International Journal of Statistics and Systems
ISSN 0973-2675 Volume 18, Number 2 (2023), pp. 105-118

http://www.ripublication.com

Estimation of Survival Probability, Death Density and Hazard Rates for Cancer and Cardiovascular Patients using Mortality Indices

Abhijeet Jadhav and S. B. Munoli

Department of Studies in Statistics, Karnatak University, Dharwad, Karnataka, India.

abhijadhav10292@gmail.com, sbmunoli@yahoo.co.in

Corresponding Author: Abhijeet Jadhav, abhijadhav10292@gmail.com

Abstract

Background: The survival experience associated with cancer and Cardio Vascular diseases (CVD) is a subject of intense scrutiny as these two diseases lead to undergoing intense medical treatment and potentially reduced survival probability and enhanced hazard rate. A survival analysis is carried out to study the impact of age, gender on deaths among the Cancer patients and CVD patients of developed and developing countries.

Methods: Survival probability, death density and hazard rate are estimated, the impacts of age, gender, economies of their country on survival outcomes of cancer and CVD are studied.

Results: Survival probability of male CVD patients starts declining at the early ages of 40's for both developing and developed countries where as for females the decline in survival probabilities is observed from early 50's. For male and female cancer patients of both developed and developing countries, decline in survival probability is observed from the age of 50 and above.

Conclusions: Compared to CVD deaths, deaths due to cancer are less at the early ages of life. In both cancer and CVD, an elderly male is most likely to succumb. However, one should note the population proportions of males to females. The population hazard rate is considerably hyped by the hazard rates due to cancer and CVD.

Keywords: Cancer, CVD, Survival Probability, Death Density, Hazard Rate.

1. Introduction:

Cancer and CVD, though not communicable diseases affect a notable proportion of world population since they are identified as diseases. It is bane to mankind when speaking in terms of quality of life that these patients have to lead. The scientific community and health care providers worldwide have evidenced stark increase in number of patients or death due to these diseases inspite of technological advances in health science. The cornucopia of data on deaths due to these diseases is used to carryout survival analysis to derive meaningful inferences which can help in our concerted fight against these diseases. This paper aims to study the mortalities reported among Cancer and CVD patients of some developed and developing economies of the world through survival analysis approach.

Huge literature is available on the topics of risk assessment for various cancers, CVD's their treatments, therapies, role of comorbidities, etc. Few of the notable references that study the association between cancer and CVD, CVD and lifestyle risk factors, cancer recovery/ relapse and treatments are:

Ossewaarde, et. al. (2005), Agop, et. al. (2008), Janssen-Hiejnen, et. al. (2010), Amundsen, et. al. (2014), Kathleen, et. al. (2019), Penson, et. al. (2019), Li, et. al. (2020), Oren, et. al. (2020), Acevedo, et. al. (2022), Nierengarten (2022). A cohort study of conditional relative survival and competing mortality in lung cancer patients, use of restricted mean survival time in analysis of time-to-event data, comparison of mortality rates and life expectancy of hemophilia patients is made respectively by Royston and Parmar (2013), Hassan, et. al. (2021), Yoo, et. al. (2021).

2. Methodology

Sources of Data

The data were retrieved from https://platform.who.int/mortality/themes/theme-details/MDB/noncommunicable-diseases, a WHO Mortality Database which is a comprehensive and authoritative source for global mortality statistics. The developed economies considered are: USA, Australia, Germany and Canada; Brazil, Egypt, Mexico, Thailand are the considered developing economies. The latest available mortality data due to Cancer and CVD is for the year 2019.

The retrieved data insights (or fields) are:

Gender: Male, Female

Age groups: 0-1, 1-4, 5-9, 10-14, … ,80-84, 85+

Death rates due to Cancer and CVD per 100,000 population.

Some basic concepts of Survival Analysis as discussed in Singh and Mukhopadhyay (2011) are:

i. Survival Distribution Function, $S(t)$:

It gives the probability that an individual survives at least for time t. That is

$$S(t) = P(X > t), t \geq 0 \quad (1)$$

The function $S(t)$ is called "survival function".

$$S(0) = 1, S(\infty) = 0$$

ii. Death Density Function, $f(t)$:

With this function, one can get the probability that the "final event", namely death will occur in an infinitesimal small interval $(t, t + dt)$, this probability is given by $f(t)dt$. The function $f(t)$ is non-negative and satisfies

$$\int_0^\infty f(t)dt = 1 \tag{2}$$

The relationship between $f(t)$ and $S(t)$ is given by,

$$f(t) = -\frac{dS(t)}{dt} \tag{3}$$

iii. Hazard Rate, $h(t)$:

