

Original Article

A Text-Mining Study on Sustainable Health Solution for India's Most Vulnerable Population

Somesh Dhamija¹, Mohammed Jamshed^{2,*}, Asad Ahmad², Aruna Dhamija¹, Waseem Khan³, Deepika Pandoi⁴, Rahisha⁴¹Institute of Legal Studies & Research, GLA University, Mathura, India²Department of Management, Jamia Hamdard, New Delhi, India³School of Business, Dr. Vishwanath Karad MIT World Peace University, Pune, India⁴Institute of Business Management, GLA University, Mathura, India

ARTICLE INFO

Received 03 October 2025

Revised 04 November 2025

Available Online 10 November 2025

Keywords:

Text-mining

Telemedicine

Rural Healthcare

Urban Healthcare

Healthcare Inequity

Mobile Health (mHealth)

Community-based Healthcare

ABSTRACT

Health disparity between urban and rural India has existed since very long back in history. Text mining techniques in the given context are used to compile 140 peer-reviewed journal articles published in the WoS database focusing more on the core theme, disparities, as well as the policy aspects related to the delivery of health services in both the urban and rural sectors. The results explain the disparities in the health infrastructure available to the country as a whole, adherence to digital health technology, as well as the financial status. Urban areas possess improved health services as well as advanced digital platforms compared to rural areas, where basic services as well as digital technologies are currently lagging. The factors identified as the bridge between the gaps in rural health services are the greater acceptance of telemedicine, mobile health (mHealth), and AI-based interventions. Community programs initiated under Accredited Social Health Activists (ASHA)s and non-governmental organizations (NGOs) worked towards adding positively to improved accessibility. Government programs like Ayushman Bharat and National Rural Health Mission (NRHM) improved health insurance coverage. However, fears about the functionality remain. This study emphasizes the importance of informed decision-making, effective financing strategies for the protection of health, or rather the development of appropriate community-based responses to health care. This particular publication contributes towards the body of knowledge in the area of inequality in health care delivery in India.

This is an Open Access journal, and articles are distributed under the terms of the [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author[s] and the source.

Introduction

Effecting the welfare and health of the population and the health status of the population as a whole, access to good health care has great significance. In the Indian scenario, the health environment found in the cities is quite different from the rural society. Better health care facilities and equipment used for health care activities and advancements in medical equipment and

*Corresponding author: Dr. Mohammed Jamshed, Assistant Professor, Department of Management, Jamia Hamdard, New Delhi, India.

<https://doi.org/10.31531/jprst.1000198>

accessibility to the latest technological advancements are found to dominate the city. This area has been covered densely by trained health staff like Specialists, Paramedical staff, and Nurses, thereby adding up to the good health care services provided to the city residents [1]. The availability of the total private health staff in the city has contributed to easy access to total health care.

Yet, rural India has some chronic health-related issues. They lack adequate health infrastructure and staff. The rural health centers lack appropriate medical equipment and old machines as well as irregular power connections. This has the effect of not allowing quality health services to reach the rural health centers [2]. Lack of specialized health services means that the rural population has to travel long distances to the urban centers for life-saving surgeries [3].

Even the health divide connects both preventive health and public health domains. People living in urban areas have access to greater opportunities in immunization campaigns, health awareness campaigns, and screening programs as the communication channels are easier to identify and the facilities are more accessible. However, the rural areas are also plagued by the lack of proper communication channels for health messages and the implementation of health programs [4]. The lack of proper road connectivity adds to all the above problems as it becomes very difficult for health practitioners to reach the rural areas.

Urban-rural health disparities continue to adversely affect the following vulnerable populations: women, children, the elderly, and the marginalized. Rural women remain inaccessible to maternal health services because of the shortage of obstetricians and gynaecologists [5]. This will inevitably lead to more maternal and infant deaths than those observed among the urban population. Children in rural areas can fall prey to preventable illnesses because of their limited immunization and diets.

The rural senior population faces other challenges related to the prevention and control of chronic illnesses like diabetes, hypertension, and heart-related issues due to the lack of specialized geriatric health services [6]. The vulnerable populations systematically find themselves in the Gandhian holes in the wall and drift further away from the health service delivery wall. Language barriers, stigma, and the lack of cultural adaptation to the health services mean that the whole rural senior population falls under the category of the under-served [7].

To overcome such imbalances, the Government of India has embarked on several programs towards the development of the health infrastructure and improved access to health in rural India. Among the most significant programs include the National Health Mission (NHM), as well as the Ayushman Bharat program, whose goal is to improve access to primary health care services, develop health infrastructure, and improve financial protection for vulnerable segments [8].

In order to understand the Ayushman Bharat Scheme, one has to recognize the transition of Health and Wellness Centers (HWCs) to provide comprehensive primary health care services in rural areas of India like health screening for non-communicable diseases and the distribution of medications as recommended. Despite the above measures being in place, the challenges still remain great. Lack of health personnel to be deployed in the rural areas forms a disincentive for service delivery. Bad working conditions and training reduce the willingness of doctors to be employed in rural areas.

Most health centers in rural areas are also suffering from inefficiencies that are bureaucratic in nature. This has created immense delay in the allocation of resources and implementation of policies [9]. Factors like income level, educational level, and living standards have also been taken in account for the deteriorating health scenario in rural areas [9]. Data analysis has to be performed to comprehend the complications present in the accessibility of health care in the country because it would be very difficult to comprehend the different trends associated with the delivery of health services through interviews and surveys for incorporating the changing demands present in the domain of health.

In this case, the use of big data analysis and text mining can assist in gaining a macroscopic vision of health inequalities. The fact that free information streams are measured in aspects of the discourse in the health information in the social media platforms, government documents, as well as the health information headlines in the newspaper industry enables scientists to monitor the upcoming health trend and the health attitudes of the people.

This enables the usage of evidence-based intelligence to develop strategies for the required zones based on health necessities. For instance, particulars concerning the factors that can withhold the implementation of telemedicine services in rural areas can be applicable when developing strategies to provide additional digital literacy skills to rural zones in the country. Moreover, the sentiment analysis of health initiatives can play a

role in recommending how trust can be built and health programs can be started. The originality of the proposed solution can be addressed based on the application of the text mining technique to handle the dissimilarities among the views considered in the unstructured information to identify patterns, issues, and core subjects concerning rural and urban solutions related to health.

The relevance of the research has immense futuristics and implications only as far as the planning of health care and formulation of policies in the Indian context. The research provides inputs towards health care challenges and aims at trying to come up with solutions towards the problems only. The entire process of text mining has produced information that can be made use of in order to plan the expenditure towards the best resources only. This refers to the health care industry towards the improved needs of the area only.

In conclusion, the above-mentioned research work offers an empirical contribution to the already available literature related to the use of unstructured information in the field of public health studies. As mentioned above, the final aim of the above-mentioned study will try to promote more equalized health care delivery systems. This will bridge the huge time difference between the health inequality disparities of urban and rural cities in countries like India. This will also try to improve the health outcomes of different classes of people in the Indian context.

Literature Review

Health Inequities in Urban and Rural India

A large volume of writing has emphasized the rural-urban disparity in the health status of the Indian population owing to variability in the delivery of services, health infrastructure, and health outcomes. The health infrastructure in the urban region can be considered very modern because of its state-of-the-art teaching institutes, greater population density of specialists among doctors, as well as the availability of advanced technological equipment. This has contributed to the general improved standard of health service. In contrast, the rural health sectors continue to grapple with understaffed Primary Health Care Centers (PHCs), as well as trained staff at the lowest level of health commodities stockout [10]. This can also be contributed to the lack of transport infrastructure and geographic inaccessibility of the rural population [11].

