

State-of-the-Art Review

Global Burden of Disease study at the World Health Organization: research methods for the most comprehensive global study of disease and underlying health policies

Lee Smith¹, Jae Il Shin², Soo Young Hwang³, Kalthoum Tizaoui⁴, Elena Dragioti⁵, Louis Jacob^{6,7}, Karel Kostev⁸, Seung Won Lee^{9,10*}, Ai Koyanagi^{6,11*}

¹ Centre for Health, Performance, and Wellbeing, Anglia Ruskin University, Cambridge, UK

² Department of Pediatrics, Yonsei University College of Medicine, Seoul, Republic of Korea

³ Yonsei University College of Medicine, Seoul, Republic of Korea

⁴ Laboratory of Microorganisms and Active Biomolecules, Sciences Faculty of Tunis, University Tunis El Manar, Tunis, Tunisia

⁵ Pain and Rehabilitation Centre and Department of Health, Medicine and Caring Sciences, Linköping University, Linköping, Sweden

⁶ Parc Sanitari Sant Joan de Deu/CIBERSAM, ISCIII, Universitat de Barcelona, Fundacio Sant Joan de Deu, Sant Boi de Llobregat, Barcelona, Spain

⁷ Faculty of Medicine, University of Versailles Saint-Quentin-en-Yvelines, Montigny-le-Bretonneux, France

⁸ University Clinic of Marburg, Marburg, Germany

⁹ Department of Data Science, Sejong University College of Software Convergence, Seoul, Republic of Korea

¹⁰ Sungkyunkwan University School of Medicine, Suwon, Republic of Korea

¹¹ ICREA, Pg. Lluís Companys 23, Barcelona, Spain

Abstract

The Global Burden of Disease (GBD) is a compass that navigates researchers and clinicians in discovering the latest knowledge and the current status of health for populations and countries worldwide. The GBD helps to present various methods in evaluating health-related quality of life and assess the health burden, economic burden, years of life lost, and healthy life-years lost from each disease. Beforehand, the details and development status of the dataset need to be identified, and the advantages and disadvantages of each research methodology should be analyzed. The aim of this paper is to help readers comprehend the above basic knowledge of GBD above a certain level. This paper also exemplifies some recent studies to help readers who yearn to understand in detail the latest research using GBD.

Keywords: Global Burden of the Disease; Methodology; GBD.

Received: date: Mar 10, 2022.

Revised date: Apr 12, 2022.

Accepted date: Apr 19, 2022.

Published date: Apr 29, 2022.

*Correspondence:

Seung Won Lee

E-mail: lsw2920@gmail.com

Ai Koyanagi

E-mail: ai.koyanagi@sjd.es

1. Introduction

The history of Global Burden of Disease (GBD) began in the early 1990s with the first landmark study: “World Development Report 1993: Investing in Health”. [1] This was able due to the commission of the World Bank. The 1993 GBD study soon became a role model that systematically measured the world's health problems, analyzing various diseases and aftereffects in representative regions and age groups. Ideas and data of GBD have inspired clinical doctors and researchers to continue expanding their research. [2] GBD 2019, the latest edition, is the global standard and provides independent estimates of 204 countries, territories,

ORCID

Seung Won Lee
<https://orcid.org/0000-0001-5632-5208>
Ai Koyanagi
<https://orcid.org/0000-0002-9565-5004>

Copyright © 2022 Life Cycle.
This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited (CC-BY-NC).

and populations around the world free of charge. Many additional factors, including mortality, life expectancy estimates, disease lists, and risk factors, have been continuously added and have contributed in improving the analytical methods[2]. GBD has grown into an international consortium of 5,500 researchers, currently measuring disability and death from various causes worldwide. This article aims to help researchers understand GBD and promote its use by explaining the history, present, and examples of GBD (Fig. 1).

2. History of GBD

2.1. Researchers and publications

“World Development Report 1993: Investing in Health” is the first GBD study authored by Dr. Christopher Murray, director of the Institute for Health Metrics and Evaluation (IHME)[1]. The GBD 1990 study had a great impact on worldwide health policies and agenda settings. In particular, the study generated global interest in hidden or neglected health issues such as the burden of road injuries. The academic paper has been cited more than 11,000 times[3].

