RESEARCH ARTICLE

Posting patterns in peer online support forums and their associations with emotions and mood in bipolar disorder: Exploratory analysis

Glorianna Jagfeld[1]*, Fiona Lobban[1], Robert Davies[2], Ryan L. Boyd[2,3,4,5,6], Paul Rayson[7], Steven Jones[1]

1 Spectrum Centre for Mental Health Research, Division of Health Research, Lancaster University, Lancaster, United Kingdom, 2 Department of Psychology, Lancaster University, Lancaster, United Kingdom, 3 Security Lancaster, Lancaster University, Lancaster, United Kingdom, 4 Data Science Institute, Lancaster University, Lancaster, United Kingdom, 5 Obelus Institute, Behavioral Science Division, Washington D.C., United States of America, 6 Department of Computer Science, Stony Brook University, Stony Brook, NY, United States of America, 7 UCREL Research Centre, School of Computing and Communications, Lancaster University, Lancaster, United Kingdom

* g.jagfeld@lancaster.ac.uk

Data Availability Statement: There are ethical restrictions which prevent the public sharing of data utilized in the course of this study. Data are available upon request from Lancaster University

Abstract

Background

Mental health (MH) peer online forums offer robust support where internet access is common, but healthcare is not, e.g., in countries with under-resourced MH support, rural areas, and during pandemics. Despite their widespread use, little is known about who posts in such forums, and in what mood states. The discussion platform Reddit is ideally suited to study this as it hosts forums (subreddits) for MH and non-MH topics. In bipolar disorder (BD), where extreme mood states are core defining features, mood influences are particularly relevant.

Objectives

This exploratory study investigated posting patterns of Reddit users with a self-reported BD diagnosis and the associations between posting and emotions, specifically: 1) What proportion of the identified users posts in MH versus non-MH subreddits? 2) What differences exist in the emotions that they express in MH or non-MH subreddit posts? 3) How does mood differ between those users who post in MH subreddits compared to those who only post in non-MH subreddits?

Methods

Reddit users were automatically identified via self-reported BD diagnosis statements and all their 2005–2019 posts were downloaded. First, the percentages of users who posted only in MH (non-MH) subreddits were calculated. Second, affective vocabulary use was compared in MH versus non-MH subreddits by measuring the frequency of words associated with

positive emotions, anxiety, sadness, anger, and first-person singular pronouns via the LIWC text analysis tool. Third, a logistic regression distinguished users who did versus did not post in MH subreddits, using the same LIWC variables (measured from users' non-MH sub-reddit posts) as predictors, controlling for age, gender, active days, and mean posts/day.

Results

1) Two thirds of the identified 19,685 users with a self-reported BD diagnosis posted in both MH and non-MH subreddits. 2) Users who posted in both MH and non-MH subreddits exhibited less positive emotion but more anxiety and sadness and used more first-person singular pronouns in their MH subreddit posts. 3) Feminine gender, higher positive emotion, anxiety, and sadness were significantly associated with posting in MH subreddits.

Conclusions

Many Reddit users who disclose a BD diagnosis use a single account to discuss MH and other concerns. Future work should determine whether users exhibit more anxiety and sadness in their MH subreddit posts because they more readily post in MH subreddits when experiencing lower mood or because they feel more able to express negative emotions in these spaces. MH forums may reflect the views of people who experience more extreme mood (outside of MH subreddits) compared to people who do not post in MH subreddits. These findings can be useful for MH professionals to discuss online forums with their clients. For example, they may caution them that forums may underrepresent people living well with BD.

Introduction

Experiences of 'Manic, Mixed, or Hypomanic episodes or symptoms' alternating with 'Depressive episodes or periods of depressive symptoms' [1] that interfere with functioning, constitute core criteria for a bipolar disorder (BD) diagnosis [1, 2]. Prevalence rate estimates for bipolar spectrum disorders range from 0.1% (India), 2% (England), to 4.4% (US) across 22 different European, American, and Asian countries [3–5]. The characteristics and outcomes of people meeting BD criteria are diverse, with some functioning on a high level [6–9], achieving long-term remission [10] and living well [11–13]. However, many individuals with a BD diagnosis experience recurring mood episodes [14, 15] and require life-long treatment [16]. Moreover, BD has the highest rate of suicide across all mental health (MH) diagnoses [16, 17]. Therefore, it is important to understand how people living with BD can best be supported to live well.

The internet, including social media and online support forums, plays an increasingly important role in sharing information and support for health and MH concerns [18, 19]. Like the general population, most people with a (BD) diagnosis are internet users [20, 21] and seek information about BD online [22]. In 2014–15, 13%-17% of people with a BD, major depressive disorder, or psychotic disorder diagnosis reported posting and seeking information on their condition in online support forums or social media [23, 24]. Since then, the importance of online MH forums has likely increased. The online discussion platform Reddit (reddit.com) is one of the most visited internet sites worldwide and hosts subforums (subreddits) for a variety of topics, ranging from general interests to specific topics, including MH [25]. In 2019, the number of members in BD subreddits (bipolar, BipolarReddit, bipolar2, bipolarSOs) increased

by more than 50%. In light of the global Covid-19 pandemic, where in-person care has become more difficult to access for many people, international BD experts have highlighted the importance of online MH support and the "great need" to "train and support clinicians to go where people are already online" [26]. People may also use online forums in addition to in-person MH care since quantitative [27, 28] and qualitative [29, 30] evidence shows that online anonymity affords personal self-disclosures and discussions of sensitive and stigmatised issues.

There are a number of content analyses of BD online forums [31–37], including subreddits [38]. Jagfeld, Lobban, Rayson, et al. [39] characterised the demographics (age, gender identity, country of origin, comorbidities) of Reddit users with a self-reported BD diagnosis. However, little is known about the users of online MH support forums and why some users chose to engage in them and others not. This knowledge is both relevant to improve online MH support forum design and to contextualise the increasing body of research relating to such forums. Because users may post in MH-focused as well as other subreddits, Reddit is uniquely useful to study who chooses to engage in online MH support forums [40]. However, only Reddit users who post both in MH and non-MH subreddits can be considered for such within-person comparisons. Ireland and Iserman [40] noted that only one-third of Reddit users who had posted at least 50 words in anxiety subreddits had also posted at least 50 words in other subreddits; no similar studies exist for other MH concerns. Therefore, the first research question investigates what proportion of Reddit users with a self-reported BD diagnosis post in MH and non-MH focused subreddits using the same account.

Research using NLP methods with online posts of people with a BD diagnosis has mainly focused on language differences between users with a BD and other or no MH diagnosis and used these to build machine-learning classifiers to identify social media users likely experiencing BD symptoms [41]. While earlier studies compared all of users' content [42, 43], more recent work excluded MH-related posts [44–46] aiming to build classification systems that could also detect users with MH issues who do not discuss MH concerns online. Yet, to date, only one study has investigated differences between posts in MH and non-MH contexts of the *same* users. Ireland and Iserman [40] found that people used more anxious language in anxiety-focused subreddits compared to their posts in other subreddits. Such a comparison could provide insights into the extent to which findings from earlier studies that compared all of users' posts were due to language differences between people with MH and without MH issues or rather due to the different contexts of the posts in MH vs. non-MH subreddits that may impact on their emotional content. However, no comparable analyses exist for mood disorders or BD. Hence, the second research question examines what differences exist in the emotions that Reddit users with a BD diagnosis express in MH and non-MH subreddit posts.

Extreme and changing mood states are the core defining features of BD [1, 2]. Yet, very little research has been conducted so far on the relationship between mood and engaging with online MH forums for people with a BD diagnosis. Gathering data on peoples' emotions and mood is difficult. Surveys require people to fit their complex emotional experiences into rigid scales or describe them with a fixed set of preselected terms. Asking about past emotions and mood might suffer from recall bias [47, 48]. In general, any kind of self-report method may be biased by social desirability [49–51]. Moreover, in online forums, users may not be active anymore and therefore cannot be invited to active study participation.

Computational text analysis approaches such as Linguistic Inquiry and Word Count (LIWC) [52] can provide non-reactive behavioural indicators of emotions and mood that sidestep a lot of the problems with self-reported mood measures. LIWC provides indicators of peoples' affective and cognitive processes based on what percentage of the words in their language falls into categories such as positive emotion, sadness, and anxiety. A number of studies have applied LIWC to online posts of people with MH issues, including BD [42–45, 53, 54]. Park

and Conway [54] provide evidence using LIWC that Reddit users' negative emotions decrease, and positive emotions increase with every post they make in a depression support subreddit. Text emotion analysis with machine learning methods [55] has shown that users' emotional tone improves during exchanges in four large MH subreddits for depression, suicidality, BD, and anxiety when comparing their initial submission to their last comment in the same thread. As these are the only two studies looking at the relationship of emotions and mood and posting in online MH communities, the third research question addresses a critical gap in answering how mood differs between Reddit users with a BD diagnosis who post in MH subreddits and those who only post in non-MH subreddits.