This causes a rural imbalance of the disease burden due to increasing cases of infectious diseases, malnutrition, and maternal deaths [12]. All this notwithstanding, the government programs such as Ayushman Bharat and

National Health Mission reveal the weaknesses that can be addressed as well as innovative ways of accessing health care within rural India.

Influence of Socioeconomic Determinants on Access to Healthcare

“Socioeconomic factors play a significant role in influencing health care access and outcome in the Indian context. Education inequality, income inequality, and employment status significantly influence the health care-seeking behavior and access to quality health services. The marginalized sections of society like women, senior citizens, and the socially disadvantaged communities also face other limitations like financial constraints and sociocultural discrediting factors that deter them from accessing health services early” [13].

Socioeconomic factors must be understood to formulate health care policies to bridge the health service accessibility divide between rural and urban areas. As per the Indian context, education inequality, income inequality, and employment status significantly influence health care behavior and access to quality health services.

Use of Text Mining in Medicine Research

The rise of big data analytics has made text mining a robust method for handling large amounts of unstructured health-related information in the field of health studies [14]. Text mining provides researchers the benefit of gaining access to a large amount of information available through health records, blog sites of social media portals like health-related forums, as well as health-related news [15].

the uses of text mining include health monitoring, health service sentiment analysis, and the detection of new trends in health-related issues. For instance, the use of text mining has been employed in the analysis of individuals' reactions during health emergencies such as the COVID-19 outbreak to note the important concerns related to the provision of health services and the adherence of the general population to the health preventive measures [16].

Comparative Analysis of Urban and Rural Health Solutions

Certain health problems peculiar to the rural and urban populations drive the need for context-specific interventions. Urban residents, as wealthy as they are, are challenged by the volume of patient flow, lifestyles, and the effects of pressure exerted on the infrastructure due to the readiness of the population [17]. The rural population experiences the shortage of health

practitioners together with the location and cultural acceptance barriers in the delivery of health services [18].

Text mining provides a clear understanding of the process of recognition and comparison of differences based on information analysis of sources such as patient opinions, health forums, and health policies. Based on the realization of the trends that have not yet happened and region-specific issues, the process of formulating region-specific health delivery and equity enhancement strategy can also be guided.

Role of Digital Health Technologies

The integration of digital health tools will reduce the health disparities between rural and urban geographic locations. Telemedicine services, mHealth solutions, and digital health tracking devices have significantly improved the coverage of health services in distant localities by reducing geographic disparities [19].

Such technologies are required in a setting where a crisis in public health exists, especially where the availability of health coverage is geographically constrained [20]. Text mining can also optimize the efficiency of the e-health program based on patient response patterns.

These two studies together establish the importance of new-age methodologies like text mining in understanding health-related aspects and drawing inference from the available data to create a link between the rural and urban health services of India. The data can be used for shaping rational policies for the greater good, thereby paving the way for equity in the health service delivery system.

Methodology

Document Search Parameters

As the researcher embarks on the study, the data has been gathered from the Web of Science (WoS), a database that encompasses the largest and most general databases used in peer-reviewed publications written under broad subjects [21]. In fact, the database encompasses the most comprehensive information in the related field of scholarly studies and has the best platform to carry out the topic modeling analysis. In the process of creating the best dataset available, the search conducted in the literatures used related to the topic at hand.

A total of 141 research papers were published this way. A Boolean operator search method was used to narrow down the field of study, and collective keywords such as “Health Solution,” “Health Access,” “Health

Innovation,” and “Sustainable Health Care” were combined by adding “Rural” or “Urban” in order to account for variation related to the problems studied. Even more innovative ways of doing the searching, like the use of the wildcard character add-ons, were used.

In order to further sharpen the search results, filters were used based on publication date (2002-2025), language (English), type (Article), and the country (India). All the above-mentioned filters were used in a bid to create a high-value and specific set of studies fitting the requirements of the chosen topic. Articles from both conceptual and empirical studies were used. Manual screening was also conducted. This involved the exclusion of irrelevant studies that were duplicated. Abstracts and keywords were filtered for relevance. This involved the inclusion of only the most relevance-yielding studies.

Text Mining Techniques

In the context of comparatively evaluating the content of the chosen literary works for the purposes of the current research, the application of techniques related to “text mining” has been used. Text mining refers to the usage of techniques to extract hidden patterns and structure from unstructured texts. Among the techniques used has been the “Latent Dirichlet Allocation” (LDA) technique. LDA has been used as the primary technique related to the analysis of texts in the context of “unsupervised” machine learning. LDA refers to a generative probabilistic model that tends to assume each document to comprise a linear combination of a fixed set of topics. Moreover, each topic tends to represent a distribution over the words [22].

This technique has been derived from the literature for the following reason: the ability to identify the “thematic structure” present in the “text” related to the “data” can sometimes lack “pre-specified” structure.

LDA was used in the research based on a set of 141 peer-reviewed documents utilizing the Web of Science database. The goal of LDA was to find recurring patterns in a set of words as well as the meaningful combinations of the words. The use of the LDA model helped in the analysis of word occurrences related to the theme of innovation and the delivery of health access through studies related to the health systems available both in urban and rural areas. Other pre-processing techniques included tokenization, removing stop words, stemming, and lemmatization. These were used in the enhancement of the performance of the models on the dataset and the generation of plots and word clouds.

The use of such techniques of text mining has enabled the exploration of the issues related to the domain of health care in the context of literary texts in a manner that adds to the analysis of the domain as a whole [23,24].

Data Visualization

Visualization of the outcome contributed significantly towards the achievement of notable results. Word clouds were useful in the visualization of the word frequency based on health, notable mention of health frequency and new health concerns based on the discourse in the related literature [25]. Frequency plots were also important towards the visualization of the detailed view of the frequency of the terms.

