

Social, Behavioral, and Cultural factors of HIV in Malawi: Semi-Automated Systematic Review

THIABAUD, Amaury, *et al.*

Abstract

Background: Demographic and sociobehavioral factors are strong drivers of HIV infection rates in sub-Saharan Africa. These factors are often studied in qualitative research but ignored in quantitative analyses. However, they provide in-depth insight into the local behavior and may help to improve HIV prevention. Objective: To obtain a comprehensive overview of the sociobehavioral factors influencing HIV prevalence and incidence in Malawi, we systematically reviewed the literature using a newly programmed tool for automatizing part of the systematic review process. Methods: Due to the choice of broad search terms ("HIV AND Malawi"), our preliminary search revealed many thousands of articles. We, therefore, developed a Python tool to automatically extract, process, and categorize open-access articles published from January 1, 1987 to October 1, 2019 in the PubMed, PubMed Central, JSTOR, Paperity, and arXiv databases. We then used a topic modelling algorithm to classify and identify publications of interest. Results: Our tool extracted 22,709 unique articles; 16,942 could be further processed. After topic modelling, 519 of [...]

Reference

THIABAUD, Amaury, *et al.* Social, Behavioral, and Cultural factors of HIV in Malawi: Semi-Automated Systematic Review. *Journal of Medical Internet Research*, 2020, vol. 22, no. 8, p. e18747

DOI : 10.2196/18747

PMID : 32795992

Available at:

<http://archive-ouverte.unige.ch/unige:149117>

Disclaimer: layout of this document may differ from the published version.



UNIVERSITÉ
DE GENÈVE

Review

Social, Behavioral, and Cultural factors of HIV in Malawi: Semi-Automated Systematic Review

Amaury Thiabaud^{1*}, PhD; Isotta Triulzi^{1,2*}, MSc; Erol Orel¹, DEA; Kali Tal^{1,3}, PhD; Olivia Keiser¹, PhD

¹Institut de Santé Globale, Université de Genève, Genève, Switzerland

²Institute of Management, Scuola Superiore Sant'Anna, Pisa, Italy

³Institute of Primary Health Care (BIHAM), University of Bern, Bern, Switzerland

*these authors contributed equally

Corresponding Author:

Amaury Thiabaud, PhD

Institut de Santé Globale

Université de Genève

Chemin des mines 9

Genève, 1202

Switzerland

Phone: 41 22 379 81 30

Email: amaury.thiabaud@unige.ch

Abstract

Background: Demographic and sociobehavioral factors are strong drivers of HIV infection rates in sub-Saharan Africa. These factors are often studied in qualitative research but ignored in quantitative analyses. However, they provide in-depth insight into the local behavior and may help to improve HIV prevention.

Objective: To obtain a comprehensive overview of the sociobehavioral factors influencing HIV prevalence and incidence in Malawi, we systematically reviewed the literature using a newly programmed tool for automatizing part of the systematic review process.

Methods: Due to the choice of broad search terms (“HIV AND Malawi”), our preliminary search revealed many thousands of articles. We, therefore, developed a Python tool to automatically extract, process, and categorize open-access articles published from January 1, 1987 to October 1, 2019 in the PubMed, PubMed Central, JSTOR, Paperity, and arXiv databases. We then used a topic modelling algorithm to classify and identify publications of interest.

Results: Our tool extracted 22,709 unique articles; 16,942 could be further processed. After topic modelling, 519 of these were clustered into relevant topics, of which 20 were kept after manual screening. We retrieved 7 more publications after examining the references so that 27 publications were finally included in the review. Reducing the 16,942 articles to 519 potentially relevant articles using the software took 5 days. Several factors contributing to the risk of HIV infection were identified, including religion, gender and relationship dynamics, beliefs, and sociobehavioral attitudes.

Conclusions: Our software does not replace traditional systematic reviews, but it returns useful results to broad queries of open-access literature in under a week, without a priori knowledge. This produces a “seed dataset” of relevance that could be further developed. It identified known factors and factors that may be specific to Malawi. In the future, we aim to expand the tool by adding more social science databases and applying it to other sub-Saharan African countries.

(*J Med Internet Res* 2020;22(8):e18747) doi: [10.2196/18747](https://doi.org/10.2196/18747)

KEYWORDS

HIV/AIDS; topic modelling; text mining; Malawi; risk factors; machine learning

Introduction

Demographic and sociobehavioral factors are strong drivers of HIV in sub-Saharan Africa, but the interactions between these factors — the way their influence shifts over time and space

and influences HIV prevalence and incidence — are poorly understood. Some epidemiological studies have reported on the spatial variability of the HIV epidemic, using statistical analyses to assess the association between the spatial distribution of HIV prevalence and potential risk factors [1-7]. They found, for

example, that high population density [6] or a short distance to a road or a clinic [7] were associated with a high HIV prevalence. Tomita et al [8] showed that behavior (sexual debut, uptake of contraception, and circumcision) and social determinants strongly influenced the risk of HIV acquisition. An analysis of 29 sub-Saharan African countries found associations between 12 demographic and sociobehavioral factors, including variables related to age, literacy, HIV knowledge, domestic violence, women's empowerment, and sexual activity [9]. The patterns of associations were complex and varied by sex and country, but the study did not include many potentially significant factors because they were absent from the data or were only available for some countries and the study did not consider subnational variation. These epidemiological studies did not draw on qualitative research, and they rarely contextualized the associations they identified.

Social scientists of various disciplines have performed qualitative studies of social and cultural factors related to HIV, providing rich detail on the perceptions and behaviors of people in specific localities. For example, medical anthropologists have examined maternal care-seeking behavior in different geographic regions and groups [10]. Cultural studies have analyzed connections between belief in witchcraft and folk epidemiological wisdom about HIV [11] and the connection between women's educational level and attendance at antenatal care [12]. Sociologists have studied the perceptions of HIV testing in rural Malawi [13] or have analyzed power structures [14] and the vulnerability of adolescent girls [15].

Qualitative research gives us an in-depth understanding of local situations and may help identify factors that quantitative analyses have not considered. Qualitative analyses often focus on individual knowledge, opinions, attitudes, and challenges, while quantitative analyses quantify relationships between various factors and between factors and outcome variables. Combining qualitative and quantitative studies may reveal how and why various factors interact across time and space in a complex and widespread epidemic.

We used the topic of HIV in Malawi as a case study for an in-depth literature review of quantitative and qualitative literature on social and behavioral factors that may influence the HIV epidemic. We designed the review to be broad and inclusive to capture all possible relevant factors and expected it to identify known factors, neglected factors, and some factors that may never have been identified or analyzed in quantitative studies. We chose Malawi for several reasons: (1) Malawi has a relatively high HIV prevalence that varies substantially between regions [16], (2) the country is socioculturally diverse because it is home to many ethnic groups (eg, Chewa, Nyanja, Tumbuka, Yao, Lomwe, Sena, Tonga, Ngoni, Ngonde, Asians, and Europeans), and (3) a preliminary search revealed that Malawi was the focus of many scientific studies on HIV, including some by our group [17-19].

Because we knew that the breadth of our topic would return too many studies to read, we developed a semi-automated literature search engine and software that automatically downloads and analyzes open-access, full-text articles; this software can be used for searches on any broad topic across any region.

