Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

Public Health 198 (2021) 44–52

Contents lists available at ScienceDirect

Public Health

journal homepage: www.elsevier.com/locate/puhe

Review Paper

What can internet users' behaviours reveal about the mental health impacts of the COVID-19 pandemic? A systematic review

Vincenza Gianfredi [a, b, *], Sandro Provenzano [c], Omar Enzo Santangelo [d]

[a] School of Medicine, University Vita-Salute San Raffaele, 20132 Milan, Italy
[b] CAPHRI Care and Public Health Research Institute, Maastricht University, 6211 Maastricht, the Netherlands
[c] Local Health Unit of Trapani, Trapani, Italy
[d] Azienda Socio Sanitaria Territoriale di Lodi, ASST di Lodi, Italy

ARTICLE INFO

Article history:
Received 8 April 2021
Received in revised form
30 May 2021
Accepted 29 June 2021
Available online 5 July 2021

Keywords:
Novel data stream
Mental health
COVID-19
Twitter
Weibo
Google trends
Baidu
YouTube
Systematic review

ABSTRACT

Objectives: At the end of 2019, an acute infectious pneumonia (coronavirus disease 2019 [COVID-19]) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) began in Wuhan, China, and subsequently spread around the world starting a pandemic. Globally, to date, there have been >118 million confirmed cases, including >2 million deaths. In this context, it has been shown that the psychological impact of the pandemic is important and that it can be associated with an increase in internet searches related to fear, anxiety, depression, as well as protective behaviours, health knowledge and even maladaptive behaviours.

Study design: This is a systematic review.

Methods: This review aims to collect, analyse and synthesise available evidence on novel data streams for surveillance purposes and/or their potential for capturing the public reaction to epidemic outbreaks, particularly focusing on mental health effects and emotions.

Results: At the end of the screening process, 19 articles were included in this systematic review. Our results show that the COVID-19 pandemic had a great impact on internet searches for mental health of entire populations, which manifests itself in a significant increase of depressed, anxious and stressed internet users' emotions.

Conclusions: Novel data streams can support public health experts and policymakers in establishing priorities and setting up long-term strategies to mitigate symptoms and tackle mental health disorders.

© 2021 The Royal Society for Public Health. Published by Elsevier Ltd. All rights reserved.

Introduction

From November to December 2019, the acute infectious pneumonia, termed COVID-19 (coronavirus disease 2019), which is caused by a new highly contagious and pathogenic respiratory coronavirus (severe acute respiratory syndrome coronavirus 2 [SARS-CoV-2]), began in China and subsequently spread causing a global pandemic.[1] To date, approximately 1 year after the World Health Organization (WHO) first declared the COVID-19 epidemic a pandemic,[2] >118 million cases and >2 million deaths have been recorded.[3] The world is facing a complex crisis not only from a health perspective but also from a social and economic point of view.[4]

According to Brooks et al., the psychological impact of the pandemic has been moderate to severe, with 17% of adults reporting moderate-to-severe depressive symptoms, 29% moderate-to-severe anxiety symptoms and 8% moderate-to-severe stress levels.[5] This was mainly due to isolation or quarantine measures, the scarce possibility of social interactions and the interruption of work for many with consequent economic loss.[5] Children and adolescents have also experienced profound life changes characterised by emotional isolation and intensive use of the internet, which may increase vulnerability to anxiety, stress and suicidal ideas or exacerbate self-inflicted violence.[6]

The perception of threat associated with a global pandemic may generate fear and lead to negative emotional reactions; fear, in turn, motivates people to initiate behavioural changes and improve their health knowledge.[7,8] Traditional and social media also tend to

* Corresponding author. University Vita-Salute San Raffaele, Via Olgettina 60, 20132, Milan, Italy.
E-mail addresses: gianfredi.vincenza@hsr.it (V. Gianfredi), provenzanosandro@hotmail.it (S. Provenzano), omarenzosantangelo@hotmail.it (O.E. Santangelo).

https://doi.org/10.1016/j.puhe.2021.06.024
0033-3506/© 2021 The Royal Society for Public Health. Published by Elsevier Ltd. All rights reserved.

report on cases of infection and mortality, while rarely mentioning cases of recovery, thereby, facilitating the spread of fear and quickly generating anxiety or distress.[9] To reduce these effects, the WHO has advised the public to focus less on COVID-19 news, since watching, reading or hearing news about COVID-19 has been shown to exacerbate fear.[10] This is particularly important as a large amount of news (and fake news) is available to the general public who may not always have the appropriate knowledge to understand scientific updates (infodemic).[11,12] Hossain et al.[13] also show that, despite demographic and social differences, the epidemiological distribution of mental health problems (due to the pandemic and associated factors) has been heterogeneously distributed across the general public, as well as among COVID-19 patients and healthcare professionals. This suggests that there is an ongoing psychiatric epidemic parallel to the COVID-19 pandemic, which requires immediate intervention from the public health community.[14,15]

In this context, it is necessary to understand how people cope with the pandemic. Internet searches could help answer this question since the internet represents a primary channel for finding health-related information.[16] It has been shown that a rise in COVID-19 cases can be associated with a surge in internet searches related to fear, anxiety, depression, as well as protective behaviours, health knowledge and even maladaptive behaviours.[17]

The digital era has given rise to new data sources and large amounts of data also known as Big Data; according to De Mauro et al., "[Big Data] represent resources/assets of an informative nature characterised by such a high volume, speed and variety as to require technology and analytical methods specific for its transformation into value".[18] The use of Big Data in scientific research is destined to grow and become increasingly present in daily healthcare practice.[19] Within Big Data, we can include the novel data streams, which are identified as those data whose "content is initiated directly by the user (patient) themselves".[20] Research studies, which are often limited to certain categories of patients, could expand their targets, including several different population subgroups, while the analyses conducted on Big Data could make health care even more personalised at the same time. The magnitude of data available on patients with the same comorbidities and specific rare diseases could facilitate the creation of *ad hoc* individual-level therapeutic plans in the future.[19]

This review aims to collect, analyse and synthesise available evidence on novel data streams for surveillance purposes and/or their potential for capturing the public reaction to epidemic outbreaks, particularly focusing on mental health effects and negative emotions. Indeed, previous research shows that nowadays any mediatic event produces a social reaction on the internet, which, in turn, generates a large volume of data, that can be collected and analysed in order to answer health-related questions and issues.[21–23] For this review, we considered the COVID-19 pandemic as the mediatic event; subsequently, the general public reaction generated a novel data stream, which has been analysed in order to identify any potential impact on mental health.

