The Development of Work-related Musculoskeletal Disorders Prediction System

U-ri Chae (chaeuri@ajou.ac.kr)¹, Seung Gi Kim (xeran75@gmail.com)², Young joon Kim (yjk4004@ajou.ac.kr)¹, Sung-Hwan Yang (shyang@knuw.ac.kr)³, Peom Park (ppark@ajou.ac.kr)¹

¹Department of Industrial Engineering, Ajou University, Suwon, 443-749
²EnValueBiz, Yeongsin-ro, Yeongdeungpo-gu, Seoul, Korea 07228
³Department of Prosthetics & Orthotics, Korea National University of Welfare, Pyeongtaek, 17738

ABSTRACT

According to the International Labor Organization, 317 million accidents occur on the work-related diseases and injuries each year(1). In addition, occupational diseases that showed a declining trend until 2011 are expected to rise again in 2013. Occupational disease, in the self-assessment, among the approximately 1.2 million workers, people who had experienced occupational diseases were judged to be ill due to their job or cause of the disease.

Overseas, there is interest and high level research on the factors that affect occupational diseases. In Europe, there is a survey of occupational diseases in the European Union countries through the European Working Conditions Survey (EWCS), which measures the working environment and correlates various aspects, and research that contributes to European policy development, in particular the quality of work and employment issues. In Korea, the EWCS is supplemented with the contents of the question according to the characteristics of Koreans, and KWCS (Korean Working Conditions Survey) is being conducted for Koreans. In this system, KWCS data is used to analyze occupation of upper limb, lower back and lower limb. We have developed a probability prediction system that can cause musculoskeletal disorders.

In this system, a decision tree, a logistic regression, and a neural network are selected as the data mining method for predicting the occurrence of detailed occupational diseases, and the decision tree algorithm which is the most reliable in the ROC curve is selected and applied.

Keywords: ESK-JES Joint Session, Decision Tree Analysis, Predictive System, Work-related Musculoskeletal disorders

1. Introduction

According to the International Labor Organization(ILO), 317 million of occupational accidents are happening around the world every year and are continuing to increase (1). In Korea, work-related musculoskeletal diseases tend to increase every year(2). In overseas countries, national survey is conducted to determine the factors attributable to these work-related diseases, such as European Working Conditions Survey(EWCS). The ‘Korean Working Conditions Survey’ data used in this paper, is a survey data benchmarked to the EWCS based on the characteristics of Koreans(3). In Europe, EWCS data serves as a contribution to European policy development and employment issues through research and development, such as correlation analysis, and social interest in working environments between various aspects(4, 5).

In Korea, there are more studies on the predictive system of diseases occurrence probability than musculoskeletal disorders. There are more studies on predictive models than system development(7, 8). In the case of other countries, studies are being conducted to develop a system for predicting the probability of musculoskeletal disorders, but it is still difficult to predict the overall probability of work-related musculoskeletal disorders by analyzing only one occupation(9, 10). Therefore, in this paper, the system was developed to analyze the actual users(physicians, occupational safety experts etc.) by applying the prediction
algorithm of work-related musculoskeletal diseases, which was developed through preprocessing and analysis of KWCS data. Work-related musculoskeletal disorders prediction system is important for improving the working environment and prevent accidents, and to reduce costs incurred by industrial accidents.

2. Method

2.1 Algorithm

To develop a predictive algorithm, the 3rd KWCS data was used and 34,788 cases were selected by weighting a total of 50,032 data. The selected data were divided into upper limb, lower back and lower limb according to body parts. The analysis methodology was conducted by selecting Regression Analysis, Artificial Neural Network and Decision Tree which are used for predictive analysis. The algorithm used to develop the system was applied to the Decision Tree Analysis with the highest reliability in the ROC curve.

A total of 50,032 data were weighted and 34,788 cases were selected and analyzed by Decision Tree. The data were analyzed by three segments, upper extremity, lower back and lower extremity depending of the body region. Also algorithm used to develop the system was applied to the Decision Tree model with the highest reliability in the ROC curve.

The tool used for the analysis was SAS E-Miner ver.14.1, which can be analyzed each analysis method as a node, and the result can be directly seen in the SAS code.

2.2 System

IIS 6.0, Visual Studio 2015, and MSSQL 2014 Express were used as development tools. The development language used HTML5, Java script, JQuery, and ASP.

The system applied a login system to limit users and prevent inadvertent data entry. This is because the accuracy of the data and the reliability of the results can be increase. Additionally, users can store information about patient or workers individually and all data are stored back into the database. These individual data are piled up to generate log data for the individual. All the log data can also be collected contributing to system reliability and improvement through algorithm improvement. We expect that we will be able to realize personalized medicine in the future.

The questionnaires applied to the system were selected form the KWCS questionnaire through variable importance analysis.
Table 1. Upper Limb variable importance

| Variable                                    | Importance |
|---------------------------------------------|------------|
| Tiring and Painful Position                 | 1          |
| Occupation                                  | 0.9729     |
| Repetitive Hand-arm Motion                  | 0.8678     |
| Moving Heavy Objects                        | 0.8424     |
| Continuous Standing                         | 0.8406     |
| Use of Internet/E-mail                      | 0.5778     |
| Computer Work                               | 0.546      |
| Education Level                             | 0.5428     |
| Absent Days in the Past Year                | 0.5181     |
| Weekly Working Time                         | 0.4758     |
| Noise                                       | 0.4114     |
| Threatening or Humiliating Behavior         | 0.4023     |
| Lifting or Moving People                    | 0.4007     |
| Contact with Angry Clients                  | 0.3859     |

