UNIVERSITY OF WOLVERHAMPTON

How many zebras are there now? An updated report on publications on rare diseases in medical literature

Item Type	Journal article
Authors	Moroz, Tara;Taylor, Michael;Schmitt, Paula;Donovan, Daniel J.;Nori, Mukund
Citation	Moroz, T., Taylor, M., Schmitt, P., Donovan, D.J. and Nori, M. (2025) How many zebras are there now? An updated report on publications on rare diseases in medical literature, Future Rare Diseases, 5:1, article number 2499402, DOI: 10.1080/23995270.2025.2499402
DOI	10.1080/23995270.2025.2499402
Publisher	Taylor & Francis
Journal	Future Rare Diseases
Download date	2026-01-15 16:02:37
License	https://creativecommons.org/licenses/by-nc/4.0/
Link to Item	https://wlv.openrepository.com/handle/2436/625942

FUTURE RARE DISEASES
2025, VOL. 5, NO. 1, 2499402
https://doi.org/10.1080/23995270.2025.2499402

Taylor & Francis
Taylor & Francis Group

REVIEW ARTICLE

OPEN ACCESS Check for updates

How many zebras are there now? An updated report on publications on rare diseases in medical literature

Tara Moroz[a], Michael Taylor[b,c], Paula Schmitt[d*], Daniel J. Donovan[e] and Mukund Nori[e]

[a]Pfizer Inc., New York, NY, USA; [b]Digital Science, London, UK; [c]University of Wolverhampton, Wolverhampton, UK; [d]Amyloidosis Support Groups, Inc., Wood Dale, IL, USA; [e]rareLife solutions, Westport, CT, USA

ABSTRACT

Previously, we reported that medical literature on rare diseases was both scarce and mostly inaccessible. In this study, we evaluated changes in the past 5 years. We performed a systematic search of the Dimensions database for rare disease-related publications from 2019 to 2023 using the same terms as in our previous paper and compared the results to our earlier findings. We also evaluated publication patterns, including open access, publication of plain language summaries, patient-reported outcomes, and social media posts. The proportion of rare disease publications remained low and relatively stable in Dimensions over the past 15 years (0.20% [2009–2013], 0.24% [2014–2018], and 0.25% [2019–2023]). The majority of recent (2019–2023) publications were research articles (65.32%), published in specialist journals (39.48%), with 72.58% published as open access. Only 0.42% of recent articles had plain language summaries. Social media mentions of recent rare disease publications (39%) were comparable with their proportion in the overall medical literature (38%). The paucity of rare disease-related publications highlights an urgent need for increasing knowledge on rare diseases through medical literature to better serve patients, their families, and the healthcare professionals who treat them. Social media increasingly is being used to share information.

ARTICLE HIGHLIGHTS

- Despite the apparent increased awareness of rare diseases, the proportion of publications on rare diseases in the medical literature has remained ~0.2% during 2019–2023 and is similar to that in the previous 10 years.
- On a positive note, the proportion of rare disease articles available as open access has increased to ~70% [please check the number] during 2019–2023.
- There is a scarcity of plain language summaries in the rare disease arena, despite their well-known added advantage of facilitating patient-focused communication.
- Articles with patients, advocates, and caregivers as authors remain a small minority indicating that their perspectives on disease and treatment are being missed by the medical community.
- Social media is becoming increasingly influential in disseminating information on rare diseases.

PLAIN LANGUAGE SUMMARY

This article provides an update of an earlier publication in which we reported on rare disease publications in PubMed during the 10-year period from 2009 to 2018. We found that publications on rare diseases were scarce, comprising about 0.2% of all publications with over 60% being inaccessible to the public. Here, we analyzed for changes in the next 5 years (2019–2023). We performed a systematic search of the Dimensions database, which includes more publications and a greater variety of publications than PubMed. For this manuscript, we used the same terms that were used for the earlier publication to make sure that the results could be compared. We extended our analysis to include plain language summaries, patient-reported outcomes, and social media posts. Unfortunately, the proportion of rare disease publications remained relatively unchanged (0.2% [2009–2013], 0.2% [2014–2018], and 0.3% [2019–2023]). The majority of recent (2019–2023) publications were research articles (65%) published in specialist journals (40%). On a positive note, 73% of rare disease articles in 2019–2023 timeframe were published as open access (up from about 33%, previously). However. only 0.4% of these articles had plain language summaries. Social media mention of recent rare disease publications was comparable with the proportion for the overall medical

ARTICLE HISTORY

Received 3 October 2024
Accepted 25 April 2025

KEYWORDS

Rare disease publications; open access; plain language summary; patients, advocates, caregivers; social media

CONTACT Mukund Nori mnori@rarelifesolutions.com rareLife solutions, Westport, CT, USA
*Patient author.

© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

literature. The low number of rare disease-related publications remains an unmet fundamental requirement that needs to be addressed urgently and could significantly benefit physicians, patients, caregivers, and payers. These publications are essential for increasing knowledge to better serve patients, their families, and the healthcare providers who treat them.

1. Introduction

According to reports, Dr. Theodore Woodward once instructed his medical student class, "When you hear hoofbeats, think horses, not zebras," meaning that the majority of the diagnoses they would be making would be common (horses), and there was very little chance that it would be a rare disease (zebras) [1]. This has formed the basis of medical education. Students learn about the ~400 diseases common to their part of the world and are required to be able to accurately diagnose at least 170 diseases for graduation [2,3]. Thus, there is very little focus on teaching about rare diseases.

Definitions of rare disease vary ranging from <5 to 40 per 10,000 individuals in the European Union and rest of the world to <200,000 affected individuals in the United States [4–8]. Rare diseases affect only a small percentage of the population, and their prevalence varies around the world [9]. Although each rare disease affects only a small number of patients, collectively they account for 263 million to 446 million people (3.5–5.9%) worldwide—comparable to the population of the United States, the world's third largest country by population [9–11].

There are an estimated 7000–10,000 rare diseases in the world [12]. Approximately 80% of them are due to genetic abnormalities, 50–75% begin in childhood, and sadly, 30% of children with a rare disease will not survive to their 5th birthday [9,13]. Moreover, early and accurate diagnosis and management of rare diseases often are challenging [14]. Despite therapeutic advancements in rare diseases, only about 5% of rare diseases have a US Food and Drug Administration (FDA)-approved treatment with <15% having a promising product for their diagnosis, treatment, or prevention [12,15]. The low prevalence of individual rare diseases and scarcity of effective treatments have contributed to inadequate knowledge and understanding of rare diseases by healthcare professionals (HCPs) and patients, advocates, and caregivers (PACs) [16].

