



Validity of self-reported cancer history in the health examinees (HEXA) study: A comparison of self-report and cancer registry records



Sooyoung Cho^a, Aesun Shin^{a,b,*}, Daesub Song^c, Jae Kyung Park^c, Yeonjung Kim^c, Ji-Yeob Choi^{b,d}, Daehee Kang^{a,b,d}, Jong-Koo Lee^{e,f,g}

^a Department of Preventive Medicine, Seoul National University College of Medicine, Seoul, Republic of Korea

^b Cancer Research Institute, Seoul National University, Seoul, Republic of Korea

^c Division of Epidemiology and Health Index, Center for Genome Science, Korea National Institute of Health, Korea Center for Disease Control and Prevention, Osong, Republic of Korea

^d Department of Biomedical Science, Seoul National University Graduate School, Seoul, Republic of Korea

^e JW LEE Center for Global Medicine, Seoul National University College of Medicine, Seoul, Republic of Korea

^f Department of Family Medicine, Seoul National University College of Medicine, Seoul, Republic of Korea

^g Seoul Center for Infectious Disease Control, Seoul, Republic of Korea

ARTICLE INFO

Article history:

Received 17 April 2017

Received in revised form 18 July 2017

Accepted 22 July 2017

Available online 29 July 2017

Keywords:

Neoplasms

Self report

Reproducibility of results

Sensitivity and specificity

Health surveys

Registries

Epidemiologic studies

ABSTRACT

Purpose: To assess the validity of the cohort study participants' self-reported cancer history via data linkage to a cancer registry database.

Methods: We included 143,965 participants from the Health Examinees (HEXA) study recruited between 2004 and 2013 who gave informed consent for record linkage to the Korean Central Cancer Registry (KCCR). The sensitivity and the positive predictive value of self-reported histories of cancer were calculated and 95% confidence intervals were estimated.

Results: A total of 4,860 participants who had at least one record in the KCCR were included in the calculation of sensitivity. In addition, 3,671 participants who reported a cancer history at enrollment were included in the calculation of positive predictive value. The overall sensitivity of self-reported cancer history was 72.0%. Breast cancer history among women showed the highest sensitivity (81.2%), whereas the lowest sensitivity was observed for liver cancer (53.7%) and cervical cancer (52.1%). The overall positive predictive value was 81.9%. The highest positive predictive value was observed for thyroid cancer (96.1%) and prostate cancer (96.1%), and the lowest was observed for cervical cancer (43.7%).

Conclusion: The accuracy of self-reported cancer history varied by cancer site and may not be sufficient to ascertain cancer incidence, especially for cervical and bladder cancers.

© 2017 Published by Elsevier Ltd.

1. Introduction

Self-reported disease history is often used to ascertain outcomes of interest in epidemiological research, as well as to obtain the characteristics of a study population. Findings from earlier studies indicate that the validity of self-reported cancer history is influenced by the study population. Results from a prospective study conducted in the United States that compared self-reported cancer diagnoses with population-based cancer registry data found a sensitivity of 79% and a positive predictive value of 75% [1].

A Swedish cohort study reported a sensitivity of 82% and a specificity of 98% [2]. In a population-based cohort study of elderly Australian woman, the overall sensitivity of self-reported cancer was 89.2% and the positive predictive value was 66.5% [3]. In contrast to results in Western populations, relatively low validity was observed in a general Japanese population (sensitivity, 52.6%; positive predictive value, 59.7%) [4]. A previous study conducted in Korea also reported low sensitivity (40%) along with high specificity (99%) [5]. This difference could be due to cultural differences with regard to disclosure of personal information.

The purpose of the present study was to assess the validity of self-reported cancer history by sociodemographic characteristics (including age, sex, education, and monthly income) and cancer site in a large-scale cohort study, the Health Examinees (HEXA) study, via data linkage to the Korean Central Cancer Registry (KCCR).

* Corresponding author at: Department of Preventive Medicine, College of Medicine, Seoul National University, 103, Daehak-ro, Jongno-gu, Seoul, 03080, Republic of Korea.

E-mail address: shinaesun@snu.ac.kr (A. Shin).

2. Methods and methods

2.1. Study population

The health examinees study (HEXA) is a population-based cohort study which aims to investigate the association between epidemiologic characteristics and major chronic diseases, including cancer, in the Korean population [6]. Since the HEXA study is within the Korean Genome and Epidemiology study, it is also called KOGES_HEXA. A total of 173,357 participants were prospectively recruited between 2004 and 2013 at 38 health examination centers and training hospitals located in eight regions in Korea, using a standardized study protocol. Details on the main objectives, rationale, study design, and baseline characteristics of the HEXA study have been published elsewhere. The study protocol was approved by the Institutional Review Board (IRB) of the Seoul National University Hospital, Seoul, Korea (IRB number 0608-018-179) and the Korea National Institute of Health (IRB number 2014-08-02-3C-A). A total of 143,965 participants gave information on resident registration number and written informed consent for data linkages with secondary data sources, such as the KCCR.