Since 1998, GBD standards regarding GBD work were made by the World Health Organization (WHO). The year 2010 saw significant progress in GBD research, with “Global Burden of Diseases, Injuries, and Risk Factors Study 2010” (GBD 2010), publishing new values for the entire time-series data from 1990 to 2010 and seven serial papers in *The Lancet* in December 2012.[4] In addition to researchers from Harvard University and the World Health Organization, a community of almost 5,000 experts across the world in statistics and epidemiology gathered.

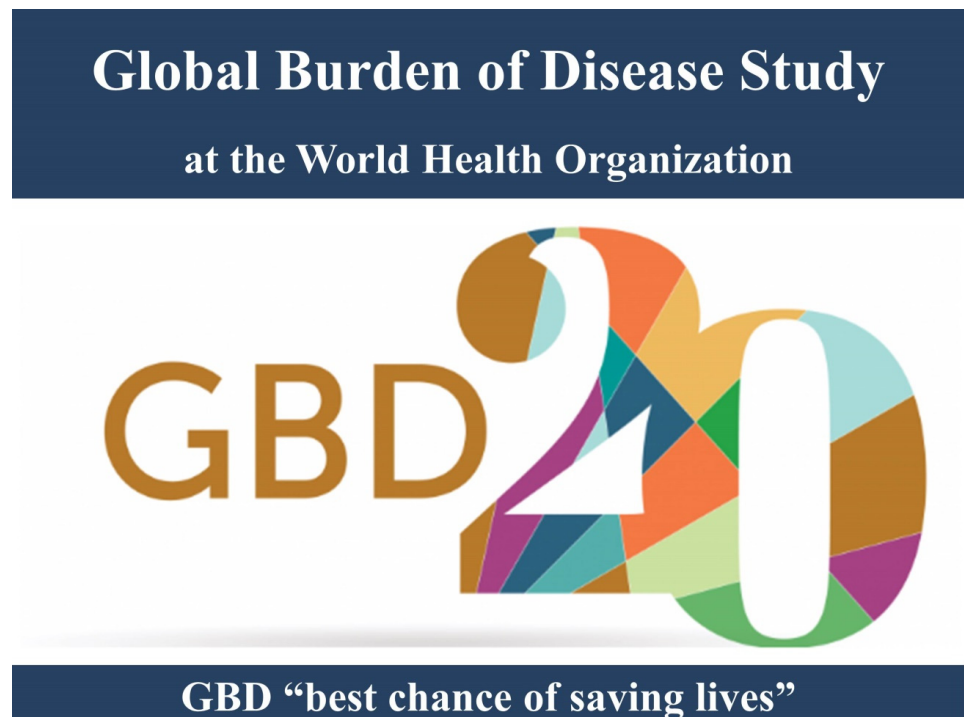


Fig. 1. A Global Burden of Disease Study at The World Health Organization: research methods for the most comprehensive global study of disease and underlying health policies.

The GBD also began to be funded by the Bill & Melinda Gates Foundation[2].

In the GBD study 2013, IHME, the coordination center for GBD contributors' international networks, reflected the work of about 1,000 researchers in more than 100 countries. The results and related research were published in a broader range of journals[5]. The GBD 2015 reflected the work of nearly 2,000 researchers in more than 120 countries[6]. In 2016, starting from 133 countries and 3 regions, 2,518 collaborators participated[7], and it soon expanded to 3,676 collaborators from 146 countries and regions in the GBD 2017[8]. Finally, in the GBD 2019, the project increased to more than 5,000 collaborators in 152 countries and regions. This latest GBD, GBD 2019, was published in the Lancet[2, 9].

2.2. Diseases and territories covered

Until the GBD study 1990, 107 diseases and 483 non-fatal complications were researched through a dataset which covered 8 regions and 5 age groups[3]. Moreover, the GBD 2010 distributed a dataset of estimates for 291 diseases and injuries, 67 risk factors, 1,160 aftereffects in 21 regions, 20 age groups, and 187 countries[4].