Methods

Search Strategy

We searched all English language articles published from the inception of the databases from January 1, 1987 to October 1, 2019 using the query "HIV AND Malawi". We performed an automated search of PubMed, PubMed Central (PMC), Paperity, and arXiv using a custom Python script and the corresponding application programming interfaces (APIs). We also sent a request for the same type of data to JSTOR since it offered no API to directly access the database.

Inclusion and Exclusion Criteria

We used a broad query designed to capture all articles about HIV and Malawi, not just those focused on health, and considered all studies that discussed social, behavioral, and cultural factors that might be associated with HIV infection in Malawi. The selection process is described in more detail in the following sections. We included original peer-reviewed articles, both quantitative and qualitative studies, and preprints. We also analyzed systematic reviews but preferred to include original publications when possible. We discarded articles that investigated the effects of HIV/AIDS (eg, [20]).

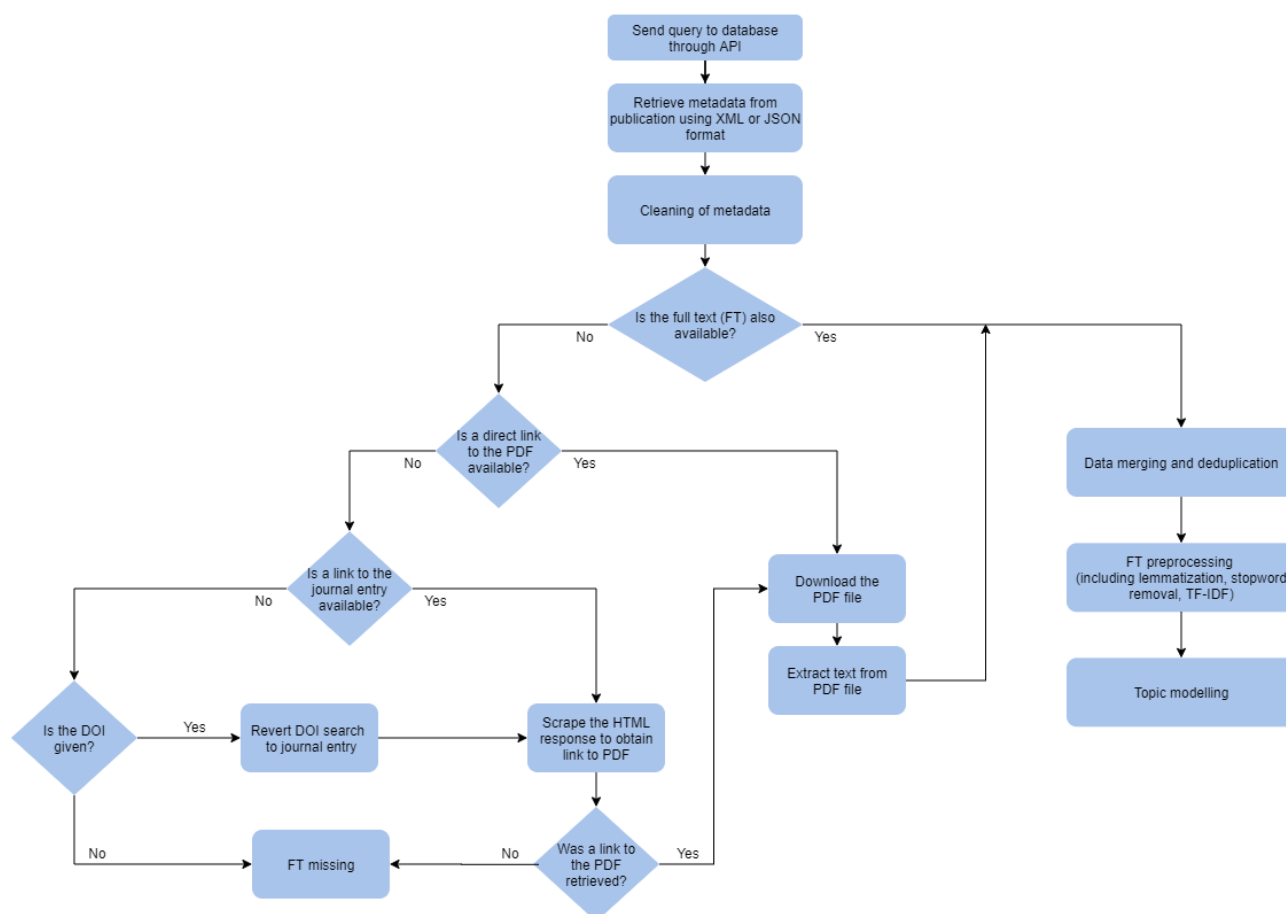
Data Collection

Figure 1 shows the flow of data collection and procession. We extracted the following information from articles: list of authors and their affiliations and MeSH keywords, Digital Object Identifier (DOI), title, abstract, full text, publication date, journal provider, and URL of the PDF version of the full-text article. Mandatory fields were title and availability of the full text; other items were retrieved if possible. If full-text articles could not be directly retrieved from the database, they were extracted from either an automatically downloaded PDF when the URL of the PDF was available or retrieved through an automatic reversed search with the DOI resolver [21]. The DOI resolver identified PDFs by scraping the internet. In some cases, the information in databases was incomplete (eg, missing DOI), or we could not access the PDF (eg, access restrictions, PDF contained only images, unavailable PDF) so we could not obtain the full text. We then checked the data for duplicates and merged them.

Each extracted full-text article was tokenized and processed further. We deleted non-ASCII characters, numbers, words under 4 characters (except some acronyms of interest like "HIV" and the abbreviation for antiretroviral therapy, "ART"), and a list of stopwords from the NLTK Python toolkit [22] with relevant additions (eg, URL, "volume," "journal"). We then lemmatized the text to avoid duplicating words with different inflectional endings.

Once cleaned, we ran term frequency-inverse document frequency from the scikit-learn Python package to extract relevant keywords [23]. Because the formatting of author affiliations was so heterogeneous, we retained only the city. We then stored original data and the generated data in a local SQLite database.

Figure 1. Flow for gathering and preprocessing data. API: application programming interface; DOI: Digital Object Identifier; TF-IDF: term frequency-inverse document frequency.



Topic Modelling

Finally, we used topic modelling to classify documents into topics based on their similarity. The process allowed us to broaden our search terms to identify relevant publications and to extract keywords relevant to a topic, providing an overview of salient and relevant terms that may best represent the data. We scored each document for a set number of topics, using the latent Dirichlet allocation (LDA) method [24]. Based on their highest topic score, publications were allocated to a topic.

To optimize computational efficiency and quickly identify potentially relevant articles, we initially chose 5 topics; for each, we used the same approach to identify 5 subtopics. We repeated the process 4 times and identified 625 topics. The resulting “tree of topics” increased its specificity at each repetition, which helped us identify our topics of interest faster than if we had started with many topics. We manually selected the number of iterations, but our selection was analysis-driven. Although more iterations would have reduced the number of potentially relevant articles that we needed to check manually, it would have raised

the risk of missing relevant articles, and the performance of the algorithm would have decreased.

The code used for this paper was frozen and is publicly available [25]. Note that this code is not up-to-date as we are improving the software, adding databases, and trying different methods of classification.

Results

The PRISMA diagram (Figure 2) summarizes the selection process for relevant articles. Based on our search term, the software extracted a total of 22,709 unique articles, of which 16,942 full-text articles (74.6%) were retrieved directly from the database or extracted from PDFs. Topic modelling automatically screened these articles, reducing our selection to a subset of 519 relevant articles that included 14 topics related to behaviors, beliefs, culture, and religion. Of these, we manually selected 119 based on titles and abstracts and after applying our inclusion and exclusion criteria. After full-text screening, 20 articles were included in the systematic review.

