

Original Paper

Exploring Psychological Trends in Populations With Chronic Obstructive Pulmonary Disease During COVID-19 and Beyond: Large-Scale Longitudinal Twitter Mining Study

Chunyan Zhang^{1,2*}, MS; Ting Wang^{3*}, PhD; Caixia Dong², MS; Duwei Dai², MS; Linyun Zhou², MPH; Zongfang Li², PhD; Songhua Xu², PhD

¹The High School Affiliated to Xi'an Jiaotong University, Xi'an, China

²Institute of Medical Artificial Intelligence, Second Affiliated Hospital of Xi'an Jiaotong University, Xi'an, China

³School of Electrical Engineering, Xi'an Jiaotong University, Xi'an, China

*these authors contributed equally

Corresponding Author:

Ting Wang, PhD

School of Electrical Engineering

Xi'an Jiaotong University

No. 28, Xianning West Road

Xi'an, 710049

China

Phone: 86 13289346632

Email: wang.ting@xjtu.edu.cn

Abstract

Background: Chronic obstructive pulmonary disease (COPD) ranks among the leading causes of global mortality, and COVID-19 has intensified its challenges. Beyond the evident physical effects, the long-term psychological effects of COVID-19 are not fully understood.

Objective: This study aims to unveil the long-term psychological trends and patterns in populations with COPD throughout the COVID-19 pandemic and beyond via large-scale Twitter mining.

Methods: A 2-stage deep learning framework was designed in this study. The first stage involved a data retrieval procedure to identify COPD and non-COPD users and to collect their daily tweets. In the second stage, a data mining procedure leveraged various deep learning algorithms to extract demographic characteristics, hashtags, topics, and sentiments from the collected tweets. Based on these data, multiple analytical methods, namely, odds ratio (OR), difference-in-difference, and emotion pattern methods, were used to examine the psychological effects.

Results: A cohort of 15,347 COPD users was identified from the data that we collected in the Twitter database, comprising over 2.5 billion tweets, spanning from January 2020 to June 2023. The attentiveness toward COPD was significantly affected by gender, age, and occupation; it was lower in females (OR 0.91, 95% CI 0.87-0.94; $P < .001$) than in males, higher in adults aged 40 years and older (OR 7.23, 95% CI 6.95-7.52; $P < .001$) than in those younger than 40 years, and higher in individuals with lower socioeconomic status (OR 1.66, 95% CI 1.60-1.72; $P < .001$) than in those with higher socioeconomic status. Across the study duration, COPD users showed decreasing concerns for COVID-19 and increasing health-related concerns. After the middle phase of COVID-19 (July 2021), a distinct decrease in sentiments among COPD users contrasted sharply with the upward trend among non-COPD users. Notably, in the post-COVID era (June 2023), COPD users showed reduced levels of joy and trust and increased levels of fear compared to their levels of joy and trust in the middle phase of COVID-19. Moreover, males, older adults, and individuals with lower socioeconomic status showed heightened fear compared to their counterparts.

Conclusions: Our data analysis results suggest that populations with COPD experienced heightened mental stress in the post-COVID era. This underscores the importance of developing tailored interventions and support systems that account for diverse population characteristics.

(*J Med Internet Res* 2025;27:e54543) doi: [10.2196/54543](https://doi.org/10.2196/54543)

KEYWORDS

COVID-19; chronic obstructive pulmonary disease (COPD); psychological trends; Twitter; data mining; deep learning

Introduction

On January 30, 2020, the World Health Organization declared the COVID-19 pandemic as a public health emergency of international concern [1]. Three years later, on May 5, 2023, given the encouraging trends observed in the COVID-19 statistics, the World Health Organization determined that COVID-19 no longer constituted a public health emergency of international concern [2]. However, this did not indicate that COVID-19 was no longer a global health threat [3]. The COVID-19 pandemic has had and continues to exert a profound and long-lasting impact on human lives worldwide, affecting public health, food systems, and global job markets [4,5]. In this situation, populations with chronic diseases, particularly chronic obstructive pulmonary disease (COPD), face great hardships [6,7] due to several reasons. First, patients with COPD are more likely to develop severe COVID-19 symptoms due to their respiratory illness [8]. Second, the overwhelming demands of managing COVID-19 may strain valuable health care resources, potentially affecting populations with COPD due to resource allocation [6]. Third, COVID-19 has greatly disrupted the daily lives of patients with COPD, leading to a higher risk of exacerbations and even mortalities. However, in contrast to the tangible physical consequences, the long-term effects of COVID-19 on the psychological conditions of patients with COPD have not drawn sufficient attention.

There are limited studies on the psychological effects of COVID-19 on patients with COPD (Table S1 in [Multimedia Appendix 1](#)) [9-14]. These studies have highlighted the high rates of psychological stress, including depression, anxiety, and fear, among patients with COPD under the threat of COVID-19. However, all these studies focused primarily on the early acute phase of COVID-19, thereby providing limited insight into the mental conditions of patients with COPD over the long term. Drawing parallels from the historical context of the SARS pandemic, wherein distinct chronological phases were discerned in the psychological conditions of SARS survivors [15], the psychological aftermath of COVID-19 survivors may exhibit similar enduring and evolving traits over the long term. Furthermore, surveys in the previous studies were conducted through traditional means, including telephone [10-12,14], video calls [14], interviews [13], and questionnaires [10,12], requiring substantial labor resources and physical contacts. This ultimately limited the amount of data that could be collected during the pandemic.

In recent years, social media data mining has been widely used in health-related domains—both physical and mental. In the physical health domain, for example, Pradeepa et al [16] proposed a methodology to identify various symptoms associated with stroke and preventive measures for stroke from

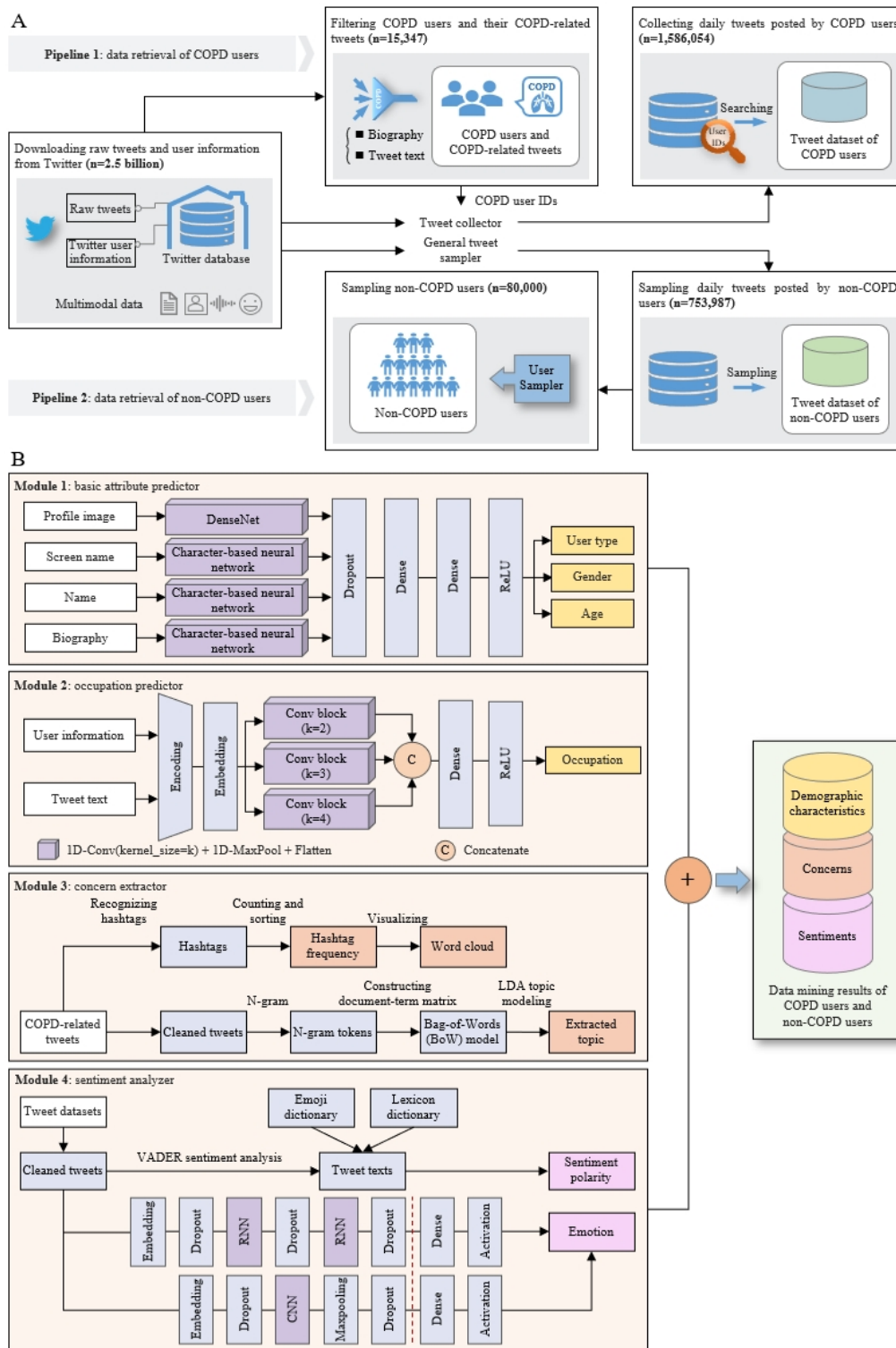
social media data. Fu et al [17] conducted a scoping review of methodologies for analyzing social media content in health care research and categorized health care areas into health practice, health services, and health education based on social media. In the mental health domain, for capturing the prevailing psychological conditions of the public during the COVID-19 pandemic, social media data mining has proven to be an effective solution. Compared with conventional survey approaches, social media data mining has significantly higher efficiency while minimizing infection risks by circumventing the need for physical contact. More importantly, the richness of social media data facilitates the discernment of long-term trends in emotional changes within a specific time period. For example, Lwin et al [18] used Twitter data to explore global trends in the 4 emotions—fear, anger, sadness, and joy—alongside their relative salience, spanning from January 28 to April 9, 2020. Similarly, through mining large-scale Twitter data, Zhang et al [19] conducted a cross-sectional study on web-based public sentiments in the early phase of the COVID-19 pandemic. However, these studies tended to either treat the populace as a monolithic entity, thus disregarding possible disparities among various population groups or only focus on general users while ignoring the existence of certain specific populations.

