Comparative analysis and machine learning predictions of cervical cancer incidence: A multi-national study

Shah Faisal 1, *, Basit Hussain 2, Munima Haque 1, Sania Zehra 1, Saliha Khalid 3 and Yumna Amjad 4

1 Biotechnology Program, Department of Mathematics and Natural Sciences, BRAC University, Dhaka, Bangladesh.
2 Department of Computer Science and Engineering, BRAC University, Dhaka, Bangladesh.
3 Department of Material Science and Nanotechnology, School of Engineering and Natural Sciences, Kadir Has University, Istanbul Turkey.
4 Department of Biochemistry, Faculty of Science and Technology, University of Central Punjab, Lahore, Pakistan.

World Journal of Biology Pharmacy and Health Sciences, 2024, 18(03), 093–104

Publication history: Received on 23 April 2024; revised on 02 June 2024; accepted on 05 June 2024

Article DOI: https://doi.org/10.30574/wjbphs.2024.18.3.0328

Abstract

This proposal presents a comprehensive investigation of cervical cancer frequency and patterns, with a focus on a multi-national perspective, particularly highlighting the circumstance in Bangladesh. Utilizing a mix of auxiliary information survey, comparative investigation, and predictive modeling, this study sheds light on the worldwide landscape of cervical cancer, emphasizing disparities in rate, screening hones, and healthcare framework. The inquire about utilizes machine learning calculations, especially linear regression, to extend future patterns of cervical cancer in Bangladesh up to 2050. Moreover, an in-depth examination of statistical, clinical, and treatment characteristics of 223 cervical cancer patients in Bangladesh offers basic bits of knowledge into components affecting results. Key discoveries uncover noteworthy fluctuations in treatment and discovery techniques over nations, underscoring the requirement for more harmonized worldwide healthcare approaches. The predictive analysis indicates a potential stabilization in cervical cancer cases in Bangladesh, suggesting a positive trend due to ongoing healthcare efforts. This proposition contributes to the existing body of information on cervical cancer, giving profitable bits of knowledge for healthcare arrangement definition and execution precisely in resource-limited settings.

Keywords: Cervical Cancer; Predictive Modeling; Machine Learning and Linear Regression; Global Health Disparities; Healthcare Infrastructure

1. Introduction

Cervical cancer continues to pose a significant global health challenge, especially impacting women in low and middle-income countries. Primarily caused by persistent infection with high-risk strains of the Human Papillomavirus (HPV), this disease remains preventable through effective screening and vaccination strategies. Despite these advancements, cervical cancer maintains high incidence and mortality rates worldwide; in 2018, it accounted for over 570,000 new cases and approximately 311,000 deaths globally [1, 2].

This research aims to conduct a detailed comparative analysis of cervical cancer incidence across multiple countries, with a specific focus on Bangladesh. The study seeks to identify global patterns, assess disparities, and apply machine learning techniques to forecast future trends in Bangladesh. These objectives are grounded in the significant need for an enhanced understanding of cervical cancer dynamics in settings that face unique healthcare challenges.
Extensive literature highlights notable disparities in cervical cancer incidence and outcomes. Globally, cervical cancer ranks as the fourth most commonly diagnosed cancer among women, with prevalence rates varying significantly across regions. Sub-Saharan Africa, Latin America, and parts of Asia experience higher incidence rates, largely due to gaps in screening and vaccination coverage [4]. In contrast, developed regions such as North America and Europe exhibit lower rates and better outcomes, thanks to their robust healthcare infrastructures and accessible preventive services [3, 5].

Research further shows that socioeconomic factors, such as access to healthcare and education, significantly affect these regional disparities. Studies indicate that early detection and education about HPV and cervical cancer can significantly improve outcomes [6, 7]. Additionally, advancements in diagnostic techniques and treatment strategies have been pivotal in managing cervical cancer effectively, though these advancements are not uniformly available across all regions [8].

The significance of this study lies in its potential to impact public health policy through the application of machine learning to predict disease trends, which could inform better resource allocation and preventive strategies tailored to the specific needs of the Bangladeshi population. By addressing the gap in comprehensive data analysis and predictive modeling for cervical cancer in Bangladesh, this research aspires to facilitate significant advancements in the management and prevention of this disease in regions similarly characterized by limited healthcare resources.

