



Googling the Lifetime Risk of Stroke Around the World

Thanh G. Phan^{1,2*†}, Anisha Haseeb^{1,2}, Richard Beare^{3,4,5}, Velandai Srikanth^{3,4}, Amanda G. Thrift² and Henry Ma^{1,2}

¹ Department of Neurology, Monash Health, Melbourne, VIC, Australia, ² Stroke and Aging Research Group, Medicine, School of Clinical Sciences, Monash University, Melbourne, VIC, Australia, ³ Department of Medicine, Frankston Hospital, Peninsula Health, Melbourne, VIC, Australia, ⁴ Central Clinical School, Monash University, Melbourne, VIC, Australia, ⁵ Developmental Imaging, Murdoch Children's Research Institute, Melbourne, VIC, Australia

Objective: We aimed to utilize the data on lifetime risk of stroke, from Global Burden of Disease (GBD) 2016, in combination with open data platforms to create an interactive map for use by clinicians and members of the public. Further, we explore the relationship between life expectancy and lifetime risk of stroke.

Design: Enhancing visual display of large volume of data.

Setting: Worldwide estimates of the lifetime risk of stroke obtained from the GBD 2016.

Participants: None.

Intervention: None.

Methods: Data were extracted from a portable document format (pdf) copy of the GBD article on the lifetime risk of stroke and exported into the R programming environment (version 3.4.4). These data were merged with (i) the world map boundary, (ii) open data platforms from the World Bank (life expectancy and income), and (ii) open data from the United Nation Population Prospects 2017. Further we plotted the relationship between the adjusted lifetime risk of stroke and life expectancy.

Outcomes: The map of the global burden of stroke shows a higher lifetime risk of stroke among high-income countries than in low-income countries (<https://gntem3.shinyapps.io/strokeglobal/>). The greatest risk was among upper-middle-income countries such as China and Eastern and Central European countries such as Latvia and Romania. The lifetime risk of stroke increased in countries with higher life expectancy ($\beta = 0.48 \pm 0.047$, 95% confidence interval = 0.390–0.574, $R^2 = 0.38$, $p < 0.01$).

Conclusion: Overall life expectancy is a major driver of the lifetime risk of stroke. The interactive map enables clinicians to search information about the lifetime risk of stroke interactively and navigate by zooming in and out, while still retaining high resolution.

Keywords: stroke, googling, map, lifetime risk, global, supplementary

OPEN ACCESS

Edited by:

Aristeidis H. Katsanos,
McMaster University, Canada

Reviewed by:

Lina Palaiodimou,
University General Hospital
Attikon, Greece
Georgios Seitidis,
University of Ioannina, Greece

*Correspondence:

Thanh G. Phan
thanh.phan@monash.edu

†ORCID:

Thanh G. Phan
orcid.org/0000-0003-3400-6323

Specialty section:

This article was submitted to
Stroke,
a section of the journal
Frontiers in Neurology

Received: 05 May 2020

Accepted: 15 June 2020

Published: 31 July 2020

Citation:

Phan TG, Haseeb A, Beare R,
Srikanth V, Thrift AG and Ma H (2020)
Googling the Lifetime Risk of Stroke
Around the World.
Front. Neurol. 11:729.
doi: 10.3389/fneur.2020.00729

INTRODUCTION

This article was inspired by a recent publication on the global lifetime risk of stroke from the Global Burden of Disease (GBD) 2016 collaborators (1). The GBD collaborators estimated that the global lifetime risk of stroke from the age of 25 years was one in four and that the risk was twice as high for ischemic than hemorrhagic stroke. Further, the authors provided a choropleth (thematic) map

illustrating the global effect of this disease, with the color on the map reflecting the lifetime stroke risk. A drawback is that the map is provided in a static image format and does not retain high resolution with zooming, nor does it easily lend itself to an interactive search. To find data for a country of interest, the reader would need to switch between the map in the main article and the 125 pages of data in the supplementary material (29.1 MB of data). The reader would struggle to easily extrapolate exact information on the lifetime risk of stroke from the map given the information is binned to a range of values. Further readers may not access the supplementary material (2, 3).

In this article, we have created a dynamic vectorized map on the web incorporating data published from the GBD 2016 global burden of stroke. Data on lifetime risk of stroke and other variables from open data sources can be obtained upon clicking the country of interest (1). Further, we bring forth in the choropleth map the remaining body of work by the GBD 2016 collaborators on ischemic and hemorrhagic stroke, available in the supplementary material.

