



Original article

Estimation of lung and bronchial cancer registry completeness via capture-recapture method using population-based cancer registry in Khuzestan province in 2011



Kourosh Etemad^a, Mohammad Javad Mohammadi^b, Seyed Saeed Hashemi Nazari^c, Homayoun Amiri^{c,*}, Abdolhassan Talaiezhadeh^d

^a Department of Epidemiology, School of Public Health, Shahid Beheshti University of Medical Sciences, Tehran, Iran

^b Department of Environmental Health Engineering, School of Public Health AND Air Pollution and Respiratory Diseases Research Center, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran

^c Department of Epidemiology, School of Public Health, Shahid Beheshti University of Medical Sciences, Tehran, Iran

^d Department of Surgery, Faculty of Medicine, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran

ARTICLE INFO

Keywords:

Capture-recapture
Estimation
Log-linear
Lung cancer
Cancer registry

ABSTRACT

Introduction: Population-based cancer registry (PBCR) plays a significant role in burden of cancer estimation. Estimating the completeness of cancer registry is one of the criteria to assess the quality of data (results). Also, Capture-recapture is one of the methods used to assess completeness of cancer registry.

Objective: This study was aimed to estimate the completeness of lung and bronchial cancer registry using three-source capture-recapture method in Khuzestan province.

Materials and methods: This study was carried out using three-source capture-recapture method and data were obtained from three sources including medical records, death certificates and pathology labs (reports) in Khuzestan. In this study, total new cases of lung and bronchial cancer registered by the three sources in 2011, were enrolled. Among the sources, the common cases were identified and the completeness rate of lung and bronchial cancer registry was calculated by log-linear method using R software, finally.

Results: Totally, 426 new cases for lung cancer had reported from the three sources in 2011. The completeness index of lung cancer registry was totally estimated 72%, as well as it was independently 40%, 25% and 37.6% for the medical records, pathology reports and death certificates sources, respectively.

Conclusion: According to the results, after collection and remove repetition of the cases related to lung cancer, it should be added their number registered by the pathology (labs), medical records and death certificates sources by 28%–40%, to estimate actual incidence of lung cancer. Also, the actual incidence of lung cancer cases related to each of medical records, pathology and death certificates was calculated by adding the reported values by 60%, 75% and 63%, respectively.

1. Introduction

The incidence of non-communicable diseases is accounted for as one of the main challenges for health and development in the 21st century. Globally, the mortality rate due to non-communicable diseases was over 38 million (68% out of total mortality), in 2012.¹

Cancer by the new cases of 14 million and the annual mortality of 8.2 million is the second leading cause for mortality after cardiovascular diseases in the world.^{2–6} Lung cancer is the most fatal and expensive cancer worldwide.⁷ Among the various cancers, lung is only the

cancer which has different incidences in Iran and world⁸ and undoubtedly its incidence in Iran has had an underestimation.⁹ The reasons for this underestimation may be lack of ease availability to lung lesions, equipment and facilities for diagnosis and sampling in some region of the country. Also, another reason is the screen of other cancers and the ease of their clinical diagnosis which causes abnormal less proportion of lung cancer cases than other cancers. In 2012, the World Health Organization (WHO) has highlighted the cancer prevention and improvement of the life quality of cancer patients. To reach the aim, the compilation of national plans to prevent and control the cancers is one

* Corresponding author. Department of Epidemiology, School of Public Health, Shahid beheshti University of Medical Sciences, Tehran, Iran. Tel.: 09161184916; fax: +986133387724.

E-mail address: amiri.homayoun54@gmail.com (H. Amiri).

<https://doi.org/10.1016/j.cegh.2020.02.017>

Received 12 April 2019; Received in revised form 26 January 2020; Accepted 17 February 2020

Available online 29 February 2020

2213-3984/ © 2020 INDIACLEN. Published by Elsevier, a division of RELX India, Pvt. Ltd. All rights reserved.