Aim, research questions, and contributions

To summarise, this study aims to shed more light on the posting patterns in online support forums and their associations with the emotions and mood of people with a BD diagnosis. By applying psychological text analysis methods to the public online posts of a large sample of Reddit users with a self-reported BD diagnosis this exploratory study investigates the following research questions:

RQ 1: What proportion of Reddit users with a BD diagnosis posts in MH and non-MH subreddits?

RQ 2: What differences exist in the emotions that Reddit users with a BD diagnosis express in MH and non-MH subreddit posts?

RQ 3: How does mood differ between Reddit users with a BD diagnosis who post in MH subreddits and those who only post in non-MH subreddits?

This study makes the following contributions to MH research:

- Provides the first quantitative evidence on the posting behaviour of people with a BD diagnosis and how demographic and emotional factors may impact posting in MH forums specifically (see Fig 1 for a visual summary of the results)

- Demonstrates differences in the emotional content of MH vs non-MH subreddit posts

- Shows how to use established NLP methods to provide quantitative data and generate findings relevant for MH researchers and professionals

- Provides a large dataset of 21M Reddit posts by 20K Reddit users with a self-reported BD diagnosis that is available for future research

Methods

Reporting of this study follows the STROBE guidelines for cross-sectional studies [56].

Reddit user identification and inclusion and exclusion criteria

Reddit users who stated on the platform that they had received a BD diagnosis from a professional were automatically identified via self-reported diagnosis statements like 'I was diagnosed with bipolar disorder' [see 39]. All posts of candidate users (id, text, timestamp created at, subreddit) between 01/2005 (site inception) and 03/2019 (data available at time of data collection in 05/2019) were downloaded via the RedditAPI praw [57]. Candidate users were filtered by matching all of their posts against 74 exclusion patterns like 'not officially diagnosed' or 'self-diagnosed'. All posts of these users were also matched against self-reported diagnosis
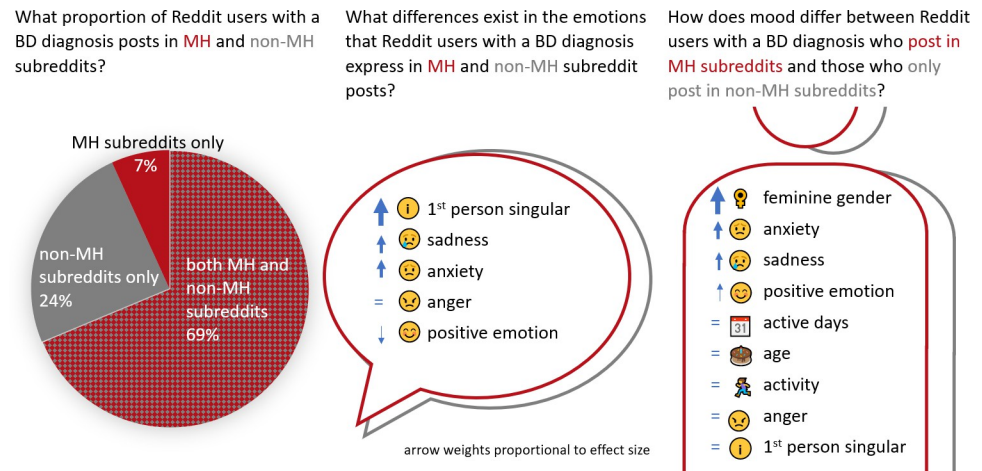
Fig 1. Visual summary of the results.

https://doi.org/10.1371/journal.pone.0291369.g001

statements for nine other MH diagnoses [39]. A major depressive disorder or psychotic disorder diagnosis constitutes an exclusion criterion for BD according to DSM-5 [2]. However, around 40% of people who are diagnosed with BD first received a major depressive disorder diagnosis [58, 59]. Therefore, only users who also reported a psychotic disorder diagnosis were removed. The Self-reported BD diagnosis (S-BiDD) dataset comprisesall posts of the remaining users.

RQ2 only concerned a subset of users with at least four posts of at least 25 words each in both MH and non-MH subreddits. Comparable studies of Reddit language using LIWC required at least four posts in depression subreddits [54] or five posts overall [60]. To address RQ3, these users who had posted in MH subreddits were compared against users who had no posts in MH subreddits and at least eight posts with at least 25 words each in non-MH subreddits. Users for which the NLP methods could not assign the control variables age or gender (see Section 'Statistical analyses' below) were excluded in the analyses to address RQ3.

MH subreddit identification

To determine whether posts were posted in subreddits with a MH focus, subreddit-to-topic assignments were amended and extended from existing resources as there is no official subreddit topic taxonomy from Reddit (see S1 Appendix for method details). The final four-level taxonomy [61] comprises 30,867 topic-categorised subreddits, including 158 MH-related subreddits, of which 37 are bipolar-specific. Only 3.6% of the posts by people with a self-reported BD diagnosis are in a subreddit not assigned to a topic. See S1 Table for an overview of the top 10 BD, MH (not BD-specific) and non-MH subreddits with most posts in the S-BiDD dataset.

Language analysis

LIWC [52] was used to obtain text-analytic measures of peoples' emotions and mood. In this study, users' emotions–usually short-lived (seconds [62]), high in intensity and in response to the evaluation of a stimulus [63]–are operationalised via the LIWC scores of single posts. Averaged LIWC scores of multiple posts over a longer timespan are assumed to represent users' mood, which is of longer duration, lower intensity and not in direct temporal relationship to a stimulus [63].

Recent NLP research focuses on emotion text analysis systems based on machine learning or deep learning methods [e.g., 64–66]. Yet, LIWC was considered the most suitable emotion analysis tool for this study for four main reasons. First, this type of language analysis via expert-created dictionaries is rooted in over 100 years of empirical research in the behavioural sciences and there is a large evidence base for the psychometric validity of this approach [e.g., 67–71]. Particularly for the LIWC emotion categories, Kahn et al. [72] demonstrated that the scores for positive emotion, negative emotion and sadness significantly differ in the expected ways in peoples' responses to discrete stimuli. The participants had higher positive emotion and lower negative emotion and sadness scores when writing about an amusing versus sad life experience and describing their feelings after exposure to a funny versus sad video clip. Newell et al. [73] showed that more negative emotion, less positive emotion, and a higher use of first-person singular pronouns correlates with peoples' subjectively reported stress level and their blood pressure. Moreover, LIWC emotion variables also correlate with established daily and seasonal mood fluctuations within individuals and for entire populations [74–76]. In contrast to this, machine learning-based NLP methods require labelled training data and usually only perform well on data that is similar to the training data. While Google recently released an emotion-labelled dataset of Reddit posts [65], it is not specific to people with MH issues and therefore not suitable to train or fine-tune an emotion classifier for the aims of this study. Second, LIWC emerged from Pennebaker's work on narrative therapy [e.g., 77], which is widely recognised amongst clinicians, while most of them are not familiar with machine learning and deep learning approaches. Applied psycholinguistic research using LIWC is very active as evidenced by many studies published just in the past two years [e.g., 78–81]. Third, LIWC yields results that are easy to interpret because they directly relate to the frequencies of words in the dictionaries. This is not the case for machine learning or deep learning systems whose predictions are often opaque and hard to trace back to the input [82]. Fourth, several previous studies have applied LIWC to online posts of people with MH issues, including BD [40, 42–45, 53, 54, 83], to the extent that most research on BD using NLP methods has used LIWC [41]. This allows for a straightforward comparison of the present results with previous work in the discussion or RQ2 and RQ 3.

Five LIWC categories were preselected for the analyses (see Table 1), including all affective processes subcategories (*positive emotion*, *anxiety*, *anger*, *sadness*). A number of previous studies used these categories to predict perceived negative/positive emotions from text [see 54]. Additionally, the linguistic variable *1st pers. sg.*, measuring the frequency of first person singular pronouns, was included because previous research pervasively links it to mental distress [84]. Only posts with at least 25 words were included in the analyses because LIWC estimates for shorter texts are less reliable [60, 85].

Table 1. LIWC variables related to emotions and mental distress selected for this study.

LIWC variable	Example vocabulary [52, Table 1]
Affective processes	
• Positive emotion (posemo)	love, nice, sweet
• Negative emotion	
○ Anxiety	worried, fearful
○ Anger	hate, kill, annoyed
○ Sadness	crying, grief, sad
Linguistic variables	
First person singular (1st pers. sg.)	I, me, mine

https://doi.org/10.1371/journal.pone.0291369.t001

Statistical analyses

To address RQ1, the percentages of users who posted only in MH subreddits, only in non-MH subreddits or in both were calculated. Previous work considered only Reddit users with at least four posts in MH subreddits as active contributors and those with fewer posts as *lurkers* [54]. Therefore, the number of users with at least four posts in both MH and non-MH subreddits was also calculated, and, for comparison, the number of users with at least eight posts who posted only in MH or non-MH subreddits.

To address RQ 2, the mean LIWC variable scores of the posts of users in MH subreddits were compared to the mean scores in non-MH subreddits via dependent t-tests with Cohen's d as the effect size measure [86].

