Development of biomarker positivity analysis system for cancer diagnosis based on clinical data

Jung Min Yun[a], Sung Jin Hwang[a], Sung Yong Anh[a], Seung Min Lee[a], Peter Kang[a], Jeong Eun Lee[b], Eun Sil Yoon[b], Jin Wook Choi[c], Sung-Hye Park[c], Jung-Wook Seo[d] and Peom Park[a,*]

[a]*Department of Industrial Engineering, Ajou University, 206 Worldcup-ro, Yeongtong-gu, Suwon, Korea*
[b]*Interdiciplinary Program for Bioengineering, Graduate School, Seoul National University, Seoul, Korea*
[c]*Department of Biomedical Engineering, College of Medicine, Seoul National University, Seoul, Korea*
[d]*Department of Pathology, Seoul National University Hospital, College of Medicine, Seoul National University, Seoul, Korea*

Abstract. In Korea, there were 224,000 new cases of cancer and 75,334 deaths caused by cancer in 2013, which was three times more than the number of death caused by heart disease, the second leading cause of death. This study proposes a biomarker positivity analysis system based on clinical data, for personalized diagnosis and therapy of cancer. Data of 78,912 cases were obtained from immunopathology and surgical pathology reports. Data on sex, age, organ, diagnosis, and biomarkers were entered into a database. To verify the reliability of the clinical data, an additional 50,450 cases from positivity-related research papers were added. The proposed biomarker positivity analysis system makes it possible to extract and combine information for searching. The positivity values are in graphical and tabular format for ease of use. With a link to the internal network of the hospital, real-time pathology reports are available. Twenty-five pathology specialists are chosen as subjects to further confirm the reliability of this system; primary assessment results demonstrate a satisfaction level of 4.7 out of 5 and a concordance rate of 79% with positive data under the same conditions as reported in the literature. In the present study, analysis methods and platforms using large volumes of clinical and literature data are developed for cancer prognoses. It is expected that these tools will benefit both healthcare professionals and non-professionals involved in cancer diagnosis and treatment.

Keywords: Cancer, biomarker, clinical data, positive rate

1. Introduction

According to the World Health Organization (WHO), the global cancer cases surpassed 14 million in 2012, and the annual number of recorded deaths caused by cancer reached 8.2 million [1]. In Korea, the number of cancer cases in 2013 was approximately 224,000, and that of cancer-induced death

---

*Address for correspondence: Peom Park, Department of Industrial Engineering, Ajou University, 206 Worldcup-ro, Yeongtong-gu, Suwon, Korea. Tel.: +82-31-219-2426; Fax: +82-31-219-2428; E-mail: ppark@ajou.ac.kr.

arrived at 75,334 [2], which was three times more than that caused by heart disease, the second leading cause of death. In addition, the social cost of cancer surpassed 14.6 trillion Won in 2013. With the increasing number of cancer patients, various studies have been conducted on ways to improve the survival rate through prevention and personalized diagnosis and therapy.

The new cancer therapy paradigm is targeted therapy, which analyzes characteristics of the cancer in each individual, and provides personalized therapy. The basic conditions for a successful targeted therapy against major cancers are accurate pathological diagnosis, immunopathological examination of tissue samples, and molecular pathology testing. Generally, biomarkers are used for pathological diagnosis; these are usually proteins, DNA, RNA, and metabolites that indicate changes in the body [3]. Using such biomarker information makes it possible to conduct personalized diagnosis and therapy, which takes into consideration characteristics of the cancer patient; this also determines the prognosis for a variety of cancer cases, thus being effective in increasing the survival rate. The market for personalized medical diagnosis was 17.7 billion USD in 2014, compared with 10.8 billion in 2009, an increase of 70% [4]; a variety of related research is being performed in areas such as pharmacogenetics, biochips, and genetic screening.

An accurate diagnosis is possible by objectively measuring pathological conditions in a cancer patient, or degree of response to a drug through biomarker positivity; in other countries, biomarker positivity values based on literature data have been already provided to users through sites like Pathpedia and GoPath [5]. In addition, there are many studies analyzing the positivity of biomarkers, or the probability of survival according to biomarkers; these studies are based on fragmented clinical experience or related data, or inferences based on disease mechanisms for the accurate diagnosis of cancer [6-13].

In Korea, as in other countries, research on biomarker positivity, using fragmented clinical tests or data on individuals or groups, is consistently performed [14-17]; however, there is a lack of effort to consolidate such research results to derive information. Since all clinical decisions must be made scientifically, based on the most recent clinical research results, there is a need to establish a clinical data-driven system.

