A Review of Major Secondary Data Resources Used for Research in Traditional Korean Medicine

Chunhoo Cheon, Bo-Hyoung Jang*, Seong-Gyu Ko*

Department of Preventive Medicine, College of Korean Medicine, Kyung Hee University, Seoul, Republic of Korea

Received: December 26, 2022 Revised: February 08, 2023 Accepted: March 08, 2023

*Corresponding authors:
Bo-Hyoung Jang
Department of Preventive Medicine, College of Korean Medicine, Kyung Hee University, Kyungheedae-ro 26, Dongdaemun-gu, Seoul 02447, Republic of Korea
Email: bhjang@khu.ac.kr
Seong-Gyu Ko
Department of Preventive Medicine, College of Korean Medicine, Kyung Hee University, Kyungheedae-ro 26, Dongdaemun-gu, Seoul 02447, Republic of Korea
Email: epiko@khu.ac.kr

https://doi.org/10.56986/pim.2023.06.002

ABSTRACT

Research in health care using secondary data is steadily increasing worldwide. In this study, secondary healthcare data was reviewed, so that the information can potentially be used for Korean medicine research. The characteristics of the data, including the variables related to Korean medicine and the method of obtaining data, were summarized. The Korean medicine variables were extracted from the Korean Medicine Utilization Survey, Korea Health Panel, Korea National Health and Nutrition Examination Survey, Medical Service Experience Survey, and the health insurance claims data. Except for health insurance claims data, the data was obtained through relatively simple procedures. There were differences in the characteristics of each secondary data and the extent to which it was used in Korean medicine research. Many Korean medicine studies using secondary data will be conducted in the future and researchers must understand the characteristics of the data and analyze it appropriately.

Keywords: complementary therapies, observational study, review

©2023 Jaseng Medical Foundation. This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

Introduction

Secondary data resources refer to data that has been collected for no specific research purposes but are often used by researchers in the healthcare sector [1]. With the development of computer technology and sophisticated analysis techniques, there has been increasing emphasis on the importance of data analysis in different medical fields. The newly emerged term “digital health” refers to the use of this information to manage illness and health risks [2]. In this context, data analysis and the interpretation of data play important roles. It has been reported that South Korea has been actively building secondary data resources and disclosing a wide range of raw data [3]. Compared with other countries, the Organization for Economic Cooperation and Development, South Korea, Denmark, and Finland showed a high level of availability of healthcare data, maturity, research utilization, and government policy support [4]. In most countries, using secondary data resources stirs up controversy between the industry sectors that intend to utilize the data and those who emphasize the importance of privacy and protection of personal information. In South Korea, three representative laws relating to healthcare data were revised in 2020 and these revisions are expected to facilitate and further promote healthcare data utilization [5].

In South Korea, doctors who majored in traditional Korean medicine (TKM) have the same legal and social status as doctors who studied conventional Western medicine [6]. There is much research into TKM, however, there is only a limited number of Korean medicine studies using secondary data resources, and most of the research topics for these studies are about the use of Korean medicine [7]. Although there has been a gradual increase in the number of published Korean medicine research articles evaluating treatment effects, more types of conditions and diseases need to be investigated. For example, musculoskeletal conditions/diseases are common reasons for patients’ visits to TKM clinics, and account for a significant proportion of consultations [8,9]. Therefore, for TKM to play a key role in this era of digital health, more studies applying various research methods for different conditions/diseases and utilizing secondary data resources are needed, and analysis of medical data should be performed with reference to electronic medical record data. Lifelog data of patients who...
have received or are receiving TKM should eventually lead to promoting health care management where TKM is used. As a first step to promoting Korean medicine research by utilizing secondary data resources, the following should be investigated: the type of data resources available, the variables in each case, and how the resources have been utilized to date.

Previous studies have reviewed Korean medicine research using secondary data resources, but these studies mainly discussed research themes of published papers or specific research methods [7,10]. Therefore, after examining major secondary data resources that can be used for Korean medicine research, each data resource’s characteristics were reviewed and variables related to Korean medicine were summarized, and strategies and methods for utilizing these data resources were presented.

Materials and Methods

This study reviewed the National Evidence-Based Healthcare Collaborating Agency report where domestic secondary data resources are introduced, published research on secondary data resources, studies that used secondary data resources, and the website, and guidelines on the use of data for various agencies that provide secondary data resources. Among different secondary data sources, those that included variables related to TKM were included in the analysis. Data provided by Korean agencies were included, and those that did not have Korean medicine-related variables were excluded from the analysis even if those resources included variables related to Korean medicine in the past.