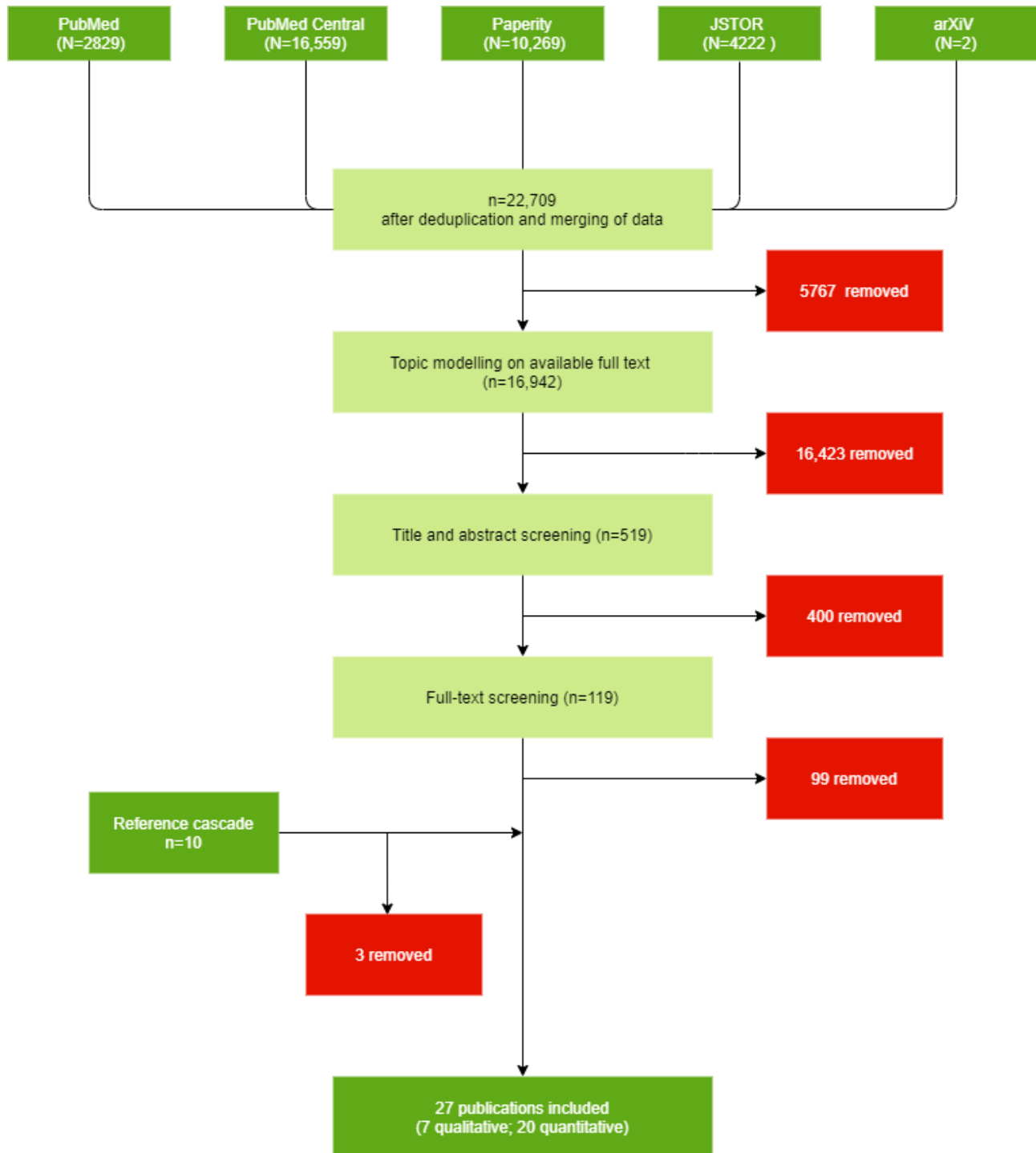
We identified 10 more articles from references and included 7 of them.

The systematic review finally included 27 publications: 20 (74.1%) were quantitative, and 7 (25.9%) were qualitative. It

took about 5 days to retrieve and preprocess the data for topic modelling. The iterative topic modelling took about a day.

In the following sections, we summarize the main findings of the included articles by topic area.

Figure 2. Study selection process.



Religion

We identified 3 publications that found religion was associated with the risk of HIV transmission or decreased risk behavior. All noted the effect of the religious leader on family planning, sexual morality, positive attitude towards condom use [26], and reducing stigma [27,28]. The studies also showed that

congregations in rural Malawi responded to the HIV epidemic by taking care of orphans, helping the sick, or sponsoring HIV-related knowledge programs [27,28]. In this way, religion shaped the HIV experience of congregation members. Yeatman and Trinitapoli [26] noticed an association between religious socialization and contraceptive use. The literature search also highlighted that Muslim men were more likely to be circumcised

and less likely to be HIV-infected than men of other faiths [24], though they were also less likely to use condoms. Being religious did not necessarily help prevent HIV, although Lau et al [29] noticed a positive association between weekly attendance of religious services and women's use of modern contraceptives. Faith was often associated with condom rejection. Anglewicz and Clark [30] found that Muslim men tended to find condom use less acceptable than did those of other faiths, while a change to Christianity increased condom use in men.

Partnerships and Gender

Reniers and Tfaily [31] investigated the role of polygyny in Rumphu, Mchinji, and Balaka. They reported more extramarital sex in polygynous marriages than in monogamous marriages. They further found evidence that in polygynous marriages, the latest wives were more likely HIV-positive than the first wives. Other studies also showed that having more than one sexual partner increased HIV prevalence, led to differences in adherence counselling and testing, or led to differences in male circumcision practices [29,32,33]. Stephenson [34] found a discrepancy between different types of relationships: Cohabiting men were less likely to have risky extramarital sex than married men.

Age asymmetry in relationships plays a role in HIV transmission [35]. One article examined the effect of an age difference between women and their partners on HIV transmission on Likoma Island. When male partners were 2-12 years older than their female partners, the women's risk of being HIV-positive was higher than for women whose partners were more than 12 years older. When female partners were more than 5 years older than their male partner, they were more likely to be HIV-positive than women whose partners were less than 5 years older. Never using condoms and being married were associated with larger age differences between men and women.

Condom use helps prevent HIV infection. Anglewicz and Clark [30] showed that marital status and women's and men's risk perceptions were associated with condom use. Getting married reduced the acceptability of using condoms during sexual intercourse. A woman's perception of her HIV status was generally more important than her real HIV status for the acceptability of condom use within marriage. Known HIV status was a more important determinant of condom use in men than in women [30].

Power structures and closeness in partnerships may strongly influence HIV transmission. Becker et al [32] showed that in women, prior HIV testing and emotional closeness to a partner were associated with acceptance of home-based HIV testing and counselling services. Schatz [36] showed that married women may be able to protect themselves from infection by communicating with their husband about HIV, by confronting their husband's sexual partners, or by refusing polygyny. Women may also seek support and advice from their friends and relatives or even ask for a divorce. This suggests that women are not always vulnerable; some can self-advocate for their protection in marriage to reduce their risk of acquiring HIV, especially in the matrilineal southern part of Malawi where a woman can tell her husband to "take your mat and go" [36].