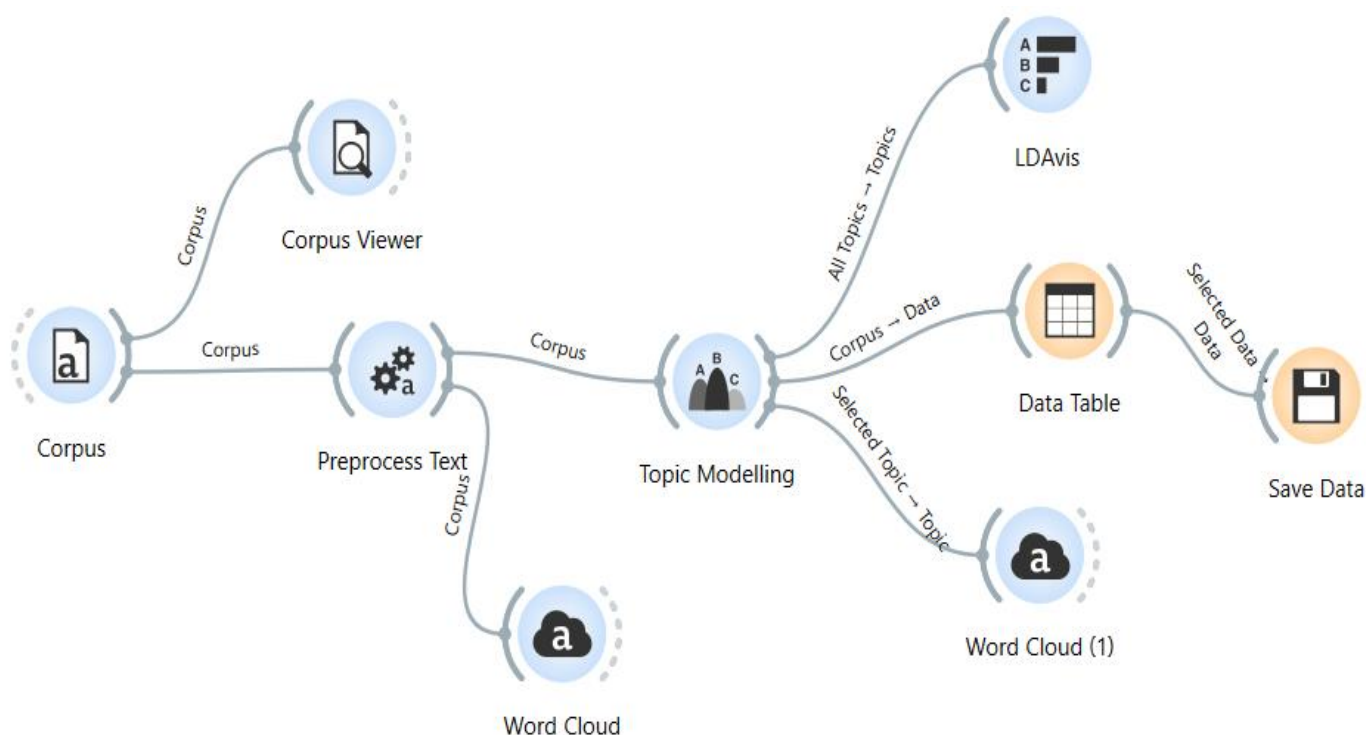


Figure 1: Text-mining workflow

Results

The outcome will be presented in two aspects. The first aspect would involve the summarization of the gathered 141 research papers through the use of bibliometric analysis. This would include the subjects covered in the studies, the publication pattern per year through journals, as well as the word cloud representation of the commonly used words. Bibliometric analysis helps in the detection of the dominant pattern present in the studies as well as the theme present in the data. Secondly, LDA helps in the detection of the theme and the visualization of the theme distribution among the collected articles. The second aspect consists of the selection of typical articles under each topic for qualitative analysis. This makes sure that the trend identified under topic modeling can be related to the details mentioned in the studies. This topic modeling and qualitative thematic analysis provides the complete picture of the disparities in the health field and proposes solutions both for urban and rural areas of India.

Trends and Trend in Research Output

In Figure 2, the graph of the publication of the research work from the year 2002-2025 clearly indicates that the rise in the publication of the academic interest of the topic has been enormous. Prior to the year 2013, very few papers were published.

However, since 2014, the trend is clearly rising with a very high peak in the publication of research output between 2018 and 2023. A total of 24 articles were published in 2023, the highest recorded output, indicating the peak level of academic interest in the area. This indicates a high level of acceptability of the relevance or timeliness of the area.

The fluctuations in the publication of articles from the period 2015 to 2018 and the gradual rise in 2019 signify the ever-changing nature of the trends in the research being conducted. This can also point towards the influence of factors outside the publication trends.

The drastic fall in 2025 represents the lack of information available for the year as yet since the year has not been completed. The general trend shown in the

graph indicates that the volume of literary output has been rising to represent the growing importance of the topic area.

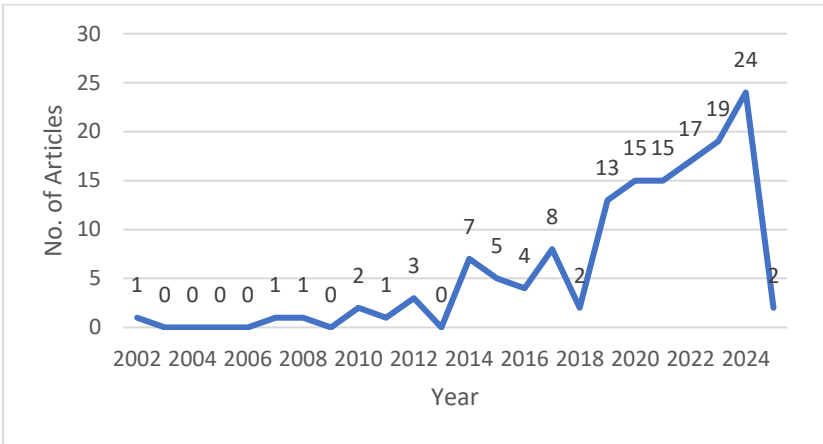


Figure 2: Articles per Year: 1992-2025

The number of publications of the research articles across the various journals shows the interdisciplinary nature of the topic (Figure 2(a)). Journal of Global Health has the highest number of publications with four articles, matching the highest concern for the topic. Indian Journal of Public Health, Clinical Epidemiology, Global Health, and BMJ Global Health follow with three articles each, signifying the highest concern for Global Health and Public Health.

In addition, five other journals such as “Studies in Health Technology and Informatics,” “Lecture Notes in Networks and Systems,” “Journal of Green Engineering,” “International Journal of Health Sciences,” and “Globalization and Health” have published two articles each. This shows the diverse areas of research the study has covered, for example, the field of health and globalization.

Analysis of the usage of the article in different journals indicates the diversified contribution of the field, confined only to the sectors of health, technology, and development across the world. From the study, the fact has been proved that despite the higher preference of the field in different journals, the research studies are spread across the world.

The fact that numerous articles related to the topic are published in numerous journals proves the interdisciplinary application of the topic in all aspects of health, tech, and development. From the study, the topic can be observed to have a widespread application despite the high interest rates in the topic in some of the journals.

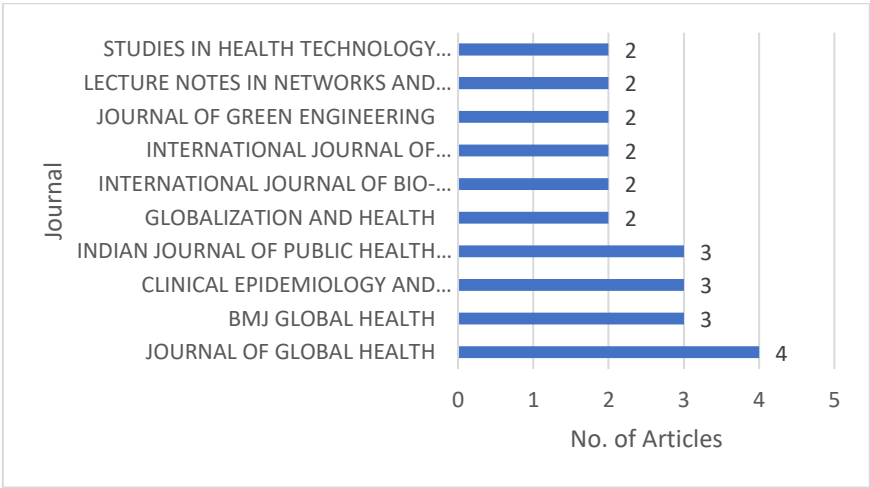


Figure 2 (a): Journal publishing distribution.

Topic Modeling

Topic modeling represents one of the popular ways of applying unsupervised machine learning to discovering hidden semantic structures in large corpora of texts [26]. Topic modeling has the capacity to help researchers identify important topics hidden in unstructured texts. Moreover, the technique has been widely used in the field of health studies to decipher patterns of language in the health domain.

In this case, topic modeling analysis used in the topic modeling process focuses on investigating alternatives available to satisfy the health requirements of marginalized sections in both urban and rural areas of India instead of finding the most popular topic.

Data Preprocessing

Prior to topic modeling, the texts were already preprocessed for a very long time to obtain the best possible performance of the model and the appropriate number of topics [27]. Texts were preprocessed in two ways. Texts were tokenized based on word form division of the sentences. Secondly, stopwords such as function words like “the,” “and,” “is” that would not add anything to the meaning of the topic were removed [28].