The agencies that provided secondary data resources included in the analysis, variables related to Korean medicine and sociodemographic information, data acquisition methods, and the brief characteristics were summarized and presented. In addition to variables directly related to Korean medicine, some indirectly related variables that were frequently used as covariates were presented in the analysis.

Results

The names and agencies that provided the data and a brief description of secondary data resources for South Korea, including variables related to Korean medicine are presented in Table 1.

1. Korean medicine utilization and herbal medicine consumption survey

The Korean medicine utilization and herbal medicine consumption survey is a national statistical data survey performed by the National Institute of Korean Medicine Development, which provides data on the status of utilization of Korean medical services, TKM recognition and demand for, and patterns in the utilization of Korean medicine. The survey is divided into the general population and users of Korean medicine services, both populations are representative of the national population aged 19 and older. The Korean Medicine Utilization and Herbal Medicine Consumption Survey did not present a list of variables

Table 1. Overview of Major Secondary Data Resources Including Korean Medicine–related Variables and the Agencies Responsible for the Surveys.

<table>
<thead>
<tr>
<th>Name of data resources</th>
<th>Responsible agencies</th>
</tr>
</thead>
</table>
| Korean Medicine Utilization Survey                  | National Institute for Korean Medicine Development  
  https://nikom.or.kr/  
  Ministry of Health and Welfare  
  https://www.mohw.go.kr |
| Korea Health Panel                                  | Korea Institute for Health and Social Affairs  
  https://www.kihasa.re.kr/  
  https://www.khp.re.kr  
  National Health Insurance Service  
  https://www.nhis.or.kr |
| Korea National Health and Nutrition Examination Survey | Korea Disease Control and Prevention Agency  
  https://www.kdca.go.kr  
  https://knhanes.kdca.go.kr |
| Survey on the Experience with Healthcare Service    | Korea Institute for Health and Social Affairs,  
  https://www.kihasa.re.kr  
  Ministry of Health and Welfare  
  https://www.mohw.go.kr |
| Health insurance claims data                        | National Health Insurance Service  
  https://www.nhis.or.kr  
  Health Insurance Review and Assessment Service  
  https://www.hira.or.kr |
because most of the survey items were related to Korean medicine. The survey further categorizes the 2 populations of respondents into outpatients and inpatients to examine the experience of Korean medicine utilization, individual perception of Korean medicine, the status of Korean medicine utilization, intention to use the service and recommendation to others, their health condition, and types of national health insurance and other private insurance. The sociodemographic variables surveyed included sex, age group, marital status, education level, occupational status, and household income. The microdata of the Korean Medicine Utilization and Herbal Medicine Consumption Survey can be acquired after agreeing to the terms of use on its website (https://www.koms.or.kr/), submitting a plan of data utilization, and obtaining approval in the review process. The survey has been conducted every three years since 2008, and the microdata generated became available to the public in 2017; the microdata of the 2017 and 2020 surveys are currently accessible to the public.

2. Korea health panel

The Korea health panel (KHP) specifically aims to build a database to analyze patterns in the utilization of healthcare services, the amount of healthcare expenditure, and influencing factors. Currently, the 2nd KHP survey is in progress, and 8,500 households nationwide are included in the survey sample based on the census registered in 2016. Since the KHP survey is a panel survey, it is conducted on the same participants yearly. In the KHP, Korean medicine-related variables are examined by asking for the types of medical institutions visited for healthcare services, expenditures related to healthcare services, and collaborative treatment. The survey also investigates the purpose of inpatient and outpatient services in Korean medicine, the conditions/diseases treated, and treatment methods (Table 2). The KHP data can be downloaded via e-mail by completing the consent form for the use of data on its website (https://www.khp.re.kr/) and sending the form to the person in charge.

3. Korea national health and nutrition examination survey

The Korea national health and nutrition examination survey (KNHANES) provides national statistics on health conditions, the prevalence of chronic conditions/diseases, and food and nutrition intake for the general population in South Korea. The survey is based on a complex sample design to ensure the representativeness of the Korean population. The survey has been conducted over eight periods (three years per period); in the 8th survey, 4,800 households were included. The Korean medicine-related variables included in the survey were: an item on checking the intake of Korean herbal medicine as a method of weight control, and the inclusion of Korean medicine hospitals/clinics when surveying the experience of inpatient or outpatient care (Table 3). The 3rd KNHANES in 2005 included the use of Korean medicine services for one year, intention for use in the future, types of Korean medicine services used,
and co-payments for Korean herbal medicine; however, these items were only included in the KNHANES 2005. The raw data from the KNHANES can be obtained from the KNHANES website (http://knhanes.kdca.go.kr/) after user information registration and consent on the use of data. Currently, data from 1998 to 2020 are publicly available.

