Arch Iran Med. January 2019;22(1):1-6

http www.aimjournal.ir

ARCHIVES OF IRANIAN MEDICINE

Open Access

Original Article

Completeness and Accuracy of Death Registry Data in Golestan, Iran

Susan Hasanpour-Heidari, MSc[1]; Nastaran Jafari-Delouei, MSc[1]; Nesa Shokoohifar, MSc[1]; Seyed Mehdi Sedaghat, MD[2]; Abbas Moghaddami, MSc[2]; Reza Hosseinpour, MSc[2]; Mohammad Poorabbasi, MD[3]; Masoomeh Gholami, PhD[3]; Shahryar Semnani, MD[1]; Mohammad Naeimi-Tabiei, MD[4]; Mohammad Reza Honarvar, PhD[2]; Abdolreza Fazel, MD[5]; Arash Etemadi, PhD[6,7]; Freddie Bray, PhD[8]; Gholamreza Roshandel, MD, PhD[1*]

[1]Golestan Research Center of Gastroenterology and Hepatology, Golestan University of Medical Sciences, Gorgan, Iran
[2]Deputy of Public Health, Golestan University of Medical Sciences, Gorgan, Iran
[3]Death Registry Unit, Deputy of Public Health, Golestan University of Medical Sciences, Gorgan, Iran
[4]Cancer Research Center, Golestan University of Medical Sciences, Gorgan, Iran
[5]Omid Cancer Research Center, Omid Preventive Medicine and Health Promotion Center, Golestan University of Medical Sciences, Gorgan, Iran
[6]Division of Cancer Epidemiology & Genetics, National Cancer Institute, Bethesda, MD, USA
[7]Digestive Oncology Research Center, Digestive Disease Research Institute, Tehran University of Medical Sciences, Shariati Hospital, Tehran, Iran
[8]Section of Cancer Surveillance, International Agency for Research on Cancer, Lyon, France

Abstract

Background: We aimed to evaluate completeness and accuracy of the Golestan Death Registry (GDR) to identify cancer-related causes of death (CCoD).

Methods: The GDR data (2004-2015) were compared with cancer data collected from clinical/pathological sources (the considered gold standard) by the Golestan Population-Based Cancer Registry (GPCR). Using a linkage method, matched cases, including subjects with CCoD and those with ill-defined cause of death (ICoD) (garbage codes), were identified and entered into the final analysis as study subjects. The completeness (percentage of study subjects with CCoD) and accuracy (number of subjects with correct CoD from the total number of study subjects) of the GDR were calculated.

Results: In total, 3,766 matched cases were enrolled. Overall, the completeness and accuracy of the GDR for identifying CCoD were 92.7% and 53.2%, respectively. There were variations by cancer site and age group, with completeness and accuracy highest for brain cancer (96.3%) and leukaemia (79.8%) while the lowest accuracy was observed for colorectal cancer (29.9%). The completeness and accuracy of GDR was higher in patients aged under 60 years (95.7% and 53.6%, respectively). We also found higher completeness (93.7%) and accuracy (55.8%) in residents of rural areas.

Conclusion: Linkage of death registry data with cancer registry data can be a significant resource for evaluating quality of the death registry data. Our findings suggested that completeness of the GDR for identifying CCoD is reasonable, but accuracy is relatively low. Access to clinical and pathological data from other sources and enhanced training of death certifiers can improve the present situation.

Keywords: Accuracy, Cancer, Completeness, Death registry, Iran

Cite this article as: Hasanpour-Heidari S, Jafari-Delouei N, Shokoohifar N, Sedaghat SM, Moghaddami A, Hosseinpour R, et al. Completeness and accuracy of death registry data in Golestan, Iran. Arch Iran Med. 2019;22(1):1–6.

Received: February 7, 2018, Accepted: December 11, 2018, ePublished: January 1, 2019

Introduction

Cancer is the third leading cause of death (CoD) in Iran, and cancer control programmes are priorities in the national health system. Robust cancer mortality data is essential for public health surveillance and providing necessary information on the burden of cancer for planning and evaluation purposes.[1-3] Accuracy and completeness are important indicators of cancer data quality.[4-6] Comparison of mortality records with other reliable and accurate information sources, especially clinical/pathological records, is a common strategy used in assessing the overall quality of death data. This can be done routinely linking death certificate information to relevant medical records available in hospital and laboratory records or to disease registries (e.g. cancer registries).[7-10]