Managing rare diseases is often difficult for HCPs due to lack of reliable sources of information and/or difficulty accessing the available information [17]. The primary source of information for HCPs remains peer-reviewed medical literature (e.g., PubMed) and some reputable curated sources (e.g., Online Mendelian Inheritance in Man [OMIM®], and Orphanet) [18–22]. The advent of COVID-19 accelerated the reliance of HCPs on the internet and social media for seeking medical information [19,23]. The lack of open access to peer-reviewed publications poses an even greater barrier to PACs searching for information on their disease and its management. Consequently, although some curated sources of information are available, it is common for PACs to use the internet as their initial and free source of information [24,25]. The rapid expansion of social media over the last few years paved the way for PACs to scour the various social media platforms for information, support, and advocacy [26,27]. However, accelerated reliance on the internet and social media for medical information makes them more vulnerable to misinformation [28–31]. Misinformation, misleading, and inaccurate information can cause confusion for the HCPs and PACs and can result in inappropriate treatment decisions with potentially drastic consequences [31–34]. On the other hand, the same internet and social media access can be harnessed for rapid and unprecedented exchange of accurate information to reach patients with rare diseases and the HCPs who treat them [28–31]. Thus, high-quality, open-access, peer-reviewed medical literature and reliable social media platforms are critical for dissemination of healthcare information [35].

Our prior publication showed that despite the large number of rare diseases, <0.2% of medical literature is devoted to these conditions [36]. It also revealed that <33% of the articles evaluated were available *via* open access. The dearth of knowledge combined with most of the articles being behind a pay wall constitutes a significant barrier to HCPs and PACs trying to access information about rare diseases.

In our current analysis, we extended the analysis to the next 5 years (2019–2023) and evaluated the rare disease share of voice in medical literature, availability of plain language summaries (PLS), inclusion of

patient-reported outcomes/quality of life (PRO/QoL), and PACs contributions. We also expanded our research to include the reach of these publications in social media to reflect the evolving patterns of information dissemination.

2. Methods

2.1. Search parameters

We searched the Dimensions database for publications on rare diseases. The Dimensions database (https://www.dimensions.ai/) is a freely available tool that widens research evaluation beyond publication citation analysis by including publications and their citations, awarded grants, clinical trials, patents, and policy papers [37,38]. Dimensions sources data from CrossRef, PubMed, PubMed Central, and preprint servers, and is enriched by publisher-supplied data. Our search strategy included articles in the 5-year period from 1 January 2019 to 31 December 2023, using the same terms as in our previous analysis related to rare or orphan diseases (Table 1) [36,38,39]. In addition, we searched for the availability of a PLS (Table 1). We excluded animal studies using the search string "rare disease" OR "orphan drug" OR "orphan disease" OR "rare disorder" OR MeSH_terms: "Rare Diseases" NOT (rat OR mouse OR chimpanzee OR monkey OR cat OR dog OR zoo OR animal) and the search was limited for the term "human." The search output was scrutinized to include only those articles that met the search criteria. For comparison and validation with our previous results, searches were also conducted for the 10-year period from 1 January 2009 to 31 December 2018, and separately for the consecutive 5-year periods from 1 January 2009 to 31 December 2013, and 1 January 2014 to 31 December 2018. For comparison, we evaluated the overall proportion of publications on diabetes published between 2009 and 2023 in PubMed and other health and clinical research not indexed in PubMed. We also evaluated the publication patterns for journal category and availability as open access (OA), as measured by both Gold/Bronze and Green. Gold OA includes articles that are freely available online; Bronze OA includes articles that are available to read on publishers' page; and Green OA means "self-archiving" of published articles in a freely accessible database [40–42].

2.2. Data analysis

The output of the search of Dimensions (abstracts and full-length papers) was evaluated for availability as OA, article type (research article, literature review, case study, editorial, letter to editor, conference abstract, conference paper, PLS, and others), inclusion of PRO/QoL, and inclusion of PACs perspectives and/or as authors. An analysis was performed in Dimensions, and journals were manually assigned to different specialties based on their title, coverage, and subject areas. Articles published in a medical journal were evaluated as social media (publications shared on platforms, such as X, Facebook pages, and Reddit), news and blog sites (mentions of publications on news and blog websites), and policy/clinical decision guidelines (CDG) metrics (growth in

Table 1. Search strategy used in dimensions database.

Particulars	Search string
Articles on rare or orphan diseases	rare disease OR rare disorder OR orphan disease OR orphan drug which included: ('rare diseases'[MeSH Terms] OR ('rare'[All Fields] AND 'diseases'[All Fields]) OR 'rare diseases'[All Fields] OR ('rare'[All Fields] AND 'disease'[All Fields]) OR 'rare disease'[All Fields]) OR ('rare diseases'[MeSH Terms] OR ('rare'[All Fields] AND 'diseases'[All Fields]) OR 'rare diseases'[All Fields] OR ('rare'[All Fields] AND 'disorder'[All Fields]) OR 'rare disorder'[All Fields]) OR ('rare diseases'[MeSH Terms] OR ('rare'[All Fields] AND 'diseases'[All Fields]) OR 'rare diseases'[All Fields] OR ('orphan'[All Fields] AND 'disease'[All Fields]) OR 'orphan disease'[All Fields]) OR ('orphan drug production'[MeSH Terms] OR ('orphan'[All Fields] AND 'drug'[All Fields] AND 'production'[All Fields]) OR 'orphan drug production'[All Fields] OR ('orphan'[All Fields] AND 'drug'[All Fields]) OR 'orphan drug'[All Fields])
Journal specialty	"rare disease" OR "orphan drug" OR "orphan disease" OR "rare disorder" OR mesh_terms: "Rare Diseases" NOT (rat OR mouse OR chimpanzee OR monkey OR cat OR dog OR zoo)' in title and abstract; Publication Year is 2013 or 2012 or 2011 or 2010 or 2009; Publication Type is Article or Preprint or Proceeding
Case study	rare disease OR "orphan drug" OR "orphan disease" OR "rare disorder" OR mesh_terms:"Rare Diseases" AND (("case study" OR "case report" OR "case series" OR "case history" OR "case analysis"~5) AND (method* OR design* OR approach* OR technique* OR framework* OR investing*)) NOT (rat OR mouse OR chimpanzee OR monkey OR cat OR dog OR zoo)
PLS	("rare disease" OR "orphan drug" OR "orphan disease" OR "rare disorder" OR mesh_terms: "Rare Diseases") AND full_text: ("plain language abstract" OR "plain language summary" OR "lay summary" OR "lay abstract") NOT (rat OR mouse OR chimpanzee OR monkey OR cat OR dog OR zoo)

PLS: plain language summaries.

probability of novel attention received by the publications) [43]. Where appropriate, one-sided t-tests were performed using GraphPad Prism v10.1.0.

