

## **Gaussian survival regression analysis of breast cancer mortality burden by deaths and years of life lost (YLLs): evidence from the global burden of disease database.**

P. Srivyshnavi<sup>1</sup>, M Darshan Teja<sup>2\*</sup>, P Shankaraiah<sup>3</sup>, Sreenivasulu T<sup>4</sup>

<sup>1</sup>Assistant Professor, Dept. of CS & E, S.P.M.V.V. Engineering College, Tirupati, India.

<sup>2, 3, 4</sup> Department of Mathematics, School of Advance Science, VIT Vellore, India.

Corresponding author: [darshanteja49@gmail.com](mailto:darshanteja49@gmail.com) \*

### **Abstract**

Breast cancer remains one of the leading causes of cancer death in women worldwide and continues to be a significant public health burden. Patterns of breast cancer mortality and premature death are important for the implementation of prevention and intervention strategies. This study utilizes two important indicators from the Global Burden of Disease (GBD) database, namely Deaths and Years of Life Lost (YLLs) to study breast cancer mortality burden. The analysis included demographic and epidemiologic variables, such as year, risk factor, location, age group, and sex. Major risk factors considered in the study included smoking, tobacco use, exposure to secondhand smoke, alcohol consumption, dietary risks, metabolic risks and high body-mass index. We modeled the association of these factors with breast cancer mortality burden using Gaussian Survival Regression framework. The results showed that the model was able to capture uncertainty related to the estimation of mortality burden. The prediction uncertainty was estimated as about 297.41 for Deaths and 261.79 for YLLs. The findings suggest variation in mortality burden across demographic groups and geographic regions. It is also indicated that behavioral and metabolic risk factors play a significant role in premature mortality associated with breast cancer. The study indicates the usefulness of Gaussian Survival Regression as a statistical framework for the analysis of population level cancer burden data. The findings could assist public health researchers and policymakers to identify high-risk populations and develop targeted strategies to reduce breast cancer mortality and premature loss of life.

**Keywords:** Breast Cancer, Deaths, Years of Life Lost, YLLs, Gaussian Survival Regression, Global Burden of Disease, Risk Factors, Mortality Analysis.

### **1. Introduction**

Breast cancer is one of the most common cancers in women worldwide and remains a major public health problem despite significant advances in diagnosis and treatment. Breast cancer is one of the most common causes of cancer mortality and a major component of the global cancer burden according to global cancer statistics. The increasing burden of breast cancer, particularly in low-

and middle-income countries, underscores the significance of routine surveillance and effective prevention measures.

The impact of breast cancer is not only a matter of mortality but also has important social, economic and healthcare implications. The Global Burden of Disease (GBD) study provides a wide range of measures of burden of disease to quantify the impact of diseases in the population. The mortality burden is usually measured by metrics such as death and Years of Life Lost (YLLs). Deaths are the total number of people who die from a disease. YLLs measure premature mortality by estimating the number of years lost because of death occurring before the expected life expectancy.

Demographic, environmental, behavioral and metabolic risk factors affect breast cancer mortality. Previous research has identified tobacco use, smoking, secondhand smoke exposure, alcohol use, dietary risks and high body mass index as major contributors to cancer burden. The prevalence and impact of these risk factors differ across countries, age groups and populations, resulting in considerable geographical differences in mortality trends.

The availability of large scale epidemiological databases has opened up opportunities to apply advanced statistical approaches for disease burden analysis. Statistical modeling techniques can aid researchers in understanding the complex relationships between demographic factors, risk factors and health outcomes. Methods of survival analysis have become important tools in medical and public health research, as they offer flexible frameworks for modeling health-related outcomes and for estimating uncertainty.

Within survival-based approaches, Gaussian Survival Regression provides a useful framework for analyzing continuous health outcomes and examining relationships between explanatory variables and disease burden indicators. This approach allows for the inclusion of multiple demographic and risk-factor variables, as well as quantification of the uncertainty in prediction.