2. Materials and Methods

2.1. Study Purpose and Research Questions

The primary aim of this study is to conduct a comparative analysis of cervical cancer incidence and mortality, with a special focus on Bangladesh. This investigation seeks to identify global patterns, examine disparities, and employ machine-learning techniques to forecast future trends of cervical cancer incidence in Bangladesh. The guiding research questions include:

- What are the global trends in cervical cancer, as illuminated by the selected data points?
- How does the cervical cancer scenario in Bangladesh align or differ from other countries under study?
- Can machine-learning tools effectively forecast future cervical cancer trends in Bangladesh?

2.2. Data Collection

Data for this study was meticulously compiled from a variety of secondary sources, including peer-reviewed research studies and official health statistics from relevant countries, focusing on parameters such as cancer treatments, survival rates, detection methods, and demographic details. Countries included in the study span a diverse geographical and economic spectrum, providing a comprehensive global perspective.

2.3. Data Analysis

2.3.1. Secondary Data Analysis

**Tools and Techniques:** Utilization of Python for data manipulation and analysis, employing libraries such as Pandas for data manipulation and Matplotlib for creating insightful visualizations.

**Process**

- **Data Extraction and Cleaning:** Extraction of relevant data organized for detailed analysis and cleaned to ensure integrity for subsequent analytical stages.
- **Data Visualization:** Visualization of data to facilitate direct comparison of cancer sizes, therapy durations, screening frequencies, and other relevant metrics across the countries studied.

2.3.2. Cervical Cancer Mortality Analysis

**Objective:** Analyze cervical cancer mortality data to examine death rates and identify trends across different regions.

- **Data Source:** The dataset includes longitudinal records from Kaggle, documenting cervical cancer mortality from 1990 to the latest available year.
- **Analysis:** Using Python, the data was prepared, analyzed, and visualized to depict trends and provide a comparative understanding of mortality rates.
2.3.3. Predictive Modeling Using Machine Learning

**Objective:** Employ a Linear Regression model to predict future trends in cervical cancer incidence up to the year 2050.

- **Data Handling:** Historical data from official health databases provided the foundation for the model, which was trained to project future incidence based on past patterns.
- **Model Assessment:** The model’s predictive accuracy was evaluated using historical data to ensure its reliability for healthcare planning.

2.3.4. Demographic, Clinical, and Treatment Characteristics Analysis

- **Data Source:** Analysis of 223 patient records from Bangladesh to explore demographic trends, clinical presentations, and treatment outcomes.
- **Analysis Process:** Employing Python-based statistical methods to identify correlations and significant factors affecting prognosis, with results visualized to aid interpretation.

3. Results

This section presents a detailed analysis of the multifaceted aspects of cervical cancer across various demographics and geographies, highlighting our key findings. We intricately unravel the threads of data, encompassing historical trends, predictive models, and demographic insights related to cervical cancer. The focus is on analyzing past and future trends in Bangladesh, comparing mortality rates across different countries, discussing the nuanced interplay between demographic factors, medical history, and screening practices. These findings provide a comprehensive picture of the current landscape and offer insights into the potential trajectory of cervical cancer incidence and management. The synthesis of results sheds light on the complex dynamics of cervical cancer, enhancing our understanding of its impact across diverse populations.

3.1. Data Set 1 Findings:

In this significant chapter, we explore through the comes about determined from a comprehensive investigation of cervical cancer treatment and discovery over numerous countries. For perusers who may have specifically gotten to this section, it is pivotal to note that the information typified here stems from a broad comparative consider laid out in past chapters. The dataset supporting these discoveries is an amalgamation of data from Pakistan, India, Nepal, Bhutan, Afghanistan, Sri Lanka, USA, and Europe with a highlight on survival rates, unusual results and screening hones, as they relate to the healthcare frameworks inside these nations.