To address RQ3, a logistic regression model was fitted via the R [87, Windows version 4.1.0] glm function to predict whether a user had posted in MH subreddits based on the mean LIWC variable scores of their non-MH subreddit posts and the control variables described below. Model fit for the model including only the control variables versus the model including control and LIWC variables was compared using a likelihood ratio test, following best practice guidance for linear mixed-effects models in psychological science [88]. See S2 Appendix for method details. The outcome variable *posted in MH* was coded as 1 for users who had posted in MH subreddits and 0 for those who had not (see 'Reddit user identification and inclusion and exclusion criteria'). Importantly, only non-MH subreddit posts were used to measure mood to provide a more comparable source of measures for users from both groups. LIWC emotion variables might be affected by the communicative context [89]. Only including posts made in non-MH subreddits avoided conflating mood and context.

The model also included the four control variables, age, gender, active days, and activity. *Age* was users' average age, considering all their posts in the S-BiDD dataset, as automatically inferred from their post texts via natural language processing (NLP) models [see 39]. There is some evidence for a decrease in perceived stigma, perceived discrimination, and stigmatising newspaper articles for MH issues in recent years [90–92]. Moreover, younger cohorts use social media differently than older adults. For example, 68% of 18–35 year old young adults with MH issues but only 54% of 36–65 year olds reported sharing personal experiences about MH issues as the main reason for their social media use [93]. Therefore, it appeared relevant to check whether younger users would be more likely to post in MH subreddits. *Gender* corresponded to users' binary gender as automatically inferred from their gender identity via self-reported statements like '*I'm 34M*' or their performed gender from the username or post texts via NLP models [see 39] (0 = feminine, 1 = masculine). Gender was included as a control variable because previous research found men were less likely to disclose and seek support for MH issues than women [94–97], although it is unknown how this plays out online. Finally, *active days* corresponded to the number of days between user's first and last post in the S-BiDD dataset and *activity* to the number of posts divided by *active days*. These two variables were included to check if users' overall activity on Reddit had an impact on their likelihood to post in MH support forums. For example, someone who is more active on Reddit generally and has been for a longer time may be more likely to have come across and post in a wider range of subreddits, including MH related ones.

Ethical considerations and transparency

The study is consistent with the Helsinki Declaration [98], the ethics guidelines for internet-mediated research of the British Psychological Society [99] and received approval from the Lancaster University Faculty of Health and Medicine Research Ethics Committee. Current ethical guidelines permit the waiving of explicit consent of internet users if only public content is

analysed without a threat to anonymity or confidentiality [e.g., 99]. In line with this, the present paper only shares aggregated results of anonymous Reddit users. See [39] for further ethical considerations in collecting the S-BiDD dataset. The purpose of this work is to better understand online users with a self-reported BD diagnosis and only comparisons within this group are conducted. Therefore, the methods and results presented here do not serve to identify online users with MH issues. People with lived experience of BD and using social media that were consulted did not raise any ethical concerns of this research (see Section Feedback from people with lived experience).

The S-BiDD dataset is available for non-commercial research [100] upon request via rdm@lancaster.ac.uk and under a data usage agreement that specifies ethical usage as detailed in [39]. The code is publicly available [101].

Involvement of people with lived experience

In preparation for this study, four volunteers with lived experience of BD and using online support forums recruited via PeopleInResearch were invited to share their relevant experiences and comment on the research plan in individual online meetings which lasted 30–60 minutes. The volunteers had a positive attitude towards the study as they were generally not aware of other work analysing online forum posts of people with a BD diagnosis. They agreed that analysis of public online posts was justifiable if users' anonymity was safeguarded. Most of the volunteers reported that they were posting in MH support forums when in low (but not very low) mood. Moreover, they all gradually reduced their engagement with online forums once they were better able to manage their BD and found a job or other projects they engaged with which may have been accompanied with more positive mood. Since sadness and anxiety were found to constitute core symptoms of low mood [102], this led to the following expectations: RQ 2: MH subreddit posts express more anxiety and sadness; RQ3: Users who post in MH subreddits have higher long-term measurements of anxiety and sadness. After completion of the analyses, the volunteers were contacted again to provide feedback on our interpretations of the results in individual online meetings.

Results

Results summary

The final Self-reported BD diagnosis (S-BiDD) dataset comprises 21,407,595 posts (both submissions (thread starts) and comments) of 19,685 users with a self-reported BD diagnosis. Fig 1 provides a visual summary of all results. Two thirds of the identified 19,685 users with a self-reported BD diagnosis posted in both MH and non-MH subreddits, while 24% never posted in MH subreddits. Users who posted in both MH and non-MH subreddits used more first-person singular pronouns and exhibited more sadness and anxiety, but less positive emotion in their MH subreddit posts. Feminine gender and expressing more anxiety, sadness, and positive emotion in non-MH subreddit posts were significant predictors for posting in MH subreddits. The following subsections present the detailed results for the three research questions.

RQ 1: What proportion of reddit users with a BD diagnosis posts in MH and non-MH subreddits?

Two thirds of the 19,685 identified Reddit users who disclosed a BD diagnosis posted in both MH and non-MH subreddits (Fig 2). One quarter never posted in MH subreddits and only 6.8% posted exclusively in MH subreddits. The majority (54.7%) of users in the S-BiDD dataset were active contributors in both MH and non-MH subreddits. Most users with less than eight
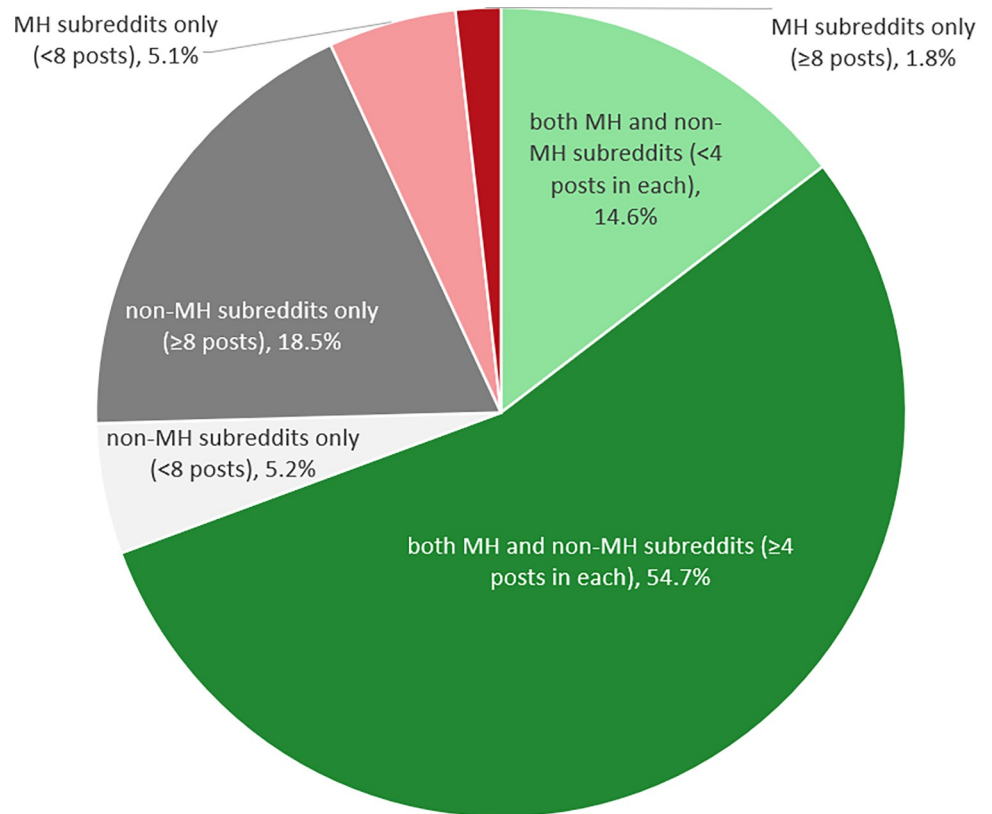
Fig 2. Reddit users with a self-reported BD diagnosis (n = 19,685) according to the subreddit types in which they post.

posts posted in both MH and non-MH subreddits (14.6%) and equally likely posted only in MH/non-MH subreddits (5.1%/5.2%).

RQ 2: What differences exist in the emotions that reddit users with a BD diagnosis express in MH and non-MH subreddit posts?

N = 9,821 users (49.9% of total) with 6,493,626 posts of at least 25 words (11.4% in MH subreddits) met the criteria set out in the 'Reddit user identification and inclusion and exclusion criteria' section as active contributors in both MH and non-MH subreddits with at least four posts in each. There is a significant difference between users' posts in MH and non-MH subreddits for each preselected LIWC variable (see Table 2). Users exhibited slightly less positive emotion in their MH subreddit posts but more anxiety and sadness, and they used much more first-person singular pronouns.

Table 2. Within user differences between MH and non-MH subreddit posts.

	MH posts		Non-MH posts		Dependent t-test, *** = p < 0.001 with Bonferroni correction for 5 comparisons		
	mean	std	mean	std	P	effect size (Cohen's d)	effect size interpretation [103]
positive emotion	3.48	1.08	3.64	0.93	< .001	-0.20	Small
anxiety	0.67	0.42	0.37	0.22	< .001	0.97	Large
anger	0.75	0.49	0.78	0.44	< .001	-0.07	Very small
sadness	0.81	0.44	0.45	0.24	< .001	1.11	Large
1st pers. sg.	8.82	2.28	6.39	1.96	< .001	1.63	Very large

Table 3. Model fit comparison for controls only and controls with LIWC variables.