In the GBD 2013 study, 188 countries including more than 300 illnesses and injuries, 79 risk factors, and 2,300 aftereffects were covered, with many landmark studies conducted on smoking[5]. Researchers expanded the GBD to 315 diseases and injuries and 79 risk factors for 195 countries in 2015[6]. The total of locations became 774 in 2016[7], including 333 diseases and injuries, 84 risk factors, and 23 age groups[8]. The GBD 2016 study included research on alcohol, gun accidents, etc[10].

In the GBD 2017, the dataset found a significant increase in temporal coverage. Estimates of mortality and life expectancy were increased compared to 1950, and a total of 359 new causes of disease and injury were added to the list of fatal and non-fatal causes[8]. Contemporarily, GBD also added one new risk, bullying victimization, and 80 new risk-result pairs[8]. The latest GBD, GBD 2019, contained the most extensive data[2, 9]. Estimated mortality and life expectancy values have been even more increased to the most detailed level to a total of 990 locations, with new causes added to the list of fatal and non-fatal causes resulting in 369 illnesses and injuries[2, 9].

2.3. Methodological advances

Compared to the early GBD, many methodological advances have been made in the modern GBD. The most representative of those advancements is the development of the Disease Burden Unit (DBU)[4]. WHO created the DBU which was able to generate GBD estimates in 2000, 2001, and 2002, and published estimates in the annual World Health Report[11].

The annual update to the entire series of GBD estimates began since GBD 2015[6]. These more frequent updates, containing the most useful and timely picture of population health, were provided to policymakers, donors, and other decision-makers. The 2015 update also introduced the Socio-demographic Index (SDI), a summary measure that extended the methodologies, datasets, and tools used in previous GBDs and identified where countries or other regions are in the development spectrum[6]. SDI, expressed through a scale from 0 to 1, is the composite

mean of per capita income, average education level, and fertility ranking in all areas of the GBD study[6].

GBD 2017 was the first to provide independent estimates of 195 countries, territories, and populations worldwide using a standardized and replicable approach, as well as comprehensive updates on fertility rates. Through GBD 2017, reviewing many Sustainable Development Goal indicators and generating forecasts up to 2030 using prediction methods became possible[8, 12, 13]. Such development enabled researchers to assess the rate of change required to achieve Sustainable Development Goals.

Published in *The Lancet* in October 2020, GBD 2019 issued a comprehensive update on fertility rates[14]. The total fertility rate is a new summary measure update representing the average number of children a woman would deliver over her lifetime.

3. Latest dataset and methods of GBD

3.1. GBD 2019 schema

The fundamental aspect of the GBD data schema is the age, sex, country, and year of the diseases. Most important estimates can be derived from these core variables. In addition, GBD 2019 demographic data include population; the analogized rate of population change (2010 to 2019), total facility rate and live births (1950, 1980, and 2019), and net productive rate[2, 9].

More detailed data are provided on deaths. They include under-five mortality rate, change in under-five mortality rate (2010 to 2019), probability of death between ages 15 and 60 years and life expectancy at birth (by sex), healthy life expectancy, the total number of deaths, the total number of deaths among children under 5 years, and observed and expected life expectancy[14].

In addition, life habits and important causes of disease like alcohol consumption and daily smoking are included[10]. There are also the current status variable and previous status variable of alcohol and smoking, with the amount of consumption. Variables that indicate accidents and injuries (i.e., from firearms) can also be extracted from the dataset. Examples of these are explained in the final part of this article.

The GBD dataset includes a variety of diseases which are classified by level. The three diseases with the highest level (i.e., level 1) of GBD are (i) communicable, maternal, neonatal, and nutritional diseases; (ii) non-communicable diseases; and (iii) injuries. These are again classified as level 2 diseases as shown in the following table (Table 1).

Level 2 diseases are subdivided into all diseases, including 169 levels 3 and 165 level 4 diseases. These 351 causes and diseases can be matched with ICD-10 codes, and these matching tables are published by Wang et al., 2019[14].