$h(t)dt$, a conditional probability defined as probability that death will occur in an infinitesimal small interval $(t, t + dt)$ given that the individual under consideration has survived up to time 't'. Hazard rate is also referred to as instantaneous death rate, force of mortality or intensity of death. This function is related to the two preceding probabilities by the relation

$$h(t) = \frac{f(t)}{S(t)}\ ; \quad S(t) > 0$$

$$= -\frac{d}{dt}\log S(t)$$

The plots of $(t, S(t))$, $(t, f(t))$, $(t, h(t))$ are respectively survival curve, death density curve and hazard rate curve. Although any of three functions $S(t), f(t)\ or\ h(t)$ defines uniquely a given survival distribution. Each function gives a different view of mortality data. For example, if the medical investigator wants to know when 50% of the individuals survive, he uses survival function or curve. On the other hand, if one is interested to know the peak age of death, one can use death density curves. The shape of the hazard curve indicates the type of risk one is exposed to. Evaluation or computation these probabilities for the patients suffering from a disease will help health care providers to estimate survival probabilities, test hypotheses, identification of risks, study of prognostic factors, etc. The insurers, also need the analysis of these functions while designing a insurance portfolio.

In order to evaluate $S(t), f(t)$ and $h(t)$ for both Cancer and Cardiovascular Disease (CVD) patients, we will utilize mortality data associated with these two diseases. The estimated $S(t), f(t)$ and $h(t)$ on the basis of number of deaths in the given age intervals are presented in Table 1 for CVD of Australia (Males). The computations of various columns of Table 1 are as follows:

i. The first two columns are self-explanatory (extracted from the website). The number of survivors at the beginning of each age interval (Column 3) is calculated by considering $l_0 = 1$ million. Subtracting number of deaths of that age interval from l_0 will lead to l_1, i.e., $l_1 = 10^6 - 71 = 999929$. On the same lines $l_5, l_{10}, \ldots, l_{85}$ are obtained.

ii. $S(0) = 1$, since all 1000,000 (1 million) individuals are alive at age '0'. $S(1)$, the probability that a particular individual survives at least one year is estimated as $999929/10^6$. $S(5)$, the probability that a individual survives at least 5 years is $999924/10^6$. Continuing this way, $S(10)$, $S(15)$ etc. are obtained and are presented in Column 4.

iii. The death densities at the mid-points of considered age-intervals of Australian CVD patients are: $f(0.5) = 71/10^6$, $f(2.5) = 5/10^6$, $f(7) = 4/10^6$, and are presented in Column 6.

iv. Finally in order to estimate hazard rates, number of deaths per unit time (Column 8) is divided by the number of survivors at the beginning of the interval in question (Column 3). That is; $h(1) = 71/10^6$ is the estimated hazard rate at the end of the first year, $h(4) = 1.25/999929$ is the estimated hazard rate at the end of the fourth year, $h(9) = 0.8/999924$, etc.

The survival probability $\left(S(t)\right)$ curve, death density $\left(f(t)\right)$ curve and hazard rate $(h(t))$ curve are plots of the points (lower bound of age interval, $S(t)$), (mid-point of age interval, $f(t)$), (upper bound of age interval, $h(t)$), respectively. Gender-specific survival analysis (computation $S(t), d(t), h(t)$) of CVD and Cancer patients of all considered countries is carried out.

3. Results

For the sake of brevity, computed survival probability, death density and hazard rates for male CVD patients of Australia only is presented in Table 1 and the same computations for female cancer patients of Egypt only is presented in Table 2. In Figures 1 and 2, gender specific estimated survival probability, death density, and hazard rates are plotted for Australian cancer and CVD patients respectively. The plots of computed survival probability for male and female patients of considered countries for CVD and Cancer patients are presented in Figures 3 and 4 respectively. Each of these two figures are self-explanatory in depicting the gender and economic differences (if any) of these two diseases. On the same lines the plots of death density and hazard rate are plotted and can be studied easily. From the study of estimated Survival probabilities, death densities and hazard rates for male and female patients of CVD and cancer of all the considered countries, following observations are made:

CVD:

The survival probabilities for female CVD patients is more for all age intervals of Thailand, Mexico, Egypt, Brazil, Australia, USA and Germany as compared to Survival probabilities of males of these countries whereas for Canada Male CVD patients survival probability is more as compared to female CVD patients. This pattern of higher survival probabilities for females and lower survival probabilities for males is in contrast to the relationship observed between gender and death density, as well as gender and hazard rates, in all of these countries.

