

Social Media Platforms Listening Study on Atopic Dermatitis: Quantitative and Qualitative Findings.

Paméla Voillot, Brigitte Riche, Michel Portafax, Pierre Foulquié, Anaïs Gedik, Sébastien Barbarot, Laurent Misery, Stéphane Heas, Adel Mebarki, Nathalie Texier, Stéphane Schück

Submitted to: Journal of Medical Internet Research
on: June 11, 2021

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressly prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript.....	5
Supplementary Files.....	24
Figures	25
Figure 1.....	26
Figure 2.....	27
Figure 3.....	28
Figure 4.....	29
Figure 5.....	30
Figure 6.....	31
Figure 7.....	32
Figure 8.....	33
Figure 9.....	34
Figure 10.....	35
Multimedia Appendixes	36
Multimedia Appendix 1.....	37
Multimedia Appendix 2.....	37

Social Media Platforms Listening Study on Atopic Dermatitis: Quantitative and Qualitative Findings.

Paméla Voillot¹; Brigitte Riche²; Michel Portafax²; Pierre Foulquié¹; Anaïs Gedik¹; Sébastien Barbarot³; Laurent Misery⁴; Stéphane Heas⁵; Adel Mebarki¹; Nathalie Texier¹; Stéphane Schück¹

¹Kap Code Paris FR

²Sanofi Genzyme Gentilly FR

³CHU de Nantes Nantes FR

⁴CHRU de Brest Brest FR

⁵Université de Rennes 2 Rennes FR

Corresponding Author:

Paméla Voillot
Kap Code
28 rue d'Enghien
Paris
FR

Abstract

Background: Atopic dermatitis (AD) is a chronic, pruritic, inflammatory disease that occurs most frequently in children but also affects many adults. Social media have become key tools for finding and disseminating medical information.

Objective: To identify the main themes of discussion, the difficulties encountered by patients with respect to atopic dermatitis and the impact of the pathology on quality of life (physical, psychological, social, or financial), but also to study the perception of patients regarding their treatment.

Methods: A retrospective study was carried out by collecting social media posts in French language written by internet users mentioning their experience with atopic dermatitis, their quality of life and their treatments. Messages related to AD discomfort posted between July 1, 2010 and October 23, 2020 were extracted from French-speaking publicly available online forums. Automatic and manual extractions were implemented to create a general corpus and two sub-corpus depending on the control of the disease.

Results: A total of 33 115 messages associated with AD were included in the analysis corpus after extraction and cleaning. These messages were posted by 15 857 separate web users, most of them being women younger than 40 years. Tips to manage eczema and everyday hygiene/treatments were among the most discussed topics for Controlled AD sub-corpus, while baby-related topics and therapeutic failure were among the most discussed topics for Insufficiently Controlled AD sub-corpus. Quality of life was discussed in both sub-corpus with a higher proportion in the Controlled AD sub-corpus. Treatments and their perception were also discussed by web users.

Conclusions: More than just emotional or peer support, AD patients turn to online forums to discuss their health. Our findings show the need for an intersection between social media and health care and the importance of developing new approaches such as Atopic Dermatitis Control Tool (ADCT) focused on patients with AD. Clinical Trial: Not applicable (not a trial)

(JMIR Preprints 11/06/2021:31140)

DOI: <https://doi.org/10.2196/preprints.31140>

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ Please make my preprint PDF available to anyone at any time (recommended).

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.

Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Original Paper

Social Media Platforms Listening Study on Atopic Dermatitis: Quantitative and Qualitative Findings.

Abstract

Background: Atopic dermatitis (AD) is a chronic, pruritic, inflammatory disease that occurs most frequently in children but also affects many adults. Social media have become key tools for finding and disseminating medical information.

Objective: To identify the main themes of discussion, the difficulties encountered by patients with respect to AD and the impact of the pathology on quality of life (physical, psychological, social, or financial), but also to study the perception of patients regarding their treatment.

Methods: A retrospective study was carried out by collecting social media posts in French language written by internet users mentioning their experience with atopic dermatitis, their quality of life and their treatments. Messages related to AD discomfort posted between July 1, 2010 and October 23, 2020 were extracted from French-speaking publicly available online forums. Automatic and manual extractions were implemented to create a general corpus and two sub-corpus depending on the level of control of the disease.

Results: A total of 33 115 messages associated with AD were included in the analysis corpus after extraction and cleaning. These messages were posted by 15 857 separate web users, most of them being women younger than 40 years. Tips to manage AD and everyday hygiene/treatments were among the most discussed topics for *Controlled AD* sub-corpus, while baby-related topics and therapeutic failure were among the most discussed topics for *Insufficiently Controlled AD* sub-corpus. Quality of life was discussed in both sub-corpus with a higher proportion in the *Controlled AD* sub-corpus. Treatments and their perception were also discussed by web users.

Conclusion: More than just emotional or peer support, AD patients turn to online forums to discuss their health. Our findings show the need for an intersection between social media and health care and the importance of developing new approaches such as Atopic Dermatitis Control Tool (ADCT) which is a patient-related disease severity assessment tool focused on patients with AD.

Introduction

Atopic dermatitis (AD) is a chronic, pruritic, inflammatory skin disease that occurs most frequently in children but that can also affect adults. The course of the disease is relapsing, and it is frequently associated with elevated levels of serum immunoglobulin E, individual or family history of food allergies, allergic rhinitis, and asthma [1–3]. According to the World Health Organization Global Burden of Diseases initiative's data, AD ranks 15th among all non-fatal disability worldwide and has the highest disease burden among skin diseases as measured by disability-adjusted life-years [4]. Childhood-onset AD begins early in life, with 50% diagnosed in the first year of life and 85% by 5 years of age [1,5]. However, AD can be present at any age, with adult-onset reported by 26% of adult patients with AD [6]. Although AD often resolves during childhood, it persists through adulthood in 20% to 50% of patients [7,8].

AD is associated with substantial morbidity and quality of life impairment. There are several comorbid health problems that occur in patients with AD, aside from the cutaneous signs and symptoms. Chronic pruritus and inflammation can lead to sleep disturbances and mental health symptoms, which are not mutually exclusive. AD may also predispose to higher risk of other atopic disorders, including asthma and allergic rhinitis [9]. Persons with AD appear to be at higher risk for multiple neuropsychiatric disorders, including depression, anxiety, attention-deficit hyperactivity disorder, speech disorders in childhood, headaches, and seizures [10]. There is also a multifactorial association of AD with osteoporosis, bone and joint injuries, infections and fractures [11,12,13,14].

Clinical presentation and severity of AD varies widely, and diagnosis is not always straightforward, especially in adults [15]. Treatment of AD follows a multifaceted, stepwise approach that is tailored according to disease severity [1]. For all patients, basic management and flare prevention consists of good skin care practices (daily showers or baths) followed immediately by the application of emollients and moisturizers, with avoidance of triggers such as irritants, aero- or food allergens, and extremes of heat, cold, or humidity [9]. In mild AD, treatment involves as needed use of low- to mid-potency topical corticosteroids and/or topical calcineurin inhibitors during flares. Patients with frequent flares may benefit from proactive application of topical anti-inflammatory therapies twice a week to the most troublesome areas. Patients with severe disease often present significant treatment challenges. Systemic therapies are usually required for severe AD but have varying degrees of success and can be associated with side-effect profiles that require counselling and close monitoring. Phototherapy has been shown to have success in treating moderate-to-severe AD, but several factors can limit its utility and efficacy including cost and access. New therapies are targeting specific

pathways relevant for AD and many others are in development. An array of topical, oral, and injectable therapies targeting specific disease pathways in AD are in development for paediatric and adult populations [16]. Dupilumab (Dupixent®) is the current first-line systemic agent for adults and children with moderate-to-severe, treatment-resistant AD. It is the only biologic treatment approved from 6 years of age. There are several emerging therapies currently in Phase III clinical trials such as JAK1 and JAK2 inhibitors or other biologic treatments (monoclonal antibodies or NK-1R antagonists) [16,17].

Social media is one of the most rapid and impactful ways of obtaining and delivering information in the modern era. In general, social media refers to forms of electronic communication (such as websites for social networking and microblogging) through which users create online communities to share information, opinions, personal messages, photos, videos and other contents within Internet applications [18]. Social media provides a readily accessible means to promote user-generated content, broaden interpersonal connections, and encourage social collaboration.

Social media has gained unprecedented worldwide popularity over the last two decades. It is estimated that there are currently over 2.3 billion active social media users internationally and this number is growing by approximately 1 million new users every day [19]. Unsurprisingly therefore, the use of social media to find, exchange, and discuss health information is growing at an unprecedented rate. Anyone with access to the internet can post or read information on a (social media) site. This means that it is directly accessible to patients, their family and friends, and all healthcare providers. They provide access to active communities of healthcare professionals and fellow patients, with whom they can share information and their experiences, raise awareness of their concerns, learn about their conditions and healthcare opportunities, and find support [18].

Artificial intelligence (AI) tools are strongly related with data mining and AI is nowadays ranked among the top-10 technology [20]. Despite their limitations, AI tools and techniques that are still in their infancy already provide substantial benefits in providing in-depth knowledge on individuals' health and predicting population health risks, and their use for medicine and public health is likely to increase substantially in the near future [21].