3. Results

3.1. Publications on rare diseases in dimensions

In the 5-year period from 1 January 2019 to 31 December 2023, there were a total of 23,588,784 articles in the Dimensions database of which 55,134 (0.23%) were on rare diseases (Table 2). The majority of articles were published in English language (86.70% in 2009–2013; 90.80% in 2014–2018; 94.40% in 2019–2023).

3.2. Use of open access

In Dimensions, the proportion of articles on rare diseases that were available as OA increased significantly from 42.30% (35.88%, Gold/Bronze; 6.42%, Green) in 2009–2013 to 56.85% (48.72%, Gold/Bronze; 8.13% Green) in 2014–2018 to 72.58% (62.48%, Gold/Bronze; 10.10% Green) in 2019–2023 ($p=0.0225$). The corresponding proportion of articles not available as OA significantly decreased from 57.70 to 27.42% ($p=0.0394$) over the same period (Figure 1).

3.3. Journal specialty

In Dimensions, the majority of the articles on rare diseases were published in specialist journals (cardiology, neurology, dermatology, oncology, etc.) across the 3 time periods (52.81%, 2009–2013; 45%, 2014–2018; 39.48%, 2019–2023; Figure 2). A small number of the articles were published in journals dedicated to rare diseases across the 3 time periods (1.91%, 2009–2013; 2.15%, 2014–2018; 2.90%, 2019–2023).

3.4. Article type

Almost two-thirds (67.32%) of the articles in the Dimensions database pertaining to rare diseases were classified as research articles and only about 1% were classified as case studies (Figure 3). Additionally, only a small

Table 2. Comparison of rare disease articles in dimensions with prior published data from PubMed.

Database	Total	2009–2013	2014–2018	2019–2023
PubMed				
Total, *N*	10,882,681	4,764,434	6,118,247	7,461,867
Articles on rare diseases, *n* (%)	14,202 (0.13)	5675 (0.12)	8527 (0.14)	13,254 (0.18)
Dimensions				
Total, *N*	23,588,784	5,784,811	7,544,289	10,259,684
Articles on rare diseases, *n* (%)	55,134 (0.23)	11,350 (0.20)	17,735 (0.24)	26,049 (0.25)
PLS	227 (0.41)	51 (0.45)	54 (0.30)	122 (0.47)
Open access	33,788 (61.28)	4801 (42.30)	10,081 (56.84)	18,906 (72.58)

PLS: plain language summaries.

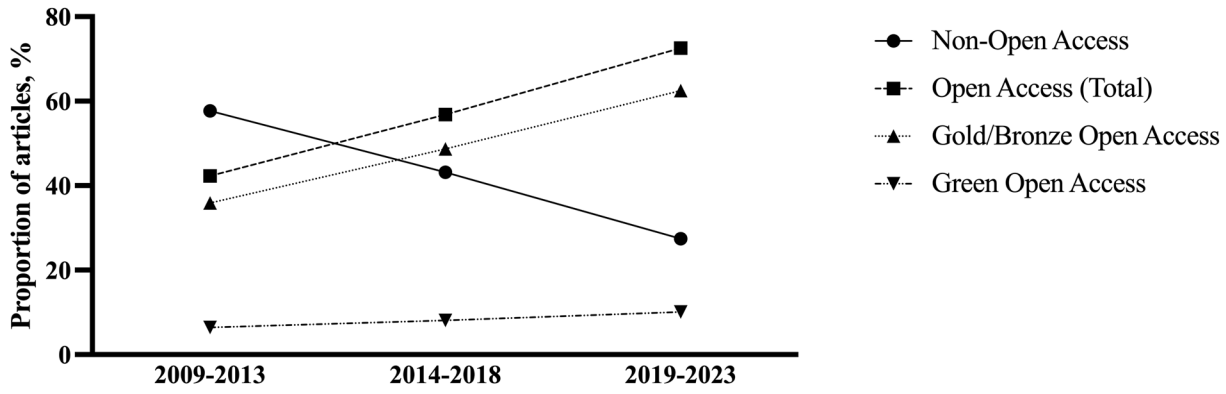

Figure 1. Proportion of articles available as open access over time.