GBD data have been used in several papers to study global breast cancer incidence, prevalence, mortality and risk factors. Most studies have concentrated, however, on descriptive epidemiological analyses and trend assessments. It is worth noting that there is a little research that has applied Gaussian Survival Regression models to analyze breast cancer mortality burden using Deaths and YLLs indicators.

Therefore, the aim of this study is to explore the breast cancer mortality burden through Deaths and Years of Life Lost (YLLs) from the Global Burden of Disease database. We assess the effect of demographic and risk-factor variables on mortality burden using a Gaussian Survival Regression model and quantify prediction uncertainty. The results are expected to help understand the patterns of breast cancer mortality and to inform evidence-based public health decision-making.

## 2. Literature Review



Breast cancer is one of the most frequently diagnosed cancers in women and the leading cause of cancer death worldwide. Its incidence, mortality, risk factors and disease burden have been studied in many research works using epidemiological and statistical methods. The Global Burden of Disease (GBD) project has generated important information on the global distribution and temporal trends of breast cancer burden in terms of deaths, Years of Life Lost (YLLs), Years Lived with Disability (YLDs) and Disability-Adjusted Life Years (DALYs).

Ferlay et al. (2015) studied the global patterns of cancer incidence and mortality and found that breast cancer represented a large proportion of deaths from cancer in women. The study highlighted significant geographic disparities in disease burden and the necessity of tailoring prevention strategies to specific regions [1]. Fitzmaurice et al. (2015) reviewed the global burden of cancer and reported increasing breast cancer incidence and mortality in some developing countries. Their findings showed the increasing public health importance of breast cancer and the need for effective healthcare interventions [2]. Global Burden of Disease Cancer Collaboration (2016) estimated cancer incidence, mortality, YLLs, YLDs, and DALYs for several cancers. In women, breast cancer was a leading contributor to the global burden of cancer, especially in middle-income and high-income regions [3]. Allemani et al. (2017) conducted a large-scale international study to assess cancer survival in different countries. Their results demonstrated great improvements in survival from breast cancer in developed countries through advances in screening and treatment programs [4]. Fitzmaurice et al. (2017) estimated the cancer burden globally, regionally and nationally from 1990 to 2015. Breast cancer was reported as one of the leading causes of cancer deaths and premature death worldwide as stated in the study [5]. Bray et al. (2018) reported the updated global cancer statistics and pointed out that the incidence of breast cancer continued to increase worldwide. The authors explained these trends by the aging of the population, changes in lifestyle and the improvement of diagnostic practice [6]. GBD 2017 Risk Factor Collaborators. 2018. "The Contribution of Behavioral, Environmental and Metabolic Risk Factors to Disease Burden." The study found that smoking, alcohol consumption, dietary risks and high body-mass index were all significant risk factors for cancer-related mortality [7]. Safiri et al. (2019) analyzed global breast cancer burden using GBD data and found large differences in mortality and YLLs between countries. The authors underscored the importance of socioeconomic development and healthcare access for disease outcomes [8]. Arnold et al. (2019) assessed future projections of breast cancer burden and predicted a significant increase in breast cancer incidence and mortality over the next decades. The study emphasized the relevance of preventive public health initiatives [9]. GBD 2019 Diseases and Injuries Collaborators (2020). Global burden of disease study 2019: a multi-country study of the global burden of disease. Breast cancer remained a leading cause of global death and YLLs, particularly among women older than 50 years [10]. GBD 2019 Risk Factors Collaborators (2020) studied the effects of modifiable risk factors on health outcomes. Their research suggested that smoking cessation, reducing alcohol consumption and obesity would significantly lower cancer-related mortality [11]. In 2020, breast cancer was the

most frequently diagnosed cancer globally, according to global cancer statistics presented by Sung et al. (2021). The study emphasized the growing burden of breast cancer and the persistent importance of strategies to decrease mortality [12]. Lei et al. (2021) studied breast cancer incidence and mortality trends and found an increasing burden of disease in younger populations. The changing patterns, they suggest, may be driven by behavioral and metabolic risk factors [13]. Morgan et al. (2021) reviewed future breast cancer burden projections and highlighted the need to improve early detection programmes and healthcare infrastructure to reduce mortality and YLLs [14]. Besides epidemiological studies, some studies have shown the utility of survival analysis methods in medical research. Kleinbaum and Klein (2012) provide extensive methods for analyzing survival data and discuss the application of the methods in the health sciences. Likewise, Hosmer, Lemeshow, and May (2008) outlined regression-based survival models that have been used extensively in epidemiological studies and health care [15].