The purpose of this retrospective study was to better understand how patients suffering from AD perceive their quality of life and their treatments, based on the assessment of web-based social media posts, which we considered as a real-life source of health information. In the present study, we aimed: (1) To identify the main themes for discussion. (2) To identify the difficulties encountered by

patients with respect to AD and the impact of the disease on quality of life (physical, psychological, social, or financial). (3) To study the perception of patients regarding their treatment.

Methods

Study Design and Population

This was a noninterventional retrospective study using a text mining approach to retrieve information from social media posts (data available in the public domain) written by French language internet users between July 1, 2010, and October 23, 2020 (Figure 1). The study was conducted in two phases: data collection using the published Detec't webcrawler [22,23] developed by Kap Code (Paris, France) to collect AD-related posts, and quantitative/qualitative analysis to identify trends and characterize key themes discussed by French speaking users.

Figure 1. Study framework

<Insert Figure 1 here>

Data Extraction

A web crawler is an engine that browses through hyperlinks and stores them for future download of associated web pages (identified by the visited hyperlinks) [24]. Scraping of messages was performed according to the HTML structure of each forum. All discussions containing at least one of the keywords or one of their synonyms were automatically retrieved with all the associated metadata, deidentified and cleaned (signature and quote withdrawal) before being stored in a study-specific database. A list of the keywords used for message retrieval is detailed in Multimedia Appendix 1 and a list of secondary filters on non-specific extraction words is detailed in Multimedia Appendix 2. These terms are searched in extracted messages with extraction words not specific to Atopic Dermatitis. If one of these words is found in the post, then the message is kept in the corpus.

The analysis corpus consisted of the corpus cleaned after the removal of messages containing predetermined keywords written in a language other than French, post containing animal-related vocabulary, and messages containing at least one of the study-specific exclusion words listed in Multimedia Appendix 2. The analysis corpus was then divided into two sub-corpus: one sub-corpus for *Controlled AD* and one sub-corpus for *Insufficiently Controlled AD* (Figure 2). The lexical fields of both sub-corpus were realized based on the ADCT questionnaire and completed with real expressions of Internet users [22].

Figure 2. Flowchart presenting the steps for creating analysis corpus and the two sub-corpus

<Insert Figure 2 here>

Data Analysis

Age and Gender: Web users' gender are determined through the identification of regular expressions for each gender. First names and gender-associated suffix and prefix are first searched in the username. Then, content of all available messages is screened for gender specific lexical fields and gender agreements of adjectives and verbs. In the end, a score is computed for each gender and a prediction for the user is obtained by comparing them.

Web users ages are identified by a two-fold method. First, regular expressions of age are identified in the total of the user posts. Then, if no expressions were found, a machine learning model predicts an age based on several features. Among them, the model considers syntactic aspects of posts as well as expressed feeling and the source on which the user expresses himself.

Topic Model: A topic model was applied to identify the themes addressed in the messages. Topic models consist of text mining approach aiming to automatically identify the abstract themes addressed in a collection of documents. Such models are based on the hypothesis that each document in the corpus corresponds to a distribution of several topics. A Bitern Topic Model (BTM) was used to identify the topics without prior knowledge. A topic is defined as a subject of discussion, which amounts to tokens that frequently appear together in a corpus. The BTM considers the whole corpus as a mixture of topics, where each co-occurring pair of tokens (the bitern) is drawn from a specific topic independently and modelled topics are probability distributions over the biterns.

Topics being probability distributions over tokens of the corpus of study, they can be characterized by the highest per-topic probability tokens. Weighting these probabilities through term-frequency inverse document frequency (TF-IDF) weighting allows to allocate a higher importance to topic-specific tokens. In this case, the per-topic probability of a token was weighted by the inverse of the probabilities of this token in other topics. For each topic, tokens were ranked from highest to lowest weighted probabilities TF-IDF value of their probability in this topic. The first 15 tokens are designated as the set of characteristic tokens and used to manually name the topic. A topic model has been applied to each of the sub-corpus.

Health-Related Quality of Life (HRQoL): A HRQoL algorithm was applied to identify and qualify expressed impact of the disease or treatments on quality of life of a patient. Types of impact are defined according to HRQoL survey categories. The algorithm is two-fold: first it indicates if an impact is expressed and, thanks to 5 specific models, it indicates the nature of the impact: Physical, Psychic, Activity-related, Relational or Financial. Features involved in the model describe expressed emotions, grammar, conjugation, and lexical fields of HRQoL related features. The HRQoL algorithm has been applied to each of the sub-corpus.

Drug Intake: To identify drug intake by the author of the message, the first step consisted of identifying the treatments cited in the two sub-corpus of messages. These were detected from the Detec't database, which contains around 2500 molecules and drugs. After that, a prediction was made on a presence couple, i.e., a product-message couple. The model bases its prediction on regular forms expressing the intake of a treatment and their distance to the drug mention. Expressions are scored according to their probability of implying a treatment intake.

Feeling Analysis: In both sub-corpus, the content of the drug intake messages was analysed (sentiment analysis) using Microsoft Azure Cognitive Services. Sentiment Analysis is part of Text Analytics and can detect the level of positive or negative sentiment of input text using a confidence score across a variety of languages. This machine learning algorithm assigns a sentiment score to each sentence of a message, then calculates the overall score to assign the message to a category (positive, negative, mixed, or neutral).

The sentiment analysis algorithm has been applied to each of the drug intake messages' sub-corpus. For drugs with few identified intakes, sentiment analysis is applied to treatment citation instead of treatment intake: this was the case for Dupixent for example.

The qualitative content of the positive and negative feelings in the messages was characterized by a manual identification with a review of posts detected with a treatment taken by the web users. The percentage of messages was calculated according to the total number of messages identified with terms related to treatments taken.

Results

Description of the Population and Posts

After cleaning and formatting, the analysis corpus contained a total of 33115 messages corresponding to 15857 different web users, with a median of 2.1 posts per patient (standard

deviation 9.55) (Figure 3). *Controlled AD* sub-corpus contained 9454 posts corresponding to 4875 web users and *Insufficiently Controlled AD* sub-corpus contained about twice as many messages and web users with 17384 posts from 9292 web users who published fewer messages than the other sub-corpus (median of 1.87 posts per patient versus 1.94).

Figure 3. Methodology. The lexical fields of *Controlled AD* and *Insufficiently Controlled AD* have been realized based on the ADCT questionnaire and completed with real expressions of internet users.

<Insert Figure 3 here>

For both subgroups, extracted data mostly came from Babycenter.fr with 889/9454 (9.40%) posts for *Controlled AD* sub-corpus and 2475/17384 (14.24%) post for *Insufficiently Controlled AD* sub-corpus (Table 1). The most frequently identified keyword was “Eczema” (7717/9 454 posts for *Controlled AD* sub-corpus and 13039/17384 posts for *Insufficiently Controlled AD* sub-corpus respectively) followed by “Dermatite” (1146/9454 posts for *Controlled AD* sub-corpus and 2386/17384 posts for *Insufficiently Controlled AD* sub-corpus respectively).

Table 1. Top 10 websites. On the left panel are presented the top ten web sites surveyed by patients concerning *Controlled AD* sub-corpus. On the right panel are presented the top ten web sites surveyed by patients concerning *Insufficiently Controlled AD* sub-corpus.

<i>Controlled AD</i>		<i>Insufficiently controlled AD</i>	
Forums (TOP10)	Posts	Forums (TOP10)	Posts
BABYCENTER.FR	889	BABYCENTER.FR	2475
AU FEMININ	307	TWITTER	2431
LAPPART-DES-	230	LAPPART-DES-	1288
SPASMOS.FR		SPASMOS.FR	
AMAZON.FR	220	AU FEMININ	893
MAMANDCO.FR	206	MAMANDCO.FR	385
TWITTER	128	BOURSORAMA.COM	282
E-SANTE	109	AMAZON.FR	228
FORUM-MELODIE.FR	88	MJEUXVIDEO.COM	197
CLUB.BEAUTE-	61	FORUM.MANUCURE.INFO	126
ADDICT.COM			
JOURNALDESFEMMES.FR	38	BOURSEREFLEX.COM	121

We observed fluctuations in the evolution of the volume of messages between July 2010 and October 2020 (Figure 4) due to either seasonality or publications of scientific results related to AD. An

apparent recurrent increase in the number of messages was observed from November to July that could be related to the increase in dust mites' allergies in winter and the discomfort associated with sweating in summer [25,26]. From February to December 2015 there was a significant increase in the number of posts which seems to be related to the publication of different scientific articles especially concerning the results of Tofacitinib citrate, an oral Janus kinase inhibitor. Another increase in messages was also seen in 2018 that can be associated with the availability of Dupixent® (dupilumab) in hospitals.

Figure 4. Fluctuations in the evolution of the volume of messages between July 2010 and October 2020.

<Insert Figure 4 here>

In both sub-corpus, most users were women (72%) with an average age of about 38 years and 19% appeared with undetermined gender (Figure 5). This is consistent with the results of many studies that point out that women express more personal issues in social networks [27,28]. In the present case, the topic of the disease is much discussed by mothers of children with AD who tended to focus on the appearance of their child's skin.

Figure 5. Age and gender distribution among *Controlled AD* sub-corpus (left panel) and *Insufficiently Controlled* sub-corpus (right panel)

<Insert Figure 5 here>

As shown in Figure 6, patients with AD are increasingly active online and use social media to participate, share their concerns, and express themselves freely regarding their disease. This trend is comparable to that of psoriasis patients, even if the number of messages is lower for AD that is not sufficiently prioritized. On the other hand, the comparison with the cancer disease area (1% of the messages) clearly highlights the lack of specific forums and/or associations for AD compared to cancer field.