Among the developing economies, males of Thailand has highest survival probability for 85 and above aged CVD patients, while Egypt has lowest survival probability for the same age group. Among developed economies, Canadian males in this age group have the highest survival probability, while the United States has the lowest survival probability. For developed economies, Australian females aged 85 and above have the highest survival probability, with other countries showing relatively equal survival probabilities. Among developing countries, females in the age group of 85 and above, Thailand boasts the highest survival probability, while Egypt finds itself at the opposite end with the lowest survival probability. Examining the hazard rate, we notice a distinct trend. It starts by decreasing during the early stages of life, then steadily increases until age 84, after which it declines again. For both males and females, survival probability (death density, hazard rate) decrease (or increase) at very slow rate up to the age of 50 and from the age of 50 and above a drastic decline (or incline) is observed.

Cancer:

Among the developing economies, males of Egypt have the highest survival probability for cancer patients aged 85 and above, while Brazil has the lowest survival probability for the same age group. Among developed economies, Canadian males in this age group have the highest survival probability, while Germany has the lowest survival probability. For developed economies, Australian females aged 85 and above have the highest survival probability, with other countries showing relatively equal survival probabilities. Among developing countries, females in the age group of 85 and above, Egypt boasts the highest survival probability, while Brazil finds itself at the opposite end with the lowest survival probability. And as we observed in CVD, other aspects remain consistent in the analysis of cancer survival probabilities, death density, and hazard rates.

4. Conclusions

In elderly population, incidence of cancer is more compared to younger population. Infants succumb to CVD more as compared to cancer. More females of age group 25-29, 30-34, …, 50-59 of countries Brazil, Mexico, Egypt, USA, Germany have cancer compared to males of these countries in respective age groups. To come out with a concrete reason for this trend, further studies need to be undertaken regarding life style, food habits, and other social practices prevailing in these countries. Economies of countries do not influence incidences of CVD and cancer, while age and gender are found to be influencing factors causing CVD and cancer. The novelty of the paper is use of death rates due to cancer (and CVD) in estimating respective survival probability, death density and hazard rate which play an important role in analyzing the influence of prognostic factors such as age, gender, economy in causing these diseases.

Limitations

One of the limitations of the study is that age-based CVD (or cancer) deaths per million population in 2019 are considered as a cohort. People who died in 2019 from cancer (or CVD) at various ages are considered as members of a cohort. The population composition by age is not taken into consideration. However, the study provides broad information about the amount of hype in hazard rates caused by these two diseases. From the data source, the prevalence of CVD among cancer deaths and vice versa is not known. CVD and cancer are identified as most dreadful diseases and hence a separate analysis for both is made.

Table 1:Survival Analysis of Australian Male CVD Patients (2019)

Age-group (1)	Number of deaths during the interval (2)	Number of Living at the beginning of the age interval (3)	Estimated Survival probability (4)	Mid-point (5)	Death Density (6)	Width of Class Interval (7)	Number of Deaths per Unit Time (8)	Estimated Hazard Rate (9)
0-1	71	1000000	1	0.5	0.000071	1	71	0.000071
01-04	5	999929	0.999929	2.5	0.000005	4	1.25	1.25009E-06
05-09	4	999924	0.999924	7	0.000004	5	0.8	8.00061E-07
10-14	7	999920	0.99992	12	0.000007	5	1.4	1.40011E-06
15-19	10	999913	0.999913	17	0.00001	5	2	2.00017E-06
20-24	13	999903	0.999903	22	0.000013	5	2.6	2.60025E-06
25-29	43	999890	0.99989	27	0.000043	5	8.6	8.60095E-06
30-34	76	999847	0.999847	32	0.000076	5	15.2	1.52023E-05
35-39	122	999771	0.999771	37	0.000122	5	24.4	2.44056E-05
40-44	294	999649	0.999649	42	0.000294	5	58.8	5.88206E-05
45-49	532	999355	0.999355	47	0.000532	5	106.4	0.000106469
50-54	765	998823	0.998823	52	0.000765	5	153	0.00015318
55-59	1275	998058	0.998058	57	0.001275	5	255	0.000255496
60-64	1742	996783	0.996783	62	0.001742	5	348.4	0.000349524
65-69	2634	995041	0.995041	67	0.002634	5	526.8	0.000529425
70-74	4103	992407	0.992407	72	0.004103	5	820.6	0.000826878
75-79	7435	988304	0.988304	77	0.007435	5	1487	0.001504598
80-84	14861	980869	0.980869	82	0.014861	5	2972.2	0.00303017
85-over	44651	966008	0.966008	92.5	0.044651	16	2790.6875	0.002888887

