Online first

Data Mining of Systematic Reviews 1934-2023: A Bibliometric Analysis

Haneen Al-Abdallat[1a], Badi Rawashdeh, M.D.[2]

[1] School of Medicine, The University of Jordan, [2] Division of Transplant Surgery, Medical College of Wisconsin

Keywords: Bibliometric, Systematic reviews, PubMed, Analysis

https://doi.org/10.59707/hymrHUHP8885

High Yield Medical Reviews

Introduction

Systematic reviews consolidate evidence and drive clinical practice guidelines, cost-effective analyses, and policy decisions; therefore, their annual publication rate has increased significantly. We used bibliometric analysis to identify research trends, the most searched topics, authors and organizations productivity and collaboration, the research network, and research gaps by examining keywords frequency and systematic reviews distribution.

Methods

We searched the PubMed database for systematic reviews using the systematic review filter described by Salvador-Oliván and coauthors, which has higher recall than the PubMed SR filter. The search period was from 1934 until February 3, 2023. Microsoft Excel and the VOSviewer application were used for analyzing yearly trends, institutions, authors, and keywords, as well as to create tables and network figures.

Results

A total of 378,685 articles were published. The number of articles published has been rising steadily during the past five years. The University of Toronto and McMaster University in Canada (n = 1415 and n = 1386) were the leading contributory universities. “Genetic predisposition to disease”, “postoperative complications”, “neoplasm”, “stroke”, and “covid-19” were the top 5 occurring keywords that are particular to a specialty in systematic reviews.

Conclusion

This bibliometric research examined systematic reviews, publication trends, the majority of publishing disciplines, authors and organizations productivity, and collaborative efforts. The results of this study could prove to be an invaluable resource for researchers, policymakers, and healthcare professionals.

INTRODUCTION

In recent years, there has been a substantial increase in the number of annual publications of systematic reviews, which reflects the critical role that systematic reviews play in synthesizing evidence, informing clinical practice guidelines, cost-effective analysis, and policy decisions.[1] Additionally, some organizations require researchers to provide systematic reviews in grant applications to support the case for proposed new research.[1]

Systematic reviews, as opposed to traditional narrative reviews, focus on a clearly defined research question, employ an explicit search strategy to find every relevant proof, assess studies using methodological standards that have been predetermined, and formally integrate the evidence-based findings.[2,3]

Therefore, a systematic review that has been properly performed is the best evidence to direct clinical practice, a foundation for the recommendations of evidence-based practice guidelines, and it should play a significant role in the planning of future research.[4,5] Consequently, we performed a bibliometric analysis to determine research

a **Corresponding author:**
Haneen Al-Abdallat, medical student from the University of Jordan. haneenabdalat@gmail.com.

trends, the number of systematic reviews published over the years, the most frequently searched topics or subject areas for searches, the productivity and collaboration of authors and organizations, the mapping of the research network, and the identification of research gaps by examining the frequency of keywords and the distribution of published systematic reviews.

This study has the potential to shed light on the inner workings of systematic review publication, thereby influencing the direction of the discipline in the future.

METHODS

DATA COLLECTION AND RETRIEVAL METHODS

We searched the PubMed database for systematic reviews using the systematic review filter presented by Salvador-Oliván and coauthors,[6] which provides higher recall than the PubMed SR filter, better at retrieving potential systematic reviews with a likely high degree of precision. The search query ranged from 1934 to February 3, 2023. The data extraction was done by the 1st author and reviewed by the 2nd author.

((systematic* [ti] AND review [ti]) OR Systematic overview* [ti] OR Cochrane review* [ti] OR systemic review* [ti] OR scoping review [ti] OR scoping literature review [ti] OR mapping review [ti] OR Umbrella review* [ti] OR (review of reviews [ti] OR overview of reviews [ti]) OR meta-review [ti] OR (integrative review [ti] OR integrated review [ti] OR integrative overview [ti] OR meta-synthesis [ti] OR meta-synthesis [ti] OR quantitative review [ti] OR quantitative synthesis [ti] OR research synthesis [ti] OR meta-ethnography [ti]) OR Systematic literature search [ti] OR Systematic literature research [ti] OR meta-analyses [ti] OR meta-analyses [ti] OR metaanalysis [ti] OR meta-analysis [ti] OR meta-analytic review [ti] OR meta-analytical review [ti] OR meta-analysis [pt] OR ((search* [tiab] OR medline [tiab] OR pubmed [tiab] OR embase [tiab] OR Cochrane [tiab] OR scopus [tiab] or web of science [tiab] OR sources of information [tiab] OR data sources [tiab] OR following databases [tiab]) AND (study selection [tiab] OR selection criteria [tiab] OR eligibility criteria [tiab] OR inclusion criteria [tiab] OR exclusion criteria [tiab])) OR systematic review [pt]) NOT (letter [pt] OR editorial [pt] OR comment [pt] OR case reports [pt] OR historical article [pt] OR report [ti] OR protocol [ti] OR protocols [ti] OR withdrawn [ti] OR retraction of publication [pt] OR retraction of publication as topic [mesh] OR retracted publication [pt] OR reply [ti] OR published erratum [pt]).