Methods

The Cochrane Collaboration[24] and the Meta-analysis Of Observational Studies in Epidemiology (MOOSE) guidelines[25] were followed in order to conduct the current systematic review. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses[26] guidelines[27] were used to report the process and results. The structured literature search was conducted on PubMed/Medline and Scopus on 17th February 2021, combining free text words and medical subject headings (MeSH). Keywords were combined using Boolean operators AND and OR. No time filter was applied. The full search strategy is available in Supplementary Table S1.

Inclusion/exclusion criteria

In order to be considered eligible, articles had to meet the following inclusion criteria: (i) article written in English; (ii) reporting original data; (iii) focusing on mental health; (iv) during the COVID-19 pandemic; and (v) addressing the usage of non-conventional data approaches in capturing public reaction to COVID-19 epidemic outbreaks, particularly focusing on adverse mental health effects and emotions. Exclusion criteria were as follows: (i) no original data (i.e. a review, opinion, book chapter, commentary or letter); (ii) article not published in English; (iii) full text not available; (iv) focusing on mental health but not during the COVID-19 pandemic; and (v) assessing outcomes not related to mental health.

Articles were firstly screened based on title and abstract and then assessed in full text. Both article screening phases and data extraction were conducted independently by two authors (VG and OES). Full texts were downloaded only for potentially eligible studies. Data extraction was conducted only for those articles that met all the inclusion criteria and was performed using a pre-defined and pre-piloted spreadsheet elaborated in Microsoft Excel® for Windows. Extracted data included author and year, data source, type of mental health outcome assessed in the study, study period, country where the study was conducted, study aims, keywords used to perform the search, type of analysis conducted and main results. Any disagreement in data extraction was resolved through discussion among the two authors; if any disagreement persisted, a third author was consulted (SP).

Results

Literature search

A total of 113 papers were retrieved on PubMed and 436 on Scopus. Of these 549 articles, 62 were duplicates and immediately removed, leaving a final sample of 487 unique papers. After preliminary screening, based on the title and abstract, 462 articles were removed for the following reasons: unrelated topic (n = 411), review (n = 25), protocol study (n = 17) and article not published in English (n = 9). Out of 25 eligible articles, six articles were removed after full-text assessment because of the following reasons: two articles did not specifically report mental health data;[28,29] one article was a feasibility study;[30] one study referred to pre-COVID-19 times;[31] one article, although using smartphone app usage data, did not correlate depressive symptoms and anxiety (self-reported symptoms) with app data, but with COVID-19 news;[32] and, the last article using passive mobile monitoring platforms, assessed the behavioural changes (i.e. time spent at home) among subjects with psychiatric disorders.[33] Fig. 1 depicts the flow diagram reporting the selection process. At the end of the screening process, 19 articles were included in the systematic review.[34–52]

Characteristics of included studies

Among the 19 retrieved articles, 15 used a single-data source, whereas four used a combination of two or more data sets. In particular, Twitter was the source most frequently used alone (n = 4),[36,46,51,52] and in combination (n = 3) with YouTube and a forum,[37] Geographic information system (GIS)[39] and Weibo.[47] Google Trends was the second most frequently used data source (n = 6);[38,41,43,44,48,49] one study investigated Weibo[40] data and another investigated Baidu[34] data alone. Forums and chats were

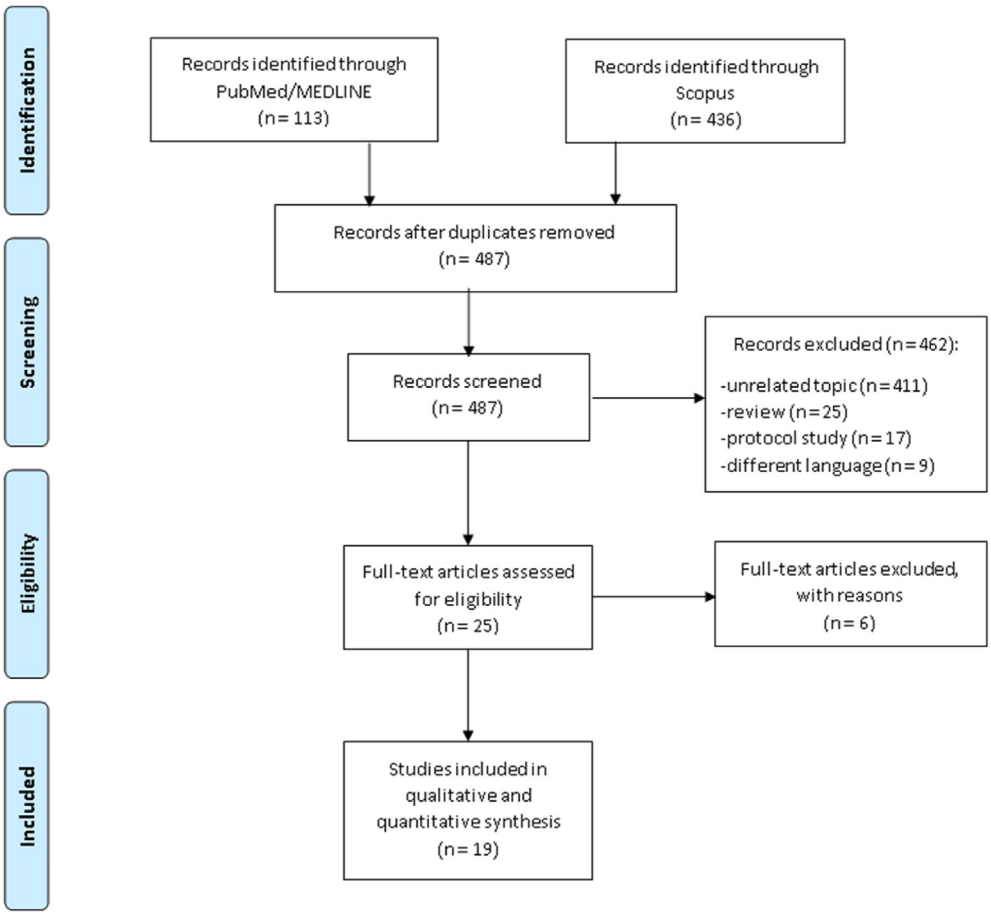

Fig. 1. Flow diagram of the study-selection process.

both used alone (n = 3)[35,42,45] and in combination with other social networks.[37]

Regarding the country where the studies were conducted, four were internationally based,[41–43,52] whereas the rest were carried out in a single country. Specifically, the US was the most frequently explored country (n = 5),[36,38,39,46,50] followed by China (n = 3)[34,40,47] (including one study that compared data between Wuhan [China] and the Lombardy region [Italy]),[47] India (n = 3),[44,48,49] Australia,[35,51] Canada[45] and Spain.[37]

The mental health outcomes assessed were frequently general mental or emotional health, or a combination of selected mental health symptoms (e.g. depression, anxiety, suicide ideation, sleep disturbance, fear and stress), while only three articles focused on a specific mental health disorder (one study focused on stress,[39] one on eating disorders[45] and one on depression[51]).