The Golestan province, located in Northern Iran, has been known as a high-risk area for upper gastrointestinal cancers since the 1970s.[11,12] Information about deaths, including cancer-related mortality, is registered using official death certificates by the Golestan Death Registry (GDR). The GDR Secretariat is located with the Deputy of Health at the Golestan University of Medical Sciences (GOUMS). Accurate and reliable information on cancer patients in Golestan province are routinely collected by the Golestan

*Corresponding Author: Gholamreza Roshandel, MD; Sayyad Shirazi Hospital, Sayyad Shirazi Blvd., Gorgan, Iran. Tel: +98 911 375 7327; Email: roshandel_md@yahoo.com

Population-Based Cancer Registry (GPCR).[13,14] The GPCR collects cancer data from all diagnostic and therapeutic centers (including pathology centers and hospitals) throughout the Golestan province. The GPCR is considered to be of high quality and taken as the gold standard in this study, having been included in the International Agency for Research on Cancer's (IARC's) *Cancer Incidence in Five Continents* series[15]; it is also a voting member of the International Association of Cancer Registries (IACR) since 2007. The overarching aim is to assess the accuracy and completeness of GDR for identifying cancer-related deaths by linking GDR to the clinical and pathology cancer data available at the GPCR.

Materials and Methods

This cross-sectional study was conducted on data registered in the GDR and GPCR from 2004 to 2015. Data on CoD is routinely collected by the GDR using the International Statistical Classification of Diseases and Related Health Problems (ICD-10) coding system,[16,17] while the GPCR collects information on cancer diagnosis using the International Classification of Diseases for Oncology (ICD-O).[18] In the first phase, the GDR data were matched with clinical and pathological cancer data registered in the GPCR from 2004–2015. GPCR cases for which the diagnosis of cancer was confirmed by clinical or pathological methods were included in this study, and the linkage method was used to identify matched cases between the two datasets on the basis of patients' demographic information (national code, first name, last name and father name). In the next phase, matched cases with specified non-cancer CoD were excluded from the study, and then GDR matched cases with cancer-related CoD (CCoD), as well as those with ill-defined CoD (ICoD) (garbage codes), were recruited as our study subjects and were entered into the final analysis. In this analysis, completeness was defined as proportion of CCoD covered by the GDR and accuracy was defined as proportion of assigning correct CoD by the GDR.

Completeness analysis explored the proportion of cases that the GDR could not assign a CCoD code for cancer cases and classified them as ICoD. That means GDR could not cover and include these cases in the cancer mortality dataset, resulting in lack of its coverage or completeness in identifying CCoD. Therefore, for calculation of GDR completeness, we divided the number of cases with CCoD by the total study subjects (the sum of CCoD and ICoD). In other words, the level of completeness of GDR was calculated using the formula: number of CCoD*100/total study subjects.

The aim of accuracy analyses was to identify the proportion of study subjects for whom a correct cancer code was assigned by the GDR. At first, the ICD-O codes of the GPCR dataset were converted into the ICD-10 codes. The accuracy of GDR data was assessed by comparing the ICD-10 codes of the GDR with those of the GPCR (considered as the gold standard). Comparison was made at three-digit main primary site level and if the ICD-10 code assigned by the GDR matched with the GPCR code, it was considered as correct CoD. Thus, the accuracy of the GDR was calculated using the formula: number of correct CoD (the right type of cancer)*100/total study subject.

According to results from previous studies,[4,19] the completeness and accuracy of mortality data were estimated to be about 90% and 80%, respectively. We used the anticipated proportions of incompleteness (10%) and inaccuracy (20%) in sample size calculation. The precision levels were considered 1% and 1.5% for completeness and accuracy, respectively. At the confidence level of 95%, the sample size was calculated at 3,458 and 2737 for completeness and accuracy, respectively. Finally, the study sample size was considered at about 3500 subjects. Considering the definition for elderly by the World Health Organization as well as the mean age of the study population, the study subjects were categorized into 0-59 years and ≥60 years old age groups. To assess how quality of the GDR data changed with time, the study period was divided into two periods; 2004–2009 and 2010–2015. The completeness and accuracy of the GDR were calculated for the most common cancer sites in our population including esophagus, stomach, colorectal, lung, leukemia, brain, prostate and female breast.[14]

Results

In total, 66005 deaths from the GDR dataset were linked to 18466 incident cancer cases from the GPCR dataset. Of the 4490 subsequently matched cases, 724 cases specified as non-cancer CoD were excluded, resulting in 3766 cases recruited as study subjects and entered into the analysis of completeness and accuracy (Figure 1). The mean (SD) and median age of the study subjects were 60.40 (17.6) and 63 years old, respectively, with 59.6% being males.