DATA ANALYSIS

Using VOSviewer version 1.6.19 and Microsoft Excel, we analyzed the annual trends, institutions, authors, keywords and generated tables and network figures.

Using PubMed, we downloaded the timeline publication (in years), and we used Excel to analyze the last 5-year trend.

In institutions analysis, we choose a minimum of one document threshold to include all organizations in the analysis. Then we choose 1000 organizations to be selected, with the highest total link strength, then using Excel we combine the departments of each organization.

In the author's analysis, we included the authors who have a minimum of 200 publications, presented them using VOSviewer 1.6.19, and used Excel duplicated was combined. The only duplicate found was "Murad, Mohammad Hassan" and "Murad, M Hassan".

The most frequently occurring keywords were studied by limiting them to a minimum of 1000 occurrences. In the keyword analysis, we manually removed words that implied the study design and demographic data, such as "systematic review," "meta-analysis," and "case-control studies," and redundant words such as "human," "male," "female," and "adult."

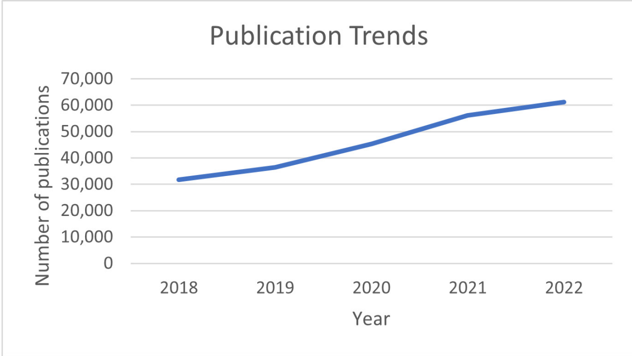

Figure 1. publication trends from (2018-2022)

RESULTS

INCLUDED STUDIES

All 378,685 articles resulting from the search filter were included and uploaded to VOSviewer 1.6.19.

ANNUAL TRENDS

We studied the trend in the last five years, from 2018-2022; the number of systematic reviews that were published in 2018, 2019, 2020, 2021, and 2022 were 31,751 (13.8%), 36,378 (15.8%), 45,252(19.6%), 56,168(24.30%), 61,237 (26.5%), respectively (Figure 1).

INSTITUTIONS

In terms of the total number of publications, of 857,388 organizations, the top contributing institutions were the University of Toronto and McMaster University from Canada (n = 1415, n = 1386), Tehran University of medical sciences (n = 1268) from Iran, University of Oxford from UK (n = 715), National University of Singapore (n = 612) from Singapore, and Harvard medical school (n = 607) from USA (Table 1).

With respect to the link strength and connections to other institutions, the University of Oxford, Harvard Medical School, National University of Singapore had the highest link strength attribution (n = 7177 and n = 6246, n = 5400), respectively.

Table 1. Top-contributing institutions according to the number of publications

Institutions-according to number of publications	Country	publications	Total link of strength
University of Toronto	Canada	1415	1631
McMaster university	Canada	1386	2285
Tehran university of medical sciences	Iran	1268	1826
University of oxford	UK	715	7177
National university of Singapore	Singapore	612	5400
Harvard medical school	USA	607	6246
Mashhad university of medical sciences	Iran	485	1860
Karolinska institutet	Sweden	453	4638
University of Bristol	UK	432	3088
Taipei medical university	Taiwan	426	418
California institute of behavioral neurosciences & psychology	USA	384	290
Iran University of medical sciences	Iran	369	337
University of Cambridge	UK	287	3886
University of Bern	Switzerland	278	473
University of York	UK	272	101
University College London	UK	259	1950
Isfahan university of medical sciences	Iran	239	432
University medical center Utrecht	The Netherlands	237	553
Tabriz university of medical sciences	Iran	236	700
Aarhus university	Denmark	192	1035