Even in the patrilineal northern region, Schatz [36] highlighted the support of women by their kin if the husband's risky behaviors cause women to return to their family.

Beliefs

We found several articles that studied beliefs and found associations between misconceptions and ignorance about HIV transmission factors and risky behavior. Authors labelled these "false beliefs" and found they stigmatized people with HIV [26,37,38]. One study mentioned beliefs about women's cleansing rituals. Some congregations believe women will be cleansed by having unprotected sex (eg, after the death of their husbands, after giving birth, or after miscarriage) [39]. The belief that HIV-infected men can be cured by having sex with a virgin woman also spreads the disease.

Personal beliefs play an important role in HIV prevention and risk of HIV infection. Three studies highlighted an important role of perceived HIV status between partners. Anglewicz and Clark [40] examined the accuracy of perceived HIV status in 768 monogamous couples, finding that partners tended to overestimate the risk of being HIV-positive; overestimation was associated with marital infidelity. But knowing one's or one's spouse's actual HIV status significantly reduced HIV risk. Fedor et al [41] showed that once HIV-negative women and HIV-positive men learned their status, they reduced risky behavior by increasing condom use and having sex with fewer partners.

Social and Behavioral Characteristics

We identified 3 studies on the effect of behavior-change interventions on HIV [42-44]. In 2 studies [42,44], the authors found that the intervention seemed to affect HIV risk behaviors and knowledge; the third study found no effect [43]. Crittenden et al [42] studied the spread of behavioral and psychological factors with peer group interventions in adults living in rural areas of central Malawi. The behavioral changes that were promoted in the intervention group (eg, partner communication, use of condoms, recent HIV test) spread to other persons in the same community. The second study [44] assessed the effect of a cash transfer program (lottery ranging from US \$1 to US \$5) in adolescents and women aged 13-22 years who attended school. The primary outcome was HIV prevalence 18 months after study enrollment; it was 1.2% in the intervention group and 3.0% in the control group. Women who received the cash transfer were less sexually active than women in the control group. In contrast, a study in northern Malawi [43] showed that behavior change interventions did not reduce the risk of HIV infection in Malawian adolescents, possibly because these interventions send contradictory messages or because adolescents are more influenced by their living environment (culture, religion, peers).

Several studies investigated the association between migration and HIV. Hellinginger et al [45] reported an association between concurrent partnerships and HIV serodiscordance among couples on Likoma Island. HIV positivity was associated with migration out of the country (circular out-migration) and sexual contact with temporary in-migrants to the island. Anglewicz and Clark [46] concluded that migrants were more at risk for HIV

infection, but migration was not the reason for the higher risk; people with HIV were more likely to migrate, thereby reversing the causality.

Low socioeconomic status often drives HIV in Malawi [39,44]. It was associated with early sexual relationships, transactional sex, and a higher probability of having sex with older men [39]. In some studies, a low socioeconomic status was also associated with coercive heterosexual relationships. Coercive sex was a strong predictor for HIV infection in male victims [47]; the likelihood of being HIV positive was 7.2 times higher among men who had been sexually coerced than among those who had not been. One publication studied the association of coercive sexual behavior with social and economic status [48] and found that unemployment was strongly associated with coercive sex in young men in Blantyre, whereas material deprivation only was strongly associated with coercive sex in young women.

HIV infection was associated with intravaginal practices and products applied by women to manage their sexual relationship, manage menstruation, and improve their health. Women used cloth or paper to wipe out their vagina, inserted products to dry or tighten the vagina, and used intravaginal cleaning soap [39,49]. The use of the injectable hormonal birth control drug medroxyprogesterone acetate was associated with HIV seroconversion in HIV-negative women during a clinical trial [33]. Among men who have sex with men, the use of water-based lubricants could lower HIV risk [50]. Being older than 25 years, not being married, and age at first sexual intercourse were associated with HIV infection [50]. Some of these variables were also identified in other studies [29,32,47].

Although studies from other sub-Saharan African countries showed an association between alcohol and drug consumption with HIV testing, HIV infection, or uptake of preventive methods, evidence was limited in our study. Lau et al [29] found no association between tobacco use and male circumcision. Conroy and Chilungo [47] found some association between alcohol use and being HIV positive, but the association was not statistically significant (odds ratio 1.56). They did find a significant association between sexual coercion of women and alcohol use among men.

Discussion

Strengths

Using a broadly inclusive search phrase and repeated topic modelling, we quickly identified a small number of highly relevant articles about sociobehavioral factors and HIV in Malawi among 16,942 open access articles from 5 different databases. Our Python tool quickly reduced the number of potentially relevant articles to 519 in a few hours. It took us 5 days to screen titles and abstracts of these 519 articles, identify 119 potential full-text articles, and include 20 remaining articles in the review. We then added 7 more articles from references.

Our software allowed us to omit the time-consuming step of devising and tuning a specific search query combining logic keywords (« AND », « OR », and « NOT ») and modifying the search to suit the different requirements of each database. Traditional systematic searches require prior knowledge of the

topic of interest (the deeper the better). While deep knowledge and tailored search strings offer benefits, they also risk missing relevant articles on topics one did not think to include and may limit the possibility of discoveries. Our software also allowed more exhaustive searching since it relied on full-text articles instead of only abstracts.

Limitations

This version of the software is limited to databases that provide free APIs for open access to full-text articles. Databases often used for systematic reviews (eg, Web of Science or Scopus), databases commonly used by social scientists (eg, SocIndex, CINAHL, ATLA Religion, ProQuest services), and preprint servers for accessing non-peer-reviewed literature (eg, medRxiv, bioRxiv) do not provide suitable APIs. For example, Web of Science has a basic API but does not provide access to full-text articles. Some databases like Scopus or ProQuest provide a free basic API from which we could retrieve basic metadata but accessing more detailed information and full-text data require a subscription. Medical preprint servers like medRxiv provide basic RSS feeds to obtain some metadata. Thus, our software is best for searching peer-reviewed, open-access medical literature, especially since most publications behind a paywall forbid text mining.

The LDA algorithm, like machine learning methods, has inherent limitations. The user must choose the number of topics; 10 is the default, but the ideal number of topics differs from corpus to corpus [24]. Our stepwise topic modelling approach mitigated this limitation. The LDA algorithm also performs badly on small corpora [24], so as our corpus was reduced, risk of incoherent topics increased. Therefore, we halted our iterations when the corpus shrank too much and analyzed the parent corpus instead.

Our software only works on PDFs, but not yet on image-based PDFs. It also does not parse HTML-only publications. We are working on integrating optical character recognition to translate images to texts. This limitation was reduced by our finding that image-based PDFs were usually less-relevant, older articles.

Our broad search strategy (“HIV AND Malawi” anywhere in the text) retrieved many irrelevant articles that mentioned Malawi only in the references, so we intend to add the option to exclude references from the search. We also plan to allow searches to be performed in specific parts of the publications, similar to a traditional systematic review (title, abstract, full text). However, some databases do not allow this, and each provider has a different approach to those searches.

Because the topics identified by the software sometimes overlapped, each article might fall under several topics, and each article was attributed only to the highest scoring topic, while ambiguous or multiple attributions were ignored. Consequently, a relevant publication might have been classified in a topic of no interest and thus was not incorporated in the final result. We checked to see if the 7 additional publications identified from references were collected by our tool but not included and, therefore, misclassified. Only one such publication was identified [36]; the other 6 were either not extracted or were

retrieved but not used for classification as the full text was missing.

