Urological Oncology

The Establishment of K-CaP (the Multicenter Korean Prostate Cancer Database)

Dong Hoon Lee, Seung Hwan Lee, Koon Ho Rha, In Young Choi[1], Ji Youl Lee[2], Sae Woong Kim[2], Sangchul Lee[3], Sung Kyu Hong[3], Seok-Soo Byun[3], In Gab Jeong[4], Jun Hyuk Hong[4], Choung-Soo Kim[4], Hwang Gyun Jeon[5], Hyun Moo Lee[5], Byung Ha Chung

Department of Urology, Urological Science Institute, Yonsei University College of Medicine, Seoul, [1]Graduate School of Management and Policy, The Catholic University of Korea, Seoul, [2]Department of Urology, Seoul St. Mary's Hospital, The Catholic University of Korea College of Medicine, Seoul, [3]Department of Urology, Seoul National University Bundang Hospital, Seongnam, [4]Department of Urology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, [5]Department of Urology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea

Purpose: The purpose of this article was to announce the establishment of the multicenter Korean Prostate Cancer Database (K-CaP) and to provide urologists with details about K-CaP's methodology.

Materials and Methods: The initial participating K-CaP institutions include five medical centers in Korea. First, we registered prostate cancer patients who underwent radical prostatectomy as the basic background data. K-CaP is poised to combine these initial observational longitudinal studies with those of other eligible institutions as the database grows. All current prostate cancer patients in Korea are able to be registered into the Web-based database system and thereby have a role in several observational studies. The structure of the database for K-CaP was developed by matching it with the respective data from different studies. The operability of the K-CaP database system was verified by using the existing databases from three participating institutions.

Results: The analysis of clinicopathologic characteristics of patients with the use of the Web-based database was successfully conducted. We confirmed the accurate operation of the Web-based database system without any difficulties.

Conclusions: We are announcing the establishment of K-CaP the first database of comprehensive observational longitudinal studies about prostate cancer in Korea. The database will be successfully maintained by sufficiently and continuously updating all patient data covering several treatments. Complete statistical results for registered prostate cancer patients are forthcoming for the basic background data to establish the database. Even though much trial and error are expected during the development process, we expect that K-CaP will eventually become one of the most powerful longitudinal observation databases.

Keywords: *Database; K-CaP; Korea; Prostate neoplasms*

Corresponding Author:
Byung Ha Chung
Department of Urology, Yonsei University College of Medicine,
50 Yonsei-ro, Seodaemun-gu, Seoul 120-752, Korea
TEL: +82-2-2019-3470
FAX: +82-2-3462-8887
E-mail: chung646@yuhs.ac

INTRODUCTION

The prevalence of prostate cancer in Korea quadrupled between 2002 and 2008, with the highest increased incidence rate in total forms of malignancy [1]. The incidence of prostate cancer in Korea increased up to 24.8 per 100,000 men in 2009 in comparison with 13 per 100,000 men in 2008 [2]. Certain environmental elements have had an effect on the increased rate of prostate cancer, including the transition to Western dietary habits among Koreans and an aging

population because of the rise in average life expectancy. In addition to environmental causes, the medical development of laboratory diagnoses and prostate-specific antigen screening campaigns by the Korean Urological Association and other health organizations have helped to raise public awareness about the increase in prostate cancer in Koreans [2,3].

Despite the rapid increase in prostate cancer incidence in Korea, no published multicenter data on practical and clinical changes in Korean patients with prostate cancer are available. In the United States, a database application project known as the Cancer of the Prostate Strategic Urologic Research Endeavor (CaPSURE) was initiated in 1995 to establish a web-based database for longitudinal observations of prostate cancer patients in natural settings [4]. The project began with 10 participating healthcare centers and increased to 26 centers in 1 year. Currently, CaPSURE is one of the most powerful prospective study groups for prostate cancer in the world consisting of approximately 14,000 registered prostate cancer patients [5]. The most recent database study in Asia, the Japan Study Group for Prostate Cancer (J-CaP), was developed in 2001 [6]. The J-CaP database comprises 17,872 prostate cancer patients from prospective studies and was developed to improve patient care.