This lemmatization process helped in the reduction of word forms to the base/root forms. This helped in the elimination of redundancies in vocabularies [29]. Bigram/trigram phrase-tagging techniques helped in the tagging of the frequently occurring paired words (for example, “health care” or “rural hospitals”) to increase the context awareness [30].

Selection of Topic Modeling Algorithm

Latent Dirichlet Allocation (LDA) has been used as the primary topic modeling technique in this research because LDA represents topics as distributions of the probability of each word in the topic. This has proven effective for the extraction of meaningful topics from large documents [31], as LDA assumes that each document can be described as a blend of a set of few topics and their word probability based on the patterns of word occurrences [32]. Nonetheless, LDA has proven more accurate than other topic modeling techniques because LDA can handle the poly

The number of topics determined the best fit for the set of documents utilizing the coherence scores. Coherence scores calculate the likeness of word meanings among

topics based on the semantic vector. Coherence has the ability to balance the interpretability and specificity of the topics at a level that enables each topic to have meaning.

In order to compare the rural and urban health-related topics between each other, two LDA topic modeling processes were performed individually on the corpora related to rural and urban health. For topic modeling in this research, the Gensim LDA library has been used. Its implementation uses the Mallet backend because the Gibbs sampling algorithm used in the LDA in the Gensim LDA library helps to form more coherent topics [33]. The K “value specifies the actual topic” determined in the topic modeling process. Its tuning has been performed based on the coherence measures and human evaluations.

Topic 1: Healthcare Innovations and Technology Adoption

The implementation of health care innovation has presented a revolutionary outlook in accessibility as well as delivery through telemedicine, mHealth, and AI-interventions. Telemedicine has significantly bridged the rural-urban divide in health as it allows greater freedom in traveling and improved health outcomes [34]. Telemedicine has been found to reduce health disparities among rural populations as physical traveling limitations were relaxed [35].

In the same way, mHealth interventions have transformed the rural health sector. According to [36], the mobile technology-based intervention involving SMS reminders for maternal health and immunization led to additional demand for health care. This follows the diagnosis innovation facilitated by AI technology, among others. Kumar and Mehta mentioned the application of AI devices such as auto-diagnostic devices for diabetic retinopathy and tuberculosis to enable diagnosis and treatment. AI-based chatbots and health buddies improve the health of the mind and general health know-how among the rural population [37].

Despite all these advancements, digital literacy, financials, and infrastructure remain the inhibitors for the widespread implementation [38]. Government programs like the Indian NDHM will bridge the gaps through the use of EHRs and health literacy conducted through digital platforms [39]. Rural health service delivery in the Indian pursuit of equal field access can bridge the infrastructure and financial gaps [40].

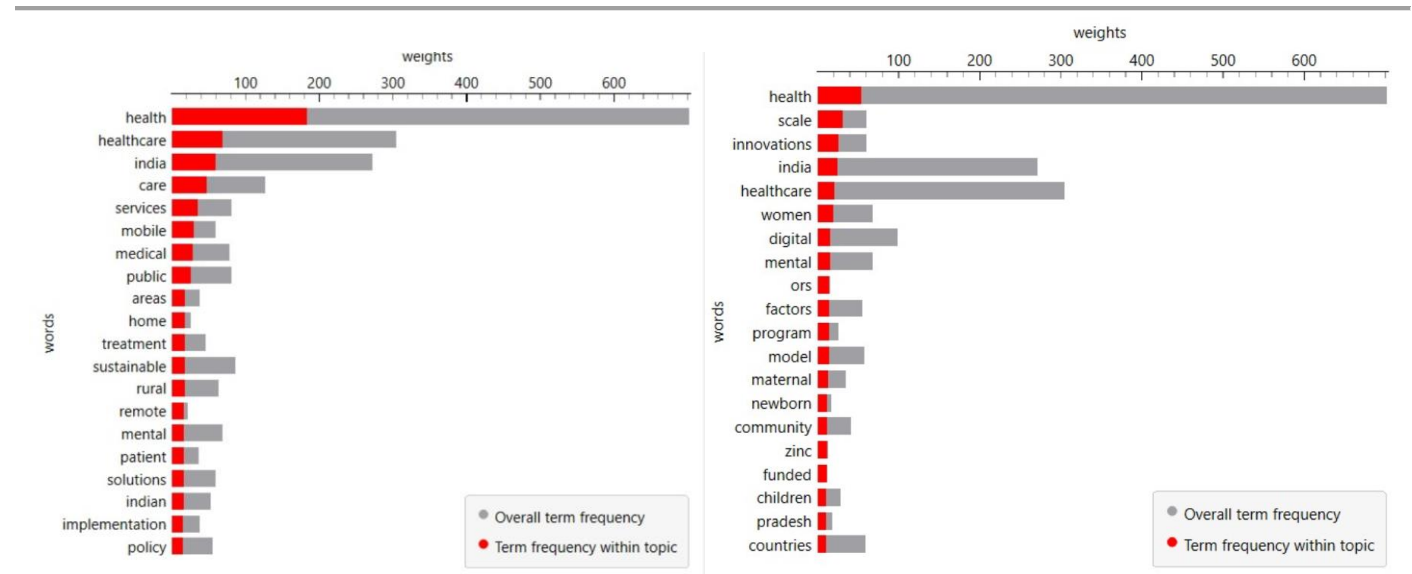


Figure 10: LDAvis for topic 1

Figure 10: LDAvis for topic 2

Topic 2: Public Health Issues and Disease Prevention

High infectious disease burden like malaria, tuberculosis, and other health related to maternal and child health are the most apparent health-related issues among the rural populations of the country. All of the above-mentioned factors become the etiological factors for the health burden due to the lack of proper health-related services and hygienic conditions as well as lack of proper health education [41].

Malaria is a massive concern in the rural zones, specifically in the tribal zones, and lack of accessibility to preventive measures such as insecticide-treated bed nets and indoor residual spraying has contributed to higher transmission rates [42]. Another massive concern in the health domain has been tuberculosis, and more cases at higher rates among rural zones have been observed because of the unhygienic living habits of the rural population as well as due to lack of knowledge about the preventive factors related to the disease [43].

Rural maternal and Neonatal health complications are related to limited antenatal care accessibility, reduced institutional delivery rates due to lack of access to antenatal services and lack of postnatal health services. The findings of Paul & Pandey (2020) suggest the role of community-based interventions like Accredited Social Health Activists (ASHA) has improved maternal and Neonatal health related outcomes due to improved antenatal checks and improved antenatal hospital delivery [44].

Public health interventions such as immunization programs, health promotion through the provision of sanitation and hygiene practices, and sanitation

programs have been employed for prevention. National government programs such as the National Vector Borne Disease Control Program (NVBDPC) and the Swachh Bharat Abhiyan have implemented policies related to the behavior of malaria and sanitation [45]. Vaccine hesitations, health education, and socio-economic factors remain the barrier in the way of the success of such programs [46].

Mass public health policies, easy accessibility of health-care services, and effective malarial control programs should be made available among rural communities in order to end such disparities [47].

Topic 3: Maternal and Child Health in India

Inequities remain widespread in the area of maternal as well as children's health in rural areas that lack access to qualified birthing assistants, antenatal care services, as well as the quality of health institutions available [39]. Rural women are also vulnerable to the risk of delay of access to appropriate maternal health services.

The figures given by Nihal and Shekhar [48] are 65% of rural women only getting four out of the four antenatal visits compared to 85% in the urban areas. Then comes the aspect of trained delivery staff, and the fact that deliveries are mainly conducted at home under the hands of untrained health staff. Government initiatives like the Janani Suraksha Yojana Scheme or the JSY scheme have promoted hospital deliveries due to incentives as well as improved newborn and maternal health [49].