3.1.1. Tumor Size and Treatment Duration

Our comparative examination has uncovered a stark variety in tumor estimate at determination over the nations, with the littlest normal estimate being 2 cm in Afghanistan and the biggest 12 cm in Europe. This fluctuation is reflected within the term of treatment, which ranges from 4 weeks in Afghanistan to 14 weeks in Europe, setting a coordinate relationship between tumor estimate at determination and the length of treatment required.

3.1.2. Screening Frequency

The recurrence of screening too displayed noteworthy aberrations, uncovering varying national rules and healthcare arrangements. In Bhutan, yearly screenings are prescribed, differentiating strongly with the USA, where the interim between screenings amplifies to 7 a long time. This dissimilarity may reflect contrasting national needs, asset accessibility, and population wellbeing methodologies.

3.1.3. Study Participation and Survival Rates

The scale of consider support shown a tremendous extend, with Europe displaying a vigorous dataset of 5,000 members, overshadowing Bhutan’s 200. Survival rates show a similarly changed scene, with the least rate of 29.6% in Afghanistan differentiating with a middle survival time of 120 months in India. Such inconsistencies emphasize the different viability of treatment conventions and get to healthcare administrations.

3.1.4. Incidence of Abnormal Results

Announcing of irregular comes about advance complements territorial contrasts in discovery and follow-up hones. Bhutan detailed the most elevated recurrence with 3 cases, whereas India, Afghanistan, and the USA each detailed zero cases, proposing varieties in screening precision or the predominance of cervical cancer in these populations.
3.1.5. **Visual Representation**

The discoveries from the consider are outwardly encapsulated in an arrangement of charts: The primary chart depicts the comparison of survival rates and the number of unusual results by nation, highlighting the relationship between healthcare practices and quiet results.

![Comparison of Survival Rates and Number of Abnormal Results by Country](image)

**Figure 1** Comparison of Survival Rates and Number of Abnormal Results by Country

Subsequent graphs present detailed comparisons of tumor size, treatment duration, screening frequency, and the number of study participants by country, offering a visual narrative of the heterogeneity in cancer care.

![Tumour Size, Treatment Duration, Screening Frequency and No of Study participants by country](image)

**Figure 2** Tumour Size, Treatment Duration, Screening Frequency and No of Study participants by country
3.1.6. Conclusion of Findings

The heterogeneity enlightened by these discoveries underscores the require for more bound together worldwide measures in cancer care, including treatment conventions and location strategies. The application of Python for information examination has illustrated its viability in overseeing broad datasets and creating quick visual comparisons that advise this study’s conclusions.

3.2. Findings of Data Set 2 - Cervical Cancer Death Rate Analysis

3.2.1. Results Interpretation

The charts given uncover compelling patterns and shifts in cervical cancer mortality rates over the countries beneath ponder. They serve as a visual declaration to the multifaceted effect of healthcare frameworks, socio-economic conditions, open wellbeing activities, and social impacts on the flay of cervical cancer.

3.2.2. Pakistan

The cervical cancer passing rate in Pakistan shows a perceivable direction that talks volumes approximately the country’s fight against the malady. The chart portrays a relentless climb, recommending a pressing call for improved healthcare mediations, supported mindfulness, and more vigorous screening programs.

3.2.3. India

India’s chart paints a striking picture of the cervical cancer challenge inside a crowded scene. The slant lines demonstrate an energetic that may be influenced by the reach of healthcare, the vigor of open wellbeing activities, and the socio-economic texture of the country.

3.2.4. Bangladesh

The design of cervical cancer passing in Bangladesh gives a window into the viability of the healthcare framework and the power of open wellbeing methodologies. The graph’s drift signals the repercussions of healthcare availability and the entrance of instructive and preventive measures.

3.2.5. Nepal

Nepal’s account, as told by the climbing and plummeting curves on the chart, reflects the nation's healthcare capacity to address cervical cancer. The forms of the chart seem well reflect the foundational quality of healthcare framework and the vital open wellbeing arrangements input.

3.2.6. Sri Lanka

In Sri Lanka, the graph’s trajectory offers a visual dialog about the nation’s policy effectiveness, preventive measures, and the level of awareness surrounding cervical cancer.