METHODS

Data From GBD 2016

Data in the *New England Journal of Medicine* article on lifetime risk of stroke were extracted from portable document format (pdf) copy of the article using *pdftables* library in the R programming environment. The data were exported as Microsoft Excel files. These data were imported into R for merging with the World Bank data described below (version 3.4.4) (1). This task was performed using *stringr* library to ensure that the names of the countries matched between the different datasets. The method for estimating the lifetime risk of stroke is detailed by the GBD 2016 investigators in the *New England Journal of Medicine* article (1). In brief, the GBD 2016 investigators (1) based their estimate of the lifetime risk of stroke on stroke incidence, cause-specific mortality, and all-cause mortality at a global level (defined by 7 GBD super regions and 21 GBD regions) and at a national level (195 countries) (4, 5).

Open Data

Data on country income and life expectancy were obtained from The World Bank and comprise an Atlas Method conversion factor for classifying economy (6, 7).

Map

The interactive thematic map was created using *leaflet* library in R (8). We added several functions to enhance user experience. The map of the world was color coded to reflect the estimated lifetime risk of stroke. Within *leaflet*, a popup function is added to enable the reader to query the data by clicking on to the country of interest.

The *Shiny* library in R is used to build a web-based app (application) with files uploaded onto the internet. We wrote a function to take the argument “country of interest” so that readers can search within the map data for that country. A “fuzzy text matching” approach is used so that even if the user mistyped “Australi” instead of “Australia,” the search path would lead the

user to the correct country. The tile map we had chosen was “Open Street Map.”

Plots

We explored the relationship between the lifetime risk of stroke and life expectancy, National income, and location of countries by latitude and continents. The plots were created within R using the *ggplot2* (9) library (Grammars of Graphics style) and expressed as a *plotly* (10) object for user interaction. This way, the plots of lifetime risk by latitude, continent, income, and life expectancy are all interactive. The reader can hover over the data point to obtain similar information on each country.

Pseudocode

```
GBD 2016 → pdf → Excel (pdftables)
World Bank → Open Data → Excel
Merge GBD 2016 and World Bank data
Excel → R → leaflet Map → Shiny → App
          → ggplot2 graphs → plotly → App.
```

Statistical Analysis

Lifetime risk of stroke was regressed against and life expectancy of each country. This step was repeated for each country income group. The (11) test of decreasing trend was performed for the countries, ordered by income groups (high to upper middle, lower middle, and low income).