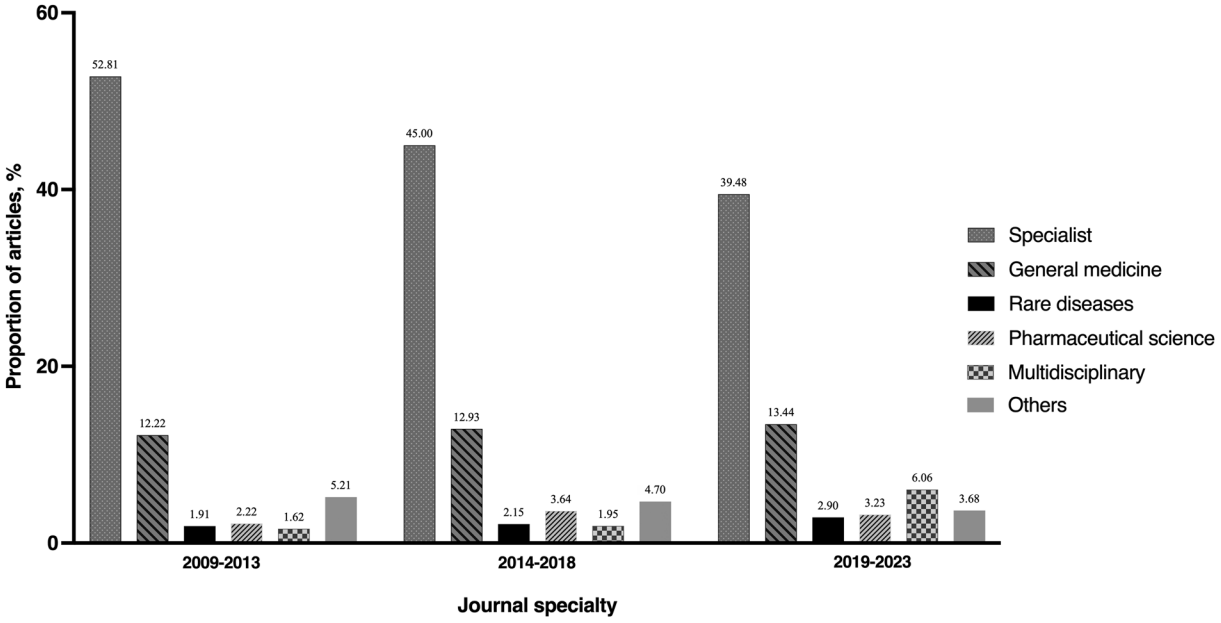

Figure 2. Specialty of medical journal in which articles on rare diseases are published.

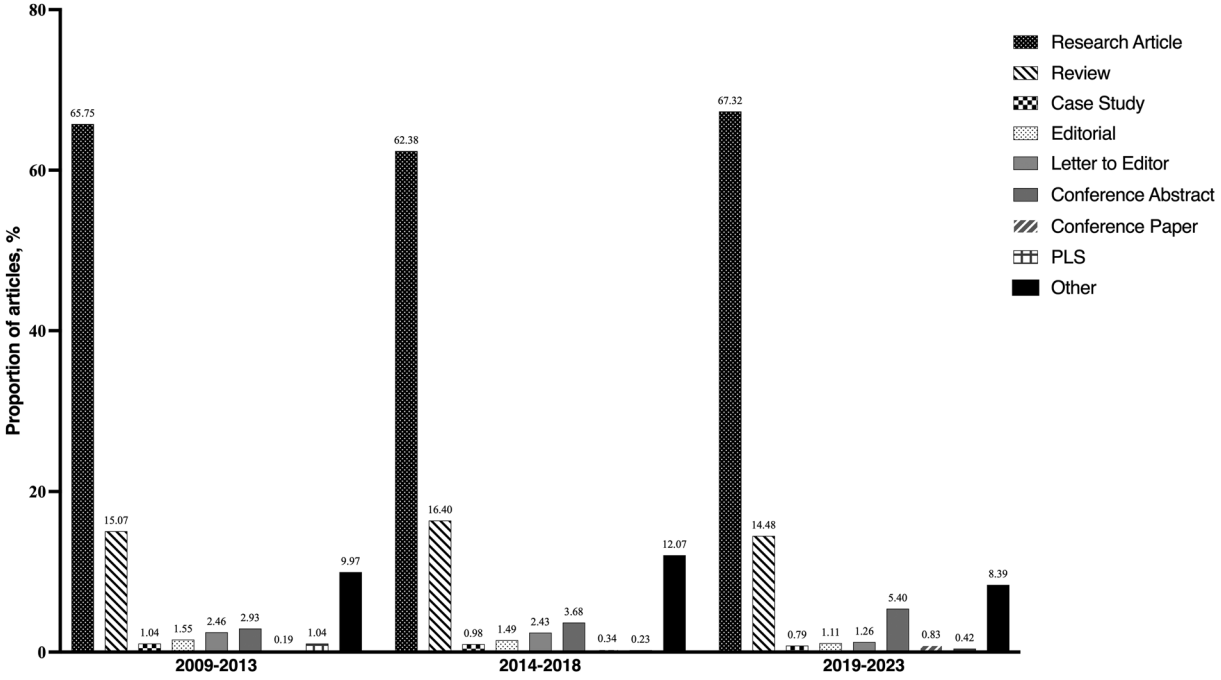

Figure 3. Evaluation of the categories of articles on rare diseases in the dimensions database across time. PLS: plain language summaries.

proportion of the articles were classified as PLS (0.42%). The relative proportions for each article type were similar across the 3 time periods.

3.5. Inclusion of PAC, PRO, and QoL

In Dimensions, articles reporting PAC contributions and inclusion of PRO/QoL have gradually increased since 2009; 44 (0.39%) articles reported PRO/QoL and 192 (1.69%) had PAC contributions between 2009 and 2013, 118 (0.66%) articles reported PRO/QoL and 441(2.48%) had PAC contributions between 2014 and 2018, and 203 (0.77%) articles reported PRO/QoL and 756 (2.90%) had PAC contributions between 2019 and 2023 (Figure 4).

3.6. Social media, news, blog sites, and policy/clinical decision guideline metrics

Social media, news, and blog sites provide means of disseminating summarized information from publications, including those on rare diseases, for quick and easy consumption by the PAC and HCPs. We added social media to this search going back to the original dates of the search in light of the significant uptick in the use of social media to consume medical literature [44]. Overall, public engagement with medical research publications on social media and the coverage of research on news and blog sites have risen dramatically since 2009 (Figure 5). Social media activity jumped dramatically from the initial 5-year period (2009–2013) when 17.63% of rare publications were mentioned in social media, to 41.13% in the second 5-year period (2014–2018), then remained flat at 39.02% into the third 5-year period (2019–2023). The lack of growth between the second and third periods may suggest social media saturation for medical literature attention. Of note, the level of social media activity for rare disease publications was comparable to those of all medical literature (17.14, 36.89, and 38.34% in 2009–2013, 2014–2018, and 2019–2023, respectively).

Citations from policy documents and CDG for rare disease publications were 2.87% in 2009–2013, 2.28% in 2014–2019, and 0.54% in 2019–2023. Citations from policy documents and CDG for rare disease publications were comparable to those of all medical literature (4.09, 2.88, and 0.97% in 2009–2013, 2014–2018, and 2019–2023, respectively).