Against this background, the multicenter Korean prostate cancer database (K-CaP) was created in 2011 by combining several urologic healthcare institutions into a nationwide multicenter database for prospective studies about prostate cancer. The purpose of K-CaP is to gather basic information about Korean prostate cancer patients and analyze the clinical and oncological outcomes of prostate cancer to improve patient care. The purpose of this article was to declare the establishment of K-CaP, to provide urologists with an overview of the K-CaP methodology, and to present pilot test results from the first umbrella database comprising patient information from three participating institutions (Gangnam Severance Hospital, Seoul St. Mary's Hospital, and Seoul National University Bundang Hospital).

MATERIALS AND METHODS

1. Organization of K-CaP
1) Participating institutions (in alphabetical order)
Participating institutions include Asan Medical Center, Samsung Medical Center, Seoul National University Bundang Hospital, Seoul St. Mary's Hospital, and Yonsei University Severance Hospital.

2) Eligible institutions
Eligible institutions include all urologic institutions in Korea. Any institution that wants to participate in K-CaP must first obtain approval from their ethics committees (or Institutional Review Board). For every eligible institution, all patients with newly diagnosed prostate cancer will be registered in the K-CaP web-based electronic server. All

registered patient information will be updated periodically.

2. Data collection, entry, follow-up, and retrieval
1) Data collection
After a diagnosis of prostate cancer, patients are invited by their urologists to join the study. All patients with biopsy-proven prostate cancer are offered enrollment in the study, regardless of disease stage, severity, or type of treatment. Included in the database is all of the physician-gathered general information about a patient, including sociodemographic data. In addition, all of the patient's clinical information at diagnosis is gathered, including diagnostic imaging, laboratory test results, pre-existing and postdiagnosis comorbidities, medical history, treatment types (e.g., active surveillance, androgen deprivation/hormonal medications, brachytherapy, cryotherapy, external beam radiation, hormone refractory and chemotherapy agents, and radical prostatectomy), neoadjuvant and adjuvant treatments, and other nontreatment-related medications. For patients who have undergone radical prostatectomy, detailed perioperative variables are measured and registered, including procedure types, intraoperative or postsurgical complications, surgical pathology results, and clinical and oncological survival outcomes. The patient files for the K-CaP database include approximately 1,000 required clinical variables.

2) Entry
The observational study of prostate cancer patients is the ultimate goal of the K-CaP database. Accordingly, the web-based electronic case report form (e-CRF) was developed by using the StarTrial system, which is a commercially developed electronic data capture system. This e-CRF system can be accessed over the Internet on a 24-hour basis by coordinators and investigators at eight participating clinics (Suppl. 1). Suppl. 1A is a screenshot of registration for prostate cancer patients. Suppl. 1B is a screenshot of the baseline parameters of the registered patients. The system was developed at the Catholic University of Korea, South Korea, and was previously tested with the help of staff at other participating sites. The database system was implemented with a Microsoft SQL Server running on Microsoft NT servers and was programmed with Java. This was accomplished by directly partnering the electronic databases of the study's participating institutions with the registered patient files. All legal requirements for privacy protection were respected in this procedure.

3) Follow-up
According to the privacy rules of K-CaP, cases for registered patients are to be updated once every 3 months in terms of test dates, changes in treatment, and progress data as a matter of course for maintaining the K-CaP patient forms until death or patient withdrawal. Follow-up with respect to morbidity and death from prostate cancer is as-

sured by cooperation with national statistics. When patients die, the date, location, and cause of death are recorded in the database.

4) Data retrieval

After completion of the data control for registered patients, analytic reports are prepared for prospective studies and research to improve the care of prostate cancer patients. These data summaries evaluate clinical occurrences, patient quality of life, economic impact, and oncological outcomes and compare types of treatment by stage and practice among patients with prostate cancer in Korea. The complex studies use well-sorted variables from the K-CaP database, together with adjustment methods including case controls, standardization, regression trees, and multivariate regression approaches for optimal analysis. The entry form contains five categories: patient demographics, clinical parameters, pathological parameters, other treatment parameters (hormone therapy or radiotherapy), and parameters associated with follow-up. The data can be extracted according to access rights. Access given to an investigator allows extraction of only his or her patient's data. However, one given data management access can extract the whole data set. These two access rights allow the ability to extract data through the web system. The data format for data extraction includes CSV format, MS Excel (xls format), or MS Access (accdb format). The System Manager is able to select data by accessing the SQL Server system directly using the SQL program (Suppl. 2).