Cause of death is obtained from vital registration, verbal autopsy, and cancer registration data. Particularly in cases of cancer data, pathology-based cancer registries with defined populations and hospital-based cancer registries exist.

3.2. GBD 2019 methodologies

The age-standardized rate is an operation to standardize mortality or the burden of death[15, 16].

Table 1. Level 1 and 2 GBD diseases

GBD diseases	Level
Communicable, maternal, neonatal, and nutritional diseases	1
HIV/AIDS and sexually transmitted infections	2
Respiratory infections and tuberculosis	2
Enteric infections	2
Neglected tropical diseases and malaria	2
Other infectious diseases	2
Maternal and neonatal disorders	2
Nutritional deficiencies	2
Non-communicable diseases	1
Neoplasms	2
Cardiovascular diseases	2
Chronic respiratory diseases	2
Digestive diseases	2
Neurological disorders	2
Mental disorders	2
Substance use disorders	2
Diabetes and kidney diseases	2
Skin and subcutaneous diseases	2
Sense organ diseases	2
Musculoskeletal disorders	2
Other non-communicable diseases	2
Injuries	1
Transport injuries	2
Unintentional injuries	2
Self-harm and interpersonal violence	2

GBD: Global Burden of Disease.

The age-based distribution is a very important factor in the number of deaths or patients in a particular population. In most cases, the group with a large number of elderly people has a high mortality and prevalence of disease per population. Therefore, it is necessary to regularize these values by standardizing them as groups with the same age distribution. Using the standard population suggested by the WHO, the age-standardized rate can be computed based on

adjusted age distribution differences in the population. This operation can be done by applying the observed age-specific mortality and prevalence for each population.

Mortality-to-incidence ratio is calculated by dividing the mortality rate for the year by the incidence rate. This is an indicator used to identify inequality in cancer prognosis. Mortality-to-incidence ratio is a higher-level comparative measure of a crude survival estimate than relative survival. Because of its simplicity, high-quality incidence and mortality data are available in most countries, so survival can be compared internationally[17].

Spatiotemporal Gaussian Process Regression (ST-GPR) is a time series model primarily used to estimate risk factor exposure. This smooths the value for each risk factor. Through this, interpolating non-linear trends are possible and only meaningful signals are extracted from noise.[2, 9]

The Healthcare Access and Quality (HAQ) Index measures the access and quality of medical benefits by the use of amenable mortality[18]. Deaths from avoidable causes in the presence of adequate medical care are calculated. HAQ Index approximates the national levels of personal health-care access and quality and is used as covariates in GBD.

The Cause of Death Ensemble model (CODEm) is an integrated cause of death modeling environment[19]. It enables trend estimation by exploring various possible models. In this process, a covariate selection algorithm is used, and this combination is executed through four model classes, including the spatial-temporal Gaussian process regression (ST-GPR). The algorithm proposes possible models by generating many plausible combinations of covariates. Finally, it combines with an ensemble with optimal non-sample validity.

Disability-Adjusted Life Year (DALY) is the number of years lost due to ill-health, disability, or early death. As a measure of overall disease burden, DALY can be calculated comparing the sum of the years of life lost (YLL) and the years lost due to disability (YLD)[20]. YLL is a product of the number of deaths (due to the specific condition) and standard life expectancy at age of deaths. YLD represents the loss of each full year of healthy life due to disability or ill-health. This can be calculated by the product of three terms: incident cases in the population, the weight of disability due to the specific condition, and the average year of the case until death (or remission).

4. Examples of GBD 2019 Study

One of our team's latest examples using GBD 2019 was carried out using the sudden infant death syndrome (SIDS) variable[2]. This article analyzed the global disease burden of SIDS and its trends from 1990 to 2019. The result was compared to the burden of SIDS according to the SDI. In this article, we collected epidemiological data from 204 countries from 1990 to 2019 using GBD, including vital statistics and civil registrations. Through this, estimates of SIDS were modeled in terms of disease burden and mortality. Mortality rates per 100,000 population, crude mortality, and DALYs were also calculated. The results showed significant reduction in the SIDS burden in regions from 1990 to 2019, and SDI levels were important factors the reduction.