Our software is intended to complement rather than replace systematic reviews. We have not yet compared our approach to a systematic review, and we expect we missed relevant articles unavailable through open access. We also expect that our software missed some topics, while also finding new topics. As the number of open access articles and preprints increases [51], and as journals and preprint sites add APIs, we expect our software to become more useful. We plan to add more databases as we gain access.

We may have missed factors that influence HIV transmission in Malawi, and it could be that broadening the search to “Malawi” would overcome this limitation and reveal possible interactions between social, political, economic, and other factors that influence the course of the HIV epidemic but have never been studied in the context of HIV.

HIV-Specific Discussion

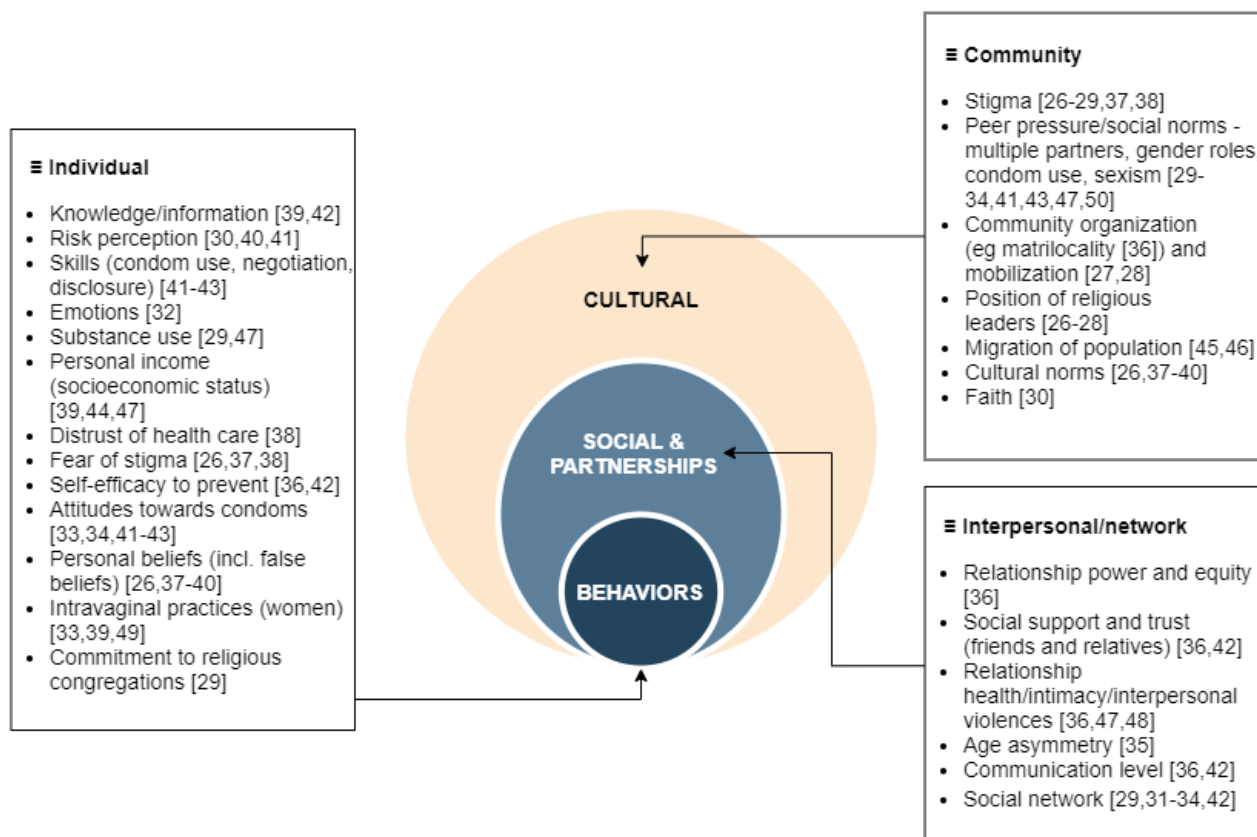
Figure 3 shows the identified factors in a 3-level diagram, modeled after that of Kaufman et al [52]. The 27 identified articles include 24 behavioral, social, and cultural factors of HIV infection. Compared to previous literature reviews on HIV-related factors observed worldwide [52-55], a few behavioral-related factors were not retrieved in our study. These include, for example, denial of HIV status, motivation and intention to change sexual behavior, reactions to stress, physical and mental health status, outcome expectancies (ie, anticipated consequences as a result of engaging in a specific behavior),

and empowerment. For social factors, relationship satisfaction and level of relationship commitment were also missing. Finally, we found no article discussing racism.

The 2015-2020 National HIV Prevention Strategy plan from the National AIDS Commission of Malawi targets specific HIV-related interventions at multiple levels for different population groups [56]. We compared the many behavioral, social, and cultural factors of HIV infection mentioned by the Prevention Strategy plan with the factors identified by our software. For the key populations of men having sex with men and sex workers, reducing the number of partners, consistent use of condoms, targeted campaigns on HIV testing and risk reduction, alcohol and substance abuse, positive health, and gender-based violence prevention programs are all behavioral interventions targeting risk factors that are present in our systematic review. Regarding the priority populations, additional interventions such as comprehensive sexuality and messages on intergenerational sex for young women at risk, stop early marriage campaigns, female support for voluntary medical male circumcision, HIV testing and counselling, and communication for couples are also addressing part of our behavioral list of HIV risk factors.

Behavioral factors, at an individual or a community level, that are not targeted by the Prevention Strategy plan for 2015-2020 are intravaginal practices. For the general population, distrust of health care, commitment to religious congregations, and the position of religious leaders are not addressed either. These are factors that should be considered when elaborating the new Malawi Prevention Strategy Plan for 2021-2026.

Figure 3. Summary of factors found within publications identified by the review.



Conclusions

From a set of articles limited by the existence of journal paywalls, our Python software quickly narrowed a set of over 16,000 articles to a small set of relevant articles. We identified sociobehavioral factors, including factors related to society and culture, such as folk beliefs, theology, and moral standards, that

may influence the course of the HIV epidemic, yet are rarely considered in the quantitative literature. Extending our approach to other countries could give researchers a more complete picture of the different drivers of the HIV epidemic in different settings and clarify the reasons for the spatial variability of HIV across sub-Saharan Africa.

Acknowledgments

The authors would like to thank Aziza Merzouki for fruitful discussions. This work was supported by the Swiss National Foundation (SNF) under grant #163878.

Authors' Contributions

AT wrote the software and obtained and analyzed the results it produced. AT further read the remaining articles after software screening and wrote the first draft of this article. IT helped read the remaining articles after software cleaning and identify relevant risk factors. IT also helped write the results section of this article. EO wrote the HIV-specific subsection of the discussion. KT provided English edits and helped identify and correct unclear sentences in the article. OK reviewed the paper and helped write the introduction.

Conflicts of Interest

None declared.