Postnatal care deficit, nutrient intake, and immunization rates continue to remain as common factors among the risk factors [50]. Improving rural health centers,

increasing the density of trained midwives, and the development of maternal health programs at the community level must be achieved in order to prevent maternal as well as infant deaths. “Nutrition, education, and health service delivery must be addressed as a holistic strategy for comprehensive scaling-up of maternal and child services in rural India” [51].

Topic 4: Socioeconomic Barriers to Accessing Healthcare in Rural India

Financial constraints, lack of health insurance coverage, and the inequality of the economy factor as serious deterrents in accessing effective health care in rural India. Among the most effective factors for the usage of health care access is poverty. Out-of-pocket expenses related to consultations, medicine, and hospitalization pose a massive economical burden to the poor [52].

Medical consultations as well as hospitalization for certain rural families are delayed or even avoided due to the cost factor, and the reality of their controlled morbidity and mortality rates. In rural India, health

insurance schemes are non-existent, and people are unaware of the schemes that already exist [53].

Although the government programs like Ayushman Bharat: PM-JAY have also attempted to bridge the financial gaps and lack of implementation and insensitization, the efficacy has been reduced due to the same [54].

The truth comes to fore that the rural population covered under the health insurance also experiences the indirect cost like transport cost, loss of income, and post-hospital treatment cost due to their proximity to the actual cost of health care [55].

Bridging the above socioeconomic lines would require stronger financial protectives measures and health infrastructure development. This will increase health insurance program awareness among the population. Improving public finance in the realm of health and subsidizing core health interventions can diminish disparities in rural and urban health-care delivery services [56].

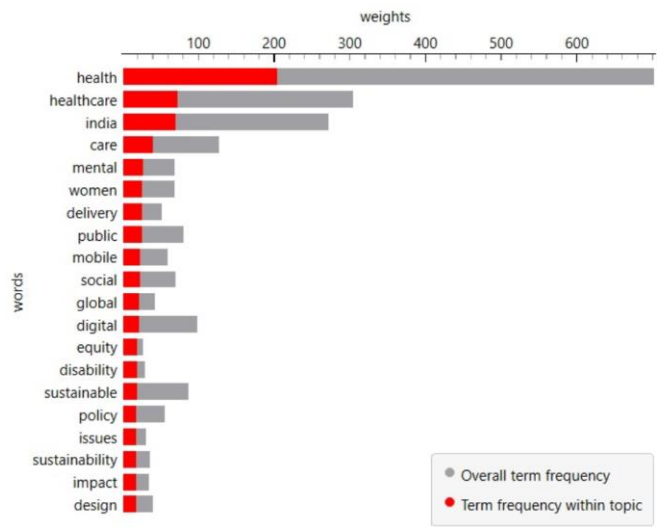


Figure 10: LDAvis for topic 3

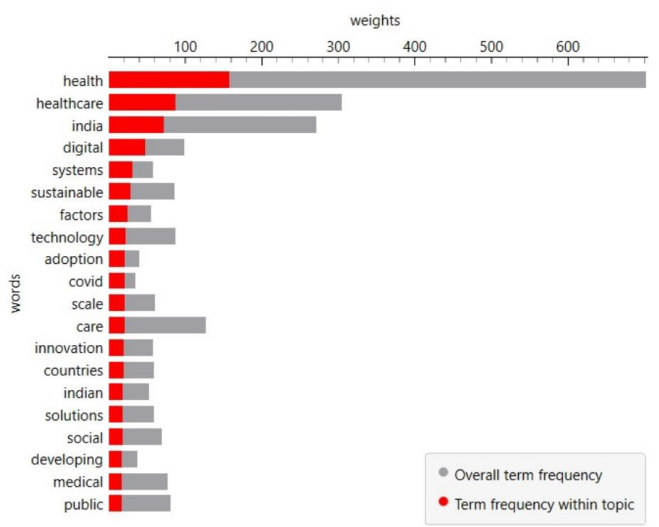


Figure 10: LDAvis for topic 4

Topic 5: The Role of Community-Based Healthcare Systems

The health service delivery gap between rural and urban India has been filled by the health systems based in the rural communities. In rural health systems, the main health service deliverers are the Accredited Social Health Activists (ASHA), health volunteers, and Self-Help Groups (SHG) in the delivery of primary health care services, maternal and child health services, and prevention services [57].

ASHA workers, further, have made a remarkable contribution to improving the scenario of institutional delivery, immunization rates, awareness at the grass roots level, and many other activities [58]. Non-governmental Organizations (NGOs) are also participating in the delivery of health services through the training of local health staff, establishment of mobile health services, and traditional medicine systems as complementary modes of the integrated modern health treatment programs [59]. Research has proven that the delivery of health services at the community level can improve access to health services

to reduce health disparities in the US and other countries as well [60]. In fact, the role of additional training, incentives, and encouragement for the greater effectiveness of health delivery and management of rural health disparities at the community health worker level should not be neglected.

Topic 6: Well-being in rural contexts and health/sustainable policy interventions

Sustainable health programs and policies for Indian health are propelled based on effective drivers for improved accessibility and affordability. The National Rural Health Mission and Ayushman Bharat programs have prominently improved primary health infrastructure access, financial protection, and preventive services [61].

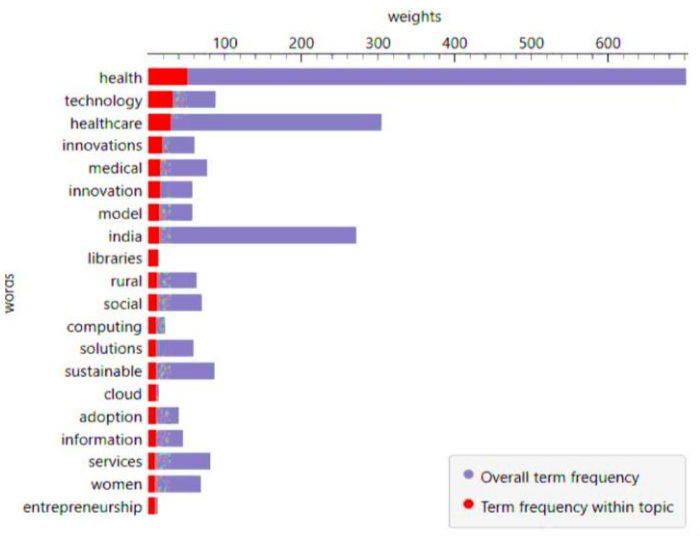


Figure 10: LDAvis for topic 5

Conclusion

The quality and availability of health care in urban and rural India are acknowledged in the study, and many challenges are due and must be addressed to avoid equal accessibility to health care. The urban region has the advantages of accessing good health services, the quality of health care, and connectivity of quality technology. The rural area has challenges related to less health service delivery, unqualified doctors, and access to primary health.

The rural disparity further diminishes the quality of exposure for rural areas, more so for women, children, senior citizens, and the poor families. All the above disparities have to be dealt with together as a whole solution related to policy changes, technological solutions, people-oriented health infrastructure development, and socio-economic safety nets.

Apart from the above, telemedicine, health units, and community interventions are cost-effective approaches to deliver health care, and they have been effective in easing rural health care issues [62].

“Policy interventions are needed to focus on improving health financing expenditure, improving last-mile delivery systems, and adopting technological interventions in order to achieve sustainable rural health outcomes,” and the aroused policy of “inclusiveness with participation with accessibility” has been described as the “engine” propelling the “removal of inequalities in health” and the “total health of the whole person” because of its focus on the “rural poor” as the “center”.