Previous studies have extensively explored breast cancer incidence, mortality and disease burden with descriptive and trend-based analyses, but relatively little attention has been paid to the use of Gaussian Survival Regression models to analyze breast cancer mortality burden via Deaths and Years of Life Lost (YLLs). Hence, the current study aims to fill this gap by employing a Gaussian Survival Regression framework to evaluate the impact of demographic and risk-factor variables on the burden of breast cancer mortality using the Global Burden of Disease data.

## **Research Gap**

Most of the previous work has focused on descriptive epidemiology, estimation of the burden and trends, and future projections of incidence and mortality of breast cancer. Studies on the application of Gaussian Survival Regression techniques to model breast cancer mortality indicators such as Deaths and YLLs are comparatively few. Moreover, few studies have included demographic and behavioral risk factors in a common survival-regression model. Therefore, this study aims to fill this gap in knowledge by using Gaussian Survival Regression to evaluate the burden of breast cancer mortality using data from GBD.

## **3. Methodology**

### **3.1 Research Framework**

This study used a quantitative research design to explore breast cancer mortality burden using data from the Global Burden of Disease (GBD). Two mortality related indicators were included in the analysis:

- Deaths
- Years of Life Lost (YLLs)

A Gaussian Survival Regression model was used to investigate the association between demographic characteristics, risk factors and burden of breast cancer mortality. The overall workflow consisted of:

1. Data Collection
2. Data Preprocessing
3. Feature Selection
4. Gaussian Survival Regression Modeling
5. Prediction and Uncertainty Estimation
6. Model Evaluation
7. Results Interpretation

### 3.2 Data Source

The dataset was obtained from the Global Burden of Disease (GBD) database, which provides comprehensive information about disease burden across countries and populations.

The study looked at information about the burden of breast cancer across various geographic regions and demographic groups.

### 3.3 Study Variables

#### Response Variables

Two outcome variables were analyzed separately:

#### Deaths

Deaths represent the total number of breast cancer-related fatalities reported within a population.

#### Years of Life Lost (YLLs)

YLLs measure premature mortality and are calculated as the difference between expected life expectancy and age at death.

#### Predictor Variables

The following explanatory variables were included:

Table1: data set description

Variable	Description
Year	Observation year
Risk Factor	Major breast cancer risk factors
Location	Country or region

Age Group	Population age category
Sex	Male or Female

### **Risk Factors**

The study considered the following risk factors:

- Smoking
- Tobacco Use
- Secondhand Smoke
- Alcohol Use
- Dietary Risks
- Metabolic Risks
- High Body-Mass Index

### **3.4 Data Preprocessing**

Several preprocessing procedures were performed before model development.

#### **Data Cleaning**

The dataset was examined for:

- Missing values
- Duplicate observations
- Inconsistent records

Any inconsistencies were corrected prior to analysis.

#### **Encoding of Categorical Variables**

Categorical variables such as:

- Risk Factor
- Location
- Age Group
- Sex

were transformed into numerical representations suitable for regression analysis.

#### **Data Standardization**

Numerical variables were standardized where necessary to improve model stability and convergence.

### 3.5 Training and Testing Dataset

The dataset was divided into:

- Training Set (80%)
- Testing Set (20%)

The training dataset was used to estimate model parameters, while the testing dataset was used to evaluate predictive performance.

### 3.6 Gaussian Survival Regression Model

A Gaussian Survival Regression model was employed to model breast cancer mortality burden.

Let:

$Y_i$  represent the mortality burden indicator (Deaths or YLLs) for the ( $i^{\text{th}}$ ) observation.

