Submitted: 11 July, 2024 Accepted: 06 August, 2024 Published: 30 December, 2024 DOI:10.22514/jomh.2024.199

JOMH
Journal of Men's Health

ORIGINAL RESEARCH

Predicting occupational musculoskeletal disorders in South Korean male office workers using a robust and sparse twin support vector machine

Haewon Byeon[1,2,*]

[1]Department of Digital Anti-Aging Healthcare (BK21), Inje University, 50834 Gimhae, Republic of Korea
[2]Department of AI-Software, INJE Medical Big Data Research Center, Inje University, 50834 Gimhae, Republic of Korea

***Correspondence**
byeon@inje.ac.kr
(Haewon Byeon)

Abstract

This study aims to investigate the prevalence and risk factors of musculoskeletal disorders (MSDs) among South Korean male office workers and to introduce a robust predictive model using the Robust and Sparse Twin Support Vector Machine (RSTSVM). A cross-sectional survey was conducted among male office workers in South Korea to assess the prevalence of MSDs and identify associated risk factors. Data on ergonomic and psychosocial factors were collected and analyzed. The RSTSVM model was developed and compared with traditional machine learning models, including Support Vector Machine (SVM) and Gradient Boosting Machine (GBM), to predict the risk of MSDs. The analysis revealed a high prevalence of MSDs among the surveyed office workers, attributed to factors such as prolonged sitting, repetitive hand/arm movements, standing posture and carrying heavy objects. Prolonged static postures were significantly linked to lower back pain and other musculoskeletal issues. Poor workstation ergonomics and psychosocial stressors, such as high job demands and low job control, were also identified as significant predictors of MSDs. The RSTSVM model demonstrated superior performance in predicting MSDs, with an Area under the Receiver Operating Characteristic Curve (AUC-ROC) value of 0.84, effectively managing high-dimensional data and maintaining robustness against outliers and noise. Furthermore, the RSTSVM model provided enhanced interpretability, making it easier to identify and understand key risk factors compared to traditional models. The study underscores the critical need for multifaceted intervention strategies to address the ergonomic and psychosocial risk factors associated with MSDs among office workers. Future research should focus on longitudinal studies to establish causal relationships and evaluate the effectiveness of various interventions across different occupational groups.

Keywords

Musculoskeletal disorders; Psychosocial factors; Office workers; RSTSVM model

1. Introduction

Occupational musculoskeletal disorders (MSDs) are a significant global health concern, affecting workers' productivity and quality of life. In 2020, the prevalence of MSDs was reported to be 6.2% among women and 4.3% among men [1]. These disorders primarily affect the lower back (52.2%) [2]. The prevalence is notably high among agricultural workers, particularly women [3]. Recognizing the severity of these issues, the Ministry of Employment and Labor in South Korea has implemented policies to automatically recognize MSDs as occupational diseases without additional investigation [4]. The economic impact of these incidents is substantial, accounting for nearly 4% of the global Gross Domestic Product (GDP), which equated to approximately $1.25 trillion in 2011 [5]. This economic burden includes reduced worker productivity, medical expenses, and long-term social costs. Given the substantial economic and health impacts, effective prediction and management strategies for MSDs are critically needed to alleviate these burdens and improve workers' overall well-being [5].

The historical context of occupational diseases dates back to ancient times with early documentation by Hippocrates and systematic categorization by Bernardo Ramazzini in the 18th century [6]. Modern studies began in the early 20th century, focusing on the causes and prevention of these diseases. The International Labour Organization (ILO) emphasizes the importance of specific hazardous substances and their causal relationships with diseases, while the World Health Organization (WHO) estimates the average cost of basic health care for workers affected by occupational diseases ranges from $18 to $60 per person [7]. MSDs have been identified as a major contributor to the economic losses associated with occupational diseases, accounting for approximately 37% of these losses [8]. Defined by the WHO, MSDs are disorders of the muscles, tendons, peripheral nerves, and vascular system caused or exacerbated by repetitive or continuous use of the body [9]. The National Institute for Occupational Safety and Health (NIOSH) in the United States defines occupational MSDs based on symptoms like pain and stiffness that persist due to work-related activities [10].

In South Korea, the prevalence of MSDs has been increasing. According to the Korea Statistical Information Service (KOSIS), the prevalence rate of MSDs increased by 26.3% in 2018 compared to 2010 [11]. This trend highlights the growing importance of awareness and management of MSDs, including those affecting the lower back. Effective chronic disease management requires the control of health risk factors, management of pre-existing conditions and early detection and treatment of major chronic diseases [12]. Recognizing the impact of MSDs, various countries and research institutions have conducted extensive studies to identify and mitigate the associated risk factors. The European Working Conditions Survey and research by The National Institute for Occupational Safety and Health (NIOSH) and Occupational Safety and Health Administration (OSHA) in the United States have categorized occupational risk factors and analyzed their impacts comprehensively [13]. Traditional regression and correlation analyses utilized in previous studies frequently fail to capture the complex, non-linear relationships between variables and are prone to overfitting. In contrast, Support Vector Machines (SVM) provide robustness and the capability to model non-linear relationships effectively. The Robust and Sparse Twin Support Vector Machine (RSTSVM), in particular, is highly advantageous as it efficiently handles high-dimensional data while maintaining robustness against outliers and noise. By incorporating sparsity, the RSTSVM model not only enhances predictive accuracy but also improves interpretability by identifying the most significant variables contributing to occupational MSDs [14]. This study proposes a fusion model of RSTSVM and logistic regression analysis for the prediction of MSDs among male office workers. Moreover, this study aims to contribute to the existing body of knowledge by offering a robust predictive tool for occupational MSDs and identifying key risk factors that can inform prevention and intervention strategies.