Another critical study is a landmark study on smoking[21], measuring tobacco use from

nationally representative sources. The article used the average relationship between different definitions to impute survey data which does not include daily tobacco smoking. ST-GPR was used to estimate age-sex-country-year observations. Cigarette consumption was utilized to estimate daily cigarettes per smoker. The research drew principle and conclusive outcomes of anti-smoking campaigns: daily smoking reduction since 1980 for all genders; net increase in the number of smokers due to population growth; net deaths, years of life lost, and disability-adjusted life-years due to tobacco.

Alcohol has also been addressed by GBD researchers[10]. This article used individual and population-level alcohol consumption. Estimates were calculated in terms of current drinking, abstention, alcohol-attributable deaths, and DALYs. The research standardized daily alcohol consumption as 10g of pure ethyl alcohol. Alcohol sales estimates were adjusted using tourist and unrecorded consumption. This research found the important conclusion that the risk of cancers and all-cause mortality has no threshold from alcohol. In other words, risks immediately increase according to consumption, and the only safe dose of alcohol is zero, which contrasts with the conventional view that small amounts of alcohol do not have negative effects on people's health.

The final research example is global mortality from firearms[22]. Researchers used deidentified aggregated data to generate estimates. These include location-years of vital registration data and rates of death. Revised proxy measures were used to evaluate the firearm injury deaths from firearms accessibility (or availability). The proxy used is a combination of per capita gun ownership and the proportion of suicides by firearm. In this article, crucial understandings of global variation in firearm mortality rates were derived, which will support future intervention and prevention policies. Also, the researchers were able to estimate the specific value of firearm injury deaths that happened worldwide in 2016, approximately 0.2 million.

5. Conclusion

GBD research greatly contributes to helping people understand important diseases in various countries. Many researchers have cooperated around the world through the long-standing effort with their precious time and even now continue to make progress in GBD collaboration. By applying the datasets, methodologies, and examples covered in this article, readers will be able to conduct their own GBD research and derive more policy, medical, and clinical implications.

Capsule Summary

The Global Burden of Disease is the most comprehensive international collaborate research to measure epidemiological levels and global trends. We summarized the substantial improvements in statistical methods and dataset of Global Burden of the Disease study adopted by the World Health Organization.

Patient and public involvement

No patients were directly involved in designing the research question or in conducting the research. No patients were asked for advice on interpretation or writing up the results. There are no plans to involve patients or relevant patient community in dissemination at this moment.

Transparency statement

The leading authors (Dr. SWL and AK) are an honest, accurate, and transparent account of the study being reported.

Acknowledgements

None

Author contribution

All authors contributed to the preparation of this review.

Funding

Research reported in this publication was supported by the Bill and Melinda Gates Foundation. The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Competing interests

The authors have no conflicts of interest to declare for this study.

Provenance and peer review

Not commissioned, externally peer reviewed.

References

1. Musgrove P. Investing in health: the 1993 world development report of the world bank. *Bulletin of the Pan American Health Organization*. 1993;27(3):284-6.
2. Park S, Han JH, Hwang J, Yon DK, Lee SW, Kim JH, et al. The global burden of sudden infant death syndrome from 1990 to 2019: a systematic analysis from The Global Burden of Disease Study 2019. *QJM : monthly Journal of the Association of Physicians*. 2022.
3. Murray CJ, Lopez AD, Organization WH. *The Global Burden of Disease: a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020: summary*: World Health Organization; 1996.
4. Murray CJ, Ezzati M, Flaxman AD, Lim S, Lozano R, Michaud C, et al. GBD 2010: design, definitions, and metrics. *Lancet (London, England)*. 2012;380(9859):2063-6.
5. Wang H, Liddell CA, Coates MM, Mooney MD, Levitz CE, Schumacher AE, et al. Global, regional, and national levels of neonatal, infant, and under-5 mortality during 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet (London, England)*. 2014;384(9947):957-79.
6. Karimkhani C, Green AC, Nijsten T, Weinstock MA, Dellavalle RP, Naghavi M, et al. The global burden of melanoma: results from the Global Burden of Disease Study 2015. *The*