3.2.7. Bhutan

Bhutan’s depiction through its chart gives bits of knowledge into the reach of healthcare administrations, the mindfulness of the illness, and the accessibility of screening and treatment Choices inside the nation.

3.2.8. Afghanistan

Afghanistan’s chart, with its interesting design, likely echoes the country's socio-political complexities, healthcare system, and open wellbeing techniques, portray a picture of the challenges and endeavors in combating cervical cancer.

3.2.9. United States (USA)

The chart for the USA exhibits a develop healthcare framework hooking with cervical cancer. It reflects on the quality and viability of screening programs, mindfulness levels, and the arsenal of treatment choices at the populace's transfer
3.2.10. Europe

The comprehensive chart speaking to Europe typifies the collective adequacy of shifted healthcare frameworks. It exhibits how distinctive approaches, preventive measures, and benefit availabilities play out over a cluster of nations.

![Cervical Cancer Death Rate in Selected Countries](image)

**Figure 3** Cervical Cancer Death rate in Bangladesh, Pakistan, India, Afghanistan, Nepal, Sri Lanka, Bhutan, USA and Europe

3.2.11. Comparative Analysis

The collective chart brings into sharp alleviation the contrasts and parallels in cervical cancer mortality over the studied nations and Europe. It underscores how dissimilar healthcare frameworks, financial conditions, and open wellbeing approaches impact the victory or disappointment of combating cervical cancer.

These discoveries, verbalized through data-driven accounts and visual pieces, shape the premise for the resulting chapter, which can dismember these designs and patterns to reveal the suggestions for future healthcare procedures, policy-making, and inquire about roads.

3.3. Findings of Data Set 3 - Predictive Analysis of Cervical Cancer in Bangladesh

Here the findings are committed to revealing the direction of cervical cancer cases in Bangladesh, based on an exhaustive examination of chronicled information and a forward-looking prescient show. This examination is basic for understanding the movement of cervical cancer and is instrumental in forming future healthcare approaches. For perusers who may straightforwardly explore to this segment, this chapter is fastidiously created to stand alone, offering a clear, comprehensive view of the study’s findings from past to future.

3.3.1. Historical Trends in Cervical Cancer (1990-2023)

The verifiable investigation set out on a travel through information from 1990 to 2023. The fastidiously sourced information uncovered a slant of cervical cancer cases over the decades, graphically delineated to supply a straightforward outline of the disease’s predominance over time. The chronicled slant shows an Increment advertising important bits of knowledge into the effect of past healthcare activities and societal changes on the frequency of cervical cancer.
3.3.2. Predictive Analysis for Future Trends (2024-2050)

Progressing from the review analysis, the focus shifted to anticipating the frequency of cervical cancer from 2024 to 2050. A direct relapse demonstrate was adeptly chosen for its demonstrated capability to explain long-term patterns. This factual demonstrative treated the year as the indicator and the number of cervical cancer cases as the result, coming about in a projection line that expands the chronicled drift into long-standing time. The prescient analysis, visualized nearby the real information, figures a proceeded marginally increment offering a foresight that's crucial for key healthcare arranging.

3.3.3. Summary of Findings

The discoveries from Information Set 3 amalgamate the observational prove from the past with a calculated look into long term, giving a story that ranges over 60 a long time. The information talks to both the flexibility and the challenges confronted by the healthcare framework in Bangladesh, portray a picture of the advancing scene of cervical cancer. The direct relapse model's expectations serve as a herald, possibly directing policymakers and healthcare suppliers in their journey to moderate the effect of this illness. The by and large discoveries propose a summarize key discoveries, e.g.,
require for enhanced screening programs, progressed treatment strategies, or expanded open wellbeing mindfulness, highlighting the basic for proactive measures within the continuous battle against cervical cancer in Bangladesh.

This expository endeavor, mixing authentic information with prescient modeling, has shed light on the designs of cervical cancer in Bangladesh, building up an establishment for educated decision-making and vigorous healthcare techniques within the chapters to come.