2. Related work

In the field of occupational health, numerous studies have focused on the identification, assessment and prediction of MSDs. These studies employ various methodologies ranging from traditional statistical techniques to advanced machine learning models. Traditional approaches often utilize regression and correlation analyses to identify factors influencing MSDs. For instance, Mekonnen (2017) [13] applied multiple regression analysis to investigate the relationship between physical workload and the incidence of lower back pain in nursing personnel. These traditional statistical methods have been instrumental in establishing the foundational understanding of MSD risk factors. While these studies provide valuable insights, they are limited by their inability to account for complex, non-linear interactions among variables and are susceptible to overfitting, particularly in high-dimensional datasets.

The advent of machine learning has introduced more sophisticated methods for analyzing and predicting MSDs. Support Vector Machines (SVM) have been widely used due to their robustness and capability to model non-linear relationships. For example, Eubank *et al.* [14] (2021) used a decision tree algorithm to predict the occurrence of shoulder pain in assembly line workers, taking into account a wide range of variables including individual characteristics, job-related factors and psychosocial factors. Despite the advancements brought by machine learning models, there is a need for methods that can handle high-dimensional data efficiently while maintaining robustness against outliers and noise. Robust and Sparse Twin Support Vector Regression (RSTSVM) has emerged as a promising approach in this regard. RSTSVM enhances predictive accuracy by incorporating sparsity, which helps in identifying the most significant variables contributing to occupational MSDs [15]. This method has been applied in various domains, including bioinformatics [16], but its application in occupational health remains relatively unexplored.

The review of related works underscores the evolution of methodologies from traditional statistical techniques to advanced machine learning models in the study of MSDs. While traditional methods have laid the groundwork for understanding the epidemiology of MSDs, machine learning models, particularly RSTSVM, offer enhanced predictive capabilities and the ability to handle complex interactions among variables. This study aims to build on this body of work by proposing a fusion model of RSTSVM and logistic regression analysis, contributing a robust predictive tool for occupational MSDs and identifying key risk factors that can inform prevention and intervention strategies.

3. Method

3.1 Data collection and preprocessing

The data for this study was sourced from the Korean Working Conditions Survey (KWCS) conducted by the Korea Occupational Safety and Health Agency (KOSHA) in 2020. The KWCS is modeled after the European Working Conditions Survey (EWCS) and has been adapted to reflect the unique work environment and cultural differences in South Korea. The survey collected responses from 50,032 individuals through face-to-face interviews at their residences, targeting one employed person per household.

From the total respondents, 6885 male office workers were selected for the final analysis. This selection was made to focus on a specific subgroup of workers who may experience unique risk factors related to MSDs. The demographic breakdown of the male office workers includes an age distribution as follows: 15–19 years (1.2%), 20–29 years (20.0%), 30–39 years (28.4%), 40–49 years (26.6%), 50–59 years (17.0%) and 60 years and above (6.8%). Data cleaning involved handling missing values, outliers and ensuring consistency across the dataset. Missing values were imputed using mean imputation for continuous variables and mode imputation for categorical variables. Outliers were identified and managed using the interquartile range (IQR) method. Feature selection was performed using a combination of domain knowledge and statistical methods. Features with high correlation to the target variable (occupational MSDs) were retained for further analysis.

3.2 Measurement

In this study, 67 factors expected to influence occupational MSDs were extracted from the overall survey items of the Working Conditions Survey through literature review. Each factor was classified into personal factors, work environment factors, social-psychological factors, and job-related factors (Table 1). Personal factors included gender, age, education level, income level, smoking frequency, drinking frequency, obesity, hypertension, number of absences in the past year and union membership. Work environment factors were divided into physical work risk factors (*e.g.*, vibration, noise, temperature, dust, steam, chemicals, tobacco smoke, infectious substances) and musculoskeletal burden factors (*e.g.*, repetitive hand/arm movements, standing posture, carrying heavy objects, fatigue or painful postures). Social-psychological factors included experiences of discrimination (*e.g.*, age, education, region of origin, gender, employment type), sexual harassment, verbal abuse, unwanted sexual attention, threats or humiliating behavior, physical violence, bullying, dealing with customers, working with angry customers, working with computers, using the internet/email, support from colleagues and supervisors, feeling of doing a good job, feeling of doing meaningful work, awareness of job expectations, emotional involvement in work, job stress and the need to hide emotions at work. Job-related factors included weekly working hours, use of personal protective equipment, provision of health and safety information, work patterns (*e.g.*, same daily working hours, shift work), frequency and timing of work schedule changes, alignment of working hours with personal life, involvement in setting work goals, participation in improving work processes, job characteristics (*e.g.*, monotony, complexity), threat level to health or safety, number of workers, years of service, job satisfaction, industry, job demands and job autonomy.

The reliability of the factors was verified using reliability coefficients, and the results showed that work environment factors (Cronbach's alpha = 0.847) and social-psychological factors (Cronbach's alpha = 0.632) met the reliability criteria. The job-related factors (Cronbach's alpha = 0.037) did not meet the reliability criteria but were included in the analysis as they were judged to be related to the job.

3.3 Model development

The Robust and Sparse Twin Support Vector Machine (RSTSVM) model was developed to handle high-dimensional data efficiently while maintaining robustness against outliers and noise. The key steps involved in developing the RSTSVM model are:

1. Twin Hyperplanes Construction: Unlike traditional SVM, which constructs a single hyperplane, RSTSVM constructs two non-parallel hyperplanes. Each hyperplane is closer to one of the classes and is designed to minimize the classification error for that class.