- British Journal of Dermatology. 2017;177(1):134-40.
7. Steel N, Ford JA, Newton JN, Davis ACJ, Vos T, Naghavi M, et al. Changes in health in the countries of the UK and 150 English Local Authority areas 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet* (London, England). 2018;392(10158):1647-61.
 8. Xia Y, Wu Q, Wang H, Zhang S, Jiang Y, Gong T, et al. Global, regional and national burden of gout, 1990-2017: a systematic analysis of the global Burden of Disease Study. *Rheumatology* (Oxford, England). 2020;59(7):1529-38.
 9. Murray CJ, Aravkin AY, Zheng P, Abbafati C, Abbas KM, Abbasi-Kangevari M, et al. Global burden of 87 risk factors in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. 2020;396(10258):1223-49.
 10. Griswold MG, Fullman N, Hawley C, Arian N, Zimsen SRM, Tymeson HD, et al. Alcohol use and burden for 195 countries and territories, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet*. 2018;392(10152):1015-35.
 11. Levav I, Rutz W. The WHO World Health Report 2001 new understanding--new hope. *The Israel Journal of Psychiatry and Related Sciences*. 2002;39(1):50-6.
 12. Lee SW. Methods for testing statistical differences between groups in medical research: statistical standard and guideline of Life Cycle Committee. *Life Cycle*. 2022;2:e1.
 13. Lee SW. Regression analysis for continuous independent variables in medical research: statistical standard and guideline of Life Cycle Committee. *Life Cycle*. 2022;2:e3.
 14. Wang H, Abbas KM, Abbasifard M, Abbasi-Kangevari M, Abbastabar H, Abd-Allah F, et al. Global age-sex-specific fertility, mortality, healthy life expectancy (HALE), and population estimates in 204 countries and territories, 1950-2019: a comprehensive demographic analysis for the Global Burden of Disease Study 2019. *The Lancet*. 2020;396(10258):1160-203.
 15. Wyper GMA, Grant I, Fletcher E, McCartney G, Fischbacher C, Stockton DL. How do world and European standard populations impact burden of disease studies? A case study of disability-adjusted life years (DALYs) in Scotland. *Archives of public health = Archives belges de sante publique*. 2020;78:1.
 16. Galles NC, Liu PY, Updike RL, Fullman N, Nguyen J, Rolfe S, et al. Measuring routine childhood vaccination coverage in 204 countries and territories, 1980-2019: a systematic analysis for the Global Burden of Disease Study 2020, Release 1. *The Lancet*. 2021;398(10299):503-21.
 17. Choi E, Lee S, Nhung BC, Suh M, Park B, Jun JK, et al. Cancer mortality-to-incidence ratio as an indicator of cancer management outcomes in organization for economic cooperation and development countries. *Epidemiol Health*. 2017;39:e2017006-e.
 18. Fullman N, Yearwood J, Abay SM, Abbafati C, Abd-Allah F, Abdela J, et al. Measuring performance on the healthcare access and quality Index for 195 countries and territories and selected subnational locations: a systematic analysis from the Global Burden of Disease Study 2016. *The Lancet*. 2018;391(10136):2236-71.
 19. Foreman KJ, Lozano R, Lopez AD, Murray CJ. Modeling causes of death: an integrated approach using CODEm. *Popul Health Metr*. 2012;10:1-.
 20. Niyibitegeka F, Thavorncharoensap M, Riewpaiboon A, Youngkong S. Does the choice of health metric, DALY or QALY, influence conclusions of health economic evaluation? a

- case study of rotavirus vaccine in burundi. *applied health economics and health policy*. 2022.
21. Ng M, Freeman MK, Fleming TD, Robinson M, Dwyer-Lindgren L, Thomson B, et al. Smoking prevalence and cigarette consumption in 187 countries, 1980-2012. *Jama*. 2014;311(2):183-92.
 22. Naghavi M, Marczak LB, Kutz M, Shackelford KA, Arora M, Miller-Petrie M, et al. Global mortality from firearms, 1990-2016. *Jama*. 2018;320(8):792-814.