Figure 6. Trend over time of the number of posts concerning atopic dermatitis compared to other diseases such as cancer, and psoriasis. It is to note that for cancer only the 1% of posts were reported.

<Insert Figure 6 here>

Discussion themes and topics. Basically, the topics around AD are roughly the same between the two sub-corpus and the observed differences are due to the stage of their disease. Patients from

Controlled AD sub-corpus have more perspective and experience, whereas patients from *Insufficiently Controlled AD* sub-corpus lack experience and are looking for information. As shown in Figure 7, the largest category of topics was related to “the exchange of tips to relieve eczema” (45.18%), followed closely by “hygiene and care adapted with the disease” (24.22%) for the *Controlled AD* sub-corpus. For the *Insufficiently Controlled AD* sub-corpus, 42.19% of the discussions were about “eczema in babies and children” with most messages related to discussions between parents who are concerned about their child's eczema and who ask questions to other web users. The second most discussed topic, with 14% of messages in the *Insufficiently Controlled AD* sub-corpus, is the therapeutic failure experienced by web users. Their concerns reflected the poor knowledge of their disease and treatments. They express their frustration and lack of satisfaction about the treatments and solutions they have tested to reduce their symptoms. Whereas we can see that in the *Controlled AD* sub-corpus, web users exchange about natural remedies and alternative medicines to treat their atopic dermatitis. They express their satisfaction with the use of essential oils, food supplements and other non-medicinal methods and the effectiveness of these solutions. They used social networks to provide advice based on their own experience.

Figure 7. Distribution of posts by overall topic category. Application of a Biterm Topic Model (BTM) that automatically identifies the different topics addressed within the message corpus. Each theme is associated with a set of characteristic words which allows to define them. A message can only be associated with one theme.

<Insert Figure 7 here>

Quality of life (QoL). The QoL is an important tool for evaluating the effect of a disease as well as effects of treatment interventions. In the present study, the impact of AD on the quality of life of individuals was assessed by monitoring physical well-being, social well-being, material well-being, emotional well-being, and development/activity.

The impact of the disease on QoL is discussed by 70% of patients in the *Controlled AD* sub-corpus (Figure 8a) and 49% for *Insufficiently Controlled AD* sub-corpus (Figure 8b). This lower proportion can be explained by the heavy burden that AD places on patients with *Insufficiently Controlled AD*, particularly those who search for a solution to improve their health rather than their quality of life. Furthermore, this trend can also be related to the lack of awareness about their disease compared to patients from the *Controlled AD* sub-corpus, who have more knowledge of AD and therefore speak about their lived experience.

Patients in both sub-corpus seem to report the same impacts in their messages, but with a time lag that corresponds to the evolution of the knowledge of their disease.

Social functioning is affected by atopic dermatitis and is the first self-reported burden in both sub-corpus. Social isolation can be seen in young children presenting AD because adults and other children avoided interacting with them. The main effect of AD on the lives of adult patients stems from embarrassment and not wanting to be seen in public. Furthermore, some web users reported that they have been teased or bullied because of their AD. This trend was also reported by the International Study on Life with Atopic Eczema (ISOLATE) that found major impacts of AD on self-esteem [29].

The impacts that follows are the emotional impact (behavioural problems, irritability, crying, problems with treatments), the impact on activities (interference with activities such as bathing, swimming, and playing, clothing restrictions, decreased productivity at work, absenteeism), the impact on physical health (including itching, scratching, pain) and the economic impact (over-the-counter pharmacy costs among others). Independently of the corpus, there are on average 3 different impacts per message which shows the importance of this disease in everyday life.

Figure 8. a) Impact of the disease on Controlled AD sub-corpus. b) Impact of the disease on Insufficiently Controlled AD sub-corpus.

<Insert Figure 8 here>

Treatments

AD treatment is targeted at both the disease and its symptoms. Among the 9454 messages retrieved from *Controlled AD* sub-corpus users, 1229 concerned discussions on the different treatments taken by the patients (i.e., 13%). A focus was made on the TOP10 and it appeared that *Controlled AD* sub-corpus web users were mainly focused on cortisone/corticosteroids and anti-inflammatory therapy. In addition, patients were prescribed antihistamines to decrease the itch-scratch cycle and antibiotics if the skin becomes superinfected (Table 2 Left panel). Among the 17384 messages retrieved from *Insufficiently Controlled AD* sub-corpus users, 1677 concerned discussions on the different treatments taken by the patients (i.e., 10%). Topically applied corticosteroids and emollients, moisturizers and bath additives are mainstay of therapy used by *Insufficiently Controlled AD* sub-corpus web users followed by anti-inflammatory therapy to control pruritus and antihistamines as a therapeutic adjunct to alleviate it. In this sub-corpus, treatments for acute phases as topical

corticosteroids for example (Diprosone®) were also discussed by *Insufficiently Controlled AD* sub-corpus patients (Table 2 Right panel).

Table 2. Top10 treatments used. Left panel/ *Controlled AD* sub-corpus. Right panel/ *Insufficiently Controlled AD* sub-corpus

<i>Controlled AD</i>		<i>Insufficiently controlled AD</i>	
Treatments taken (TOP10)	Web users	Treatments taken (TOP10)	Web users
Cortisone/Corticoids	131	Cortisone/Corticoids	222
Anti-inflammatories	32	Liniment	40
Dexeryl	28	Dexeryl	39
Calendula	28	Anti-inflammatories	26
Liniment	26	Calendula	26
Antibiotics	26	Antibiotics	17
Vitamins	12	Diprosone	16
Atarax	8	Bepanthen	11
Bepanthen	6	Aerius	10
Aerius	5	Vitamins	10

Then, the perception of treatments is found in 63% of the messages related to treatments taken by *Controlled AD* sub-corpus (778/1229 messages) and in 72% of the messages dealing with treatments taken by *Insufficiently Controlled AD* sub-corpus (1209/1677 messages). In both sub-corpus, the most expressed treatment perception is cortisone.

Mixed perception is found in 56% of the *Controlled AD* and 66% of the *Insufficiently Controlled AD* with some concerns on the risks/side-effects of cortisone emphasized by its efficacy. Then the perception is mostly negative (25% for *Controlled AD* and 23% for *Insufficiently Controlled AD* sub-corpus) and positive in only 8% of *Controlled AD* and in 7% of *Insufficiently Controlled AD* web users (Figure 9). It is to note that in the lower panel, Dupixent is not a medication taken by *Insufficiently Controlled AD* patients but only cited in their messages on which an analysis of perception was applied.

Figure 9. Treatments' perception. Upper panel/ *Controlled AD* sub-corpus. Lower panel/ *Insufficiently Controlled AD* sub-corpus

<Insert Figure 9 here>

Concerning cortisone, the *Controlled AD* sub-corpus and the *Insufficiently Controlled AD* sub-corpus expressed safety concerns associated with its long-term use with negative feelings and beliefs (25% and 22% of the messages, respectively) while 8% of the *Controlled AD* sub-corpus and 7% of the *Insufficiently Controlled AD* sub-corpus felt a positive perception following the effectiveness they had experienced. Positive and negative perception was quite counter balanced between side effects and satisfaction for the texture and the use as a preventive treatment in *Controlled AD* patients using Dexeryl (9% of positive and negative messages), while patients from the *Insufficiently Controlled AD* sub-corpus reported more concerns about safety (32%). The ineffectiveness of anti-inflammatories to calm the flare-ups was a concern for 19% of *Controlled AD* and 33% for *Insufficiently Controlled AD* whereas only 5% and 17% respectively had a positive efficacy feeling. It is to note that for Dupixent, there was an absence of messages from internet users who have taken or are aware of the treatment that may be due to its restricted prescription by dermatologists or directly from hospital and because it is not prescribed as first intention AD therapy. The only messages found were from patients seeking information mainly on the conditions of access (8%) and on the effectiveness of the drug (8%) (Figure 10).

Figure 10. Treatments' perception. 10a/ *Controlled AD* sub-corpus. 10b/ *Insufficiently Controlled AD* sub-corpus

<Insert Figure 10 here>

Discussion

Principal Findings

This retrospective study aimed at assessing how patients suffering from AD and/or their parents perceive their quality of life and their treatments following the analysis of web forums in France over the last 10 years. AD is increasingly discussed in French web forums as shown in this study (Figure 6) with most users being women between the ages of 30-40 (Figure 5).

We were able to segment web users into two sub-corpus corresponding to *Controlled AD* patients and *Insufficiently Controlled AD* patients. These sub-corpus are characterized by similarities in the

expressions. Web users from the *Controlled AD* sub-corpus and those from the *Insufficiently Controlled AD* sub-corpus have almost the same preoccupations but we observed a time lag due in part to the stage and the knowledge of their disease. The main topics were the lack of information and solutions leading to some sort of therapeutic wandering, the high impact of AD on the quality of life and the well-being of patients, and the dissatisfaction with the available treatments.

Limitations of the Study

Given the inherent observational nature of social media data, their analyses are subject to many limitations.

Selection bias was the first limitation because analyses were restricted to French data sources, and social media users and non-users may differ. Thus, results are not generalizable at a worldwide scale. Furthermore, using social media to analyse patients' reactions excludes patients who do not have access to the Internet or who are not familiar with the use of online discussions.

Extraction bias was the second limitation. Keyword selection in social media studies can induce varying levels of extraction bias.