2. Robustness Integration: To mitigate the impact of outliers, robust loss functions are used. These functions reduce the influence of data points that lie far from the hyperplane, thereby enhancing the model's robustness.

3. Sparsity Induction: L1 regularization is applied to induce sparsity. This regularization technique penalizes the absolute values of the coefficients, leading to a model that relies on a smaller number of features, making it more interpretable and efficient.

Subject to the constraints:

$$[y_i\left(w_1^T x_i + b_1\right) \geq 1 - \xi_i]$$
$$[y_i\left(w_2^T x_i + b_2\right) \leq -1 + \xi_i^*]\ [\xi_i,\ \xi_i^* \geq 0]$$

Where (w_1) and (w_2) are the weight vectors, (b_1) and (b_2) are the bias terms, (ξ_i) and (ξ_i^*) are the slack variables.

Logistic regression was used in conjunction with RSTSVM to enhance the interpretability of the model. Logistic regression is a widely used statistical method for binary classification problems and provides a clear understanding of the relationship between the predictors and the outcome. The logistic regression model was formulated to predict the probability of the occurrence of MSDs based on the features selected.

The logistic regression model was formulated as follows:

$$\left[log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n\right]$$

Where (p) is the probability of the occurrence of the event (MSD), (β_0) is the intercept, and ($\beta_1, \beta_2, \ldots, \beta_n$) are the coefficients of the predictors ($X_1, X_2, \ldots, X_n$).

3.4 Comparative models

To evaluate the effectiveness of the RSTSVM model, its performance was compared with two widely used machine learning models: Support Vector Machine (SVM) and Gradient Boosting Machine (GBM).

1. Support Vector Machine (SVM): SVM is a powerful classification algorithm that finds the optimal hyperplane separating the classes by maximizing the margin between them. It uses kernel functions to handle non-linear relationships.

2. Gradient Boosting Machine (GBM): GBM is an ensemble learning technique that builds multiple weak learners (typically decision trees) in a sequential manner. Each subsequent model corrects the errors of the previous models, resulting in a strong predictive model. GBM is known for its high accuracy and ability to handle complex datasets.

3.5 Model evaluation

The models were evaluated using k-fold cross-validation (k = 10) to ensure robustness and to prevent overfitting. The dataset was divided into 10 subsets, and each model was trained on 9 subsets and tested on the remaining subset. This process was repeated 10 times, and the results were averaged to obtain the final performance metrics.

TABLE 1. Key variables of study.

Category	Factors
Personal Factors	Age, Education level, Income level, Smoking frequency, Drinking frequency, Obesity, Hypertension, Number of absences in the past year, Union membership
Work Environment Factors	Physical Work Risk Factors: Vibration, Noise, Temperature, Dust, Steam, Chemicals, Tobacco smoke, Infectious substances Musculoskeletal Burden Factors: Repetitive hand/arm movements, Standing posture, Carrying heavy objects, Fatigue or painful postures
Social-Psychological Factors	Experiences of discrimination (age, education, region of origin, employment type), Sexual harassment, Verbal abuse, Unwanted sexual attention, Threats or humiliating behavior, Physical violence, Bullying, Dealing with customers, Working with angry customers, Working with computers, Using the internet/email, Support from colleagues and supervisors, Feeling of doing a good job, Feeling of doing meaningful work, Awareness of job expectations, Emotional involvement in work, Job stress, Need to hide emotions at work
Job-Related Factors	Weekly working hours, Use of personal protective equipment, Provision of health and safety information, Work patterns (same daily working hours, shift work), Frequency and timing of work schedule changes, Alignment of working hours with personal life, Involvement in setting work goals, Participation in improving work processes, Job characteristics (monotony, complexity), Threat level to health or safety, Number of workers, Years of service, Job satisfaction, Industry, Job demands, Job autonomy

The performance of the models was assessed using several metrics: accuracy, precision, recall, F1-score and the area under the receiver operating characteristic curve (AUC-ROC). Accuracy measured the proportion of correctly classified instances. Precision calculated the proportion of true positive instances among the instances classified as positive. Recall determined the proportion of true positive instances among all actual positive instances. The F1-score provided the harmonic mean of precision and recall. The AUC-ROC assessed the model's ability to distinguish between classes.

3.6 Variable importance analysis

The importance of each variable was analyzed to understand its contribution to the prediction of occupational MSDs. In logistic regression, the magnitude of the coefficients was used to determine the importance of each predictor. In RSTSVM, the weights assigned to each feature were analyzed to understand their contribution to the model. Additionally, permutation importance was used, where the change in the model's performance was observed by randomly shuffling each feature.

3.7 Odds ratio and confidence intervals

For the six most significant predictors, odds ratios and 95% confidence intervals (CI) were calculated to quantify the strength and direction of the association between the predictors and the occurrence of MSDs. The odds ratio was calculated as the exponentiation of the coefficient of the predictor in the logistic regression model. Confidence intervals were calculated using the standard error of the coefficients, providing a range within which the true odds ratio is expected to fall with 95% confidence.

By integrating these methodologies, the study aims to develop a robust, sparse, and interpretable model for predicting occupational MSDs, providing valuable insights for prevention and intervention strategies.