Considering the limitations of existing studies [9-19], this study aims to investigate the long-term psychological trends and patterns of populations with COPD with different characteristics during and after the COVID-19 pandemic. A new deep learning framework has been proposed to intelligently extract and analyze the characteristics of populations with COPD based on large-scale Twitter data.

Methods**Study Design**

This study was conducted following the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines [20] ([Multimedia Appendix 2](#)). In this study, raw Twitter data were collected with our self-designed program via the Twitter application programming interface [21]. These data are composed of multimodal information, including texts, emojis, images, and hashtags and are stored in a chronological order in our Twitter database. To date, over 2.5 billion tweets have been collected, spanning from January 2020 to June 2023. Based on these data, we designed a 2-stage framework, titled Deep-COPD, incorporating a series of deep learning algorithms to track COPD Twitter users (hereafter referred to as COPD users). The flowcharts for the data retrieval (stage 1) and data mining (stage 2) procedures are shown in [Figure 1](#).

Figure 1. Overall design of the Deep-COPD framework. (A) Data retrieval procedure. (B) Data mining procedure. CNN: convolutional neural network; Conv: convolutional; COPD: chronic obstructive pulmonary disease; LDA: latent Dirichlet allocation; RNN: recurrent neural network; ReLU: rectified linear unit; VADER: valence aware dictionary and emotional reasoner.



Data Retrieval

Figure 1A shows the procedure of data retrieval, which involves 2 pipelines: pipeline 1, retrieving the data of COPD users, including user information, COPD-related tweets, and their daily tweets; and pipeline 2, retrieving the data of non-COPD users, including user information and their daily tweets. In pipeline 1, we used a filter pattern implemented as a regular

expression composed of a set of COPD keywords to identify COPD users and tweets by searching 2 sources: user biographies and tweet texts (see Multimedia Appendix 1). The specific construction process of this regular expression can be found in the build_match_word_regex() function of the filter_copd.py file (lines 17-35) in the stage 1 folder of our GitHub repository. We collected the daily tweets posted by these COPD users

through their user IDs. To limit the size of the data retrieved in pipeline 2, we sampled the tweets posted by non-COPD users at a rate of 0.05%. Through data retrieval, we obtained raw data that were used for the sentiment and psychological analysis. Moreover, we extracted data representing COPD and non-COPD characteristics from the daily tweets, enabling targeted research on the psychological evolution patterns associated with these 2 distinct groups. Specifically, this approach enables a more focused investigation into how the psychological trends may differ between users with COPD and those without COPD, thereby providing valuable insights into the impact of the condition on mental health over time. By analyzing these 2 groups separately, we aim to uncover specific psychological patterns and trends that could inform tailored interventions and support strategies for each population.

Data Mining

The data mining procedure aims to extract distinct information from multimodal Twitter data by using a variety of deep learning algorithms, including natural language processing, image processing, and symbol processing (Figure 1B). It consists of the following 4 parallel intelligent modules. Module 1 and module 2 are used to extract demographic characteristics, including user type, gender, age, and occupation, from both COPD and non-COPD users. Module 3 extracts concern information from COPD-related tweets, while module 4 extracts sentiment information from the daily tweet datasets of both COPD and non-COPD users.

Module 1: Basic Attribute Predictor

Module 1 is an attribute predictor that predicts the probabilities of 3 demographic attributes, that is, user type, gender, and age, through the analysis of profile images, screen names, display names, and biographies. It was implemented with an open-source package of the M3 (multimodal, multilingual, and multiattribute) model [22], a multimodal deep neural system trained on a massive dataset. It is mainly composed of a dense convolutional network (DenseNet) [23] for image processing, 3 character-based neural networks for text processing, and 2 fully connected dense layers for result classification. The training and validating M3 model uses 5 distinct datasets: a large dataset of Twitter profiles whose gender and age are heuristically identified, a curated dataset of Twitter accounts belonging to organizations, an image dataset of faces from IMDb (internet movie database) and Wikipedia, a massive unlabeled set of users, and a crowdsourced dataset for all 3 attributes. For more detailed information, readers can refer to the original literature [22]. This module was tested with a subset of our tweet dataset, achieving the following benchmark performance: 99.07% accuracy for user type, 95.88% accuracy for gender, and 77.65% accuracy for age attributes.

Module 2: Occupation Predictor

Module 2 is a self-designed deep learning model that derives the 2-class occupations of Twitter users from their profile information and tweet texts. It employs a word-based convolutional neural network that consists of an encoding layer, an embedding layer, 3 parallel convolutional blocks with different kernel sizes, and a dense layer. The training and

validation data are obtained from a publicly available dataset [24] with 5191 Twitter users annotated with 9 occupation categories (OCs), which are defined according to the standard occupation classification [25]. In our study, the 9 OCs were merged into 2 OCs, abbreviated as OC1 and OC2 (see Table S4 in Multimedia Appendix 1). Individuals in OC1 are those of relatively higher socioeconomic status (such as managers, directors, and senior officials), and individuals in OC2 are those of relatively lower socioeconomic status (such as process, plant, and machine operatives). Using a 10-fold cross-validation approach, this module achieved an accuracy of 93.15% on the 2-class occupational dataset.

Module 3: Concern Extractor

Module 3 is a concern extractor that captures hashtags and latent Dirichlet allocation (LDA) [26] topics. Hashtags were identified in the tweet texts using a regular expression ($\#[a-zA-Z0-9\-_]{1,}$), and then the frequencies of these hashtags were calculated. LDA topic modeling was employed to identify the main topics expressed in the tweets. The original tweets underwent preprocessing, n-gram processing, and document-term matrix constructing, yielding a bag-of-words, which was fed into LDA models. The coherence score calculated by Gensim CoherenceMode [27] was used to evaluate the quality of the extracted topics. To pinpoint the optimal number of LDA topics, we performed a grid search within a range of 2-20 with a step of 1 for each year. The selection principle is to choose the number that makes the coherence score of the extracted topics reach the highest value. Through searching, the optimal number of topics and the coherence score for each year were obtained, as shown in Table S5 in Multimedia Appendix 1.

Module 4: Sentiment Analyzer

Module 4 is composed of a sentiment analyzer and an emotion detector to perceive sentiments from 2 different perspectives. The sentiment analyzer utilizes a widely used open-source tool known as the valence aware dictionary and emotional reasoner (VADER) [28]. VADER requires no training data, but it is constructed from a generalizable, valence-based, human-curated gold standard sentiment lexicon. It is highly effective in analyzing social media sentiments, achieving an F_1 classification accuracy of 96%, outperforming human raters of 84% at classifying the sentiment of tweets. The emotion detector was built upon an open-source algorithm for emotion recognition on Twitter, which employs a character-based trained recurrent neural network [29]. The training and validation dataset is a set of tweets containing emotional hashtags and satisfying filtering requirements that searched from 73 billion tweets. Specifically, this study adopted the well-known Plutchik's emotion wheel [30] model, which consists of 8 basic emotions: joy, sadness, fear, anger, trust, disgust, surprise, and anticipation. This emotion detector achieved a macro F_1 -score of 0.6950 for Plutchik's emotion model [29].

Specifically, the deep learning algorithms chosen in this study were selected for their strengths in handling various data types. DenseNet was employed for image processing, while character-based and word-based neural networks were used for the text classification of demographic attributes and occupations. LDA's topic modeling capability made it suitable for extracting

concerns from large text datasets. VADER’s high accuracy ensured robust social media sentiment detection. Additionally, a character-based recurrent neural network enabled the emotion detector to effectively capture and classify complex emotional states.