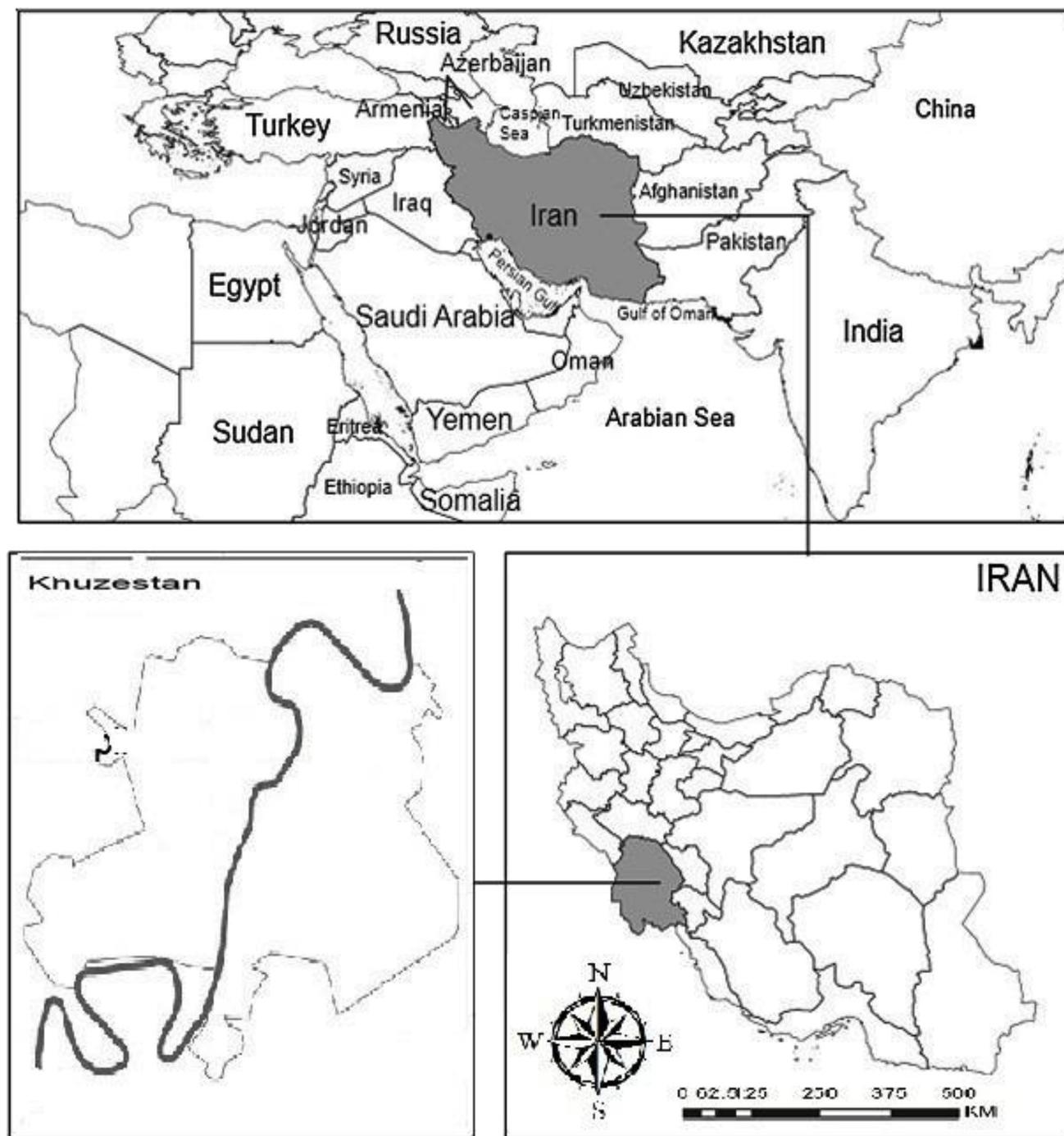


Fig. 1. Location of the Khuzestan province in southwest of Iran.

of the health necessities as well as the most effective factor to reduce the incidence and burden of cancer in each community.^{9,10}

One of the problems to compile such programs is lack of adequate data on the cancer behavior and its incidence trend over various decades and geographies, especially in developing countries such as Iran. To design cancer registry systems, extensive efforts have been made in various countries including Iran, in recent years.^{9,10} The aim for population-based cancer registry is estimation of the cancer burden in the covered population in order to monitor its trend and regional variations, as well as to supply a database to perform epidemiological researches.¹¹

Health politicians need to know how to rely on information which forms their policies and decisions basis.¹² The assessment of data

quality for cancer registry is included two parts, the completeness estimation and the possibility to compare the data, its reliability and timeliness. The registry completeness index is defined as the ratio of incidence cases which has been recorded.¹² The Methods used for completeness estimation are included two groups of quantitative and qualitative methods. Qualitative methods are included date-based data methods, the ratio of mortality to incidence, the number of sources and their respective reports, and diagnostic tissue effect method; as well as quantitative methods are included independent case comparison (ascertainment), capture-recapture method, and death certificates method and follow up method.¹³ The capture-recapture studies can be carried out directly and indirectly. In the direct method, a case is selected, marked and afterwards re-captured. In indirect method, the current

data such as data registry lists or surveys is used. A viewed case is considered as a capture and then a total estimation of the whole population is made using estimators.¹⁴