4. Results

4.1 General characteristics of the subjects

The study analyzed data from 6885 male office workers, extracted from the Korean Working Conditions Survey (KWCS) conducted in 2020. The age distribution of the subjects is as follows: 15–19 years: 1.2% (83), 20–29 years: 20.0% (1377), 30–39 years: 28.4% (1955), 40–49 years: 26.6% (1832), 50–59 years: 17.0% (1171) and 60 years and above: 6.8% (467). The largest age group is 30–39 years, followed closely by the 40–49 years' group. These two groups together constitute more than half of the sample, indicating that the majority of the male office workers are in their prime working years. It was 17% (n = 1170) as a result of calculating the prevalence of MSDs in male office workers.

4.2 Model performance

The performance of the Robust and Sparse Twin Support Vector Machine (RSTSVM) model was compared with two widely used machine learning models: Support Vector Machine (SVM) and Gradient Boosting Machine (GBM). The models were evaluated using 10-fold cross-validation to ensure robustness and to prevent overfitting. The average performance metrics are presented in the following Table 2. The RSTSVM model outperformed both the SVM and GBM models in terms of accuracy, precision, recall, F1-score (Fig. 1), and AUC-ROC. The high AUC-ROC value of 0.81 (Fig. 2) indicates that the RSTSVM model has a strong ability to distinguish between workers with and without MSDs.

4.3 Variable importance

The importance of each variable was analyzed to understand its contribution to the prediction of occupational MSDs. The Fig. 3 summarizes the importance of the top factors based on the analysis. The top factor, "fatigue or painful postures",

TABLE 2. The average performance metrics.

Model	Accuracy	Precision	Recall	F1-Score	AUC-ROC
RSTSVM	85.2%	84.5%	83.8%	84.1%	0.84
SVM	83.7%	82.9%	82.1%	82.5%	0.80
GBM	84.9%	84.3%	83.5%	83.9%	0.82

AUC-ROC: area under the receiver operating characteristic curve; RSTSVM: Robust and Sparse Twin Support Vector Machine; SVM: Support Vector Machine; GBM: Gradient Boosting Machine.

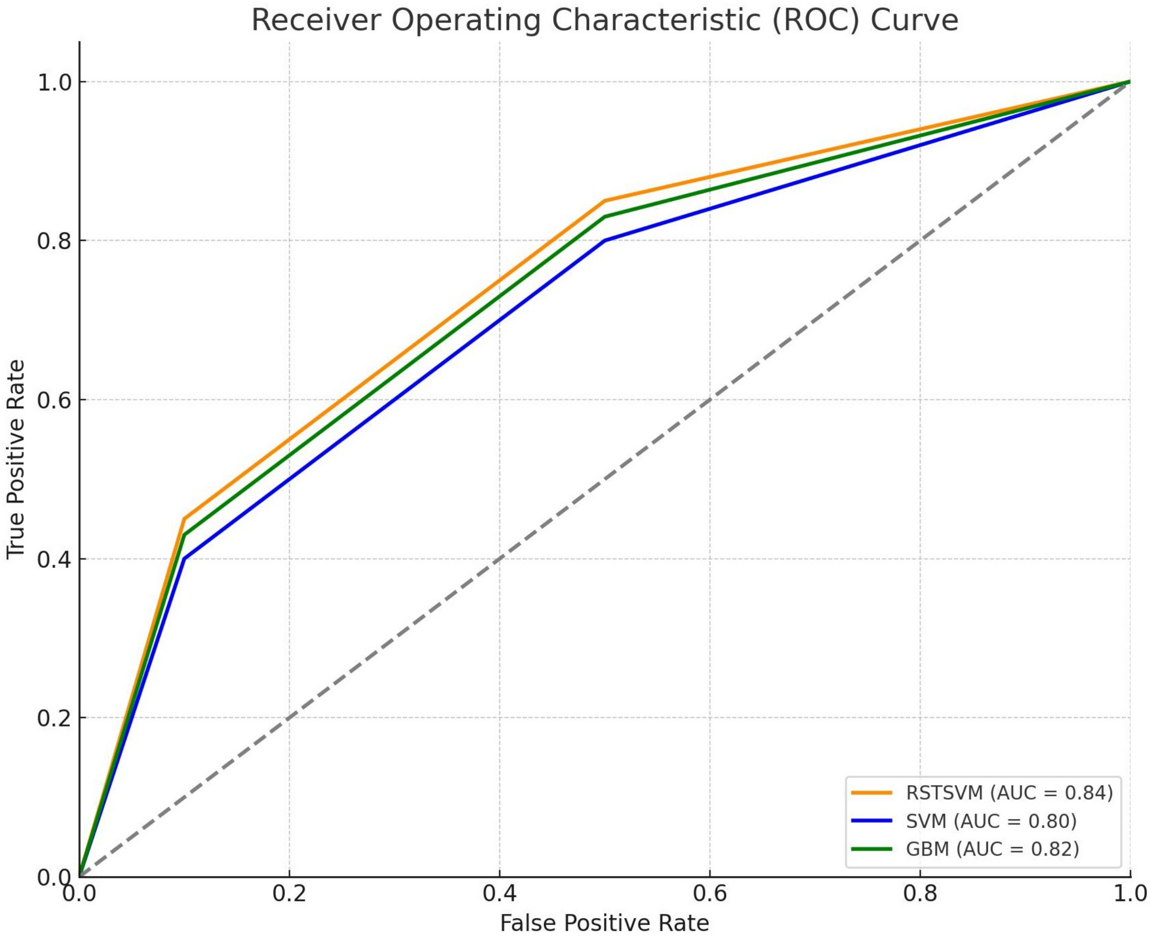

FIGURE 1. ROC curve of RSTSVM. RSTSVM: Robust and Sparse Twin Support Vector Machine; SVM: Support Vector Machine; GBM: Gradient Boosting Machine; AUC: area under the curve.

has the highest importance score of 0.204, indicating a strong influence on the prediction of MSDs. Other significant factors include repetitive hand/arm movements, standing posture, and carrying heavy objects, which are all related to physical strain and ergonomic risks. Social-psychological factors such as working with angry customers and exposure to threats or humiliating behavior also show considerable importance.