Almost all included studies performed an analysis of the internet activity content/sentiment in order to explore the predictive value of these searches in forecasting mental health disorders. Some studies assessed how online searches changed throughout the pandemic,[37,38,47,49] whereas, six articles assessed the association between COVID-19 cases or deaths and internet search volume related to mental health.[39,41,43,44,48,51]

All included articles found a significant association with at least one of the explored mental health disorders/symptoms; however, three studies found conflicting results for specific mental health outcomes. In particular, the study conducted using data from Baidu (China) found a significant association between the COVID-19 pandemic and panic and fear, but failed to detect an association with depression, sadness and obsessive-compulsive disorders.[34] An international study performed on Google Trends found an association between insomnia and COVID-19 deaths, but not with depression and suicidal thoughts.[41] Lastly, even an international (24 countries) study found a significant association between COVID-19 cases and deaths and suicide and depression, but no association with insomnia.[43] These differences may be explained by considering the different countries and the length of the study period. Characteristics of included studies and main results are provided in Table 1.

Discussion

Results of our systematic review revealed that the COVID-19 pandemic has significantly impacted internet searches on mental health of entire populations. Indeed, in almost all included articles, a statistically significant increase in negative (seen as depressed, anxious and stressed) internet and social network users' emotions was recorded. These data were confirmed both by associating the number of COVID-19 cases or deaths with internet search volume for specific keywords, as well as by content and sentiment analysis of posts published on social networks. This phenomenon was confirmed in all studied countries, including in those articles where several countries were studied together.

It is interesting to note the parallels drawn by Su et al.,[47] who assessed the psychological states of social network users (Twitter and Weibo) in Lombardy (Italy) and Wuhan (China), the two regions first affected by SARS-CoV-2. Chinese users significantly

Table 1
Characteristics and the main results of included studies (reported in alphabetical order of first author).

First author, year (Reference)	Data Source	Mental Health outcome	Study period	Country	Study aims	Keywords	Analysis	Main results
Chen, 2020[34]	Baidu	Depression, anxiety, sleeping problem, obsessive-compulsive disorder	24 Jan to 22 Feb, 2020	China	To understand psychological public reaction to the COVID-19 pandemic	Depressed, sad, panic, fear, insomnia, obsessive-compulsive disorder, psychological counselling	Non-parametric Mann –Whitney U test	Internet search for panic/fear and psychological counselling were more frequent in 2020 than in 2019 ($P < 0.05$), depressed/sad and obsessive-compulsive disorder were the opposite ($P < 0.05$). No differences for insomnia.
Chivers, 2020[35]	"New mum forum" in Google (831 posts)	Distress, despair	27 Jan to 12 May, 2020	Australia	To examine the public discourse in the context of COVID-19 pandemic	COVID, corona, pandemic	Thematic analysis, sentiment analysis and word frequency calculations	Distress related to external high risk; despair for lack of social and family support. Half of the posts have a negative sentiment. The most frequent word was worry (and similar).
Guntuku, 2020[36]	Twitter	Stress, anxiety and loneliness	Jan to May 2020 vs the same period in 2019	US	To analyse the content of tweets with a specific focus on mental health during the COVID-19 pandemic	Sentiment, stress, anxiety, loneliness and COVID-19-related symptoms	Pre-trained data-driven machine learning models	Stress, anxiety and loneliness were more frequent in 2020 ($P < 0.001$), whereas sentiment was lower in 2020 ($P < 0.001$).
Iglesias-Sánchez, 2020[37]	Twitter, YouTube, Instagram, official press websites and Internet forums (80,091 posts)	Emotional health	Mar to May 2020, split in three stages (characterised by enforcement of containment measures of the first two, and mitigation of the last one)	Spain	To understand emotional health during the COVID-19 confinement time	Anger, fear, joy, sadness, disgust and uncertainty	Content analysis	Anger, fear, sadness and uncertainty were statistically significantly different throughout the three studied periods, but not joy and disgust.
Jacobson, 2020[38]	Google Trends	Mental health symptoms	16–23 Mar 2020 (stay-at-home orders in 11 states)	US	To examine if COVID-19-containment measures produced changes in mental health symptoms	Anxiety, depression, obsessive-compulsive disorder, hopeless, angry, afraid, apathy, worthless, worried, restless, irritable, tense, scattered, tired, avoiding, procrastinate, insomnia, suicidal, suicide	Generalised additive mixed models	Topics related to anxiety, negative thoughts about oneself and the future, sleep disturbances, and suicidal ideation increased prior to stay-at-home orders.
Li, 2020[39]	Twitter and GIS (Geographic Information System)	Stress symptoms	Jan to Apr 2020	US	To detect COVID-19 related stress symptoms at a spatiotemporal scale	Patient Health Questionnaire lexicon	Machine learning	Stress-related tweets rapidly increased from March 8th until April 5th when the number of COVID-19 cases started to decrease.
Li, 2020[40]	Weibo	Psychological traits	13–26 Jan, 2020	China	To explore the predictive value of Weibo in forecasting mental health	Words of positive emotion, negative emotion, anxiety, anger, health, leisure, family, friend, money, death and religion	Sentiment analysis	Between 20 and 26 January, negative emotional indicators of psychological traits increased in anxiety, depression, and indignation, while positive emotional indicators decreased.
Lin, 2020[41]	Google Trends	Insomnia, depression, and suicide	20 Mar to 19 Apr, 2020	International	To explore the predictive value of Google Trends in forecasting mental distress	Insomnia, depression, suicide	Mathematical model	Iran, Spain, US and Italy were the countries with the highest insomnia research volume. COVID-19 death numbers were associated with days at higher insomnia searches, but not for depression and suicide.
Low, 2020[42]	Reddit Mental Health Dataset	General mental health	1 Jan to 20 Apr, 2020	International	To assess changes in the natural language of public commentary	Lexicon built on Linguistic Inquiry and Word Count	Machine learning	Attention-deficit/hyperactivity disorder, eating disorders and anxiety showed the