Model name	Nested / simpler Model	Predictors added	Model fit				LRT against nested	
			AIC	BIC	LL	df	df	X^2
Controls	-	age, gender, active days, activity	10464	10500	-5227	10153		
Controls + LIWC	Controls	posemo, anxiety, anger, sadness, 1ˢᵗ pers. sg.	10403	10476	-5192	10148	5	70.12, $P < .001$

N subjects = 10,309; table key: LRT: likelihood ratio test, AIC: Akaike Information Criterion, BIC: Bayesian Information Criterion, LL: log likelihood, df: degrees of freedom, X^2: Analysis of deviance (deviance $\boldsymbol{X^2}$, $\boldsymbol{P}$ ($>\boldsymbol{X^2}$))

RQ 3: How does mood differ between users who post in MH subreddits and those who only post in non-MH subreddits?

N = 10,158 users (51.6% of total) met the criteria stated in the 'Reddit user identification and inclusion and exclusion criteria' section and were included in the regression model. Of these 2,312 (22.8%) had only posted in non-MH subreddits. S2 Table provides descriptive statistics of the dependent variables. Logistic regression assumptions were checked according to Field et al. [104] (see S3 Appendix for details). The regression model including the LIWC variables as predictors fit the data better than the model with the control variables only (see Table 3).

According to the regression results for the full model in Table 4, Reddit users with feminine gender are 33% more likely to post in MH subreddits compared to users with masculine gender (see S4 Appendix for how this was calculated from the odds ratios of the coefficients). Moreover, higher levels of mean positive emotions, anxiety, and sadness were associated with increased odds to post in MH subreddits. Although the timespan for how long users were active on Reddit was a significant predictor, the coefficient is very small, thus practically negligible. Users' age, their average posts per day, anger, and first person singular pronoun use were not significantly related to whether someone posted in MH subreddits.

Table 4. Results for the glm regression model including controls and LIWC variables.

Model Controls + LIWC	Est/Beta	SE	95% CI	z	P	Odds ratio (exp(coeff))	
						Est/Beta	95% CI
(Intercept)	0.56	0.17	0.22–0.90	3.21	0.001	1.75	1.24–2.47
age	$< .001$	$< .001$	-0.01 – $< .001$	-0.24	0.81	1.00	0.99–1.00
gender	-0.79	0.06	-0.90 – -0.68	-14.33	$< .001$	0.45	0.41–0.51
active days	$< .001$	$< .001$	$< .001$– $< .001$	11.18	$< .001$	1.00	1.00–1.00
activity	-0.01	0.01	-0.02–0.01	-1.17	0.24	0.99	0.98–1.01
posemo	0.05	0.01	0.03–0.08	4.94	$< .001$	1.06	1.03–1.08
anxiety	0.62	0.15	0.32–0.91	4.08	$< .001$	1.85	1.39–2.51
anger	-0.09	0.04	-0.17 – -0.01	-2.10	0.04	0.92	0.85–0.99
sadness	0.41	0.12	0.18–0.63	3.49	$< .001$	1.50	1.20–1.89
1ˢᵗ pers. sg.	0.02	0.02	-0.02–0.05	0.97	0.33	1.02	0.98–1.05
			Model fit				
Pseudo R^2		Hosmer and Lemeshow				Cox and Snell	Nagelkerke
		0.047				0.049	0.075

Key: p-values for fixed effects calculated using Satterthwaites approximations.

Confidence Intervals have been calculated using the Wald method.

Model equation: glm(posted_in_MH ~ age + gender + active days + activity + posemo + anxiety + anger + sadness + 1ˢᵗ pers. sg., family = binomial(link = "logit"))

Feedback from people with lived experience

Two of the four people with lived experience who commented on the study design agreed to provide feedback on the results. Overall, the results resonated with their experience of using online support forums, particularly the finding that people who experienced more anxiety and sadness were more likely to post in these forums. One of them shared that initially she had the expectation that online support forums would portray a holistic picture of living with BD. However, over time she noticed a strong focus on difficult experiences in the posts while users who posted that they were doing well got almost no reactions. Both volunteers agreed that it might be beneficial to encourage more discussions around living well with BD and aspects that can make life meaningful to balance out the strong focus on symptoms and diagnosis that they perceived in the forums. Neither of them reported concerns about potential for misuse of the research methods or misinterpretation of the results in a manner that could be harmful for people living with BD.

Discussion

Strengths of this study

This study shed more light on the posting behaviour of Reddit users with a self-reported BD diagnosis and provides the first quantitative evidence on the relationship between mood and posting in BD support forums. A strength of this study was The automatic user identification via self-reported diagnosis statements, which resulted in a large sample of participants. Since the data collection did not require direct interaction with the social media users, there was no researcher influence on the behaviour of the participants. Moreover, the psychological text analysis tool LIWC provided easily interpretable measures for in-the-moment emotion assessments, unbiased by self-report. Its wide use facilitated comparison with related work and its psychometric validity to measure emotions has been well established [71]. However, due to a lack of previous research, the study was necessarily exploratory. Future research should be carried out to confirm the findings and explore possible causal relationships between posting behaviour and mood.

RQ1: What proportion of reddit users with a BD diagnosis posts in MH and non-MH subreddits?

RQ 1 results showed that most (> 50%) Reddit users with a BD diagnosis frequently post both in MH and non-MH subreddits. This exceeds the one third of Reddit users who had posted at least 50 words in both anxiety and non-anxiety-related subreddits identified by Ireland and Iserman [40].

A specific feature of Reddit are throwaway accounts. Since Reddit accounts only require a username for registration, users may create single-use online identities to discuss sensitive topics such as MH issues even more anonymously than via the main account through which they usually engage on the platform [27]. Separating an online identity from one's offline identity or a more holistic online identity (the main Reddit account) may be a main factor of online disinhibition [105, 106], causing people to act differently than in identifiable online settings [107]. De Choudhury and De [27] found that 4.5% of the 1,209 Reddit users who had posted in one of eight MH subreddits in their dataset used a throwaway account as identified by usernames matching *throw*; 61% of these user accounts had made exactly one post. Similarly, in the present study, 3.9% of users had made exactly one post. Moreover, user accounts with less than eight posts equally likely posted only in MH or non-MH subreddits. This suggests that single-use accounts are not necessarily favoured for posting about MH issues.

The fact that so few Reddit users with a self-reported BD diagnosis seem to make use of throwaway accounts and (by inference) to compartmentalise their Reddit identity into only MH vs. non-MH-related concerns is encouraging for further research on the S-BiDD dataset. While [40] considered users with anxiety concerns who post in both MH and non-MH subreddits to be atypical, this does not seem to be the case for users who disclose a BD diagnosis. The present findings imply that most user profiles with a self-disclosed BD diagnosis can be identified as users' main profiles: these profiles are associated with posts covering several aspects of their interests and concerns, which may more closely reflect them compared to throwaway accounts.

The RQ 1 findings also demonstrate that the eligibility criteria for the subsequent RQs that require users to have at least four posts in MH and non-MH subreddits are not too restrictive but capture more than half of the Reddit users who disclose a BD diagnosis.

RQ 2: What differences exist in the emotions that Reddit users with a BD diagnosis express in MH and non-MH subreddit posts?

The comparison between users' posts in MH and non-MH subreddit posts showed that, in MH subreddits, users expressed somewhat fewer positive emotions, more anxiety and sadness, and used much more first-person singular pronouns. These findings aligned with the expectations of people with lived experience that people may mainly post in MH support forums about MH struggles and when experiencing mental distress. They also aligned with the finding from Ireland and Iserman [40] that people used more anxious language in anxiety-focused subreddits compared to their posts in other subreddits. Increased use of first person singular pronouns has been mainly linked to mental distress [84], although positive I-talk may also occur [84].

The results demonstrate a higher presence of negative emotions and self-focus in MH subreddit posts, which previous research links to rumination and low mood [108, 109]. However, this analysis cannot determine why these differences occur. Possible explanations that future research needs to test are that Reddit users in a lower mood state may be more likely to post in MH subreddits or that they may feel more comfortable expressing negative emotions there.

RQ 3: How does mood differ between users who post in MH subreddits and those who only post in non-MH subreddits?

Users who exhibited more anxiety and sadness in their non-MH subreddit posts, behaviours that constitute core symptoms of low mood [102], were more likely to post in MH subreddits. This was expected by people with lived experience who reported that they mainly used online support forums when experiencing low mood. Interestingly, exhibiting more positive emotions also increased the odds of posting in MH subreddits, although the effect was an order of magnitude smaller than for anxiety and sadness. Taken together, these findings could indicate that users who experienced more extreme mood, hence stronger BD symptoms, are more likely to post in MH subreddits.