5) Quality control

To ensure the quality of the database, it is most important that accurate data be provided by the investigators. However, if a strange code is entered into the database, the database system automatically checks the input code and will not process the input of the strange data set. The standard input code is determined under the participating investigators' agreement. Investigators are also able to compare their own raw data with data in the K-CaP database at anytime, because each institution's data can be extracted if the investigator has access rights.

3. The e-CRF system for K-CaP

A pilot test was performed to determine whether to use the standard e-CRF system based on the limited perioperative information of prostate cancer patients who underwent radical prostatectomy. Between January 2006 and December 2010, 858 prostate cancer patients who underwent radical prostatectomy at three institutions (Gangnam Severance Hospital, Seoul St. Mary's Hospital, or Seoul National University Bundang Hospital) were registered in the e-CRF system of K-CaP. The database collected preoperative variables, including patient age, prostate-specific antigen level at diagnosis, preoperative Gleason score, clinical stage, and the initial procedure chosen for radical prostatectomy. Postoperative variables were also registered, including pathologic Gleason score and stage. For

each patient, the T stage was determined in accordance with tumor-node-metastasis (TNM) categories, as published in 2010 (American Joint Committee on Cancer, 7th edition). All of the pathologic parameters that influence the determination of TNM staging were registered in detail. This process is important to prepare for the periodic renewal of TNM categories. The first step was to gather the prostate cancer database files from each institution using the Excel format and to send these files to an e-CRF system manager controlling the data input/output of the web-based database. Subsequently, the manager converted the coded Excel files for use in the e-CRF system. Finally, we were able to obtain outcomes from the multicenter e-CRF database according to several conditions that we initially established to fit the goals of K-Cap.

RESULTS

The multicenter K-CaP test database, which contains data for a total of 858 prostate cancer patients who underwent radical prostatectomy at three institutions, represents a single source for comprehensive information about the patients and the disease. A Java program was developed to upload Excel data into the e-CRF database by matching each column name. If the uploaded Excel file executes the Java program, the data is automatically entered into the designated column. Verification of the upload process can be conducted in two steps. The first step is to compare the number of records between the uploaded Excel file and the SQL Server database. The second step is to transform data from the SQL Server into Excel format and then compare it with the original data set.

We established accurate operation of the web-based database system without any difficulties. We confirmed that the process of converting data to the e-CRF system from Excel files is easy and exact and that output data from the web-based database system are quickly retrievable according to many weighted limitations.

DISCUSSION

A longitudinal observation database has many advantages for prospective studies and research. Complete treatment histories and patient information allow for comparison of different outcomes. As the number of patients enrolled in the database increases, the database will help to categorize patients and identify historical controls for various purposes. In existing research about prostate cancer, patients have been selected with indiscriminate consideration of treatment type or stage, regardless of the varying aims of specific studies.

For example, the life expectancy of Korean prostate cancer patients with multiple bone metastases is not yet known. In addition, we do not yet know the exact clinical course of Korean prostate cancer patients after treatment. Even though we reviewed the results of many studies and made conclusions about disease progression [7-12], the re-

sults were from single-institution cohort studies with small sample sizes. Thus, these studies insufficiently represent the characteristics of prostate cancer in Koreans. Furthermore, despite the many studies of prostate cancer in Koreans, most of our knowledge is still based on studies from a Western database of prostate cancer. Thus, we needed a prostate cancer database of our own, and we believe that K-CaP is that database. However, under a system of longitudinal observation, the collection of visible results for analysis will require more time than is required for studies focusing on retrospective design. We expect that at least 5 years will be required to obtain tangible and qualitative results from the K-CaP database. If the initial results are positive, however, lasting products will exist in the database on an ongoing basis. In the case of CaPSURE, the first published results about trends in prostate cancer treatment (2003) were published 7 years after initial development of the CaPSURE project. Currently, a PubMed search with the term "CaPSURE" delivers more than 100 relevant articles, including articles on the J-CaP project in Japan.