RESULTS

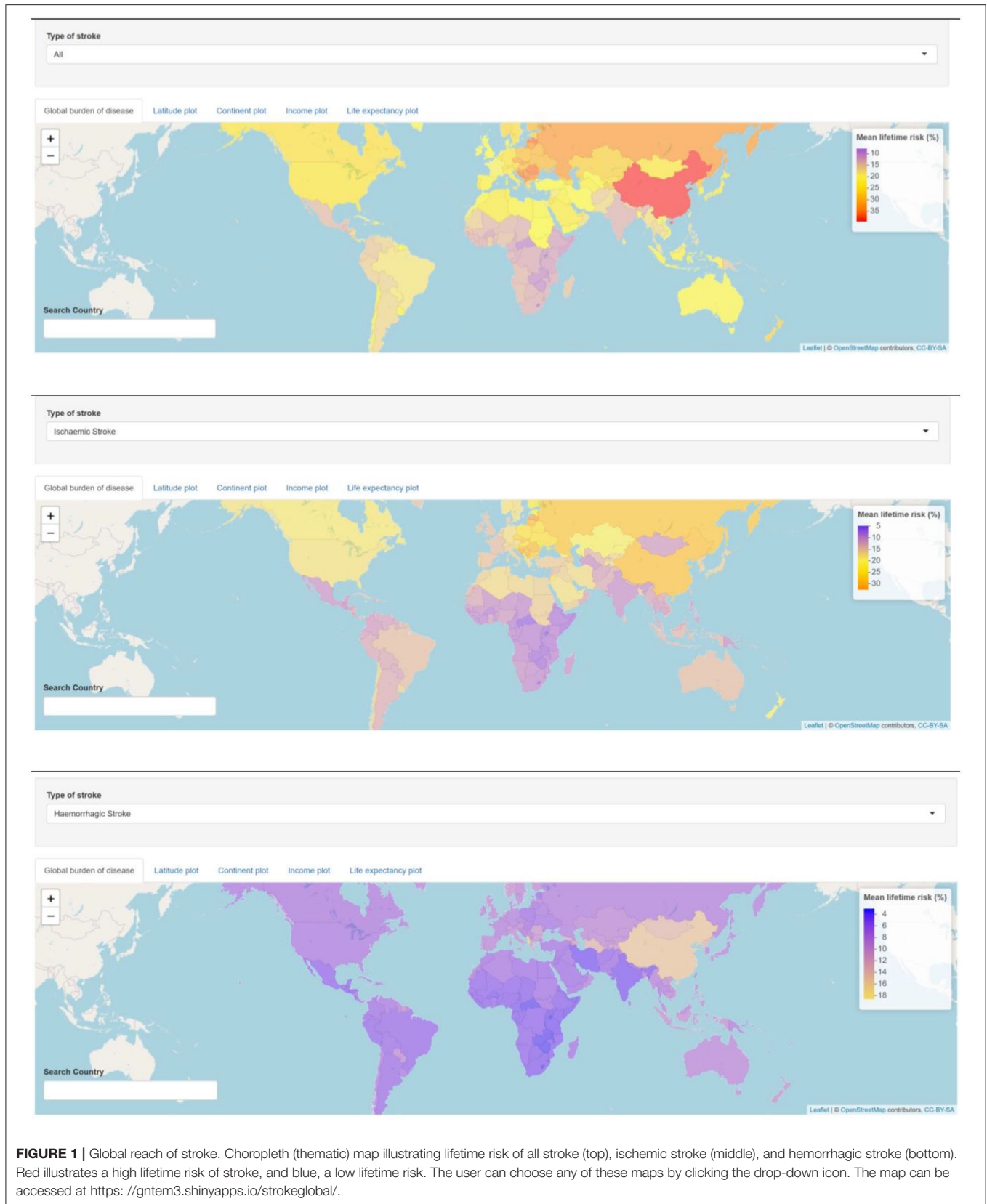
The interactive map of global burden of stroke shows high estimated lifetime risk among countries with high income and Eastern European countries with upper middle income (<https://gntem3.shinyapps.io/strokeglobal/>). The lowest lifetime risk of all stroke is observed in African countries with low income (**Figure 1**). The risk of all stroke was greatest in countries with a higher life expectancy [$\beta = 0.48 \pm 0.047$, 95% confidence interval (CI) = 0.390–0.574, $R^2 = 0.380$, $p < 0.01$]; for every year increase in life expectancy, the lifetime stroke risk increases by $\sim 0.48\%$ (**Figure 2**). Outliers in this figure include upper-middle-income countries, such as China, and Eastern and Central European countries, such as Latvia and Romania. The identity of these countries can be easily accessed by hovering on the plot.