Another finding was that users with a feminine gender identity were 33% more likely to post in MH subreddits compared to users with a masculine gender identity. While biological men and women are equally likely to meet BD criteria [2, 5], Jagfeld et al. [39] already established that slightly more than half of the Reddit users who self-report a BD diagnosis had a feminine gender, although two thirds of Reddit users in general are male [110]. Of the users who never posted in MH subreddits, 60.5% had a masculine gender, while 59.2% of the users who posted in MH subreddits had a feminine gender. This proportion is in line with an average of 58% of women enrolled in 55 trials included in a large meta-analysis of psychological

interventions for BD [111]. Overall, the finding that users with a feminine gender identity are more likely to post in MH forums is in line with previous research that found men were less likely to disclose and seek support for MH issues [94–96].

Also using LIWC, De Choudhury et al. [83] found that feminine-identified users who disclosed MH issues on Twitter exhibit more anxiety and sadness in their posts and less positive emotions compared to masculine-identified users who disclosed MH issues. This mirrors our findings on the differences between Reddit users who post or do not post in MH subreddits. To check whether our findings of mood differences between Reddit users who post versus do not post in MH subreddits could be attributed to the gender imbalance in the outcome groups, an additional regression model was evaluated with gender-balanced samples for both outcome groups. However, the results for the gender-balanced sample of 8,092 users did not change substantially compared to the original sample (see S5 Appendix).

Implications

Implications for research.   In general, a better understanding of people with a BD diagnosis who post on Reddit is important to contextualise other research on this kind of data, with Reddit being one of the most frequently used data sources for research on BD using NLP methods [41]. More specifically, the findings for RQ 2 demonstrated substantial differences in LIWC scores between posts in MH and non-MH subreddits. Although expected, this had not been empirically tested so far. Providing quantitative evidence for this is important because what seems 'common sense' is actually only so when it has been empirically demonstrated [112]. The RQ2 findings provide strong support for all studies relying on emotion text analysis, specifically comparative language analyses and detection systems of people with and without MH issues, to only consider posts in a comparable context as done in more recent research [44–46], compared to earlier work [42, 43]. For studies that consider recruiting participants from online support forums, it is relevant to know that mainly users who may experience stronger mood symptoms tend to be active there. There is also a higher proportion of feminine-identified users who post in these forums, which is similar to the proportion of women recruited into clinical trials. Relying only on online support forums for recruitment may thus introduce sampling biases.

Implications for practice.   In general, this study provides information to MH professionals about online support forums for BD–something for which international BD experts highlighted a currently critical need [26]. Since the analyses aggregated several thousands of Reddit users and found mostly small to medium effect sizes, the findings are relevant on the group level and do not permit conclusions for individual people. It may be helpful to inform people living with BD who (want to) use online support forums that people living well with MH issues might be underrepresented there and recommend additional information sources to obtain a more balanced picture of living with BD. Forum providers and moderators might consider encouraging more discussions around living well with BD to potentially balance some of the content exhibiting higher levels of anxiety and sadness. On the other hand, online support forums may be a useful outlet where people living with BD feel safe and welcome to express their more intense emotions, particularly anxiety and sadness. Spaces where people can express intense negative emotions are important because they counteract suppression and avoidance of such emotions. These are considered maladaptive emotion-regulation strategies and have been linked to depression and anxiety [113].

Limitations and future work

This study has three limitations that future research could address. It was not possible to check the veracity of self-reported diagnoses, although it appears unlikely that people would

deliberately unfaithfully identify with a stigmatised MH condition. Cluss et al. [114] verified that 93% of 804 individuals who self-reported BD diagnosis met criteria for a lifetime BD diagnosis according to an in-person structured diagnostic interview. Nevertheless, future complementary research could conduct diagnostic interviews to determine the eligibility of participants, which would also allow inclusion of people who meet criteria for BD but have not received a diagnosis yet. LIWC is only one tool to study the mood of social media users. Future work could triangulate the present results with emotion text analysis systems trained via machine learning that may be more accurate in context but less interpretable than the dictionary-based LIWC tool [82], as well as with self-reported questionnaires and qualitative interviews. Finally, his study only considered data until March 2019. While online MH forum usage has likely increased since then, fuelled by the global Covid-19 pandemic [115], there are no apparent reasons why the usage patterns found here would have changed. However, future research could replicate these analyses with more recent data and consider changes over time. Future research should also explore whether the present findings are specific to BD or mood or MH issues more general.

Additionally, future research could extend the present findings by looking at what effects engaging with online MH support forums has on users' mood as well as on person-centred holistic wellbeing in longitudinal studies. Recommendations for people living with BD for how to use social media and online support forums to achieve good long-term quality of life [116] or personal recovery [117] was something that people with lived experience of BD considered as particularly helpful to develop. The ongoing Improving Peer Online Forums study, which aims to find out how online MH forums work, why some work better than others, and why some people find them helpful and others do not, may yield promising evidence for this [118].

Conclusions

This exploratory study was the first quantitative investigation to provide information on forum posting behaviour and its relationship with emotions and mood in Reddit users with a self-reported BD diagnosis. The study had three main findings: First, Reddit users with a self-reported BD diagnosis frequently post in both MH and non-MH subreddits. Second, for the same users, there are large differences in the expressed emotions in MH versus non-MH subreddit posts. Third, feminine gender, higher anxiety, sadness and positive emotions were associated with significantly increased odds that a Reddit user posted to MH subreddits. These findings have important implications for research on Reddit and online MH forums in general, as well as for the design of online MH forums, and for recommendations concerning their use to people living with BD.

Supporting information

S1 Table. Top 10 subreddits with most posts in the dataset for subreddits in the BD or MH subreddit list (excl. BD) and subreddits not in the MH subreddit list. The number of non-MH subreddits is the number of unique subreddits in which users in the dataset posted that are not in the MH subreddit list. The users in the dataset only posted in 36 of the 37 pre-identified BD-specific subreddits and in 116 of the 121 (158–37) not BD-specific subreddits in the MH subreddit list.
(DOCX)

S2 Table. Descriptive statistics and between user differences for users in the regression model.
(DOCX)

S1 Appendix. Details on the creation of the BD and MH subreddit lists.
(DOCX)

S2 Appendix. Method details for the logistic regression model with controls and LIWC variables.
(DOCX)

S3 Appendix. Assumptions checking for the logistic regression model with controls and LIWC variables.
(DOCX)

S4 Appendix. Probability calculation for the gender coefficient.
(DOCX)

S5 Appendix. Regression results for subsampled dataset with gender-balanced outcome groups.
(DOCX)

Acknowledgments

Author Contributions

Conceptualization: Glorianna Jagfeld, Fiona Lobban, Paul Rayson, Steven Jones.

Data curation: Glorianna Jagfeld.

Formal analysis: Glorianna Jagfeld.

Funding acquisition: Fiona Lobban, Paul Rayson, Steven Jones.

Investigation: Glorianna Jagfeld.

Methodology: Glorianna Jagfeld, Fiona Lobban, Robert Davies, Ryan L. Boyd, Paul Rayson, Steven Jones.

Project administration: Glorianna Jagfeld.

Resources: Glorianna Jagfeld, Ryan L. Boyd.

Software: Glorianna Jagfeld.

Supervision: Fiona Lobban, Paul Rayson, Steven Jones.

Validation: Glorianna Jagfeld.

Visualization: Glorianna Jagfeld.

Writing – original draft: Glorianna Jagfeld.

Writing – review & editing: Glorianna Jagfeld, Fiona Lobban, Robert Davies, Ryan L. Boyd, Paul Rayson, Steven Jones.

References

1. World Health Organization. Bipolar or related disorders. International classification of diseases for mortality and morbidity statistics (11th revision). 2018. Available: https://icd.who.int/browse11/l-m/en#/http%3A%2F%2Fid.who.int%2Ficd%2Fentity%2F613065957

2. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, fifth edition. 2013.

3. McManus S, Bebbington P, Jenkins R, Brugha T. Mental Health and Wellbeing in England: Adult Psychiatric Morbidity Survey 2014. NHS Digit. 2016; 1–405. https://dx.doi.org/10.1302/0301-620X.95B2.28383

4. Merikangas KR, Jin R, He J, Kessler RC, Lee S, Sampson NA, et al. Prevalence and correlates of bipolar spectrum disorder in the world mental health survey initiative. Arch Gen Psychiatry. 2011; 68: 241–251. https://doi.org/10.1001/archgenpsychiatry.2011.12 PMID: 21383262

5. Pini S, De Queiroz V, Pagnin D, Pezawas L, Angst J, Cassano GB, et al. Prevalence and burden of bipolar disorders in European countries. Eur Neuropsychopharmacol. 2005; 15: 425–434. https://doi.org/10.1016/j.euroneuro.2005.04.011 PMID: 15935623

6. Coryell W, Turvey C, Endicott J, Leon AC, Mueller T, Solomon D, et al. Bipolar I affective disorder: Predictors of outcome after 15 years. J Affect Disord. 1998; 50: 109–116. https://doi.org/10.1016/s0165-0327(98)00043-3 PMID: 9858070

7. Goldberg JF, Harrow M. Consistency of remission and outcome in bipolar and unipolar mood disorders: A 10-year prospective follow-up. J Affect Disord. 2004; 81: 123–131. https://doi.org/10.1016/S0165-0327(03)00161-7 PMID: 15306137