The overall completeness of GDR for identifying CCoD was 92.72% (95% CI: 91.85–93.51). Thus, the CoDs in about 7.28% (95% CI: 6.49–8.15) of cancer cases were identified as garbage or ill-defined codes during the study period. Our results suggested the highest completeness of GDR for cancer sites of brain (96.3%), stomach (95.2%), breast (95.1%), colorectal (94.9%) and leukaemia (94.2%). We found that the completeness of GDR was higher in patients aged under 60 years (95.7%) compared to older subjects (90.4%). The results also suggested that completeness of GDR was higher during earlier years (2004–2009) (94.7%) than the later years (91.4%) of this study period. We found higher completeness of GDR in residents of rural areas (93.7%) than the urban population (91.6%). Table 1 shows distribution of completeness for GDR by age group, residence area, time interval and main primary site in males and females.

The overall estimate of GDR accuracy at the main primary site level was 53.19% (95% CI: 51.59–54.78). Our findings suggested highest accuracy of GDR for cancer sites of the hematopoietic system (79.8%), prostate (74.6%), lung

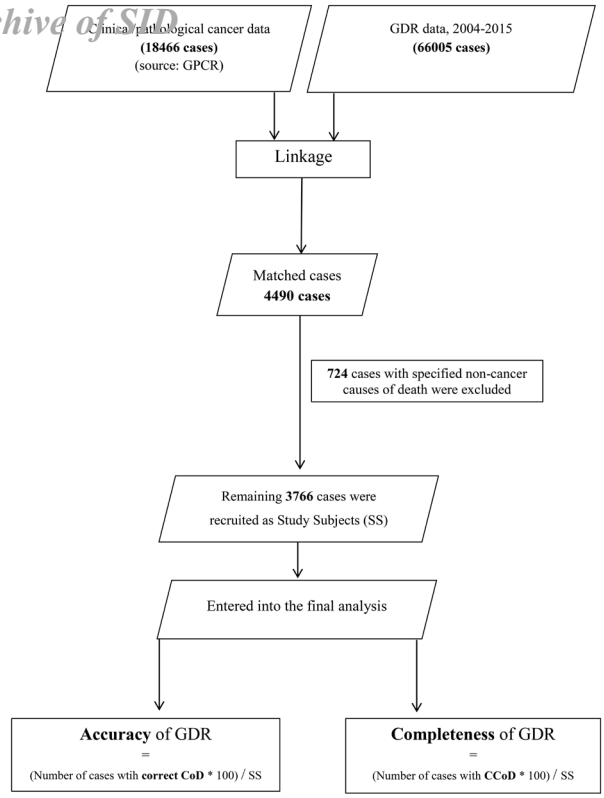

Figure 1. Completeness and Accuracy of Golestan Death Registry (GDR): Study Flowchart. (GPCR=Golestan Population-based Cancer registry; CoD=cause of death; CCoD=Cancer-related cause of death).

(67.9%), stomach (64.9%), breast (61.4%) and oesophagus (54.0%). We found higher accuracy for GDR among death registrations in patients aged under 60 years (53.6%) than older subjects (52.9%). Our results suggested higher accuracy for GDR in more recent years of the study period (2010–2015) (54.6%) compared to earlier years (51.3%). Accuracy of GDR was higher in rural subjects (55.8%) than those residing in urban areas (50.2%). Distribution of GDR accuracy by age group, residence area, time interval and main primary site in males and females is shown in Table 2.

Discussion

Mortality statistics play a critical role in public health surveillance for monitoring health and disease status at the population level and identifying priorities for action. Reliable mortality statistics are essential in this respect,[2,19,20] and a comparison of death certificate information with equivalent clinical and pathology data may help evaluate death certification quality[7,21] and pinpoint means to improve its completeness and accuracy. Using such an approach, this study assessed deaths registry data quality in identifying CCoD in Northern Iran.