4.4 Odds ratio and confidence intervals

The odds ratio was calculated as the exponentiation of the coefficient of the predictor in the logistic regression model. Confidence intervals were calculated using the standard error of the coefficients, providing a range within which the true odds ratio is expected to fall with 95% confidence. The Table 3 summarizes the odds ratios and confidence intervals for the six most significant predictors.

The odds ratios indicated that exposure to fatigue or painful postures, repetitive hand/arm movements, standing posture, carrying heavy objects, using the internet/email, and working with computers were associated with a higher likelihood of developing MSDs. Fatigue or painful postures had the highest odds ratio of 1.97, indicating a strong association with MSDs.

5. Discussion

The study identified several key risk factors that significantly contribute to the likelihood of developing MSDs. The most influential factors included fatigue or painful postures, repetitive hand/arm movements, standing posture and carrying heavy objects. These factors are consistent with the established

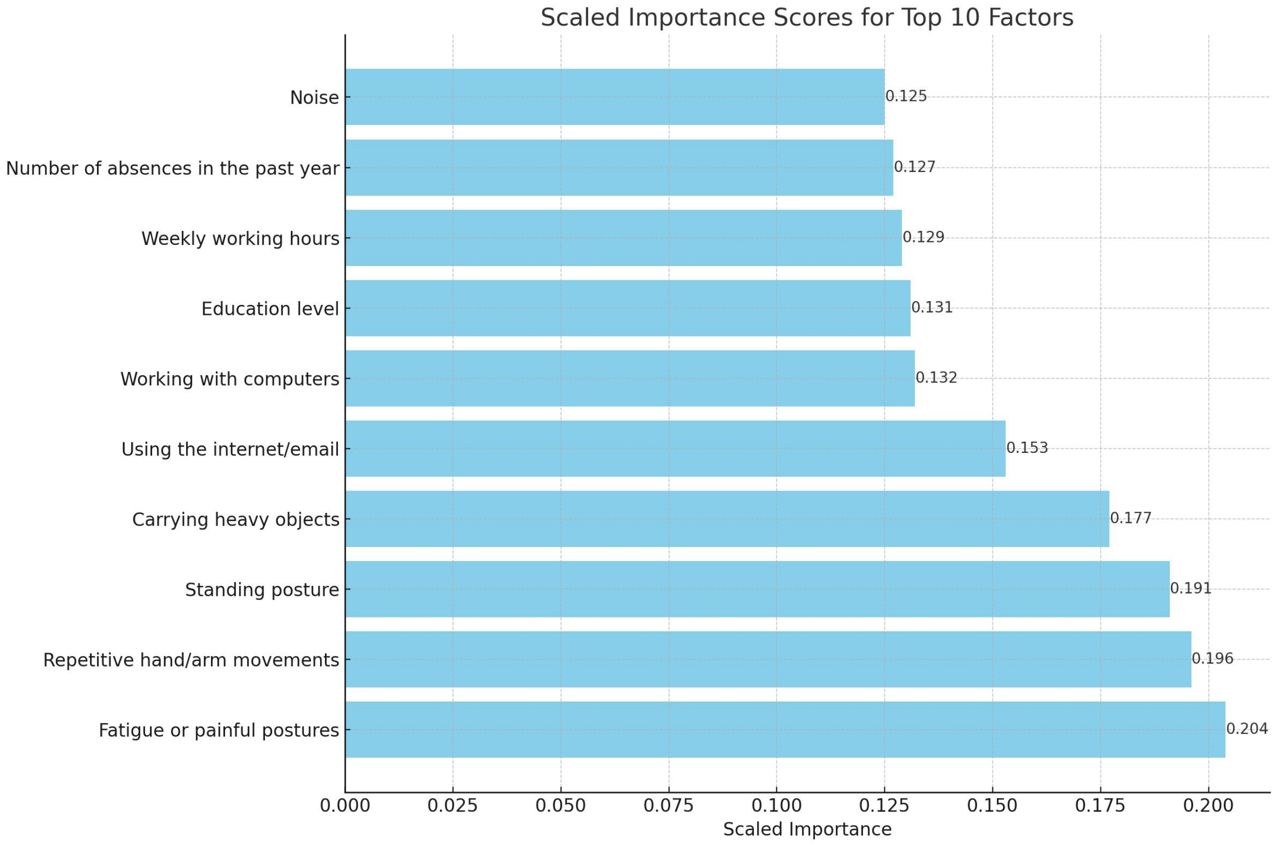

FIGURE 2. Scaled importance scores of RSTSVM.