References

1. HIV/AIDs in Namibia: Behavioural and Contextual Factors Driving the Epidemic. Government of the Republic of Namibia. Ministry of Health and Social Services. 2009. URL: <http://www.infocenter.nercha.org.sz/sites/default/files/NamibiaHIVDrivers.pdf> [accessed 2020-03-23]
2. Zulu LC, Kalipeni E, Johannes E. Analyzing spatial clustering and the spatiotemporal nature and trends of HIV/AIDS prevalence using GIS: the case of Malawi, 1994-2010. *BMC Infect Dis* 2014 May 23;14:285 [FREE Full text] [doi: [10.1186/1471-2334-14-285](https://doi.org/10.1186/1471-2334-14-285)] [Medline: [24886573](https://pubmed.ncbi.nlm.nih.gov/24886573/)]
3. Ngesa O, Mwambi H, Achia T. Bayesian spatial semi-parametric modeling of HIV variation in Kenya. *PLoS One* 2014;9(7):e103299 [FREE Full text] [doi: [10.1371/journal.pone.0103299](https://doi.org/10.1371/journal.pone.0103299)] [Medline: [25061669](https://pubmed.ncbi.nlm.nih.gov/25061669/)]
4. Poulin M, Muula AS. An inquiry into the uneven distribution of women's HIV infection in rural Malawi. *Demogr Res* 2011;25:869-902 [FREE Full text] [doi: [10.4054/DemRes.2011.25.28](https://doi.org/10.4054/DemRes.2011.25.28)] [Medline: [28936119](https://pubmed.ncbi.nlm.nih.gov/28936119/)]
5. Zulu LC, Kalipeni E, Johannes E. Analyzing spatial clustering and the spatiotemporal nature and trends of HIV/AIDS prevalence using GIS: the case of Malawi, 1994-2010. *BMC Infect Dis* 2014 May 23;14:285 [FREE Full text] [doi: [10.1186/1471-2334-14-285](https://doi.org/10.1186/1471-2334-14-285)] [Medline: [24886573](https://pubmed.ncbi.nlm.nih.gov/24886573/)]
6. Manda S, Masenyetse L, Cai B, Meyer R. Mapping HIV prevalence using population and antenatal sentinel-based HIV surveys: a multi-stage approach. *Popul Health Metr* 2015;13:22 [FREE Full text] [doi: [10.1186/s12963-015-0055-z](https://doi.org/10.1186/s12963-015-0055-z)] [Medline: [26336361](https://pubmed.ncbi.nlm.nih.gov/26336361/)]
7. Feldacker C, Emch M, Ennett S. The who and where of HIV in rural Malawi: Exploring the effects of person and place on individual HIV status. *Health Place* 2010 Sep;16(5):996-1006 [FREE Full text] [doi: [10.1016/j.healthplace.2010.06.004](https://doi.org/10.1016/j.healthplace.2010.06.004)] [Medline: [20598623](https://pubmed.ncbi.nlm.nih.gov/20598623/)]
8. Tomita A, Vandormael AM, Bärnighausen T, de Oliveira T, Tanser F. Social Disequilibrium and the Risk of HIV Acquisition: A Multilevel Study in Rural KwaZulu-Natal Province, South Africa. *J Acquir Immune Defic Syndr* 2017 Jun 01;75(2):164-174 [FREE Full text] [doi: [10.1097/QAI.0000000000001349](https://doi.org/10.1097/QAI.0000000000001349)] [Medline: [28291049](https://pubmed.ncbi.nlm.nih.gov/28291049/)]
9. Baranczuk Z, Estill J, Blough S, Meier S, Merzouki A, Maathuis MH, et al. Socio-behavioural characteristics and HIV: findings from a graphical modelling analysis of 29 sub-Saharan African countries. *J Int AIDS Soc* 2019 Dec;22(12):e25437 [FREE Full text] [doi: [10.1002/jia2.25437](https://doi.org/10.1002/jia2.25437)] [Medline: [31854506](https://pubmed.ncbi.nlm.nih.gov/31854506/)]
10. Zamawe C. Factors that Affect Maternal Care Seeking Behaviour and the Choice of Practitioner(s) during Complications: the Case of Mang'anja Tribe in Malawi. *Research on Humanities and Social Sciences* 2013;3(18):18-26 [FREE Full text]
11. Ashforth A. When the Vampires Come for You: A True Story of Ordinary Horror. *Social Research* 2014;81:851-882 [FREE Full text]
12. Babalola S. Women's education level, antenatal visits and the quality of skilled antenatal care: a study of three African countries. *J Health Care Poor Underserved* 2014 Feb;25(1):161-179. [doi: [10.1353/hpu.2014.0049](https://doi.org/10.1353/hpu.2014.0049)] [Medline: [24509018](https://pubmed.ncbi.nlm.nih.gov/24509018/)]
13. Conroy AA. 'It means there is doubt in the house': perceptions and experiences of HIV testing in rural Malawi. *Cult Health Sex* 2014;16(4):397-411 [FREE Full text] [doi: [10.1080/13691058.2014.883645](https://doi.org/10.1080/13691058.2014.883645)] [Medline: [24580127](https://pubmed.ncbi.nlm.nih.gov/24580127/)]