In Korea, hospital work has seen ongoing computerization since 1972. For example, at Seoul National University Hospital (SNUH), all records have been computerized since a full electronic medical record (EMR) was introduced in 2004. However, in a conventional EMR, data is recorded in an unstructured free text format, and therefore it requires a great amount of time and manpower to extract information, making it difficult for actual use in the medical field.

Therefore, in the present study, with reference to foreign literature, a biomarker positivity analysis system was developed, by extracting information such as sex, age, and details on organ involvement and disease, based on a massive amount of clinical data from Korean cancer patients. In addition, comparison and validation on positivities was conducted using a group of pathologists and prior literature data, to increase the reliability of research results.

2. Materials and methods

In the present study, with the goal of providing cancer biomarker positivities based on characteristics of cancer in Koreans, clinical data on cancer patients from SNUH, as well as validated domestic and international literature data, were used to extract related information, create a database, and develop a biomarker positivity analysis system.
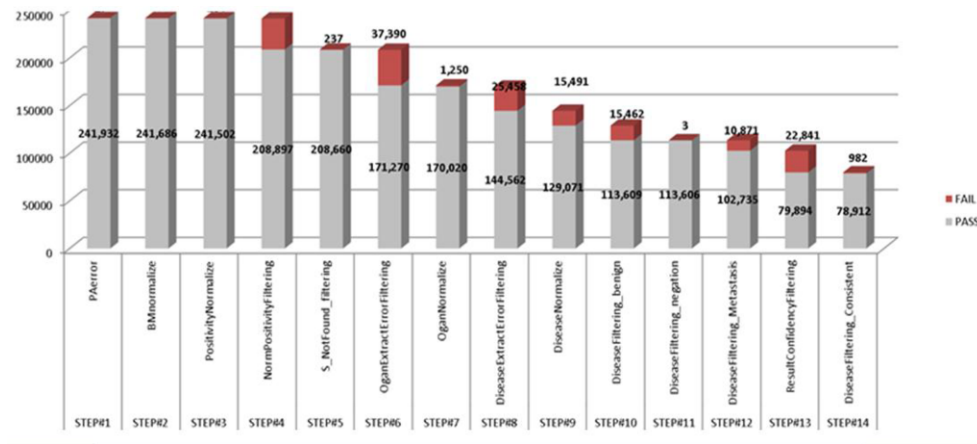
Fig. 1. Clinical data screening process.

2.1. Acquisition and preprocessing of data

2.1.1. Clinical data

As shown in Figure 1, immunopathology reports on 47,451 cases, and surgical reports on 46,084 cases, prepared from January 1st, 2007 to November 30th, 2012 in the Pathology Department of SNUH, were mapped; as a result, data on a total of 241,932 cases related to biomarker positivity tests were extracted. The data included 39 types of detailed organ information (an organ classification system that matches with Korean clinical data was defined, with reference to organ classification systems of WHO and the American Joint Committee on Cancer [AJCC]), 244 types of diagnostic details, and 166 types of biomarker information currently being used at SNUH.

The extracted data from the biomarker positivity test results on 241,932 cases were screened through 14 stages of error removal and normalization, with filtering for positivity, organ name, and diagnosis, as shown in Figure 1; ultimately, 78,912 cases were selected from the clinical data for further analysis.

2.1.2. Literature data

To compare and verify the reliability of clinical data, research articles related to biomarker positivity were collected, and the literature data were used to conduct a positivity analysis. Initially, literature data from studies involving Korean subjects were chosen preferentially; however, analogous case studies from other countries were inevitably included to secure biomarker positivity data obtained under similar conditions as the extracted clinical data. From 325 dissertation reviews involving 38 types of organs, positivity experimental results involving 153 types of biomarkers and 50,450 cases were extracted. However, the literature data differed from clinical data, as the former lacked specific information such as sex, age, organ type, and diagnosis. As a result, positivity result values were primarily extracted based on biomarker data, which was as much as possible mapped to construct a database (Table 1).