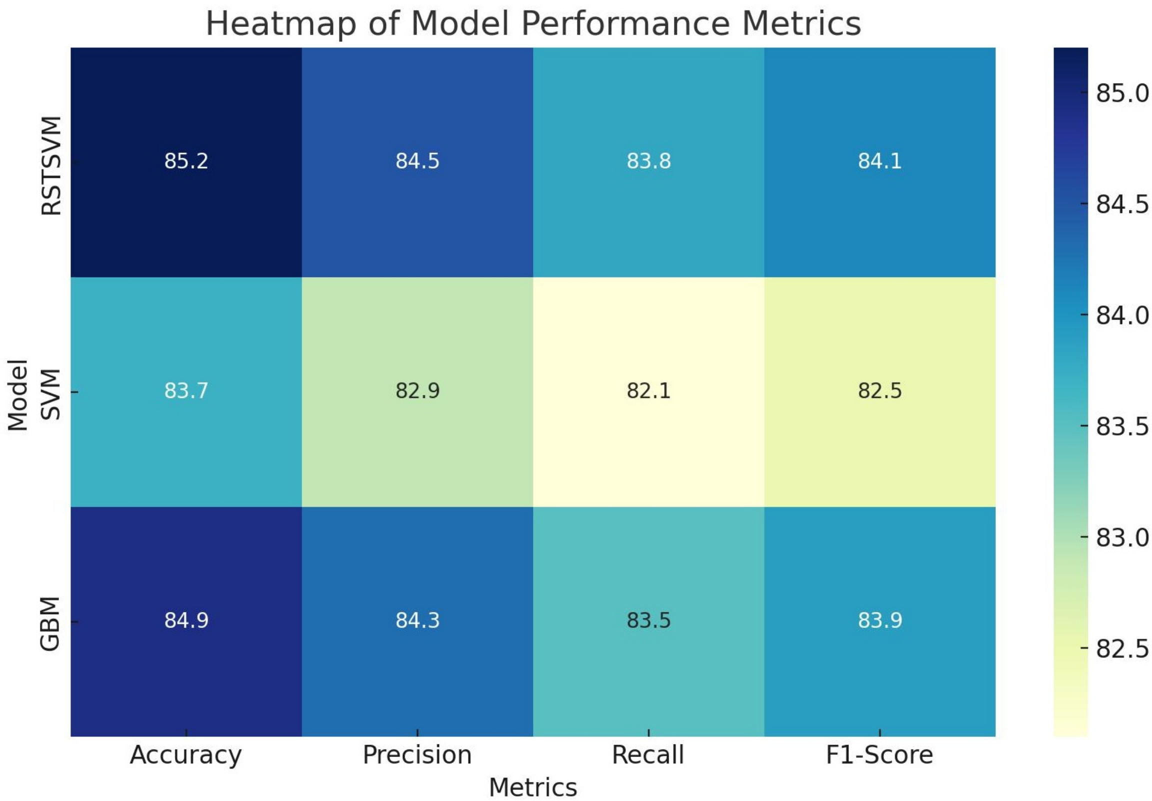

FIGURE 3. Heatmap of model performance. RSTSVM: Robust and Sparse Twin Support Vector Machine; SVM: Support Vector Machine; GBM: Gradient Boosting Machine; AUC: area under the curve.

TABLE 3. The odds ratios for the six most significant predictors.

Predictor	Odds Ratio	95% CI
Fatigue or painful postures	1.97	1.83–2.11
Repetitive hand/arm movements	1.85	1.72–1.99
Standing posture	1.82	1.69–1.95
Carrying heavy objects	1.77	1.64–1.91
Using the internet/email	1.58	1.45–1.72
Working with computers	1.56	1.43–1.69

CI: confidence intervals.

understanding of MSD risk factors [17–26]. For example, Kazeminasab *et al.* [21] (2022) emphasized the role of repetitive tasks and awkward postures in the development of MSDs among office workers. The findings from this study reinforce the importance of addressing these ergonomic risk factors to prevent MSDs.

Moreover, polonged static postures, such as extended periods of sitting, have been linked to lower back pain and other musculoskeletal issues [27–33]. Besharati *et al.* [32] (2020) found that extended periods of sitting contribute to an increased risk of lower back disorders, underscoring the need for dynamic sitting and regular movement breaks. Poor workstation ergonomics, such as non-adjustable chairs and improperly positioned monitors, can lead to increased muscle strain and discomfort. Afroz & Haque (2021) [33] reported that ergonomic interventions, such as adjustable chairs and properly positioned monitors, significantly reduce the incidence of MSDs among office workers.

Psychosocial factors also play a critical role in the development of MSDs. High job demands, low job control and insufficient social support have been identified as significant predictors of musculoskeletal pain. This is consistent with the findings of Tang (2020) [32], who emphasized the interplay between physical and psychosocial stressors in the workplace. Addressing these factors requires a comprehensive approach that includes both ergonomic and organizational interventions.

Intervention strategies should therefore be multifaceted, addressing both physical and psychosocial risk factors. Implementing ergonomic improvements, such as sit-stand desks, ergonomic chairs, and proper monitor placement, can alleviate physical strain. Additionally, promoting a supportive work environment and providing resources for stress management can help mitigate the impact of psychosocial stressors. Furthermore, regular training and education on proper ergonomic practices are essential. Employees should be educated on the importance of maintaining neutral body postures, taking regular breaks, and performing exercises to reduce muscle tension. As demonstrated by Choobineh *et al.* [34] (2021), educational interventions can significantly improve ergonomic behaviors and reduce the risk of MSDs. This study underscores the critical need for targeted interventions to address the multifaceted risk factors associated with MSDs among office workers. By adopting a comprehensive approach that includes ergonomic adjustments, psychosocial support, and regular training, organizations can significantly reduce the incidence and impact of MSDs, thereby improving overall employee well-being and productivity.