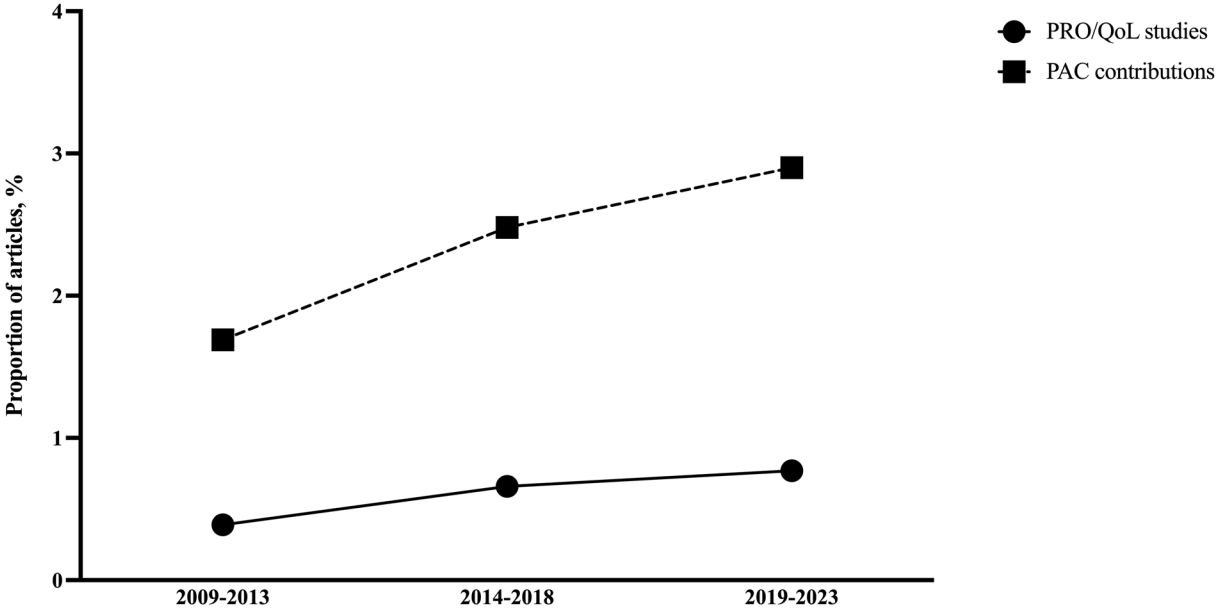

Figure 4. Proportion of articles on rare diseases with patients, advocates, or caregivers as authors and inclusion of PRO/QoL by time. PAC: patients, advocates, and caregivers; PRO/QoL: patient-reported outcomes/quality of life.

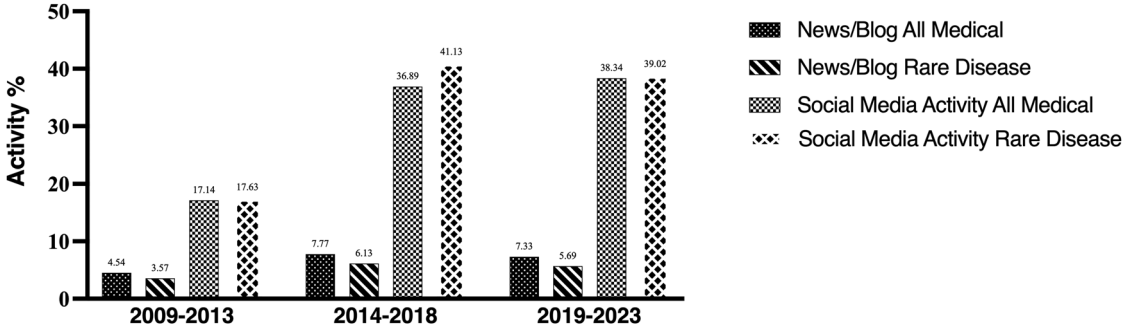

Figure 5. Relative attention to rare disease publications in social media and news/blogs across time.

4. Discussion

The current study provided an update on the representation of rare disease publications in the medical literature database and the extended reach of those publications *via* social media. We selected the Dimensions database as it is freely available and covers a greater context of research, including awarded grants, patents, clinical trials, policy documents, and altmetric information in addition to the standard publication ecosystem [38]. The Dimensions database is currently the most robust database available to search for medical literature and includes articles that are not available *via* standard PubMed search. Thus, we believe that our analysis provides a robust sample of the share of voice of rare disease literature and its reach through social media.

Our analysis revealed that despite rare diseases directly affecting about 400 million people, which is only slightly less than that for diabetes (~529 million people globally [45]), the overall proportion of publications on rare diseases published between 2009 and 2023 is an order of magnitude below that of publications in diabetes alone (0.23% compared with 2.84% for diabetes). We reported <0.2% of rare disease-related publications in our previous analysis [36]. The small difference in the proportion of rare disease-related publications could be due to the English language limitation in the original publication. We also found that the proportion of publications on rare diseases has remained relatively flat for each of the 5-year periods during the last 15 years. Our analysis underscores a continued dearth of research publications on rare diseases and an urgent need to increase such publications in medical literature.

Peer-reviewed medical publications for accurate information on disease treatment remain the preferred source of information for HCPs [19,46–48]. Therefore, the scarcity of publications on rare diseases may impact a physician's ability to recognize symptoms early, make timely diagnoses, and initiate treatment early enough to improve the patient's QoL. PACs often become experts in their own disease, and they often share their information with HCPs [49,50]. Trusted sources of information, especially peer-reviewed journal articles, are essential for HCP and PAC education, and for shared-decision making [49].

Articles behind journal paywalls pose an access barrier, especially to PACs [49]. Thus, PACs frequently rely on the internet for medical information despite apparent lack of credibility or reliability of the source [49]. Freely accessible publications are also imperative in areas where regional economics make it impossible for HCPs and PACs to pay for access [36]. Our previous publication showed that a substantial proportion of articles were published behind the paywall [36]. A positive trend uncovered by the current analysis is a large increase in the number of OA publications across each 5-year period, with over 70% of rare disease publications freely accessible in the most recent time period (2019–2023). This is a welcome trend, as it eliminates at least one barrier by making the scant rare disease literature widely accessible to those who need it most.

Reflecting our previous article, our current data show that most articles on rare diseases continue to be published in specialist journals. This may be due to familiarity with established journals or lack of awareness of rare disease journals, many of which are new and not as well established. For some rare diseases, e.g., rare cancers, publishing in specialist journals may be the most effective way to disseminate the information as these are the journals most widely read by HCPs. We suggest that these articles be open access so that they are equally available to PACs. Articles on ultra-rare diseases or those for which there are limited treatment options may be best served by being published in journals focused on rare diseases where they are more likely to reach their intended audience.