3.4. Findings of Data Set 4 - Demographic and Clinical Characteristics of Cervical Cancer Patients in Bangladesh

Data set 4 gives an explanatory dissection of the information collected from 223 patients over different clinics in Bangladesh, advertising basic experiences into the statistic and clinical characteristics related with cervical cancer. This examination is urgent for perusers looking for to comprehend the scope and subtleties of cervical cancer cases and reactions to screening inside the Bangladeshi setting. The discoveries are laid out deliberately, supplemented by visual helps to upgrade understanding, especially centering on how these components relate with the statistic variable of age.

3.4.1. Medical History and Screening Tests

The information uncovered designs within the restorative history of respondents, demonstrating a generally tall event of cervical cancer analyze and medications for precancerous conditions inside certain age bunches. The overview reactions too recommended that cervical cancer screening, such as Pap tests or HPV tests, is predominant, reflecting a degree of mindfulness and proactive wellbeing administration among the population.

3.4.2. Lifestyle Factors and Risk Behaviors

The investigation of way of life components highlighted the nearness of chance behaviors, such as tobacco utilize, which appeared a relationship with cervical wellbeing results, fortifying the require for focused on wellbeing instruction. Moreover, the consider inspected the suggestions of verbal prophylactic utilize and the history of sexually transmitted infections (STIs), shedding light on their potential affiliation with cervical cancer dangers.

3.4.3. Demographic Comparisons

The study reactions changed over age socioeconomics, disclosing critical bits of knowledge. For case, more seasoned age bunches appeared a better probability of having experienced cervical cancer screening tests and medicines for precancerous conditions. This change underscores the impact of age on healthcare behaviors and encounters.

3.4.4. Statistical Significance

A few correlations between the collected reactions were measurably noteworthy, showing basic designs that justify encourage investigate. These discoveries point to potential ranges for healthcare interventions and public wellbeing procedures.

3.4.5. Graphical Representation of Data

The survey data was aptly translated into a series of bar charts, providing a clear visual representation of the variations in responses across different age categories. These charts not only distilled complex data into digestible visuals but also highlighted the disparities and commonalities in the experiences of the respondents.

Figure 6 Medical History and Screening Tests by Age of patients
Critical analysis and medical history questionnaires from patients depend on following aspects:

- Do you currently smoke tobacco?
- Have you ever used oral contraceptives?
- Have you ever been diagnosed with any sexually transmitted infection (STI)?

Implications of Findings:
- The findings from Data Set 4 carry significant implications for public health in Bangladesh:
  - They emphasize the necessity for cervical health education and screening programs tailored to specific demographic segments.
  - The associations found between lifestyle factors and cervical cancer underscore the need for risk factor modification and prevention strategies.

The demographic differences revealed through the study could inform targeted interventions to improve health outcomes and enhance the effectiveness of cervical cancer prevention and treatment programs.

In summary, the analysis of Data Set 4 has painted a detailed portrait of cervical cancer demographics and clinical characteristics in Bangladesh, providing invaluable insights that could drive future healthcare policies and interventions aimed at reducing the burden of this disease.
The findings detailed in this section provide a robust response to the research questions outlined at the onset of this study. We have explored the historical and predictive landscapes of cervical cancer in Bangladesh, alongside comparative mortality analyses across different countries. Additionally, we have examined demographic and clinical characteristics of patients within the Bangladeshi context. Through detailed quantitative and qualitative analysis, this study has identified significant patterns, trends, and associations that enhance our understanding of the prevalence, risk factors, and effectiveness of screening and prevention strategies for cervical cancer. These insights contribute profoundly to the global discourse on cervical cancer management and set the stage for focused interventions aimed at mitigating its impact.

4. Discussion

This study provides a comprehensive analysis of cervical cancer across various demographics and geographies, with a special focus on Bangladesh. By leveraging extensive datasets and advanced machine learning techniques, our findings offer new insights into the global and regional dynamics of cervical cancer, with implications for public health policy and practice.

4.1. Global Trends and Regional Disparities:

Our findings confirm the significant global burden of cervical cancer, particularly in regions with limited healthcare resources. The comparative analysis of mortality rates across different countries reveals a pronounced disparity, aligning with previous reports by WHO and other studies that underscore the critical role of robust public health infrastructures in managing this disease [1]. This disparity highlights the persistent global health inequity, where developed countries benefit from advanced screening and vaccination programs, while developing countries, like Bangladesh, struggle with basic healthcare needs.