The data mining procedure extracts valuable information from multimodal Twitter data by using advanced deep learning algorithms. This process involves 4 parallel intelligent modules, of which the outputs are as follows: module 1 (basic attribute predictor) predicts demographic attributes such as user type, gender, and age by analyzing user profile images, names, and biographies; module 2 (occupation predictor) classifies Twitter users’ occupations into 2 groups based on their profiles and tweet content; module 3 (concern extractor) identifies user concerns by extracting hashtags and topics by using LDA; and module 4 (sentiment analyzer) analyzes sentiments from tweets by using VADER for general sentiments and a character-based recurrent neural network for emotion detection. Together, these modules provide a comprehensive analysis of the demographic, psychological, and emotional factors affecting both COPD and non-COPD Twitter users.

The outputs from data retrieval and data mining are crucial for assessing the psychological trends in populations with COPD, particularly the changes in sentiments, prevalent emotions, and evolving concerns (eg, health-related topics) during and after COVID-19. Through data retrieval, raw Twitter data are

collected, allowing for data extraction of COPD and non-COPD groups, which enables focused analysis of psychological patterns within the target population. The data mining process further analyzes these data, extracting key demographic attributes, concerns, and sentiments by using deep learning algorithms. The concern extraction identifies health-related topics, while the sentiment and emotion analyzer tracks shifts in emotional states and sentiments, providing critical insights into the psychological evolution of populations with COPD in response to the pandemic. This comprehensive data processing directly supports the study’s goal of understanding the psychological trends of COVID-19 in populations with COPD.

Data Analysis Tools

Based on the data mining results from our Deep-COPD framework, we conducted attentiveness analysis, concern analysis, and sentiment analysis. In these analyses, multiple statistical tools were used, namely, odds ratio (OR), difference-in-difference (DiD), and emotion pattern method.

Statistical Measure OR

OR is a statistical measure widely used in medical literature to evaluate the relevance between an event (such as lung cancer) and a particular exposure (such as smoking) [31]. In this study, the event refers to COPD, and the exposure factors include gender, age, and occupation. The ORs are used to quantitatively assess the levels of attentiveness toward COPD under these different exposure factors (Table 1).

Table 1. A 2x2 table for the chronic obstructive pulmonary disease event and the gender exposure factor (a, b, c, and d represent the number of users).

	Chronic obstructive pulmonary disease event	No event of chronic obstructive pulmonary disease
Female exposure	a	b
Male exposure	c	d

Taking gender as the exposure factor as an example, the 2x2 table shown in Table 1 is used to calculate the attentiveness OR toward a COPD event in females compared to males. The formula is as follows: $OR = (a/b)/(c/d) = ad/bc$

An OR of 1 indicates that the gender factor does not affect the level of attentiveness toward COPD. When $OR > 1$, it suggests that the attentiveness toward COPD is higher in females than in males. Conversely, when $OR < 1$, it indicates that the attentiveness toward COPD is lower in females than in males. The confidence interval in OR provides a range of values within which the true OR is likely to fall, offering an indication of the precision and reliability of the estimate. As a typical value, 95% CI is commonly used, indicating that there is a 95% probability that the true OR lies within the specified range [31]. The following formula is used for 95% CI:

$$Upper\ 95\% \ CI = e^{[\ln(OR) + 1.96\sqrt{(1/a) + (1/b) + (1/c) + (1/d)}]}$$

$$Lower\ 95\% \ CI = e^{[\ln(OR) - 1.96\sqrt{(1/a) + (1/b) + (1/c) + (1/d)}]}$$

In addition, the P values were obtained using the chi-square test to evaluate the significance levels of the differences between groups under different exposure factors. A lower P value indicates a more statistically significant difference between the groups [32].

DiD Model

The DiD model is typically used to estimate the effect of a specific intervention or treatment by comparing the changes in outcomes over time between a population that is enrolled in a program (the treated group) and a population that is not (the control group) [33,34]. In this study, we took COPD users as the treated group and non-COPD users as the control group and adopted the DiD model to estimate the relative sentiment trends of COPD users. The formula of the DiD model is as follows:

$$Y_{it} = \alpha + \sum_{k=1}^{K-1} \beta_k(G_i * D_t) + \gamma G_i + \sum_{k=1}^{K-1} \delta_k D_t + \varepsilon_{it}$$

where Y_{it} is the dependent variable representing the sentiment status for group G_i and time t ; G_i is a dummy variable for the treated group (G_0) or control group (G_1); D_t is a dummy variable for time t ; α , β_k , γ , δ_k are the regression parameters; K is the total time period ($k \in [1, K]$); and ε_{it} is an error item.

The coefficients of interest are those of the interaction term, β_k , which can reflect the relative sentiment trends of COPD users compared to those of non-COPD users. A value of $\beta_k = 0$ signifies that both groups have parallel sentiment trends, whereas

a negative value indicates the extent to which the sentiment trend of the COPD group declines compared to that of the non-COPD group and vice versa.

Emotion Pattern

We proposed an emotion pattern method to calculate the values of the sentiment components of Plutchik’s 8 emotions. The emotion pattern of population p is composed of the sentiment component scores of Plutchik’s 8 emotions, which is expressed as follows:

$$S(p) = \sum_e S(p)|_e$$

$$S(p)|_e = f(x_1, \dots, x_{m_e})|_e * \frac{m_e}{N}$$

where $e \in [\text{joy, trust, surprise, anticipation, sadness, disgust, anger, fear}]$, $S(p)|_e$ is the sentiment component of population p on emotion e ,

$$f(x_1, \dots, x_{m_e})|_e$$

is the calculated sentiment scores of tweets

$$(x_1, \dots, x_{m_e})$$

expressing emotion e via the VADER method f , m_e is the number of tweets expressing emotion e , and N is the total number of tweets from population p .

Ethical Considerations

This analysis did not receive approval from an institutional research board. Analysis of large bodies of text written in social media such as the Twitter social network is generally not considered human subjects research [35], and therefore, ethics approval was not sought. We confirm that all data used in this study have been deidentified, and no individual user is identifiable in any of the images included in the manuscript or supplementary data.

Results

Demographic Characteristics of the Participants

In total, 15,347 COPD users were identified as participants in this study. From this cohort, 1,586,054 daily tweets were retrieved, spanning from January 2020 to June 2023. For comparison, 80,000 non-COPD users and 753,987 general tweets were also selected in the same period. Table 2 presents the statistical data on the demographic characteristics of these two groups. To obtain a comprehensive understanding of the psychological conditions of populations with COPD during COVID-19 and beyond, we analyzed the data mined by the Deep-COPD framework along 3 dimensions: attentiveness analysis, concern analysis, and sentiment analysis.

Table 2. Statistical data of chronic obstructive pulmonary disease and non-chronic obstructive pulmonary disease users.

Characteristics	COPD ^a users (n=15,347), n (%)	Non-COPD users (n=80,000), n (%)
User type		
Individual	13,375 (87.15)	75,468 (94.34)
Organization	1972 (12.85)	4532 (5.67)
Gender		
Male	7854 (58.72)	42,509 (56.33)
Female	5521 (41.28)	32,959 (43.67)
Age (years)		
≤18	1179 (8.81)	29,810 (39.50)
19-29	2373 (17.74)	25,086 (33.24)
30-39	2091 (15.63)	8551 (11.33)
≥40	7732 (57.81)	12,021 (15.93)
Occupation category		
OC1 ^b	7356 (55)	50,566 (67)
OC2 ^c	6019 (45)	24,902 (33)

^aCOPD: chronic obstructive pulmonary disease.

^bOC1: occupation category 1 (relatively higher socioeconomic status such as managers, directors, and senior officials).

^cOC2: occupation category 2 (relatively lower socioeconomic status such as process, plant, and machine operatives).

Attentiveness Analysis

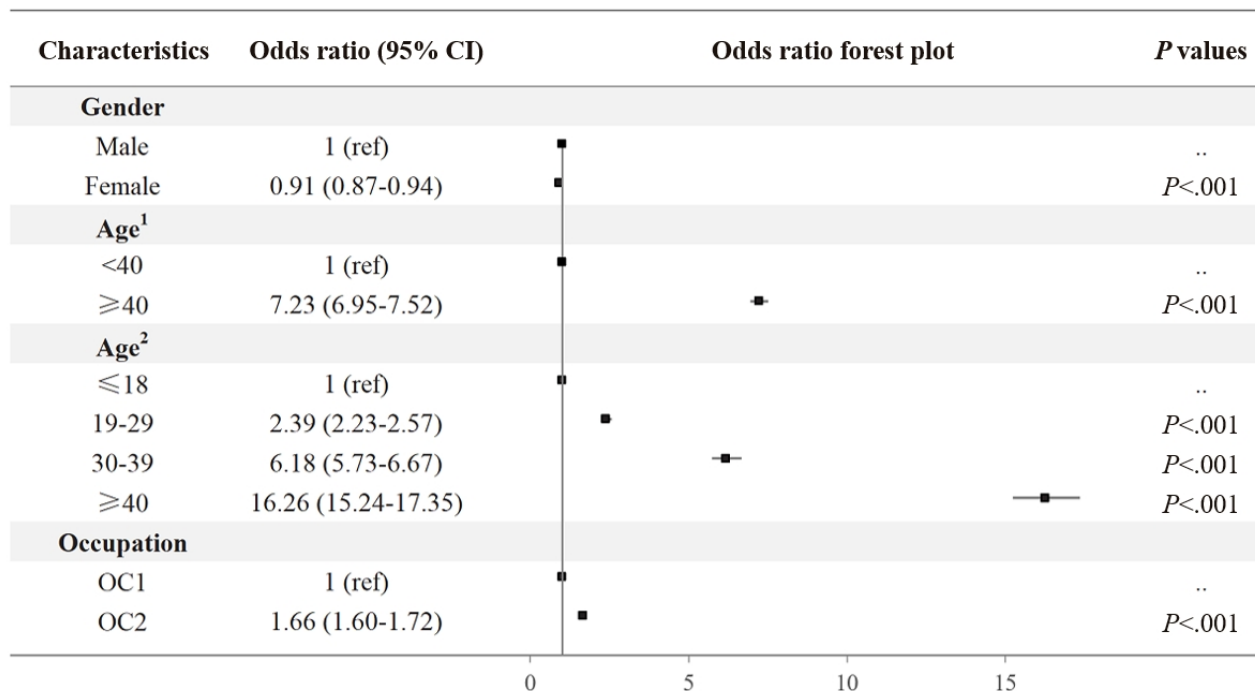
Based on the statistics in different demographic groups (Table 2), a comparative study was performed between COPD and