2.2. Database design and creation

The database of the present system is composed of a raw data pool, code data, data pool for analyses, and an algorithm for analysis, based on Microsoft-SQL. Extracted raw data are continuously updated

Table 1

Data feature

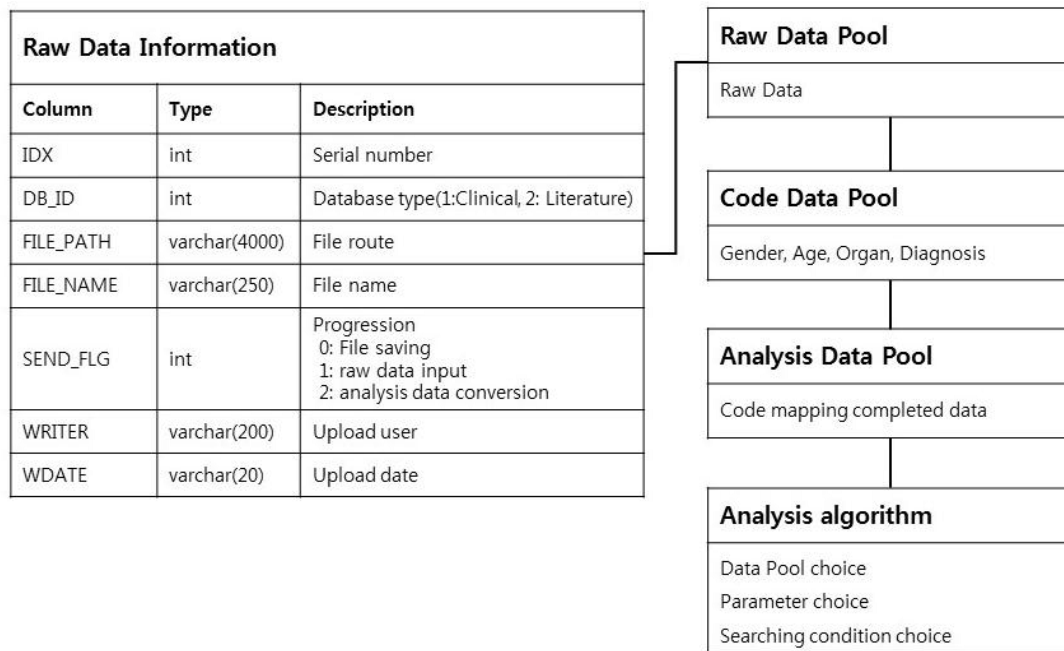| Variable	Clinical Data		Literature Data
	Count	Percentage (%)	Count
Gender			N/A
Male	34,393	43.58	
Female	44,519	56.42	
Age range			N/A
0~10	29	0.03	
11~20	954	1.21	
21~30	3,377	4.28	
31~40	8,487	10.76	
41~50	16,698	21.16	
51~60	20,212	25.61	
61~70	17,913	22.70	
71~80	10,036	12.72	
81~	1,188	1.50	
Organ Type	39		38
Diagnosis Type	244		244
Biomarker Type	166		153

Fig. 2. Data extracting process.

in the system to comprise a raw data pool, and the data that are cleared through coding are transmitted to the data pool for analysis. Going through such a process facilitates management of the data pool for analysis, even if changes are introduced to the format of the raw data, or new codes are added later on.

In addition, flags were attached, according to the type of each original data file, e.g. storage file (0), raw data input (1), or analysis data transmission (2), so that the administrator could manage the data

pool for each file (Figure 2).

Table 2
Analysis algorithm parameter

Parameter	Description
@DB_ID	1: Clinical, 2: Literature
@selParam	Cross analysis code group(ex: diseases, gender, organ, diagnosis)
@whereCondition	Restriction on search query
@data_display	Indication (1:Positivity%, 2: positive number/total number, 3: Positivity% & positivenumber/total number)

2.3. Algorithm for analysis

Positivity rate in Pathpedia.com, and other biomarker positivity rate-related studies, can be defined as the rate of positive cases per total number of cases; this value can change according to the selection of a specific variable. Therefore, this study first analyzed conventional methods of deriving biomarker positivity in literature, to understand which processes the researchers used in setting variables and the level of variables (Figures 3 and 4), and then designed the algorithm for analysis [18-21]. It was wondered, for example, what approaches would be necessary if one wanted to find a biomarker that had a high positivity for cancer from a specific disease, or for a specific age group. The parameters of the algorithm for analysis were defined through the analysis of the behavior of the researchers, as shown in Table 2, and the algorithm for analysis was coded as shown in Figure 5.