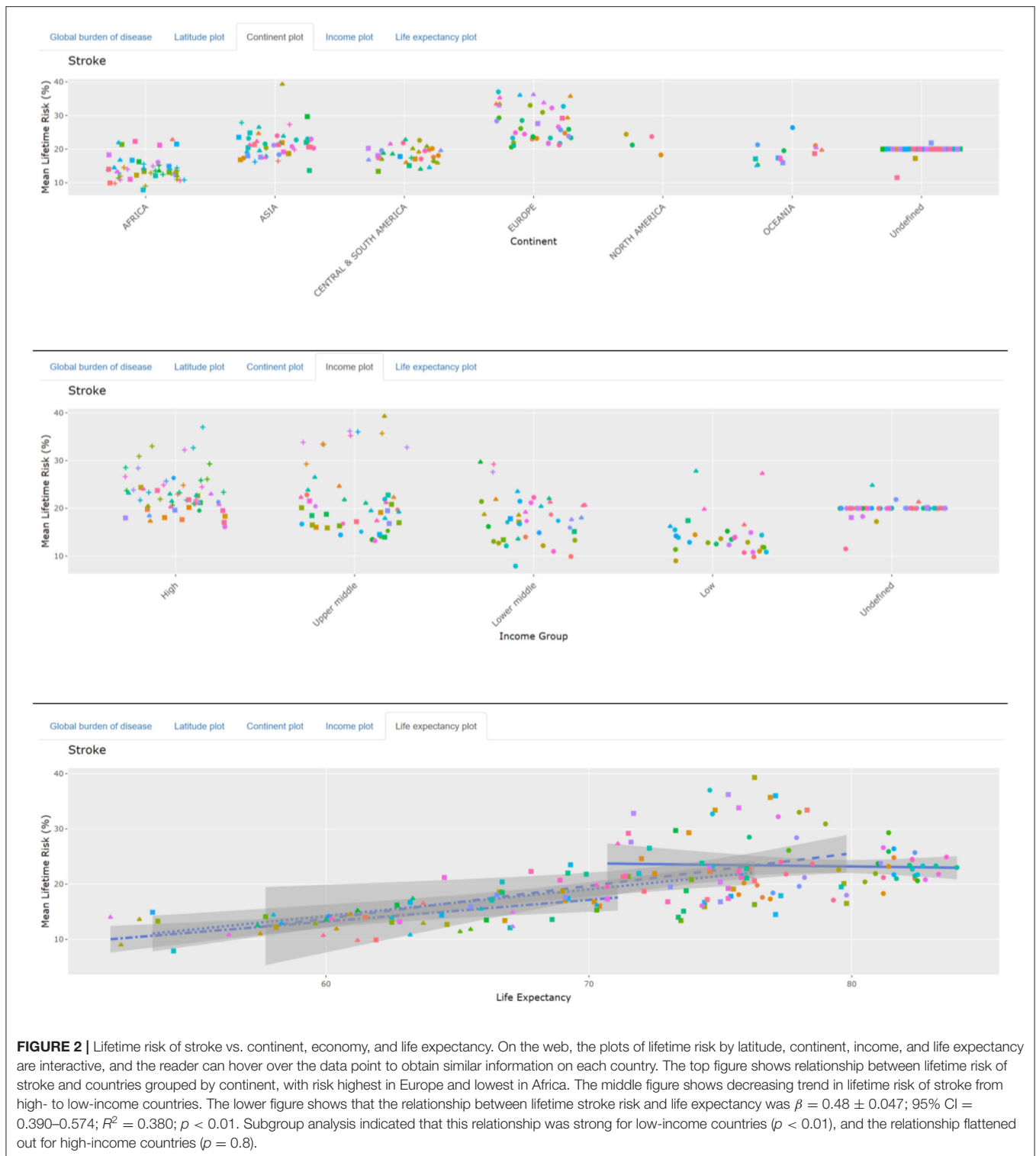
Continents

The mean in lifetime stroke risk is highest in Europe (27.7 ± 5.1), North America (21.9 ± 2.8), Asia (21.5 ± 4.2), Oceania (18.7 ± 3.0), Central and South America (18.2 ± 2.3), and Africa (14.2 ± 3.5).

World Bank Income

The relationship between lifetime stroke risk and life expectancy for low-income countries was $\beta = 0.392 \pm 0.108$, 95% CI = 0.357–0.426, $R^2 = 0.322$, $p < 0.01$; for lower-middle-income countries, it was $\beta = 0.482 \pm 0.102$, 95% CI = 0.281–0.683, $R^2 = 0.362$, $p < 0.01$; for upper-middle-income countries, it was $\beta = 0.590 \pm 0.216$, 95% CI = 0.216–1.000, $R^2 = 0.140$, $p = 0.01$. The relationship between lifetime stroke risk and life expectancy was





not significant for high-income country ($\beta = -0.0568 \pm 0.194$, 95% CI = -0.438 – 0.325 , $R^2 = 0.085$, $p = 0.8$).

Hemorrhagic Stroke

The estimated lifetime risk of ischemic and hemorrhagic stroke can be viewed by clicking on the tab “Type of Stroke.” China

has the greatest lifetime risk of both ischemic (28.5%) and hemorrhagic (15.5%) stroke (**Figure 1**). With the exception of China and Mongolia, the lifetime risk of hemorrhagic stroke in Asia is similar to Oceania or Europe and $\sim 8\%$. The pattern of high lifetime ischemic stroke risk in Eastern and Central Europe did not exist for hemorrhagic stroke. In Europe, the only

countries with a high lifetime risk of hemorrhagic stroke were Montenegro (18.6%) and Albania (17.1%).

The map functions on web engines such as Microsoft Edge, Google Chrome, and Mozilla. An internet connection is required to use the map.

DISCUSSION

We have created an interactive choropleth map of the global lifetime risk of all stroke in 2016 in an accessible format that can be viewed by anyone with internet access. Lifetime risk of stroke denotes the likelihood of suffering from a stroke in one's lifetime. Several additional features were added to explain the observation of high lifetime risk of stroke among high-income countries and selected countries in Eastern Europe with upper middle income. It is proposed that the readers view this article as a visual display of the large body of work by the GBD 2016 collaborators and refer to the original articles regarding in-depth discussion of the method and results for generating statistics (1).