Table 2:Survival Analysis of Egypt Female CANCER Patients (2019)

Age-group (1)	Number of deaths during the interval (2)	Number of Living at the beginning of the age interval (3)	Estimated Survival probability (4)	Mid-point (5)	Death Density (6)	Width of Class Interval (7)	Number of Deaths per Unit Time (8)	Estimated Hazard Rate (9)
0-1	38	1000000	1	0.5	0.000038	1	38	0.000038
01-04	17	999962	0.999962	2.5	0.000017	4	4.25	4.25016E-06
05-09	14	999945	0.999945	7	0.000014	5	2.8	2.80015E-06
10-14	13	999931	0.999931	12	0.000013	5	2.6	2.60018E-06
15-19	18	999918	0.999918	17	0.000018	5	3.6	3.6003E-06
20-24	20	999900	0.9999	22	0.00002	5	4	4.0004E-06
25-29	30	999880	0.99988	27	0.00003	5	6	6.00072E-06
30-34	54	999850	0.99985	32	0.000054	5	10.8	1.08016E-05
35-39	92	999796	0.999796	37	0.000092	5	18.4	1.84038E-05
40-44	145	999704	0.999704	42	0.000145	5	29	2.90086E-05
45-49	216	999559	0.999559	47	0.000216	5	43.2	4.32191E-05
50-54	334	999343	0.999343	52	0.000334	5	66.8	6.68439E-05
55-59	493	999009	0.999009	57	0.000493	5	98.6	9.86978E-05
60-64	689	998516	0.998516	62	0.000689	5	137.8	0.000138005
65-69	974	997827	0.997827	67	0.000974	5	194.8	0.000195224
70-74	937	996853	0.996853	72	0.000937	5	187.4	0.000187992
75-79	1077	995916	0.995916	77	0.001077	5	215.4	0.000216283
80-84	976	994839	0.994839	82	0.000976	5	195.2	0.000196213
85-over	1153	993863	0.993863	92.5	0.001153	16	72.0625	7.25075E-05