If @selParam is not designated with a value in the algorithm for analysis, biomarker positivity values of all data within given search conditions will be calculated; when setting search conditions as shown in Figure 6, the system is designed so that the accurate code could be designated using synonyms (thesaurus) of diseases, organs, and biomarkers.
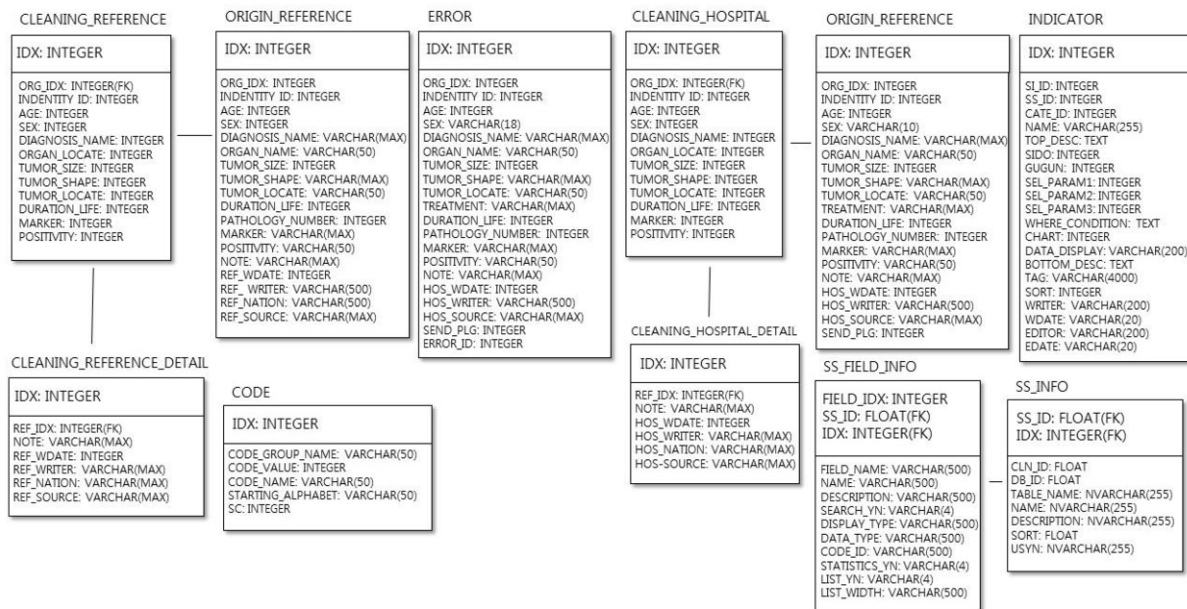

Fig. 3. Database structure.

IDX	DB_ID	ORG_IDX	EXCEL_RN	MARKER	POSITIVITY	ORGAN_NAM	DIAGNOSIS_I	SEX	AGE	Size_Tumor	NOMED_Sect	NOMED_Tes	OMED_Diagn	SEND_FLG
1683177	1	50	1	Genes, erb	POSITIVE	Breast	Invasive duc	F	74	0.3 x 0.2 x 1	breast	ectomy	infiltrating d	1
1683178	1	50	2	Carcinoem	NEGATIVE	Salivary glan	Pleomorphic	F	63	0.7 x 0.5 x 0	parotid glan	ectomy	mixed tumo	1
1683179	1	50	3	Proto-Onc	POSITIVE	Soft Tissue	MALIGNAN	M	77		soft tissue	specimen	malignant ly	1
1683180	1	50	4	Genes, bcl-	POSITIVE	Soft Tissue	MALIGNAN	M	77		soft tissue	specimen	malignant ly	1
1683181	1	50	5	Keratin-5	POSITIVE	Breast	Invasive duc	F	74	0.3 x 0.2 x 1	breast	ectomy	infiltrating d	1
1683182	1	50	6	MUM-1	POSITIVE	Soft Tissue	MALIGNAN	M	77		soft tissue	specimen	malignant ly	1
1683183	1	50	7	Calponin	POSITIVE	Breast	Invasive duc	F	74	0.3 x 0.2 x 1	breast	ectomy	infiltrating d	1
1683184	1	50	8	CDX 2	POSITIVE	Lung	ADENOCAR	M	56			T28000	P1140	M81403	1

ID	Code_Group_Name	Code_Value	Code_Name	Starting_Alphabet
4321	Diagnosis	D00062	Adenoma, ret	A
4322	Diagnosis	D00063	Adenoma, ser	A
4323	Diagnosis	D00064	Adenoma, tub	A
4324	Diagnosis	D00065	Adenoma, tub	A
4325	Diagnosis	D00066	Adenoma, tub	A
4326	Diagnosis	D00067	Adenoma, vill	A
4327	Diagnosis	D00068	Adenomatoid	A
4328	Diagnosis	D00069	Adenomatoid	A
4329	Diagnosis	D00070	Adenomyoep	A
4330	Diagnosis	D00071	Adenomyoep	A