AUTHORS

A total of 963,261 authors were included in our analysis. Among them, 48 meet the threshold of 200 minimum number of documents per author; of those, some were not connected to each other, 44 were most connected to each other; the visualization of each author's contribution and their interconnections is demonstrated in 4 clusters, (Figure 2)

7 authors have had at least 300 publications of systematic reviews. The top 5 were Wang W., Zhang W., Murad MH, Wang Y., and Zhang L. A complete list of the top-publishing authors is presented in Table (2)

Regarding the total link strength between authors, Pasty BM., Rotter JI. And Hofman A. had the highest total link strength with authors with total link strength equal to n =341, n = 326, n = 310, respectively.

KEYWORDS

After excluding the nonspecific key words such as "humans," "female," "male." The top 10 most occurring keywords that are specific to a specialty in systematic reviews are shown in Table (3). All of them occurred over 5000 times: "genetic predisposition to disease" (n = 9747), "postoperative complications" (n = 9215), "neoplasm " (n = 8791), "stroke" (n = 7792), "covid-19" (n = 7651), "polymorphism, single nucleotide" (n = 7605), "biomarkers" (n = 7021), "depression" (n = 6921), "cardiovascular diseases" (n = 6141) and "diabetes mellitus, type 2" (n = 5973).

Figure (3) shows the top 50 specific keyword that represents a field of interest, grouped into 5 clusters, cluster number one contains 21 medical-related items: "anti-bacterial agents," "asthma," "atrial fibrillation," "biomarkers," "blood glucose," "blood pressure," "cardiovascular disease," "chronic disease," "coronary artery disease," "diabetes mellitus," "diabetes mellitus, type 2," "diet," "dietary supplements," "dose-response relations," "drug therapy, combinations," "exercise therapy," "heart failure," "HIV infections," "hypertension," "myocardial infarction" and "stroke." The second cluster includes 13 items that are related to neoplasms: "antineoplastic agents," "antineoplastic combined chemotherapy protocols," "biomarkers, tumor," "breast neoplasms," "cancer," "colorectal neoplasms," "laparoscopy," "lung neoplasms," "neoplasms recurrence, local," "neoplasms staging," "neoplasms," "postoperative complications," "prostatic neoplasms." Cluster 3 is psychiatry-related: "anxiety," "brain," "cognition," "depression," "mental disorders," "mental health," and "schizophrenia." Cluster 4 is genetic related: "genetic association studies," "genetic predisposition to disease," "genome-wide association study," "genotype," "polymorphism, genetic," "polymorphism, single nucleotide." The last cluster, which is COVID, pandemic related: "COVID-19, "pandemics," and "SARS-COV-2."

DISCUSSION

Although locating systematic reviews quickly and accurately can be challenging for clinicians and researchers due