non-COPD users. As mentioned in the data analysis tools, we employed ORs to quantitatively assess the levels of attentiveness toward COPD in different groups (Figure 2). The ORs were calculated between COPD and non-COPD users under different

exposure factors, including gender, age, and occupation. Results showed that the attentiveness toward COPD in females (OR 0.91, 95% CI 0.87-0.94; $P < .001$) was lower than that of males. The cohort aged 40 years and older exhibited a high OR of 7.23 (95% CI 6.95-7.52; $P < .001$), which was significantly higher than those younger than 40 years, and it showed a significant increase with age across 4 age groups. As for occupation, users

in OC2 occupations exhibited a notably higher OR of 1.66 (95% CI 1.60-1.72; $P < .001$) compared to those in OC1 occupations. These results suggested that gender, age, and occupation are 3 important factors influencing attentiveness. Specifically, older male users with lower socioeconomic status (OC2) were the most concerning group.

Figure 2. Attentiveness odds ratios toward chronic obstructive pulmonary disease in different demographic groups. Age 1: 2 age groups; age 2: 4 age groups; OC: occupation category; ref: reference.

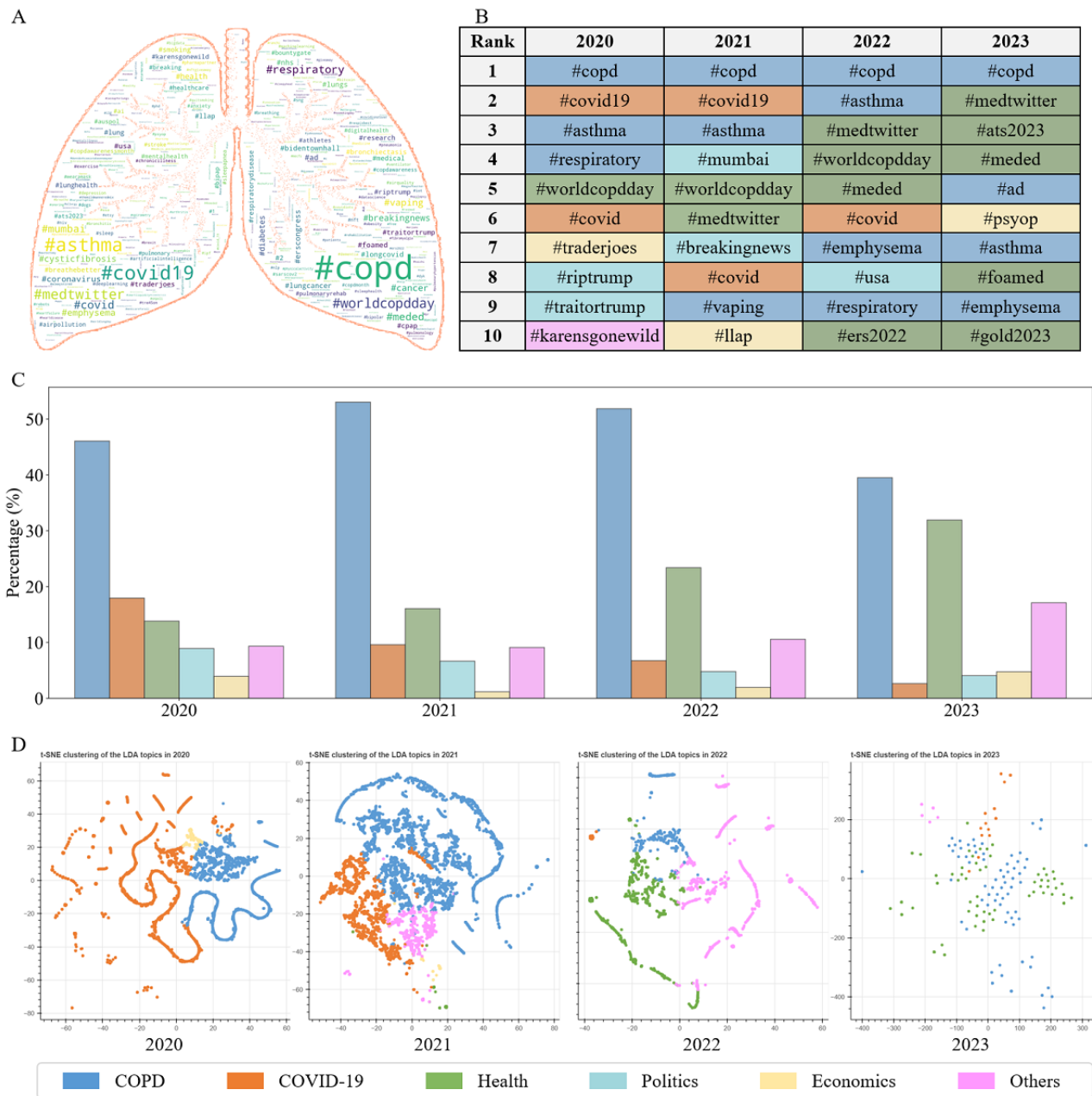


Concern Analysis

The word cloud of the top 500 hashtags in the tweets shows that #COPD and #COVID-19 were the most frequently used hashtags throughout the 4 years (Figure 3A). Among the top 10 hashtags in each year, #COPD consistently remained the primary concern throughout the 4-year period (Figure 3B). In 2020 and 2021, #COVID-19 ranked second but later gave way to health-related hashtags such as #medtwitter and #meded. To study the trends in the evolving concerns of COPD users, we categorized the 100 most frequently used hashtags from each year into 6 categories: COPD, COVID-19, health, politics,

economics, and others. It is evident that the COPD category maintained the highest prevalence among COPD users, peaking at over 50% in 2021 (Figure 3C). The COVID-19 category held the second position in 2020, accounting for nearly 20%, but substantially decreased in 2021. On the other hand, the health category experienced continuous growth over the 4-year period, surpassing 30% in 2023 (Figure 3C). Figure 3D shows that the concerns from the perspective of LDA topics (see detailed topics in Table S6 in Multimedia Appendix 1) over the years follow the same trends as the hashtags. These statistical results highlight that, among COPD users, personal health became an even more significant concern as the pandemic receded.

Figure 3. Results of the concern analysis. (A) Word cloud of the top 500 hashtags. (B) Top 10 most used hashtags in 4 years. (C) The percentages of the 6 categories of hashtags in 4 years. (D) Visualization of the latent Dirichlet allocation topics via t-distributed stochastic neighbor embedding clustering. COPD: chronic obstructive pulmonary disease; LDA: latent Dirichlet allocation; tSNE: t-distributed stochastic neighbor embedding clustering.