8. Tohen M, Zarate CA, Hennen J, Khalsa HMK, Strakowski SM, Gebre-Medhin P, et al. The McLean-Harvard first-episode mania study: Prediction of recovery and first recurrence. Am J Psychiatry. 2003; 160: 2099–2107. https://doi.org/10.1176/appi.ajp.160.12.2099 PMID: 14638578

9. Akers N, Lobban F, Hilton C, Panagaki K, Jones SH. Measuring social and occupational functioning of people with bipolar disorder: A systematic review. Clin Psychol Rev. 2019; 74. https://doi.org/10.1016/j.cpr.2019.101782 PMID: 31751878

10. Goldberg JF, Garno JL, Harrow M. Long-term Remission and Recovery in Bipolar Disorder: A Review. Curr Psychiatry Rep. 2005; 7: 456–461. https://doi.org/10.1007/s11920-005-0067-7 PMID: 16318824

11. Warwick H, Tai S, Mansell W. Living the life you want following a diagnosis of bipolar disorder: A grounded theory approach. Clin Psychol Psychother. 2019. https://doi.org/10.1002/cpp.2358 PMID: 30746808

12. Russell SJ, Browne JL. Staying well with bipolar. Aust New Zeal J Psychiatry. 2005; 39: 187–193.

13. Crowe M, Inder M. Staying well with bipolar disorder: A qualitative analysis of five-year follow-up interviews with young people. J Psychiatr Ment Heal Nurs (John Wiley Sons, Inc). 2018; 25: 236–244. https://doi.org/10.1111/jpm.12455 PMID: 29421858

14. Treuer T, Tohen M. Predicting the course and outcome of bipolar disorder: A review. Eur Psychiatry. 2010; 25: 328–333. https://doi.org/10.1016/j.eurpsy.2009.11.012 PMID: 20444581

15. Gignac A, McGirr A, Lam RW, Yatham LN. Recovery and recurrence following a first episode of mania: a systematic review and meta-analysis of prospectively characterized cohorts. J Clin Psychiatry. 2015; 76: 1241–1248. https://doi.org/10.4088/JCP.14r09245 PMID: 25845021

16. Goodwin GM, Haddad PM, Ferrier IN, Aronson JK, Barnes TRH, Cipriani A, et al. Evidence-based guidelines for treating bipolar disorder: Revised third edition recommendations from the British Association for Psychopharmacology. J Psychopharmacol. 2016; 30: 495–553. https://doi.org/10.1177/0269881116636545 PMID: 26979387

17. Novick DM, Swartz HA, Frank E. Suicide attempts in bipolar I and bipolar II disorder: a review and meta-analysis of the evidence. Bipolar Disord. 2010; 12: 1–9. https://doi.org/10.1111/j.1399-5618.2009.00786.x PMID: 20148862

18. Sundar SS, Rice RE, Kim H, Sciamanna CN. Online health information: conceptual challenges and theoretical opportunities. 2nd ed. In: Thompson TL, Parrott R, Nussbaum JF, editors. The Routledge Handbook of Health Communication. 2nd ed. Taylor & Francis Group; 2011. pp. 181–202.

19. Bujnowska-Fedak MM, Waligóra J, Mastalerz-Migas A. The internet as a source of health information and services. Adv Exp Med Biol. 2019; 1211: 1–16. https://doi.org/10.1007/5584_2019_396 PMID: 31273574

20. Martini T, Czepielewski LS, Fijtman A, Sodré L, Wollenhaupt-Aguiar B, Pereira CS, et al. Bipolar disorder affects behavior and social skills on the internet. PLoS One. 2013; 8. https://doi.org/10.1371/journal.pone.0079673 PMID: 24244541

21. Bauer R, Conell J, Glenn T, Alda M, Ardau R, Baune BT, et al. Internet use by patients with bipolar disorder: Results from an international multisite survey. Psychiatry Res. 2016; 242: 388–394. https://doi.org/10.1016/j.psychres.2016.05.055 PMID: 27391371

22. Conell J, Bauer R, Glenn T, Alda M, Ardau R, Baune BT, et al. Online information seeking by patients with bipolar disorder: results from an international multisite survey. Int J Bipolar Disord. 2016; 4. https://doi.org/10.1186/s40345-016-0058-0 PMID: 27552813

23. Bauer R, Conell J, Glenn T, Alda M, Ardau R, Baune BT, et al. International multi-site survey on the use of online support groups in bipolar disorder. Nord J Psychiatry. 2017; 71: 473–476. https://doi.org/10.1080/08039488.2017.1334819 PMID: 28696841

24. Naslund JA, Aschbrenner KA, Bartels SJ. How People with Serious Mental Illness Use Smartphones, Mobile Apps, and Social Media. Psychiatr Rehabil J. 2016; 39: 364–367. https://doi.org/10.1037/prj0000207 PMID: 27845533

25. Alexa Internet. reddit.com. 2020 [cited 27 Feb 2020]. Available: https://www.alexa.com/siteinfo/reddit.com

26. Stefana A, Youngstrom EA, Chen J, Hinshaw S, Maxwell V, Michalak E, et al. The COVID-19 pandemic is a crisis and opportunity for bipolar disorder. Bipolar Disord. 2020; 22: 641–643. https://doi.org/10.1111/bdi.12949 PMID: 32511859

27. De Choudhury M, De S. Mental Health Discourse on reddit: Self-Disclosure, Social Support, and Anonymity. Proc 8th Int AAAI Conf Weblogs Soc Media. 2014; 71–80.

28. Pavalanathan U, De Choudhury M. Identity management and mental health discourse in social media. WWW 2015 Companion—Proc 24th Int Conf World Wide Web. 2015; 315–321. https://doi.org/10.1145/2740908.2743049 PMID: 27376158

29. Wright K. Communication in health-related online social support groups/communities: A review of research on predictors of participation, applications of social support theory, and health outcomes. Rev Commun Res. 2016; 4: 65–87. https://doi.org/10.12840/issn.2255-4165.2016.04.01.010

30. Smith-Merry J, Goggin G, Campbell A, McKenzie K, Ridout B, Baylosis C. Social connection and online engagement: Insights from interviews with users of a mental health online forum. JMIR Ment Heal. 2019; 6: 1–13. https://doi.org/10.2196/11084 PMID: 30912760

31. McDonald D, Woodward-Kron R. Member roles and identities in online support groups: Perspectives from corpus and systemic functional linguistics. Discourse Commun. 2016; 10: 157–175. https://doi.org/10.1177/1750481315615985

32. Kramer ADI, Fussell SR, Setlock LD. Text analysis as a tool for analyzing conversation in online support groups. Conf Hum Factors Comput Syst—Proc. 2004; 1485–1488. https://doi.org/10.1145/985921.986096

33. Vayreda A, Antaki C. Social support and unsolicited advice in a bipolar disorder online forum. Qual Health Res. 2009; 19: 931–942. https://doi.org/10.1177/1049732309338952 PMID: 19556400

34. Latalova K, Prasko J, Kamaradova D, Ivanova K, Jurickova L. Bad on the net, or bipolars' lives on the web: Analyzing discussion web pages for individuals with bipolar affective disorder. Neuroendocrinol Lett Lett. 2014; 35: 206–212. Available: http://www.nel.edu PMID: 24977970

35. Poole R, Smith D, Simpson S. How Patients Contribute to an Online Psychoeducation Forum for Bipolar Disorder: A Virtual Participant Observation Study. JMIR Ment Heal. 2015; 2. https://doi.org/10.2196/mental.4123 PMID: 26543925

36. Bauer R, Bauer M, Spiessl H, Kagerbauer T. Cyber-support: An analysis of online self-help forums (online self-help forums in bipolar disorder). Nord J Psychiatry. 2013; 67: 185–190. https://doi.org/10.3109/08039488.2012.700734 PMID: 22817138

37. Campbell IH, Campbell H. Ketosis and bipolar disorder: controlled analytic study of online reports. BJPsych Open. 2019; 5: 1–6. https://doi.org/10.1192/bjo.2019.49 PMID: 31530294

38. Sahota PKC, Sankar PL. Bipolar Disorder, Genetic Risk, and Reproductive Decision-Making: A Qualitative Study of Social Media Discussion Boards. Qual Health Res. 2019. https://doi.org/10.1177/1049732319867670 PMID: 31409193

39. Jagfeld G, Lobban F, Rayson P, Jones SH. Understanding who uses Reddit: Profiling individuals with a self-reported bipolar disorder diagnosis. Proceedings of the Seventh Workshop on Computational Linguistics and Clinical Psychology: Improving Access at NAACL 2021. 2021. pp. 1–14.