Generally, the capture-recapture methods are used in epidemiology in three parts: estimating the number of population, estimating the incidence of a condition or disease based on performed surveys and evaluation of the completeness of data registry systems.¹⁵ There are four basic presumption in capture-recapture method: 1) the population be closed, 2) finding common cases between two or more lists be possible, and 3) the lists be independent to each other (it means that the presence of a person in a list should not cause to increase or decrease his presence possibility in other list), 4) the presence possibility of individuals in the lists be independent to their characteristics.¹⁶

The three-source capture-recapture method has been used to estimate the completeness of cancers registry in the overseas, for example in Seoul, South Korea in 2012,¹⁷ Gambia in 2012,¹⁸ Riyadh, Saudi Arabian in 2003,¹⁹ in Seoul by Kim and colleagues in 1999²⁰ and in Virginia, United State in 2004 (to estimate completeness of lung cancer registry).²¹ Also, two-source method has been used to estimate lung cancer registry in the Tuscany, Italian in 2001.²² In Iran, three-source capture-recapture method has been used to estimate the completeness of lung cancer registry by Khodadost and colleagues in 2014²³ as well as to estimate the coverage of mortality data caused by lung cancer by Isaac Dortaj and colleagues in 2011.²⁴ Furthermore, estimation of cancer incidence in northwest of Iran by Morteza Ghajarzadeh and colleagues in 2013,²⁵ estimation of the cancer cases in Golestan province by Ali Reza Zomestani and colleagues in 2013²⁶ as well as Roxana Sharifian and colleagues researches which aimed to assess the completeness of cancer registry in 2009, were all conducted using two-source capture-recapture method.²⁷ The current study aimed to estimate the completeness of lung and bronchial cancer registry in Khuzestan province by compare three sources for data registry (including pathology reports, death certificates and medical records) using capture-recapture method. Due to the difference between the incidence rate of lung cancer in Iran and the world, as well as considering the higher incidence of lung and bronchial cancer in Khuzestan province than other provinces (following the Markazi province), the lung cancer was selected to study in this research.

2. Materials and Methods

2.1. Area of study

The study was carried out by three-source capture-recapture method using data obtained from three sources including medical reports, death certificates and the pathology reports of respective labs in Khuzestan province, southwest of Iran. Total new cases of lung and bronchial cancer registered by the three sources in 2011, were enrolled in the study. Khuzestan is bounded to the west by Iraq and to the south by the Persian Gulf. Khuzestan province has an area of 63,238 km². Ahvaz is center of this province and located in the Middle East, between 48° and 29' east of the Greenwich meridian, 31° and 45' minutes north of the equator.^{3,28–36} Location of the Khuzestan province in southwest of Iran is presented in Fig. 1.

3. Method

First, the data from each source were separately entered into Excel (Office 2010), and then repeated cases of each source were identified and deleted. To find common cases between these three lists, the patient data related to all of the sources were merged (each source was marked by a certain color). The four characteristics of the patients including the name, surname, father's name and their residence city were considered. Common cases were identified by sorting the patients in terms of different characteristics (name, surname, father's name and residence) using the Sort order as well as compare each case with others

which were registered in another source.

In this study, it was used a new methodology such a way that the death data related to the year of study (2011) were compared to the previous years-at least five years-as well as to the pathological and medical sources. Although, some of the death data in 2011 are common with the medical and pathological data from same year, but in some cases of death, the incidence of disease may be related to previous years. So in this study, death data were compared with the pathology and medical data from previous years and then if the names of these persons were found, these common names were excluded from the list of death data related to the year of study. Also, considering the survival rate of lung cancer, patients who have been registered in the pathology or medical lists in the year of study (and so they are not present in the source of death list of that year) may be die in the next years and be recorded in the death list of the future years. Therefore, their presence were also compared with death data of next years after the year of the study-at least five years- and if there were (common) their names in the future lists, they were added in to the death list of the study year. If the mentioned methodology is not followed, it will lead to an over-estimation in the size of population.

Afterwards, a cross table (contingency table) with 8 cell ($2^3 = 8$) was drawn. 7 of these cells were defined (including the number of individuals presented in each source and the number of common individuals among the sources) and one cell (including the number of individuals who were not present in any of the three sources) was not marked. In Venn Figure (Fig. 2), the number of cases presents in each resource has been shown as well as how the resources overlaps.