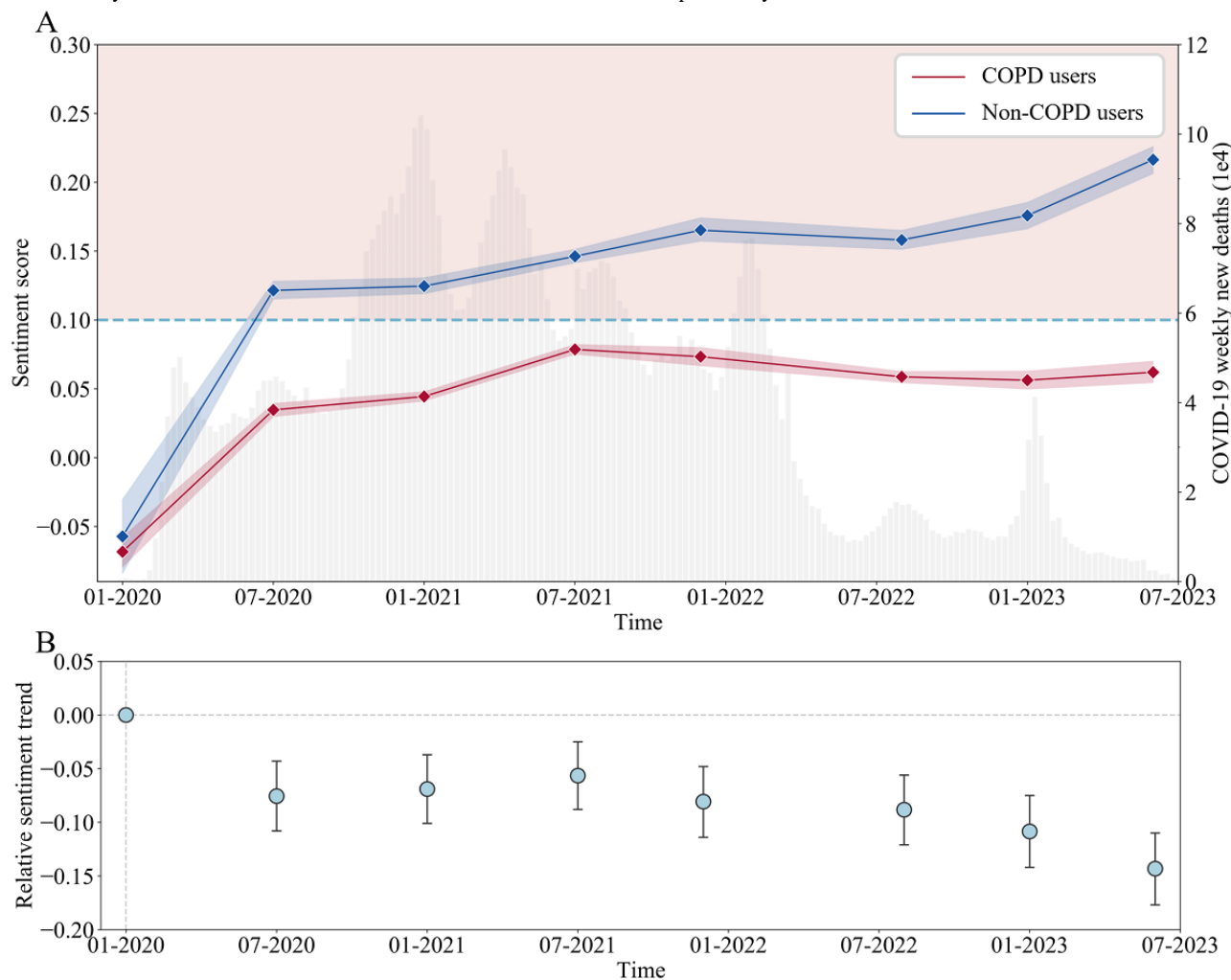
Sentiment Analysis

To gain an insight into the long-term trends and disparities of sentiments among different groups, we conducted a sentiment analysis spanning from January 2020 to June 2023. The sentiment score, calculated by VADER, was used to quantify the sentiment intensity of a tweet. This score ranges from -1 (most negative) to 1 (most positive), with [-1, -0.1] indicating negative, [-0.1, 0.1] indicating neutral, and (0.1, 1] indicating positive [36].

Sentiment Trends of the Populations With COPD

The sentiment scores from both COPD and non-COPD users are shown in Figure 4A, revealing a consistent pattern of lower sentiments in the long term among tweets posted by COPD users compared to those posted by non-COPD users. At the outbreak of COVID-19 in January 2020, the sentiment scores of both the population groups were below -0.05. Between July 2020 and June 2023, the sentiment scores of COPD users remained below 0.1, whereas those of non-COPD users exceeded 0.1. Notably, as the pandemic gradually subsided in 2023, the sentiment scores of non-COPD users increased progressively. In contrast, the sentiment scores of the COPD group still stagnated at a low level.

Figure 4. Results of the sentiment trend analysis. (A) Mean sentiment scores with 95% CIs of daily tweets from COPD and non-COPD users over time. The horizontal dashed line (sentiment score = 0.1) is the dividing line between positive (pink area) and neutral (white area). The background bar chart shows the number of COVID-19 weekly deaths. (B) Relative sentiment trends of COPD users, as reflected by the estimated coefficients with 95% CIs calculated by the difference-in-difference model. COPD: chronic obstructive pulmonary disease.

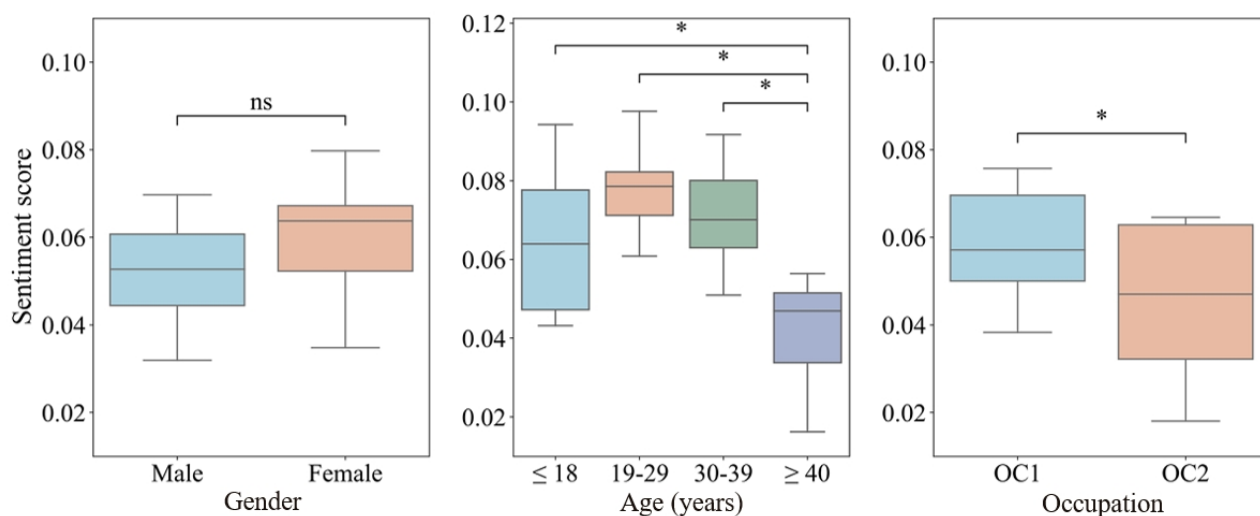


To accurately understand the relative sentiment trends in the COPD group in contrast with those in the non-COPD group, we employed a DiD model [34] to calculate the estimated coefficients between them. The results calculated by the DiD model can be found in Table S7 in [Multimedia Appendix 1](#). [Figure 4B](#) shows that these estimated coefficients remained constantly negative, except for the initial time point, which was set to 0. These coefficients gradually approached 0 until July 2021 but significantly decreased afterward, reaching the lowest point in June 2023. This indicated a continuing relatively downward sentiment trend in the COPD group compared to that in the non-COPD group, and this trend was more pronounced after July 2021.

Sentiment Disparities Among Different Population Groups

As shown in [Figure 5](#), sentiment scores differ among demographic groups. The average sentiment score of females was higher than that of males; yet, the difference was statistically insignificant ($P=.08$). Among age groups, those aged 40 years and older exhibited much lower sentiments than all the other 3 younger groups. As for the occupation, COPD users with OC2 occupations received lower sentiment scores than those with OC1 occupations ($P=.03$). In short, male COPD users aged 40 years and older with lower socioeconomic status (OC2) had the lowest sentiment in their daily tweets.

Figure 5. Sentiment comparisons of COPD users among population groups. COPD: chronic obstructive pulmonary disease. OC: occupation category; ns: not significant. * $P < .05$.