The model is expressed as:

$$Y_i = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p + \varepsilon_i$$

where:

- $Y_i$  = Response variable
- $X_1$  = Predictor variables
- $\beta$  = Regression coefficients
- $\varepsilon_i$  = Random error

### 3.8 Prediction Process

After model fitting, new observations were generated using combinations of:

- Years
- Risk Factors
- Locations
- Age Groups
- Sex Categories

The trained Gaussian model was then used to estimate:

- Predicted Mortality Burden
- Prediction Uncertainty

for each observation.

### 3.9 Model Evaluation Metrics

Model performance was assessed using standard statistical measures.

#### Root Mean Square Error (RMSE)

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

#### Mean Absolute Error (MAE)

$$MAE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$

#### Coefficient of Determination (R<sup>2</sup>)

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}$$

where:

- $y_i$  = Observed value
- $\hat{y}_i$  = Predicted value
- $\bar{y}$  = Mean observed value

Higher R<sup>2</sup> values and lower RMSE and MAE values indicate better model performance.

### 3.10 Statistical Software

The analysis was conducted using Python programming language.

The following libraries were utilized:

- NumPy
- Pandas
- Scikit-learn
- Matplotlib
- Seaborn

These tools were used for data preprocessing, model development, prediction, visualization, and performance evaluation.

## 5. Results and Discussion

### 5.1 Overview of Breast Cancer Mortality Burden

In this study, the burden of breast cancer mortality was analyzed with two important GBD indicators, Deaths and Years of Life Lost (YLLs). To investigate the effect of demographic characteristics and important risk factors on breast cancer burden, a Gaussian Survival Regression model was used.

The analysis included several countries, age groups, sexes and categories of risk factor. Risk factors assessed included smoking, tobacco use, secondhand smoke, alcohol use, dietary risks, metabolic risks and high body-mass index. These factors have been previously identified as important factors for breast cancer mortality and premature death in epidemiological studies.

### 5.2 Deaths Prediction Analysis

The Gaussian Survival Regression model was applied to breast cancer mortality data represented by Deaths indicator. Predictor variables included in the model were year, risk factor, location, age group and sex. The prediction uncertainty obtained from the model was approximately:

Table 2. Sample Predictions for Breast Cancer Deaths Using Gaussian Survival Regression

Years	Metric	Risk Factor	Location	Age Group	Sex	Predicted Value	Prediction Uncertainty
2020	0.931215	Alcohol use	India	50–74 years	Male	0.0	297.41416
2022	0.435774	Dietary risks	Canada	50–74 years	Female	0.0	297.41416
2022	0.409839	Tobacco	Bangladesh	50–74 years	Female	0.0	297.41416
2021	0.942263	High body-mass index	New Zealand	50–74 years	Male	0.0	297.41416
2021	0.728419	Secondhand smoke	Bangladesh	15–49 years	Male	0.0	297.41416
2022	0.536212	High body-mass index	Singapore	50–74 years	Male	0.0	297.41416

2020	0.944391	Alcohol use	Singapore	50–74 years	Female	0.0	297.41416
2021	0.935317	Smoking	Bhutan	50–74 years	Female	0.0	297.41416
2022	0.559841	Metabolic risks	India	50–74 years	Female	0.0	297.41416
2021	0.724029	Alcohol use	Singapore	50–74 years	Male	0.0	297.41416

The estimated uncertainty indicates heterogeneity in the breast cancer mortality burden across populations. This variability might be linked to differences in access to health care, screening programmes, socio-economic status, lifestyle factors and demographic characteristics.

The model has shown its ability to also estimate the uncertainty of the mortality with several covariates at the same time. This feature is especially useful in public health research, where disease burden often varies greatly across geographical regions and population groups.

### 5.3 Years of Life Lost (YLLs) Analysis

We also fitted the Gaussian Survival Regression model to Years of Life Lost (YLLs) which is the number of years lost due to premature death caused by breast cancer.