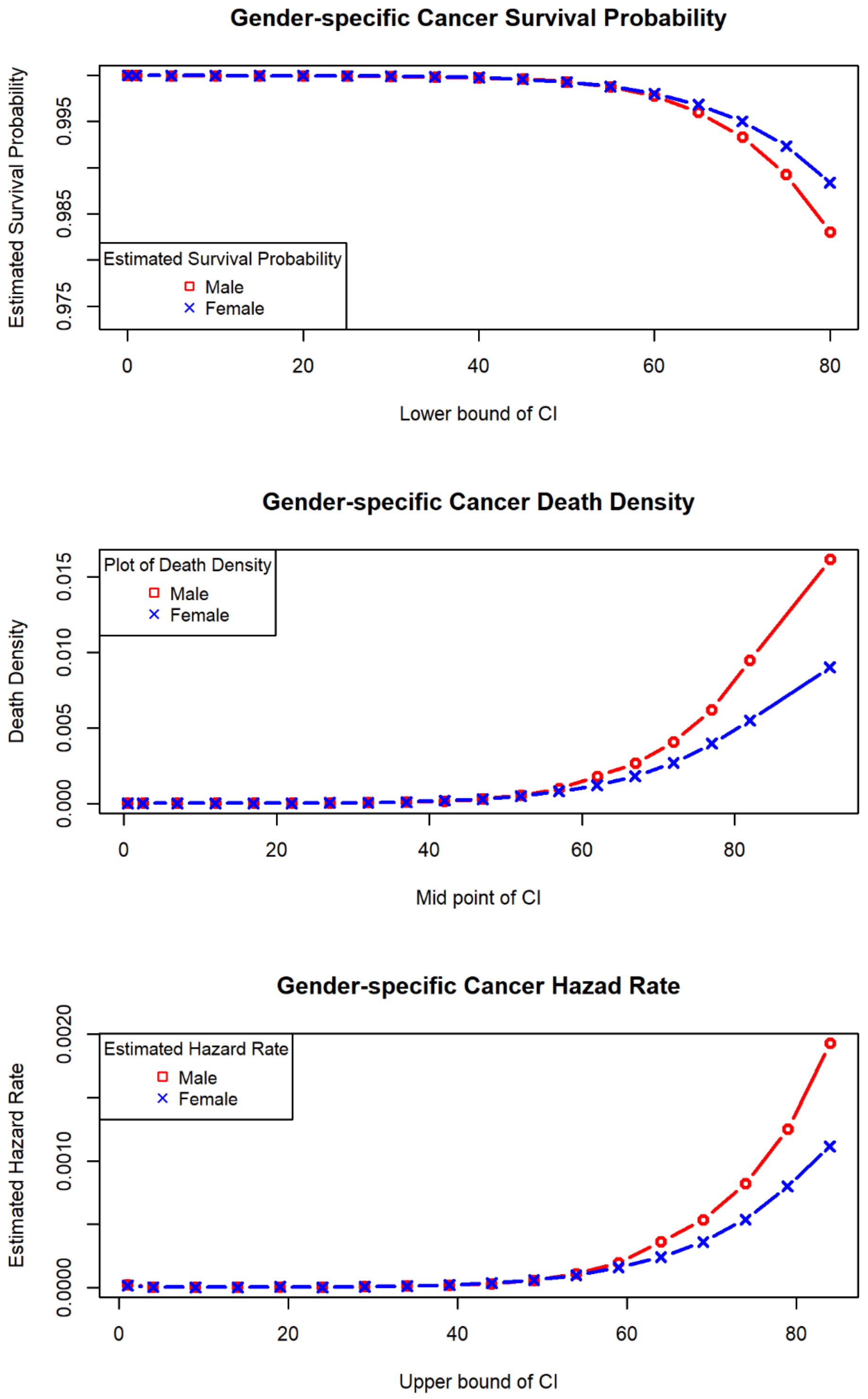

Figure 1. Survival Analysis of Australian Cancer Patients

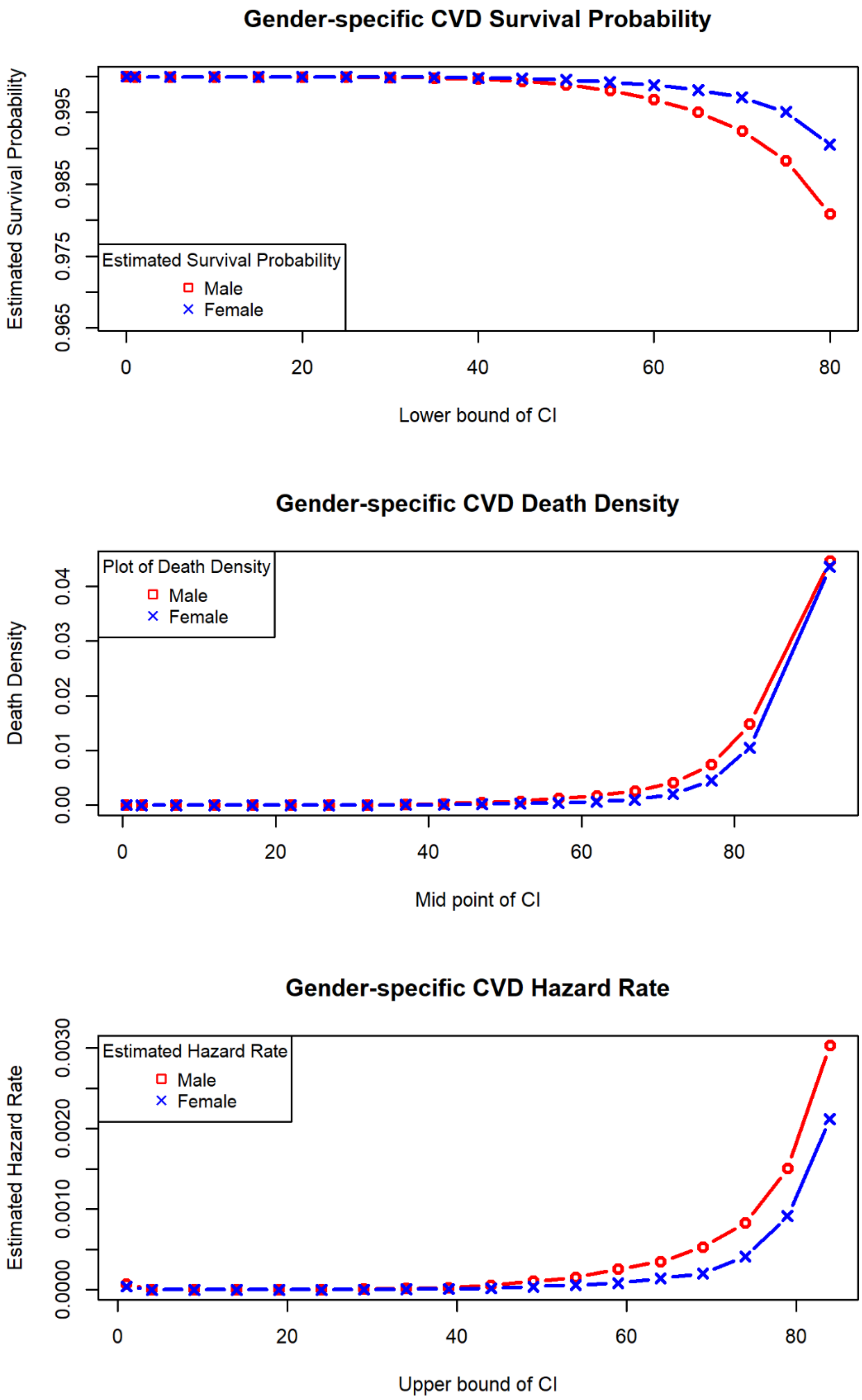

Figure 2. Survival Analysis of Australian CVD Patients