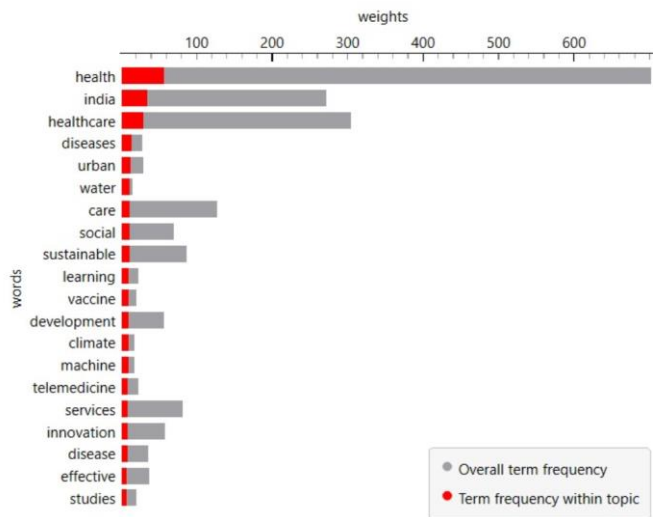


Figure 10: LDAvis for topic 6

The towns in the urban area are adequately equipped with new-age hospitals and health care centers because of which the health status of the people improves. As health service tops the list when given priority in the case of urban towns, the availability of mother care services, childcare services, and elder care services improves. Urbanized people are more educated and digitally penetrated, and their awareness related to health improves due to which the usage of health services and preventive health services goes up.

On the other hand, rural areas are plagued by chronic shortages of doctors and health centers. In rural parts of the world, the lack of specialist services, diagnostic services, and casualty services means that the patient has to go long distances. This affects their health. According to the study, rural Indian infant and maternal health causes a serious concern since most mothers are not given proper antenatal care due to the lack of trained birthing attendants. For the elderly living in rural areas

across the world, chronic health issues like diabetes, blood pressure, and heart disease are beyond their control because the required services from geriatric specialties are not available.

Among the most promising techniques available in the rural health sector today are the technological advancements related to telemedicine, mHealth solutions, and AI-related interventions. Telemedicine has significantly impacted the delivery of health services in rural areas as it has made it possible for health services to be accessed remotely. Indeed, the penetration of telemedicine has been confirmed to promote greater access to health services related to psychiatric health and the management of chronic illnesses. Nevertheless, the lack of infrastructure appropriateness and cost have been considered as barriers.

Likewise, mHealth technologies play a pivotal role in the improved health of the rural population through the provision of education services under health care delivery, appointment reminders, and maternal health services conducted through the use of mobile phones. Medical diagnostic equipment equipped with AI can detect health conditions like tuberculosis and diabetic retinopathy at a very early age to react accordingly. While digital health solutions can offer inexpensive and scalable solutions as well, the pressing concern related to connectivity issues, affordability, and Trust in health services facilitated by digital solutions should be addressed for their success in the long run.

Financial requirements are one of the fundamental factors influencing health access in rural India. Expenditure on health services out-of-pocket, the lack of health insurance coverage, and meager incomes dissuade rural families from accessing health services immediately. Findings have shown that rural families postpone and forego accessing health services due to financial constraints and thereby suffer from higher rates of morbidity and mortality. Though government initiatives such as Ayushman Bharat-Pradhan Mantri Jan Arogya Yojana (PMJAY) aim to overcome financial barriers in the form of health insurance coverage among the economically weaker sections of the society.

Rural-based health systems can also play a very important role in reducing the health disparities found in rural India. Government-accredited Social Health Activists (ASHA), volunteer community health workers, Self Help Groups (SHG), and Non-Government Organizations (NGOs) play important roles in the delivery of primary health care services, maternal and child health services, and prevention

programs. All the SHAs have contributed significantly to the acceleration of institutional delivery, immunization rates, and health awareness.

The role of the NGOs has been significant in the capacity-building activities of health personnel, establishment of mobile health units, and integration of traditional medicine. From the available information, the health systems based on the concept of the rural access to health care services model demonstrate the priority of rural access to health care services, improve health literacy, and improve the relationship between rural communities and health service providers. Scaling up the health programs based on community-oriented health concepts can increase their effectiveness.

Government policies and programs have been important factors in improving the coverage of health care and the delivery of health services in rural areas of India. NRHM and Ayushman Bharat schemes have been substantially strengthened in regard to the upgradation of primary health centers, financial components, and the promotion of preventive health behavior. Weaknesses in implementation, financial constraints, as well as disparities between regions, remain as barriers.

Text mining for purposes of evidence-based policy intervention can also be applied when dealing with the needs at the regional level of health care. For example, the factors that hamper the adoption of telemedicine services among rural zones can be employed to promote advancement in the field of connectivity. Moreover, knowledge of personal perceptions concerning interventions in health, immunization programs, and health programs for the community can assist in the effective use of communication designs to persuade individuals to benefit from health programs.

Findings from the study highlight the contrast between access to quality health care in India and access in rural health care in the country because of the challenge factors pertaining to the structure of the country and innovative technological advancements. Urban areas are endowed and beneficiaries of upgraded medical hospitals and health specialist centers and innovative technologies compared to the rural areas that lack non-existent medical health structures and health specialist doctors and appropriate health procedures to the nearest accessible health centers. Moreover, the health disparity has made the rural areas more vulnerable compared to the other areas because of their weaknesses in health security. All the imbalances above can be addressed only by the interdisciplinary endeavor of health reform restructuring, technological invention development, health Grass Root structures creation, and socio-economic welfare programs.

Implications

Strong policy implications of the finding can be drawn related to health and the role of technological interventions. In order to reverse the situation created due to disparities, the rural health infrastructure has to be set up first. Secondly, additional investment has to be made in the health centers of rural areas. This will ensure that the health chain can be set up and more availability of more rural health experts can be achieved. Public-private partnerships can also be complemented in order to get more health services because of the increase in technological development.

Secondly, the merging of telemedicine systems, AI-powered diagnostic systems, and mHealth systems has the potential to create a solution for the lack of access to health services in rural locations. For the implementation of this solution to occur, investment in digital literacy capability and the availability of the internet can play a significant role. Policymakers must come together to develop appropriate digital health policy for the acceptance of the solution among rural populations.

Third, grassroot health models should also be integrated in such a way as to promote grassroot health delivery. Self-help groups, ASHAs, as well as NGOs together contribute to maternal services, immunization services, and preventive services for disorders. Thus, supplying additional resources in the form of training and infrastructure to the grassroot health delivery personnel can greatly contribute towards realizing the ultimate health benefits for the poor classes. Finally, grassroot health models should cater to the aspect of enhancing the protection continuum in the form of universal schemes like universal protection and insurance to reduce the toll of cost in the rural areas. This enhanced level of protection will promote greater consumption behavior related to health services and health practices belonging to the category of poverty.

Limitations and Future Directions

Although this topic is thought-provoking as far as the disparities in the healthcare field and how the disparities can be addressed, there are some limitations that must also be considered. In this case, the use of the techniques of text mining could have created potential biases.

Though the conclusions drawn from the text mining process are fact-based and up-to-date observations, sometimes the picture presented does not represent the everyday experience of rural people. Inevitably, the use of other techniques for qualitative research appears.

Second, the heterogeneity existing at the state level can carry the potential to influence the effectiveness of health care interventions. The rural problems are geographically and politically very diverse at the regional level. For example, rural problems at the regional level are so diverse that it is very hard to draw a definitive conclusion about the whole region. This demands the use of region-specific methodologies in future studies.

Third, barriers in the use of digital health aspects such as language barriers, technophobia, and confidentiality issues hinder the widespread use of telemedicine and mHealth interventions. They can be addressed through culturally appropriate digital health educational interventions and improved cybersecurity competences to build confidence and ensure the best practices of ethical health delivery.