All participants were questioned by trained interviewers and responded to a questionnaire that asked about sociodemographic status and lifestyle factors. Information on cancer history was obtained by asking: ‘Have you ever been told by a doctor that you have cancer?’, and then, as follow-up questions, ‘‘If yes, is the site of the cancer: stomach, colon/rectum, liver, lung, breast (among women), prostate, thyroid, cervix uteri, or bladder?’’ and ‘‘How old were you when you were diagnosed with cancer?’’. Participants can report their information on diagnosis of up to three cancers. If participants reported two different cancer histories, both cancers were counted. Cancer sites that were included in the questionnaire were the most common cancers in Korea during the period 1999–2001 [7].

2.2. Cancer registry

The KCCR is responsible for collecting and managing cancer registry data nationwide and provides nationwide cancer statistics annually [7]. The Cancer Act provides the legal basis for cancer registration. Data on cancer incidences are collected from hospitals, 11 population-based registries, and additional medical chart reviews. The completeness of the data was around 97.8% in 2014 [8]. The definition of multiple primary cancers follows the rules recommended by the International Association of Cancer Registries [9]. The percentage of cases registered with unknown primary sites was 0.9%, and the proportion of death certificate only (DCO) was 1.2% in 2014 [10]. We compared the self-reported cancer history of the HEXA study participants at enrollment with records obtained from the KCCR data for the years 1999 through 2012.

We linked the HEXA data with the KCCR data using a resident registration number which is a unique number assigned to individuals for identification in Korea. The KCCR database provides information on cancer cases, including date at cancer diagnosis and the site of the cancer. We selected the following cancers using codes from the International Classification of Diseases (ICD), 10th revision [11]: stomach (C16), colon/rectum (C18–C21), liver (C22), lung (C33–C34), female breast (C50), prostate (C61), thyroid (C73), cervix uteri (C53), bladder (C67) and all cancers combined (C00–C96). We included multiple primary cancers in this study. So if the participant had two different primary cancers, both cancers were counted. All participants who enrolled in 2013 did not have a record of cancer diagnosis in 2013. We excluded participants who reported a cancer diagnosis before 1999 from analysis for positive predictive value, because of incompleteness of the KCCR data before 1999 [7]. We calculated the year of cancer diagnosis based

on age and year at enrollment and age at cancer diagnosis. Cancer diagnoses registered in the KCCR were considered the gold standard.

2.3. Statistical analyses

Because most participants are true negatives, we present sensitivity and positive predictive value – which do not use true negative values – as the main results. The validity of self-reported cancer history was assessed by cancer site, age at enrollment, education level, and the time period from cancer diagnosis to enrollment in the HEXA study. Sensitivity was calculated as the proportion of people who reported a prior cancer diagnosis among people who were registered in the KCCR to determine the likelihood of a true cancer history being obtained by the survey questionnaire. Positive predictive value was calculated as the proportion of people who were registered in the KCCR among people who reported a prior cancer diagnosis to determine the likelihood that participants accurately reported their cancer history. The 95% confidence intervals for both sensitivity and positive predictive value were estimated using the Wilson score method for binomial distribution. All analyses were conducted using SAS software version 9.4 (SAS Institute, Inc., Cary, North Carolina).

3. Results

The distribution of baseline characteristics of the entire HEXA study population is shown in Table 1. Overall, there were 143,965 participants, including 49,947 men and 94,018 women. More than half the participants were in their 40s and 50s at enrollment. Table 2 presents the sensitivity of self-reported cancer history. We

Table 1
Sociodemographic characteristics of HEXA study participants, 2004–2013.

	N (%)	
Sex		
Male	49,947	(34.7)
Female	94,018	(65.3)
Age at enrollment (years)		
35–39	2,981	(2.1)
40–49	48,841	(33.9)
50–59	56,078	(39.0)
60–69	33,108	(23.0)
70–79	2,957	(2.1)
Enrollment year		
2004	381	(0.3)
2005	3,003	(2.1)
2006	11,658	(8.1)
2007	21,924	(15.2)
2008	23,650	(16.4)
2009	24,522	(17.0)
2010	25,385	(17.6)
2011	21,328	(14.8)
2012	7,174	(5.0)
2013	4,940	(3.4)
Highest level of education		
Middle school or below	45,840	(31.8)
High school graduate	59,405	(41.3)
College degree or higher	36,720	(25.5)
Unknown	2,000	(1.4)
Monthly income (10,000 Korean won)		
<200	41,178	(28.6)
200–400	55,035	(38.2)
≥400	32,628	(22.7)
Unknown	15,124	(10.5)

Download English Version:

<https://daneshyari.com/en/article/5524747>

Download Persian Version:

<https://daneshyari.com/article/5524747>

[Daneshyari.com](https://daneshyari.com)