Another limitation of using social media is that complete information about individual cases may be harder to obtain, unlike traditional epidemiological studies. There is also the problem of discovering demographic information—only limited or no information regarding individual user demographics may be available. Demographic information such as age and race need to be determined via automated techniques.

Health misinformation is significant on social media, nevertheless, recent reviews showed that misinformation or fake information related to health are most prevalent in studies dealing with the safety of tobacco, vaccines, and drugs such as opioids and marijuana. Finally, the lowest levels of misinformation were observed in studies evaluating medical treatments [30,31]. Nevertheless, results from Pulido et al. [32] indicated that messages focused on fake health information are mostly aggressive and that messages with evidence of social impact overcome fake information.

Finally, social media represent an ideal place where patients can freely and spontaneously discuss their experiences with their therapy, thus providing valuable information on their QoL. However, this observation should be interpreted cautiously, since social media data may include a higher frequency of erroneous information, and patients posting on social media forums may not be representative of the wider patient population.

Implications and Future Research

Messages published on social networks should be integrated into the assessment of patient's quality of life, as they can help to characterise the patient's experience in a more individualised and spontaneous way. Furthermore, it seems important to explore the specific areas of AD patients' QoL and potentially enrich the existing standard questionnaires with new elements more relevant for these patients in their daily confrontation with disease and treatment.

To the extent that the most dominant topics can be interpreted as unmet informational needs, our study highlights the refinement of practical implications such as the improvement of available tools, and further analysis on the perception of treatments and the evolution of the quality of life under treatment. This study suggested to carry out a study segmented by year to see the evolution of the different subjects (QoL, topics, treatments) over the years and to observe the global evolutions.

Conclusion

Social media listening offers the opportunity to consider behaviour and interactions that are difficult to assess through traditional research methods. Millions of microblogs act as online communities, dealing with topics from the impact of AD on quality of life through the lack of information on possible treatments and their effectiveness. Of particular interest is the exponential growth in recent years of patient support groups, and the high potential of users disseminating materials and opinions relating to AD through their posts. These forums are increasingly popular and have become an additional source of evidence, therapy, or support for the patients. Our study illustrates the current situation and the evolution over time of AD on social media platforms. Discussions on AD are becoming a frequent purpose on social media but they are still in contrast with the high frequency reported for the use of social media in research with psoriatic patients.

All these data highlight the importance of clearly defining the role and limitations of these platforms for orienting future information campaigns and developing new models such as Atopic Dermatitis Control Tool centred on patients with AD [33]. This validated 6-item ADCT facilitates patient-physician communication on disease control. Such tool might better inform health care professionals and patients with individualized measures covering quality of life more in depth than existing standard questionnaires.

Acknowledgments

The medical writing support was provided by Nathalie Baudry of the microenterprise NB-

Pharmaconsult. The study and the medical writing support have been funded by Sanofi.

Conflicts of Interest

BR and MP have disclosed that they are employees of Sanofi Genzyme. May hold shares and/or stock options in the company.

PV, PF, AG, AM, NT and SS have disclosed that they are employees of Kap Code, a CRO, with contracted with Sanofi Genzyme to carry out this study. SB, LM and SH have nothing to declare.

Abbreviations

AD: atopic dermatitis

ADCT: atopic dermatitis control tool

AI: artificial intelligence

BTM: biterm topic model

HRQoL: health-related quality of life

ISOLATE: international study on life with atopic eczema

QoL: quality of life

TF-IDF: term-frequency inverse document frequency

REFERENCES

1. Eichenfield LF, Tom WL, Berger TG, et al. Guidelines of care for the management of atopic dermatitis: section 2. Management and treatment of atopic dermatitis with topical therapies. *J Am Acad Dermatol*. 2014;71(1):116-132. doi:10.1016/j.jaad.2014.03.023
2. Berke R, Singh A, Guralnick M. Atopic dermatitis: an overview. *Am Fam Physician*. 2012;86(1):35-42.
3. Drucker AM, Wang AR, Li W-Q, Sevetson E, Block JK, Qureshi AA. The Burden of Atopic Dermatitis: Summary of a Report for the National Eczema Association. *J Invest Dermatol*. 2017;137(1):26-30. doi:10.1016/j.jid.2016.07.012
4. Laughter MR, Maymone MBC, Mashayekhi S, et al. The global burden of atopic dermatitis: lessons from the Global Burden of Disease Study 1990-2017. *Br J Dermatol*. 2021;184(2):304-309. doi:10.1111/bjd.19580
5. Hanifin JM, Reed ML, Eczema Prevalence and Impact Working Group. A population-based survey of eczema prevalence in the United States. *Dermatitis*. 2007;18(2):82-91. doi:10.2310/6620.2007.06034
6. Lee HH, Patel KR, Singam V, Rastogi S, Silverberg JI. A systematic review and meta-analysis of the prevalence and phenotype of adult-onset atopic dermatitis. *J Am Acad Dermatol*. 2019;80(6):1526-1532.e7. doi:10.1016/j.jaad.2018.05.1241
7. Silverberg JI, Gelfand JM, Margolis DJ, et al. Patient burden and quality of life in atopic dermatitis in US adults: A population-based cross-sectional study. *Ann Allergy Asthma Immunol*. 2018;121(3):340-347. doi:10.1016/j.anai.2018.07.006
8. Margolis JS, Abuabara K, Bilker W, Hoffstad O, Margolis DJ. Persistence of mild to moderate atopic dermatitis. *JAMA Dermatol*. 2014;150(6):593-600. doi:10.1001/jamadermatol.2013.10271
9. Silverberg JI. Selected comorbidities of atopic dermatitis: Atopy, neuropsychiatric, and musculoskeletal disorders. *Clin Dermatol*. 2017;35(4):360-366. doi:10.1016/j.clindermatol.2017.03.008
10. Kage P, Zarnowski J, Simon J-C, Treudler R. Atopic dermatitis and psychosocial comorbidities - What's new? *Allergol Select*. 2020;4:86-96. doi:10.5414/ALX02174E
11. Silverberg JI. Association between childhood atopic dermatitis, malnutrition, and low bone mineral density: a US population-based study. *Pediatr Allergy Immunol*. 2015;26(1):54-61. doi:10.1111/pai.12315
12. Garg NK, Silverberg JI. Eczema is associated with osteoporosis and fractures in adults: a US population-based study. *J Allergy Clin Immunol*. 2015;135(4):1085-1087.e2. doi:10.1016/j.jaci.2014.10.043
13. Narala S, Hata TR. Adult Atopic Dermatitis with Comorbid Atopic Disease is Associated with Increased Risk of Infections: A Population-Based Cross-Sectional Study. *Dermatol Ther*