Another finding of this study is RSTSVM model exhibited superior performance relative to traditional machine learning models such as the SVM and GBM. This superior performance can be attributed to the RSTSVM model's capability to efficiently manage high-dimensional data while maintaining robustness against outliers and noise [17]. The integration of sparsity within the RSTSVM model further enhanced its interpretability by identifying the most significant variables that contribute to MSDs [34]. This feature is particularly pertinent in occupational health research, where it is critical to understand specific risk factors to develop targeted interventions.

The high AUC-ROC value of 0.84 achieved by the RSTSVM model underscores its robust ability to differentiate between workers with and without MSDs. This level of accuracy suggests that the model could serve as a valuable tool for the early identification of individuals at risk of developing MSDs, thereby facilitating timely interventions. This study is significant in that it validated the robustness and efficacy of the RSTSVM model in predicting MSDs through comparative analysis with SVM and GBM.

The significance of this study lies in its comprehensive approach to identifying and addressing both physical and psychosocial risk factors associated with MSDs. However, this study has several limitations. First, the cross-sectional design does not allow for the establishment of causal relationships between the identified risk factors and the development of MSDs. Longitudinal studies are required to validate these associations over time. Second, the reliance on self-reported data is a limitation, as it can be subject to recall bias and may not accurately reflect the true prevalence of MSDs or the extent of exposure to risk factors. Future research should incorporate objective measures of both MSD symptoms and ergonomic risk factors to enhance the reliability of the findings. Third, missing values in the dataset were imputed using mean or mode imputation. While this method is commonly used, it may introduce bias by not fully capturing the variability in the data. Future studies should consider using more sophisticated imputation methods, such as multiple imputation or machine learning-based approaches, to address this issue.

6. Conclusions

This study underscores the significant burden of MSDs among South Korean male office workers and highlights the critical risk factors contributing to these disorders. The development and validation of the RSTSVM model provide a robust predictive tool for early identification of individuals at risk of MSDs. By addressing both ergonomic and psychosocial risk factors, workplace interventions can effectively reduce the prevalence of MSDs and improve the overall health and productivity of office workers. Future research should continue to explore innovative approaches to prevent and manage MSDs, ensuring a healthier and more sustainable workforce.

AVAILABILITY OF DATA AND MATERIALS

The data presented in this study are provided at the request of the corresponding author. The data is not publicly available because researchers need to obtain permission from the Korea Centers for Disease Control and Prevention. Detailed information can be found at: http://knhanes.cdc.go.kr.

AUTHOR CONTRIBUTIONS

HB—conceptualization, software, methodology, validation, investigation, writing-original draft preparation, formal analysis, writing-review and editing, visualization, supervision, project administration, funding acquisition; contributed to editorial changes in the manuscript; read and approved the final manuscript.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Before conducting the survey, written informed consent was acquired from all participants or their legal guardians. This study employed only pre-existing, anonymized data. It adhered to the principles outlined in the Declaration of Helsinki. The protocol for the Panel Study of Worker's Compensation Insurance received approval from the Institutional Review Board (IRB) of the KNHANES (IRB approval numbers: 2018-01-03-5C-A). All study participants provided written informed consent.

ACKNOWLEDGMENT

Not applicable.

FUNDING

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-RS-2023-00237287, NRF-2021S1A5A8062526) and local government-university cooperation-based regional innovation projects (2021RIS-003).

CONFLICT OF INTEREST

The author declares no conflict of interest. Haewon Byeon is serving as one of the Guest Editors of this journal. We declare that Haewon Byeon had no involvement in the peer review of this article and has no access to information regarding its peer review. Full responsibility for the editorial process for this article was delegated to Xinxin Ye.

REFERENCES

[1] Korea Occupational Safety and Health Agency. Prevalence of among South Korean workers. Journal of Occupational Health. 2022; 64: 123–130.

[2] Korea Occupational Safety and Health Agency. Health survey of agricultural workers in South Korea. Journal of Agricultural Safety and Health. 2021; 27: 56–64.

[3] Ministry of Employment and Labor, South Korea. Updated policies on occupational diseases in South Korea. Journal of Occupational Medicine. 2022; 58: 210–220.

[4] International Labour Organization. Updated facts on safe work practices. 2021. Available at: https://www.ilo.org/topics/safety-and-health-work (Accessed: 15 June 2024).

[5] Hannu T, Kritzinger K. Morbi artificum—a post-ramazzinian academic dissertation on occupational diseases by a pupil of carl von linné. Acta Medico-Historica Adriatica. 2021; 19: 195–220.

[6] World Health Organization. Global health expenditure database. 2021. Available at: https://apps.who.int/nha/database (Accessed: 15 June 2024).

[7] World Health Organization. Protecting workers' health in the 21st century. 2021. Available from: https://www.who.int/news-room/fact-sheets/detail/occupational-health--health-workers (Accessed: 15 June 2024).

[8] National Institute for Occupational Safety and Health (NIOSH). Current research on occupational musculoskeletal disorders. 2021. Available at: https://www.cdc.gov/niosh/ergonomics/index.html/ (Accessed: 15 June 2024).

[9] Korea Statistical Information Service (KOSIS). Updated national health statistics. 2021. Available at: https://kosis.kr/index/index.do/ (Accessed: 15 June 2024).

[10] Ministry of Health and Welfare, South Korea. Comprehensive chronic disease management strategies. 2021. Available at: http://www.mohw.go.kr/eng/ (Accessed: 15 June 2024).

[11] European Foundation for the Improvement of Living and Working Conditions. Sixth European working conditions survey. 2020. Available at: https://www.eurofound.europa.eu/en/surveys/european-working-conditions-surveys-ewcs (Accessed: 15 June 2024).

