The Gulf Journal of Oncology

GJO

Indexed By PubMed and Medline Database

Issue 41, January 2023
ISSN No. 2078-2101

The Official Journal of the Gulf Federation For Cancer Control

Table of Contents

Original Article

How does Linkage to the National Death Index Affect Population–Based Net Survival Estimates for Women with Cervical Cancer in Saudi Arabia?

Eman Alkhalawi[1,2], Claudia Allemani [2], Ali Saeed Al–Zahrani[3] , Michel P. Coleman[2]

[1] Department of Family and Community Medicine, King Abdulaziz University, Jeddah, Saudi Arabia.
[2] Cancer Survival Group, Department of Non–Communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, UK.
[3] Gulf Centre for Cancer Control and Prevention, King Faisal Special Hospital and Research Centre, Riyadh, Saudi Arabia.

Abstract

Background: Population–based cancer survival is a key metric for the assessment of cancer control strategies. Accurate estimation of cancer survival requires complete follow–up data for all patients.

Aim: To explore the impact of linking national cancer registry data to the national death index on net survival estimates for women diagnosed with cervical cancer in Saudi Arabia during 2005–2016.

Methods: We acquired data on 1,250 Saudi women diagnosed with invasive cervical cancer during the 12–year period 2005–2016 from the Saudi Cancer Registry. These included the woman's last known vital status and the date of last known vital status, but this was restricted to information from clinical records and death certificates that mention cancer as a cause of death ("registry follow–up"). We submitted available national ID numbers to the National Information Center (NIC) of the Ministry of Interior, to ascertain the date of death, from any cause of death, for women who had died up until 31 December 2018 ("NIC follow–up"). We estimated age–standardised 5–year net survival using the Pohar–Perme estimator under five different scenarios using the two sources of follow–up, and censoring at the date of last contact with the registry versus extending survival until the closing date if no information on death was obtained.

Results: 1,219 women were eligible for survival analysis. Five–year net survival was lowest when using NIC follow–up only (56.8%; 95%CI 53.5 – 60.1%), and highest when registry follow–up only was used and survival time was extended until closure date for those with no information on death (81.8%; 95%CI 79.6 – 84%).

Conclusion: Reliance solely on information from deaths certified as due to cancer and clinical records leads to a high proportion of missing deaths in the national cancer registry. This is probably due to low quality of certification of the cause of death in Saudi Arabia. Linkage of the national cancer registry to the national death index at the NIC identifies virtually all deaths, providing more reliable survival estimates, and it eliminates the ambiguity in determining the underlying cause of death. Therefore, this should become the standard approach to estimating cancer survival in Saudi Arabia.

Keywords: cancer registry, cervical cancer, survival, Saudi Arabia

Background:

Population–based cancer survival is a key metric in assessing the overall performance of the health services in managing cancer, including timely diagnosis and referral, and adequate access to treatment. It also enables countries to monitor the impact of policy changes. Population–based cancer survival differs from clinical trials, which aim to assess the direct effect of interventions through applying strict inclusion criteria and rigorous protocols, and are often carried out in high–resource tertiary care centres. Therefore, their results do not reflect the experience of the general population.

Corresponding author: Dr. Eman Alkhalawi, Department of Family and Community Medicine, King Abdulaziz University, Jeddah, Saudi Arabia. P.O. Box : 80200, Zip Code : 21589, Jeddah 23617–7174, Saudi Arabia Email: ealkhalawi@kau.edu.sa

High–quality population–based cancer registries, which record every cancer diagnosis in a defined country or region, provide the backbone for monitoring cancer incidence and survival. But accurate estimation of cancer survival also requires complete follow–up data for all registered patients.

In population–based cancer survival, we are interested in isolating the hazard of death due to cancer. This can be achieved by two different means: estimating survival in the cause–specific or the relative survival setting. In the former, deaths due to cancer are the outcome of interest while deaths due to any other cause are censored, and survival can be estimated using the Kaplan–Meier method. In the relative survival setting, all deaths are counted as an event, and survival is derived from the hazard of death in the population with the cancer of interest *relative* to the age– and sex–matched hazard in the general population from which they come. If a person's vital status cannot be verified, they are considered lost to follow–up and their survival time is censored at the date they were last known to be alive.