IDX	ORG_IDX	idx_Original	SEX	AGE_PERIOD1	MARKER	POSITIVITY	ORGAN	DIAGNOSIS
582246	50	1683177	2	8	P00055	3	F01	D00654
582247	50	1683178	2	7	P00054	1	A06	D01060
582248	50	1683179	1	8	P00015	3	D02	D00756
582249	50	1683180	1	8	P00014	3	D02	D00756
582250	50	1683181	2	8	P00060	3	F01	D00654
582251	50	1683182	1	8	P00128	3	D02	D00756
582252	50	1683183	2	8	P00022	3	F01	D00654
582253	50	1683184	1	6	P00053	3	B01	D00026
582255	50	1683186	1	6	P00177	1	B01	D00026
582256	50	1683187	1	6	P00059	1	B01	D00026
582257	50	1683188	1	6	P00061	3	B01	D00026

Fig. 4. Raw data pool (Top) and code data (Bottom left), analysis data (Bottom right).

```
Result Table = SP_INDICATOR_DATA(
                @DB_ID int,
                @selParamnvarchar(max),
                @whereConditionnvarchar(max),
                @data_display int )
```

Fig. 5. Example of analysis algorithm.

Fig. 6. Searching condition input screen.

Table 3

Current top 10 number of clinical data for organ (left), diagnosis (middle) and biomarker (right)

Rank	Organ	Count	Diagnosis	Count	Biomarker	Count
1	Breast	18920	Invasive ductal carcinoma, NOS	14007	Genes, erbB-1	5517
2	Stomach	14549	Adenocarcinoma, gastric foveolar type	11849	Genes, erbB-2	5147
3	Brain, spinal cord	7591	Adenocarcinoma	7174	Tumor Suppressor Protein p53	4925
4	Lung	6301	Malignant lymphoma	6705	Genes, bcl-2	3447
5	Soft Tissue	5958	Poorly differentiated carcinoma	2354	Estrogen Receptor alpha	2872
6	Liver, Intrahepatic bile duct	4844	Ductal carcinoma insitu	2261	Receptors, Progesterone	2854
7	Bone marrow and lymph node	4256	Clear cell renal cell carcinoma	2085	Keratin-5	2716
8	Kidney	2972	Spindle cell tumor with thymus-like differentiation	1682	Keratin-7	2244
9	Thyroid and Parathyroid	2009	Glioblastoma	1590	S100 Proteins	2149
10	Colon and Rectum	1760	Small cell carcinoma	1361	Thymidylate Synthase	1907

	Marker	Positivity		Marker	D00408	D00489	D00529
1	MGMT	33.23% (107/322)	1	MGMT	100% (2/2)	25% (1/4)	33.8% (48/142)
2	Olig2	92.73% (51/55)	2	Olig2	0% (0/1)	100% (2/2)	96% (24/25)
3	Synaptophysin	65.14% (880/1351)	3	Synaptophysin	20% (1/5)	91.67% (11/12)	23% (23/100)

Fig. 7. Example of outcome for single analysis (left) and cross analysis (right).

3. Results

As shown in Table 3, the characteristics of cancer in Korean were investigated using the numbers of extracted data items related to organs, diagnoses, and biomarkers from the clinical database as the main criteria. In general, breast, stomach, brain, and lung cancers were found to have the highest incidence among Koreans. A closer look at detailed diagnoses revealed that breast, stomach, brain, and lung are the most prominent, followed by adenocarcinoma (gastric foveolar type), adenocarcinoma, and malignant lymphoma. Finally, an investigation on the number of biomarker data items related to clinical data items revealed that biomarkers used for organs with high cancer incidence rates are observed frequently during diagnosis and treatment.

Finally, using the previously constructed clinical/literature database and analysis algorithm, a biomarker positivity analysis system comprising a search criteria input section (Figure 6) and an analysis results section (Figures 7 and 8) was developed. When inputting search criteria, the user can choose which database (clinical or literature) is used for the positivity analysis. Subsequently, a single or cross analysis was also implemented in the search criteria (e.g., biomarker, organ, diagnosis). Single analysis is a method of calculating positivity according to biomarkers within the user's designated limits, whereas cross analysis is a method of calculating positivity according to biomarkers in the designated category and within the limits set by the user.

Although various combinations of search criteria could be generated using the clinical data, which

included all search criteria information, it was impossible to apply some search criteria to the literature data because the information only existed partially or not at all. Accordingly, the system was designed so that single and multiple cross analyses with respect to sex, age, organ, diagnosis, and biomarker were only possible when implementing a positivity analysis based on clinical data. Criteria with relatively few categories, such as age and sex, were configured as buttons so that the desired criteria could be selected directly; for those with many categories, such as organ, diagnosis, and biomarker, a direct input or selection could be performed using a search autocomplete function. Additionally, the search criteria input query was continuously updated in the user library, a design that allows the user to make comparisons between the analysis results obtained during previous criteria searches.

