Journal of Social Media Research
ISSN: 3062-0945
2026, 3(1), 30-47
https://jsomer.org

REVIEW ARTICLE OPEN ACCESS

Quality, reliability and misinformation in mental health and neurodivergence content on social media: a systematic review

Alice Carter[1]*, Fergus Gracey[1], Joanna Moody[2], Amber Ovens[1], and Eleanor Chatburn[1, 3]

1 University of East Anglia, Norwich Medical School, Department of Clinical Psychology and Psychological Therapies, United Kingdom.
2 Norfolk and Suffolk NHS Foundation Trust, United Kingdom
3 University of Cambridge, Department of Public Health and Primary Care, United Kingdom
* **Correspondence**: Alice Carter, **Address**: University of East Anglia, Norwich, Norfolk, England, NR4 7TJ, **E-Mail**: alicevictoriacarter@gmail.com

Article Info
DOI: 10.29329/jsomer.84
Article History:
Received: 30/11/2025
Revised: 10/02/2026
Accepted: 25/02/2026
Keywords:
Social media,
Mental health,
Neurodivergence,
Misinformation,
Systematic review

Highlights:

- Significant variability and frequent inaccuracies found in mental health and neurodivergence information on social media.
- Findings highlight urgent need for credible content, clearer misinformation standards, and improved platform moderation.

Abstract

Social media is increasingly used for health information seeking, yet no systematic review has assessed the quality of mental health or neurodivergence-related information on social media. This systematic review aimed to assess the quality, reliability, and prevalence of misinformation in such content, comparing findings across platforms and topics. Searches were performed in MEDLINE Ultimate, APA PsycINFO, CINAHL, and Scopus. Studies were eligible if they evaluated the accuracy, quality, or reliability of mental health or neurodivergent-related information on social media platforms. Twenty-seven studies met the inclusion criteria and were critically appraised using a bespoke tool assessing the search, evaluation, and methodological quality. Due to heterogeneity, the findings were synthesized narratively. Across the 27 included studies, 5057 social media posts were analysed. Misinformation prevalence ranged from 0% to 56.9% and was higher on TikTok than YouTube, and neurodivergence-related content showed higher misinformation prevalence than mental health topics. Information quality and reliability varied widely but were generally higher for professionally created content. These findings highlight a clear need for action: mental health and neurodivergence organizations should create and share accurate, evidence-based content to counter misinformation, and clinicians should be supported to do the same. There is a need for strengthened content moderation, as well as consistent definitions and measures of mental health misinformation. Addressing these issues is vital to protect public mental health and improve the reliability of online information.

Citation: Carter, A., Gracey, F., Moody, J., Ovens, A., & Chatburn, E. (2026). Quality, reliability and misinformation in mental health and neurodivergence content on social media: A systematic review. *Journal of Social Media Research, 3*(1), 30-47. https://doi.org/10.29329/jsomer.84

1. Introduction

Social media platforms, such as YouTube, Facebook, Instagram, and TikTok, are increasingly utilized for health information-seeking purposes (Falgoust et al., 2022; Neely et al., 2021; Sumayyia et al., 2019). The interactive and dynamic nature of social media promotes the sharing of experiences and peer support, which can be beneficial for those seeking health information (Chen & Wang, 2021; Zhao & Zhang, 2017). Social media algorithms serve users content that aligns with their existing interests and beliefs, while limiting exposure to different perspectives, creating an "echo chamber" (Cinelli et al., 2021). Echo chambers limit users' exposure to differing perspectives and can therefore reinforce misleading claims. These processes align with theories of selective exposure and confirmation bias, which suggest that individuals preferentially engage with content that confirms their existing beliefs, thereby reinforcing misinformation within algorithmically curated platforms (Cinelli et al., 2021; Nickerson, 1998). The lack of verification and regulation on social media platforms raises further concerns about the accuracy and reliability of the 'infinite scroll' of information users consume (Girardi et al., 2022; Wang et al., 2019).

Commentators have highlighted TikTok's algorithmic model as a key factor in the spread of misinformation (Grandinetti & Bruinsma, 2022). To encompass the variety of definitions outlined in the literature, the present review defines misinformation as a claim which is based on anecdotal, false, or misleading information due to a lack of scientific evidence (Suarez-Lledo & Alvarez-Galvez, 2021). This definition aligns with the broader "information disorder" framework, which conceptualizes misinformation as information that is false, but not created with the intent on causing harm (Wardle & Derakhshan, 2017). Misinformation is particularly prominent on social media; one review found that up to 80% of health-related content was misinformation, while another found that health-related misinformation was more popular than accurate health information (Suarez-Lledo & Galvez, 2021; Wang et al., 2019). Health misinformation has been linked to harmful consequences, such as promoting misinformed behaviors and heightening distress during health emergencies and pandemics (Borges do Nascimento et al., 2022; Kim & Tandoc, 2022).

Mental health misinformation can perpetuate stigma, which can lead to discrimination and delays in people seeking professional help (Corrigan et al., 2014). Inaccurate beliefs about the causes of mental illness, such as that mental health problems are due to weakness, can reinforce negative stereotypes and discourage individuals from accessing treatment (Henderson et al., 2013). Furthermore, research has found that misinformation that portrays mental illnesses as dangerous or untreatable leads to public fear (Clement et al., 2015; Knaack et al., 2017). In addition, misinformation about treatment options, such as promoting non-evidence-based treatments, can delay people from receiving appropriate care and ultimately result in poorer outcomes (McVay, 2023; Starvaggi et al., 2024).

Studies have reported an increase in young people self-diagnosing with mental health conditions and neurodivergence following information they have seen on social media (Gilmore et al., 2022; Hasan, 2023). This is concerning as incorrect self-diagnosis based on misinformation could conceivably result in delayed or inappropriate treatment and contribute to the pathologization of behaviors, although there is a lack of existing research on this topic. This trend can be partially understood through a mental health literacy framework, which suggests that low public understanding of diagnostic criteria could increase vulnerability to misinformation (Jorm, 2000)

In this review, neurodivergence specifically refers to neurodevelopmental conditions such as autism and ADHD, which are conceptually distinct from mental health conditions. These constructs are examined together as they commonly co-occur within mental health discourse on social media and are similarly affected by misinformation on these platforms (Starvaggi et al., 2024). Despite these potentially significant implications for public (mental) health and increasing research in this area, a comprehensive systematic review that assesses the quality, reliability, and prevalence of mental-health and neurodivergence-related misinformation across social media platforms has yet to be conducted. While one literature review summarized recent findings related to this topic (Starvaggi et al., 2024), it did not utilize systematic search or appraisal methods and was limited to a small number of studies and therefore does not provide a full picture of the present issue. The present systematic review addresses this gap by assessing the quality and accuracy of mental health and neurodivergence-related information across different social media platforms. Through a narrative synthesis of the data, this review sought to answer the following questions: (1) What is the prevalence of mental health and neurodivergence-related misinformation on social media? (2) What is the quality and reliability of mental health and neurodivergence-related information on social media? (3) Does the accuracy, quality, and reliability of mental health and neurodivergence-related information vary across social media platforms and topics?

2. Method

2.1. Guidelines

This review was prospectively registered on Open Science Framework (OSF) (https://doi.org/10.17605/OSF.IO/EHJBK). The review was then conducted and reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (Page et al., 2021) guidelines and the Synthesis Without Meta-analysis (SWiM) guidelines (Campbell et al., 2020).

2.2. Search Strategy

The databases included in the search were MEDLINE Ultimate, APA PsycINFO, CINAHL, and Scopus. These databases were searched on the 1st October 2024 using the following search terms: ("mental health" OR "mental illness" OR "mental disorder" OR psychosis OR dissociation OR schizophreni* OR adhd OR CBT OR "cognitive behavioural" OR "eating disorde*" OR anorexi* OR bulimi* OR OCD OR autism OR ASD OR BPD OR "personality disorder" OR depression OR bipolar OR "obsessive compulsive" OR anxiety OR "attention-deficit hyperactivity disorder" OR ptsd OR "post-traumatic" OR phobia OR "body dysmorphic disorder" OR psychotherapy) AND (misinformation OR disinformation OR accurate OR accuracy OR "fake news" OR useful OR quality OR reliable OR reliability OR credibility OR credible OR trustworth* OR DISCERN OR misleading) AND ("social media" OR youtube OR reddit OR facebook OR twitter OR instagram OR tiktok OR pinterest OR tumblr) NOT (covid-19 or coronavirus or pandemic). An academic librarian approved the strategy and databases. Google Scholar and reference lists of eligible articles were also searched.