(*continued on next page*)

Table 1 (*continued*)

First author, year (Reference)	Data Source	Mental Health outcome	Study period	Country	Study aims	Keywords	Analysis	Main results
					posted to mental health support groups on the Reddit website			largest sum of change in negative semantic features during the pandemic. Anxiety emerged as a general theme across Reddit.
Misiak, 2020[43]	Google Trends	General mental health	18 Feb to 13 Apr, 2020	International (24 countries)	To assess if COVID-19 cases and death are associated with mental health-related internet search volume	Suicide, depression, anxiety, insomnia	Spearman rank correlation	Suicide and depression were both associated with COVID-19 deaths and cases, anxiety was only associated with deaths and insomnia was not associated with cases neither deaths.
Rana, 2020[44]	Google Trends	Mental health indicators	4 Mar to 25 Jul, 2020	India	To estimate the correlation between COVID-19 cases and internet search	Suicide, autism, anxiety, depression, dementia	Pearson correlation	Suicide, anxiety, depression and dementia were highly correlated with the number of daily infectious cases.
Richardson, 2020[45]	National Eating Disorder Information Centre instant chat service	Eating disorders	1 Mar to 30 Apr, 2020	Canada	To describe the impact of the COVID-19 pandemic on help-seeking behaviours	Over-eating/binge eating, over-exercising, dieting/restriction, weight preoccupation, perfectionism, purging, anxiety, and depression	Content analysis	The number of contacts during the pandemic period was significantly higher compared to the two previous years. Eating disorder symptoms, anxiety, and depression were higher in 2020. Content analysis detected four themes: 1) lack of access to treatment, 2) worsening of symptoms, 3) feeling out of control and 4) need for support.
Saha, 2020[46]	Twitter (59,096,694 tweets in 2020 compared with 40,875,185 in 2019)	Anxiety, depression, stress, suicidal ideation and emotional support	24 Mar to 24 May, 2020	US	To study the temporal and linguistic changes in symptomatic mental health expressions during the COVID-19 pandemic	Classification based on Diagnostic and Statistical Manual of Mental Disorders, 5th Edition	Machine learning	Anxiety tweets showed the most significant increase in 2020 compared to 2019, followed by suicidal ideation, depression and stress, with a steady decline during the study period.
Su, 2020[47]	Twitter (3,650,380 tweets) and Weibo	Psychological states	23 Feb to 21 Mar, 2020 (Twitter); 9 Jan to 5 Feb, 2020 (Weibo)	Wuhan (China) and Lombardy (Italy) region	To examine and compare the impact of COVID-19 lockdown on individuals' psychological states in China and Italy	Language Inquiry and Word Count dictionary	Psychometric analysis	In Weibo posts, (China) users increased the use of first-person plural pronoun, religion, social, negative emotion and home significantly after the lockdown. Twitter (Italy) users increased the wording discrepancy and home and decreased anxiety.
Talbot, 2021[52]	Twitter (192 tweets)	Mental health of pregnant women	1 Mar to 31 May, 2020	International mostly from US and UK	To characterise the content of an international sample of tweets related to pregnancy and mental health	Several keywords related to COVID-19, pregnancy and mental health	Sentiment and thematic analysis	Individual and company tweets had the same negative tonality. Among the individual tweets, stress about being isolated, depressive symptoms and sleep difficulties were the most frequent. Among the tweets from companies, stress and depressive symptoms that pregnant women may experience during the pandemic, as well as the services offered were the most frequent.
Uvais, 2020[48]	Google Trends	General mental health	12 Mar to 13 Jun, 2020	India	To explore the association of internet search	Depression, anxiety, insomnia, suicide	Spearman rank correlation	COVID-19 cases and deaths in India were significantly associated

Table 1 (*continued*)

First author, year (Reference)	Data Source	Mental Health outcome	Study period	Country	Study aims	Keywords	Analysis	Main results
					volume and COVID-19 cases			with searches for depression, anxiety and suicide. Moreover, suicide and depression were the most frequently searched terms.
Uvais, 2020[49]	Google Trends	General mental health	26 Jan to 24 May, 2020	India	To understand the changes in patient interest in psychiatric search terms during the lockdown	Depression, anxiety, suicide	T-test	Depression, anxiety, and suicide search increased from the pre-lockdown to the lockdown period.
Zhang, 2020[50]	Google Search and YouTube	Depression and anxiety	Jan 2020 vs May 2020	US	To explore the predictive value of YouTube and Google Trends for forecasting depression and anxiety	Not available	Machine learning	PHQ-9 and GAD-7 scores increased during the study period. Some internet behaviours such as late-night online activity, short YouTube intervals between videos and videos containing anxiety and sadness keywords were associated with scores increment.
Zhou, 2020[51]	Twitter (94,707,264 tweets)	Depression	1 Jan to 22 May, 2020	Australia	To explore the predictive value of Twitter in forecasting depression	Classification based on Diagnostic and Statistical Manual of Mental Disorders, 5th Edition	Mathematical model	Depression level increased following the increasing COVID-19 cases and reinforcement of government containment measures.

GIS, Geographic information system; GAD-7, General Anxiety Disorder-7; PHQ-9, Patient Health Questionnaire-9.

increased the use of first-person plural pronouns, religion, social, negative emotion and home after the lockdown. Italian users increased the use of the words 'discrepancy' and 'home' while searching less frequently for anxiety. These results show how the same event can impact differently on communities based on cultural and societal aspects. Despite the posts' common negative and fearful content, the wording between cultures differed, at least partially.

Moreover, it should be considered that, in the e-health era, the diffusion of information by the internet and social media is faster than ever before, even surpassing the velocity of spread of a highly infectious virus such as SARS-CoV-2. Furthermore, information is not always correct or validated, promoting the spread of misleading or incorrect facts on unfiltered internet platforms, which include all major social networks (i.e. Facebook, Twitter, YouTube and several others).