Analysis results are provided in table and graph forms, as shown in Figure 8. In the table, the
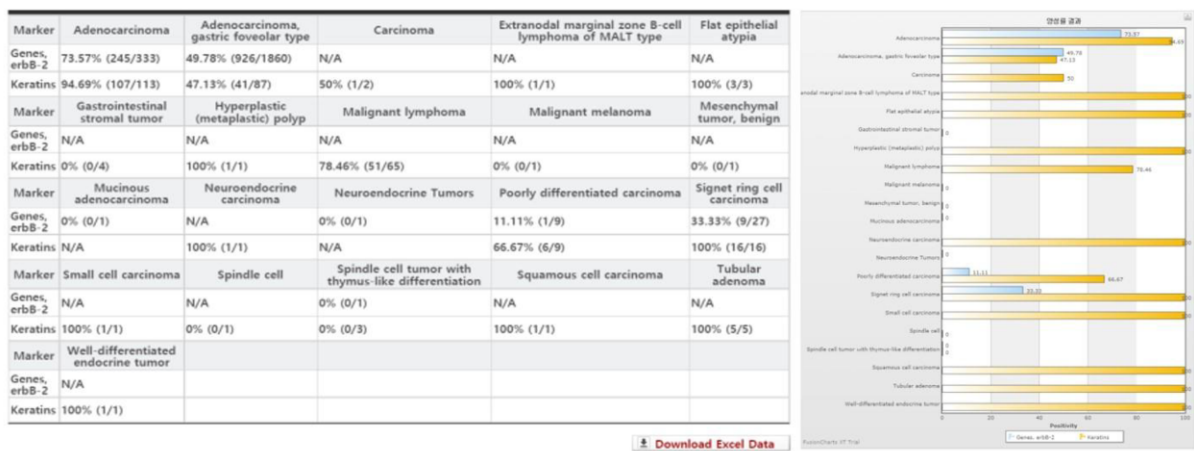

Fig. 8. Analysis result screen.

Biomarker	Clone	Catalog	Subjects	Positive site	Description
Parathyroid Hormone	105G7	NCL-PTH-488	Mouse Monoclonal	Human	The parathyroid glands are small, oval, endocrine glands closely associated with the thyroid gland. The glands regulate serum calcium and phosphate levels via parathyroid hormone (parathormone). Parathyroid hormone raises serum calcium levels by a direct mechanism on the bone increasing the rate of osteoclastic resorption and promoting breakdown of the bone matrix, by increasing the renal tubular reabsorption of calcium ions and inhibiting the reabsorption of phosphate ions from the glomerular filtrate and finally by promoting the absorption of calcium from the small intestine, involving vitamin D. Parathyroid hormone is the most important regulator of blood calcium levels and is essential to life whereas calcitonin appears only to provide a complementary mechanism for fine adjustment. Chief cells are the most abundant cells in the parathyroid gland and are responsible for the secretion of parathyroid hormone.

Fig. 9. Example of biomarker information.

positivity values of factors such as biomarker, organ, and diagnosis are shown quantitatively. For example, in the analysis results shown in the table (left) in Figure 8, which features keratins as the biomarker and adenocarcinoma as the diagnosis, 107 of 113 identified positivity test cases were positive, yielding a positivity rate of 94.69% based on the current data. In other words, the positivity value represents the number of positive cases divided by the total number of cases retrieved for the search criteria. For search criteria involving cases without positive responses, the results are displayed as (0/total number of cases), whereas search criteria without any relevant cases are displayed as N/A. In addition, depending on the selected condition, biomarker positivity is displayed in bar graph form, allowing for easy and intuitive comparisons of biomarker positivity rates according to particular conditions. Experts from the pathology department of Seoul National University Hospital served as the subjects for pilot tests. Bar graphs were used more often for general situations, whereas tables were preferred in situations requiring a more detailed understanding. Additionally, clicking on the biomarker within the table, as shown in Figure 9, reveals biomarker information such as the clone, subjects, and positive site. Furthermore, on the Seoul National University Hospital internal network, real-time EMR-associated pathology reports can be accessed by clicking the positivity result. The user may also use this platform to easily share information and obtain opinions regarding their chosen pathology cases.