Achievement of this long-term plan requires verification of the web-based database system as the first step toward the establishment of K-CaP. The pilot test with prostate cancer patient databases from three institutions showed the feasibility of a web-based system for K-CaP. The important checkpoint is that the results of the pilot test are not intended to be used to make analytical conclusions about the specific subjects from the three institutions. The intent of the pilot test was to ensure that the K-CaP database system was operating accurately before other participating institutions provided proprietary data.

As this article goes to press, patient data from the remaining institutions in the Korean healthcare system are being registered. We estimate that approximately 4,000 cases from participating institutions will be registered as the basic background data for the K-CaP database (620 patients from Asan Medical Center, J.H.H., C.S.K.; 1,400 patients from Samsung Medical Center, H.M.L.; 340 patients from Seoul National University Bundang Hospital, S.K.H., S.S.B.; 300 patients from Seoul St. Mary's Hospital, J.Y.L., S.W.L.; and 1,235 patients from Yonsei University Severance Hospital, K.H.R., B.H.C.). The patients with newly diagnosed prostate cancer from the participating institutions will be registered, including the required entry information for K-CaP. According to the patients' privacy agreement, basic information about the patients will be encrypted by using a program unique to the K-CaP database and will remain unknown to our team and even to the database manager. Thus, the privacy of the patients registered in K-CaP is guaranteed.

Developing a uniform format for the K-CaP database was one of the main concerns of the project staff at each of the participating institutions. For example, because of different policies at each institution, differences in preoperative prostate volume measurements (by transrectal ultrasonography) existed between the participating institu-

tions. Some institutions measured prostate volume preoperatively in all patients, but the procedure was performed at different clinical stages at some institutions. Furthermore, each institution had different postoperative follow-up periods over which to conduct patient orientation surveys, such as the International Index of Erectile Function questionnaire. In addition to policy differences between the institutions, internal problems concerning a uniform database format also existed. For example, pathologists at different institutions reported tumor volumes after radical prostatectomy differently. Some pathologists reported the precise weights of the tumors in grams, whereas others reported approximate volumes using relative measurements (e.g., V1, V2, and V3). Importantly, we previously identified several problems, such as those mentioned above, and have discussed solutions intended to achieve a uniform format for the database. We will continue to revise and develop the K-CaP format in an effort to provide the flexibility needed to achieve the long-term goals of the database.

Another important consideration in the management of the K-CaP database is how best to maintain continuous updates about patients. Data maintenance is the most practical problem faced by every participant in the database. It is difficult to predict how many new prostate cancer patients from each participating institution will be enrolled in K-CaP. Likewise, it is difficult to update K-CaP data periodically if the number of enrolled patients is too large, even before the K-CaP settings are ready. However, it is well known that the design of a longitudinal observational study ensures that the quality of the data is more important than the quantity of the data. To solve this problem, we enrolled a limited number of prostate cancer patients representative of a particular institution rather than all of the prostate cancer patients at that institution. After determining the appropriate K-CaP settings through trial and error, we intend to enroll all prostate cancer patients in the K-CaP database and to include patients from any institution that wants to participate in K-CaP.

One last consideration is the importance of K-CaP's collaboration with CaPSURE and J-CaP. K-CaP is intended to be used not only for domestic investigation but also for international participation and collaboration. Even though K-CaP is still in the preparatory stage, we have already received valuable advice and support from the working staff and chairmen of CaPSURE and J-CaP. They have shown sincere interest in K-CaP and a willingness to collaborate. Collaboration will best be achieved if the format of the K-CaP database is analogous to that of the other databases. The K-CaP database was developed considering the original Excel database of CaPSURE and J-CaP. The CaPSURE database has many variables, for which each treatment of prostate cancer was equally taken into account. In contrast, the J-CaP database focuses on hormonal treatment. This probably reflects the different trends in prostate cancer treatment between the United States and Japan. The ability to reference these different data-

bases will help to enrich the K-CaP database, so that our database can cover many fields.