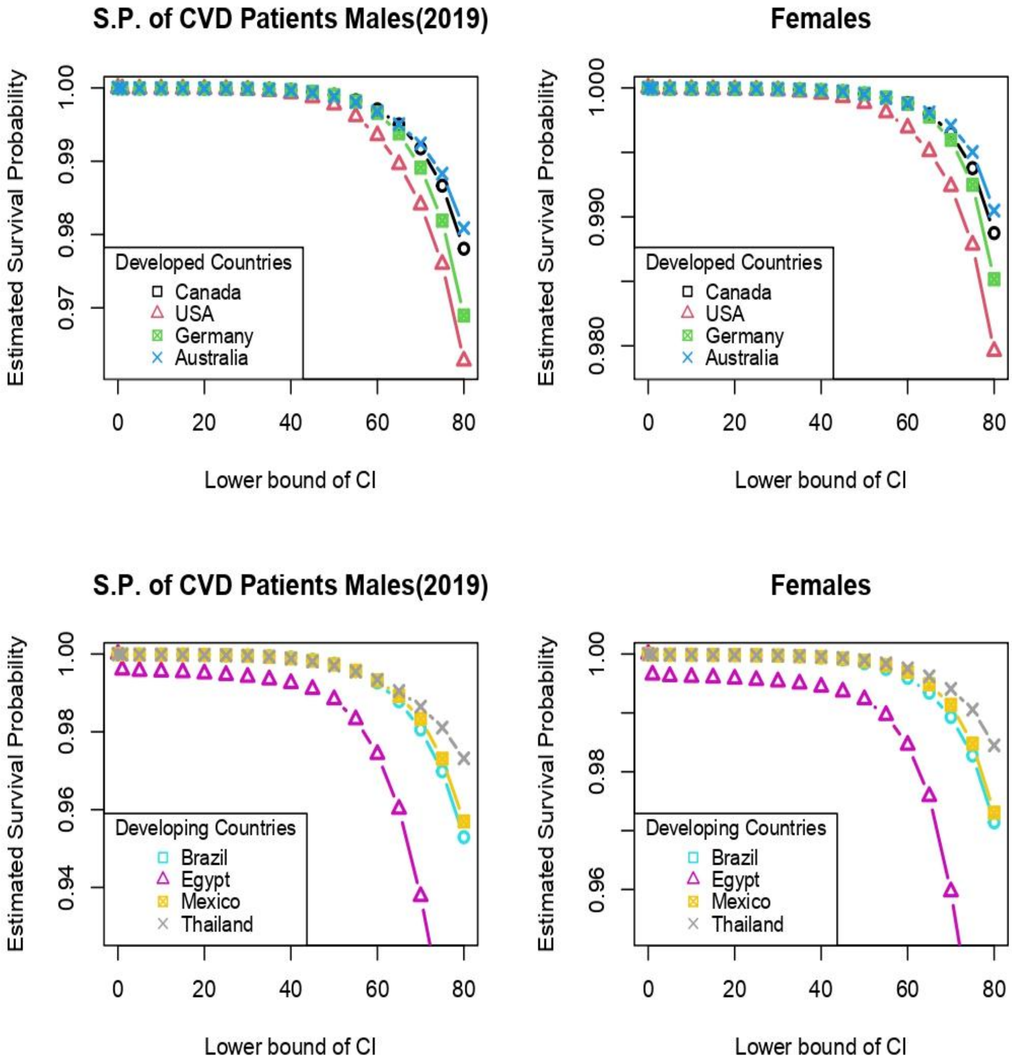

Figure 3. Gender-Based Survival Probability of Cardiovascular Disease (CVD) Patients in Developed and Developing Countries