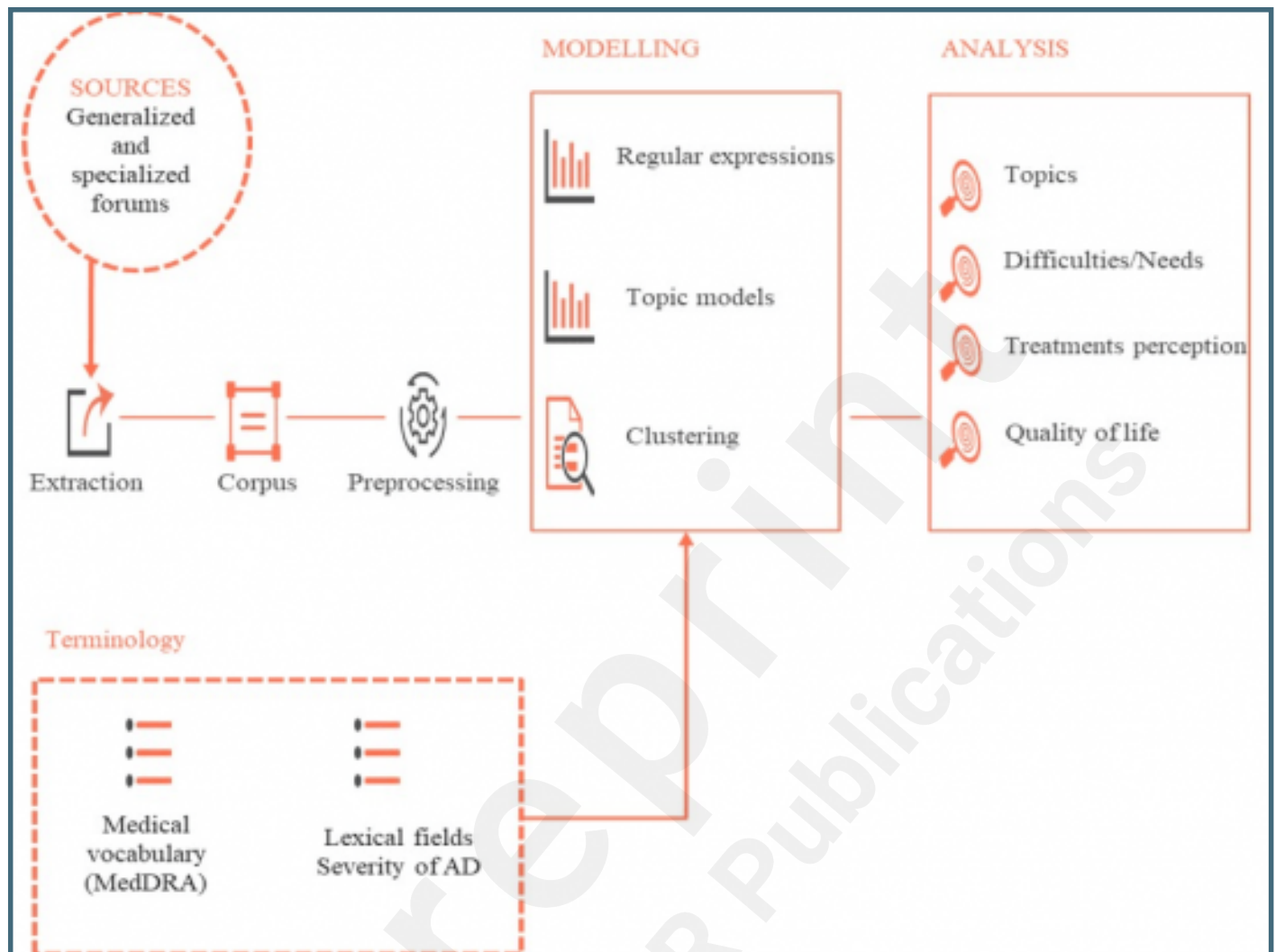
- (Heidelb). 2017;7(1):111-121. doi:10.1007/s13555-017-0172-7
14. Narla S, Silverberg JI. Association between atopic dermatitis and serious cutaneous, multiorgan and systemic infections in US adults. *Ann Allergy Asthma Immunol*. 2018;120(1):66-72.e11. doi:10.1016/j.anai.2017.10.019
 15. Silvestre Salvador JF, Romero-Pérez D, Encabo-Durán B. Atopic Dermatitis in Adults: A Diagnostic Challenge. *J Investig Allergol Clin Immunol*. 2017;27(2):78-88. doi:10.18176/jiaci.0138
 16. Johnson BB, Franco AI, Beck LA, Prezzano JC. Treatment-resistant atopic dermatitis: challenges and solutions. *Clin Cosmet Investig Dermatol*. 2019;12:181-192. doi:10.2147/CCID.S163814
 17. Dupixent SPC. DUPIXENT: SUMMARY OF PRODUCT CHARACTERISTICS.
 18. von Muhlen M, Ohno-Machado L. Reviewing social media use by clinicians. *J Am Med Inform Assoc*. 2012;19(5):777-781. doi:10.1136/amiajnl-2012-000990
 19. Donegan K, Ovelgonne H, Flores G, Fuglerud P, Georgescu A. Social Media and M-Health Data. :29.
 20. 10 Tech & IT Buzzwords For 2021 You Won't Be Able To Avoid. BI Blog | Data Visualization & Analytics Blog | datapine. Published December 1, 2020. Accessed February 10, 2021. <https://www.datapine.com/blog/technology-buzzwords/>
 21. Shaban-Nejad A, Michalowski M, Buckeridge DL. Health intelligence: how artificial intelligence transforms population and personalized health. *NPJ Digit Med*. 2018;1:53. doi:10.1038/s41746-018-0058-9
 22. Abdellaoui R, Schück S, Texier N, Burgun A. Filtering Entities to Optimize Identification of Adverse Drug Reaction From Social Media: How Can the Number of Words Between Entities in the Messages Help? *JMIR Public Health and Surveillance*. 2017;3(2):e6577. doi:10.2196/publichealth.6577
 23. Kürzinger M-L, Schück S, Texier N, et al. Web-Based Signal Detection Using Medical Forums Data in France: Comparative Analysis. *Journal of Medical Internet Research*. 2018;20(11):e10466. doi:10.2196/10466
 24. Najork M. Web Crawler Architecture. In: LIU L, ÖZSU MT, eds. *Encyclopedia of Database Systems*. Springer US; 2009:3462-3465. doi:10.1007/978-0-387-39940-9_457
 25. Fleischer AB. Atopic dermatitis: the relationship to temperature and seasonality in the United States. *Int J Dermatol*. 2019;58(4):465-471. doi:10.1111/ijd.14289
 26. Dhar S, Kanwar AJ. Epidemiology and clinical pattern of atopic dermatitis in a North Indian pediatric population. *Pediatr Dermatol*. 1998;15(5):347-351. doi:10.1046/j.1525-1470.1998.1998015347.x
 27. Rollero C, Daniele A, Tartaglia S. Do men post and women view? The role of gender, personality and emotions in online social activity. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*. 2019;13(1). doi:10.5817/CP2019-1-1

28. Gender-Specific Behaviors on Social Media and What They Mean for Online Communications. Social Media Today. Accessed February 17, 2021. <https://www.socialmediatoday.com/social-networks/gender-specific-behaviors-social-media-and-what-they-mean-online-communications>
29. Zuberbier T, Orlow SJ, Paller AS, et al. Patient perspectives on the management of atopic dermatitis. *Journal of Allergy and Clinical Immunology*. 2006;118(1):226-232. doi:10.1016/j.jaci.2006.02.031
30. 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
31. 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
32. Pulido CM, Ruiz-Eugenio L, Redondo-Sama G, Villarejo-Carballido B. A New Application of Social Impact in Social Media for Overcoming Fake News in Health. *Int J Environ Res Public Health*. 2020;17(7):E2430. doi:10.3390/ijerph17072430
33. Pariser DM, Simpson EL, Gadkari A, et al. Evaluating patient-perceived control of atopic dermatitis: design, validation, and scoring of the Atopic Dermatitis Control Tool (ADCT). *Curr Med Res Opin*. 2020;36(3):367-376. doi:10.1080/03007995.2019.1699516