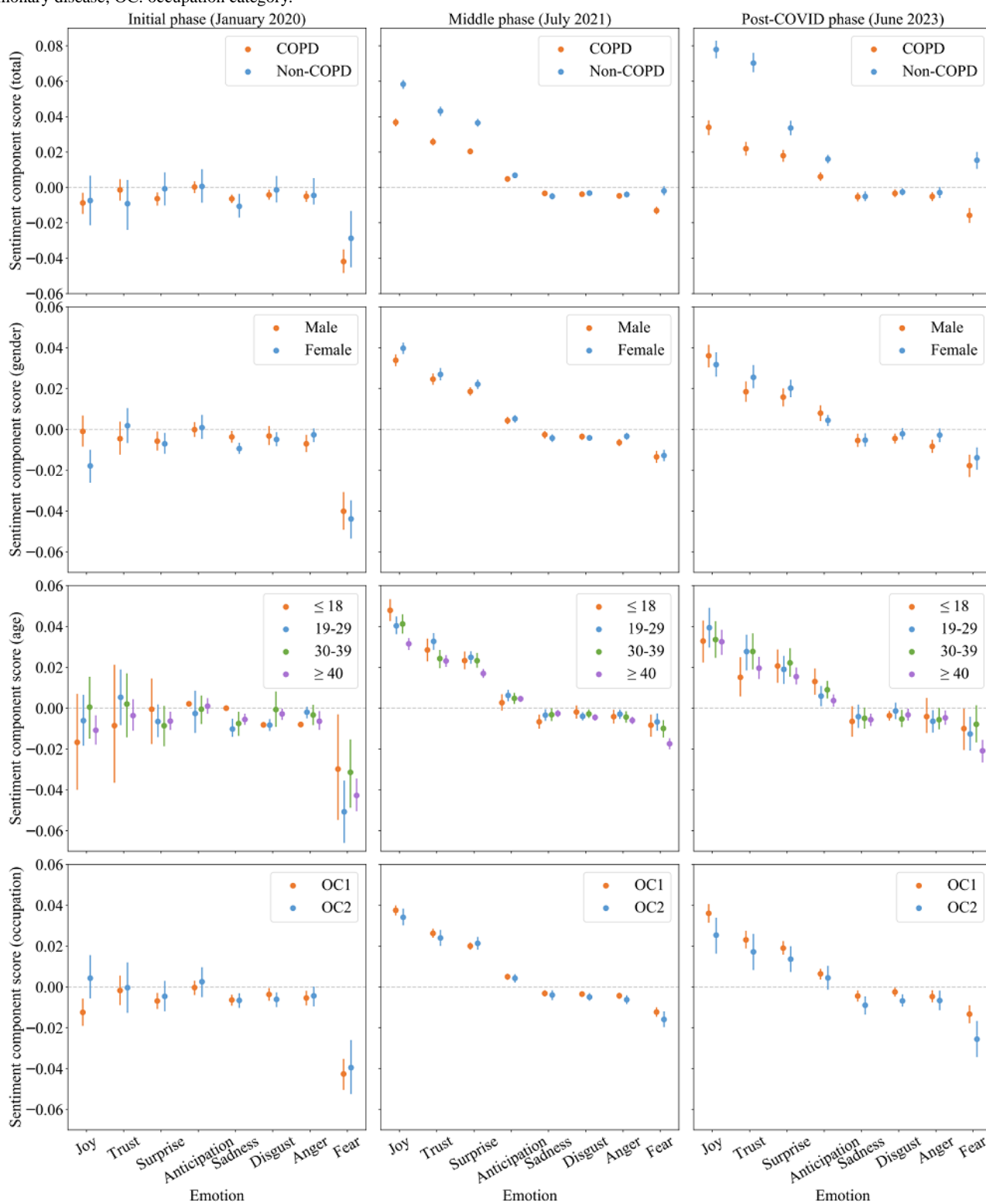


To identify emotion patterns in the sentiments of various population groups, we proposed a new method to calculate the values of the sentiment components of Plutchik's 8 emotions, denoted as sentiment component scores. Detailed descriptions of the emotion pattern method can be found in the Data Analysis Tools. We selected 3 representative timepoints corresponding to the different phases of COVID-19: the initial phase in January 2020, when the pandemic was declared public health emergency of international concern by the World Health Organization; the middle phase in July 2021, when the sentiments of COPD users reached the highest point; and the post-COVID phase in June 2023, after the pandemic was declared over. In the initial phase, fear was the predominant negative emotion among both COPD and non-COPD users. In the middle phase, the sentiment component scores of positive emotions such as joy, trust, and surprise of non-COPD users were significantly higher than those

of COPD users. In the post-COVID era, the sentiment component scores of joy and trust of non-COPD users exceeded those of COPD users by far, while the sentiment component score of fear of COPD users became even more negative.

As shown in Figure 6, we studied the sentiment component scores of different demographic groups across the 3 COVID-19 phases. The results indicated that females exhibited significantly lower scores of joy than males in the initial phase (January 2020), whereas no obvious difference was observed in the middle phase (July 2021) and post-COVID phase (June 2023). Regarding age, those aged 40 years and older displayed lower sentiment component scores in most emotions, especially fear, compared to those younger than 40 years. As for occupations, those with lower socioeconomic status (OC2) demonstrated lower sentiment component scores than those with higher socioeconomic status (OC1).

Figure 6. Sentiment component scores of Plutchik’s 8 emotions for different population groups across 3 COVID-19 phases. COPD: chronic obstructive pulmonary disease; OC: occupation category.



Discussion

Principal Findings

In this study, we designed a deep learning framework, Deep-COPD, incorporating a series of data retrieval and deep learning methods to investigate the psychological conditions of populations with COPD by using large-scale Twitter data collected during COVID-19 and beyond. First, we computed

the attentiveness toward COPD among populations with different demographic features. Subsequently, we identified the popular concerns by hashtags and LDA topics in tweets posted by COPD users in 4 years. Finally, we examined the sentiment trends and prevalent emotions of populations with COPD in contrast to those of populations without COPD.

To the best of our knowledge, this study is the first research effort dedicated to exploring the psychological trends of

populations with COPD during and after COVID-19. Unlike in the existing investigations into the psychological impacts of COVID-19 on patients with COPD via conventional survey approaches, where the participant cohorts comprised only a few hundred individuals [10-14], our study intelligently recognized a cohort of 15,347 COPD users as participants. Moreover, by leveraging large-scale Twitter data spanning from January 2020 to June 2023, our study offers invaluable insights into the long-term psychological trends among populations with COPD in the pandemic.

By executing a comparative study between COPD and non-COPD users, this study shows that gender, age, and occupation affected the attentiveness toward COPD. The attentiveness toward COPD was lower in females than in males, significantly higher in adults aged 40 years and older than those younger than 40 years, and significantly higher in individuals of lower socioeconomic status than those of higher socioeconomic status. These interesting findings are aligned with the scientific statistics on the prevalence of COPD. In terms of gender, the prevalence of COPD is reported to be about 5% higher in males than in females [37]. As for age, the prevalence of COPD is appreciably higher in those aged 40 years and older compared to that of COPD in those younger than 40 years, with the incidence progressively increasing with age [38]. For occupation, individuals with lower socioeconomic status (OC2) were more likely to have poorer COPD outcomes than those with higher socioeconomic status (OC1) [8].

Through concern analysis, a notable decline in the popularity of COVID-related concerns among COPD users was observed after 2021. This phenomenon can be attributed to the fact that the SARS-CoV-2 virus undergoes constant mutations. With the emergence of the Omicron variant in November 2021, there has been a decline in the death rate [39]. As more knowledge of the virus has been gained, the associated attention has subsided. Concurrently, a steady rise in health-related concerns among COPD users was noted throughout the course of the pandemic. This trend can be attributed to the higher risk of patients with COPD to more severe diseases under COVID-19 (63%) compared to 33.4% among patients without COPD [8].

By tracking the sentiment trends from January 2020 to June 2023, we found that the sentiments of COPD users were much lower than those of non-COPD users across our study period. Remarkably, through DiD analysis, we noticed an increasingly significant downward trend in sentiments among COPD users compared to non-COPD users after the middle phase of COVID-19 (July 2021). This is an important warning that deserves our close attention. By analyzing the emotion patterns of 3 representative timepoints, this study further demonstrates that negative emotions, including fear, sadness, disgust, and anger, all contributed to the low sentiments of COPD users. Notably, in the post-COVID era (June 2023), COPD users exhibited decreased levels of joy and trust and increased fear compared to their levels of joy and trust during the middle phase of COVID-19. In particular, males, older adults, and individuals with lower socioeconomic status showed heightened fear compared to their counterparts. These results indicated that the populations with COPD experience a much longer impact from COVID-19 than the healthy population.

Last but not the least, similar psychological challenges have also been reported in studies on other chronic conditions under COVID [40-42]. Among all the chronic disease groups studied, including COPD, diabetes, and cardiovascular disease, the patients experienced long-term emotional distress during COVID-19, reinforcing the importance of psychological interventions tailored to different demographic and health characteristics.

Implications

Considering the implications for practical scenarios, it is clear that patients with COPD need targeted mental intervention and support. The findings of this work reveal a notable surge of fear among COPD users in the post-COVID era, stemming from their health-related concerns. In view of this, we suggest improving the situation in 2 ways to address the emotional challenges faced by patients with COPD. First, offline health care organizations are encouraged to actively offer easy-to-use mental health screenings, interventions, and self-management measures tailored to different population groups in addition to COPD treatments. For example, tools such as the Hospital Anxiety Depression Scale [43] or the Depression Anxiety Stress Scale [44] can be employed to identify patients with COPD who may have mental health concerns, thereby providing suitable guidance and periodic monitoring. Second, by utilizing the potential of social media platforms, web-based health care organizations and influential users can disseminate accessible psychological content, including videos, games, and helpful resources, to raise awareness of mental health and promote positive lifestyles. This approach involves targeted delivery to COPD users according to the distinct demographic and psychological patterns of different groups. By synergizing these offline and web-based approaches, health care organizations and professionals may effectively tailor their intervention measures to alleviate the psychological toll on populations with COPD.

Limitations

This study has several limitations. First, the identification of COPD users was solely based on self-reported data from Twitter, which may not fully capture all individuals with COPD. Second, this study relies on the VADER tool for sentiment analysis, which, despite its high accuracy, has limitations due to its static lexicon. This restricts its ability to adapt to evolving language trends or capture domain-specific terminology, potentially impacting the accuracy of sentiment classification. Finally, this study is limited to English-language tweets, potentially excluding non-English-speaking populations with COPD and missing cultural variations in psychological responses. These factors should be considered when interpreting the generalizability of our findings. Future research could address these limitations by incorporating data from multiple social media platforms, using more comprehensive methods to verify user health conditions, analyzing multilingual content to capture a broader and more diverse population with COPD, and exploring adaptive or domain-specific sentiment analysis tools to enhance sentiment detection accuracy.