2.3. Eligibility Criteria

2.3.1. Inclusion Criteria

Studies were included if the objectives were to evaluate the quality and/or accuracy of mental health and neurodivergence-related information on social media platforms. All methods of assessing the quality and/or accuracy of this information were eligible. Studies were included if they were written in English and no date restrictions were implemented.

2.3.2. Exclusion Criteria

Studies were excluded if they did not assess the quality and/or accuracy of mental health or neurodivergence-related information on social media. For example, studies were excluded if they explored attitudes towards mental health misinformation on social media or if they evaluated mental health information on standard websites.

2.4. Study Selection

Once the search was conducted, duplicates were identified and removed. The remaining articles were screened based on title and abstract by the lead reviewer (AC), who also screened the full text of all potentially eligible articles to identify studies which met the criteria for inclusion. A second rater (AO) reviewed 25% of articles at each stage. While formal inter-rater reliability statistics were not calculated, the reviewers discussed and aligned the inclusion criteria in advance, and discrepancies were resolved through discussion. Screening was conducted in Rayyan.

2.5. Data Extraction

Data extraction was conducted on Microsoft Excel using a pre-determined data extraction template by the lead reviewer (AC), which was piloted on a sample of the included studies. A second reviewer (JM) extracted data for 25% of the included studies, and any discrepancies were discussed and resolved. Extracted data included key study characteristics such as author(s), year, study design, sample size (i.e., number of social media posts evaluated), social media platform, topic(s), how misinformation was defined, evaluation method(s), and results (i.e., misinformation prevalence, information quality, and reliability). From piloting the extraction template, the DISCERN scale (Charnock et al., 1999) and Global Quality Scale (GQS) (Bernard et al., 2007) appeared to be frequently used and were therefore added to the extraction table.

2.6. Quality Ratings

Due to the nature of the review questions, the research designs utilized within this topic differ from the standard experimental, observational or qualitative designs usually seen within the field of psychology and did not contain participants. This meant standard tools for assessing the quality of studies, such as the Mixed Methods Appraisal Tool (MMAT) (Hong et al., 2018) and Critical Appraisal Skills Programme (CASP) (CASP, 2018) checklist were not applicable. A quality appraisal tool created and used in a similar review exploring health misinformation on social media was therefore used, which fit with the study designs used in this topic (Suarez-Lledo & Alvarez-Galvez, 2021) (Table 1). This tool assessed the quality of the search strategy used (S-score) (eight items), how rigorous the evaluation was (E-score) (six items) and a general evaluation of the quality of the research process, such as the methodology, reporting of the results and discussion, for either quantitative (9 items) or qualitative (6 items) studies (G-score). Each score was calculated as the sum of each of the items by equating "yes" or "good" as 1 point, "fair" as 0.5 points, and "no" or "poor" as 0 points. A higher score indicates that a study is of good quality, while a lower score indicates a poor-quality study, with scores <50% classed as low quality. As the original tool does not yield a category for high quality studies, the authors of the present review took the decision to classify studies with a rating of over 75% as high quality. This threshold was selected to enable transparent categorisation and comparison of study quality. All of the included studies were critically appraised by AC, with a second reviewer (JM) assessing 25% of these. Any discrepancies were discussed and resolved.

Table 1. Quality assessment tool.

Dimension	Items
Search Quality (SQ)	1. Was search date/period mentioned? 2. Was search tools mentioned? 3. Was more than 1 search tool used? 4. Was search terms mentioned? 5. Was user engagement mentioned? 6. Was initial hits reported? 7. Was posts in more than 1 language assessed? 8. Was interrater reliability for post selection determined
Evaluation Quality (EQ)	1. Raters blinded for the source 2. Number of raters reported 3. More than 1 rater 4. Interrater reliability figure for evaluation determined 5. A priori criteria defined for accuracy / A priori criteria defined for evaluation 6. Criterion standard for evaluation stated and different from personal opinion
Scoring system for methodological quality of quantitative included studies (GQ)	1. Did the study address a clearly focused issue? 2. Did the authors use an appropriate method to answer their question? 3. Was the study population clearly specified and defined? 4. Were measures taken to accurately reduce measurement bias? 5. Were the study data collected in a way that addressed the research issue? 6. Did the authors take sufficient steps to assure the quality of the study data? 7. Was the data analysis sufficiently rigorous? 8. How complete is the discussion? 9. To what extent are the findings generalizable to other international contexts?
Scoring system for methodological quality of qualitative included studies (GQ)	1. Were steps taken to increase rigour in the analysis of the data? 2. Were the findings of the study grounded in/ supported by the data? 3. Please rate the findings of the study in terms of their breadth and depth. 4. To what extent does the study privilege the perspectives and experiences of health care professionals and patients/carers that are relevant to comparable health systems 5. Overall, what weight would you assign to this study in terms of the reliability/ trustworthiness of its findings? 6. What weight would you assign to this study in terms of the usefulness of its findings for this review?

2.7. Synthesis

A meta-analysis was deemed to be inappropriate for the analysis due to significant heterogeneity across the included studies. This included variation in the social media platforms examined and differences in topic focus (i.e., mental health and neurodivergence). Studies also varied in both the evaluation methods and reporting formats, which limited the comparability of results. Therefore, a narrative synthesis of the findings from the included studies is provided, adopting the Synthesis without Meta-Analysis (SWiM) guidelines (Campbell et al., 2020), structured around the quality of the research, findings related to the quality and reliability of the information, including the misinformation prevalence and whether this varies by social media platform or topic.

3. Results

3.1. Identification of Studies

Of 2772 abstracts identified for screening, 46 full text papers were retrieved and screened, yielding 26 studies which met the criteria for inclusion. An additional paper was identified through searching the reference lists of included papers, resulting in a total of 27 studies. A PRISMA flowchart is presented in Figure 1.

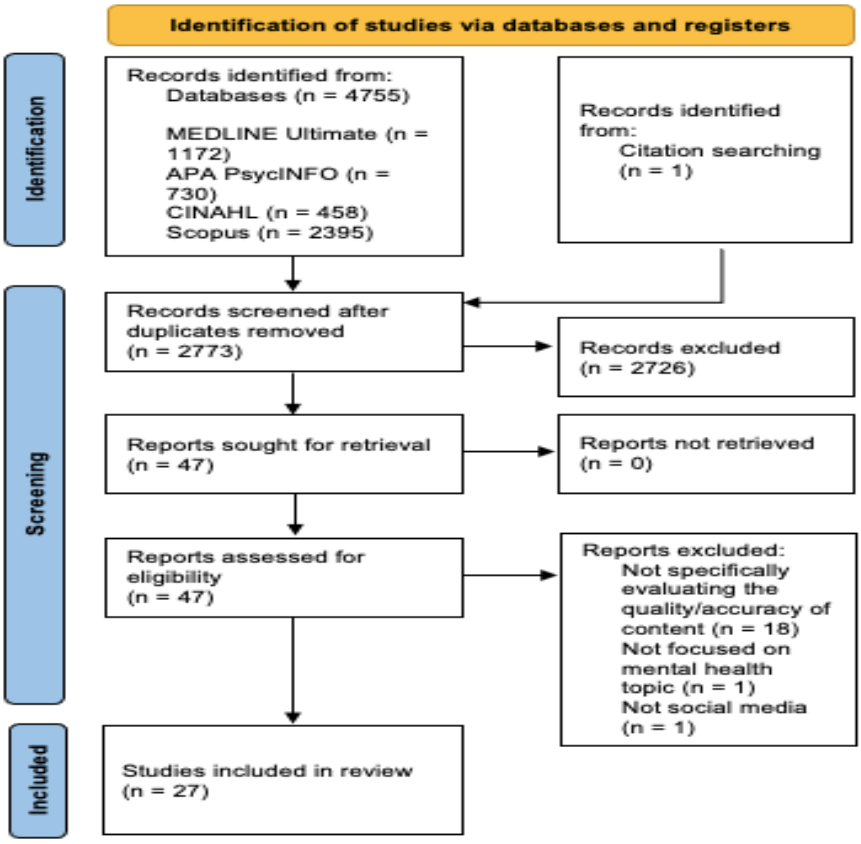

Figure 1. Prisma flow diagram.