Data analysis was performed using log-linear model in Stata 13 and R software. 8 various log-linear models were fitted on these seven cells in the cross table to estimate the abundance of the 8th cell (individuals which not registered in any of the three sources) in the table. Then expected abundance of 8th cell was estimated using each of these models.

These 8 models were different due to the number of interactions which was existed: one of them is independent model (no interaction); three models had one interaction; three models had two interactions and one model had three interactions (the saturation model). The log-linear model considers dependency and heterogeneity of the lists to calculate, and acts very powerful when several sources be present.¹⁵ The Models for the 3 sources of registry (source A with i levels, source B with j levels and source C with k levels) are as follows:

Statistical models

$$\begin{aligned} \log \mu_{ijk} &= \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C \\ \log \mu_{ijk} &= \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} \\ \log \mu_{ijk} &= \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ik}^{AC} \\ \log \mu_{ijk} &= \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{jk}^{BC} \\ \log \mu_{ijk} &= \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} \\ \log \mu_{ijk} &= \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{jk}^{BC} \\ \log \mu_{ijk} &= \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ik}^{AC} + \lambda_{jk}^{BC} \\ \log \mu_{ijk} &= \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC} + \lambda_{ijk}^{ABC} \end{aligned}$$

Where $\log \mu_{ijk}$ is the logarithm of expected abundance for all table's cells, λ is the natural logarithm of the expected number of individuals who are present in all lists, λ^A , λ^B and λ^C are the main effect of variables, λ^{AB} , λ^{AC} and λ^{BC} are the effect of two-ways interactions and λ^{ABC} is the effect of three-ways interactions (the highest desired effect).

To assess fitness of the models to available data and to select the best model, the logarithmic likelihood ratio test (G^2), degree of Freedom, Akaike information criterion (AIC) and Bayesian information criterion (BIC) is used.

The AIC and BIC are tested by likelihood ratio tests, and the lower values represent the better fitness.¹⁵ Their calculation is as follows³⁷:

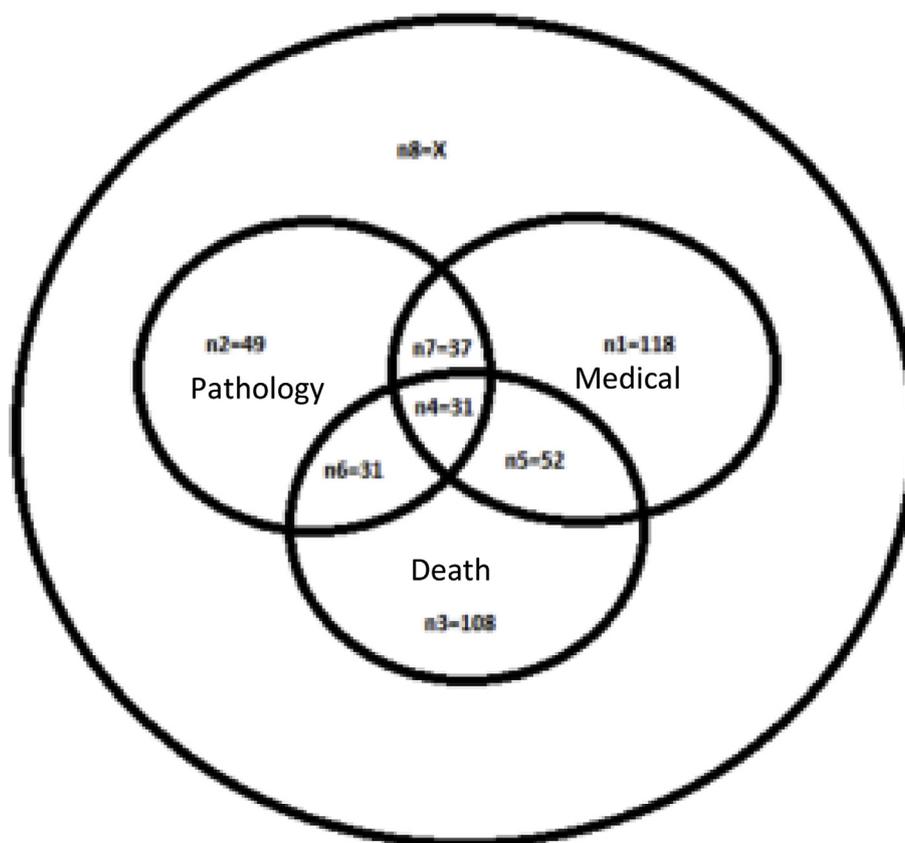


Fig. 2. Venn Figure related to the number of common cases of lung cancer registered in three sources of medical records, pathology reports and death certificates in Khuzestan province in 2011.