Conclusions

This study proposes a novel approach to tracking the psychological trends of COPD users via large-scale Twitter mining. This study reveals that although COVID-19 has been declared to be ended, COPD users remain emotionally distressed, with much higher levels of fear than the healthy population. Therefore, strategies for the prevention and treatment of psychological problems in patients with COPD should be

tailored to the characteristics of different population groups. The findings in this study offer valuable guidance to health care providers, policymakers, and researchers on the necessary measures to address the psychological needs of patients with COPD in the post-COVID era. The method implemented through our deep learning approach can be easily repurposed to study the psychological effects of other social events from a long-term perspective.

Data Availability

The source code for the deep learning framework is available at [45]. In compliance with Twitter's agreements and policy, we can only share data about Twitter user IDs and tweet IDs for academic purposes. These data are available through data access agreements from the corresponding authors.

Authors' Contributions

Conceptualization: CZ (lead), TW (equal)

Data curation: CZ (lead), CD (supporting), DD (supporting), SX (supporting)

Formal analysis: CZ (lead), TW (equal), CD (supporting), DD (supporting), LZ (supporting), ZL (supporting), SX (supporting)

Methodology: CZ (lead), TW (equal)

Investigation: CZ (lead), TW (equal), CD (supporting), DD (supporting), LZ (supporting), ZL (supporting)

Software: CZ

Writing – original draft: CZ (lead), TW (equal)

Writing – review & editing: CZ (lead), TW (equal)

Conflicts of Interest

None declared.

Multimedia Appendix 1

Supplementary information on the survey of related studies, details of some methods used in this study, and the results of some methods.

[DOCX File, 33 KB-Multimedia Appendix 1]

Multimedia Appendix 2

STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines.

[DOCX File, 33 KB-Multimedia Appendix 2]

References

1. COVID-19 Public Health Emergency of International Concern (PHEIC). World Health Organization. Feb 12, 2020. URL: <https://www.who.int/publications/m/item/covid-19-public-health-emergency-of-international-concern-%28pheic%29-global-research-and-innovation-forum> [accessed 2023-10-01]
2. Statement on the fifteenth meeting of the IHR (2005) Emergency Committee on the COVID-19 pandemic. World Health Organization. May 5, 2023. URL: <https://tinyurl.com/bdj6rm46> [accessed 2023-10-01]
3. Wise J. COVID-19: WHO declares end of global health emergency. *BMJ*. May 09, 2023;381:1041. [doi: [10.1136/bmj.p1041](https://doi.org/10.1136/bmj.p1041)] [Medline: [37160309](https://pubmed.ncbi.nlm.nih.gov/37160309/)]
4. Impact of COVID-19 on people's livelihoods, their health and our food systems. World Health Organization. Oct 13, 2020. URL: <https://www.who.int/news/item/13-10-2020-impact-of-covid-19-on-people-s-livelihoods-their-health-and-our-food-systems> [accessed 2023-10-01]
5. Brodeur A, Gray D, Islam A, Bhuiyan S. A literature review of the economics of COVID-19. *J Econ Surv*. Sep 2021;35(4):1007-1044. [FREE Full text] [doi: [10.1111/joes.12423](https://doi.org/10.1111/joes.12423)] [Medline: [34230772](https://pubmed.ncbi.nlm.nih.gov/34230772/)]
6. Press VG, Gershon AS, Scirba FC, Blagev DP. Concerns about coronavirus disease-related collateral damage for patients with COPD. *Chest*. Sep 2020;158(3):866-868. [FREE Full text] [doi: [10.1016/j.chest.2020.05.549](https://doi.org/10.1016/j.chest.2020.05.549)] [Medline: [32473947](https://pubmed.ncbi.nlm.nih.gov/32473947/)]
7. Leung JM, Niikura M, Yang CWT, Sin DD. COVID-19 and COPD. *Eur Respir J*. Aug 2020;56(2):2002108. [FREE Full text] [doi: [10.1183/13993003.02108-2020](https://doi.org/10.1183/13993003.02108-2020)] [Medline: [32817205](https://pubmed.ncbi.nlm.nih.gov/32817205/)]

8. Alqahtani JS, Oyelade T, Aldhahir AM, Alghamdi SM, Almeahadi M, Alqahtani AS, et al. Prevalence, severity and mortality associated with COPD and smoking in patients with COVID-19: a rapid systematic review and meta-analysis. *PLoS One*. 2020;15(5):e0233147. [FREE Full text] [doi: [10.1371/journal.pone.0233147](https://doi.org/10.1371/journal.pone.0233147)] [Medline: [32392262](https://pubmed.ncbi.nlm.nih.gov/32392262/)]
9. Yohannes AM. COPD patients in a COVID-19 society: depression and anxiety. *Expert Rev Respir Med*. Jan 2021;15(1):5-7. [doi: [10.1080/17476348.2020.1787835](https://doi.org/10.1080/17476348.2020.1787835)] [Medline: [32578464](https://pubmed.ncbi.nlm.nih.gov/32578464/)]
10. Pedrozo-Pupo JC, Campo-Arias A. Depression, perceived stress related to COVID, post-traumatic stress, and insomnia among asthma and COPD patients during the COVID-19 pandemic. *Chron Respir Dis*. 2020;17:1479973120962800. [FREE Full text] [doi: [10.1177/1479973120962800](https://doi.org/10.1177/1479973120962800)] [Medline: [33000648](https://pubmed.ncbi.nlm.nih.gov/33000648/)]
11. Daynes E, Gerlis C, Briggs-Price S, Jones P, Singh SJ. COPD assessment test for the evaluation of COVID-19 symptoms. *Thorax*. Feb 2021;76(2):185-187. [doi: [10.1136/thoraxjnl-2020-215916](https://doi.org/10.1136/thoraxjnl-2020-215916)] [Medline: [33148809](https://pubmed.ncbi.nlm.nih.gov/33148809/)]
12. Hume E, Armstrong M, Manifold J, McNeillie L, Chambers F, Wakenshaw L, et al. Impact of COVID-19 shielding on physical activity and quality of life in patients with COPD. *Breathe (Sheff)*. Sep 2020;16(3):200231. [FREE Full text] [doi: [10.1183/20734735.0231-2020](https://doi.org/10.1183/20734735.0231-2020)] [Medline: [33447283](https://pubmed.ncbi.nlm.nih.gov/33447283/)]
13. Sykes DL, Faruqi S, Holdsworth L, Crooks MG. Impact of COVID-19 on COPD and asthma admissions, and the pandemic from a patient's perspective. *ERJ Open Res*. Jan 2021;7(1):00822-2020. [FREE Full text] [doi: [10.1183/23120541.00822-2020](https://doi.org/10.1183/23120541.00822-2020)] [Medline: [33575313](https://pubmed.ncbi.nlm.nih.gov/33575313/)]
14. Scarlata S, Cardaci V, Santangelo C, Matarese M, Cesari M, Antonelli Incalzi R. Distancing measures in COVID-19 pandemic: loneliness, more than physical isolation, affects health status and psycho-cognitive wellbeing in elderly patients with chronic obstructive pulmonary disease. *COPD*. Aug 2021;18(4):443-448. [FREE Full text] [doi: [10.1080/15412555.2021.1941834](https://doi.org/10.1080/15412555.2021.1941834)] [Medline: [34180766](https://pubmed.ncbi.nlm.nih.gov/34180766/)]
15. Gardner PJ, Moallem P. Psychological impact on SARS survivors: Critical review of the English language literature. *Canadian Psychology / Psychologie canadienne*. Feb 2015;56(1):123-135. [doi: [10.1037/a0037973](https://doi.org/10.1037/a0037973)]
16. Pradeepa S, Manjula KR, Vimal S, Khan MS, Chilamkurti N, Luhach AK. DRFS: detecting risk factor of stroke disease from social media using machine learning techniques. *Neural Process Lett*. Jun 09, 2020;55(4):3843-3861. [doi: [10.1007/s11063-020-10279-8](https://doi.org/10.1007/s11063-020-10279-8)]
17. Fu J, Li C, Zhou C, Li W, Lai J, Deng S, et al. Methods for analyzing the contents of social media for health care: scoping review. *J Med Internet Res*. Jun 26, 2023;25:e43349. [FREE Full text] [doi: [10.2196/43349](https://doi.org/10.2196/43349)] [Medline: [37358900](https://pubmed.ncbi.nlm.nih.gov/37358900/)]
18. Lwin MO, Lu J, Sheldenkar A, Schulz PJ, Shin W, Gupta R, et al. Global sentiments surrounding the COVID-19 pandemic on Twitter: analysis of Twitter trends. *JMIR Public Health Surveill*. May 22, 2020;6(2):e19447. [FREE Full text] [doi: [10.2196/19447](https://doi.org/10.2196/19447)] [Medline: [32412418](https://pubmed.ncbi.nlm.nih.gov/32412418/)]
19. Zhang C, Xu S, Li Z, Liu G, Dai D, Dong C. The evolution and disparities of online attitudes toward COVID-19 vaccines: year-long longitudinal and cross-sectional study. *J Med Internet Res*. Jan 21, 2022;24(1):e32394. [FREE Full text] [doi: [10.2196/32394](https://doi.org/10.2196/32394)] [Medline: [34878410](https://pubmed.ncbi.nlm.nih.gov/34878410/)]
20. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, et al. STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol*. Apr 2008;61(4):344-349. [doi: [10.1016/j.jclinepi.2007.11.008](https://doi.org/10.1016/j.jclinepi.2007.11.008)] [Medline: [18313558](https://pubmed.ncbi.nlm.nih.gov/18313558/)]
21. Volume streams. Twitter Developer Platform. URL: <https://developer.twitter.com/en/docs/twitter-api/tweets/sampled-stream/introduction> [accessed 2023-07-12]
22. Wang Z, Hale S, Adelani D, Grabowicz P, Hartmann T, Flöck F, et al. Demographic inference and representative population estimates from multilingual social media data. 2019. Presented at: The World Wide Web Conference; May 13-17:2056-2067; San Francisco, CA, USA.
23. Huang G, Liu Z, Maaten L, Weinberger K. Densely connected convolutional networks. 2017. Presented at: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); July 21-26:4700-4708; Honolulu, HI, USA. [doi: [10.1109/CVPR.2017.243](https://doi.org/10.1109/CVPR.2017.243)]
24. Preoțiuc-Pietro D, Lampos V, Aletras N. An analysis of the user occupational class through Twitter content. 2015. Presented at: 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing; July 1:1754-1764; Beijing, China. [doi: [10.3115/v1/P15-1169](https://doi.org/10.3115/v1/P15-1169)]
25. Standard occupational classification. Office for National Statistics. URL: <https://www.ons.gov.uk/methodology/> [accessed 2023-07-12]
26. Blei DM, Ng AY, Jordan MI. Latent dirichlet allocation. *Journal of Machine Learning Research*. 2003;3:993-1022. [doi: [10.5555/944919.944937](https://doi.org/10.5555/944919.944937)]
27. models.coherencemodel. Gensim. URL: <https://radimrehurek.com/gensim/models/coherencemodel.html> [accessed 2023-07-12]
28. Hutto C, Gilbert E. VADER: A parsimonious rule-based model for sentiment analysis of social media text. In: *ICWSM*. 2014. Presented at: Eighth International AAAI Conference on Weblogs and Social Media; June 1-4:216-225; Ann Arbor, MI, USA. [doi: [10.1609/icwsm.v8i1.14550](https://doi.org/10.1609/icwsm.v8i1.14550)]
29. Colneric N, Demsar J. Emotion recognition on Twitter: Comparative study and training a unison model. *IEEE Trans Affective Comput*. Jul 1, 2020;11(3):433-446. [doi: [10.1109/taffc.2018.2807817](https://doi.org/10.1109/taffc.2018.2807817)]
30. Plutchik R. A general psychoevolutionary theory of emotion. In: *Theories of Emotion*. New York. Elsevier; 1980:3-33.
31. Tenny S, Hoffman MR. Odds ratio. *StatPearls*. 2023:1-10. [Medline: [28613750](https://pubmed.ncbi.nlm.nih.gov/28613750/)]