The uncertainty of the YLLs prediction was estimated as:

Table 3. Sample Predictions for Breast Cancer YLLs Using Gaussian Survival Regression

Years	Metric	Risk_Factor	Location	Age_Group	Sex	Predicted_Value	Prediction_Uncertainty
2020	0.256940	Behavioral risks	Nepal	50-74 years	Male	0.0	261.789141
2021	0.718482	Metabolic risks	India	50-74 years	Female	0.0	261.789141
2020	0.167316	Alcohol use	Australia	50-74 years	Female	0.0	261.789141
2020	0.930632	Alcohol use	Australia	50-74 years	Male	0.0	261.789141

2020	0.944290	Smoking	Japan	15-49 years	Female	0.0	261.789141
2022	0.412312	Smoking	Pakistan	15-49 years	Female	0.0	261.789141
2020	0.244486	Metabolic risks	Nepal	15-49 years	Female	0.0	261.789141
2021	0.640039	Dietary risks	Pakistan	50-74 years	Female	0.0	261.789141
2022	0.925417	Alcohol use	Canada	15-49 years	Male	0.0	261.789141
2022	0.897333	Dietary risks	Japan	15-49 years	Female	0.0	261.789141
2022	0.931860	Alcohol use	Singapore	50-74 years	Female	0.0	261.789141
2020	0.799412	Smoking	Canada	15-49 years	Male	0.0	261.789141
2020	0.986307	Secondhand smoke	Nepal	50-74 years	Male	0.0	261.789141
2021	0.021101	Behavioral risks	Japan	15-49 years	Female	0.0	261.789141
2021	0.708185	Alcohol use	Pakistan	15-49 years	Female	0.0	261.789141

YLLs had relatively lower uncertainty than mortality burden. This suggests that the variation in the burden of premature deaths may be less variable than the variation in total deaths.

YLLs are an important measure of population health because they reflect the impact of premature deaths before the expected life expectancy. Thus YLL reductions may reflect better early diagnosis, better treatment outcomes and better access to health care.

#### 5.4 Risk Factor Assessment

This analysis included several behavioral and metabolic risk factors associated with breast cancer burden.

### **Smoking and Tobacco Use**

The importance of smoking and tobacco exposure as a major risk factor for the development and death from cancer has long been established. Previous research has found positive associations between tobacco-related exposures and higher cancer burden. The current analysis provides support for inclusion of these factors in models for mortality burden assessment

### **Alcohol Consumption**

Alcohol intake is a well-known, important modifiable risk factor for breast cancer. Epidemiologic evidence indicates that long-term alcohol intake may influence breast cancer risk through hormonal and metabolic pathways. Alcohol exposure was an important explanatory variable in the model for estimating the burden of mortality.

### **Dietary Risks**

Unhealthy dietary patterns have been associated with obesity and metabolic disorders which may increase the risk of breast cancer. Dietary risks influence disease progression through several biological mechanisms, leading to indirect effects on the mortality burden.

### **Metabolic Risks and High Body-Mass Index**

Metabolic disorders and obesity have become major public health problems worldwide. High body-mass index is associated with increased breast-cancer incidence and mortality, especially in postmenopausal women. Statistical models improve their ability to explain variation in breast cancer burden if metabolic risk factors are incorporated.

## **5.5 Geographical and Demographic Variations**

The dataset included observations from many countries such as India, Bangladesh, Nepal, Pakistan, Bhutan, Canada, Japan, Singapore, New Zealand and Australasia.

Differences in mortality burden among countries may reflect variations in:

- Healthcare infrastructure
- Screening coverage
- Early detection programs
- Treatment availability
- Socioeconomic conditions
- Lifestyle factors

Similarly, age-group variations were observed between:

- 15–49 years

- 50–74 years

As disease incidence increases with age and health conditions become more complex, the burden of breast cancer mortality is generally higher among the older age groups.