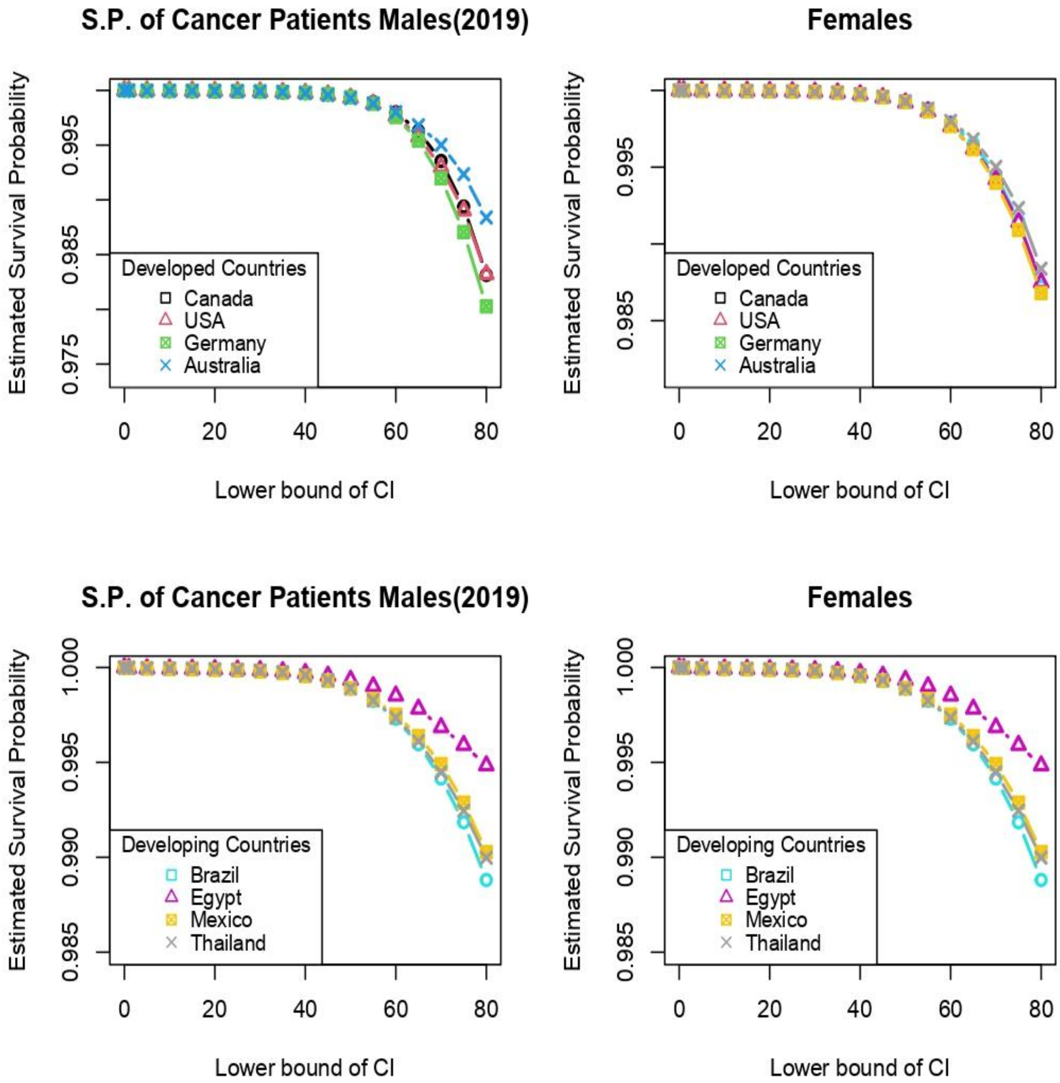

Figure 4. Gender-Based Survival Probability of Cancer Patients in Developed and Developing Countries

Acknowledgement: The first author is thankful to Karnatak University, Dharwad, Karnataka, India for financial support.

References

[1] Acevedo, F., Ip, T., Orellana, M., Martínez, G., Gabrielli, L., Andia, M., ... & Merino, T. 2022, “Oncological Benefit versus Cardiovascular Risk in Breast

Cancer Patients Treated with Modern Radiotherapy," *Journal of Clinical Medicine, 11*(13), pp. 3889. https://doi.org/10.3390/jcm11133889.