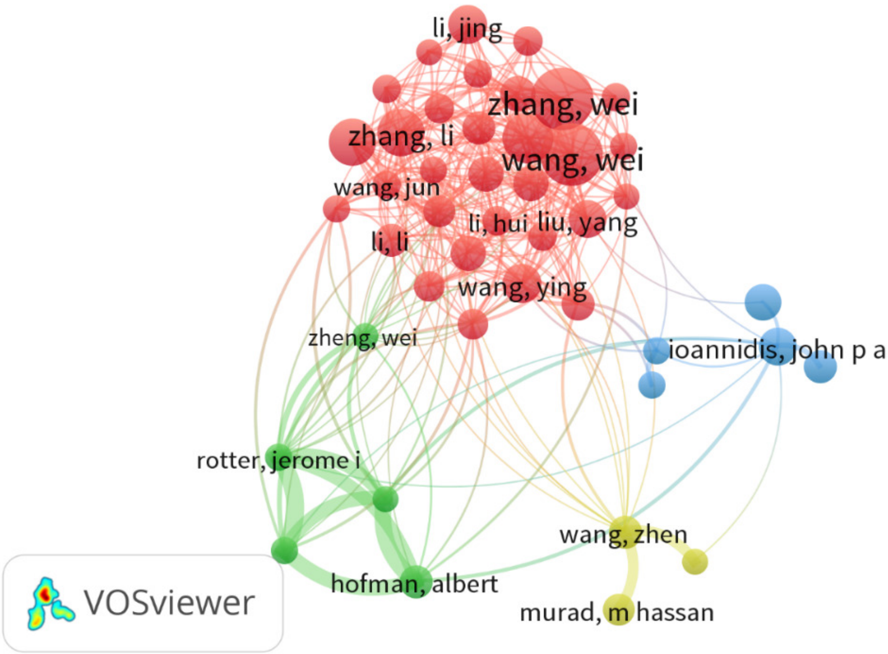

Figure 2. visualization of the most publishing authors and their connections

Table 2. A list of the top-publishing authors

author	documents	total link strength	author	documents	total link strength
Wang W.	484	78	Chen W.	239	42
Zhang W.	484	62	Lee YH.	239	0
Murad MH	451	99	Li H.	232	36
wang Y.	393	43	Liu W.	232	27
Zhang L.	367	58	Craig JC.	229	0
Wang J.	366	52	Wang H.	226	20
Li J.	304	28	Wang J.	225	19
Wang Y.	299	35	Maffulli N.	222	0
Ioannidis JPA.	294	35	Wang B.	221	18
Liu Y.	290	42	Wang L.	221	28
Stubbs B.	288	6	Wang X.	221	30
Li w.	279	40	Wang J.	217	45
Li y.	278	34	Li J.	215	35
Zhang l.	273	38	Pasty BM.	213	341
Zhang j.	271	34	Liu J.	212	39
Sahebkar A.	266	0	Takagi H.	212	0
Li L.	260	24	Bhandari M.	209	15
Wang L.	259	42	Liu J.	209	38
Zhang Y.	259	25	Rotter JI.	208	326
Cuijpers P.	257	12	Thabane L.	208	21
Hofman A.	254	310	Zhang Y.	206	25
Wang Z,	253	112	Zheng W.	206	53
Zhang Y.	248	45	Yang K.	205	26
Zhang Y.	240	28	van Duijn CM.	201	304

Table 3. The top 10 most occurring keywords that are specific to a specialty.

Keyword	Occurrences
Genetic predisposition to disease	9747
Postoperative complications	9215
Neoplasms	8791
Stroke	7792
Covid-19	7651
Polymorphism, single nucleotide	7605
Biomarkers	7021
Depression	6921
Cardiovascular Diseases	6141
Diabetes Mellitus, type 2	5973

to lack of time and expertise in working with search strategies and database field.[7,8] Systematic review publishing began in the late 1980s and has shown remarkable interest and growth ever since.[9] The number of systematic reviews published over the last five years has increased from 31,751 to 61,237 in 2018 and 2022, with an average annual publication rate of 46,157 systematic reviews per year or roughly 128 systematic reviews per day.

There has been a recent uptick in the publication of systematic reviews, which can be traced back to the realization in the scientific and medical communities that there is so much research available that it needs to be integrated and that literature reviews written by experts are more trustworthy than relying on the results of individual studies.[10,11] Moreover, some nations have developed a research culture that places a heavy emphasis on the production of systematic reviews,[12] and there is a trend among some funding agencies to require applicants to support their research funding applications with a systematic review.[13]

Despite the wide range of numerous systematic reviews, our research found that the majority of systematic reviews are concentrated in four primary fields: genetics, cancer, mental illnesses, and cardiovascular diseases.

We found that the genetic field is the one that publishes systematic reviews the most, with the keyword “genetic predisposition to disease” ranking first with 9,747 occurrences. This can be correlated with the fact that systematic reviews have become a popular method for summarizing gene-disease connections.[14] The growing interest in the genetic basis of common diseases has been observed during the past ten years, along with quick developments in high-throughput genotyping technologies.[15] As the most common tumors to cause death worldwide, lung and colorectal cancers are the main topics of study in the field of cancer research.[16]

Within mental disorders, schizophrenia and depression are the most extensively reviewed topics. This finding was consistent with previous research in 2009, that found neurological and mental disorders to rank the largest number of systematic reviews,[17] with depression and mental disorders being a major contributor to the disease burden in high-income countries[18]; such findings are expected.