$$G^2 = 2 \sum n_{ijk} \log \left(\frac{n_{ijk}}{\mu_{ijk}} \right)$$

Where n_{ijk} is the number of observed individuals and μ_{ijk} is the expected number of individuals which shows that when G^2 is lower, the fitness model will be more appropriate.²⁶

When the interactions is entered into the model and the aim is selection of best model, Akaike will be practical and its formula is³⁸:

$$AIC = G^2 - 2(df)$$

The G^2 term is a criterion which represents how the model is fitted to the data, and the second term “2(df)” is due to additional parameters and also the complexity of the model.²⁶

BIC itself has two different types³⁸:

$$SIC = G^2 - (\ln N_{obs})(df)$$

$$DIC = G^2 - (\ln N_{obs}/2\pi)(df)$$

AIC is the most common criterion to select appropriate model.²⁶

Usually, the most fitness model using AIC, is the saturated model-a model with the degrees of freedom by zero and the interactions by K-1. But, in this model the population size has an overestimation and gives more extensive reliability. In the cases where sample size is small, the two other criteria especially DIC represent better results.¹⁵ In the present study, AIC, BIC and G^2 were used to select a better fitness model. Finally, according to AIC- which has been mostly used by researchers-a model including the interactions between the medical and death sources and an independent effect of pathology source was selected.

In this study, log-linear method was used using three sources to control bias due to lack of sources independency. The benefit of this method is that by mediating the interactions (between different sources) it can be consider the effect of the dependency (positive or negative) between the resources in estimations as well as can be greatly

overcome on the bias caused by lack of the presumption of sources independency.¹⁶

4. Results

The cases of three sources were compared and the common cases between them deleted (the common cases considered once) and then 426 new cases of lung cancer were totally registered in 2011 from three sources. The medical, pathology and death certificates sources were registered 238, 148 and 222 cases, respectively. Among them, there were 280 men (65.7%) and 146 women (34.3%) with the men to women ratio of 1.9. The most abundant age group was 65–80 years-old with 35% followed by 45–64 years-old with 32% of total cases.

Table 1 represents the demographic characteristics of patients registered by the three sources of cancer registry.

The three-source analysis was performed using the Log-linear model and its results are given in Table 2. Among the 8 fitted models based on AIC, the third model including an interaction between the medical and death sources as well as an independent effect of pathology source with the highest fitness, was selected.

Table 1 Comparison of demographic data of patients obtained from three sources of lung cancer registry in 2011.

Gender	Age groups				
	45 >	45–64	65–80	80 <	
man	43	148	174	61	
woman	10	34.7	40.8	14.5	
year	number	280	146		
1390	Percentage	65.7	34.3		

Table 2
Log-linear models fitted to the data obtained from three sources of lung and bronchial cancer registry in 2011.

Log-linear model	AIC	Df	Estimated number of lung cancer cases registered in none of the sources	Estimated total number of lung cancer cases	95% confidence interval	infit
h/p/d	53.6	2	201	627	568–701	OK
hd/p	52.7	1	163	589	529–667	OK
Hp/d	53.2	1	217	643	575–729	OK
Pd/h	55.3	1	208	634	570–715	OK
hp/pd/hd	54.6	0	324	750	NA	Warning 2
hp/hd	54.6	0	170	596	524–703	OK
hp/pd	54.6	0	345	671	587–785	OK
hd/pd	54.6	0	156	582	518–674	OK

h: medical source, p: pathology source, d: death source, Infit: represents existence an error in model.

According to this model, the number of lung cancer cases which was registered in none of these sources was estimated to be 163. Consequently, the total number of lung cancer cases in 2011 was estimated to be 589 (with a confidence interval of 667-529).

Accordingly, the total completeness index for lung cancer registry after delete repeated cases was 72% for all three sources and it was separately for each medical (with 238 cases), pathology (with 148 cases) and death certificates (with 222 cases) source was 40%, 25% and 37.6%, respectively. The incidence rate which was reported by three sources of medical, pathology, and death records following the deletion of repeated cases, was 9.4 of 100000, in 2011. Also, using log-linear model, the incidence rate was estimated to be 13 of 100000. The incidence rates in age subgroups and the completeness of cancer registry were separately calculated based on gender and age subgroups (Table 3).