[2] Amundsen, T., Hatlen, P., & Hummel, T. M., 2014, "Survival in patients with lung cancer, with and without self-reported cardiovascular disease the HUNT study," *European Respiratory Journal, 44*(Suppl 58).

[3] Bedikian, A. Y., Johnson, M. M., Warneke, C. L., Papadopoulos, N. E., Kim, K., Hwu, W. J., ... & Hwu, P., 2008, "Prognostic factors that determine the long-term survival of patients with unresectable metastatic melanoma," *Cancer investigation, 26*(6), pp. 624-633. DOI:10.1080/07357900802027073.

[4] Hassan, S., Monahan, R. C., Mauser-Bunschoten, E. P., van Vulpen, L. F., Eikenboom, J., Beckers, E. A., ... & Gouw, S. C., 2021, "Mortality, life expectancy, and causes of death of persons with hemophilia in the Netherlands 2001–2018," *Journal of Thrombosis and Haemostasis, 19*(3), pp. 645-653.

[5] Janssen-Heijnen, M. L., Szerencsi, K., van de Schans, S. A., Maas, H. A., Widdershoven, J. W., & Coebergh, J. W. W., 2010, "Cancer patients with cardiovascular disease have survival rates comparable to cancer patients within the age-cohort of 10 years older without cardiovascular morbidity," *Critical reviews in oncology/hematology, 76*(3), pp. 196-207. https://doi.org/10.1016/j.critrevonc.2009.11.004.

[6] Kathleen M Sturgeon, Lei Deng, Shirley M Bluethmann, Shouhao Zhou, Daniel M Trifiletti, Changchuan Jiang, Scott P Kelly, Nicholas G Zaorsky, 2019, "A population-based study of cardiovascular disease mortality risk in US cancer patients," *European Heart Journal, Volume 40,* Issue 48, pp. 3889–3897, https://doi.org/10.1093/eurheartj/ehz766.

[7] Li, Y., Schoufour, J., Wang, D. D., Dhana, K., Pan, A., Liu, X., ... & Hu, F. B., 2020, "Healthy lifestyle and life expectancy free of cancer, cardiovascular

disease, and type 2 diabetes: prospective cohort study," *bmj,* pp. 368. doi: 10.1136/bmj.l6669|BMJ 2020;368:l6669.

[8] Nierengarten, M. B., 2022, "Cancer survivors have an increased risk of cardiovascular disease," *128*(21), pp. 3756, https://doi.org/10.1002/cncr.34501.

[9] Oren, O., Yang, E. H., Molina, J. R., Bailey, K. R., Blumenthal, R. S., & Kopecky, S. L., 2020, "Cardiovascular health and outcomes in cancer patients receiving immune checkpoint inhibitors," *The American journal of cardiology*, *125*(12), pp. 1920-1926.

[10] Ossewaarde, M. E., Bots, M. L., Verbeek, A. L., Peeters, P. H., van der Graaf, Y., Grobbee, D. E., & van der Schouw, Y. T., 2005, "Age at menopause, cause-specific mortality and total life expectancy," *epidemiology*, pp. 556-562.

[11] Penson, P., Long, L., Howard, G., Howard, V. J., Jones, S. R., Martin, S. S., ... & Banach, M., 2018, "Associations between cardiovascular disease, cancer and very low HDL cholesterol in the Reasons for Geographical And Racial Differences in Stroke (REGARDS) study," *Cardiovascular Research, 115*(1), pp. 204-212. DOI: 10.1093/cvr/cvy198. PMID: 30576432; PMCID: PMC6302258.

[12] Royston, P., Parmar, M.K., 2013, "Restricted mean survival time: an alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome. BMC Med Res Methodol," 13, 152. https://doi.org/10.1186/1471-2288-13-152

[13] Singh, R., & Mukhopadhyay, K., 2011, "Survival analysis in clinical trials: Basics and must know areas," *Perspectives in clinical research, 2*(4), pp. 145.

[14] Yoo, J. E., Han, K., Shin, D. W., Park, S. H., Cho, I. Y., Yoon, D. W., ... & Jung, K. W., 2021, "Conditional relative survival and competing mortality

in patients who underwent surgery for lung cancer: a nationwide cohort study," *International Journal of Cancer, 148*(3), pp. 626-636. DOI: 10.1002/ijc.33239.