Our research also uncovered important authors and institutions that are well-connected and influential in their respective fields. The University of Toronto and McMaster University in Canada are the top contributing institutions. The University of Oxford and Harvard Medical School have the strongest links and connections to other institutions,

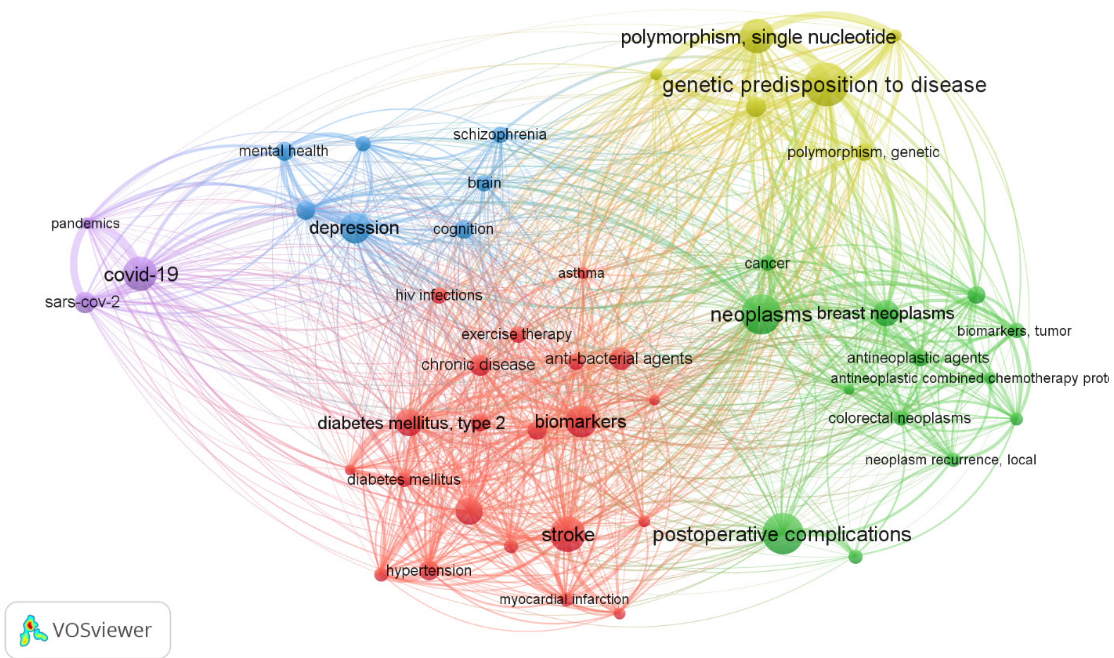

Figure 3. the top 50 specific keyword that represents a field of interest, grouped in 5 clusters, each color represents a cluster of related items.

which could be attributed to their strong research financial support.[19]

It's important to note that Pasty BM. is the author with the most co-authorship relationships and impact among other researchers, while Wang W. is determined to have the most systematic review publications.

Our bibliometric study, like other bibliometric research, had limitations. First, we searched only using the PubMed database; we did not search in the Web of Science Core Collection or Google Scholar. However, we utilized a search filter with a high degree of sensitivity in order to incorporate a maximum number of systematic reviews. Second, we did not include an analysis of citations or publications in our work. Despite this, our findings offer important insights into the present state of systematic reviews across a variety of fields and have the potential to help inform future research directions and collaborations.

CONCLUSION

Systematic reviews play a critical role in synthesizing evidence, informing clinical practice guidelines, cost-effective analysis, and policy decisions. Thus, this bibliometric analysis was conducted to add to the current body of knowledge regarding the features of systematic reviews, publication trends, the majority of publishing specialties, the productivity of authors and organizations, and the nature of their collaborative efforts. Within the realm of systematic reviews, the findings of this research have the potential to become an extremely helpful resource for academics, decision-makers, and healthcare practitioners. In addition, we encourage conducting systematic reviews in areas where there is limited or outdated evidence.

CONFLICT OF INTEREST

We declare that we have no competing interest to disclose. This research was not funded by any organization, and we have no conflict of interest to report.