Sex-based differences were also incorporated into the analytical framework, allowing for a more comprehensive assessment of the population-level burden

## **5.6 Comparison with Previous Studies**

The results of the present study are in line with previous studies carried out on Global Burden of Disease data . Breast cancer is still one of the main causes of death from cancer in the world (Fitzmaurice et al., 2017; Bray et al., 2018; Safiri et al., 2019; Sung et al., 2021).

The earlier work has been primarily focused on descriptive epidemiologic analysis, estimation of burden and evaluation of trends. Conversely, this study utilized a Gaussian Survival Regression framework to model mortality burden and estimate uncertainty associated with breast cancer Deaths and YLLs.

Another advantage of the Gaussian framework is that it can provide uncertainty estimates for public health decision making and risk assessment.

## **5.7 Public Health Implications**

This study’s findings have some practical implications.

- Identification of high-risk population may help in supporting targeted intervention programmes.
- Identifying patterns in the burden of mortality can help policymakers to better allocate healthcare resources.
- Among the risk factors it is necessary to monitor smoking, alcohol consumption and obesity, which may reduce the mortality from breast cancer.
- Statistical prediction models may be useful for future disease burden prediction and public health planning.

In conclusion, the findings imply that the Gaussian Survival Regression offers a user-friendly approach for studying the breast cancer mortality burden with Deaths and YLLs data from the Global Burden of Disease database

## **6. Conclusion and Future Scope**

### **6.1 Conclusion**

Breast cancer remains one of the most common causes of cancer-related death and early death worldwide. To develop effective prevention and healthcare strategies, it is vital to understand the

factors influencing the burden of mortality. In this study, we explored the burden of breast cancer by two main Global Burden of Disease (GBD) indicators, Deaths and Years of Life Lost (YLLs).

We used a Gaussian Survival Regression framework to assess the association between demographic characteristics and risk factors and the burden of breast cancer mortality. The analysis included several explanatory variables, including year, location, age group, sex, smoking, tobacco use, secondhand smoke exposure, alcohol consumption, dietary risks, metabolic risks and high body-mass index.

Behavioral and metabolic risk factors were found to have a substantial contribution to the mortality burden of breast cancer. The Gaussian model could estimate the uncertainty in the mortality burden estimation. The prediction uncertainty was estimated to be ~297.41 for Deaths and 261.79 for YLLs, illustrating the variability in disease burden across populations and geographical regions.

The study also points to the importance of including demographic and risk factor information when examining patterns of breast cancer mortality. The application of Gaussian Survival Regression is a practical statistical approach to analyze disease burden at a population level, and could support researchers and professionals in public health to interpret outcomes related to mortality.

In conclusion, the study shows the usefulness of Global Burden of Disease data and Gaussian-based methods for modeling survival to assess the burden of mortality from breast cancer. The results can inform evidence-based decision making, resource allocation and the development of targeted prevention programs to reduce breast cancer-related deaths and premature mortality.

## 6.2 Future Scope

While the present study has provided valuable insights into the burden of breast cancer mortality, there are several opportunities for future research.

- Future studies may include other GBD indicators such as Disability-Adjusted Life Years (DALYs) and Years Lived with Disability (YLDs) to get a more comprehensive assessment of breast cancer burden.
- Advanced survival analysis methods such as Cox Proportional Hazards models, Accelerated Failure Time models, Frailty models and Bayesian survival models can be considered and compared with Gaussian Survival Regression.
- We can combine machine learning and deep learning methods such as Random Forest, XGBoost, Artificial Neural Networks, Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU) to improve prediction accuracy.
- Future studies should incorporate additional risk factors including genetic predisposition, family history, reproductive characteristics, environmental exposures and indicators of health care accessibility.
- Spatio-temporal analysis can also be performed to evaluate the geographical and temporal variations of the burden of breast cancer mortality at country and regional levels.

- Explainable Artificial Intelligence (XAI) techniques like SHAP and LIME can be used to improve the interpretability of the model and identify the key risk factors that are associated with breast cancer mortality.
- The long-term forecasting studies could be carried out to project future trends in breast cancer deaths and YLLs under different public health intervention scenarios.
- Hybrid statistical/machine learning frameworks can be developed to improve disease burden estimation and uncertainty quantification.
- Cross-cancer comparison studies could offer a broader perspective on patterns of global cancer burden and health-care priorities.
- Our findings could inform the development of data-driven decision-support systems to assist policymakers and healthcare organizations in designing effective breast cancer prevention and control programs.