Future studies should examine the long-term outcome effects of digital health interventions, the effectiveness of impact policies, and the economic sustainability of health programs in rural areas of the Indian context. The most important areas under the topic that should be studied further are as follows:

Firstly, the impact of the availability of access and quality of care in rural India must be measured and studied. In future studies, the ability of telemedicine to bridge the gaps in health services and improve patient outcomes must also be studied. Secondly, a nationwide and rural community comparative analysis will highlight the health disparities between different states and suggest health policy interventions for the concerned region. Thirdly, the measures of financial protection must also be analyzed since studies must aim at determining the delivery efficiency of health insurance programs in containing out-of-pocket health expenses and improving access to health services among rural populations. In the fourth aspect, a review of the use of AI and analytics in rural health sectors has the potential to explain the efficiency of the use of artificial intelligence and analytics-based health planning in maximizing health surveillance, predictive health planning strategies, and health planning among the rural population. In the final aspect, more studies must happen to review the effectiveness of health model integration at the rural community levels towards improving health outcomes in rural Indian populations through rural health practitioners like the ASHA program among other self-help health programs. Efficiently addressing the above-mentioned areas, future studies should try to develop sustainable health systems and health programs that bridge the health gaps between the rural and urban health sectors in the Indian

states. Efficiently addressing the above-mentioned areas.

Acknowledgement

The authors would like to thank the Indian Council of Social Science Research (ICSSR), Delhi, for the assistance extended to the authors in completing the research project “Transforming Lives: Sustainable health solution for the most vulnerable section of the Indian population in the selected cities”. The authors would also like to thank the assistance extended to the authors by GLA University Mathura, Jamia Hamdard and Dr. Vishwanath Karad MIT World Peace University, Pune.

Funding

The project is funded by Indian Council of Social Science Research (ICSSR), Delhi.

Conflict of Interest

None declared.

Author Contributions

All the authors contributed to the study.

References

1. Balarajan, Y., Selvaraj, S., & Subramanian, S. V. (2011). Health care and equity in India. *The Lancet*, 377(9764), 505-515.
2. Selvaraj, S., Karan, A. K., Srivastava, S., Bhan, N., & Mukhopadhyay, I. (2022). India health system review.
3. Sharma, M. D., Mishra, P., Ali, A., Kumar, P., Kapil, P., Grover, R., ... & Kulshrestha, S. (2023). Microbial waterborne diseases in India: status, interventions, and future perspectives. *Current Microbiology*, 80(12), 400.
4. Tagat, A., & Kapoor, H. (2020). Go Corona Go! Cultural beliefs and social norms in India during COVID-19. *Journal of Behavioral Economics for Policy*, 4(S), 9-15.
5. Kornelsen, J., Lin, S., Williams, K., Skinner, T., & Ebert, S. (2023). System interventions to support rural access to maternity care: an analysis of the rural surgical obstetrical networks program. *BMC pregnancy and Childbirth*, 23(1), 621.
6. Ramesh, J., Mehta, A., & Jayanna, K. (2025). The Impact of Chronic Diseases on Cognitive Impairment in Rural Population of India: A Focus on Diabetes, Hypertension, Cardiovascular Disease, and Stroke. *Brain Behavior and Immunity Integrative*, 100107.
7. Theodosopoulos, L., Fradelos, E. C., Panagiotou, A., Dreliozi, A., & Tzavella, F. (2024). Delivering Culturally Competent Care to Migrants by Healthcare Personnel: A Crucial Aspect of Delivering Culturally Sensitive Care. *Social Sciences*, 13(10), 530.
8. Kumar, J. S., & Shobana, D. (2024). A study on assessment of rural health care system in India: schemes and implications. *Economics*, 12(2), 24-34.
9. Dasgupta, A., & Kapur, D. (2020). The political economy of bureaucratic overload: Evidence from rural development officials in India. *American Political Science Review*, 114(4), 1316-1334.
10. Heunis, C., & Kigozi-Male, G. (2024). Exploring Managers' Insights on Integrating Mental Health into Tuberculosis and HIV Care in the Free State Province, South Africa. *International Journal of Environmental Research and Public Health*, 21(11), 1528.
11. Banerjee, A., & Duflo, E. (2011). Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty. *PublicAffairs*.
12. Kathuria, K. (2023). Impact of maternal health and disease on neonatal outcome. In *Clinical Anesthesia for the Newborn and the Neonate* (pp. 11-27). Singapore: Springer Nature Singapore.
13. Baker, K., Adams, J., & Steel, A. (2022). Experiences, perceptions and expectations of health services amongst marginalized populations in urban Australia: a meta-ethnographic review of the literature. *Health Expectations*, 25(5), 2166-2187.
14. Kumari, S., Vani, V., Malik, S., Tyagi, A. K., & Reddy, S. (2021). Analysis of text mining tools in disease prediction. In *Hybrid Intelligent Systems: 20th International Conference on Hybrid Intelligent Systems (HIS 2020)*, December 14-16, 2020 (pp. 546-564). Springer International Publishing.
15. Hassani, H., Beneki, C., Unger, S., Mazinani, M. T., & Yeganegi, M. R. (2020). Text mining in big data analytics. *Big Data and Cognitive Computing*, 4(1), 1.
16. Hu, N. (2022). Sentiment analysis of texts on public health emergencies based on social media data mining. *Computational and mathematical methods in medicine*, 2022(1), 3964473.
17. Bereitschaft, B., & Scheller, D. (2020). How might the COVID-19 pandemic affect 21st century urban design, planning, and development?. *Urban science*, 4(4), 56.
18. Peprah, P., Abalo, E. M., Agyemang-Duah, W., Budu, H. I., Appiah-Brempong, E., Morgan, A. K., & Akwasi, A. G. (2020). Lessening barriers to healthcare in rural Ghana: providers and users'