The highest underestimation was occurring in the age group of older than 80.

5. Discussion

In this study, the completeness of lung cancer registry was estimated using three capture-recapture method and log-linear models. The completeness of lung cancer registry in Khuzestan province was estimated 72% using data obtained from three sources of medical records, pathology reports and death certificates following deletion of repeated cases.

This estimation is not consistent with Khodadost and colleagues study in 2011 which had reported the completeness of lung cancer registry by 24.22% in Ardabil province in 2006 and 2008 using capture-recapture method.¹² The study of Dortaj and colleagues (2011), which was conducted on the data obtained from 2004 to 2006 years in Fars province using three-source method, has been reported the completeness of lung cancer registry by 50% that is closer to the results of current study.²⁴ Sharifian and colleagues (2009) in Shiraz using two-source method found that the completeness of lung cancer registry by Chapman method was 70.4% and by Chao method was 59.9% which these results are consistent with the present study.²⁷

The completeness of lung cancer registry which reported in the present study is lower than other countries as well as almost similar to

Table 3
Estimation of lung cancer cases based on gender and age groups in 2011.

	subgroups	Reported cases	Estimated cases	95% confidence interval	Completeness of cancer registry (%)
Age groups	man	280	397	353–456	70.5
	woman	146	251	203–322	58
	45 >	43	73	51–119	59
	45–64	148	191	169–221	77
	65–80	174	253	218–302	68
	80 <	66	185	103–436	33

them in some cases. For example, in a study entitled "Completeness of cases registry in the national program of cancer registry in Ireland" was published by Brin and colleagues in 2014, the completeness of lung cancer registry had reported 98.7%.²⁰ In a study which has been carried out by Crowsti and colleagues in Tuscany (Italian) in 2001 using the two-source capture-recapture method (in order to estimate completeness of cancer registry of Tuscany), completeness of lung cancer registry was estimated to be 96.6%.²⁷ In the study entitled "estimation of non-registered cases in cancer registry program using three-source capture-recapture method", the completeness of lung cancer registry was reported 74% in Virginia, United State in 2004.³⁹ In the present study, the highest completeness of lung cancer registry (40%) was related to the medical source, but in the Khodadost's study, the death source had the highest completeness of registry by 55.12%.

The incidence rate which was reported based on the three sources of medical, pathology, and death records following the deletion of repeated cases, was 9.4 of 100000 in 2011. Also, using log-linear model, the incidence rate was estimated to be 13 of 100000.

Based on the national report in 2009, the incidence of lung cancer in Khuzestan province in men and women had been reported 10 and 3.9 of 100000, respectively. But, based on the present study, the reported incidence in men was 12.2 of 100000 and in women was 6.5 of 100000. These estimations using log-linear model were calculated for men and women 17.3 and 9.4 of 100000, respectively.

The estimated incidence based on the age subgroups represents an increase in disease incidence as the age increases. Also in the current study, the highest incidence is related to the age group of 65 years and older which is in agreement with the study was done in Ardebil province in 2004 and 2006. The highest underestimation was observed in the age group of 80 and older.

The mean age in men and women was 64.3 and 62 years. The age distribution of disease incidence in men and women was not significantly (greatly) different.

6. Strengths and limitations

In spite of the mentioned considerations, there are still death cases which have not been registered in pathological and medical sources in the same year or in the past. These cases may be related to the year of study or previous years which it could affect the results of the study. Therefore, it is suggested that this limitation should be considered in the next researches.

7. Conclusion

Findings of this study represent essential use of capture-recapture method to determine underestimations in lung cancer incidence and also to estimate its incidence more accurately which be closer to reality.

The results of this study showed that to estimate actual incidence of lung cancer in Khuzestan province, it should be added 28%–40% to the numbers of case registered by the pathology (labs), medical records and

death certificates sources, after delete repeated cases. Also, if one of each medical records, pathology and death certificates was considered it should be respectively added 60%, 75% and 63% to each source (data) in order to estimate actual incidence of lung cancer. In current study considering the used methodology is important such a way that the death data related to the year of study were compared to the cases registered in pathological and medical sources in the previous years and the common cases were excluded from the death data. Also, pathological and medical cases in the year of study were compared to death data in the next years and the common cases were added in to the death data. If this methodology is not followed, it will lead to an over-estimation.