In this article, we announced the establishment of K-CaP on behalf of all the participants in the K-CaP project. The main purpose of K-CaP is to establish and maintain a large database of randomized clinical trials and observational longitudinal studies of Korean prostate cancer patients. Several retrospective studies about prostate cancer in Korea have been published in domestic and international journals. However, we have always been keenly aware of the limitations of retrospective studies. We believe that mutual cooperation is needed to establish the K-CaP project as being equal to the CaPSURE and J-CaP projects in the United States and Japan, respectively. Although our beginnings have been humble, we are confident that our dedication and ongoing efforts will result in K-CaP becoming a world-renowned longitudinal observation database, along with CaPSURE and J-CaP.

CONCLUSIONS

This article announces the development and establishment of K-CaP as the first database for comprehensive data collection about prostate cancer patients in Korea for the purposes of research and improved patient care. This study tested the web-based system of the K-CaP database to analyze coded Excel files from three institutions. The system operated precisely, and the pilot test verified that the web-based database system is suitable for K-CaP. The system processes will run successfully as long as sufficient and updated data is continuously provided to the system manager. As soon as possible, complete statistical results of registered prostate cancer patients will be reported for basic background data.

CONFLICTS OF INTEREST
The authors have nothing to disclose.

SUPPLEMENTARY MATERIALS

Scan this QR code to see the supplementary materials, or visit http://kjurology.org/src/sm/kju-54-229-s001.pdf.

REFERENCES

1. Statistics Korea [Internet]. Daejeon: Statistics Korea; [2012 Mar 10]. Available from: http://kostat.go.kr/portal/english/index.action.
2. Jung KW, Park S, Kong HJ, Won YJ, Lee JY, Seo HG, et al. Cancer statistics in Korea: incidence, mortality, survival, and prevalence in 2009. Cancer Res Treat 2012;44:11-24.
3. Park SK, Sakoda LC, Kang D, Chokkalingam AP, Lee E, Shin HR, et al. Rising prostate cancer rates in South Korea. Prostate 2006;66:1285-91.
4. Lubeck DP, Litwin MS, Henning JM, Stier DM, Mazonson P, Fisk R, et al. The CaPSURE database: a methodology for clinical practice and research in prostate cancer. CaPSURE Research Panel. Cancer of the Prostate Strategic Urologic Research Endeavor. Urology 1996;48:773-7.
5. Cooperberg MR, Broering JM, Litwin MS, Lubeck DP, Mehta SS, Henning JM, et al. The contemporary management of prostate cancer in the United States: lessons from the cancer of the prostate strategic urologic research endeavor (CapSURE), a national disease registry. J Urol 2004;171:1393-401.
6. Akaza H, Usami M, Hinotsu S, Ogawa O, Kagawa S, Kitamura T, et al. Characteristics of patients with prostate cancer who have initially been treated by hormone therapy in Japan: J-CaP surveillance. Jpn J Clin Oncol 2004;34:329-36.
7. Lee DH, Lee SH, Lee DH, Chung MS, Chung BH. Are men who undergo radical prostatectomy with lower urinary tract symptoms at an increased risk for aggressive prostate cancer? Korean J Urol 2011;52:819-23.
8. Park S, Kim SC, Kim W, Song C, Ahn H. Impact of adjuvant androgen-deprivation therapy on disease progression in patients with node-positive prostate cancer. Korean J Urol 2011;52:741-5.
9. Kim YJ, Kim SO, Ryu KH, Hwang IS, Hwang EC, Oh KJ, et al. Prostate cancer can be detected even in patients with decreased PSA less than 2.5 ng/ml after treatment of chronic prostatitis. Korean J Urol 2011;52:457-60.
10. Lee SH, Park KK, Chung MS, Chung BH. Clinical features of familial or hereditary prostate cancer in korean men: a pilot study. Korean J Urol 2011;52:9-12.
11. Seo HS, Lee NK. Predictors of PSA screening among men over 40 years of age who had ever heard about PSA. Korean J Urol 2010;51:391-7.
12. Kwon JK, Chang IH, Kim TH, Myung SC. Changes in prostate cancer pattern according to prostate-specific antigen screening test. Korean J Urol 2009;50:439-44.