14. Jacques G. "Coming Out" or Coming in? Social Exclusion of Sexual Minorities in Africa: Challenges for Social Work Education and Practice. *Journal of Gay & Lesbian Social Services* 2014 Feb 24;26(1):91-110. [doi: [10.1080/10538720.2013.829396](https://doi.org/10.1080/10538720.2013.829396)]
15. Mkandawire P, Luginaah I, Baxter J. Growing up an orphan: vulnerability of adolescent girls to HIV in Malawi. *Trans Inst Br Geogr* 2013 Mar 08;39(1):128-139. [doi: [10.1111/tran.12002](https://doi.org/10.1111/tran.12002)]
16. Palk L, Blower S. Geographic variation in sexual behavior can explain geospatial heterogeneity in the severity of the HIV epidemic in Malawi. *BMC Med* 2018 Feb 09;16:22 [FREE Full text] [doi: [10.1186/s12916-018-1006-x](https://doi.org/10.1186/s12916-018-1006-x)] [Medline: [29422096](https://pubmed.ncbi.nlm.nih.gov/29422096/)]
17. Phiri N, Haas AD, Msukwa MT, Tenthani L, Keiser O, Tal K. "I found that I was well and strong": Women's motivations for remaining on ART under Option B+ in Malawi. *PLoS One* 2018;13(6):e0197854 [FREE Full text] [doi: [10.1371/journal.pone.0197854](https://doi.org/10.1371/journal.pone.0197854)] [Medline: [29874247](https://pubmed.ncbi.nlm.nih.gov/29874247/)]
18. Merzouki A, Styles A, Estill J, Baranczuk Z, Petrie K, Keiser O. Identifying groups of people with similar socio-behavioural characteristics in Malawi to inform HIV interventions: a Latent Class Analysis. medRxiv 2019 Dec 27 (forthcoming)(forthcoming) [FREE Full text] [doi: [10.1101/2019.12.26.19015560](https://doi.org/10.1101/2019.12.26.19015560)]
19. Tenthani L, Haas AD, Tweya H, Jahn A, van OJJ, Chimbandira F, et al. Retention in care under universal antiretroviral therapy for HIV-infected pregnant and breastfeeding women ('Option B+') in Malawi. *AIDS* 2014 Feb 20;28(4):589-598 [FREE Full text] [doi: [10.1097/QAD.000000000000143](https://doi.org/10.1097/QAD.000000000000143)] [Medline: [24468999](https://pubmed.ncbi.nlm.nih.gov/24468999/)]
20. Hellandendu J. Contributory factors to the spread of HIV/AIDS and its impacts in sub-Saharan African countries. *European Scientific Journal* 2012 Jun 29;8(14):144-156 [FREE Full text] [doi: [10.19044/esj.2012.v8n14p%25p](https://doi.org/10.19044/esj.2012.v8n14p%25p)]
21. DOI solver. URL: <https://dx.doi.org> [accessed 2020-07-24]
22. Natural Language Toolkit. NLTK 3.5 documentation. 2020 Apr 13. URL: <https://www.nltk.org> [accessed 2020-07-16]
23. Machine Learning in Python. scikit-learn. URL: <https://scikit-learn.org/stable/> [accessed 2020-07-16]
24. Blei D, Ng A, Jordan, MI. Latent Dirichlet Allocation. *Journal of Machine Learning Research* 2003;3:993-1022. [doi: [10.1162/jmlr.2003.3.4-5.993](https://doi.org/10.1162/jmlr.2003.3.4-5.993)]
25. A freeze of the ASR as it was used for the paper "Social, behavioural, and cultural factors of HIV in Malawi: a semi-automated systematic review" submitted to JMIR. ASR_freeze_HIVMalawiJMIR. URL: https://gitlab.com/Pezam/asr_freeze_hivmalawijmir [accessed 2020-07-16]
26. Yeatman SE, Trinitapoli J. Beyond Denomination: The Relationship between Religion and Family Planning in Rural Malawi. *Demogr Res* 2008 Oct 24;19(55):1851-1882 [FREE Full text] [doi: [10.4054/DemRes.2008.19.55](https://doi.org/10.4054/DemRes.2008.19.55)] [Medline: [20463916](https://pubmed.ncbi.nlm.nih.gov/20463916/)]
27. Trinitapoli J. Religious Responses to AIDS in Sub-Saharan Africa: An Examination of Religious Congregations in Rural Malawi. *Review of Religious Research* 2006;47(3):253.
28. Adams J, Trinitapoli J. The Malawi Religion Project: Data collection and selected analyses. *Demogr Res* 2009 Sep 01;21(4):255-288 [FREE Full text] [doi: [10.4054/DemRes.2009.21.10](https://doi.org/10.4054/DemRes.2009.21.10)] [Medline: [20072664](https://pubmed.ncbi.nlm.nih.gov/20072664/)]
29. Lau FK, Jayakumar S, Sgaier SK. Understanding the socio-economic and sexual behavioural correlates of male circumcision across eleven voluntary medical male circumcision priority countries in southeastern Africa. *BMC Public Health* 2015 Aug 22;15:813 [FREE Full text] [doi: [10.1186/s12889-015-2135-1](https://doi.org/10.1186/s12889-015-2135-1)] [Medline: [26297202](https://pubmed.ncbi.nlm.nih.gov/26297202/)]
30. Anglewicz P, Clark S. The effect of marriage and HIV risks on condom use acceptability in rural Malawi. *Soc Sci Med* 2013 Nov;97:29-40 [FREE Full text] [doi: [10.1016/j.socscimed.2013.06.024](https://doi.org/10.1016/j.socscimed.2013.06.024)] [Medline: [24161086](https://pubmed.ncbi.nlm.nih.gov/24161086/)]
31. Reniers G, Tfaily R. Polygyny and HIV in Malawi. *Demogr Res* 2008 Oct 10;19(53):1811-1830 [FREE Full text] [doi: [10.4054/DemRes.2008.19.53](https://doi.org/10.4054/DemRes.2008.19.53)] [Medline: [21949484](https://pubmed.ncbi.nlm.nih.gov/21949484/)]
32. Becker S, Taulo FO, Hindin MJ, Chipeta EK, Loll D, Tsui A. Pilot study of home-based delivery of HIV testing and counseling and contraceptive services to couples in Malawi. *BMC Public Health* 2014 Dec 20;14(1):1309 [FREE Full text] [doi: [10.1186/1471-2458-14-1309](https://doi.org/10.1186/1471-2458-14-1309)] [Medline: [25526799](https://pubmed.ncbi.nlm.nih.gov/25526799/)]
33. Kumwenda J, Makanani B, Taulo F, Nkhoma C, Kafulafula G, Li Q, et al. Natural History and Risk Factors Associated with Early and Established HIV Type 1 Infection among Reproductive - Age Women in Malawi. *Clin Infect Dis* 2008 Jun 15;46(12):1913-1920. [doi: [10.1086/588478](https://doi.org/10.1086/588478)]
34. Stephenson R. Community-level gender equity and extramarital sexual risk-taking among married men in eight African countries. *Int Perspect Sex Reprod Health* 2010 Dec;36(4):178-188 [FREE Full text] [doi: [10.1363/3617810](https://doi.org/10.1363/3617810)] [Medline: [21245024](https://pubmed.ncbi.nlm.nih.gov/21245024/)]
35. Beauclair R, Helleringer S, Hens N, Delva W. Age differences between sexual partners, behavioural and demographic correlates, and HIV infection on Likoma Island, Malawi. *Sci Rep* 2016 Nov 02;6:36121 [FREE Full text] [doi: [10.1038/srep36121](https://doi.org/10.1038/srep36121)] [Medline: [27805053](https://pubmed.ncbi.nlm.nih.gov/27805053/)]
36. Schatz E. "Take your mat and go!": rural Malawian women's strategies in the HIV/AIDS era. *Cult Health Sex* 2005;7(5):479-492. [doi: [10.1080/13691050500151255](https://doi.org/10.1080/13691050500151255)] [Medline: [16864217](https://pubmed.ncbi.nlm.nih.gov/16864217/)]
37. Regnerus MD, Salinas V. Religious Affiliation and Aids-Based Discrimination in Sub-Saharan Africa. *Review of Religious Research* 2007;48(4):385-400 [FREE Full text]
38. Westercamp N, Bailey RC. Acceptability of male circumcision for prevention of HIV/AIDS in sub-Saharan Africa: a review. *AIDS Behav* 2007 May;11(3):341-355 [FREE Full text] [doi: [10.1007/s10461-006-9169-4](https://doi.org/10.1007/s10461-006-9169-4)] [Medline: [17053855](https://pubmed.ncbi.nlm.nih.gov/17053855/)]