4. Discussion

In this study, a biomarker positivity analysis system is developed by utilizing clinical data from the Korean population for cancer diagnosis and treatment. Previously, to conduct studies related to biomarker positivity, researchers would first establish a hypothesis and verify it through clinical experiments, or conduct a long process of directly searching for, extracting, and collecting data through the hospital EMR and parsing the data before analysis [22-25]. Researchers often experience difficulties in securing such large volumes of data. In the present study, however, natural language processing techniques are used to automatically extract a large amount of clinical data from Seoul National University Hospital's pathology reports. Experts only need to conduct a verification process through random sampling, and users are thus able to deduce the positivity values under their desired conditions more quickly and easily.

To evaluate the level of accuracy and satisfaction associated with this clinical data-based biomarker positivity analysis system, 25 pathology experts from Seoul National University Hospital serve as subjects for an assessment that yields a satisfaction level of 4.7 out of 5. In particular, when the opinions of the experts are gathered through a survey, the rating of 4.5 out of 5 given to positivity values based on search criteria differs little from that for actual diagnoses. However, as the survey is considered to reflect the experts' subjective opinions, the system constructed in our study is utilized to provide further quantitative verification by comparing 78,912 clinical data cases and 50,450 literature data cases using the same inputted search criteria for both databases. The comparison results demonstrate that 10,854 cases, or roughly 79.47%, of the 13,658 cases existing in both the clinical and literature data are of concordance. This concordance rate is not as high as initially expected because the 13,658 cases existing in both databases only comprise 17.31% and 27.07% of the clinical and literature data, respectively. Additionally, the clinical data only includes Korean data samples from Seoul National University Hospital pathology reports, whereas the literature data also include the data extracted from other countries besides Korea. To resolve this issue, biomarker positivity-related clinical and literature data from Korean subjects must be continually acquired; it is expected that in the

future, Korean researchers can input their own experimental results and research directly, and there are plans to develop and provide an additional program that will enable real-time linkage with the system proposed in our study.

As this is only one method of acquiring a biomarker-related data pool, additional studies are needed to reduce the amount of data omitted during the data refinement process. A total of 241,932 cases of positivity data are collected from the Seoul National University Hospital pathology reports, but 163,020 (roughly 67%) are omitted during the data refinement process. Refinement is done because clinical documents use a mixture of Korean and English, which are not grammatically precise, express terms using abbreviations, and contain various synonyms. For this study, a thesaurus is initially constructed to deal with synonyms, but for the results that had been changed to typical expressions, only cases above a certain level of reliability are selected, leading to the omission of a large amount of data. Accordingly, there are plans to conduct research regarding the improvement of natural language processing algorithms for clinical texts in order to increase the thesaurus coverage and decrease the proportion of omitted data.

Finally, in this study, new analysis methods and platforms are proposed for predicting cancer diagnoses and prognoses based on a large volume of clinical and literature data. It is expected that these tools will ultimately provide greater benefits to cancer-associated healthcare professionals and non-professionals.

Acknowledgments

References

[1] World Health Organization, Globocan 2012 IARC, 2012, Available at: http://www.who.int/cancer/en/, last accessed: April 21, 2015.

[2] National Cancer Information Center, Cancer statistics; 2013, Available at: http://www.cancer.go.kr/mbs/cancer/subview.jsp?id=cancer_040101000000, last accessed: April 21, 2015.

[3] Wikipedia, Biomarker, 2014, Available at: http://en.wikipedia.org/wiki/Biomarker, last accessed: April 21, 2015.

[4] PR Newswire, Personalized medicine market, 2015, Available at: http://www.prnewswire.com/news-releases/232-billion-personalized-medicine-market-to-grow-11-percent-annually-says-pricewaterhousecoopers-78751072.html, last accessed: April 21, 2015.

[5] Pathpedia, Immunopedia, 2015, Available at: http://www.pathpedia.com, last accessed: April 21, 2015.

[6] Jian Zhou, Jan Halámek, Vera Bocharova, Joseph Wang and Evgeny Katz, Bio-logic analysis of injury biomarker patterns in human serum samples, Talanta **83** (2011), 955-959.

[7] M. MalvezziLevi, C. La Vecchia and E. NegriBertuccio, European cancer mortality predictions for the year 2013, Annals of Oncology **24** (2013), 792-800.

[8] Mads Grønborg Zakarias Kristiansen, Akiko Iwahori and Rubens Chang Troels, Biomarker Discovery from Pancreatic Cancer Secretome Using a Differential Proteomic Approach, Molecular & Cellular Proteomics **5** (2006), 157-171.

[9] Nobuyoshi KosakaIguchi and Takahiro OchiyaHaruhisa, Circulating microRNA in body fluid: A new potential biomarker for cancer diagnosis and prognosis, Cancer Science **101** (2010), 2087-2092.