Our previous publication showed that a large number of rare disease articles published in 2009–2018 and listed in PubMed were case reports or case series [36]. That analysis was conducted manually on 200 randomly selected articles or 1.4% of the 14,202 articles listed. In our current study, all articles in the Dimensions database could be electronically evaluated for type of article. It is notable that most of the publications in this database are curated as research articles with reviews being the second most common type. The difference in results suggests that curation criteria may be different between the 2 databases and that the 200 randomly selected articles in our earlier analysis may not have been a representative sample, despite a sensitivity analysis performed to test for bias. In the current analysis, we also could capture the proportion of articles with PLS in the database. Surprisingly, in spite of efforts made by various organizations, including the FDA, to include patients in rare disease research and shared decision-making, the number of PLS has not changed appreciably in the last 15 years and remains low at <0.05% of rare disease publications [51]. PLS have an added advantage in patient-focused communication as they are written in a non-technical language that enables a diverse readership, including laymen, to understand the ongoing research on different diseases [52].

Our current study highlights the scarcity of PLS in the rare disease arena and emphasizes a remaining need for more concerted efforts to publish them.

In contrast to our previous findings, PACs contributions and PRO/QoL inclusions in rare disease-related publications in Dimensions have gradually increased since 2009. Our current results are consistent with the ongoing efforts of the FDA encouraging stakeholders to collect and report patient experience data (PED) from PACs [51]. PED collected *via* natural history studies, clinical trials, patient registries, patient surveys, and other patient input techniques on benefits and risks can be critical as it may support therapy development [51]. In addition, the safety and efficacy of a new therapy are also evaluated by assessing PRO/QoL in clinical trials that broadly include the way a patient feels, functions, or survives [51]. Although there has been a noticeable increase in the number of PACs contributions and PRO/QoL inclusions in rare disease-related publications, a welcome trend in the right direction, the number still is very low when compared with the total number of rare disease articles published. Thus, it underscores the continuing need for the inclusion of PED across all stages of research and therapy development.

The COVID-19 pandemic accelerated the use of social media for the identification and consumption of medical literature by HCPs and PACs [23,24]. Our findings suggest that public engagement with medical research and rare disease-related publications on social media has substantially increased over the past decade. Notably, the social media attention for rare disease publications was found comparable to attention for other medical research publications. This finding suggests that rare disease publications, in spite of being infrequent, received social media attention comparable to publications of medical conditions in general. This trend may be the result of most journals setting up dedicated social media accounts, such as X (formerly Twitter), to share pertinent content with their audiences [53]. Additionally, publishers often urge researchers to post their findings on their personal social media accounts [54]. However, sharing through personal accounts may involve channels that lack credibility or reliability [23]. Moreover, the sources of information PACs rely on for their knowledge can be a matter of concern [24]. As society moves toward gathering information in small bites *via* the internet and specifically from social media platforms, it is important to have a comprehensive and reliable social media plan for rare publications to try to reach as many people in the community as possible to increase awareness of the disease and to educate the patients, families, and physicians who seek information *via* these channels.

Citations from policy documents and CDG are known to be amongst the slowest indicator to accumulate. The proportion of rare disease research that gets citations from policy documents and CDG is much lower than the baseline of all medical research for the older time periods, whereas it's broadly in line for the more recent time period. This may indicate an uptick in the translation of research findings into clinical practice and guidelines, or it may indicate an increase in the number of policy documents and CDG related to rare diseases.

5. Limitations

Although the Dimensions database is robust, extensive and includes publications and categories not listed in PubMed, our analysis is limited by how these entries are curated. This is a limitation of all investigations involving databases. Including additional databases, e.g., Scopus and Web of Science, would make our analysis more robust. However, PubMed and Dimensions are free and therefore more accessible than the other databases which require payment for access, one of the central issues we raise in this paper. Nevertheless, both PubMed and Dimensions require access to the internet, which is not globally ubiquitous. Dimensions requires registration for use, but registration is free. Furthermore, both databases require a working knowledge of the English language, which many people may not have. HCPs and PACs across most of the world prefer receiving information in their own native language. Although this barrier can be overcome by translating published articles into other languages with appropriate attribution to the original article, this analysis is beyond the scope of our current work. A comparison of the volume of basic, clinical, and policy literature on rare diseases over time was also beyond the scope of this manuscript. Such an analysis may provide additional insights into barriers and opportunities for rare disease knowledge generation and communication with all stakeholders.

6. Conclusions

Our current analysis shows that the proportion of articles on rare diseases has stayed relatively flat for the past 15 years, is underrepresented compared with diseases affecting a similar number of people (e.g., diabetes), and

is a very small fraction of the total medical literature database. HCPs and PACs are reliant on scarce medical literature as their source of primary information for rare diseases, so it is imperative that all reliable information be published and freely accessible (open access). Inclusion of PACs in rare disease research, including publications, is critical for fuller understanding of the burden of rare diseases and patient journey. More needs to be done to ensure the PAC perspective is included in rare disease research. The impact of social media is evolving and still to be realized in full measure but should be considered as part of the publication process. The urgent need remains to increase the knowledge and understanding of rare diseases for all stakeholders or else the zebras will remain lost in a herd of horses.

7. Future perspectives

As the number of rare diseases identified increases, we anticipate an increase in HCPs specializing in these diseases. We hope this dynamic will be associated with a corresponding increase in medical literature focused on rare diseases to address the gaps in knowledge on the natural history of rare diseases, their manifestations, and underlying etiologies, and potential treatments and management. As technology advances, one can only hope and await cures for at least a few of the rare diseases for which etiology is already known. Repeating this analysis at 5-year intervals could provide insights into the progress, or lack thereof, in rare diseases communications in medical literature.

Author contributions

All authors contributed to the manuscript concept, data analysis and interpretation, critical review of content, and provided final approval of the manuscript.

Disclosure statement

TM is an employee of Pfizer Inc. and holds stock in the company. MT is an employee of Digital Science. PS has no conflicts to disclose. DJD and MN are employees of rareLife solutions and hold stock in the company. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

Medical writing support was provided by Sonali K. Kalra, PhD, of rareLife solutions, Westport, CT, USA and funded by Pfizer, Inc.

Ethical declaration

Mukund Nori is a member of the Future Rare Diseases Editorial Board. They were not involved in any editorial decisions related to the publication of this article, and author details were not made available to the article's peer reviewers as per the journal's double-anonymized peer review policy.