Estimating survival in the cause–specific setting relies on accurate attribution and coding of the underlying cause of death. Often, the cause of death among cancer patients is subjective, due to the complex nature of the disease, which is frequently accompanied by serious comorbidity or complications of treatment. This can affect the accuracy of cancer mortality rates, [1–6] and it renders the cause–specific setting less reliable, especially for international comparisons of survival. In Saudi Arabia, less than 30% of causes of death during the period 2010–2017 were considered well–certified according to the Global Burden of Disease. [7]

Deriving survival estimates from excess mortality, that is, in the relative survival setting, overcomes this issue, by comparing the survival of the cancer patients with that of the general population from which they come, without needing the cause of death, which may be inaccurate or inaccessible. Net survival is the survival probability up until a given time since diagnosis, after correcting for background mortality and the increasing risk of death at older ages. This is given by life tables of all–cause mortality in the general population by age, sex and calendar year. It can be estimated with the Pohar–Perme estimator. [8] Net survival estimates are independent of competing causes of death, and they enable monitoring survival over time and between populations.

The Saudi Cancer Registry (SCR) is a population–based registry with national coverage which has been operating since 1994. It has legislative access to all cancer records in hospitals, clinics and laboratories. Active case–finding is carried out by certified tumour registrars, relying primarily on pathology reports. Other sources include cytology, radiology, radiotherapy, haematology and outpatient departments. Registrars regularly search through hospital death notifications for those issued with a mention of cancer. Death certificates with a mention of cancer have been acquired periodically from the regional departments of vital statistics since 2005. The latter two sources are used to update the vital status of registered patients and to capture cases that have been missed during the patient's lifetime. If there is no record of that person in the registry, traceback to find a hospital record to confirm or refute the diagnosis is initiated after six months. For each registered patient, the registry keeps a record of the date of last contact and the vital status on that date. This is initially set as the date of abstraction, and is updated when more recent data are obtained. If the registry receives information on the death of a patient, the date of death is updated as the date of last contact and the vital status is updated as dead, either due to cancer or due to another cause. In most cases, however, this information is obtained from a death certificate with a mention of cancer.

We aimed to explore the impact on net survival estimates of linking national cancer registry data to the national death index, in order to obtain more complete follow–up. We have done this for women diagnosed with cervical cancer in Saudi Arabia during 2005–2016, as an example. We compared the estimates to those obtained using only the data on vital status available to the registry, mostly deaths certified as due to cancer. We also explored the effect on survival estimates of censoring follow–up time at the date women were last known to be alive to the registry versus extending survival time until the end of the study period, in order to determine whether a patient who was last known to be alive at a date earlier than the closing date should be considered lost to follow–up, or assumed to be alive if no information on her death was received.

Methods:

Data sources

We acquired data from the Saudi Cancer Registry for 1,250 Saudi women diagnosed with invasive cervical cancer during the 12–year period 2005–2016. The registry supplied, among other variables, women's dates of birth, dates of diagnosis and follow–up data including last known vital status (alive, dead, not known), date of last contact, and cause of death for those known to have died ("registry follow–up").

Registry staff then submitted available national ID numbers to the National Information Center (NIC) of the Ministry of Interior, to ascertain the date of death, from any cause of death, for women who had died up until the date of record linkage, 29th of August 2019 ("NIC follow–up"). In order to verify the identity of the women, name, sex and date of birth were also requested. Registry staff verified women's identities by manually comparing full names obtained from the NIC to names in the registry (first name, father's name, grandfather's name, family name as recorded in Saudi Arabia's official records). Complete 10–digit national ID numbers were available for 928 women (76%).

For data quality assurance, we used the eligibility criteria of the CONCORD programme for the global surveillance of cancer survival. [9] Eight women were considered ineligible for these survival analyses: four because they were younger than 20 or aged 100 years or more at diagnosis, and four for having ineligible morphology codes. Further, eight women were excluded because of site–morphology mismatch, one because her date of diagnosis preceded her date of birth, and 14 because they were death certificate only registrations, where no earlier date of diagnosis had been found. This resulted in 1,219 women being available for the analysis.