[12] Vapnik V. The nature of statistical learning theory. 3rd edn. Springer-Verlag: New York. 2021.

[13] Mekonnen TH. Work-related factors associated with low back pain among nurse professionals in east and west Wollega Zones, Western Ethiopia, 2017: a cross-sectional study. Pain and Therapy. 2019; 8: 239–247.

[14] Eubank BHF, Lackey SW, Slomp M, Werle JR, Kuntze C, Sheps DM. Consensus for a primary care clinical decision-making tool for assessing, diagnosing, and managing shoulder pain in Alberta, Canada. BMC Family Practice. 2021; 22: 201.

[15] Tanveer M, Rajani T, Rastogi R, Shao YH, Ganaie MA. Comprehensive review on twin support vector machines. Annals of Operations Research. 2024; 339: 1223–1268.

[16] Wang S, Taha AF, Abokifa AA. How effective is model predictive control in real-time water quality regulation? State-space modeling and scalable control. Water Resources Research. 2021; 57: e2020WR027771.

[17] Bongers PM, Ijmker S, van den Heuvel S, Blatter BM. Epidemiology of work related neck and upper limb problems: psychosocial and personal risk factors (part I) and effective interventions from a bio behavioural perspective (part II). Journal of Occupational Rehabilitation. 2006; 16: 279–302.

[18] Huysmans MA, IJmker S, Blatter BM, Knol DL, van Mechelen W, Bongers PM, *et al*. The relative contribution of work exposure, leisure time exposure, and individual characteristics in the onset of arm-wrist-hand and neck-shoulder symptoms among office workers. International Archives of Occupational and Environmental Health. 2012; 85: 651–666.

[19] Hartvigsen J, Leboeuf-Yde C, Lings S, Corder EH. Is sitting-while-at-work associated with low back pain? A systematic, critical literature review. Scandinavian Journal of Public Health. 2000; 28: 230–239.

[20] Bongers PM, de Winter CR, Kompier MA, Hildebrandt VH. Psychosocial factors at work and musculoskeletal disease. Scandinavian Journal of Work, Environment & Health. 1993; 19: 297–312.

[21] Kazeminasab S, Nejadghaderi SA, Amiri P, Pourfathi H, Araj-Khodaei M, Sullman MJ, *et al*. Neck pain: global epidemiology, trends and risk factors. BMC Musculoskeletal Disorders. 2022; 23: 26.

[22] Basakci Calik B, Yagci N, Oztop M, Caglar D. Effects of risk factors related to computer use on musculoskeletal pain in office workers. International Journal of Occupational Safety and Ergonomics. 2022; 28: 269–274.

[23] Bonlioli R, Caraballo-Arias Y, Salmen-Navarro A. Epidemiology of work-related musculoskeletal disorders. Current Opinion in Epidemiology and Public Health. 2022; 1: 18–24.

[24] Bao S, Howard N, Lin JH. Are work-related musculoskeletal disorders claims related to risk factors in workplaces of the manufacturing industry? Annals of Work Exposures and Health. 2020; 64: 152–164.

[25] Kulkarni M, Yadav T. Prevalence of cumulative trauma disorder of wrist joint in auto mechanical workers. Indian Journal of Public Health Research & Development. 2020; 11: 619–623.

[26] van der Molen HF, Nieuwenhuijsen K, Frings-Dresen MH, de Groene G. Work-related psychosocial risk factors for stress-related mental disorders: an updated systematic review and meta-analysis. BMJ Open. 2020; 10: e034849.

[27] El Bannany M, Khedr AM, Sreedharan M, Kanakkayil S. Financial distress prediction based on multi-layer perceptron with parameter optimization. IAENG International Journal of Computer Science. 2021; 48: 1–12.

[28] Argus M, Pääsuke M. Musculoskeletal disorders and functional characteristics of the neck and shoulder: comparison between office workers using a laptop or desktop computer. Work. 2023; 75: 1289–1299.

[29] Khan SH, Mohan TRC, Abed AAAA, KB S, Bhumik A. Posture related musculoskeletal disorders (MSDs) among computer users in higher education sectors of Malaysia. Malaysian Journal of Medicine & Health Sciences. 2020; 16: 71–78.

[30] Jain R, Meena ML, Rana KB. Risk factors of musculoskeletal symptoms among mobile device users during work from home. International Journal of Occupational Safety and Ergonomics. 2022; 28: 2262–2268.

[31] Tang KHD. A review of psychosocial models for the development of musculoskeletal disorders and common psychosocial instruments. Archives of Current Research International. 2020; 20: 9–19.

[32] Besharati A, Daneshmandi H, Zareh K, Fakherpour A, Zoaktafi M. Work-related musculoskeletal problems and associated factors among office workers. International Journal of Occupational Safety and Ergonomics. 2020; 26: 632–638.

[33] Afroz S, Haque MI. Ergonomics in the workplace for a better quality of work life. In Muzammil M, Khan AA, Hasan F (eds.) Ergonomics for improved productivity (pp. 503–511). 1st edn. Springer: Singapore. 2021.

[34] Choobineh A, Shakerian M, Faraji M, Modaresifar H, Kiani J, Hatami M, *et al*. A multilayered ergonomic intervention program on reducing musculoskeletal disorders in an industrial complex: a dynamic participatory approach. International Journal of Industrial Ergonomics. 2021; 86: 103221.

How to cite this article: Haewon Byeon. Predicting occupational musculoskeletal disorders in South Korean male office workers using a robust and sparse twin support vector machine. Journal of Men's Health. 2024; 20(12): 41-49. doi: 10.22514/jomh.2024.199.