In general, completeness of the GDR for identifying CCoD was 92.7%. This means that CoDs in about 7% of cancer cases in Golestan (matched between GDR and GPCR) could be considered coded as 'garbage' or were 'ill-defined'. A South African study demonstrated that less than 10% of CoDs were reported with such codes and completeness of mortality data was reported higher than 90%.[19] In another study from the same country, 12.8% of registered deaths during a 10-year period were assigned to ill-defined or garbage codes.[22] In a study examining death registrations in Brazil, Franca et al reported a wide range of completeness (72%–97%) at the regional level.[23] Several

Table 1. Distribution of Completeness of the Golestan Death Registry in Identifying Cancer-Related Causes of Death in Golestan, Iran (2004–2015), by Age Group, Residence Area, Time Interval and Main Primary Site

	Male			Female		
	Total Number	Completeness		Total Number	Completeness	
		No.	%		No.	%
Age group (y)						
<60	875	838	95.8	808	772	95.5
≥60	1370	1239	90.4	712	643	90.3
Residence area						
Urban	1002	911	90.9	712	659	92.6
Rural	1244	1166	93.7	808	756	93.6
Time interval						
2004–2009	921	866	94.0	595	570	95.8
2010–2015	1324	1211	91.5	925	845	91.4
Main primary site						
Oesophagus	324	300	92.6	215	201	93.5
Stomach	439	414	94.3	139	136	97.8
Colorectal	175	166	94.9	119	113	95.0
Lung	214	195	91.1	66	61	92.4
Leukaemia	260	246	94.6	185	173	93.5
Breast	—	—	—	243	231	95.1
Prostate	118	110	93.2	—	—	—
Brain	109	106	97.2	79	75	94.9

Table 2. Distribution of Accuracy of the Golestan Death Registry in Identifying Cancer-Related Causes of Death in Golestan, Iran (2004–2015), by Age Group, Residence Area, Time Interval and Main Primary Site

	Male			Female		
	Total Number	Completeness		Total Number	Completeness	
		No.	%		No.	%
Age group (y)						
<60	875	474	54.2	808	428	53.0
≥60	1370	748	54.6	712	354	49.7
Residence area						
Urban	1002	515	51.4	712	345	48.5
Rural	1244	707	56.8	808	437	54.1
Time interval						
2004–2009	921	476	51.7	595	301	50.6
2010–2015	1324	746	56.3	935	481	52.0
Main primary site						
Oesophagus	324	173	53.4	118	215	54.9
Stomach	439	283	64.5	92	139	66.2
Colorectal	175	52	29.7	36	119	30.3
Lung	214	148	69.2	42	66	63.6
Leukaemia	260	201	77.3	154	185	83.2
Breast	—	—	—	151	243	62.1
Prostate	118	88	74.6	—	—	—
Brain	109	48	44.0	29	79	36.7

Nordic studies have evaluated the quality of death certificates. A systematic investigation from Denmark suggested that only 2–3 per 1000 deaths per year were coded as ill-defined CoD.[24] Therefore, although the completeness of the GDR in our study was relatively good, it could be improved by considering available evidence from other sources.

The accuracy of the GDR was as low as 53.19%. A Swedish study also reported a low comparative accuracy (almost 50%, similar to that report for GDR in our study) of death certificates and hospital records and recommended use of case summaries – which contains brief medical records of a patient – to help certifiers make accurate CoD certification[7,25]; cases lacking these complementary documents have been shown to have lesser accuracy.[6] An American study in 2011 reported a reasonable overall accuracy of 82.8% and this may have been an overestimate due to miscoding of CoD for certain cancers under investigation.[4] Therefore, in order to improve accuracy of GDR, clinical and pathology documents should be carefully investigated by death certifiers.

Completeness and accuracy of the GDR were lower among older subjects. This replicates the poorer accuracy seen in death certification in the older age group observed in Sweden.[7] Low completeness among this group could be associated with multiple disease conditions often seen among the aged, making it difficult to diagnose the underlying CoD and leading to a greater number of garbage codes.[19] Therefore, certifiers should be careful to scrutinize these records thoroughly before assigning CoD.

The completeness of the GDR decreased during the study period; this might be partly explained by a decline in quality of the GDR in recent years resulting in more garbage or ill-defined codes in death certificates. Concurrently, there was an increase in quality of GPCR,[14] resulting from collection of clinical and pathology documents for cancer patients from a wider range of sources. Consequently, more GDR cases with ill-defined CoD were registered by the GPCR with defined diagnosis of cancer. Thus, improving GPCR quality may be a possible explanation for decreasing GDR completeness during recent years. We found increased GDR accuracy in more recent years of the study period. This may reflect improvement in expertise of certifiers and/or availability of supportive documents due to increase in diagnostic medical services in the study area. Further studies may serve to clarify these points.