Funding

Medical writing support was funded by Pfizer, Inc.

ORCID

Mukund Nori http://orcid.org/0000-0001-6165-1825

References

Papers of special note have been highlighted as either of interest (*) or of considerable interest (**) to readers.

[1] Hart AM, Moore KS. When you hear hoofbeats, think horses—but be prepared for zebras. J Nurse Pract. 2019;15(6):A10. doi: 10.1016/j.nurpra.2019.03.025

[2] Babac A, Litzkendorf S, Schmidt K, et al. Shaping an effective health information website on rare diseases using a group decision-making tool: inclusion of the perspectives of patients, their family members, and physicians. Interact J Med Res. 2017;6(2):e23. doi: 10.2196/ijmr.7352
[3] Urushibara-Miyachi Y, Kikukawa M, Ikusaka M, et al. Lists of potential diagnoses that final-year medical students need to consider: a modified Delphi study. BMC Med Educ. 2021;21(1):234. doi: 10.1186/s12909-021-02652-5
[4] Moliner AM, Waligora J. The European Union policy in the field of rare diseases. Adv Exp Med Biol. 2017;1031:561–587. doi: 10.1007/978-3-319-67144-4_30
[5] United States Food and Drug Administration. Rare diseases: common issues in drug development: guidance for industry. Silver Spring (MD): United States Food and Drug Administration; 2019.
[6] Rare Diseases Act of 2002. Washington (DC): US Congress; 2002.
[7] Larizza L, Cubellis MV. Rare diseases: implementation of molecular diagnosis, pathogenesis insights and precision medicine treatment. Int J Mol Sci. 2023;24(10):9064. doi: 10.3390/ijms24109064
[8] Haendel M, Vasilevsky N, Unni D, et al. How many rare diseases are there? Nat Rev Drug Discov. 2020;19(2):77–78. doi: 10.1038/d41573-019-00180-y
[9] Chung CCY, Chu ATW, Chung BHY. Rare disease emerging as a global public health priority. Front Public Health. 2022;10:1028545. doi: 10.3389/fpubh.2022.1028545
[*10] Nguengang Wakap S, Lambert DM, Olry A, et al. Estimating cumulative point prevalence of rare diseases: analysis of the Orphanet database. Eur J Hum Genet. 2020;28(2):165–173. doi: 10.1038/s41431-019-0508-0Important reference for prevalence of rare diseases.
[11] Auvin S, Irwin J, Abi-Aad P, et al. The problem of rarity: estimation of prevalence in rare disease. Value Health. 2018;21(5):501–507. doi: 10.1016/j.jval.2018.03.002
[**12] Fermaglich LJ, Miller KL. A comprehensive study of the rare diseases and conditions targeted by orphan drug designations and approvals over the forty years of the Orphan Drug Act. Orphanet J Rare Dis. 2023;18(1):163. doi: 10.1186/s13023-023-02790-7A reference that provides an overview of the huge unmet need in treatments for rare diseases.
[13] Narayanan G. Rare diseases: a common problem. J Rare Dis Diag Ther. 2015;1(15):1–3.
[14] Zanello G, Chan CH, Pearce DA. Recommendations from the IRDiRC Working Group on methodologies to assess the impact of diagnoses and therapies on rare disease patients. Orphanet J Rare Dis. 2022;17(1):181. doi: 10.1186/s13023-022-02337-2
[15] Han Q, Fu H, Chu X, et al. Research advances in treatment methods and drug development for rare diseases. Front Pharmacol. 2022;13:971541. doi: 10.3389/fphar.2022.971541
[16] von der Lippe C, Diesen PS, Feragen KB. Living with a rare disorder: a systematic review of the qualitative literature. Mol Genet Genomic Med. 2017;5(6):758–773. doi: 10.1002/mgg3.315
[17] McMullan J, Crowe AL, Bailie C, et al. Improvements needed to support people living and working with a rare disease in Northern Ireland: current rare disease support perceived as inadequate. Orphanet J Rare Dis. 2020;15(1):315. doi: 10.1186/s13023-020-01559-6
[18] Salinas GD. Trends in physician preferences for and use of sources of medical information in response to questions arising at the point of care: 2009–2013. J Contin Educ Health Prof. 2014;34 Suppl 1:S11–S16. doi: 10.1002/chp.21224
[19] Physicians are online—here's the what, why, and how. Elsevier Pharma & Life Sciences Solutions; 2024 [cited 2024 Jun 22]. Available from: https://www.elsmediakits.com/physicians-are-online-2024-thank-you?cookieredirect=1
[20] Amberger JS, Bocchini CA, Scott AF, et al. OMIM.org: leveraging knowledge across phenotype–gene relationships. Nucleic Acids Res. 2019;47(D1):D1038–D1043. doi: 10.1093/nar/gky1151
[21] Rath A, Olry A, Dhombres F, et al. Representation of rare diseases in health information systems: the orphanet approach to serve a wide range of end users. Hum Mutat. 2012;33(5):803–808. doi: 10.1002/humu.22078
[22] Vandeborne L, van Overbeeke E, Dooms M, et al. Information needs of physicians regarding the diagnosis of rare diseases: a questionnaire-based study in Belgium. Orphanet J Rare Dis. 2019;14(1):99. doi: 10.1186/s13023-019-1075-8
[23] Bessler JB, Fazzone W, Sladicka N, et al. Understanding US healthcare provider preferences for consumption of publication content: opportunities to leverage omnichannel approaches. Curr Med Res Opin. 2023;39(9):1271–1277. doi: 10.1080/03007995.2023.2240116
[24] Hesse BW, Nelson DE, Kreps GL, et al. Trust and sources of health information: the impact of the Internet and its implications for health care providers: findings from the first Health Information National Trends Survey. Arch Intern Med. 2005;165(22):2618–2624.
[25] Litzkendorf S, Babac A, Rosenfeldt D, et al. Information needs of people with rare diseases—what information do patients and their relatives require? J Rare Sis Diag Ther. 2016;2(2):40.
[26] Miller EG, Woodward AL, Flinchum G, et al. Opportunities and pitfalls of social media research in rare genetic diseases: a systematic review. Genet Med. 2021;23(12):2250–2259. doi: 10.1038/s41436-021-01273-z
[27] Davies W. Insights into rare diseases from social media surveys. Orphanet J Rare Dis. 2016;11(1):151. doi: 10.1186/s13023-016-0532-x
[28] Smailhodzic E, Hooijsma W, Boonstra A, et al. Social media use in healthcare: a systematic review of effects on patients and on their relationship with healthcare professionals. BMC Health Serv Res. 2016;16(1):442. doi: 10.1186/s12913-016-1691-0