3.2. Study Characteristics

The characteristics and findings for each included study are outlined in Table 2 and a narrative summary is also provided. In terms of platforms, included studies evaluated information on Instagram (n=1), YouTube Kids (n=1), Facebook (n=2), X (formerly Twitter) (n=1), TikTok (n=5), and YouTube (n=18). One of these studies evaluated both TikTok and YouTube.

Nearly a third of the included studies focused on neurodivergence, specifically autism (n=4) and attention-deficit hyperactivity disorder (ADHD) (n=4). The remaining studies focused on a variety of mental health diagnoses and treatments, including anorexia nervosa (n=3), post-partum depression (PPD) (n=2), bipolar disorder (n=2), obsessive-compulsive disorder (OCD) (n=2), substance abuse (n=1), dissociative identity disorder (DID) (n=1), schizophrenia (n=1), psychosocial interventions for schizophrenia (n=1), MDMA-assisted psychotherapy for PTSD (n=1), agoraphobia (n=1), MRI claustrophobia (n=1), electroconvulsive therapy (ECT) (n=1), and "mental health" (n=1). One study also explored anxiety, depression and ADHD within the same paper. In total, 5057 social media posts were analyzed across the 27 included studies. The smallest sample was 38 YouTube videos, while the largest was 1532 Facebook statements.

Table 2. Study characteristics and results.

Publication	Platform	Topic	Sample	Definition of Misinformation	Prevalence of Misinformation	DISCERN	GQS	Quality Rating (%)
Abhishek et al. (2021)	YouTube	OCD	82 videos	Grossly deviating from DSM-5 descriptions.	Meaning of obsessions - 8.5% Meaning of compulsions - 4.9%	N/A	N/A	67.40
Abu Sabra & Al Kalaldeh (2024)	YouTube	ECT	250 videos	Not recorded	N/A	3 (Median) *mDISCERN*	3 (Median)	71.74
Alsabhan et al. (2024)	YouTube	Bipolar Disorder	58 videos	Not recorded	N/A	35.8 (Median) *Full DISCERN*	N/A	65.22
Aragon-Guevara et al. (2023)	TikTok	Autism	133 videos	Lack of consistency with existing scientific knowledge related to causes, presentation, diagnostic criteria, evidence-based interventions, and other relevant areas of research.	41%	N/A	N/A	76.101
Bizzotto et al. (2023)	Facebook	Mental Health	1532 statements	Not recorded	26.1%	N/A	N/A	43.48
Brown et al. (2024)	TikTok	Autism	100 videos	Videos containing any factually untrue or scientifically unsubstantiated claims about any aspect of ASD.	40%	N/A	N/A	60.87
Cavalcante et al. (2023)	YouTube	Autism	216 videos	Not recorded	N/A	3 (Median) *mDISCERN*	3 (Median)	71.74
Chakrabarty et al. (2024)	YouTube	Autism	41 videos	Not recorded	N/A	Type of uploader: Doctors – 3 Hospitals – 3 Healthcare organization - 2.5 News channel - 2 Parent of patient – 3 Patient - 3 (Medians) *mDISCERN*	Type of uploader: Doctors – 5 Hospitals – 3 Healthcare organization - 3 News channel - 2 Parent of patient – 4 Patient – 4 (Medians)	67.40

Dobosz et al. (2023)	YouTube	Body Dysmorphic Disorder	38 videos	Not recorded	N/A	32.89 (Mean) *Full DISCERN*	Total: 2.84 Type of uploader: Healthcare – 3.83 Non-healthcare – 2.53 (Medians)	71.74
Joseph et al. (2015)	Twitter	Schizophrenia	685 Tweets	Medically inappropriate Tweets make direct reference to inaccurate facts about the illness.	"Schizophrenia" – 6.6% "Schizophrenic" – 30.1% Total – 18.76%	N/A	N/A	54.35
Kaya et al. (2021)	YouTube	OCD	131 videos	Not recorded	N/A	33.5 (Median) *Full DISCERN*	N/A	73.90
Kaya & Azturk (2023)	YouTube	Agoraphobia	50 videos	Videos containing false or unreliable content.	22%	Total: 3.55 Type of uploader: Professionals – 3.78 Non-professionals – 2.89 (Means) *mDISCERN*	Total: 3.4 Type of uploader: Professionals – 3.57 Non-professionals – 2.92 (Means)	69.60
Kumar & Jha (2018)	YouTube	Psychosocial interventions for Schizophrenia	49 videos	Not recorded	Psychosocial interventions – 0% Interventions in general – 8% CBTp – 0% CR – 12%	N/A	N/A	58.70
Kyarunts et al. (2022)	YouTube	MDMA-assisted psychotherapy for PTSD	100 videos	Not recorded	N/A	31.30 (Mean) *Full DISCERN*	2.3 (Mean)	56.52
Liu-Zarzuela et al. (2023)	YouTube	Postpartum Depression	64 videos	Videos containing: 1) An inaccurate definition of PPD, 2) At least one inaccurate statement about the mechanism of PPD, and 3) At least one inaccurate statement about the treatment/management of PPD.	Total: 7.81% Type of Uploader: Psychiatrist – 0% Other healthcare provider – 0% Health organization – 11.11% Television clip – 0% News channel – 8.33% Other organization – 11.11% Independent user – 12.50%	1.87 (Mean) *mDISCERN*	2.73 (Mean)	80.43
Liu-Zarzuela et al. (2024)	Facebook	Postpartum Depression	69 videos	Videos containing: 1) an inaccurate statement about PPD, 2) provided an inaccurate explanation of the mechanism/pathogenesis of PPD, 3) included an	Total: 3% Type of Uploader: Healthcare provider – 0% Healthcare organization – 0%	2.32 (Mean) *mDISCERN*	2.48 (Mean)	76.10

				inaccurate statement about the treatment/management of PPD.	Television clip – 0% News channel – 0% Other organization – 7.69% Independent user – 7.69%			
Lookingbill et al. (2023)	TikTok	Anorexia	200 videos	Pro-anorexia content.	29.5%	N/A	N/A	58.70
Mallya et al. (2024)	YouTube (Kids)	Depression, Anxiety & ADHD	163 videos	Videos containing an inaccurate definition of the condition, an inaccurate statement about the condition's mechanism, and/or an inaccurate statement about the condition's treatment/management.	Depression – 0% Anxiety – 0% ADHD – 8.89%	N/A	N/A	52.17
Munoz et al. (2024)	YouTube, TikTok	Dissociative Identity Disorder	60 YouTube videos 97 TikTok videos	1) An inaccurate definition on at least one portion of the definition of DID per the DSM-V-TR, 2) An inaccurate statement on at least one portion on the mechanism of DID, and 3) An inaccurate statement on at least one portion on the treatment/ management of DID.	YouTube 6.7% TikTok 10.3%	YouTube 1.7 TikTok 0.4 (Means) *mDISCERN*	YouTube 1.8 TikTok 1.1 (Means)	73.90
Mutlu & Arik (2023)	YouTube	MRI Claustrophobia	65 videos	Videos containing scientifically inaccurate content that misleads patients regarding decisions or treatment.	56.92%	Type of uploader: Professionals – 4.06 Non-professionals – 2.90 (Means) *mDISCERN*	Type of uploader: Professionals – 4.13 Non-professionals – 2.08 (Means)	76.10
Niu & Reed (2023)	YouTube	Substance Abuse	100 videos	Not recorded	25%	N/A	N/A	41.30
Patel et al. (2023)	Instagram	Bipolar Disorder	196 posts	Not recorded	N/A	Type of uploader: Medical professionals - 1 Healthcare organization -2 Patient – 1 Others – 1 (Medians) *mDISCERN*	Type of uploader: Medical professionals - 3 Healthcare organization -2 Patient – 2 Others – 2 (Medians)	58.70
Suresh et al. (2023)	YouTube	Anorexia	59 videos	Not recorded	N/A	Type of uploader: Doctors – 4 Hospital/healthcare organization – 4	Type of uploader: Doctors – 4 Hospital/healthcare	54.35