In addition, it is important to note that the COVID-19 pandemic is sustained by an infection with a completely novel and still partially unknown virus.[53,54] This characteristic highly impacts several aspects. Firstly, due to the unknown nature of the virus, even trustworthy sources made mistakes in communication at the beginning of the epidemic, fuelling confusion and a hyperreactive information search among the general public.[11] Secondly, the lack of appropriate therapies and preventive measures (such as vaccines) during the first year of the pandemic incited a sense of powerlessness and fear in the population, which, in turn, drove the search for information and fuelled a widespread sense of stress, uncertainty, anxiety and depression.[55] Moreover, in pandemic times, social deprivation or lack of a well-structured communication of information may strengthen prejudice and stigma towards vulnerable individuals, which, in turn, reinforces a sense of despair, loneliness, depression and anxiety.[15] Also, the implementation of containment measures, such as social distancing, national lockdowns and absence from the workplace, may have pushed people to spend a lot of time on the internet, not only to search for information but also to connect with friends and relatives.[56]

All of these human behaviours generate a significant amount of data on numerous different platforms (Big Data and novel data streams), which contribute to a newly emerging research field aimed at supporting traditional surveillance systems (especially for infectious diseases)[22,57,58] and at understanding human behaviours,[21,23] with the aim of guiding public health decisions.[59,60]

Systematically collecting and examining health-related data generated by these new sources might carry the potential to understand, investigate and interpret events relevant for public health purposes.[61] Several cross-sectional studies, based on online surveys administered to convenience samples, have been published throughout 2020, without any information on the real generalisability and representativeness of results.[62–69] In contrast, these novel data streams might reduce several biases typical of surveys, such as social desirability and recall bias, as well as improving the above-mentioned representativeness of the sample, since thousands or millions of users' data can be analysed together. Moreover, Big Data and novel data streams reduce the time lag between traditional epidemiological data collection and data analysis, as well as the data entry burden and potential errors that can occur with traditional epidemiological systems. Furthermore, novel data streams and Big Data can improve spatial resolution and forecast unanticipated outcomes of interest (such as the current study looking at the mental health impacts of the pandemic), especially for outcomes not routinely surveyed.[20] Big data and novel data streams can help to better understand many disease transmission/process aspects that are not usually captured by traditional surveillance systems. Last, but not least, Big Data and novel data streams are important for improving data dissemination.[20]

Results of this systematic review are extremely relevant because they offer an overview of the general public's psychological reaction to the COVID-19 outbreak, showing how fear, anxiety, stress,

suicidal thoughts and depression are strongly associated with the pandemic. Previous systematic reviews on cross-sectional studies assessing the mental health status of the general public during the COVID-19 pandemic showed that the prevalence of depression increased from 7.2% to 14.6–48.3% during the pandemic.[70] Moreover, in addition to the previously established risk factors for depression, such as female gender, presence of chronic/psychiatric illnesses and unemployment, frequent exposure to news concerning COVID-19 was identified as one of the most important risk factors for depression during the pandemic.[70]

In light of the disease burden already caused by mental disorders,[71] and considering the prolonged exposure to this stressful pandemic event, we might expect long-term psychological consequences. This is especially true considering that, during the pandemic, many healthcare activities have been postponed[72] or substituted by telemedicine.[73] Indeed, telemedicine offers a great opportunity to deliver mental health services remotely, without borders, on a large scale and in a safe way.[73,74] However, telemedicine, in certain cases, has shown some limitations.[5] Even if telemedicine is a more readily accessible option, overcoming the problems related to physical accessibility to services (e.g. remote location, lack of transportation or funds, stigma associated with seeking treatment and physical mobility-related health issues),[75] not all patients have the necessary technology, even just a simple internet line, to be able to use telemedicine services, and some individuals may not have the skills to use these tools.[76] From this perspective, there is a risk that a service designed to reach a large number of people creates inequalities in the population, with the possibility of 'leaving behind' some already disadvantaged groups, such as the poorest, the elderly or those who live in disadvantaged social contexts.[77]

Moreover, evidence has found that during the COVID-19 pandemic, a reduced number of mental health services has coincided with an increased mental health burden, potentially indicating an escalating number of untreated individuals with mental health problems.[78]

We should be aware of the above-mentioned elements in planning future healthcare services. From this perspective, Big Data and novel data streams can support public health experts and policymakers in establishing priorities and setting up long-term strategies to mitigate the symptoms and tackle mental health disorders, with the final aim of facilitating the implementation of future preventive interventions.

Strengths and limitations

Before generalising the results of the current systematic review, some limitations need to be taken into account. Firstly, we limited our search to articles published in English. This might have reduced the total number of potentially eligible studies; however, since English is the most commonly used language in the scientific community, we believe this did not significantly affect our results. Secondly, some authors divided their analyses across several short papers, multiplying the evidence volume but reducing the quality of their discussion and the interpretation of results. Lastly, internet search spikes and contents of tweets may be influenced by different triggers, such as the increased number of cases, increased attention given by mass media or as a reaction to the containment measures adopted by governments, making it difficult to precisely identify a cause–effect association. In light of this, it is not possible to differentiate between the mental health effects of fear of the virus itself from the containment measures adopted, neither to differentiate between negative feelings and diagnosis of mental health disorders. However, previous cross-sectional studies conducted during the pandemic confirmed a variable increase in the prevalence of depression, anxiety and stress in the general population.[70,72,79,80]

Despite the above-mentioned limitations, our review has some important strengths. Firstly, this is a systematic and extensive review offering an exhaustive overview of available evidence on Big Data and mental health during the COVID-19 pandemic. With this overview, we not only synthesised and analysed results obtained by each included study but also provided a picture of the data sources used, the aims of the original studies, as well as the types of analyses performed. Secondly, our search strategy was developed considering several keywords, including both MeSH and free text terms. Lastly, as far as we know, this is the first systematic review addressing the usage of non-conventional data approaches in capturing public reaction to the ongoing COVID-19 pandemic, particularly focusing on mental health and emotions.