Funding/Support

The authors gratefully acknowledge Abadan School of Medical Sciences for financial support and providing necessary facilities to accomplish this research with project number of 92/p/p.A/919.

Declaration of competing interest

There are no conflicts of interests among the authors.

Acknowledgment

This paper is derived out from thesis of Homayoun Amiri for post-graduate degree of Infectious Diseases specialist (Grant; 92/p/p.A/919). The authors wish to thank the chief and personnel of Shahid Beheshti University of Medical sciences for supporting of this study as well as the research affair deputy of medical college for approving the study. This study was supported by Shahid Beheshti University of Medical Sciences for providing ethical issue: (IR.SBMU.RAM.REC.1395.91) of this research.

References

1. WHO. Global. *Status Report on Noncommunicable Diseases 2014* 2014; 2014 (vii).
2. IARC. *reportWorld Cancer Report 2014*2014.
3. Goudarzi G, Geravandi S, Idani E, et al. An evaluation of hospital admission respiratory disease attributed to sulfur dioxide ambient concentration in Ahvaz from 2011 through 2013. *Environ Sci Pollut Res*. 2016;23:22001–22007.
4. Geravandi S, Sicard P, Khaniabadi YO, et al. A comparative study of hospital admissions for respiratory diseases during normal and dusty days in Iran. *Environ Sci Pollut Res*. 2017;24:18152–18159.
5. Dastoorpoor M, Sekhvatpour Z, Masoumi K, et al. Air pollution and hospital admissions for cardiovascular diseases in Ahvaz, Iran. *Sci Total Environ*. 2019;652:1318–1330.
6. Goudarzi G, Geravandi S, Alavi N, et al. Association between cancer risk and polycyclic aromatic hydrocarbons' exposure in the ambient air of Ahvaz, southwest of Iran. *Int J Biometeorol*. 2018;62:1461–1470.
7. Kanodra Neeti M, Tanner Nichole T. *Screening and Early Detection Efforts in Lung Cancer*. vol. 1. 2015; 2015.
8. Saba V. Estimation of age standardized ratio of lung cancer in Iran in 2014 and 2030. *Paramed Sci Mil Health*. 2015;10:17–23.
9. Radmard AR. Five common cancers in Iran. *Arch Iran Med*. 2010;13:143.
10. Office CfDCPnDUC. *reportIranian Annual of National Cancer Registration Report 2009-2010*. Tehran2012:2.
11. Schmidtman I. Estimating completeness in cancer registries—comparing capture-recapture methods in a simulation study. *Biom J*. 2008;50:1077–1092.
12. Khodadost M, Yavari P, Babaei M, et al. Estimating the completeness of gastric cancer registration in Ardabil/Iran by a capture-recapture method using population-based cancer registry data. *Asian Pac J Cancer Prev*. 2015;16:1981–1986.
13. Castro CSFcd. *Estimating Completeness of Cancer Registration-Na Application to Gastric Cancer in the North Region Registry of portugal*. 2011; 2011.
14. Brittain S, Böhning D. Estimators in capture-recapture studies with two sources. *AStA Adv Stat Anal*. 2009;93:23–47.
15. Hassan Zadeh J, Nasehi M, Rajaeifard A, Roshani D, Ghaderi E. Introduction to capture-recapture studies. *Iran J Epidemiol*. 2014;10:62–70.
16. Motevalian AHNK, et al. *Size Estimation of Deaths Caused by Traffic Accidents in the City of Kerman with Capture-Recapture Method*. vol. 2. 2007; 2007.
17. Song M, Cho I-S, Li ZM, Ahn Y-O. Completeness of cancer case ascertainment in Korea radiation effect and epidemiology cohort study. *J Kor Med Sci*. 2012;27:489–494.
18. Shimakawa Y, Bah E, Wild CP, Hall AJ. Evaluation of data quality at the Gambia national cancer registry. *Int J Canc*. 2013;132:658–665.
19. Al-Zahrani A, Baomer A, Al-Hamdan N, Mohamed G. Completeness and validity of cancer registration in a major public referral hospital in Saudi Arabia. *Ann Saudi Med*. 2003;23:6–9.
20. Kim D-S, Lee M-S, Kim D-H, et al. Evaluation of the completeness of cancer case ascertainment in the Seoul male cohort study: application of the capture-recapture method. *J Epidemiol*. 1999;9:146–154.