- perspectives on the role of mHealth technology. A qualitative exploration. *BMC medical informatics and decision making*, 20, 1-12.
19. Paniagua, A. (2023). Geographies of marginality in Europe: space, people and politics. In *SDGs in the European Region* (pp. 1-21). Cham: Springer International Publishing.
 20. von Humboldt, S., Low, G., & Leal, I. (2022). Health service accessibility, mental health, and changes in behavior during the COVID-19 pandemic: a qualitative study of older adults. *International Journal of Environmental Research and Public Health*, 19(7), 4277.
 21. Prancutė, R. (2021). Web of Science (WoS) and Scopus: The titans of bibliographic information in today's academic world. *Publications*, 9(1), 12.
 22. Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3, 993–1022.
 23. Chauhan, U., & Shah, A. (2021). Topic modeling using latent Dirichlet allocation: A survey. *ACM Computing Surveys (CSUR)*, 54(7), 1-35.
 24. Kousis, A., & Tjortjis, C. (2023). Investigating the key aspects of a smart city through topic modeling and thematic analysis. *Future Internet*, 16(1), 3.
 25. Boon-Itt, S., & Skunkan, Y. (2020). Public perception of the COVID-19 pandemic on Twitter: sentiment analysis and topic modeling study. *JMIR public health and surveillance*, 6(4), e21978.
 26. Albalawi, R., Yeap, T. H., & Benyoucef, M. (2020). Using topic modeling methods for short-text data: A comparative analysis. *Frontiers in artificial intelligence*, 3, 42.
 27. Maslej-Krešňáková, V., Sarnovský, M., Butka, P., & Machová, K. (2020). Comparison of deep learning models and various text pre-processing techniques for the toxic comments classification. *Applied Sciences*, 10(23), 8631.
 28. Sarica, S., & Luo, J. (2021). Stopwords in technical language processing. *Plos one*, 16(8), e0254937.
 29. Abidin, Z., & Junaidi, A. (2024). Text Stemming and Lemmatization of Regional Languages in Indonesia: A Systematic Literature Review. *Journal of Information Systems Engineering and Business Intelligence*, 10(2), 217-231.
 30. Ogada, K., Mwangi, W., & Cheruiyot, W. (2015). N-gram based text categorization method for improved data mining. *Journal of Information Engineering and Applications*, 5(8), 35-43.
 31. Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., & Zhao, L. (2019). Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey. *Multimedia tools and applications*, 78, 15169-15211.
 32. Rahimi, M., Zahedi, M., & Mashayekhi, H. (2022). A probabilistic topic model based on short distance co-occurrences. *Expert Systems with Applications*, 193, 116518.
 33. Li, H., Khomh, F., & Openja, M. (2021, September). Understanding quantum software engineering challenges an empirical study on stack exchange forums and github issues. In *2021 IEEE International Conference on Software Maintenance and Evolution (ICSME)* (pp. 343-354). IEEE.
 34. Kelkar, S. (2021). India's Public Health Care Delivery. *India's Public Health Care Delivery; Indian Journal of Psychological Medicine*.
 35. Palozzi, G., Schettini, I., & Chirico, A. (2020). Enhancing the sustainable goal of access to healthcare: findings from a literature review on telemedicine employment in rural areas. *Sustainability*, 12(8), 3318.
 36. Bangal, V. B., Borawake, S. K., Gavhane, S. P., & Aher, K. H. (2017). Use of mobile phone for improvement in maternal health: a randomized control trial. *Int J Reprod Contracept Obstet Gynecol*, 6(12), 5458.
 37. Anisha, S. A., Sen, A., & Bain, C. (2024). Evaluating the potential and pitfalls of AI-powered conversational agents as humanlike virtual health carers in the remote management of noncommunicable diseases: scoping review. *Journal of Medical Internet Research*, 26, e56114.
 38. Asrani, C. (2022). Spanning the digital divide in India: Barriers to ICT adoption and usage. *Journal of Public Affairs*, 22(4), e2598.
 39. Singh, R., Raj, R. S., & Radha. (2025). Ensuring Health for All: Examining the Role of Digital Health in India. In *Intersecting Realities of Health Resilience and Governance in India: Emerging Domestic and Global Perspectives* (pp. 201-227). Singapore: Springer Nature Singapore.
 40. Sindakis, S., & Showkat, G. (2024). The digital revolution in India: bridging the gap in rural technology adoption. *Journal of Innovation and Entrepreneurship*, 13(1), 29.
 41. Bhandari, L., & Dutta, S. (2007). Health infrastructure in rural India. *India infrastructure report, 2007*, 265-85.
 42. Vijayakumar, K. N., Gunasekaran, K., Sahu, S. S., & Jambulingam, P. (2009). Knowledge, attitude and practice on malaria: a study in a tribal belt of Orissa state, India with reference to use of long lasting treated mosquito nets. *Acta tropica*, 112(2), 137-142.
 43. Bhargava, A., Bhargava, M., & Juneja, A. (2021). Social determinants of tuberculosis: context,

- framework, and the way forward to ending TB in India. *Expert Review of Respiratory Medicine*, 15(7), 867-883.
44. Paul, P. L., & Pandey, S. (2020). Factors influencing institutional delivery and the role of accredited social health activist (ASHA): a secondary analysis of India human development survey 2012. *BMC Pregnancy and Childbirth*, 20, 1-9.
 45. Kumar, R. (2024). Epidemiological Challenges in India. In *Handbook of Epidemiology* (pp. 1-27). New York, NY: Springer New York.
 46. Kashyap, A., Shrivastava, S., & Krishnatray, P. (2019). Vaccine hesitancy: the growing parent-provider divide. *Asia Pacific Media Educator*, 29(2), 259-278.
 47. Balarajan, Y., Selvaraj, S., & Subramanian, S. V. (2011). Health care and equity in India. *The Lancet*, 377(9764), 505-515.
 48. Nihal, S., & Shekhar, C. (2024). An assessment of adequate quality antenatal care and its determinants in India. *BMC Pregnancy and Childbirth*, 24(1), 698.
 49. Randive, B., San Sebastian, M., De Costa, A., & Lindholm, L. (2014). Inequalities in institutional delivery uptake and maternal mortality reduction in the context of cash incentive program, Janani Suraksha Yojana: results from nine states in India. *Social Science & Medicine*, 123, 1-6.
 50. Neethi Mohan, V., Vaidyanathan, G., Dash, U., & Muraleedharan, V. R. (2024). Wealth and education-related inequalities in the utilisation of reproductive, maternal, newborn, and child health interventions within scheduled tribes in India: an analysis of Odisha and Jharkhand. *BMC Public Health*, 24(1), 1-15.
 51. O'Neil, S., Naeve, K., & Ved, R. (2017). An examination of the maternal health quality of care landscape in India. *Mathematica policy research*, 2.
 52. Rahman, A., & Pingali, P. (2024). Public Health Insurance: Reducing Poverty or Access to Equitable Health Care?. In *The Future of India's Social Safety Nets: Focus, Form, and Scope* (pp. 203-244). Cham: Springer International Publishing.
 53. Saprii, L., Richards, E., Kokho, P., & Theobald, S. (2015). Community health workers in rural India: analysing the opportunities and challenges Accredited Social Health Activists (ASHAs) face in realising their multiple roles. *Human resources for health*, 13, 1-13.
 54. Dehury, R. K., Dehury, P., Sripathi, N., Acharyulu, G. V. R. K., Behera, M. R., & Neeragatti, S. (2023). Health sector development in India: an account from Bhole Committee 1946 to National Health Policy 2017. *Journal of Development Policy and Practice*, 8(2), 209-242.
 55. Sriram, S., & Khan, M. M. (2020). Effect of health insurance program for the poor on out-of-pocket inpatient care cost in India: evidence from a nationally representative cross-sectional survey. *BMC Health Services Research*, 20, 1-21.
 56. Sahoo, P. M., & Rout, H. S. (2023). Analysis of public health-care facilities in rural India. *Facilities*, 41(13/14), 910-926.
 57. Wahid, S. S., Munar, W., Das, S., Gupta, M., & Darmstadt, G. L. (2020). 'Our village is dependent on us. That's why we can't leave our work'. Characterizing mechanisms of motivation to perform among Accredited Social Health Activists (ASHA) in Bihar. *Health Policy and Planning*, 35(1), 58-66.
 58. Khan, A. M. (2020). Evaluation study of ASHA in providing access to health services and community awareness under national rural health mission (Doctoral dissertation, Tilak Maharashtra Vidyapeeth).
 59. Fletcher-Brown, J., Carter, D., Pereira, V., & Chandwani, R. (2021). Mobile technology to give a resource-based knowledge management advantage to community health nurses in an emerging economies context. *Journal of Knowledge Management*, 25(3), 525-544.
 60. Bhan, N., Madhira, P., Muralidharan, A., Kulkarni, B., Murthy, G. V. S., Basu, S., & Kinra, S. (2017). Health needs, access to healthcare, and perceptions of ageing in an urbanizing community in India: a qualitative study. *BMC geriatrics*, 17, 1-11.
 61. Kadarpetta, R. S. R., Anand, J. S., & Achungura, G. (2024). Strengthening Public Financing of Primary Healthcare in India: A Perspective. *Health Services Insights*, 17, 11786329241249289.
 62. Sherlin, S., & Karunanidhi, K. (2024). Community Health Workers in India: Bridging the Gap in Healthcare Access and Equity. In *Case Studies on Holistic Medical Interventions* (pp. 506-510). CRC Press.