The collaboration between data and/or citizen scientists and health care researchers was made possible by the availability of data from GBD 2016 investigators (1) and open data from the United Nations (7) and World Bank (6, 12). The key aspect of this article is the conversion from pdf to Microsoft Excel and uploading of data onto the web. This aspect enables seamless access to the tabular results from GBD 2016 investigators (1). The *pdftables* library in R is not the only package that can perform this task. Other available libraries within R for extracting such data include *pdftools* and *tabulizer*. The *Shiny* library serves as a framework within R environment for creating a web-based app and without the need to learn languages such as HTML (hypertext markup language) or CSS (cascading style sheets). The approach used here is one way of bringing the vast data hidden in supplementary material out into the open (2, 3). We were fortunate to be given permission by GBD 2016 investigators to show their results in map form. While it would help readers to have access to the full data, these data are already publicly available in the original manuscript. Here, we provide a different approach to visualize the body of work.

Data on life expectancy help to explain the relationship between a low lifetime risk of stroke and low-income countries. The life expectancy in low-income countries is <60 years, some 20 years less than that in high-income countries (~80 years). As the risk of stroke increases with age, stroke will be more common in countries with a greater life expectancy than in countries where people die at younger ages. For example, in Japan, the high lifetime risk of stroke is at least partially attributable to a high life expectancy (84 years). In contrast, in Somalia, the relatively low lifetime risk of stroke (10.8%) is likely largely attributable as people do not live to ages where stroke most commonly occurs, as life expectancy is only 56.3 years. Indeed, the low risk observed in the low-income countries (such as Sub-Saharan Africa including Somalia) might be explained by the high risk of death from any cause other than stroke at younger ages.

There are some worrying signs of very high lifetime risk of stroke among selected high- and upper-middle-income countries in Eastern and Central Europe. In this region, countries such as Bosnia, Latvia, and Romania have a lifetime risk of stroke >35%, whereas the risk is <30% in Poland, Moldova, and Ukraine (Figures 1, 2) (6). Overall, these risks are much higher than those in Western and Southern European countries. The exception in Eastern Europe is Estonia, which has similar lifetime stroke risk as Western Europe. In an earlier study from GBD 2013 investigators, it was proposed that 74.2% of stroke burden was due to behavioral factors such as smoking, poor diet, and low physical activity (13). It is not clear if additional factors are at play in these countries such as sociopolitics and organization of health services (14, 15). Elsewhere there has been an 11.8% decline in the all-cause burden of disease attributable to smoking in high-income countries (15).

The pattern described above is driven by the lifetime risk of ischemic stroke. Globally, it appears that the lifetime risk of hemorrhagic stroke is similar between countries, with exceptions being China, Mongolia, Albania, Montenegro, and Turkmenistan. This pattern is somewhat different from the previously reported high risk of intracerebral hemorrhage in Asians (16). One possible explanation is that the high lifetime risk of hemorrhagic stroke may be limited to Chinese and Mongolians, the former being the world's most populous nation (15). The GBD 2016 had combined subarachnoid hemorrhage and intracerebral hemorrhage into hemorrhagic stroke data. However, this should not significantly affect the quality of the data as subarachnoid hemorrhage forms a very minor subset of all stroke (14).

Limitations

The GBD 2016 collaborators provided lifetime risk on most country but noted that data were not available in some countries (1). The GBD 2016 collaborators had acknowledged this limitation in their article and had discussed the assignment of values of neighboring countries and data on country-level risk exposure (1). It should be noted that the World Bank has changed the status of some high-income countries over time, and this change has affected countries such as Russia and Venezuela; these countries are now listed as upper-middle-income countries (6). One advantage of this dynamic map is that it can easily be updated should the GBD 2016 collaborators wish to make changes.

CONCLUSION

An interactive thematic map extends the value of the excellent work by the GBD 2016 collaborators on the lifetime risk of stroke.

ARTICLE SUMMARY

Strength

(1) Data are based on the adjusted lifetime risk of stroke that come from GBD 2016 collaborators.

(2) The interactive web display is available at <https://gntem3.shinyapps.io/strokeglobal/>.

(3) The plots of lifetime risk by latitude, continent, income, and life expectancy are interactive.

PERMISSION

On 20th April 2020, the GBD team provided written permission to use the data. They noted that “any data made available for download on IHME Websites can be used, shared, modified, or built upon by non-commercial users via the Open Data Commons Attribution License. Our full Terms and Conditions can be found here.”