We obtained life tables of all–cause mortality rates for Saudi Arabia by single calendar year, sex and 5–year age group from the United Nations Population Division. [10] We interpolated them and extended them to age 99 years using the Elandt–Johnson method in order to obtain mortality rates by single year of age. [11, 12] Life tables are available online from the Cancer Survival Group website. [13]

Ethical approval was obtained from the London School of Hygiene and Tropical Medicine.

Statistical analysis

We estimated survival for women under five scenarios:

1. Restricting the analysis to women with complete NIC follow–up (n=928).
2. Using NIC follow–up for women for whom it was available (n=928), complemented with registry follow–up for women with no available ID number (n=291), while censoring their survival time at their date of last known vital status.
3. Using NIC follow–up for women for whom it was available, complemented with registry follow–up for women with no available ID number, and extending their survival time until the end of follow–up (i.e., assuming they had remained alive to that date if no information on their death was obtained).
4. Using registry follow–up only for all women, while censoring at the date of last known vital status.
5. Using registry follow–up only for all women, and extending their survival time until the closing date (i.e., assuming they had remained alive to that date if no information on their death was obtained).

Except for the first scenario, where the analysis is restricted to the subset of 928 women with complete NIC follow–up, all 1,219 women were included in each analysis.

We estimated 5–year net survival using the Pohar–Perme estimator, and women with NIC follow–up were considered alive until the closing date if no information on their death had been received. Survival time for women diagnosed during 2014–2016 was censored on 31 December 2018, the end of the last complete calendar year preceding the date of record linkage.

Estimates were not age–standardised because the sole aim was an internal comparison, using the same study population.

Results

1,219 women (97.5% of those eligible) were included in the survival analyses, with a mean age of 52.8 years (SD 14.5, range 20 – 93). Of 928 women with NIC follow–up, 411 (44.3%) were found to have died during the follow–up period. This was more than double the proportion reported to have died by registry follow–up (18.3%, n=223). Of these, 216 (96.9%) deaths were reported as being due to cancer. Only three women (0.25%) were considered to have an unknown vital status in the registry records, while the remaining 933 (81.5%) were alive at the date of last contact.

Women with NIC follow–up were slightly younger on average, were more likely to be diagnosed at a regional and distant stage and at grade 2 and 3, but less likely to have an unknown stage and grade (Table 1).

5–year net survival was very similar when restricting to NIC follow–up (scenario 1) and when including women with registry follow–up only, censoring their survival time at the date they were last known to be alive (scenario 2). However, when assuming all women were alive until the closing date if no information on their death was obtained during that period (scenario 3), the 5–year net–survival estimate increased from 57.2% (95%CI 54.0–60.4) to 64.3% (95%CI 61.5–67.0) (Table 2)

When using deaths captured by the registry, and censoring the survival time at the date the woman was last known to be alive to the registry (scenario 4), the 5–year net survival estimate was slightly higher than the that derived using NIC follow–up (scenario 2), but had large confidence intervals, and was unstable over the 5 years examined (Figure 1).

Extending the follow–up time until the closing date for women who were not known to be dead in the registry increased the 5–year survival by more than 20% (scenario 5).

Discussion:

We explore 5–year net survival for women diagnosed with cervical cancer in Saudi Arabia using two available sources of follow–up (complete follow–up by linkage to the national death index and registry follow–up), and for the latter source, examine the assumption that women for whom no information on death was available are alive at least until the end of follow–up.

Available follow–up	Registry only		NIC		Total	
	Mean	SD	Mean	SD	Mean	SD
Age	53.7	14.1	52.5	13.3	52.8	13.5
	N	%	N	%	N	%
Stage						
Localised	108	37.1	270	29.1	378	31.0
Regional	82	28.2	398	42.9	480	39.4
Distant	43	14.8	180	19.4	223	18.3
Unknown	58	19.9	80	8.6	138	11.3
Grade						
I	30	10.3	68	7.3	98	8.0
II	83	28.5	313	33.7	396	32.5
III	71	24.4	272	29.3	343	28.0
IV	13	4.5	28	3.0	41	3.4
Unknown	94	32.3	247	26.6	341	28.0
Basis of diagnosis						
Pathology	287	98.6	920	99.1	1,207	99.0
Clinical/ imaging	1	0.3	3	0.3	4	0.3
Unknown	3	1.0	5	0.5	8	0.7

Table 1: Characteristics of women included in the survival analysis (n=1,219).