Conclusions

To conclude, the results of this systematic review found a statistically significant increase in searches for information about mental health issues by the general public throughout the COVID-19 pandemic. This was observed in both single- and multi-country studies, which analysed data in up to 24 different national contexts. The most-searched-for terms included anxiety, fear, suicide, depression, despaired, stress, solitude and loneliness. While some papers aimed to examine how online expressions of mental health changed throughout the pandemic, others specifically investigated whether containment measures or COVID-19 cases and deaths were associated with mental health internet search volumes. Internet searches and users' behaviour on social networks generated a large amount of data (Big Data) and novel data streams that were used by numerous researchers all over the world. Results obtained from these analyses might prove extremely useful in informing policymakers and health authorities on the implementation of healthcare services and policies, with a focus beyond treating COVID-19 patients, to also provide care to those affected by direct and indirect mental health consequences.

Author statements

Ethical approval

Not required.

Funding

None declared.

Competing interests

None declared.

Author contributions

VG conceptualised, designed the study and performed the literature search. VG and OES performed resource analysis and data extraction. VG and OES wrote the first draft. All authors have read and agreed to the published version of the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.puhe.2021.06.024.

References

1. Perrella A, Carannante N, Berretta M, Rinaldi M, Maturo N, Rinaldi L. Novel Coronavirus 2019 (Sars-CoV2): a global emergency that needs new approaches? *Eur Rev Med Pharmacol Sci* 2020;**24**(4):2162–4.
2. World Health Organization. *WHO announces COVID-19 outbreak a pandemic*. 2020. Available from: http://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/news/news/2020/3/who-announces-covid-19-outbreak-a-pandemic.
3. World Health Organization. *Coronavirus disease (COVID-19) dashboard*. 2021. Available from: https://covid19.who.int/?gclid=CjwKCAiAyc2BBhAaEiwA44-wWyljMo8PMXghPyqWZ6fNQIR9fFi3mcnQGT2YxNJUD_V3Rfp6Av_TdxoC4gwQAvD_BwE.
4. Nicola M, Alsafi Z, Sohrabi C, Kerwan A, Al-Jabir A, Iosifidis C, et al. The socio-economic implications of the coronavirus pandemic (COVID-19): a review. *Int J Surg* 2020;**78**:185–93.
5. Brooks SK, Webster RK, Smith LE, Woodland L, Wessely S, Greenberg N, et al. The psychological impact of quarantine and how to reduce it: rapid review of the evidence. *Lancet* 2020;**395**(10227):912–20.
6. Deslandes SF, Coutinho T. The intensive use of the internet by children and adolescents in the context of COVID-19 and the risks for self-inflicted violence. *Ciência Saúde Coletiva* 2020;**25**(suppl 1):2479–86.
7. Kok G, Bartholomew LK, Parcel GS, Gottlieb NH, Fernandez ME. Finding theory- and evidence-based alternatives to fear appeals: intervention Mapping. *Int J Psychol* 2014;**49**(2):98–107.
8. Bavel JJV, Baicker K, Boggio PS, Capraro V, Cichocka A, Cikara M, et al. Using social and behavioural science to support COVID-19 pandemic response. *Nat Hum Behav* 2020;**4**(5):460–71.
9. Depoux A, Martin S, Karafillakis E, Preet R, Wilder-Smith A, Larson H. The pandemic of social media panic travels faster than the COVID-19 outbreak. *J Trav Med* 2020;**27**(3).
10. World Health Organization. *Mental health and psychosocial considerations during the COVID-19 outbreak*. 2020. Available from: https://www.who.int/docs/default-source/coronaviruse/mental-health-considerations.pdf.
11. Patel MP, Kute VB, Agarwal SK. Nephrology C-WGoISo. "Infodemic" COVID 19: more pandemic than the virus. *Indian J Nephrol* 2020;**30**(3):188–91.
12. Provenzano S, Santangelo OE, Armetta F, Pesco G, Allegro A, Lampasona M, et al. COVID-19 infection: comparing the knowledge, attitude and practices in a sample of nursing students. *Acta Biomed* 2020;**91**(12-S):e2020001.
13. Hossain MM, Tasnim S, Sultana A, Faizah F, Mazumder H, Zou L, et al. Epidemiology of mental health problems in COVID-19: a review. *F1000Research* 2020;**9**:636.
14. Amerio A, Aguglia A, Odone A, Gianfredi V, Serafini G, Signorelli C, et al. Covid-19 pandemic impact on mental health of vulnerable populations. *Acta Biomed* 2020;**91**(9-S):95–6.
15. Amerio A, Odone A, Aguglia A, Gianfredi V, Bellini L, Bucci D, et al. La casa de papel: a pandemic in a pandemic. *J Affect Disord* 2020;**277**:53–4.
16. Du H, Yang J, King RB, Yang L, Chi P. COVID-19 increases online searches for emotional and health-related terms. *Appl Psychol Health Well-being* 2020;**12**(4):1039–53.
17. Twenge JM, Joiner TE. Mental distress among U.S. adults during the COVID-19 pandemic. *J Clin Psychol* 2020;**76**(12):2170–82.
18. De Mauro A, Greco M, Grimaldi M. What is Big Data? A consensual definition and a review of key research topics. *AIP Conf Proc* 2016;**1644**(1):97–104.
19. Murdoch TB, Detsky AS. The inevitable application of big data to health care. *J Am Med Assoc* 2013;**309**(13):1351–2.
20. Althouse BM, Scarpino SV, Meyers LA, Ayers JW, Bargsten M, Baumbach J, et al. Enhancing disease surveillance with novel data streams: challenges and opportunities. *EPJ Data Sci* 2015;**4**.
21. Bragazzi NL, Barberis I, Rosselli R, Gianfredi V, Nucci D, Moretti M, et al. How often people google for vaccination: qualitative and quantitative insights from a systematic search of the web-based activities using Google Trends. *Hum Vaccines Immunother* 2017;**13**(2):464–9.
22. Gianfredi V, Bragazzi NL, Mahamid M, Bisharat B, Mahroum N, Amital H, et al. Monitoring public interest toward pertussis outbreaks: an extensive Google Trends-based analysis. *Publ Health* 2018;**165**:9–15.
23. Mahroum N, Bragazzi NL, Sharif K, Gianfredi V, Nucci D, Rosselli R, et al. Leveraging google trends, twitter, and wikipedia to investigate the impact of a celebrity's death from rheumatoid arthritis. *J Clin Rheumatol* 2018;**24**(4):188–92.
24. Higgins JP, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011;**343**:d5928.
25. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis of Observational Studies in Epidemiology (MOOSE) group. *J Am Med Assoc* 2000;**283**(15):2008–12.
26. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009;**6**(7):e1000097.
27. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Ann Intern Med* 2009;**151**(4):W65–94.
28. Lucini D, Gandolfi CE, Antonucci C, Cavagna A, Valzano E, Botta E, et al. #stay home stay fit: UNIMI's approach to online healthy lifestyle promotion during the covid-19 pandemic. *Acta Biomed* 2020;**91**(3):1–7.
29. Valdez D, Ten Thij M, Bathina K, Rutter LA, Bollen J. Social media insights into US mental health during the COVID-19 pandemic: longitudinal analysis of twitter data. *J Med Internet Res* 2020;**22**(12):e21418.
30. Alabdulkreem E. Prediction of depressed Arab women using their tweets. *J Decis Syst* 2021. https://doi.org/10.1080/12460125.2020.1859745.
31. Hwang J, Toma CL, Chen J, Shah DV, Gustafson D, Mares ML. Effects of web-based social connectedness on older adults' depressive symptoms: a two-wave cross-lagged panel study. *J Med Internet Res* 2021;**23**(1):e21275.
32. Huckins JF, da Silva AW, Wang W, Hedlund E, Rogers C, Nepal SK, et al. Mental health and behavior of college students during the early phases of the COVID-19 pandemic: longitudinal smartphone and ecological momentary assessment study. *J Med Internet Res* 2020;**22**(6).
33. Jagesar RR, Roozen MC, van der Heijden I, Ikani N, Tyborowska A, Penninx B, et al. Digital phenotyping and the COVID-19 pandemic: capturing behavioral change in patients with psychiatric disorders. *Eur Neuropsychopharmacol : J Eur Coll Neuropsychopharmacol* 2021;**42**:115–20.
34. Chen H, Zhang K. Insight into the psychological problems on the epidemic of COVID-19 in China by online searching behaviors. *J Affect Disord* 2020;**276**:1093–4.
35. Chivers BR, Garad RM, Boyle JA, Skouteris H, Teede HJ, Harrison CL. Perinatal distress during COVID-19: thematic analysis of an online parenting forum. *J Med Internet Res* 2020;**22**(9):e22002.
36. Guntuku SC, Sherman G, Stokes DC, Agarwal AK, Seltzer E, Merchant RM, et al. Tracking mental health and symptom mentions on twitter during COVID-19. *J Gen Intern Med* 2020;**35**(9):2798–800.
37. Iglesias-Sánchez PP, Vaccaro Witt GF, Cabrera FE, Jambrino-Maldonado C. The contagion of sentiments during the COVID-19 pandemic crisis: the case of isolation in Spain. *Int J Environ Res Publ Health* 2020;**17**(16).
38. Jacobson NC, Lekkas D, Price G, Heinz MV, Song M, James O'Malley A, et al. Flattening the mental health curve: COVID-19 stay-at-home orders are associated with alterations in mental health search behavior in the United States. *JMIR Mental Health* 2020;**7**(6).
39. Li D, Chaudhary H, Zhang Z. Modeling spatiotemporal pattern of depressive symptoms caused by COVID-19 using social media data mining. *Int J Environ Res Publ Health* 2020;**17**(14).
40. Li S, Wang Y, Xue J, Zhao N, Zhu T. The impact of covid-19 epidemic declaration on psychological consequences: a study on active weibo users. *Int J Environ Res Publ Health* 2020;**17**(6).
41. Lin YH, Chiang TW, Lin YL. Increased internet searches for insomnia as an indicator of global mental health during the COVID-19 pandemic: multinational longitudinal study. *J Med Internet Res* 2020;**22**(9):e22181.
42. Low DM, Rumker L, Talkar T, Torous J, Cecchi G, Ghosh SS. Natural language processing reveals vulnerable mental health support groups and heightened health anxiety on reddit during COVID-19: observational study. *J Med Internet Res* 2020;**22**(10):e22635.
43. Misiak B, Szcześniak D, Koczanowicz L, Rymaszewska J. The COVID-19 outbreak and Google searches: is it really the time to worry about global mental health? *Brain Behav Immun* 2020;**87**:126–7.
44. Rana U, Singh R. Emotion analysis of Indians using google trends during COVID-19 pandemic. *Diabetes Metab Synd: Clin Res Rev* 2020;**14**(6):1849–50.
45. Richardson C, Patton M, Phillips S, Paslakis G. The impact of the COVID-19 pandemic on help-seeking behaviors in individuals suffering from eating disorders and their caregivers. *Gen Hosp Psychiatr* 2020;**67**:136–40.
46. Saha K, Torous J, Caine ED, de Choudhury M. Psychosocial effects of the COVID-19 pandemic: large-scale quasi-experimental study on social media. *J Med Internet Res* 2020;**22**(11).
47. Su Y, Xue J, Liu X, Wu P, Chen J, Chen C, et al. Examining the impact of COVID-19 lockdown in wuhan and Lombardy: a psycholinguistic analysis on weibo and twitter. *Int J Environ Res Publ Health* 2020;**17**(12).
48. Uvais NA. Association between the COVID-19 outbreak and mental health in India: a google trends study. *Prim Care Companion CNS Disord* 2020;**22**(6).
49. Uvais NA, Rasmina V. Analysis of psychiatry-related search engine trends during the COVID-19 pandemic in India. *Prim Care Companion CNS Disord* 2020;**22**(5).
50. Zhang B, Zaman A, Silenzio V, Kautz H, Hoque E. The relationships of deteriorating depression and anxiety with longitudinal behavioral changes in google and youtube use during COVID-19: observational study. *JMIR Mental Health* 2020;**7**(11).
51. Zhou L, Xie RH, Yang X, Zhang S, Li D, Zhang Y, et al. Feasibility and preliminary results of effectiveness of social media-based intervention on the psychological well-being of suspected COVID-19 cases during quarantine. *Can J Psychiatr Revue Canadienne de Psychiatrie* 2020;**65**(10):736–8.
52. Talbot J, Charron V, Konkle AT. Feeling the void: lack of support for isolation and sleep difficulties in pregnant women during the covid-19 pandemic revealed by twitter data analysis. *Int J Environ Res Publ Health* 2021;**18**(2):1–12.
53. Signorelli C, Odone A, Gianfredi V, Bossi E, Bucci D, Oradini-Alacreu A, et al. COVID-19 mortality rate in nine high-income metropolitan regions. *Acta Biomed* 2020;**91**(9-S):7–18.
54. Signorelli C, Odone A, Gianfredi V, Bossi E, Bucci D, Oradini-Alacreu A, et al. The spread of COVID-19 in six western metropolitan regions: a false myth on the excess of mortality in Lombardy and the defense of the city of Milan. *Acta Biomed* 2020;**91**(2):23–30.