The completeness and accuracy of the GDR data were higher among deaths from rural areas compared with urban areas. This is likely due to better coverage by primary health care networks in rural areas, where death certificates are mainly issued by local physicians or healthcare workers who are in close contact with patients within their catchment population and may easily access patients' medical documents. This may result in preparing more accurate death certificates in residents of rural areas; differences in death registration quality in different areas should be investigated.

Our results also showed discrepancies in accuracy for different cancer sites. The GDR had relatively high accuracy for the hematopoietic system, as well as prostate, lung, stomach, breast and oesophagus, while the least accurate was colorectal cancer. The difference in accuracy between sites may partly be explained by the frequency of deaths for a given cancer[4]; for cancers with higher incidence and/or mortality in Golestan (e.g. oesophageal and stomach

cancer), there were more certificates to be assigned a CoD by certifiers, resulting in a greater workload and expertise for certifiers regarding, for example, the clinical presentation of these cancers.[14] In addition, the possibility of an extension or metastasis of a tumour to neighbouring organs (e.g. colorectal cancers) may make it difficult for a certifier to identify the primary site of cancer and to assign a correct code for CoDs.[25] Tailored training programs for certifiers that includes cancer epidemiology and clinical presentation, especially rarer cancers, may result in increased accuracy of the GDR.[19,25]

In conclusion, linkage of death registry data with cancer registry data can be a significant resource for evaluating quality of death registry data. Our findings suggest that completeness of the GDR in identifying CCoD was reasonable, but accuracy was relatively low. Our results indicated higher rates of completeness and accuracy in patients under 60 years old and for specific cancer sites, while quality of data was also relatively better for deaths occurring in rural versus urban areas. Access to clinical and pathological data from multiple sources, especially from other disease registries, may help improve quality of the death registry data. Death certifiers should be trained and made aware of factors related to the quality of death registry data and should consider scrutiny of wide range of information sources when issuing death certificates.

Authors' Contribution

SH-H: collaborated in data processing; performed statistical analysis; wrote the manuscript; NJ-D, NS: collaborated in data processing; edited and critically reviewed manuscript; SMS, AM, RH: collaborated in collection of cancer incidence data; edited and critically reviewed manuscript; MP, MG: collaborated in collection of cancer mortality data; collaborated in quality control; SS, MN-T, MH: interpreted results; critically reviewed manuscript; AE, FB, AF: edited and critically reviewed manuscript; collaborated in quality control; GR: conceptualized and designed the study; performed statistical analysis; wrote the manuscript.

Conflict of Interest Disclosures

The authors have no conflicts of interest.

Ethical Statement

This study was approved by ethics committee of Golestan University of Medical Sciences (ir.goums.rec.1396.166).

Acknowledgement

This study was supported by the Golestan University of Medical Sciences (GOUMS). The data were obtained from Golestan Death Registry (GDR) and Golestan Population-based Cancer Registry (GPCR). The authors would like to thank the staff of the GPCR and GDR for their efforts in the collection and processing of the cancer incidence and mortality data.