Overall, the combination of Global Burden of Disease data and sophisticated statistical modelling approaches has the potential to enhance our understanding of the burden of breast cancer mortality and to inform evidence-based public health interventions..

## References

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray F. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *International journal of cancer*. 2015 Mar 1;136(5):E359-86..
2. Global Burden of Disease Cancer Collaboration, Fitzmaurice C, Dicker D, Pain A, Hamavid H, Moradi-Lakeh M, MacIntyre MF, Allen C, Hansen G, Woodbrook R, Wolfe C. The global burden of cancer 2013. *JAMA oncology*. 2015 Jul;1(4):505-27..
3. Fitzmaurice C, Abate D, Abbasi N. Global burden of disease cancer collaboration. global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 29 cancer groups, 1990 to 2017: A systemic analysis for the global burden of disease study (vol 5, pg 1749, 2019). *JAMA oncology*. 2020 Mar 1;6(3):444-.
4. Allemani C. The importance of global surveillance of cancer survival for cancer control: the CONCORD programme. *Cancer control*. 2017;27:10-9.
5. Fullman N, Barber RM, Abajobir AA, Abate KH, Abbafati C, Abbas KM, Abd-Allah F, Abdulkader RS, Abdulle AM, Abera SF, Aboyans V. Measuring progress and projecting attainment on the basis of past trends of the health-related Sustainable Development Goals in 188 countries: an analysis from the Global Burden of Disease Study 2016. *The Lancet*. 2017 Sep 16;390(10100):1423-59.
6. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*. 2018 Nov;68(6):394-424.



7. GBD 2017 Risk Factor Collaborators. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet*. 2018;392(10159):1923-94.
8. Force LM, Abdollahpour I, Advani SM, Agius D, Ahmadian E, Alahdab F, Alam T, Alebel A, Alipour V, Allen CA, Almasi-Hashiani A. The global burden of childhood and adolescent cancer in 2017: an analysis of the Global Burden of Disease Study 2017. *The Lancet Oncology*. 2019 Sep 1;20(9):1211-25.
9. Zhao SW, Wang JW, Szeto IM, Meng LP, Wang Y, Li T, Zhang YM. *WJG*. *World J Gastroenterol*. 2019 Sep 14;25(12):1513-23.
10. GBD 2017 Child and Adolescent Health Collaborators, Reiner Jr RC, Olsen HE, Ikeda CT, Echko MM, Ballestreros KE, Manguerra H, Martopullo I, Milllear A, Shields C, Smith A. Diseases, injuries, and risk factors in child and adolescent health, 1990 to 2017: findings from the global burden of diseases, injuries, and risk factors 2017 study. *JAMA pediatrics*. 2019 Jun;173(6):e190337.
11. Mensah GA, Roth GA, Fuster V. The global burden of cardiovascular diseases and risk factors: 2020 and beyond. *Journal of the American college of cardiology*. 2019 Nov 19;74(20):2529-32..
12. Siegel RL, Torre LA, Soerjomataram I, Hayes RB, Bray F, Weber TK, Jemal A. Global patterns and trends in colorectal cancer incidence in young adults. *Gut*. 2019 Dec 1;68(12):2179-85.
13. Lin L, Yan L, Liu Y, Yuan F, Li H, Ni J. Incidence and death in 29 cancer groups in 2017 and trend analysis from 1990 to 2017 from the Global Burden of Disease Study. *Journal of hematology & oncology*. 2019 Sep 12;12(1):96..
14. Lin G, Qu M. Smart use of state public health data for health disparity assessment. Productivity Press; 2018 Sep 3..
15. Allison PD. Event history and survival analysis. In *The reviewer's guide to quantitative methods in the social sciences* 2018 Nov 15 (pp. 86-97).