40. Ireland M, Iserman M. Within and Between-Person Differences in Language Used Across Anxiety Support and Neutral Reddit Communities. Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic,. 2018. pp. 182–193. https://doi.org/10.18653/v1/w18-0620

41. Harvey D, Lobban F, Rayson P, Warner A, Jones S. Natural Language Processing Methods and Bipolar Disorder: Scoping Review. JMIR Ment Heal. 2022; 9. https://doi.org/10.2196/35928 PMID: 35451984

42. Coppersmith G, Dredze M, Harman C. Quantifying Mental Health Signals in Twitter. Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality. 2014. pp. 51–60. https://doi.org/10.3115/v1/W14-3207

43. Coppersmith G, Dredze M, Harman C, Hollingshead K. From ADHD to SAD: Analyzing the Language of Mental Health on Twitter through Self-Reported Diagnoses. Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality. 2015. https://doi.org/10.3115/v1/w15-1201

44. Cohan A, Desmet B, Yates A, Soldaini L, MacAvaney S, Goharian N. SMHD: A Large-Scale Resource for Exploring Online Language Usage for Multiple Mental Health Conditions. Proceedings of the 27th International Conference on Computational Linguistics (COLING). Santa Fe; 2018. pp. 1485–1497. Available: https://www.aclweb.org/anthology/C18-1126

45. Sekulić I, Gjurković M, Šnajder J. Not Just Depressed: Bipolar Disorder Prediction on Reddit. 9th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis (WASSA). Brussels; 2018. pp. 72–78. https://doi.org/10.18653/v1/P17

46. Sekulić I, Strube M. Adapting Deep Learning Methods for Mental Health Prediction on Social Media. Proceedings of the 2019 EMNLP Workshop W-NUT: The 5th Workshop on Noisy User-generated Text. 2019. pp. 322–327. https://doi.org/10.18653/v1/d19-5542

47. Thomas DL, Diener E. Memory accuracy in the recall of emotions. J Pers Soc Psychol. 1990; 59: 291–297.

48. Colombo D, Suso-Ribera C, Fernández-Álvarez J, Cipresso P, Garcia-Palacios A, Riva G, et al. Affect Recall Bias: Being Resilient by Distorting Reality. Cognit Ther Res. 2020; 44: 906–918. https://doi.org/10.1007/s10608-020-10122-3

49. Tanaka-Matsumi J, Kameoka VA. Reliabilities and Concurrent Validities of Popular Self-Report Measures of Depression, Anxiety, and Social Desirability. J Consult Clin Psychol. 1986; 54: 328–333. https://doi.org/10.1037//0022-006x.54.3.328 PMID: 3722561

50. Barrett LF. Hedonic Tone, Perceived Arousal, and Item Desirability: Three Components of Self-reported Mood. Cogn Emot. 1996; 10: 47–68. https://doi.org/10.1080/026999396380385

51. Van de Mortel TF. Faking it: Social desirability response bias in selfreport research. Aust J Adv Nurs. 2008; 25: 40–48.

52. Pennebaker JW, Boyd RL, Jordan K, Blackburn K. The Development and Psychometric Properties of LIWC2015. Austin; 2015. https://doi.org/10.15781/T29G6Z

53. Fineberg SK, Leavitt J, Deutsch-Link S, Dealy S, Landry CD, Pirruccio K, et al. Self-reference in psychosis and depression: A language marker of illness. Psychol Med. 2016; 46: 2605–2615. https://doi.org/10.1017/S0033291716001215 PMID: 27353541

54. Park A, Conway M. Longitudinal changes in psychological states in online health community members: Understanding the long-term effects of participating in an online depression community. J Med Internet Res. 2017; 19. https://doi.org/10.2196/jmir.6826 PMID: 28320692

55. Silveira B, Silva HS, Murai F, da Silva APC. Predicting user emotional tone in mental disorder online communities. Futur Gener Comput Syst. 2021; 125: 641–651. https://doi.org/10.1016/j.future.2021.07.014

56. Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. PLoS Med. 2007; 4: 1623–1627. https://doi.org/10.1371/journal.pmed.0040296 PMID: 17941714

57. Boe B. PRAW: The Python Reddit API Wrapper. 2023 [cited 11 May 2023]. Available: https://praw.readthedocs.io/

58. Ghaemi SN, Sachs GS, M. Chiou A, Pandurangi AK, Goodwin FK. Is bipolar disorder still underdiagnosed? Are antidepressants overutilized? J Affect Disord. 1999; 52: 135–144. https://doi.org/10.1016/s0165-0327(98)00076-7 PMID: 10357026

59. Ghaemi SN, Boiman EE, Goodwin FK. Diagnosing bipolar disorder and the effect of antidepressants: A naturalistic study. J Clin Psychiatry. 2000; 61: 804–808. https://doi.org/10.4088/jcp.v61n1013 PMID: 11078046

60. Seraj S, Blackburn KG, Pennebaker JW. Language left behind on social media exposes the emotional and cognitive costs of a romantic breakup. Proc Natl Acad Sci U S A. 2021; 118. https://doi.org/10.1073/pnas.2017154118 PMID: 33526594

61. Jagfeld G. Subreddit topics. 2022 [cited 23 Feb 2023]. Available: https://github.com/glorisonne/reddit_bd_mood_posting_mh/blob/main/data/subreddit_topics.csv

62. Ekman. An argument for basic emotions. Cogn Emot. 1992; 6: 169–200.

63. Ekkekakis P. The Measurement of Affect, Mood, and Emotion. Panteleimon Ekkekakis. Cambridge University Press; 2013. https://doi.org/10.1176/pn.39.2.0031b

64. Köper M, Kim E, Klinger R. IMS at EmoInt-2017: Emotion Intensity Prediction with Affective Norms, Automatically Extended Resources and Deep Learning. Proceedings of the 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis (WASSA). 2017. pp. 50–57. https://doi.org/10.18653/v1/W17-5206

65. Demszky D, Jeongwoo DM, Cowen A, Nemade G, Ravi S. GoEmotions: A Dataset of Fine-Grained Emotions. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. 2020. pp. 4040–4054.

66. Pérez JM, Giudici JC, Luque F. pysentimiento: A Python Toolkit for Sentiment Analysis and SocialNLP tasks. 2021; 1–4. Available: http://arxiv.org/abs/2106.09462

67. Kennedy B, Ashokkumar A, Boyd R, Dehghani M. Text Analysis for Psychology: Methods, Principles, and Practices. Guilford Press. 2022; 3–33.

68. Entwistle C, Marceau E, Boyd RL. Personality disorder and verbal behavior. Handb Lang Anal Psychol. 2022; 335–356. Available: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=psyc18&NEWS=N&AN=2022-08338-017

69. Boyd RL, Pennebaker JW. Language-based personality: a new approach to personality in a digital world. Curr Opin Behav Sci. 2017; 18: 63–68. https://doi.org/10.1016/j.cobeha.2017.07.017

70. Pennebaker JW. Computer-Based Language Analysis as a Paradigm Shift. Handb Lang Anal Psychol. 2022; 576–587.

71. Tausczik YR, Pennebaker JW. The psychological meaning of words: LIWC and computerized text analysis methods. J Lang Soc Psychol. 2010; 29: 24–54. https://doi.org/10.1177/0261927X09351676

72. Kahn JH, Tobin RM, Massey AE, Anderson JA. Measuring emotional expression with the Linguistic Inquiry and Word Count. Am J Psychol. 2007; 120: 263–286. https://doi.org/10.2307/20445398 PMID: 17650921

73. Newell EE, McCoy SK, Newman ML, Wellman JD, Gardner SK. You Sound So Down: Capturing Depressed Affect Through Depressed Language. J Lang Soc Psychol. 2018; 37: 451–474. https://doi.org/10.1177/0261927X17731123

74. Golder SA, Macy MW. Diurnal and Seasonal Mood Vary with Work, Sleep, and Daylength Across Diverse Cultures. Science (80-). 2011; 333: 1878–1881. https://doi.org/10.1126/science.1202775 PMID: 21960633

75. Dzogang F, Lightman S, Cristianini N. Diurnal variations of psychometric indicators in twitter content. PLoS One. 2018; 13. https://doi.org/10.1371/journal.pone.0197002 PMID: 29924814

76. Dzogang F, Lansdall-Welfare T, Cristianini N. Seasonal Fluctuations in Collective Mood Revealed by Wikipedia Searches and Twitter Posts. IEEE Int Conf Data Min Work ICDMW. 2016; 0: 931–937. https://doi.org/10.1109/ICDMW.2016.0136

77. Pennebaker JW, Beall SK. Confronting a Traumatic Event. Toward an Understanding of Inhibition and Disease. J Abnorm Psychol. 1986; 95: 274–281. https://doi.org/10.1037/0021-843X.95.3.274 PMID: 3745650

78. Tay D, Qiu H. Modeling Linguistic (A)Synchrony: A Case Study of Therapist–Client Interaction. Front Psychol. 2022; 13. https://doi.org/10.3389/fpsyg.2022.903227 PMID: 35677134

79. O'Dea B, Boonstra TW, Larsen ME, Nguyen T, Venkatesh S, Christensen H. The relationship between linguistic expression in blog content and symptoms of depression, anxiety, and suicidal thoughts: A longitudinal study. PLoS One. 2021; 16: 1–17. https://doi.org/10.1371/journal.pone.0251787 PMID: 34010314

80. Stone JA, Ryerson NC. Tweeting about alcohol: Exploring differences in Twitter sentiment during the onset of the COVID-19 pandemic. PLoS One. 2022; 17: e0276863. https://doi.org/10.1371/journal.pone.0276863 PMID: 36327323

81. Andy A, Sherman G, Guntuku SC. Understanding the expression of loneliness on Twitter across age groups and genders. PLoS One. 2022; 17: 1–19. https://doi.org/10.1371/journal.pone.0273636 PMID: 36170276

82. Canales L, Martínez-Barco P. Emotion Detection from text: A Survey. Proceedings of the Workshop on Natural Language Processing in the 5th Information Systems Research Working Days. Association for Computational Linguistics; 2014. pp. 37–43.