[10] G. Richard, MooreMac Laughlan, C. Robert and Bast Jr. Shannon, Current state of biomarker development for clinical application in epithelial ovarian cancer, Gynecologic Oncology **116** (2010), 240-245.

[11] E. Sergey, IlyinM. Belkowski and Carlos R. Plata-Salama´n0Stanley, Biomarker discovery and validation: Technologies and integrative approaches, Trends in Biotechnology **22** (2004), 411-416.

[12] YanXiong Peng, Wenyuan Li, and Ying Liu, A hybrid approach for biomarker discovery from microarray Gene Expression Data for Cancer Classification, Cancer Informatics **2** (2006), 301-311.

[13] Yu-Jen JouLin, Chih-Ho Lai, Chih-Hsin Tang, Su-Hua Huang and Ming-Hsui Tsai Chia-Der, Salivary zinc finger protein 510 peptide as a novel biomarker for detection of oral squamous cell carcinoma in early stages, Clinica Chimica Acta **412** (2011), 1357-1365.

[14] Chul-Kee ParkHye Jung, Min Jeong Moon, Young-Yim Kim, Jin Hyun Kim, Sung-Hye Park, Chae-Yong Kim, Sun Ha Paek, Dong Gyu Kim, Hee-Won Jung and Byung-Kyu ChoJi, Tissue expression of manganese superoxide dismutase is a candidate prognostic marker for glioblastoma, Oncology **77** (2009), 178-181.

[15] Hye Sook MinYeoun Lee, Seung-Ki Kim and Sung Hye ParkJi, Genetic grouping of medulloblastomas by representative markers in pathologic diagnosis, Translational Oncology **6** (2013), 265-272.

[16] Ji Hoon PhiPark, Chun Kee Chung, Kyu-Chang Wang, Byung-Kyu Cho and Seung-Ki KimSung-Hye, Atypical cell clusters expressing both neuronal and oligodendrocytic markers: Novel histological pattern of glioneuronal tumors?, Pathology International **59** (2009), 735-743.

[17] Roh-Eul Yoo Seung Hong Choi, Hye Rim Cho, Tae Min Kim, Se-Hoon Lee, Chul-Kee Park and Sung-Hye Park, Tumor blood flow from arterial spin labeling perfusion MRI: A key parameter in distinguishing high-grade gliomas from primary cerebral lymphomas, and in predicting genetic biomarkers in high-grade gliomas, Journal of Magnetic Resonance Imaging **38** (2013), 852-860.

[18] L.M. Sullivan, M.E. Smolkin, H.F. Frierson Jr. and M.T. Galgano, Comprehensive evaluation of CDX2 in invasive cervical adenocarcinomas: Immunopositivity in the absence of overt colorectal morphology, American Journal of Surgical Pathology **32** (2008), 1715–1720.

[19] T. Tot, Cytokeratins 20 and 7 as biomarkers: Usefulness in discriminating primary from metastatic adenocarcinoma, European Journal of Cancer **38** (2002), 758–763.

[20] J.J. Brooks, Immunohisto chemistry of soft tissue tumors: Myoglobin as a tumor marker for rhabdomyosarcoma, Cancer **50** (1982), 1757–1763.

[21] L.B. De Lott, C. Morrison, S. Suster, D.E. Cohn and W.L. Frankel, CDX2 is a useful marker of intestinal-type differentiation: A tissue microarray-based study of 629 tumors from various sites, Archives of Pathology & Laboratory Medicine **129** (2005), 1100–1105.

[22] S.H. Kim, J.Y. Choe, Y. Jeon, J. Huh, H.R. Jung, Y.D. Choi, H.J. Kim, H.J. Cha and W.S. Park and J.E. Kim, Frequent expression of follicular dendritic cell markers in Hodgkin lymphoma and anaplastic large cell lymphoma, Journal of Clinical Pathology **66** (2013), 89–96.

[23] K.M. Kurian, H.R. Haynes, C. Crosby, K. Hopkins and M. Williams, IDH mutation analysis in gliomas as a diagnostic and prognostic biomarker, British Journal of Neurosurgery **27** (2013), 442–445.

[24] B.H. Lee, J.L. Hecht, J.L. Pinkus and G.S. Pinkus, WT1: Estrogen receptor, and progesterone receptor as markers for breast or ovarian primary sites in metastatic adenocarcinoma to body fluids, American Journal of Clinical Pathology **117** (2002), 745–750.

[25] J.L. Hecht, J.L. Pinkus, L.J. Weinstein and G.S. Pinkus, The value of thyroid transcription factor-1 in cytologic preparations as a marker for metastatic adenocarcinoma of lung origin, American Journal of Clinical Pathology **116** (2001), 483–488.