39. Ramjee G, Daniels B. Women and HIV in Sub-Saharan Africa. *AIDS Res Ther* 2013 Dec 13;10(1):30 [FREE Full text] [doi: [10.1186/1742-6405-10-30](https://doi.org/10.1186/1742-6405-10-30)] [Medline: [24330537](https://pubmed.ncbi.nlm.nih.gov/24330537/)]
40. Anglewicz PA, Bignami-Van Assche S, Clark S, Mkandawire J. HIV risk among currently married couples in rural Malawi: what do spouses know about each other? *AIDS Behav* 2010 Feb;14(1):103-112 [FREE Full text] [doi: [10.1007/s10461-008-9497-7](https://doi.org/10.1007/s10461-008-9497-7)] [Medline: [19051003](https://pubmed.ncbi.nlm.nih.gov/19051003/)]
41. Fedor TM, Kohler H, Behrman JR. The Impact of Married Individuals Learning HIV Status in Malawi: Divorce, Number of Sexual Partners, and Condom Use With Spouses. *Demography* 2015 Feb;52(1):259-280 [FREE Full text] [doi: [10.1007/s13524-014-0364-z](https://doi.org/10.1007/s13524-014-0364-z)] [Medline: [25582891](https://pubmed.ncbi.nlm.nih.gov/25582891/)]
42. Crittenden KS, Kaponda CPN, Jere DL, McCreary LL, Norr KF. Participation and diffusion effects of a peer-intervention for HIV prevention among adults in rural Malawi. *Soc Sci Med* 2015 May;133:136-144 [FREE Full text] [doi: [10.1016/j.socscimed.2015.03.055](https://doi.org/10.1016/j.socscimed.2015.03.055)] [Medline: [25864150](https://pubmed.ncbi.nlm.nih.gov/25864150/)]
43. Mwale M, Muula AS. Effects of adolescent exposure to behaviour change interventions on their HIV risk reduction in Northern Malawi: a situation analysis. *SAHARA J* 2018 Dec;15(1):146-154 [FREE Full text] [doi: [10.1080/17290376.2018.1529612](https://doi.org/10.1080/17290376.2018.1529612)] [Medline: [30278823](https://pubmed.ncbi.nlm.nih.gov/30278823/)]
44. Baird SJ, Garfein RS, McIntosh CT, Ozler B. Effect of a cash transfer programme for schooling on prevalence of HIV and herpes simplex type 2 in Malawi: a cluster randomised trial. *Lancet* 2012 Apr 07;379(9823):1320-1329. [doi: [10.1016/S0140-6736\(11\)61709-1](https://doi.org/10.1016/S0140-6736(11)61709-1)] [Medline: [22341825](https://pubmed.ncbi.nlm.nih.gov/22341825/)]
45. Helleringer S, Mkandawire J, Kalilani-Phiri L, Kohler H. Cohort Profile: The Likoma Network Study (LNS). *Int J Epidemiol* 2014 Apr;43(2):545-557 [FREE Full text] [doi: [10.1093/ije/dyt001](https://doi.org/10.1093/ije/dyt001)] [Medline: [23543589](https://pubmed.ncbi.nlm.nih.gov/23543589/)]
46. Anglewicz P. Migration, marital change, and HIV infection in Malawi. *Demography* 2012 Feb;49(1):239-265 [FREE Full text] [doi: [10.1007/s13524-011-0072-x](https://doi.org/10.1007/s13524-011-0072-x)] [Medline: [22109083](https://pubmed.ncbi.nlm.nih.gov/22109083/)]
47. Conroy AA, Chilungo A. Male victims of sexual violence in rural Malawi: the overlooked association with HIV infection. *AIDS Care* 2014;26(12):1576-1580 [FREE Full text] [doi: [10.1080/09540121.2014.931562](https://doi.org/10.1080/09540121.2014.931562)] [Medline: [24992179](https://pubmed.ncbi.nlm.nih.gov/24992179/)]
48. Kamndaya M, Kazembe LN, Vearey J, Kabiru CW, Thomas L. Material deprivation and unemployment affect coercive sex among young people in the urban slums of Blantyre, Malawi: A multi-level approach. *Health Place* 2015 May;33:90-100 [FREE Full text] [doi: [10.1016/j.healthplace.2015.03.001](https://doi.org/10.1016/j.healthplace.2015.03.001)] [Medline: [25814337](https://pubmed.ncbi.nlm.nih.gov/25814337/)]
49. Low N, Chersich MF, Schmidlin K, Egger M, Francis SC, van de Wijgert JHHM, et al. Intravaginal practices, bacterial vaginosis, and HIV infection in women: individual participant data meta-analysis. *PLoS Med* 2011 Feb 15;8(2):e1000416 [FREE Full text] [doi: [10.1371/journal.pmed.1000416](https://doi.org/10.1371/journal.pmed.1000416)] [Medline: [21358808](https://pubmed.ncbi.nlm.nih.gov/21358808/)]
50. Wirtz AL, Jumbe V, Trapence G, Kamba D, Umar E, Ketende S, et al. HIV among men who have sex with men in Malawi: elucidating HIV prevalence and correlates of infection to inform HIV prevention. *J Int AIDS Soc* 2013;16 Suppl 3:18742 [FREE Full text] [Medline: [24321110](https://pubmed.ncbi.nlm.nih.gov/24321110/)]
51. Trends for open access to publications. European Commission. URL: https://ec.europa.eu/info/research-and-innovation/strategy/goals-research-and-innovation-policy/open-science/open-science-monitor/trends-open-access-publications_en [accessed 2020-03-04]
52. Kaufman MR, Cornish F, Zimmerman RS, Johnson BT. Health behavior change models for HIV prevention and AIDS care: practical recommendations for a multi-level approach. *J Acquir Immune Defic Syndr* 2014 Aug 15;66 Suppl 3:S250-S258 [FREE Full text] [doi: [10.1097/QAI.0000000000000236](https://doi.org/10.1097/QAI.0000000000000236)] [Medline: [25007194](https://pubmed.ncbi.nlm.nih.gov/25007194/)]
53. Coates TJ, Richter L, Caceres C. Behavioural strategies to reduce HIV transmission: how to make them work better. *Lancet* 2008 Aug 23;372(9639):669-684 [FREE Full text] [doi: [10.1016/S0140-6736\(08\)60886-7](https://doi.org/10.1016/S0140-6736(08)60886-7)] [Medline: [18687459](https://pubmed.ncbi.nlm.nih.gov/18687459/)]
54. Faust L, Yaya S. The effect of HIV educational interventions on HIV-related knowledge, condom use, and HIV incidence in sub-Saharan Africa: a systematic review and meta-analysis. *BMC Public Health* 2018 Nov 13;18(1):1254 [FREE Full text] [doi: [10.1186/s12889-018-6178-y](https://doi.org/10.1186/s12889-018-6178-y)] [Medline: [30424761](https://pubmed.ncbi.nlm.nih.gov/30424761/)]
55. Bekker L, Beyrer C, Quinn TC. Behavioral and biomedical combination strategies for HIV prevention. *Cold Spring Harb Perspect Med* 2012 Aug 01;2(8):a007435 [FREE Full text] [doi: [10.1101/cshperspect.a007435](https://doi.org/10.1101/cshperspect.a007435)] [Medline: [22908192](https://pubmed.ncbi.nlm.nih.gov/22908192/)]
56. National HIV Prevention Strategy 2015-2020. National AIDS Commission. 2014. URL: http://www.aidsmalawi.org.mw/viewhiv/Malawi%20HIV%20Prevention%20%20Strategy%202015%20-%202020_Final.pdf [accessed 2020-03-04]

Abbreviations

- API:** application programming interface
- DOI:** Digital Object Identifier
- LDA:** latent Dirichlet allocation
- PMC:** PubMed Central
- TF-IDF:** term frequency-inverse document frequency

Edited by G Eysenbach; submitted 23.03.20; peer-reviewed by C Nweneka, C Xu; comments to author 09.05.20; revised version received 20.05.20; accepted 04.06.20; published 14.08.20

Please cite as:

Thiabaud A, Triulzi I, Orel E, Tal K, Keiser O

Social, Behavioral, and Cultural factors of HIV in Malawi: Semi-Automated Systematic Review

J Med Internet Res 2020;22(8):e18747

URL: <http://www.jmir.org/2020/8/e18747/>

doi: [10.2196/18747](https://doi.org/10.2196/18747)

PMID:

©Amaury Thiabaud, Isotta Triulzi, Erol Orel, Kali Tal, Olivia Keiser. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 14.08.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.