Table 2. Lower Back variable importance

| Variable                                    | Importance |
|---------------------------------------------|------------|
| Occupation                                  | 1          |
| Tiring and Painful Position                 | 0.961      |
| Continuous Standing                         | 0.855      |
| Moving Heavy Objects                        | 0.8479     |
| Repetitive Hand-arm Motion                  | 0.8479     |
| Work Environment Satisfaction               | 0.6814     |
| Feeling of Doing Good Job                   | 0.5899     |
| Use of Internet/E-mail                      | 0.5873     |
| Computer Work                               | 0.5602     |
| Education Level                             | 0.5567     |
| Working Hours and Compatibility with Personal Life | 0.5513 |
| Contact with Angry Clients                  | 0.5359     |
| Verbal Abuse                                | 0.5347     |
| Feeling of Doing Useful Work                | 0.526      |
| Tenure                                      | 0.5157     |
| Absent Days in the Past Year                | 0.4964     |

Table 3. Lower Limb variable importance

| Variable                                    | Importance |
|---------------------------------------------|------------|
| Occupation                                  | 1          |
| Tiring and Painful Position                 | 0.9489     |
| Moving Heavy Objects                        | 0.8971     |
| Repetitive Hand-arm Motion                  | 0.8151     |
| Continuous Standing                         | 0.8084     |
| Use of Internet/E-mail                      | 0.7134     |
| Computer Work                               | 0.7114     |
| Educational Level                           | 0.6575     |
| Weekly Working Time                         | 0.5485     |
| Threatening of Humiliating Behavior         | 0.4571     |
| Complex Work                                | 0.4507     |
| Work Demand                                 | 0.4372     |
| High Temperatures                           | 0.4291     |
| Physical Violence                           | 0.4285     |
| Feeling of Doing Useful Work                | 0.4247     |
| Recognition of what is Expected at Work     | 0.4207     |

3. Results

The schematic diagram of the system is shown below.

![Schematic Diagram](image)

All users are required to register. When they sign up, they can pre-set individual codes, such as their number, on the server so that they can select their patients.

When the user logs in, the main menu is displayed. The main menu contains ‘All Data’, ‘Past Results’, and ‘New Data’.

3.1 All Data tab

All data is divided into upper limb, lower back, and lower limb where users can see existing data stored in the entire DB.

3.2 Past Results tab

In this tab, users can see the results that they entered. In the case of hospital, the patients’ results can be selected.
alone. In case of the workplace, the user can check the data of the worker entered. Also, one log data can be shown separately by inputting basic information of a worker.

### 3.3 New Data tab

New data can be input. Enter basic information of the patient or worker, or search for and select existing data, and start the survey.

![New Data tab](image)

**Figure 8. Mobile web-based system**

**Figure 9. Web-based system**

### 4. Conclusion

The third data used in this research is difficult to confirm the results over time according to the characteristics of the survey data. In addition, since only the 3rd data is analyzed, it is necessary to apply and analyze the latest data additionally. Therefore, we will analyze the total 1st-3rd data, including 4th data which is the latest, to develop an algorithm that can predict work-related musculoskeletal disorders, thereby developing an algorithm for work-related diseases that can occur by work. In addition, studies on aging and musculoskeletal disorders are already underway in foreign countries. If the data of this system accumulates continuously, it is expected that it will be a stepping stone of research that can separate and predict disease and work-related diseases according to time flow through time series analysis. Furthermore, since it is difficult to see the result screen on the existing mobile web, we will develop a mobile application so that the users can understand the image in a bar graph format and display a brief explanation.

Finally, due to the characteristics of the system, in which personal information of patients and workers are entered, a thorough review of these issues will be needed.

### Acknowledgements

This work was funded by grants from Korean Federation of Science and Technology Societies (Grant-#KFSTS000-0000).

### References

International Labour Organization, *Safety and health at work*, http://www.ilo.org/global/topics/safety-and-health-at-work/lang-en/index.htm (retrieved April 12, 2017).

Korea Ministry of Employment and Labor, *Analysis of the status of industrial accidents*, 2016.

Youngsun Kim, Wanyoun Yang, Jaiwook Baik, *Comparative study of working conditions of Korea and Europe*, Journal of the Korean Data & Information Science Society, 27(1), 45-65, 2016.

Publication Office of the European Union, *5th European Working conditions Survey: Overview report*, 2012.

Diego Montano, *Upper body and lower limbs musculoskeletal symptoms and health inequalities in Europe: an analysis of cross-sectional data*, BMC Musculoskeletal Disorders, 15:285, 2014.

Korea Ministry of Employment and Labor, *Development of web-based integrated information system for distribution and dissemination of work-related cancer-related system*, 2013.

Korea Ministry of Employment and Labor, *Improvement Measures of Risk Factors for Musculoskeletal Disorders of Typical Work – Focused on Healthcare Industry and Accommodation Industry*, 2008.

Korea Ministry of Employment and Labor, *Development of Risk Assessment Guideline (Risk Level Model) to Investigate Harmful Factors of Musculoskeletal Burden Work*, 2008.

Bhornsawan Thanathornwong, Siriwan Suebnukan, Yupin Songpaisan, Kan Ouivirach, *A system for predicting and preventing work-related musculoskeletal disorders among dentists*, Computer Methods in Biomechanics and Biomedical Engineering, 17(2), 177-185, 2014.

Bhornsawan Thanathornwong, Siriwan Suebnukan, Kan Ouivirach, *A System for Predicting Musculoskeletal Disorders Among Dental Students*, International Journal of Occupational Safety and Ergonomics, 20(3), 463-475, 2014.

Health and Safety Executive, *Ageing and work-related musculoskeletal disorders*, 2010.