21. McClish D, Penberthy L. Using Medicare data to estimate the number of cases missed by a cancer registry: a 3-source capture-recapture model. *Med Care*. 2004;1111–1116.
22. Crocetti E, Miccinesi G, Paci E, Zappa M. An application of the two-source capture-recapture method to estimate the completeness of the Tuscany Cancer Registry, Italy. *Eur J Canc Prev*. 2001;10:417–423.
23. Khodadost M, Yavari P, Khodadost B, et al. Estimating the esophagus cancer incidence rate in Ardabil, Iran: a capture-recapture method. *Iran J Cancer Prev*. 2016;9:e3972.
24. Dortag E, Bahrampour A, Haghdoost A, Zendedel K, Jaberipour M, Marzaban M. Completeness of fars province deaths registry on cancer death using capture-recaptures method. *J North Khorasan Univ Med Sci*. 2011;3:33–44.
25. Ghojzadeh M, Mohammadi M, Azami-Aghdash S, Sadighi A, Piri R, Naghavi-Behzad M. Estimation of cancer cases using capture-recapture method in Northwest Iran. *Asian Pac J Cancer Prev APJCP*. 2013;14:3237–3241.
26. Zemestani AR, Mahmoudi M, Keshkkaar AA, Majdzadeh SR, Foroozfar MH, Semnani S. Estimation of cancer cases in golesan province between 2004–2006 by using capture-recapture method. *Med J Tabriz Univ Med Sci Health Serv*. 2013;35:26–33.
27. Sharifian R, SedaghatNia MH, Nematolahi M, Zare N, Barzegari S. Estimation of completeness of cancer registration for patients referred to Shiraz selected centers through a two source capture re-capture method, 2009 data. *Asian Pac J Cancer Prev APJCP*. 2015;16:5549–5556.
28. Davar H, Taghaviirad SS, Mohammadi MJ. The investigation of effects of silica on the environment and prevention of release of the silica particles with simulation of gas-solid flow in a gas cyclone. *Res J Chemi Environ*. 2014;18:28–30.
29. Dobaradaran S, Geravandi S, Goudarzi G, et al. Determination of cardiovascular and respiratory diseases caused by PM10 exposure in bushehr, 2013. *J Mazandaran Univ Med Sci*. 2016;26:42–52.
30. Geravandi S, Goudarzi GR, Vousoghi Niri M, Mohammadi Mj, Saeidimehr S, Geravandi S. Estimation of the cardiovascular and respiratory mortality rate resulted from exposure to sulfur dioxide pollutant in ahvaz. *J Environ Stud*. 2015;41:341–350.
31. Goudarzi G, Geravandi S, Forouzandeh H, et al. Cardiovascular and respiratory mortality attributed to ground-level ozone in Ahvaz, Iran. *Environ Monit Assess*. 2015;187:1–9.
32. Khaefi M, Goudarzi G, Yari AR, et al. An association between ambient pollutants and hospital admitted respiratory cases in Ahvaz, Iran. *Fresenius Environ Bull*. 2016;25:3955–3961.
33. Neisi A, Goudarzi G, Akbar Babaei A, et al. Study of heavy metal levels in indoor dust and their health risk assessment in children of Ahvaz city, Iran. *Toxin Rev*. 2016;35:16–23.
34. Yari AR, Goudarzi G, Geravandi S, et al. Study of ground-level ozone and its health risk assessment in residents in Ahvaz City, Iran during 2013. *Toxin Rev*. 2016;35:201–206.
35. Khaefi Mehran GS, Ghasem Hassani, Ahmad Reza Yari, et al. Association of particulate matter impact on prevalence of chronic obstructive pulmonary disease in ahvaz, southwest Iran during 2009–2013. *Aerosol Air Qual Res*. 2017;17:230–237.
36. Nashibi R, Afzalzadeh S, Mohammadi MJ, Yari AR, Yousefi F. Epidemiology and treatment outcome of mucormycosis in khuzestan, southwest of Iran. *Clin Infect Dis*. 2017;12:e37221.
37. Aghaei A1 A-JT, Baiki O, Mosavi-Jarrahi A. Estimation of the gastric cancer incidence in Tehran by two-source capture-recapture. *Asian Pac J Cancer Prev*. 2013;14:673–677.
38. Hook EB, Regal RR. Validity of methods for model selection, weighting for model uncertainty, and small sample adjustment in capture-recapture estimation. *Am J Epidemiol*. 1997;145:1138–1144.
39. Billimoria KY, Stewart AK, Winchester DP, Ko CY. The National Cancer Data Base: a powerful initiative to improve cancer care in the United States. *Ann Surg Oncol*. 2008;15:683–690.