						News channel – 3 Patient – 3 Other – 3 (Medians) *mDISCERN*	organization – 4 News channel – 4 Patient – 4 Other – 4 (Medians)	
Syed-Abdul et al. (2013)	YouTube	Anorexia	140 videos	Pro-anorexia content.	29.3%	N/A	N/A	73.90
Thapa et al. (2018)	YouTube	ADHD	159 videos	Not recorded	38.36%	N/A	N/A	67.40
Ward et al. (2020)	YouTube	ADHD	120 videos	Not recorded	N/A	Total: 2.03 Type of uploader: Neurologists, pediatricians, psychiatrist (MD) – 2.63 Other medical professional (non-MD) - 3.40 Nonmedical professional – 1.69 PhD – 2.13 Company/advertiser - 2.40 (Means) *mDISCERN*	N/A	52.17
Yeung et al. (2022)	TikTok	ADHD	100 videos	Videos containing information lacking scientific evidence (e.g., unsubstantiated claims about ADHD).	Total: 52% Type of uploader: HCP – 3% Non-HCP – 55.1%	N/A	N/A	76.10

Note. DISCERN (Charnock et al., 1999); modified DISCERN (mDISCERN) (Singh et al., 2012); Global Quality Scale (GQS) (Bernard et al., 2007). Abbreviations: ADHD, attention-deficit hyperactivity disorder; ECT, electroconvulsive therapy; GQS, Global Quality Scale; HCP, healthcare professional; MDMA, 3,4-Methylenedioxymethamphetamine; MRI, magnetic resonance imaging; OCD, obsessive-compulsive disorder; PTSD, post-traumatic stress disorder

3.3. Quality Ratings

The mean quality rating for included studies was 64.82%, with a range of 41.30% (Niu & Reed, 2023) to 80.43% (Liu-Zarzuela, 2023), indicating variation in quality. A summary table of the quality ratings for all included studies is outlined in Table 3. Amongst the included studies, Aragon-Guevara et al. (2023), Liu-Zarzuela et al. (2023), Liu-Zarzuela et al. (2024), Mutlu et al. (2023), and Yeung et al. (2022) were deemed to be of the highest quality. The studies rated as low quality were Bizzotto et al. (2023) and Niu and Reed (2023). A common weakness was search quality: 25/27 studies included social media content in only one language, and 24/27 did not assess interrater reliability for post selection. Studies also frequently lacked measures to reduce measurement bias, with 16/27 studies failing to report an interrater reliability figure for the evaluation.

3.4. Synthesis

It was not possible to group findings by social media platform and topic due to differences in measurement and reporting methods, which would not have allowed for sufficiently large groups to coherently synthesize the findings. Findings are grouped by the measurement method used and comparisons drawn between platforms and topics within this.

Table 3. Summary of quality ratings.

Study	SQ	EQ	GQ	Total
Abishek et al. (2021)	50.00	75.00	77.78	67.40
Abu Sabra et al. (2024)	62.50	66.67	83.33	71.74
Alsabhan et al. (2024)	50.00	66.67	77.78	65.22
Aragon-Guevara et al. (2023)	62.50	75.00	88.89	76.10
Bizzotto et al. (2023)	18.75	41.67	66.67	43.48
Brown et al. (2024)	50.00	58.33	72.22	60.87
Cavalcante et al. (2023)	50.00	75.00	88.89	71.74
Chakrabarty et al. (2024)	62.50	58.33	77.78	67.40
Dobosz et al. (2023)	43.75	83.33	88.89	71.74
Joseph et al. (2015)	25.00	66.67	72.22	54.35
Kaya et al. (2021)	50.00	83.33	88.89	73.90
Kaya et al. (2023)	50.00	83.33	77.78	69.60
Kumar et al. (2018)	56.25	50.00	66.67	58.70
Kyarunts et al. (2022)	50.00	33.33	77.78	56.52
Liu-Zarzuela et al. (2023)	62.50	83.33	94.44	80.43
Liu-Zarzuela et al. (2024)	50.00	83.33	94.44	76.10
Lookingbill et al. (2023)	37.50	58.33	77.78	58.70
Mallya et al. (2024)	31.25	66.67	61.11	52.17
Munoz et al. (2024)	56.25	75.00	88.89	73.90
Mutlu et al. (2023)	50.00	83.33	94.44	76.10
Niu et al. (2023)	31.25	33.33	55.56	41.30
Patel et al. (2023)	50.00	33.33	83.33	58.70
Suresh et al. (2023)	37.50	33.33	83.33	54.35
Syed-Abdul et al. (2013)	68.75	75.00	77.78	73.90
Thapa et al. (2018)	50.00	75.00	77.78	67.40
Ward et al. (2020)	31.25	66.67	61.11	52.17
Yeung et al. (2022)	62.50	75.00	88.89	76.10

3.5. Definitions of Misinformation

Thirteen studies outlined their definitions of misinformation, which varied across studies. Most studies defined misinformation as content which contained factually inaccurate and/or scientifically unsubstantiated claims (e.g., Brown et al., 2024; Liu-Zarzuela et al., 2024; Yeung et al., 2022). Other studies used more specific criteria to define misinformation, such as deviation from the DSM-V (Abishek et al., 2021), pro-anorexia content (Lookingbill et al., 2023; Syed-Abdul et al., 2013), and content which misleads patients regarding treatment decisions (Mutlu et al., 2023).

3.6. Tools Used to Evaluate Information

Three approaches were used to evaluate the reliability, quality and accuracy of mental health and neurodivergence-related information on social media.

Seventeen studies reported the percentage of misinformation, demonstrating the accuracy of the information. The prevalence of misinformation depended on how each study defined it, rather than using a validated tool, and was calculated as the percentage of content assessed that contained misinformation. Fifteen studies used the DISCERN measure, a validated tool for assessing the reliability of written health information (Charnock et al., 1999). Four studies used the full 16-item DISCERN, with scores between 63-75 considered “excellent”, 51-62 “good”, 39-50 “fair”, 27-38 “poor”, and 16-26 “very poor”. Eleven studies used the 5-item

modified DISCERN (mDISCERN) (Singh et al., 2012), in which a score of three or more indicates highly reliable information.

The Global Quality Scale (GQS), a validated five-point Likert scale designed to assess the quality of online health information (Bernard et al., 2007), was used in 12 studies. A score of one indicates poor quality and a score of five indicates excellent quality.

Studies varied in how they reported the (m)DISCERN and GQS, with some reporting means and others reporting medians. A narrative summary was used to compare findings across topics and social media platforms.

3.7. Prevalence of Misinformation

Misinformation prevalence was reported in 17/27 studies and varied across social media platforms and topics. Overall, misinformation prevalences ranged from 0% for videos on anxiety and depression on YouTube Kids (Mallya et al., 2024), to 56.92% for videos on MRI claustrophobia on YouTube (Mutlu et al., 2023). The mean misinformation prevalence across all studies was 26.41%.

In terms of platform, misinformation prevalence was consistently higher on TikTok, including prevalences of 52% for ADHD-related TikTok videos (Yeung et al., 2022) and 41% for autism-related TikTok videos (Aragon-Guevara et al., 2023), while another study reported higher misinformation prevalence on TikTok (10.3%) than on YouTube (6.7%) for content about DID (Munoz et al., 2024). The mean misinformation prevalence for mental health and neurodivergence-related information on TikTok was 34.56%. The prevalence of misinformation for YouTube videos varied depending on the topic and was lowest in videos about DID at 6.7% (Munoz et al., 2024), and highest in videos about MRI claustrophobia at 56.92% (Mutlu et al., 2023), with a mean of 21.99% misinformation. YouTube Kids had the lowest misinformation prevalence, reporting no misinformation for both anxiety and depression and 8.89% for ADHD (Mallya et al., 2024). Two studies investigated misinformation on Facebook, with a mean prevalence of 14.55%. Only one study investigated misinformation on X (formerly Twitter), reporting a prevalence of 18.76%, while the single study on Instagram did not report a prevalence.

In terms of topic, social media content on PPD contained the least amount of misinformation, ranging from 3% on Facebook (Liu-Zarzuela et al., 2024) to 7.81% on YouTube (Liu-Zarzuela et al., 2023), while content on MRI claustrophobia was found to contain the most misinformation at 56.92% (Mutlu et al., 2023). Content on neurodivergence consistently contained a higher misinformation prevalence than content on mental health conditions and treatments, with prevalences of 40% (Brown et al., 2024) and 41% (Aragon-Guevara et al., 2023) for autism, and 38.6% (Thapa et al., 2018) and 52% (Yeung et al., 2022) for ADHD.