References

1. Cancer Office. Cancer Incidence in Iran, According to State Cancer Registries in 2008. Tehran: Iranian Ministry of Health and Medical Education; 2010.
2. Fink AK, German RR, Heron M, Stewart SL, Johnson CJ, Finch JL, et al. Impact of using multiple causes of death codes to compute site-specific, death certificate-based cancer mortality statistics in the United States. Cancer Epidemiol. 2012;36(1):22-8. doi: 10.1016/j.canep.2011.07.004.
3. Marashi Hoseini L, Azizi A, Norouzi M. The survey of neoplasm classification systems in educational hospitals of Iran, Tehran, Shahid Beheshti and Ahvaz Jundishapur Universities of Medical Sciences. Jentashapir Journal of Health Research. 2011;2(3):111-7.
4. German RR, Fink AK, Heron M, Stewart SL, Johnson CJ, Finch JL, et al. The accuracy of cancer mortality statistics based on death certificates in the United States. Cancer Epidemiol. 2011;35(2):126-31. doi: 10.1016/j.canep.2010.09.005.
5. Percy C, Stanek E, 3rd, Gloeckler L. Accuracy of cancer death certificates and its effect on cancer mortality statistics. Am J Public Health. 1981;71(3):242-50.
6. Perez-Gomez B, Aragones N, Pollan M, Suarez B, Lope V, Llacer A, et al. Accuracy of cancer death certificates in Spain: a summary of available information. Gac Sanit. 2006;20 Suppl 3:42-51.
7. Johansson LA, Bjorkenstam C, Westerling R. Unexplained differences between hospital and mortality data indicated mistakes in death certification: an investigation of 1,094 deaths in Sweden during 1995. J Clin Epidemiol. 2009;62(11):1202-9. doi: 10.1016/j.jclinepi.2009.01.010.
8. Li SQ, Cass A, Cunningham J. Cause of death in patients with end-stage renal disease: assessing concordance of death certificates with registry reports. Aust N Z J Public Health. 2003;27(4):419-24.
9. Mahapatra P, Chalapati Rao PV. Cause of death reporting systems in India: a performance analysis. Natl Med J India. 2001;14(3):154-62.
10. Nagle CM, Purdie DM, Webb PM, Green AC, Bain CJ. Searching for cancer deaths in Australia: National Death Index vs. cancer registries. Asian Pac J Cancer Prev. 2006;7(1):41-5.
11. Kmet J, Mahboubi E. Esophageal cancer in the Caspian littoral of Iran: initial studies. Science. 1972;175(4024):846-53.
12. Mahboubi E, Kmet J, Cook PJ, Day NE, Ghadirian P, Salmasizadeh S. Oesophageal cancer studies in the Caspian Littoral of Iran: the Caspian cancer registry. Br J Cancer. 1973;28(3):197-214.
13. Roshandel G, Sadjadi A, Aarabi M, Keshtkar A, Sedaghat SM, Nouraie SM, et al. Cancer incidence in Golestan Province: report of an ongoing population-based cancer registry in Iran between 2004 and 2008. Arch Iran Med. 2012;15(4):196-200. doi: 012154/aim.004.
14. Roshandel G, Semnani S, Fazel A, Honarvar M, Taziki M, Sedaghat S, et al. Building cancer registries in a lower resource setting: The 10-year experience of Golestan, Northern Iran. Cancer Epidemiol. 2018;52:128-33. doi: 10.1016/j.canep.2017.12.014.
15. Bray F, Colombet M, Mery L, Pineros M, Znaor A, Zanetti R, et al. Cancer Incidence in Five Continents, Vol. XI. Lyon: International Agency for Research on Cancer; 2017.
16. Jafari N, Kabir MJ, Motlagh ME. Death registration system in IR Iran. Iran J Public Health. 2009;38 Suppl 1:127-9.
17. World Health Organization (WHO). The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines. Vol 1. Geneva: WHO; 1992.
18. Fritz A, Percy C, Jack A. International classification of diseases for oncology: ICD-O. Geneva: World Health Organization; 2000.
19. Bradshaw D, Pillay-Van Wyk V, Laubscher R. Cause of death statistics for South Africa: Challenges and possibilities for improvement. Cape Town: Medical Research Council: Burden of Disease Research Unit; 2010.
20. Murray CJ, Rajaratnam JK, Marcus J, Laakso T, Lopez AD. What can we conclude from death registration? Improved methods for evaluating completeness. PLoS Med. 2010;7(4):e1000262. doi: 10.1371/journal.pmed.1000262.
21. Hong TT, Phuong Hoa N, Walker SM, Hill PS, Rao C.

Completeness and reliability of mortality data in Viet Nam: Implications for the national routine health management information system. PLoS One. 2018;13(1):e0190755. doi: 10.1371/journal.pone.0190755.
22. Joubert J, Rao C, Bradshaw D, Vos T, Lopez AD. Evaluating the quality of national mortality statistics from civil registration in South Africa, 1997-2007. PLoS One. 2013;8(5):e64592. doi: 10.1371/journal.pone.0064592.
23. Franca E, de Abreu DX, Rao C, Lopez AD. Evaluation of cause-of-death statistics for Brazil, 2002-2004. Int J Epidemiol. 2008;37(4):891-901. doi: 10.1093/ije/dyn121.
24. Ylijoki-Sorensen S, Sajantila A, Lalu K, Boggild H, Boldsen JL, Boel LW. Coding ill-defined and unknown cause of death is 13 times more frequent in Denmark than in Finland. Forensic Sci Int. 2014;244:289-94. doi: 10.1016/j.forsciint.2014.09.016.
25. Schaffar R, Rapiti E, Rachet B, Woods L. Accuracy of cause of death data routinely recorded in a population-based cancer registry: impact on cause-specific survival and validation using the Geneva Cancer Registry. BMC Cancer. 2013;13:609. doi: 10.1186/1471-2407-13-609.