55. Su Z, McDonnell D, Wen J, Kozak M, Abbas J, Segalo S, et al. Mental health consequences of COVID-19 media coverage: the need for effective crisis communication practices. *Glob Health* 2021;**17**(1):4.
56. Saud M, Mashud Mi, Ida R. Usage of social media during the pandemic: seeking support and awareness about COVID-19 through social media platforms. *J Publ Aff* 2020;**20**(4):e2417.
57. Provenzano S, Gianfredi V, Santangelo OE. Insight the data: wikipedia's researches and real cases of arboviruses in Italy. *Publ Health* 2021;**192**:21–9.
58. Gianfredi V, Santangelo OE, Provenzano S. Correlation between flu and Wikipedia's pages visualization. *Acta Biomed* 2021;**92**(1):e2021054.
59. Gianfredi V, Balzarini F, Gola M, Mangano S, Carpagnano LF, Colucci ME, et al. Leadership in public health: opportunities for young generations within scientific associations and the experience of the "academy of young leaders. *Front Publ Health* 2019;**7**:378.
60. Gianfredi V, Bragazzi NL, Nucci D, Martini M, Rosselli R, Minelli L, et al. Harnessing big data for communicable tropical and sub-tropical disorders: implications from a systematic review of the literature. *Front Publ Health* 2018;**6**:90.
61. Zhang T. Data mining can play a critical role in COVID-19 linked mental health studies. *Asian J Psychiatr* 2020;**54**:102399.
62. Al Zabadi H, Alhroub T, Yaseen N, Haj-Yahya M. Assessment of depression severity during coronavirus disease 2019 pandemic among the Palestinian population: a growing concern and an immediate consideration. *Front Psychiatr* 2020:11.
63. Al-Ajlouni YA, Park SH, Alawa J, Shamaileh G, Bawab A, El-Sadr WM, et al. Anxiety and depressive symptoms are associated with poor sleep health during a period of COVID-19-induced nationwide lockdown: a cross-sectional analysis of adults in Jordan. *BMJ Open* 2020;**10**(12):e041995.
64. Badr H, Oluyomi A, Adel Fahmideh M, Raza SA, Zhang X, El-Mubasher O, et al. Psychosocial and health behavioural impacts of COVID-19 pandemic on adults in the USA: protocol for a longitudinal cohort study. *BMJ Open* 2020;**10**(12):e044642.
65. Bendau A, Petzold MB, Pyrkosch L, Mascarell Maricic L, Betzler F, Rogoll J, et al. Associations between COVID-19 related media consumption and symptoms of anxiety, depression and COVID-19 related fear in the general population in Germany. *Eur Arch Psychiatr Clin Neurosci* 2020;**271**(2):283–91.
66. Deng CH, Wang JQ, Zhu LM, Liu HW, Guo Y, Peng XH, et al. Association of web-based physical education with mental health of college students in wuhan during the COVID-19 outbreak: cross-sectional survey study. *J Med Internet Res* 2020;**22**(10).
67. Guo Y, Cheng C, Zeng Y, Li Y, Zhu M, Yang W, et al. Mental health disorders and associated risk factors in quarantined adults during the COVID-19 outbreak in China: cross-sectional study. *J Med Internet Res* 2020;**22**(8).
68. Hasan MJ, Tabssum T, Ambia NE, Zaman MS, Rahman M, Khan AS. Mental health of the COVID-19 patients in Bangladesh. *Mymensingh Med J : MMJ* 2021;**30**(1):189–95.
69. Yan S, Xu R, Stratton TD, Kavcic V, Luo D, Hou F, et al. Sex differences and psychological stress: responses to the COVID-19 pandemic in China. *BMC Publ Health* 2021;**21**(1):79.
70. Xiong J, Lipsitz O, Nasri F, Lui LMW, Gill H, Phan L, et al. Impact of COVID-19 pandemic on mental health in the general population: a systematic review. *J Affect Disord* 2020;**277**:55–64.
71. The Lancet. *Global burden of disease study 2015 assesses the state of the world's health*. 2016. Available from: http://www.thelancet.com/gbd.
72. Shevlin M, McBride O, Murphy J, Miller JG, Hartman TK, Levita L, et al. Anxiety, depression, traumatic stress and COVID-19-related anxiety in the UK general population during the COVID-19 pandemic. *BJPsych Open* 2020;**6**(6):e125.
73. Rauschenberg C, Schick A, Hirjak D, Seidler A, Paetzold I, Apfelbacher C, et al. Evidence synthesis of digital interventions to mitigate the negative impact of the COVID-19 pandemic on public mental health: rapid meta-review. *J Med Internet Res* 2021;**23**(3):e23365.
74. Figueroa CA, Aguilera A. The need for a mental health technology revolution in the COVID-19 pandemic. *Front Psychiatr* 2020;**11**:523.
75. Marton K, Kanas N. Telehealth modalities for group therapy: comparisons to in-person group therapy. *Int J Group Psychother* 2016;**66**(1):145–50.
76. Nimrod G. Not good days for technophobes: older internet users during the COVID-19 pandemic. *Educ Gerontol* 2021;**47**(4):160–71.
77. Turgoose D, Ashwick R, Murphy D. Systematic review of lessons learned from delivering tele-therapy to veterans with post-traumatic stress disorder. *J Telemed Telecare* 2017;**24**(9):575–85.
78. Carr MJ, Steeg S, Webb RT, Kapur N, Chew-Graham CA, Abel KM, et al. Effects of the COVID-19 pandemic on primary care-recorded mental illness and self-harm episodes in the UK: a population-based cohort study. *Lancet Publ Health* 2021;**6**(2):e124–35.
79. Blithikioti C, Nuno L, Paniello B, Gual A, Miquel L. Impact of COVID-19 lockdown on individuals under treatment for substance use disorders: risk factors for adverse mental health outcomes. *J Psychiatr Res* 2021;**139**:47–53.
80. Nigatu YT, Elton-Marshall T, Wells S, Jankowicz D, Wickens CM, Hamilton HA. The association between COVID-19 diagnosis or having symptoms and anxiety among Canadians: a repeated cross-sectional study. *Hist Philos Logic* 2021:1–10.