4.2. Predictive Insights and Healthcare Implications

The application of machine learning to predict future trends in cervical cancer incidence in Bangladesh represents a novel aspect of this research. The model suggests a potential stabilization in the incidence rates, offering a hopeful outlook for the future. This contrasts with the static or worsening projections seen in earlier studies and suggests that recent healthcare initiatives might be starting to have an impact. Such predictive insights are invaluable for healthcare planning and resource allocation, emphasizing the need for continued investment in health infrastructure and public health strategies.

4.3. Socio-Demographic and Clinical Insights

The detailed examination of demographic, clinical, and treatment characteristics of cervical cancer patients in Bangladesh has provided critical insights. Our analysis reveals that early detection and tailored healthcare interventions are crucial. The demographic data underscore the importance of targeted screening programs that address specific community needs, particularly in areas with high incidence rates. These findings resonate with the literature emphasizing the significance of personalized healthcare approaches to improve outcomes in cancer care [3].

4.4. Echoing and Expanding Existing Research

Our study not only echoes the findings of previous research regarding the importance of screening and early detection but also expands on these themes by integrating machine learning to forecast how these factors might change in the future. The alignment of our findings with existing research confirms the robustness of our methodologies and the relevance of our results to ongoing global efforts to combat cervical cancer.

4.5. Implications for Public Health Policy

The insights gained from this research have significant implications for public health policy, particularly in resource-constrained settings like Bangladesh. The evidence supports the need for policies that prioritize the expansion of HPV vaccination programs and the accessibility of regular screening. Public health campaigns should also focus on education and awareness to combat the stigma associated with cervical cancer and to promote early detection strategies.

5. Conclusion

This study has provided comprehensive insights into the dynamics of cervical cancer across various global contexts, highlighting significant disparities and the potential impact of targeted interventions. Through detailed analysis and
predictive modeling, we have illuminated the stark realities and disparities that characterize cervical cancer management worldwide.

## 5.1. Key Findings

### 5.1.1. Global Trends and Comparisons

The research underscores significant disparities in cervical cancer incidence and mortality rates. Developed countries exhibit lower rates, attributable to structured screening programs and widespread HPV vaccination. In contrast, developing nations, including Bangladesh, experience higher rates, which reflect the critical impact of limited healthcare accessibility and preventive infrastructure.

### 5.1.2. Predictive Analysis and Future Trends

Our application of machine learning techniques to predict future trends in cervical cancer incidence in Bangladesh suggests a potential stabilization of rates. This optimistic outlook indicates that ongoing improvements in healthcare access and awareness may start to yield positive outcomes in the foreseeable future.

### 5.1.3. Demographic, Clinical, and Treatment Insights

Analysis of patient-level data from Bangladesh has shed light on demographic and clinical factors crucial to cervical cancer outcomes, such as age and stage at diagnosis. These findings emphasize the importance of early detection and timely intervention.

## 5.2. Implications and Forward Path

The conclusions drawn from this research not only highlight the current state of cervical cancer but also validate the efficacy of preventive measures and the transformative potential of predictive analytics in healthcare. The insights provided by this study advocate for a continued focus on enhancing healthcare infrastructure, particularly in low-resource settings, to make significant strides in combating cervical cancer.

Future efforts should prioritize the expansion of HPV vaccination programs, the accessibility of regular screening, and public health education to address and reduce the burden of cervical cancer globally. Moreover, the promise of data science and predictive modeling as potent tools in the fight against this disease underscores the need for sustained collaborative efforts. These findings serve as a guiding light for future research and actions aimed at transforming cervical cancer from a common threat to a preventable and manageable condition.

Through collaborative efforts, ongoing innovation, and a committed approach to improving women’s health, we can aspire to significantly mitigate the impact of cervical cancer across the world.

## Compliance with ethical standards

**Disclosure of conflict of interest**

No conflict of interest to be disclosed.

## References