4. Medical service experience survey

In the Medical service experience survey, various items related to healthcare services were investigated according to the health condition, experience with healthcare service, perception of the healthcare system, awareness of burden in medical expenditure, experience of health screening/examination, personal characteristics, and household characteristics. Using the survey, it was possible to identify whether the medical institution that the respondent recently visited was a Korean medicine hospital/clinic and whether the respondent used the service frequently (Table 3). Thus, the survey serves as a data resource that allows comparative analysis of the participants’ healthcare service experience between Korean and conventional Western medicine. The raw data from the Survey on the Medical Service Experience Survey can be obtained by completing an application form, pledge, and research proposal from the Health and Welfare Data Portal (https://data.kihasa.re.kr/) of the Korea Institute for Health and Social Affairs and upon approval of the submitted forms. Data from 2017 to 2020 are currently available to the public.

5. Health insurance claims data

The health insurance claims data of South Korea includes treatment history information, prescription history information, and health screening information for users of medical institutions in Korea. Since the National Health Insurance Service (NHIS) is a single insurer for national health insurance, it holds the data for almost the entire national population, making it the largest data set among domestic data related to healthcare. Variables related to Korean medicine include whether the medical institution used for healthcare services were Korean medicine

<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Category</td>
</tr>
<tr>
<td>Korea National Health and Nutrition Examination Survey</td>
</tr>
<tr>
<td>Socio-demographic variables</td>
</tr>
<tr>
<td>Medical Service Experience Survey</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Health insurance claims data</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
hospitals or clinics and whether the type of services used were inpatient or outpatient care from institutions of Korean medicine. Since the survey includes most of the treatments related to the claim of health insurance coverage, details of the treatments can be checked, such as acupuncture, moxibustion, cupping, Chuna manual therapy, Korean medicine physical therapy, and prescription of herbal medicines (Table 3). Various analyzes on Korean medicine are possible by determining Korean medicine-related variables, other medical use-related variables, and user characteristics variables in the insurance data (Table 3). Health insurance claims data is largely divided into data provided by the NHIS and those provided by the Health Insurance Review and Assessment Service (HIRA). There are customized data that provides data on the total population according to the requests, and sample cohort data, and health screening cohort that provide data of the sample representative of the entire population of South Korea. In addition, there are the elderly cohort, inpatient cohort, and pediatrics and adolescent cohorts that provide data by extracting a sample that can represent the target cohort. The working women cohort and health screening of infants and children cohort provided by the NHIS are no longer available. Detailed descriptions of the currently available sample cohort data are shown in Table 4. The sample cohort of the HIRA consists of general information on specification, medical treatment history, diagnosis statements, outpatient prescription history, and information on medical care institutions. There are two routes to acquiring the data. Data provided by the HIRA can be obtained through the Healthcare Bigdata Hub (https://opendata.hira.or.kr/); data provided by the NHIS are obtained through the National Health Insurance Sharing Service (https://nhiss.nhis.or.kr/). Before the application for data, the Institutional Review Board approval of the researcher’s affiliated institution must be obtained, and the study protocol and Institutional Review Board review result notification must be submitted for deliberation. For customized data, a data request statement with a selection of the variables should also be completed and submitted. The method of access to health insurance claims data frequently changes. Currently, customized data can be accessed for analysis by directly visiting the analysis center which is operated by the NHIS and the HIRA. Other cohort data can be accessed remotely on a virtual PC built by the data provider and used for analysis. The characteristics of each database are summarized as follows. Similar to cohort data, the KHP acquires data from the same participants every year. Using this, it is possible to observe changes in the status of the participants’ medical use. Since health insurance claims data is data for the entire population of South Korea, it can be used as a cohort data source. Among the health insurance claims data, the