While the source of follow–up did not have a large effect on the 5–year net survival estimate, using only registry follow–up led to a large loss in total follow–up time, and consequentially, in precision. However, assuming women remained alive if no information of their death was received led to a substantial overestimation of survival.

A large proportion of loss to follow–up can bias survival estimates if the reason for a patient being lost to follow–up is related to their risk of dying (e.g., related to age, stage or comorbidity). Mean age was similar for women who did and did not have NIC follow–up, while women with only registry follow–up were slightly more likely to be diagnosed at a regional or distant stage and with grade 2 and 3 tumours. However, these women also had more unknown stage and grade. The registry does not collect data on comorbidity. It is therefore difficult to determine whether there were true differences in women's characteristics that could lead to biased survival estimates.

Despite the proportion of women who were dead by registry follow–up being much lower than those dead by NIC follow–up, the survival estimates when including women with only registry follow–up and censoring their survival time at the date of last contact (scenario 2) were similar to those obtained from NIC follow–up alone (scenario 1). This is probably due to these women contributing minimally to the total follow–up time. Although women with only registry follow–up make up 24% of the study population, they contributed less than 6% of the follow–up time and 8.5% of deaths. Similarly, using only registry follow–up led to a minimal increase (2%) in the survival estimate when censoring at the date of last known vital status (scenario 4). Survival estimates reflect both the number of individuals dying during the study period and their follow–up time during which they are at risk. In the case of registry follow–up, both the total follow–up time and number of deaths were much lower, which led to a less smooth survival function and

Follow–up scenario	n	Analysis time	Failures	5–year net survival	95% confidence intervals
1: NIC follow–up only	928	4,456	411	56.8	53.5 – 60.1
2: NIC follow–up complemented with registry follow–up, censoring at date of last contact	1,219	4,728	449	57.2	54.0 – 60.4
3: NIC follow–up complemented with registry follow–up, extending survival until closure date for those with no information on death	1,219	6,672	449	64.3	61.5 – 67.0
4: Registry follow–up only, censoring at date of last contact	1,219	1,693	223	58.5	51.7 – 65.4
5: Registry follow–up only, extending survival until closure date for those with no information on death	1,219	7,800	223	81.8	79.6 – 84.0

Table 2: 5–year net survival estimates for women diagnosed with invasive cervical cancer in Saudi Arabia during 2005–2016