83. De Choudhury M, Logar T, Sharma SS, Eekhout W, Nielsen RC. Gender and Cross-Cultural Differences in Social Media Disclosures of Mental Illness. Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing. 2017. pp. 353–369. https://doi.org/10.1145/2998181.2998220

84. Tackman AM, Sbarra DA, Carey AL, Donnellan MB, Horn AB, Holtzman NS, et al. Depression, Negative Emotionality, and Self-Referential Language: A Multi-Lab, Multi-Measure, and Multi-Language-Task Research Synthesis. J Pers Soc Psychol. 2019; 116: 817–834. https://doi.org/10.1037/pspp0000187 PMID: 29504797

85. Boyd RL. Psychological Text Analysis in the Digital Humanities. Multimedia. In: Hai-Jew S, editor. Data Analytics in Digital Humanities. Multimedia. Springer; 2017. pp. 161–190.

86. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Erlbaum; 1988.

87. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria; 2022. Available: https://www.r-project.org

88. Meteyard L, Davies RAI. Best practice guidance for linear mixed-effects models in psychological science. J Mem Lang. 2020; 112: 104092. https://doi.org/10.1016/j.jml.2020.104092

89. Hunt D, Brookes G. Corpus, discourse and mental Health. Teubert W, Mahlberg M, editors. Discourse and Mental Health. Bloomsbury Academic; 2020. https://doi.org/10.4324/9780203701928

90. Angermeyer MC, Matschinger H, Carta MG, Schomerus G. Changes in the perception of mental illness stigma in Germany over the last two decades. Eur Psychiatry. 2014; 29: 390–395. https://doi.org/10.1016/j.eurpsy.2013.10.004 PMID: 24321774

91. Goulden R, Corker E, Evans-Lacko S, Rose D, Thornicroft G, Henderson C. Newspaper coverage of mental illness in the UK, 1992–2008. BMC Public Health. 2011; 11. https://doi.org/10.1186/1471-2458-11-796 PMID: 21992410

92. Corker E, Hamilton S, Henderson C, Weeks C, Pinfold V, Rose D, et al. Experiences of discrimination among people using mental health services in England 2008–2011. Br J Psychiatry. 2013; 202: 58–63. https://doi.org/10.1192/bjp.bp.112.112912 PMID: 23553696

93. Naslund JA, Aschbrenner KA, McHugo GJ, Unützer J, Marsch LA, Bartels SJ. Exploring opportunities to support mental health care using social media: A survey of social media users with mental illness. Early Interv Psychiatry. 2019; 13: 405–413. https://doi.org/10.1111/eip.12496 PMID: 29052947

94. Addis ME, Mahalik JR. Men, masculinity, and the contexts of help seeking. Am Psychol. 2003; 58: 5–14. https://doi.org/10.1037/0003-066x.58.1.5 PMID: 12674814

95. Rosenfield S, Mouzon D. Gender and mental health. In: Aneshensel CS, Phelan JC, Bierman A, editors. Handbook of the Sociology of Mental Health. Springer Science+Business Media; 2013. pp. 277–296. https://doi.org/10.1057/9781137295408_12

96. Wang PS, Aguilar-Gaxiola S, Alonso J, Angermeyer MC, Borges G, Bromet EJ, et al. Use of mental health services for anxiety, mood, and substance disorders in 17 countries in the WHO world mental health surveys. Lancet. 2007; 370: 841–850. https://doi.org/10.1016/S0140-6736(07)61414-7 PMID: 17826169

97. Brown JSL, Sagar-Ouriaghli I, Sullivan L. Help-Seeking Among Men for Mental Health Problems. The Palgrave Handbook of Male Psychology and Mental Health. Springer International Publishing; 2019. pp. 397–415. https://doi.org/10.1007/978-3-030-04384-1

98. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. 2013. https://doi.org/10.1093/acprof:oso/9780199241323.003.0025

99. The British Psychological Society. Ethics guidelines for internet-mediated research. 2021.

100. Jagfeld G, Jones SH, Rayson P, Lobban F. Self-reported BD diagnosis (S-BiDD) dataset. 2022. https://doi.org/10.17635/lancaster/researchdata/589

101. Jagfeld G. Code accompanying the paper "Posting patterns in peer online support forums and their associations with emotions and mood in bipolar disorder: exploratory analysis." 2022 [cited 15 Aug 2022]. Available: https://github.com/glorisonne/reddit_bd_mood_posting_mh

102. Fried EI, Epskamp S, Nesse RM, Tuerlinckx F, Borsboom D. What are "good" depression symptoms? Comparing the centrality of DSM and non-DSM symptoms of depression in a network analysis. J Affect Disord. 2016; 189: 314–320. https://doi.org/10.1016/j.jad.2015.09.005 PMID: 26458184

103. Sawilowsky SS. New Effect Size Rules of Thumb. J Mod Appl Stat Methods. 2009; 8: 597–599. https://doi.org/10.22237/jmasm/1257035100

104. Field A, Miles J, Field Z. Discovering Statistics Using R. SAGE Publications Inc.; 2012.

105. Lapidot-Lefler N, Barak A. Effects of anonymity, invisibility, and lack of eye-contact on toxic online disinhibition. Comput Human Behav. 2012; 28: 434–443.

106. Lapidot-Lefler N, Barak A. The benign online disinhibition effect: Could situational factors induce self-disclosure and prosocial behaviors? Cyberpsychology J Psychosoc Res Cybersp. 2015; 9. https://doi.org/10.5817/CP2015-2-3

107. Chester A, O'Hara A. Image, Identity and Pseudonymity in Online Discussions. Int J Learn Annu Rev. 2007; 12: 193–204. https://doi.org/10.18848/1447-9494/CGP/v13i12/45145

108. Brockmeyer T, Zimmermann J, Kulessa D, Hautzinger M, Bents H, Friederich HC, et al. Me, myself, and I: Self-referent word use as an indicator of self-focused attention in relation to depression and anxiety. Front Psychol. 2015; 6: 1–10. https://doi.org/10.3389/fpsyg.2015.01564 PMID: 26500601

109. Mor N, Winquist J. Self-focused attention and negative affect: A meta-analysis. Psychol Bull. 2002; 128: 638–662. https://doi.org/10.1037/0033-2909.128.4.638 PMID: 12081086

110. Barthel M, Stocking G, Holcomb J, Mitchell A. Nearly Eight-in-Ten Reddit Users Get News on the Site. 2016. Available: www.pewresearch.org.

111. Oud M, Mayo-Wilson E, Braidwood R, Schulte P, Jones SH, Morriss R, et al. Psychological interventions for adults with bipolar disorder: Systematic review and meta-analysis. Br J Psychiatry. 2016; 208: 213–222. https://doi.org/10.1192/bjp.bp.114.157123 PMID: 26932483

112. Furnham A. Social psychology as common sense. Bull Br Psychol Soc. 1983; 36: 105–109.

113. Aldao A, Nolen-Hoeksema S, Schweizer S. Emotion-regulation strategies across psychopathology: A meta-analytic review. Clin Psychol Rev. 2010; 30: 217–237. https://doi.org/10.1016/j.cpr.2009.11.004 PMID: 20015584

114. Cluss PA, Marcus SC, Kelleher KJ, Thase ME, Arvay LA, Kupfer DJ. Diagnostic certainty of a voluntary bipolar disorder case registry. J Affect Disord. 1999; 52: 93–99. https://doi.org/10.1016/s0165-0327(98)00049-4 PMID: 10357022

115. Sorkin DH, Janio EA, Eikey E V., Schneider M, Davis K, Schueller SM, et al. Rise in use of digital mental health tools and technologies in the United States during the COVID-19 pandemic: Survey study. J Med Internet Res. 2021; 23: 1–12. https://doi.org/10.2196/26994 PMID: 33822737

116. Morton E, Michalak EE, Murray G. What does quality of life refer to in bipolar disorders research? A systematic review of the construct's definition, usage and measurement. J Affect Disord. 2017; 212: 128–137. https://doi.org/10.1016/j.jad.2017.01.026 PMID: 28160685

117. Jagfeld G, Lobban F, Marshall P, Jones SH. Personal recovery in bipolar disorder: Systematic review and "best fit" framework synthesis of qualitative evidence–a POETIC adaptation of CHIME. J Affect Disord. 2021; 292: 375–385. https://doi.org/10.1016/j.jad.2021.05.051 PMID: 34139411

118. Lobban F, Rayson P, Coole M. iPOF: Improving Peer Online Forums. Healthcare text analytics conference, HealTAC 2022. 2022.