3.8. Reliability of Information

The reliability of information, demonstrated by (m)DISCERN scores, varied considerably across platforms, topics, and uploader types. For the full DISCERN, possible scores range from 16 (very poor reliability) to 75 (excellent reliability). Only YouTube studies used the full DISCERN and reported scores ranging from a median of 31.3 (Kyarunts et al., 2022) to 35.8 (Alsabhan et al., 2024), indicating poor reliability across different topics.

For studies that utilized the 5-item mDISCERN, overall mean scores for YouTube ranged from 1.7 (Munoz et al., 2024) to 3.55 (Kaya et al., 2023), indicating the reliability of YouTube videos varies from poor to high across different topics. YouTube videos were reported to be more reliable than TikTok videos, as demonstrated by Mutlu et al. (2023), who found that TikTok videos on DID had a mean mDISCERN score of 0.4, compared to 1.7 for YouTube videos on the same topic. When looking at the mean mDISCERN scores for PPD content, Facebook was reported to be more reliable than YouTube, with means of 2.32 and 1.87, respectively. Meanwhile, YouTube videos demonstrated higher reliability than Instagram videos across different topics.

Professionally created content was usually found to be more reliable than content created by non-professionals when considering the mDISCERN scores (Kaya & Azturk, 2023; Mutlu et al., 2023; Suresh et al., 2023), although some studies reported the reliability of professional and patient-created content to be of equal reliability (Chakrabarty et al., 2024; Patel et al., 2023). Content by professionals was reported to be more reliable on YouTube than on Instagram, with studies reporting median scores of 4 (Suresh et al., 2023) and 1 (Patel et al., 2023), respectively.

Three of the high-quality studies assessed the reliability of the information, all of which used the mDISCERN. The overall mean mDISCERN score for the higher-quality studies ranged from 1.87 (Liu-Zarzuela et al., 2023) to 2.32 (Liu-Zarzuela et al., 2024), whereas reliability was not evaluated in the lower-quality studies.

Studies that reported both misinformation prevalence and overall mean mDISCERN scores had lower misinformation prevalence. Some studies with low misinformation prevalence also demonstrated low

mDISCERN scores, indicating low reliability (Liu-Zarzuela et al., 2023; Munoz et al., 2024), whereas a study with a higher misinformation prevalence was rated as highly reliable (Kaya et al., 2021). All three of these studies assessed information on YouTube but varied by topic.

3.9. Quality of Information

The quality of the information was demonstrated by GQS scores, which varied by platform, topic, and uploader type. For YouTube, the overall mean GQS scores ranged from 1.8 for DID content (Munoz et al., 2024) to 3.4 for agoraphobia content (Kaya et al., 2023), demonstrating that the quality and flow of YouTube content vary from poor to moderate across topics. YouTube videos were reported to be of slightly higher quality than TikTok videos, as demonstrated by Mutlu et al. (2023) in which TikTok videos on DID had a mean GQS score of 1.1, compared to 1.8 for YouTube videos on the same topic. Content for PPD was reported to be of poor quality across both Facebook and YouTube, with mean GQS scores of 2.48 (Liu-Zarzuela, 2024) and 2.73 (Liu-Zarzuela, 2023), respectively.

Professionally created content was mostly reported to be of higher quality than content created by non-professionals when considering the GQS scores (Chakrabarty et al., 2024; Dobosz et al., 2023; Kaya & Azturk, 2023; Mutlu et al., 2023; Patel et al., 2023). However, for YouTube videos on autism, content uploaded by patients and parents had higher median GQS scores than content uploaded by hospitals and healthcare organizations (Chakrabarty et al., 2024), whereas there was no difference in median GQS scores between uploader types for YouTube videos about anorexia.

Some studies that reported low misinformation prevalence also reported low mean GQS scores, indicating the information was of poor quality (Liu-Zarzuela et al., 2023; Liu-Zarzuela et al., 2024; Munoz et al., 2024), while a study reporting comparatively higher misinformation prevalence also reported a higher mean GQS score, which indicated that the information was of moderate quality (Kaya et al., 2021). Some studies reported information as both of poor reliability and poor quality (Munoz et al., 2024) and of high reliability and moderate quality (Kaya et al., 2021), although this finding was inconsistent.

4. Discussion

This systematic review aimed to identify the quality and accuracy of mental health and neurodivergence-related information on social media, including the prevalence of misinformation. A total of 27 papers of varying quality were included in this review, and a synthesis of the included studies compared the quality, accuracy and reliability of information by social media platform and topic.

This review highlights considerable variation in the accuracy, reliability and quality of information across social media platforms, topics, and uploader types. TikTok was found to have the highest prevalence of misinformation, while Facebook was reported to have the lowest, with no prior research having compared misinformation prevalences between the two. The accuracy, reliability and quality of information on YouTube varied across different topics but scored consistently better than TikTok, aligning with previous research which found that YouTube content contained more credible information and less misinformation than TikTok (Tam et al., 2022). This variability suggests that platform-specific factors, such as algorithmic systems and content moderation, may influence the spread of misinformation, therefore supporting theories of selective exposure and confirmation bias (Cinelli et al., 2021; Nickerson, 1998). This review's findings in relation to TikTok align with previous research suggesting the role of TikTok's algorithm in spreading misinformation (Grandinetti & Bruinsma, 2022). Contrastingly, Facebook and YouTube's search-based designs have less focus on rapid engagement and viral trends, with YouTube often favoring more established channels and Facebook prioritizing content by existing connections and followed pages (UIDesignz, 2024; QuickFrame, 2023), which perhaps deliver information from more established and trusted sources than TikTok's 'For You Page'. YouTube Kids was the only platform to report findings of no misinformation for some topics, which is likely due to the implementation of stricter content moderation and prioritization of child-friendly content (YouTube Kids, n.d.).

The variability of findings for YouTube across different topics indicates that platform-specific factors are not the only influence in the spread of misinformation. Low health literacy has been linked with the spread of misinformation on health topics (Borges do Nascimento et al., 2022), which suggests a lack of (mental) health literacy on certain topics may contribute to increased misinformation. While most studies included in this review did not evaluate whether misinformation prevalence differed by type of uploader, the studies which did report this supported previous research which found that misinformation usually originates from individual users with no official or institutional affiliations (Wang et al., 2019). This suggests the proportion of professional vs non-professional uploaders may vary depending on the topic within a certain social media platform, with some topics

involving more discourse among lay people which may impact the misinformation prevalence, although this cannot be concluded as part of the present review. These findings may partly be explained by mental health literacy frameworks, in which conditions characterised by overlapping symptoms, less clearly defined diagnostic boundaries, and high public discourse may be particularly vulnerable to oversimplification and misinterpretation on social media (Jorm, 2000).

Studies using the (m)DISCERN and GQS measures also consistently reported that professionally created content contained more reliable and higher quality information than non-professional content across most topics, aligning with findings in the field of health misinformation in which content by medical professionals were of higher quality than non-medical influencers (Dimitroyannis et al., 2024). However, anorexia content by doctors and patients were found to be of equally high quality, while content on bipolar disorder by medical professionals and patients were found to be of equally poor reliability. This may indicate a clearer public understanding and reduced stigma for anorexia than bipolar disorder, and there is scope for future research to investigate the role of stigma and public perceptions on the quality of information shared across different topics. However, this finding also raises concerns regarding the quality of information on bipolar disorder being shared by professionals and the role this may play in the spread of misinformation, particularly as health professionals are viewed as trusted sources on social media (Freeman et al., 2023).

Interestingly, studies reported information on mental health and neurodivergence as being accurate but unreliable and of low quality, and of being inaccurate but with high reliability and moderate quality. The purpose of this review was not to identify relationships between measures, which would not have been possible regardless due to the limited number of studies. However, this finding demonstrates the importance of studies using more than one measure when conducting research on this topic to thoroughly evaluate and provide a fuller picture of mental health and neurodivergence-related information on social media.

Many of the included studies used the (m)DISCERN to evaluate the reliability of the information. While this is a validated tool, it was developed for the evaluation of written health information, and the suitability for its use in evaluating videos on social media platforms is therefore questionable (Azer, 2020). This research poses a need for a tool specifically designed to assess the reliability of mental health and neurodivergence-related content on various social media platforms, such as including criteria for short video content rather than purely written information.