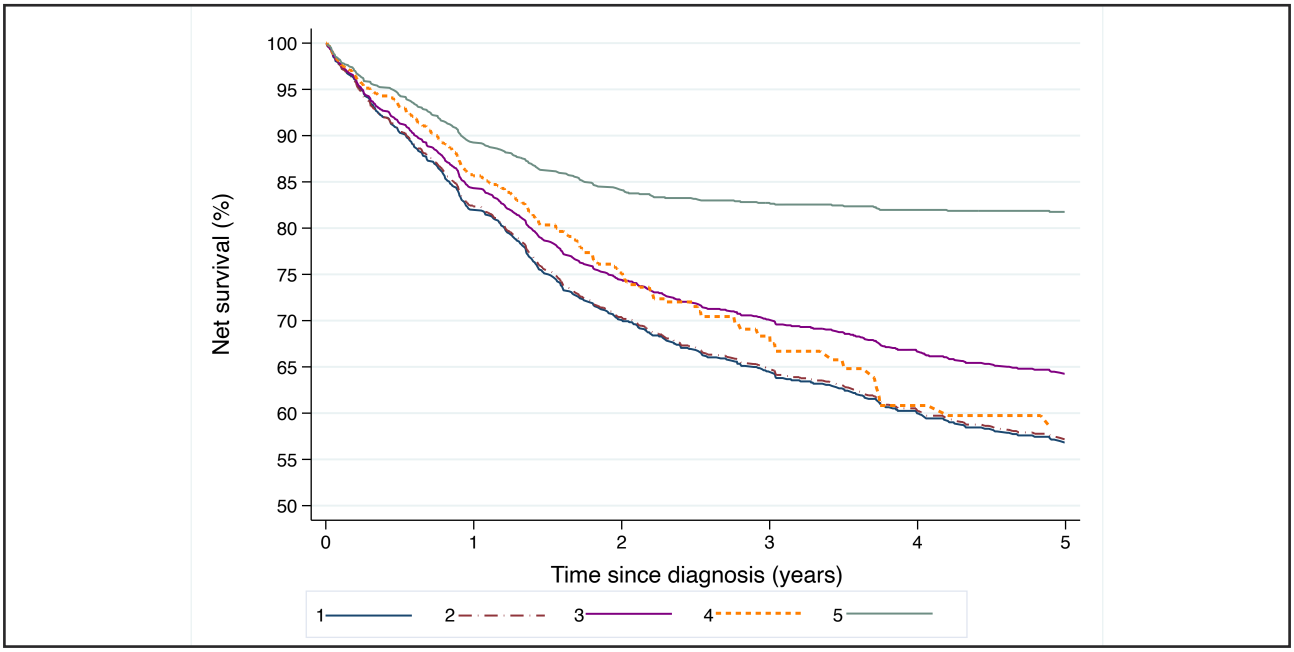

Figure 1: 5–year net survival for women diagnosed with invasive cervical cancer in Saudi Arabia during 2005–2016 (n=1,219) under 5 follow–up scenarios

1: NIC follow–up only, censoring at date of last contact
2: NIC follow–up complemented with registry follow–up, censoring at date of last contact
3: NIC follow–up complemented with registry follow–up, extending survival until closure date for those with no information on death
4: Registry follow–up only, censoring at date of last contact
5: Registry follow–up only, extending survival until closure date for those with no information on death

large confidence intervals. The mean follow–up time for these women was short (1.6 years). Given the high level of censoring, this leads to large uncertainty.

Assuming that women remained alive beyond the date of last contact if no information on their death was received by the registry led to an overestimation of survival (scenario 3 and 5). There was a 25% absolute difference in 5–year survival between the estimate that used only complete death ascertainment (scenario 1) and that relying solely on deaths due to cancer that were captured by the registry (scenario 5). This clearly shows that registry follow–up did not capture all deaths due to cancer. This is expected given the weak cause of death attribution in Saudi Arabia and difficulty in obtaining all death notifications with a mention of cancer which requires extensive manual effort.

In a previous analysis of SEER (Surveillance, Epidemiology and End Results) data in the US, the authors compared the effect of varying degrees of incomplete death ascertainment and follow–up. [14] Starting with a dataset which contained both complete death ascertainment (through linkage to the national death index) and complete follow–up (through active follow–up and linkage to state vital statistics records), the authors randomly assigned different proportions of patients found to be deceased by death ascertainment an “alive” vital status, and randomly assigned earlier dates of last contact to patients who are known to be alive through follow–up. Their results showed that ascertainment of death only is sufficient to obtain accurate survival estimates, but that this was more sensitive to missed deaths. A 10% “missingness” of deaths resulted in a 3.1% absolute overestimation in survival (death ascertainment only scenario), while 10% loss to follow–up resulted in a 0.64% decrease in survival. This differed by cancer location, but the reduction was more pronounced for cancers with intermediate survival.

Limitations:

Net survival may be subject to bias if life tables are inappropriate for the cancer cohort under study. However, net survival estimates have been shown to be less sensitive to bias due to inaccuracies in life tables than survival estimates in the cause–specific setting with errors in attribution of the cause of death. [15]

We considered linkage to the NIC dataset as the “gold standard”. However, incomplete or delayed death reporting cannot be ruled out.

Conclusion

Optimal estimation of net survival requires confidence in the completeness of follow–up for vital status. In Saudi

Arabia, linkage of cancer registry records to the National Death Index provides this confidence, but that is only possible when the patients' national ID numbers are available.

Where ID numbers are not available, the date of last known vital status should be taken as the date of death, if known, or the date of the patient's last clinical encounter. If patients without an ID number who are not known to be dead are assumed to be alive until the end of follow–up, survival will be overestimated.

