(6):e21820.
 49. Mourad M, Brendebach J, Alkilzy M, Splieth C. Choosing a paediatric dentist: Factors influencing parents' decision. *Eur J Paediatr Dent.* 2020;21(1):74-9.
-
-