Another key issue highlighted by this review is the lack of consistency in how information on mental health and neurodivergence is evaluated and reported in social media studies. While there was variation in the methods used to evaluate the information (i.e., percentage of misinformation, mDISCERN, DISCERN, GQS), the way these findings were reported also lacked consistency, with some studies reporting means and others medians. This limits the ability to coherently compare findings across platforms and topics and highlights a need for a consistent methodology in the evaluation and reporting of mental health and neurodivergence-related information quality on social media.

4.1. Limitations

There are several limitations to this review which should be considered, the first of which relates to the bias in the number of studies evaluating each platform. Most studies focused on YouTube, while minimal studies evaluated X (formerly Twitter), Facebook or Instagram. This prevented comparisons from being made within each platform as they were for YouTube, limiting the conclusions which can be made regarding potential factors which impact the quality of information on these platforms.

As this review focused on the quality of mental health and neurodivergence-related information on social media, studies were only included if they utilised tools which specifically measured this, i.e., whether information was accurate or reliable. However, this meant that studies exploring other qualities of this social media content were excluded. This excluded studies which utilised the Patient Education Materials Assessment Tool (PEMAT) (Shoemaker et al., 2014), which measures the understandability and actionability of content. Although outside this review's scope, this limited broader insights into mental health and neurodivergence-related content, particularly whether content is presented in a way that is understandable and promotes action, which would be an important topic to explore within its own context.

A further limitation related to the use of multiple evaluation methods (i.e., DISCERN, mDISCERN, GQS, and misinformation prevalence). While this allowed for a more comprehensive assessment of accuracy, quality, and reliability, this limited direct comparability and complicated synthesis across platforms and topics.

Another limitation was the quality appraisal tool used. Due to the novel topic area, there were no existing validated tools which would have been appropriate due to the designs of the included studies. While the tool

used was appropriate for the study designs and had been used in a previous review (Suarez-Lledo & Alvarez-Galvez, 2021), it is limited in its validity in assessing the quality of the studies in this review. Furthermore, this tool does not outline a cut-off score for high quality studies, and therefore the conclusions made about the quality of the studies were limited to whether they were higher or lower quality than one another, and an arbitrary value was required to establish the highest quality studies. Furthermore, while some studies were of considerably lower quality, indicating a higher risk of bias, they were included in this review. While this is a limitation of the present review, this decision was made as removing lower-rated studies would have made comparing findings by social media platform and topic even more challenging.

Finally, the generalizability of these findings is limited by aspects of the studies included in the review, particularly that many studies did not report inter-rater reliability, and that most included studies evaluated content in only one language. Furthermore, while this review included some studies from the Global South, the majority came from the Global North, which presents a risk of Westernised bias and limits the present review's generalizability to other international contexts.

5. Implications and Future Directions

These findings have important implications, particularly relating to public (mental) health. Given some of the concerning findings related to the accuracy, reliability and quality of mental health and neurodivergence-related information on social media, organisations for mental health and neurodivergence should disseminate more credible content to counteract misinformation. Individual clinicians should also actively engage with content and share accurate information to users, with organisations and services providing support with this, including guidance on effective and responsible information sharing on social media. Considering the fast-moving nature of social media, it would be beneficial for future research to understand the social media literacy of clinicians and identify areas of learning to support with the sharing of credible information. The accurate interpretation and responsible sharing of social media content for both clinicians and the public could be further supported by social media literacy initiatives.

Future research should assess social media user demographics by platform to identify whether this influences the accuracy, quality and reliability of information across platforms. This review also demonstrates the need for consistency within this topic, first with the conceptual definition of (mental) health misinformation, and second with the measurement and reporting of information. Future research should therefore focus on the development of validated and standardised tools for assessing the quality, accuracy, and reliability of assessing this content on social media. From a policy perspective, these findings support the development of clearer standards or recommendations for online information on mental health and neurodivergence. Finally, these findings highlight a need for platforms to review algorithmic designs and strengthen content moderation strategies to prioritise accurate information on mental health and neurodivergence and reduce the spread of misinformation, particularly for trending or highly engaged content.

Statement of Researchers	
Researchers' contribution rate statement:	**Alice Carter:** Conceptualization, data curation, formal analysis, investigation, methodology, project administration, supervision, validation, visualization, writing (original draft), writing (review & editing). **Fergus Gracey:** Second Author: Conceptualization, methodology, supervision, writing (review & editing). **Joanna Moody:** Investigation, validation, writing (review & editing) **Amber Ovens:** Investigation, validation, writing (review & editing) **Eleanor Chatburn:** Conceptualization, methodology, supervision, writing (review & editing)
Conflict statement:	The authors declare that they have no conflict of interest.
Data Availability Statement:	Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.
Funding:	This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.
Ethical Considerations:	This is a review study, and ethical approval was therefore not needed.
Author Biographies	
First Author,	**Alice Carter** is a registered clinical psychologist working in an adult community mental health service in Surrey, UK, where she provides psychological assessment and intervention for adults experiencing a range of mental health difficulties. She completed her BSc in Psychology at the University of Surrey in 2020 and went on to complete the Doctorate in Clinical

Second Author,	Psychology at the University of East Anglia in 2025. Her doctoral thesis portfolio focused on the topics of social media and youth mental health. **Fergus Gracey** joined the Department of Clinical Psychology at UEA in 2014 and is now a Clinical Associate Professor in Clinical Psychology, Postgraduate Director of Research for the Clinical Psychology Programme and co-lead of the Neuropsychology module. Fergus was formerly the Clinical Lead for the Cambridge Centre for Paediatric Neuropsychological Rehabilitation and lead psychologist at the Oliver Zangwill Centre for Neuropsychological Rehabilitation. Fergus' clinical and research interests are in predominantly in the psychosocial aspects of neuropsychological rehabilitation. Fergus is currently involved in research relating to family and caregivers of people with brain injuries (children young people and adults) or dementia, and social connection and wellbeing following a brain injury, mainly from a qualitative perspective. Fergus holds a number of honorary positions being a member of the BPS Division of Neuropsychology's Policy Unit, the Board of Associate Editors of Neuropsychological Rehabilitation, and is a committee member for AnchorPoint, the WFNR Organisation for Psychological Research in Stroke (OPSYRIS) and the Norfolk and Waveney Acquired Brain Injury Network.
Third Author,	**Joanna Moody** graduated in 2022 with a Bachelor's degree in Psychology, followed by a Master's in Applications of Psychology at University of Suffolk. Joanna hopes to continue furthering her education and research experience by pursuing a career as a clinical psychologist. Her first Assistant Psychologist role was in Adult Learning Disability and Autism service in Suffolk which sparked her passion for this population. She is currently an Assistant Psychologist in an adult community mental health team in Norfolk, UK.
Fourth Author,	**Amber Ovens** is a registered clinical psychologist currently working in a youth community mental health team in Norfolk, UK. She completed her BSc in Psychology at the University of East Anglia in 2019, before completing the Doctorate of Clinical Psychology at the University of East Anglia in 2025. Her doctoral thesis portfolio focused on the topics of social media and anxiety.
Fifth Author,	**Eleanor Chatburn** is a registered clinical psychologist and mental health researcher with a special interest in global child and youth mental health. She is a Senior Lecturer in Clinical Psychology at Anglia Ruskin University, Cambridge, and holds honorary academic positions as a Visiting Researcher at the University of Cambridge and Honorary Lecturer at Norwich Medical School, University of East Anglia. Eleanor completed her professional doctorate in clinical psychology in the UK and is accredited in cognitive behavioural therapy (CBT). Alongside her academic work, she practices clinically within the NHS, providing mental health support to medical students at the University of Cambridge.