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here (15): the GBD 2016 investigators indicated that data can be found on the IHME website.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for

participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

TP: design. AH, TP, and RB: data extraction, statistical analysis, and mapping. TP, AH, RB, HM, VS, and AT: manuscript writing. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

The figure in this article was created using data from the ©OpenStreetMap contributors and tiles from osm.org. The data are available under the Open Database License while the tiles are available under Creative Commons Attribution-Share Alike 2.0 license. Details are available at www.openstreetmap.org/copyright. While the maps were created using the OpenStreetMap tiles, it does not suggest that the licensor endorses the use of this map. This material was presented as a poster at the Stroke Society of Australasia in 2019 in Canberra, Australia and published in Abstract form in the International Journal of Stroke on 1/9/2019.

REFERENCES

- Feigin VL, Nguyen G, Cercy K, Johnson CO, Alam T, Parmar PG, et al. Global, regional, and country-specific lifetime risks of stroke, 1990 and 2016. *N Engl J Med.* (2018) 379:2429–37. doi: 10.1056/NEJMoa1804492
- Pop M, Salzberg SL. Use and mis-use of supplementary material in science publications. *BMC Bioinform.* (2015) 16:237. doi: 10.1186/s12859-015-0668-z
- Price A, Schroter S, Clarke M, McAneney H. Role of supplementary material in biomedical journal articles: surveys of authors, reviewers and readers. *BMJ Open.* (2018) 8:e021753. doi: 10.1136/bmjopen-2018-021753
- GBD 2016 Risk Factors Collaborators. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2016: a systematic analysis for the global burden of disease study 2016. *Lancet.* (2017) 390:1345–422. doi: 10.1016/S0140-6736(17)32366-8
- GBD 2015 Risk Factors Collaborators. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: a systematic analysis for the global burden of disease study 2016. *Lancet.* (2017) 390:1151–210. doi: 10.1016/S0140-6736(17)32152-9
- The World Bank. *The World by Income and Region.* (2018). Available online at: <http://datatopics.worldbank.org/world-development-indicators/the-world-by-income-and-region.html> (accessed July 19, 2020).
- Department of Economic and Social Affairs. *World Population Prospects 2019* (2017).
- Cheng J, Xie Y, Wickham H, Agafonkin V. *Leaflet: Create Interactive Web Maps With the Javascript 'Leaflet' Library.* Available online at: <https://cran.r-project.org/web/packages/leaflet/leaflet.pdf> (accessed Jan 24, 2017).
- Wickham H. *Ggplot2: Elegant Graphics for Data Analysis.* New York, NY: Springer-Verlag (2016). doi: 10.1007/978-3-319-24277-4_9
- Plotly Technologies Inc. *Collaborative Data Science.* (2015). Available online at: <https://plot.ly> (accessed May 25, 2020).
- Jonckheere AR. A distribution-free k-sample test against ordered alternatives. *Biometrika.* (1954) 41:133–45. doi: 10.1093/biomet/41.1-2.133
- The World Bank. *Life Expectancy at Birth, Total (Years).* (2019).
- Feigin VL, Roth GA, Naghavi M, Parmar P, Krishnamurthi R, Chugh S, et al. Global burden of stroke and risk factors in 188 countries, during 1990–2013: a systematic analysis for the global burden of disease study 2013. *Lancet Neurol.* (2016) 15:913–24. doi: 10.1016/S1474-4422(16)30073-4
- Budincevic H, Tiu C, Bereczki D, Korv J, Tsiskaridze A, Niederkorn K, et al. Management of ischemic stroke in central and eastern Europe. *Int J Stroke.* (2015) 10(Suppl.A100):125–7. doi: 10.1111/ij.s.12575
- GBD 2015 Risk Factors Collaborators. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: a systematic analysis for the global burden of disease study 2015. *Lancet.* (2016) 388:1659–24. doi: 10.1016/S0140-6736(16)31679-8
- van Asch CJ, Luitse MJ, Rinkel GJ, van der Tweel I, Algra A, Klijn CJ. Incidence, case fatality, and functional outcome of intracerebral haemorrhage over time, according to age, sex, and ethnic origin: a systematic review and meta-analysis. *Lancet Neurol.* (2010) 9:167–76. doi: 10.1016/S1474-4422(09)70340-0

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Phan, Haseeb, Beare, Srikanth, Thrift and Ma. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.