32. P-value: comprehensive guide to understand, apply, and interpret. GeeksforGeeks. URL: <https://www.geeksforgeeks.org/p-value/> [accessed 2024-12-28]
33. Callaway B, Sant'Anna PH. Difference-in-differences with multiple time periods. *Journal of Econometrics*. Dec 2021;225(2):200-230. [doi: [10.1016/j.jeconom.2020.12.001](https://doi.org/10.1016/j.jeconom.2020.12.001)]
34. Greenwood BN, Agarwal R. Matching platforms and HIV incidence: An empirical investigation of race, gender, and socioeconomic status. *Management Science*. Aug 2016;62(8):2281-2303. [doi: [10.1287/mnsc.2015.2232](https://doi.org/10.1287/mnsc.2015.2232)]
35. Do I need ethics approval for social media research? JMIR. URL: <https://support.jmir.org/hc/en-us/articles/115001620728-Do-I-need-ethics-approval-for-social-media-research> [accessed 2024-09-29]
36. Jain R, Kumar A, Nayyar A, Dewan K, Garg R, Raman S, et al. Explaining sentiment analysis results on social media texts through visualization. *Multimed Tools Appl*. 2023;82(15):22613-22629. [FREE Full text] [doi: [10.1007/s11042-023-14432-y](https://doi.org/10.1007/s11042-023-14432-y)] [Medline: [36747895](https://pubmed.ncbi.nlm.nih.gov/36747895/)]
37. Varmaghani M, Dehghani M, Heidari E, Sharifi F, Moghaddam SS, Farzadfar F. Global prevalence of chronic obstructive pulmonary disease: systematic review and meta-analysis. *East Mediterr Health J*. Mar 19, 2019;25(1):47-57. [FREE Full text] [doi: [10.26719/emhj.18.014](https://doi.org/10.26719/emhj.18.014)] [Medline: [30919925](https://pubmed.ncbi.nlm.nih.gov/30919925/)]
38. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. Global Initiative for Chronic Obstructive Lung Disease. URL: <https://goldcopd.org/2023-gold-report-2/> [accessed 2023-07-12]
39. Markov PV, Ghafari M, Beer M, Lythgoe K, Simmonds P, Stilianakis NI, et al. The evolution of SARS-CoV-2. *Nat Rev Microbiol*. Jun 2023;21(6):361-379. [doi: [10.1038/s41579-023-00878-2](https://doi.org/10.1038/s41579-023-00878-2)] [Medline: [37020110](https://pubmed.ncbi.nlm.nih.gov/37020110/)]
40. Wang C, Pan R, Wan X, Tan Y, Xu L, Ho CS, et al. Immediate psychological responses and associated factors during the initial stage of the 2019 coronavirus disease (COVID-19) epidemic among the general population in China. *Int J Environ Res Public Health*. Mar 06, 2020;17(5):1729. [FREE Full text] [doi: [10.3390/ijerph17051729](https://doi.org/10.3390/ijerph17051729)] [Medline: [32155789](https://pubmed.ncbi.nlm.nih.gov/32155789/)]
41. Joensen LE, Madsen KP, Holm L, Nielsen KA, Rod MH, Petersen AA, et al. Diabetes and COVID-19: psychosocial consequences of the COVID-19 pandemic in people with diabetes in Denmark-what characterizes people with high levels of COVID-19-related worries? *Diabet Med*. Jul 2020;37(7):1146-1154. [FREE Full text] [doi: [10.1111/dme.14319](https://doi.org/10.1111/dme.14319)] [Medline: [32392380](https://pubmed.ncbi.nlm.nih.gov/32392380/)]
42. O'Neil A, Nicholls SJ, Redfern J, Brown A, Hare DL. Mental health and psychosocial challenges in the COVID-19 pandemic: food for thought for cardiovascular health care professionals. *Heart Lung Circ*. Jul 2020;29(7):960-963. [FREE Full text] [doi: [10.1016/j.hlc.2020.05.002](https://doi.org/10.1016/j.hlc.2020.05.002)] [Medline: [32561126](https://pubmed.ncbi.nlm.nih.gov/32561126/)]
43. Bjelland I, Dahl AA, Haug TT, Neckelmann D. The validity of the Hospital Anxiety and Depression Scale. An updated literature review. *J Psychosom Res*. Feb 2002;52(2):69-77. [doi: [10.1016/s0022-3999\(01\)00296-3](https://doi.org/10.1016/s0022-3999(01)00296-3)] [Medline: [11832252](https://pubmed.ncbi.nlm.nih.gov/11832252/)]
44. Lovibond SH, Lovibond PF. *Manual for the Depression Anxiety Stress Scales*. Sydney, Australia. Psychology Foundation of Australia; 1995.
45. Two-stage framework: Deep-COPD. GitHub. URL: https://github.com/cy Zhang87/COPD_analysis_code [accessed 2024-10-01]

Abbreviations

COPD: chronic obstructive pulmonary disease

DenseNet: dense convolutional network

DiD: difference-in-difference

IMDb: internet movie database

LDA: latent Dirichlet allocation

M3: multimodal, multilingual, and multiattribute

OC: occupation category

OR: odds ratio

STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

VADER: valence aware dictionary and emotional reasoner

Edited by T de Azevedo Cardoso; submitted 15.11.23; peer-reviewed by U Lokala, R Rupasingha; comments to author 18.07.24; revised version received 28.10.24; accepted 20.01.25; published 05.03.25

Please cite as:

Zhang C, Wang T, Dong C, Dai D, Zhou L, Li Z, Xu S

Exploring Psychological Trends in Populations With Chronic Obstructive Pulmonary Disease During COVID-19 and Beyond: Large-Scale Longitudinal Twitter Mining Study

J Med Internet Res 2025;27:e54543

URL: <https://www.jmir.org/2025/1/e54543>

doi: [10.2196/54543](https://doi.org/10.2196/54543)

PMID: [40053739](https://pubmed.ncbi.nlm.nih.gov/40053739/)

©Chunyan Zhang, Ting Wang, Caixia Dong, Duwei Dai, Linyun Zhou, Zongfang Li, Songhua Xu. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 05.03.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research (ISSN 1438-8871), is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.