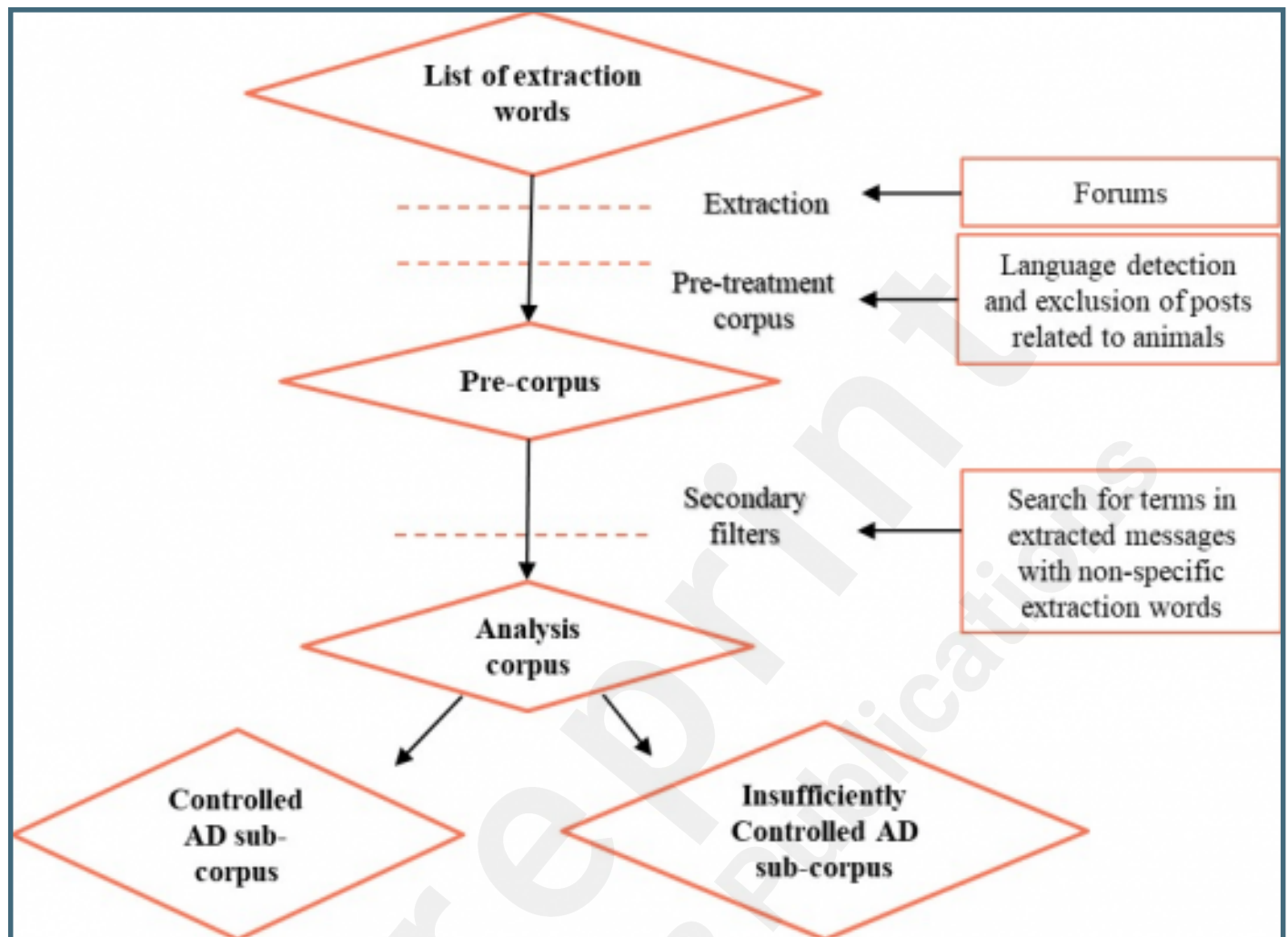
Supplementary Files

Figures

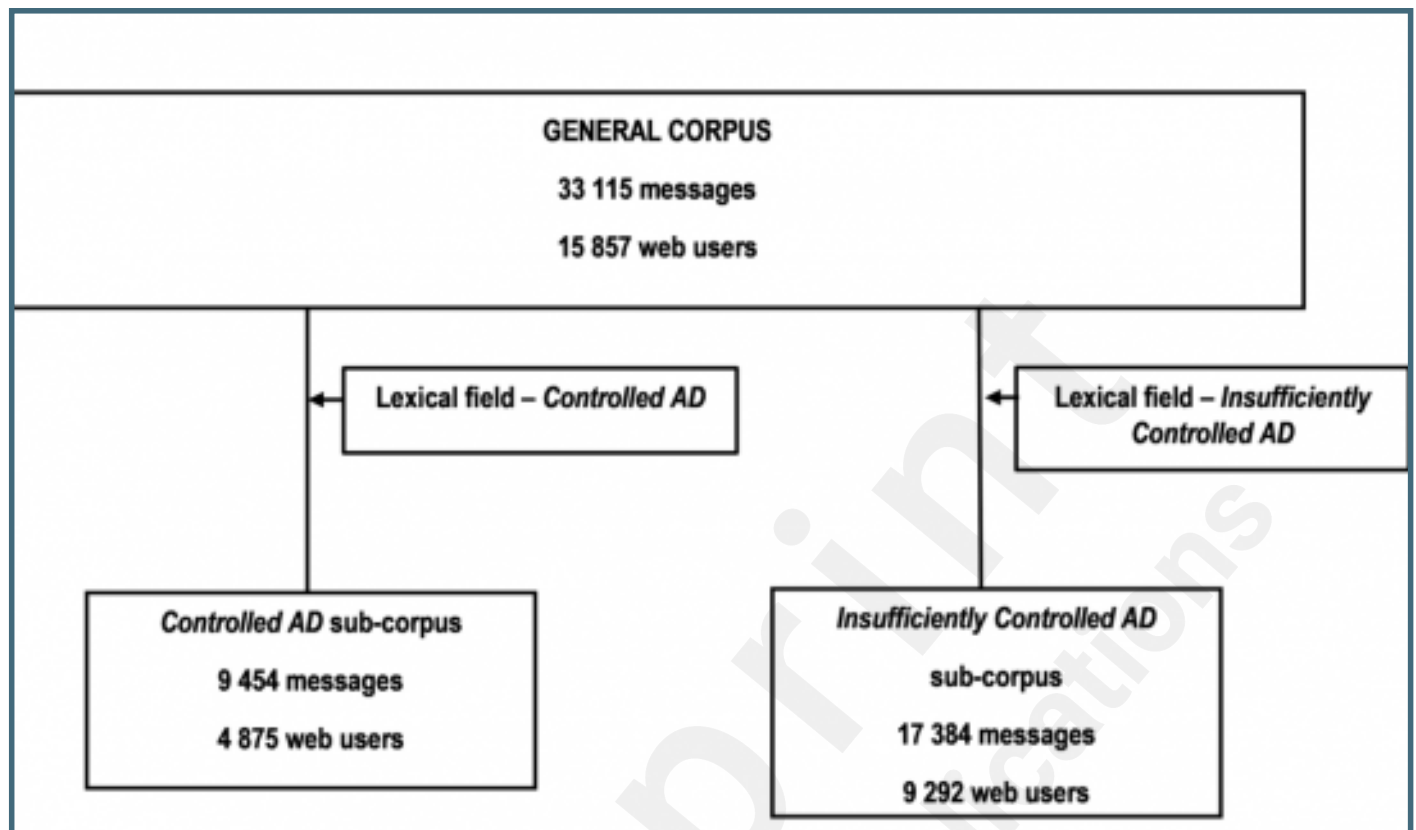
Study framework.



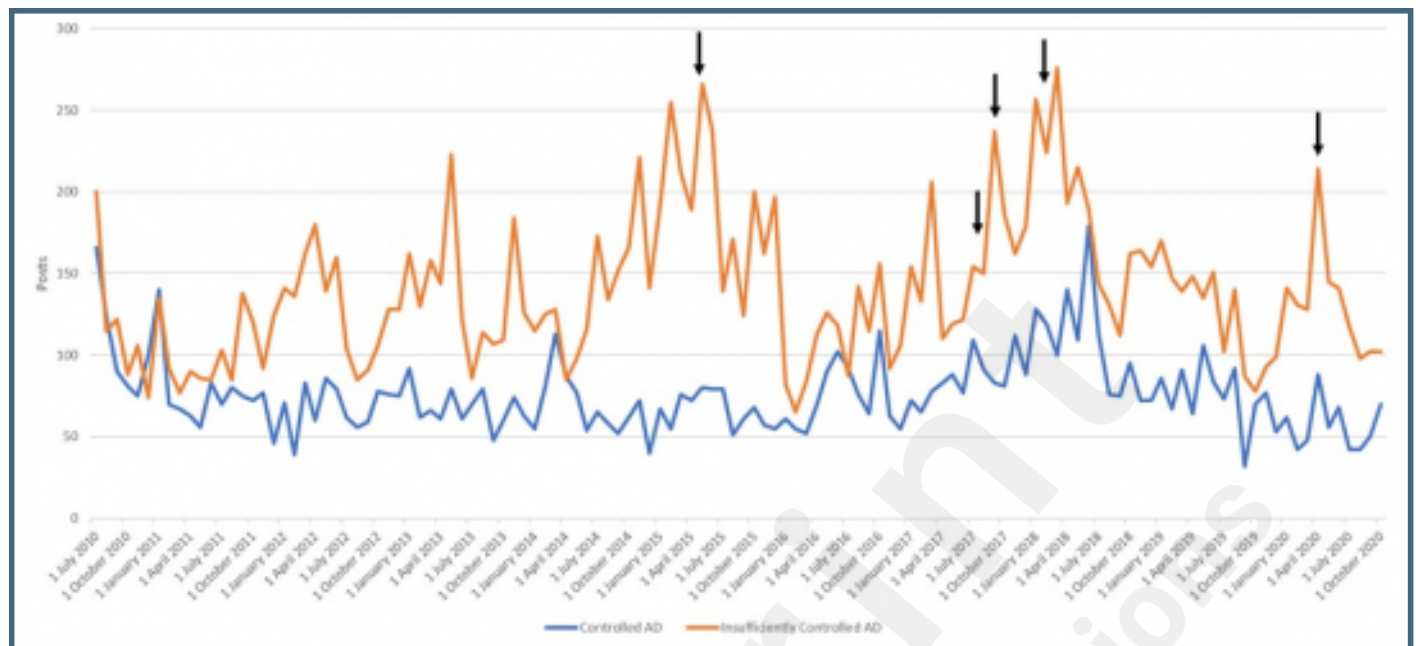
Flowchart presenting the steps for creating analysis corpus and the two sub-corpus.



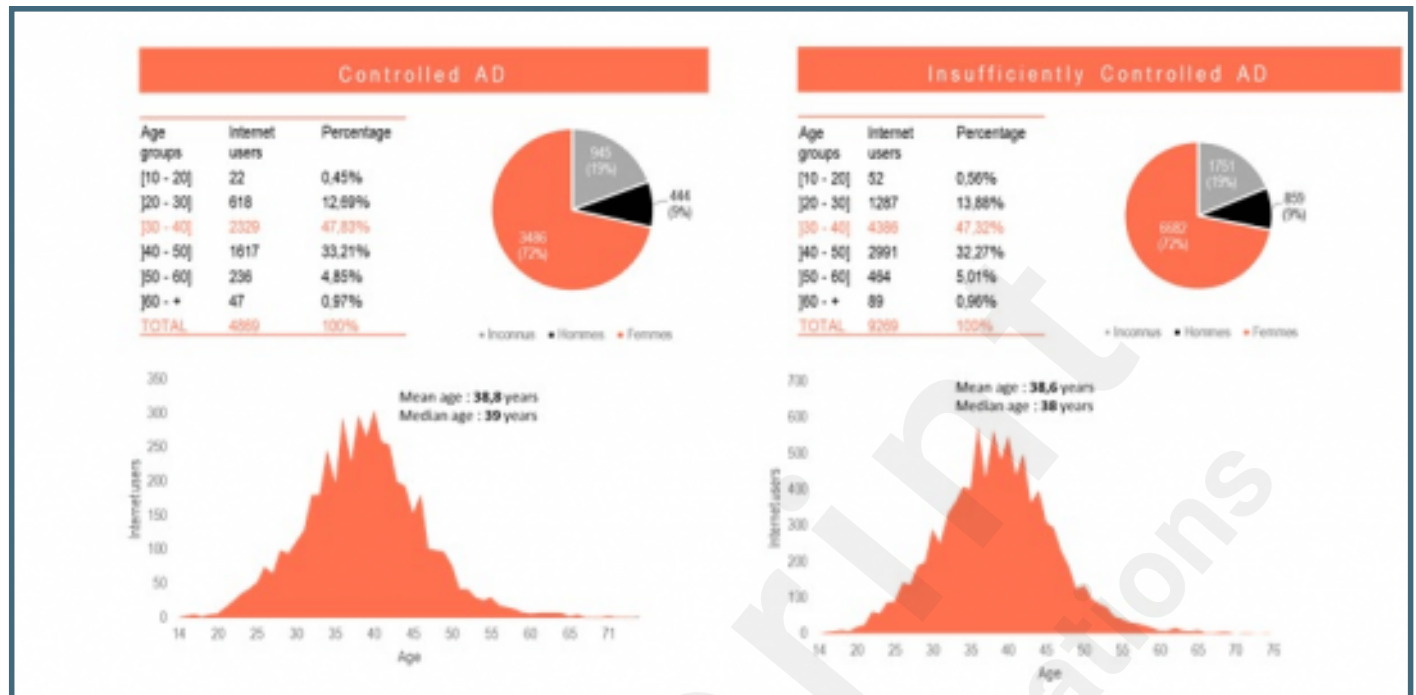
Methodology.