[29] Chen J, Wang Y. Social media use for health purposes: systematic review. J Med Internet Res. 2021;23(5):e17917. doi: 10.2196/17917
[30] Kanchan S, Gaidhane A. Social media role and its impact on public health: a narrative review. Cureus. 2023;15(1):e33737. doi: 10.7759/cureus.33737
[31] Patrick M, Venkatesh RD, Stukus DR. Social media and its impact on health care. Ann Allergy Asthma Immunol. 2022;128(2):139–145. doi: 10.1016/j.anai.2021.09.014
[32] Chou W-YS, Gaysynsky A, Cappella JN. Where we go from here: health misinformation on social media. Am J Public Health. 2020;110(S3):S273–S275. doi: 10.2105/AJPH.2020.305905
[33] Vraga EK, Bode L. Defining misinformation and understanding its bounded nature: using expertise and evidence for describing misinformation. Polit Commun. 2020;37(1):136–144. doi: 10.1080/10584609.2020.1716500
[34] Pauer F, Litzkendorf S, Göbel J, et al. Rare diseases on the internet: an assessment of the quality of online information. J Med Internet Res. 2017;19(1):e23. doi: 10.2196/jmir.7056
[35] Wang ML, Britton OJ, Beard J. The call for science communication and public scholarship. Transl Behav Med. 2023;13(3):156–159. doi: 10.1093/tbm/ibac096
[**36] Walewski J, Donovan D, Nori M. How many zebras are there, and where are they hiding in medical literature? A literature review of publications on rare diseases. Expert Opin Orphan Drugs. 2019;7(11):513–519.The original reference establishing the paucity of publications on rare diseases in medical literature.
[37] Herzog C, Hook D, Konkiel S. Dimensions: bringing down barriers between scientometricians and data. Quant Sci Stud. 2020;1(1):387–395. doi: 10.1162/qss_a_00020
[38] Hook DW, Porter SJ, Herzog C. Dimensions: building context for search and evaluation [technology report]. Front Res Metr Anal. 2018;3. doi: 10.3389/frma.2018.00023
[39] Mouratidis RW. Dimensions. J Med Libr Assoc. 2019;107(3):459–461.
[40] Stauss M, Floyd L, Woywodt A. Weighing up Open Access Publishing in Nephrology—bronze, platinum, or fools' gold? Kidney360. 2023;4(11):1637–1640. doi: 10.34067/KID.0000000000000282
[41] Piwowar H, Priem J, Larivière V, et al. The state of OA: a large-scale analysis of the prevalence and impact of Open Access articles. PeerJ. 2018;6:e4375. doi: 10.7717/peerj.4375
[42] Pearce JM. The rise of platinum open access journals with both impact factors and zero article processing charges. Knowledge. 2022;2(2):209–224. doi: 10.3390/knowledge2020013
[43] Taylor M. Slow, slow, quick, quick, slow: five altmetric sources observed over a decade show evolving trends, by research age, attention source maturity and open access status. Scientometrics. 2023;128(4):2175–2200. doi: 10.1007/s11192-023-04653-1
[44] St. Aubin C, Liedke J. Social media and news fact sheet. Washington (DC): Pew Research Center; 2024 [cited 2025 Feb 6]. Available from: https://www.pewresearch.org/journalism/fact-sheet/social-media-and-news-fact-sheet/
[45] Ong KL, Stafford LK, McLaughlin SA, et al. Global, regional, and national burden of diabetes from 1990 to 2021, with projections of prevalence to 2050: a systematic analysis for the Global Burden of Disease Study 2021. Lancet. 2023;402(10397):203–234. doi: 10.1016/S0140-6736(23)01301-6
[46] Haug JD. Physicians' preferences for information sources: a meta-analytic study. Bull Med Libr Assoc. 1997;85(3):223–232.
[47] Kosteniuk JG, Morgan DG, D'Arcy CK. Use and perceptions of information among family physicians: sources considered accessible, relevant, and reliable. J Med Libr Assoc. 2013;101(1):32–37. doi: 10.3163/1536-5050.101.1.006
[48] Weller FS, Hamming JF, Repping S, et al. What information sources do Dutch medical specialists use in medical decision-making: a qualitative interview study. BMJ Open. 2023;13(10):e073905. doi: 10.1136/bmjopen-2023-073905
[**49] Crowe AL, McKnight AJ, McAneney H. Communication needs for individuals with rare diseases within and around the healthcare system of Northern Ireland. Front Public Health. 2019;7:236. doi: 10.3389/fpubh.2019.00236Important reference reporting on the unmet need of communicating information to patients with rare diseases and their caregivers.
[50] McMullan J, Lohfeld L, McKnight AJ. Needs of informal caregivers of people with a rare disease: a rapid review of the literature. BMJ Open. 2022;12(12):e063263. doi: 10.1136/bmjopen-2022-063263
[51] Guide to Patient Involvement in Rare Disease Therapy Development. A publication of the rare disease PFDD Compendium Workshop Series; 2022. Available from: https://everylifefoundation.org/wp-content/uploads/2022/01/Guide-to-Patient-Involvement-FINAL-COMPLETE-GUIDE-Rev.pdf
[*52] Rosenberg A, Walker J, Griffiths S, et al. Plain language summaries: enabling increased diversity, equity, inclusion and accessibility in scholarly publishing. Learn Publ. 2023;36(1):109–118. doi: 10.1002/leap.1524Reference highlighting the role of plain language summaries in communicating medical information in a manner easily understood even by those uninitiated in medical communications.
[53] Kim D, Jung W, Jiang T, et al. An exploratory study of medical journal's Twitter use: metadata, networks, and content analyses. J Med Internet Res. 2023;25:e43521. doi: 10.2196/43521
[54] Search engine optimization (SEO) for your article. Wiley Author Services; 2024 [cited 2024 Jun 18]. Available from: https://authorservices.wiley.com/author-resources/Journal-Authors/Prepare/writing-for-seo.html