6. References

Abhishek, P., Gogoi, V., & Borah, L. (2021). Depiction of obsessive-compulsive disorder in YouTube videos. *Informatics for Health and Social Care, 46*(3), 256-262. https://doi.org/10.1080/17538157.2021.1885036

Abu Sabra, M., & Al Kalaldeh, M. (2024). Is the electroconvulsive therapy video on YouTube supportive information for patients and their primary caregivers? *Journal of ECT, 40*(2), 124-128. https://doi.org/10.1097/yct.0000000000000996

Alsabhan, J., Almalag, H., Abanmy, N., Aljadeed, Y., Alhassan, R., & Albaker, A. (2024). A content-quality and optimization analysis of YouTube as a source of patient information for bipolar disorder. *Saudi Pharmaceutical Journal, 32.* https://doi.org/10.1016/j.jsps.2024.101997

Aragon-Guevara, D., Castle, G., Sheridan, E., Vivanti, G. (2023). The reach and accuracy of information on autism on TikTok. *Journal of Autism and Developmental Disorders.* https://doi.org/10.1007/s10803-023-06084-6

Azer, S. (2020). Are DISCERN and JAMA suitable instruments for assessing YouTube videos on thyroid cancer? Methodological concerns. *Journal of Cancer Education, 35,* 1267-1277. https://doi.org/10.1007/s13187-020-01763-9

Bernard, A., Langille, M., Hughes, S., Rose, C., Leddin, D., & Veldhuyzen van Zanten. (2007). A systematic review of patient inflammatory bowel disease information resources on the World Wide Web. *The American Journal of Gastroenterology, 102*(9), 2070-2077. https://doi.org/10.1111/j.1572-0241.2007.01325.x

Bizzotto, N., Schulz, P., & Bruijn, G. (2023). The "loci" of misinformation and its correction in peer- and expert-led online communities for mental health: content analysis. *Journal of Medical Internet Research, 25.* http://dx.doi.org/10.2196/44656

Borges do Nascimento, I., Pizarro, A., Almeida, J., Azzopardi-Muscat, N., Gonçalves, M., Björklund, M., & Novillo-Ortiz, D. (2022). Infodemics and health misinformation: a systematic review of reviews. *Bulletin of the World Health Organization, 100*(9), 544-561. https://doi.org/10.2471/BLT.21.287654

Brown, E., Kuzmiak, F., Singh, A., Monga, V., Bell, T., Nolan, J., Schlenker, M., Moore, J., & Kashyap, R. (2024). A cross-sectional analysis of TikTok autism spectrum disorder content quality. *Emerging Trends in Drugs, Addictions and Health, 4*. https://doi.org/10.1016/j.etdah.2024.100150

Campbell, M., McKenzie, J., Sowden, A., Katikireddi, S., Brennan, S., Ellis, S., Hartmann-Boyce, J., Ryan, R., Shepperd, S., Thomas, J., Welch, V., & Thomson, H. (2020). Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. *Research Methods & Reporting.* https://doi.org/10.1136/bmj.l6890

Cavalcante, J., Sales, M., de Sousa Junior, R., Souto, D., Gonçalves, R., Resende, C., & Leite, H. (2024). Analysis of the Brazilian-Portuguese content on autism spectrum disorder available on YouTube videos. *Physiccal & Occupational Therapy in Pediatrics, 44*(1), 128-142. https://doi.org/10.1080/01942638.2023.2199843

Chakrabarty, D., Arora, N., Ahmed, I., & Satish, S. (2024). An objective comparison of the quality and reliability of information regarding autism spectrum disorder on YouTube. *Cureus, 16*(5). https://doi.org/10.7759/cureus.60904

Charnock, D., Sheppard, S., Needham, G., & Gann, R. (1999). DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. *Journal of Epidemiology and Community Health, 53*(2), 105-111. https://doi.org/10.1136/jech.53.2.105

Chen, J., & Wang, Y. (2021). Social Media Use for Health Purposes: Systematic Review. *Journal of Medical Internet Research, 23*(5). https://doi.org/10.2196/17917

Cinelli, M., Morales, G., Galeazzi, A., & Starnini, M. (2021). The echo chamber effect on social media. *Computer Sciences, 118*(9). https://doi.org/10.1073/pnas.2023301118

Clement, S., Schauman, O., Graham, T., Maggioni, F., Evans-Lacko, S., Bezborodovs, N., Morgan, C., Rüsch, N., Brown, J., & Thornicroft, G. (2015). What is the impact of mental health-related stigma on help-seeking? A systematic review of quantitative and qualitative studies. *Psychological Medicine, 45*(1), 11-27. https://doi.org/10.1017/s0033291714000129

Corrigan, P., Druss, B., & Perlick, D. (2014). The impact of mental illness stigma on seeking and participating in mental health care. *Psychological Science in the Public Interest: A Journal of the American Psychological Society, 15*(2), 27-70. https://doi.org/10.1177/1529100614531398

Critical Appraisal Skills Programme. (2018). *CASP systematic review checklist.* https://casp-uk.net/casp-tools-checklists/

Dimitroyannia, R., Fenton, D., Cho, S., Nordgren, R., Pinto, J., & Roxbury, C. (2024). A social media quality review of popular sinusitis videos on TikTok. *Otolaryngology – Head and Neck Surgery, 170*(5), 1456-1466. https://doi.org/10.1002/ohn.688

Dobosz, M., Lewandowski, M., Świerczewska, Z., Barańska-Rybak, W., & Cubała, W. (2023). Are YouTube videos a reliable source of information about body dysmorphic disorder? *Advances in Psychiatry & Neurology, 32*(2), 76-82. https://doi.org/10.5114/ppn.2023.128706

Falgoust, G., Winterlind, E., Moon, P., Parker, A., Zinzow, H., & Madathil, K. (2022). Applying the uses and gratifications theory to identify motivational factors behind young adult's participation in viral social media challenges on TikTok. *Human Factors in Healthcare, 2*. https://doi.org/10.1016/j.hfh.2022.100014

Freeman, J., Caldwell, P., & Scott, K. (2023). How adolescents trust health information on social media: A systematic review. *Academic Pediatrics, 23*(4), 703-719. https://doi.org/10.1016/j.acap.2022.12.011

Gilmore, R., Beezhold, J., Selwyn, V., Howard, R., Bartolome, I., & Henderson, N. (2022). Is TikTok increasing the number of self-diagnoses of ADHD in young people? *European Psychiatry, 65*(1). https://doi.org/10.1192/j.eurpsy.2022.1463

Girardi, A., Singh, N., Boyd, C. (2022). Using Social Media in Health Care Research Should Proceed with Caution. Comment on "The Use of Social Media for Health Research Purposes: Scoping Review". *Journal of Medical Internet Research, 24*(1). https://doi.org/10.2196/35286

Grandinetti, J., & Bruinsma, J. (2023). The affective algorithms of conspiracy TikTok. *Journal of Broadcasting & Electronic Media, 67*(3), 274-293. https://doi.org/10.1080/08838151.2022.2140806

Hasan, B. (2023). The relationship between social media and the increase in mental health problems. *International Journal of Environmental Research and Public Health, 20*(3). https://doi.org/10.3390/ijerph20032383

Henderson, C., Evans-Lacko, S., Thornicroft, G. (2013). Mental illness stigma, help seeking, and public health programs. *American Journal of Public Health, 103*(5), 777-780. https://doi.org/10.2105/ajph.2012.301056

Hong, Q., Pluye, P., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M., Griffiths, F., Nicolau, B., O'Cathain, A., Rousseau, M., & Vedel, I. (2018). *Mixed Methods Appraisal Tool (MMAT) Version 2018. User Guide.* McGill Department of Family Medicine. https://doi.org/10.3233/EFI-180221

Jorm, A. (2000). Mental Health Literacy. *The British Journal of Psychiatry, 177*(5), 396-401. https://doi.org/10.1192/bjp.177.5.396

Joseph, A., Tandon, N., Yang, L., Duckworth, K., Torous, J., Seidman, L., Keshavan, M. (2015). #Schizophrenia: Use and misuse on Twitter. *Schizophrenia Research, 165*, 111-115. http://dx.doi.org/10.1016/j.schres.2015.04.009

Kaya, A., & Akturk, B. (2023). Quality and content analysis: Can YouTube videos on agoraphobia be considered a reliable source? *Cureus, 15*(8). https://doi.org/10.7759/cureus.43318

Kaya, H., Bulbul, A., & Kahve, A. (2021). YouTube as a source of information: How good is the quality and reliability of videos related to obsessive compulsive disorder? *Dusunen Adam The Journal of Psychiatry and Neurological Sciences, 34*, 368-374. https://dx.doi.org/10.14744/DAJPNS.2021.00159

Kim, H., & Tandoc, E. (2022). Consequences of online misinformation on COVID-19: two potential pathways and disparity by eHealth literacy. *Frontiers in Psychology, 13*. https://doi.org/10.3389/fpsyg.2022.783909

Knaack, S., Mantler, E., & Szeto, A. (2017). Mental illness-related stigma in healthcare: Barriers to access and care and evidence-based solutions. *Healthcare Management Forum, 30*(2), 111-116. https://doi.org/10.1177/0840470416679413