<table>
<thead>
<tr>
<th>Provider</th>
<th>Cohort name</th>
<th>Characteristics</th>
</tr>
</thead>
</table>
| NHIS             | National sample cohort       | Target population: The whole nation who have maintained their qualifications as health insurance beneficiaries and medical aid beneficiaries for one year in 2006  
Duration: 2002 to 2019  
Size: 1,000,000  
Variables: Socioeconomic status (including disability and death), medical utilization, information on medical care institutions, status of elderly long-term care |
|                  | National health screening cohort | Target population: Examinees of general health examination aged 40 to 79 years from 2002 to 2003  
Duration: 2002 to 2019  
Size: 510,000  
Variables: Socioeconomic status (including disability and death), medical utilization, information on medical care institutions |
|                  | Senior cohort database       | Target population: People aged 60 to 80 in January 2008  
Size: 510,000  
Variables: Socioeconomic status (including disability and death), medical utilization, information on medical care institutions, status of elderly long-term care |
| HIRA             | National patient sample      | 2009-2016: 3% of all patients (approximately 1.4 million)  
After 2015: 2% of all patients (approximately 1 million) |
|                  | National inpatient sample    | 2009-2016: 13% of inpatients (approximately 1 million)  
After 2016: 10% of inpatients (approximately 750,000)  
Includes 1% of outpatients (excluding 2016 and 2019 and beyond) |
|                  | Elderly patient sample       | 2009-2016: 20% of elderly patients (over 65) (approximately 1 million)  
After 2017: 10% of elderly patients (approximately 700,000) |
|                  | Pediatric patient sample     | After 2009: 10% of pediatric patients (under 20) (approximately 1 million) |

HIRA = Health Insurance Review and Assessment Service; NHIS = National Health Insurance Service.
national sample cohort provided by the NHIS can be used as a cohort data source because the included participants were followed up for a certain period of time. However, the national patients sample provided by the HIRA is not cohort data because the participants are newly selected every year. In other data, such as data generated by the KNHANES and the Korean Medicine Utilization Survey, participants are newly selected to represent the entire population each time. The most detailed data available for Korean medicine research is the Korean Medicine Utilization Survey. This is because the treatment and level of satisfaction received from a Korean medicine institution for the treatment of a specific condition/disease are also investigated. However, it has the disadvantage that the sample size is not large enough to study specific conditions/diseases. Since health insurance claims data has a large number of participants and treatment method data, it can be used when conducting research on specific conditions/diseases. The source of data for the HIRA and the NHIS is the national health insurance data of South Korea, which is the same. It is provided in different forms by the two institutions. The most significant difference is that the NHIS can provide cohort data and the HIRA cannot.

**Discussion**

Among secondary healthcare data resources available in South Korea, this study selected those with variables related to Korean medicine that can be utilized for Korean medicine research and investigated the Korean medicine-related variables including the resources, data characteristics, and data acquisition methods. In addition to the differences in the surveyed variables, each data resource had different characteristics, and these differences should be well understood and utilized for research.

The healthcare data that was associated with the most diverse range of variables relating to the utilization of Korean medical services was obtained from the Korean Medicine Utilization and Herbal Medicine Consumption Survey. However, the disadvantage of this survey is that it is a survey of service users which although it was conducted on 1,010 outpatients and 904 inpatients, the sample size was too small when specific conditions/diseases were analyzed [11]. Therefore, most studies using this data have mainly investigated characteristics, or satisfaction of the general public, or service users [12,13]. A previous study did analyze the current status of treatment methods for low back pain, a condition with a relatively high proportion of participants, under the amount of expenditure for herbal medicines. It should be noted that the total cost of purchasing herbal medicines as reported by the KHP is the question of the cost spent at “pharmacy or Korean medicine pharmacy.” Until 2019, the cost of herbal medicines prepared at Korean medicine clinics or hospitals was separately investigated under the amount of expenditure for herbal medicines. However, since the KHP survey in 2020, this information was included in the item of medical expenditure of each household, and analysis should take this change into account. Recently, a study focused on a specific disease, cancer, and the results of analyzing the incidence of cancer on the effects from the use of healthcare services, including Korean medicine, were published [19]. The KHP data allows for the most extensive analysis of the use of healthcare services and medical expenses and has the advantage of enabling the analysis of the dynamics through repeated questions for the same household over time. Therefore, continuous interest from Korean medicine researchers is essential for the KHP survey.