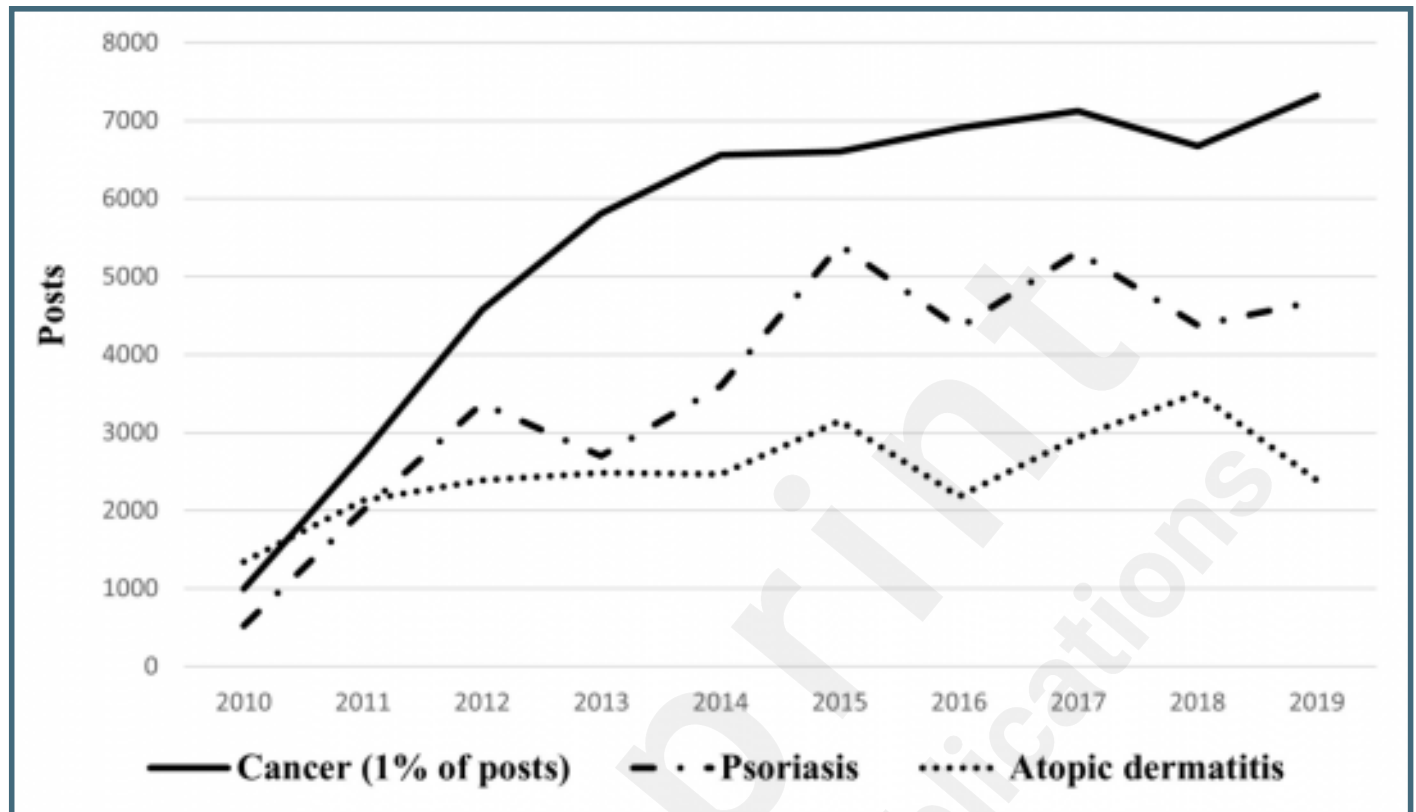
Fluctuations in the evolution of the volume of messages between July 2010 and October 2020.



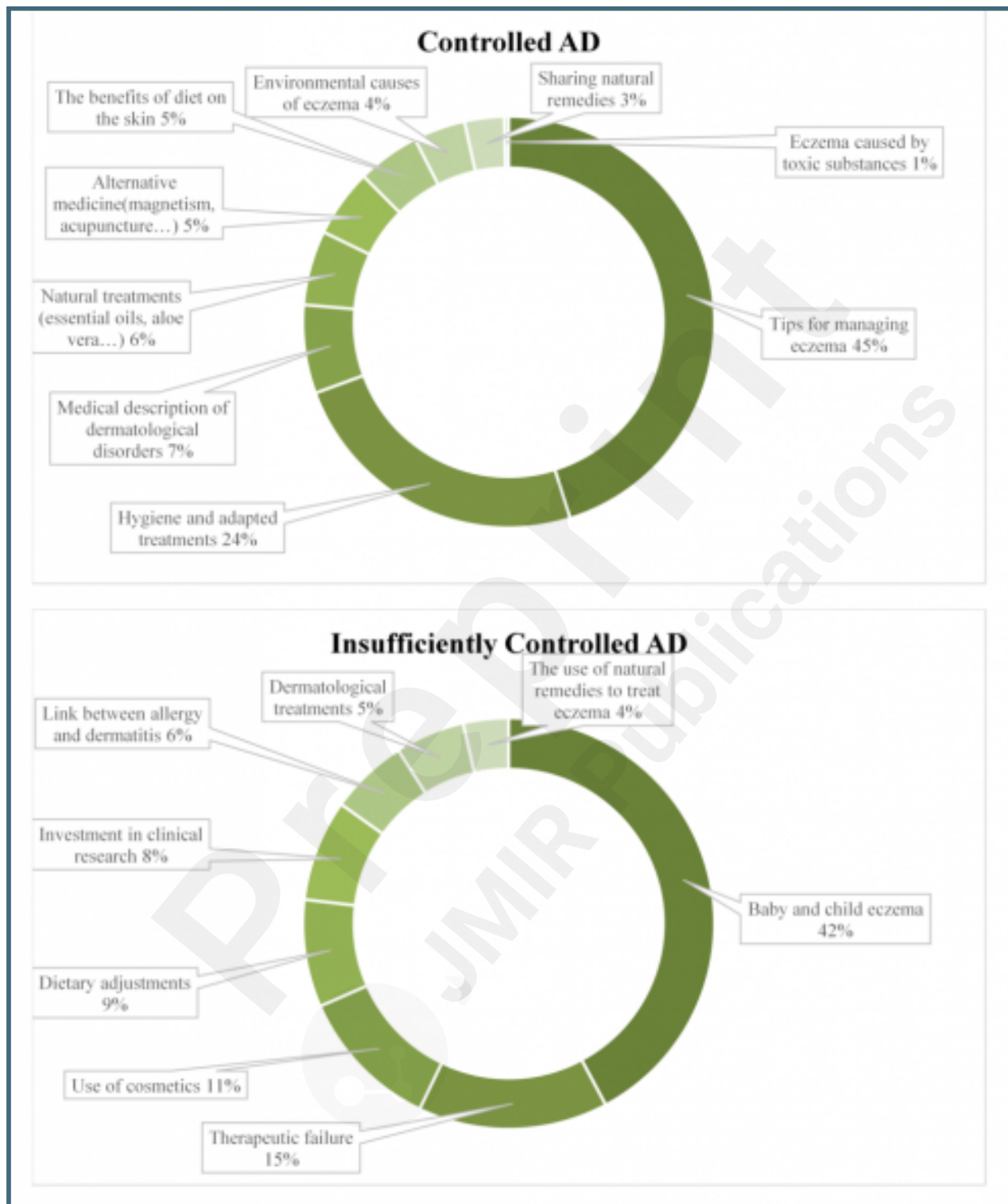
Age and gender distribution among Controlled AD sub-corpus (left panel) and Insufficiently Controlled sub-corpus (right panel).