Kumar, D., & Jha, M. (2018). YouTube videos on psychosocial interventions for schizophrenia. *Psychosis, 10*(3), 220-224. https://doi.org/10.1080/17522439.2018.1497694

Kyarunts, M., Mansukhani, M., Loukianova, L., & Kolla, B. (2022). Assessing the quality of publicly available videos on MDMA-assisted psychotherapy for PTSD. *The American Journal on Addictions, 31*, 502-507. https://doi.org/10.1111/ajad.13325

Liu-Zarzuela, J., Mallya, M., Munoz, I., & Grayson, K. (2023). Evaluating YouTube as a source of medical information for postpartum depression. *Journal of Affective Disorders Reports, 13.* https://doi.org/10.1016/j.jadr.2023.100599

Liu-Zarzuela, J., Munoz, I., Truong, N., Zahra, A., Mallya, M., & Grayson, K. (2024). Postpartum depression and Facebook: A cross-sectional study of content quality and reliability. *Journal of Affective Disorders Reports, 16*. https://doi.org/10.1016/j.jadr.2024.100745

Lookingbill, V., Mohammadi, E., & Cai, Y. (2023). Assessment of accuracy, user engagement, and themes of eating disorder content in social media short videos. *Jama Network Open, 6*(4). https://doi.org/10.1001/jamanetworkopen.2023.8897

Mallya, M., Liu-Zarzuela, J., Munoz, I., & Shotwell, J. (2024). Assessing the educational value of YouTube Kids videos related to anxiety, depression, and ADHD. *Journal of the American Academy of Child & Adolescent Psychiatry.* https://doi.org/10.1016/j.jaac.2024.05.026

McVay, E. (2023). *Social Media and Self-diagnosis*. John Hopkins Medicine. https://www.hopkinsmedicine.org/news/articles/2023/08/social-media-and-self-diagnosis#:~:text=This%20can%20also%20delay%20access,represent%20their%20experience%20or%20symptoms.

Munoz, I., Liu-Zarzuela, J., Oorjitham, N., & Jacob, D. (2024). YouTube and TikTok as a source of medical information on dissociative identity disorder. *Journal of Affective Disorders Reports, 15*. https://doi.org/10.1016/j.jadr.2023.100707

Mutlu, F., & Arik, E. (2023). Quality and reliability analysis of YouTube videos on magnetic resonance imaging claustrophobia. *Cureus, 15*(4). https://doi.org/10.7759/cureus.37648

Neely, S., Eldredge, C., & Sanders, R. (2021). Health Information Seeking Behaviours on Social Media During the COVID-19 Pandemic Among American Social Networking Site Users: Survey Study. *Journal of Medical Internet Research, 23*(6). https://doi.org/10.2196/29802

Nickerson, R. (1998). Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. *Review of General Psychology, 2*(2), 175-220. https://doi.org/10.1037/1089-2680.2.2.175?urlappend=%3Futm_source%3Dresearchgate.net%26utm_medium%3Darticle

Niu, S., & Reed, K. (2023). Towards understanding substance abuse misinformation in YouTube videos. *CSCW '23 Companion: Companion Publication of the 2023 Conference on Computer Supported Cooperative Work and Social Computing*, 64-68. https://doi.org/10.1145/3584931.3606974

Page, M., McKenzie, J., Bossuyt, P., Boutron, I., Hoffmann, T., Mulrow, C., Shamseer, L., Tetzlaff, J., Akl, E., Brennan, S., Chou, R., Glanville, J., Grimshaw, J., Hróbjartsson, A., Lalu, M., Li, T., Loder, E., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Research Methods & Reporting.* https://doi.org/10.1136/bmj.n71

Patel, P., Nagare, M., Randhawa, J., Ali, A., & Olivieri, L. (2023). Bipolar disorder in social media: An examination of Instagram's role in disseminating accurate information. *Cureus, 15*(9). https://doi.org/10.7759/cureus.46296

QuickFrame. (2023). *TikTok vs Facebook: Key Differences for Brands to Know.* QuickFrame. https://quickframe.com/blog/tiktok-vs-facebook-key-differences-to-know/

Shoemaker, S., Wolf, M. & Brach, C. (2014). Development of the Patient Education Materials Assessment Tool (PEMAT): a new measure of understandability and actionability for print and audiovisual patient information. *Patient Education and Counselling, 96*(3), 395-403. https://doi.org/10.1016/j.pec.2014.05.027

Singh, A., Singh, H., & Singh, P. (2012). YouTube for Information on Rheumatoid Arthritis — A Wakeup Call? *The Journal of Rheumatology, 39*(5), 899-903. https://doi.org/10.3899/jrheum.111114

Starvaggi, I., Dierckman, C., & Lorenzo-Luaces. (2024). Mental health misinformation on social media: Review and future directions. *Current Opinion in Psychology, 56.* https://doi.org/10.1016/j.copsyc.2023.101738

Suarez-Lledo, V., & Alvarez-Galvez, J. (2021). Prevalence of health misinformation on social media: systematic review. *Journal of Medical Internet Research, 23*(1), https://doi.org/10.2196/17187

Sumayyia, M., Al-Madaney, M., & Almousawi, F. (2019). Health information on social media. Perceptions, attitudes, and practices of patients and their companions. *Saudi Medical Journal, 40*(12), 1294-1298. https://doi.org/10.15537/smj.2019.12.24682

Suresh, A., Pallempati, L., Saxena, P., Ansari, A., Bassi, R., & Bhandari, A. (2023). Exploring YouTube videos about anorexia nervosa on the basis of reliability, popularity, and contributions of healthcare professionals: A cross-sectional study. *Cureus, 15*(11), https://doi.org/10.7759/cureus.48095

Syed-Abdul, S., Fernandez-Luque, L., Jian, W., Li, Y., Crain, S., Hsu, M., Wang, Y., Khandregzen, D., Chuluunbaatar, E., Nguyen, P., Liou, D. (2013). Misleading health-related information promoted through video-based social media: Anorexia on YouTube. *Journal of Medical Internet Research, 15*(2). http://dx.doi.org/10.2196/jmir.2237

Tam, J., Porter, E., & Lee, U. (2022). Examination of Information and Misinformation about Urinary Tract Infections on TIkTok and YouTube. *Urology, 168,* 35-40. https://doi.org/10.1016/j.urology.2022.06.030

Thapa, P., Thapa, A., Khadka, N., Bhattarai, R., Jha, S., Khanal, A., & Basnet, B. (2018). YouTube lens to attention deficit hyperactivity disorder: a social media analysis. *BMC Research Notes, 11*(854). https://doi.org/10.1186/s13104-018-3962-9

UIDesignz. (2024). *TikTok Live vs. YouTube Live: Which Platform is Better for You?* Medium. https://medium.com/@uidesign0005/tiktok-live-vs-youtube-live-which-platform-is-better-for-you-38512120dc74#:~:text=If%20you're%20looking%20for,a%20sustainable%20live%2Dstreaming%20career.

Wang, Y., McKee, M., Torbica, A., & Stuckler, D. (2019). Systematic literature review on the spread of health-related misinformation on social media. *Social Science & Medicine, 240.* https://doi.org/10.1016/j.socscimed.2019.112552

Ward, M., Ward, B., Warren, C., Silverstein, S., Ray, C., Paskhover, B., Kornitzer, J. (2020). The quality of YouTube videos as an educational resource for attention-deficit/hyperactivity disorder. *Paediatric Neurology, 103*, 84-85. https://doi.org/10.1016/j.pediatrneurol.2019.04.001

Wardle, C., & Derakhshan, H. (2017). *Information Disorder: Toward an interdisciplinary framework for research and policy making.* Council of Europe.

Yeung, A., Ng, E., & Abi-Jaoude, E. (2022). TikTok and attention-deficit/hyperactivity disorder: A cross-sectional study of social media content quality. *The Canadian Journal of Psychiatry, 67*(12), 899-906. https://doi.org/10.1177/07067437221082854

YouTube Kids. (n.d.). A safer online experience for children. https://www.youtube.com/kids/safer-experience/

Zhao, Y., & Zhang, J. (2017). Consumer health information seeking in social media: a literature review. *Health Information & Libraries Journal, 34*(4), 268-283. https://doi.org/10.1111/hir.12192