Submitted: April 09, 2023 AST, Accepted: May 06, 2023 AST

REFERENCES

1. Shojania KG, Sampson M, Ansari MT, et al. Updating Systematic Reviews. *Health (San Francisco)*. 2010;(16). Accessed March 4, 2023. http://europepmc.org/books/NBK44099

2. Bero LA, Grilli R, Grimshaw JM, Harvey E, Oxman AD, Thomson MA. Closing the gap between research and practice: an overview of systematic reviews of interventions to promote the implementation of research findings. *BMJ*. 1998;317(7156):465-468. doi:10.1136/bmj.317.7156.465

3. Oxman AD, Cook DJ, Guyatt GH, et al. Users' Guides to the Medical Literature: VI. How to Use an Overview. *JAMA*. 1994;272(17):1367-1371. doi:10.1001/jama.1994.03520170077040

4. Scholten RJPM, Kremer LCM. [Systemic reviews as a basis for guidelines]. *Ned Tijdschr Geneeskd*. 2004;148(6):262-264. Accessed March 4, 2023. https://europepmc.org/article/med/15004951

5. Haines A, Jones R. Implementing findings of research. *BMJ*. 1994;308(6942):1488-1492. doi:10.1136/bmj.308.6942.1488

6. Salvador-Oliván JA, Marco-Cuenca G, Arquero-Avilés R. Development of an efficient search filter to retrieve systematic reviews from pubmed. *J Med Libr Assoc*. 2021;109(4):561-574. doi:10.5195/jmla.2021.1223

7. Boluyt N, Tjosvold L, Lefebvre C, Klassen TP, Offringa M. Usefulness of Systematic Review Search Strategies in Finding Child Health Systematic Reviews in MEDLINE. *Arch Pediatr Adolesc Med*. 2008;162(2):111-116. doi:10.1001/archpediatrics.2007.40

8. Ely JW, Osheroff JA, Ebell MH, et al. Obstacles to answering doctors' questions about patient care with evidence: qualitative study. *BMJ*. 2002;324(7339):710. doi:10.1136/bmj.324.7339.710

9. Ioannidis JPA. The Mass Production of Redundant, Misleading, and Conflicted Systematic Reviews and Meta-analyses. *Milbank Q*. 2016;94(3):485-514. doi:10.1111/1468-0009.12210

10. Moher D, Tetzlaff J, Tricco AC, Sampson M, Altman DG. Epidemiology and Reporting Characteristics of Systematic Reviews. *PLoS Med*. 2007;4(3):e78. doi:10.1371/journal.pmed.0040078

11. Page MJ, Shamseer L, Altman DG, et al. Epidemiology and Reporting Characteristics of Systematic Reviews of Biomedical Research: A Cross-Sectional Study. *PLoS Med*. 2016;13(5):e1002028. doi:10.1371/journal.pmed.1002028

12. Ioannidis JPA, Chang CQ, Lam TK, Schully SD, Khoury MJ. The Geometric Increase in Meta-Analyses from China in the Genomic Era. *PLoS One*. 2013;8(6):e65602. doi:10.1371/journal.pone.0065602

13. Moher D, Glasziou P, Chalmers I, et al. Increasing value and reducing waste in biomedical research: who's listening? *Lancet*. 2016;387(10027):1573-1586. doi:10.1016/s0140-6736(15)00307-4

14. Ioannidis JPA, Gwinn M, Little J, et al. A road map for efficient and reliable human genome epidemiology. *Nat Genet*. 2006;38(1):3-5. doi:10.1038/ng0106-3

15. Khoury MJ, Little J, Gwinn M, Ioannidis JP. On the synthesis and interpretation of consistent but weak gene-disease associations in the era of genome-wide association studies. *Int J Epidemiol*. 2007;36(2):439-445. doi:10.1093/ije/dyl253

16. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin*. 2021;71(3):209-249. doi:10.3322/caac.21660

17. Hoffmann T, Erueti C, Thorning S, Glasziou P. The scatter of research: cross sectional comparison of randomised trials and systematic reviews across specialties. *BMJ*. 2012;344(7861):e3223. doi:10.1136/bmj.e3223

18. Whiteford HA, Degenhardt L, Rehm J, et al. Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. *Lancet*. 2013;382(9904):1575-1586. doi:10.1016/s0140-6736(13)61611-6

19. Oliver S, Bangpan M, Stansfield C, Stewart R. Capacity for conducting systematic reviews in low- and middle-income countries: A rapid appraisal. *Heal Res Policy Syst*. 2015;13(1):1-8. doi:10.1186/S12961-015-0012-0/FIGURES/2