The KNHANES contains a wide range of variables related to health and nutrition, thus drawing much interest from researchers in various fields such as medicine, dentistry, public health, and nutrition [20-23]. However, because there are few variables related to Korean medicine, this limits the interest of Korean medicine researchers in the KNHANES. Until the 2015 survey, the medical institutions used were classified by type; thus, analyzing the use of services by Korean medicine clinics/hospitals was possible. A previous study analyzed the determinants of the combined treatment of Korean and Western medicine for patients with hypertension using the 2010-2014 data [24]. Since the KNHANES in 2016, inpatient and outpatient care included the treatment received from Korean medicine institutions, and no distinction was made between institutional types; therefore, analyzing the utilization of Korean medicine was
impossible.

The Community Health Survey data for evaluating health-related factors by region are similar to those of the KNHANES, but no variables are related to Korean medicine. Since all of these healthcare data serve as basic data for national health-related policy-making, Korean medicine-related variables should be included in the KNHANES and Community Health Survey to incorporate Korean medicine into the national healthcare policy actively. Although these two surveys have few variables related to Korean medicine, they include various lifestyle factors in their surveys. Therefore, Korean medicine researchers with an interest in lifestyle factors could analyze these surveys. One study analyzed various lifestyle habits based on the KNHANES data in association with Yangsaeng therapy, a type of Korean medicine, and another study analyzed the parameters of herbal medicine intake during an investigation into weight control strategies [25,26].

Since The Medical Service Experience Survey classifies medical institutions that have been visited into conventional Western medicine hospitals and clinics, Korean medicine hospitals/clinics, and dental hospitals/clinics, the survey allows an understanding of each institution’s level of service, and the public’s perception of that healthcare system. Since most variables focus on healthcare services, researchers have little interest in the Medical Service Experience Survey compared with other data resources. However, some studies have analyzed the determinants of satisfaction with healthcare services [27,28]. As for Korean medicine-related papers, a study analyzed factors influencing the use of Korean and Western medicine [29]. Although this is a data resource with few health-related variables, since the level of satisfaction with the service provided by medical institutions also affects treatment outcomes, more research should be conducted with an interest in different healthcare services.

Data from health insurance claims draws much interest and attention from many researchers in the field of medicine because it contains an extensive range of items and has data for the entire Korean population. Since healthcare data of this scale is rare, even on a global level, the research utilizing the data and the health insurance claims data has attracted much attention [30-32]. However, there are some drawbacks. Firstly, since the data is administrative information for insurance coverage claims, it may differ from the actual health information of the users. Secondly, the data does not include the treatment information not covered by health insurance. For example, the decoction, which accounts for a significant portion of Korean medicine services, cannot be analyzed using the health insurance claims data because the national health insurance does not cover it. As for Korean medicine research based on the health insurance claims data, a previous study reported that receiving acupuncture was associated with a low lumbar surgery rate, and another study reported that receiving acupuncture was associated with a low rate of knee surgery for osteoarthritis [33,34]. Some studies have compared the status of using traditional medicine between South Korea and Taiwan, a country that has traditional medicine doctors, operates a health insurance system, and discloses claims data, similar to South Korea [35,36]. Since many comparative studies on treatment effects are conducted using health insurance data, researchers of Korean medicine must continue to pay attention to this type of data. Compared with other secondary data resources, requesting data from health insurance claims to analysis takes a substantial amount of time; therefore, a research plan considering this aspect is required. It has become possible to statistically analyze the health insurance claims data in combination with other healthcare data, drawing the interests of many researchers.

This study had several limitations. Among the secondary data resources in South Korea, resources that include Korean medicine-related variables may not have been included in this analysis. The data resources for this review were selected considering the ease of obtaining data for analysis and the applicability to Korean medicine research. Variables of data resources may vary depending on the timepoint of access, and for health insurance data, variables that can be analyzed may vary depending on policies. However, these factors were not considered in this study. Researchers must consider the accessibility of healthcare data and the variables of the data resources during research planning. It is hoped that this review will serve as a useful guide for researchers to conduct on Korean medicine research based on secondary data resources.

This study reviewed secondary data resources that can be used for Korean medicine research, and the methods for utilizing each type of data resource have been discussed. Since each data resource has its advantages and disadvantages, it is necessary to identify the characteristics of each data resource and select the most appropriate data resource for the research topic to conduct analysis.

Author Contributions

Conceptualization: CC, BHJ, and SGK. Methodology: CC, BHJ, and SGK. Writing – Original Draft: CC. Writing – Review & Editing: CC, BHJ, and SGK.

Conflicts of Interest

The authors declare that there was no conflict of interest.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.
Ethical Statement

Ethical approval was not applicable because this was a review article.

Data Availability

Data sharing was not applicable to this article as no datasets were generated or analyzed during the current study.

References