Passive follow–up through linkage to the National Death Index provides robust estimates of population–based net survival, because it enables complete follow–up for vital status, and it avoids reliance on accurate attribution of the cause of death. It also requires fewer resources than for active follow–up, in which the last known vital status must be established by direct contact with the physician, the hospital or even directly with the patient. However, active follow–up is indispensable for survivorship research, for which cancer survivors need to be asked about their quality of life.

These results have important policy implications in Saudi Arabia, and in other countries where linkage to a national death index is not yet readily available for the cancer registry or for researchers. Legislation allowing the cancer registry to update the vital status of registered patients on a routine basis would facilitate timely and accurate monitoring of cancer survival.

Acknowledgement

We thank the Saudi Cancer Registry staff, including the cancer registrars for their hard work and for sharing their knowledge and experience. We especially acknowledge Ms Wafaa Al–Zahrani's effort in obtaining data on vital status from the National Information Center and verifying them.

Funding and Conflict of Interest

The authors have no conflict of interest to disclose. This work was funded by a scholarship from the Saudi Arabian Ministry of Education.

References

1. Percy C, Ries LG and Van Holten VD. The accuracy of liver cancer as the underlying cause of death on death certificates. Public Health Rep 1990; 105: 361–367.
2. Percy CL, Miller BA and Gloeckler Ries LA. Effect of changes in cancer classification and the accuracy of cancer death certificates on trends in cancer mortality. Annals of the New York Academy of Science 1990; 609: 87–97. DOI: 10.1111/j.1749–6632.1990.tb32059.x.
3. Percy CL and Muir CS. The international comparability of cancer mortality data: results of an international death certificate study. Am J Epidemiol 1989; 129: 934–946.
4. Percy CL, Stanek E and Gloeckler Ries LA. Accuracy of cancer death certificates and its effect on cancer mortality statistics. Am J Public Health 1981; 71: 242–250.
5. Pavillon G, Coleman MP, Johansson LA, et al. Coding of causes of death in European Community: a report to INSERM. 1998. Luxembourg: Statistics Office of the European Communities.
6. Coleman MP and Aylin P. Death certification and mortality statistics: an international perspective (Studies on Medical and Population Subjects No. 64). In: Coleman MP and Aylin P, (eds.). London: The Stationery Office, 2000, p. 58.
7. GBD Causes of Death Collaborators. Global, regional, and national age–sex–specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018; 392: 1736–1788. 2018/11/30. DOI: 10.1016/S0140–6736(18)32203–7.
8. Pohar–Perme M, Stare J and Estève J. On estimation in relative survival. Biometrics 2012; 68: 113–120. 2011/06/22. DOI: 10.1111/j.1541–0420.2011.01640.x.
9. Allemani C, Matsuda T, Di Carlo V, et al. Global surveillance of trends in cancer survival 2000–14 (CONCORD–3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population–based registries in 71 countries. The Lancet 2018; 391: 1023–1075. DOI: 10.1016/S0140–6736(17)33326–3.
10. UN Population Division. World population prospects: the 2012 revision. Volume I: comprehensive tables. . 2013. New York: UN Department of Economic and Social Affairs.
11. Elandt–Johnson RC, Johnson NL. *Survival models and data analysis (Wiley Series in Probability and Mathematical Statistics)*. Indianapolis: John Wiley & Sons, Inc.; 1980.
12. Spika D, Bannon F, Bonaventure A, et al. Life tables for global surveillance of cancer survival (the CONCORD programme): data sources and methods. BMC Cancer 2017; 17: 159. 2017/03/01. DOI: 10.1186/s12885–017–3117–8.
13. life tables from the CONCORD programme, http://csg.lshtm.ac.uk/life–tables/ (accessed 4 May 2021).
14. Johnson CJ, Weir HK, Yin D, et al. The impact of patient follow–up on population–based survival rates. J Registry Manag 2010; 37: 86–103. 2011/04/06.
15. Schaffar R, Rachet B, Belot A, et al. Estimation of net survival for cancer patients: Relative survival setting more robust to some assumption violations than cause–specific setting, a sensitivity analysis on empirical data. Eur J Cancer 2017; 72: 78–83. DOI: https://doi.org/10.1016/j.ejca.2016.11.019.