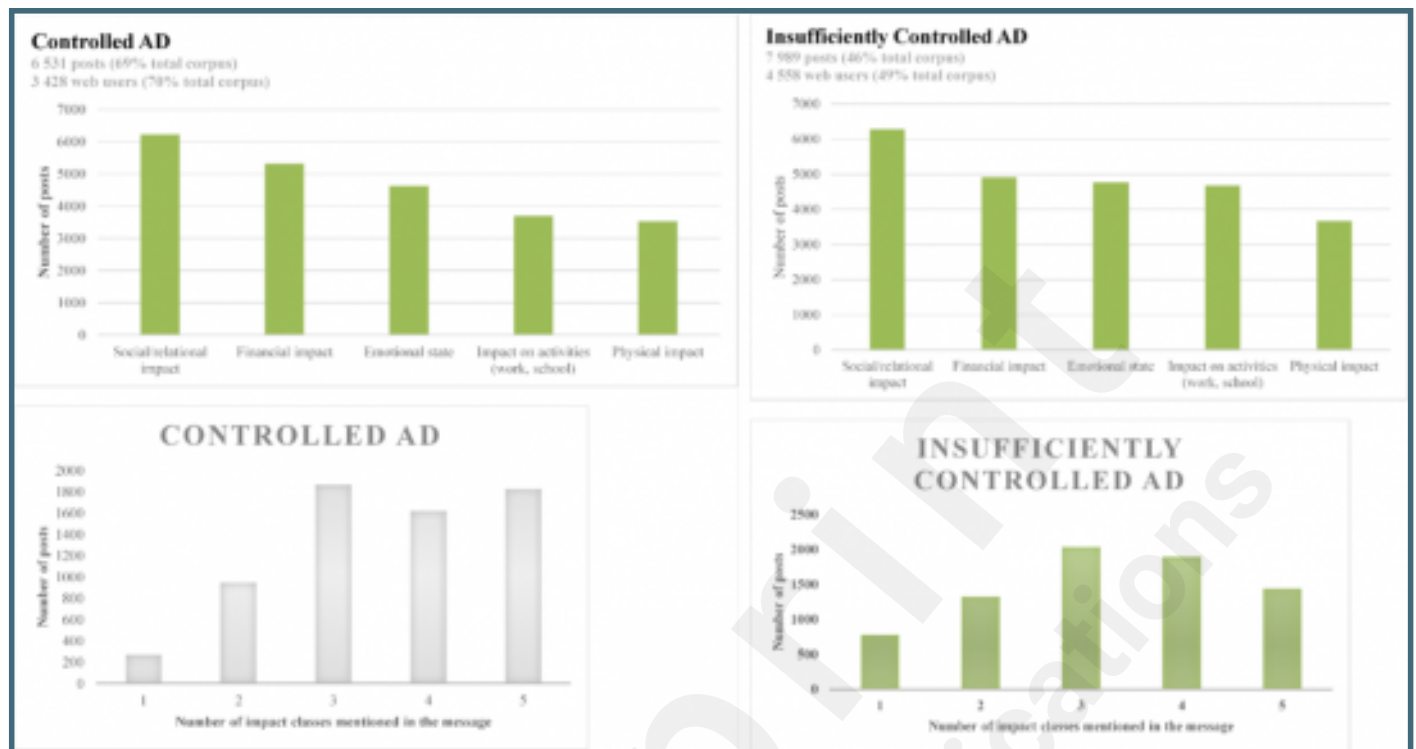
Trend over time of the number of posts concerning atopic dermatitis compared to other diseases such as cancer, and psoriasis.



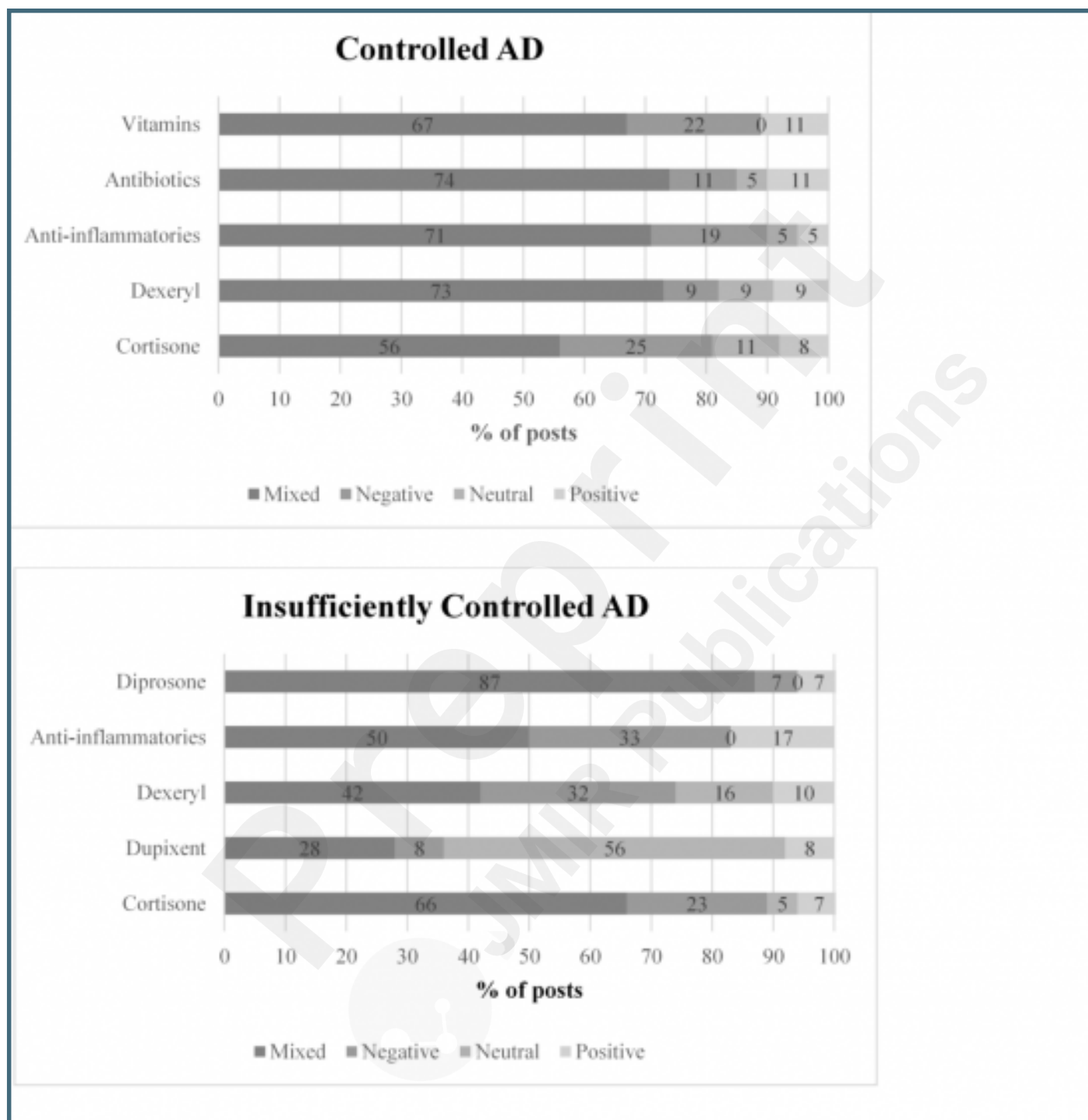
Distribution of posts by overall topic category.



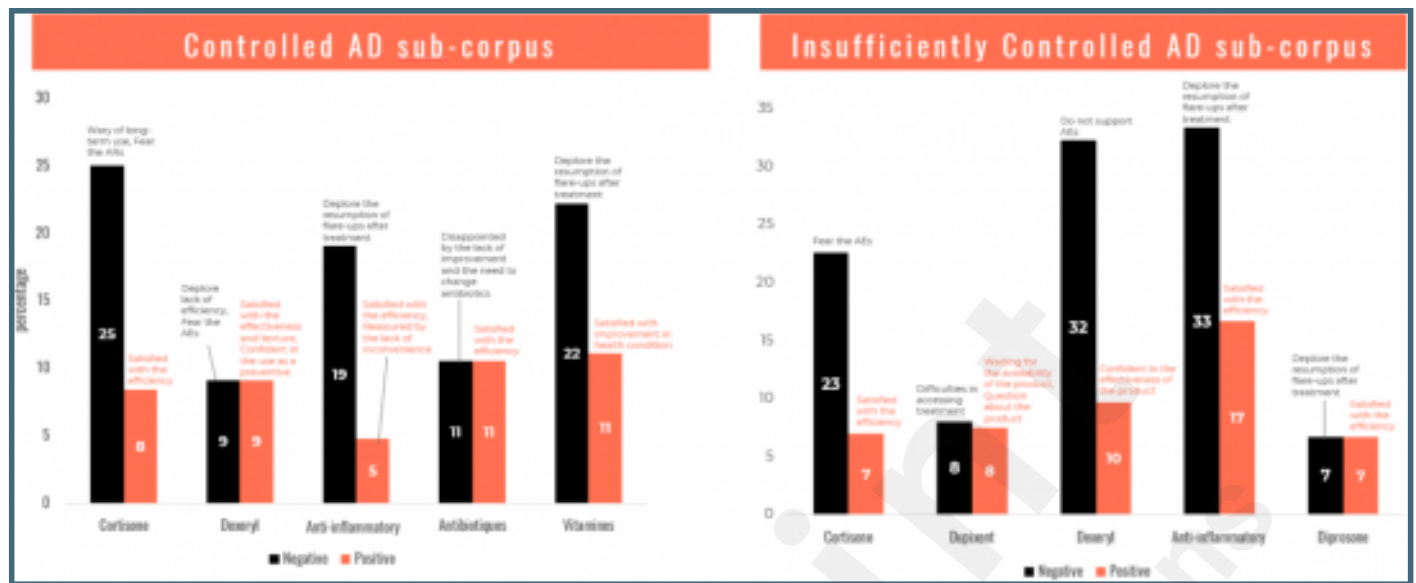
a) Impact of the disease on Controlled AD sub-corpus. b) Impact of the disease on Insufficiently Controlled AD sub-corpus.



Treatments' perception. Upper panel/ Controlled AD sub-corpus. Lower panel/ Insufficiently Controlled AD sub-corpus.



Treatments' perception. 10a/ Controlled AD sub-corpus. 10b/ Insufficiently Controlled AD sub-corpus.



Multimedia Appendixes

List of keywords used for the extraction of messages.

URL: <http://asset.jmir.pub/assets/7b2d357e339c5a120be27f0766cd59bd.docx>

Number of extracted messages and associated number of web users per data source.

URL: <http://asset.jmir.pub/assets/c626509d22ade257376b88d93af3e0ad.docx>

