Supplementary Table 2. Comprehensive comparison of findings of studies included

Author/Country of study	Sample size	Target population/ Working sector	Screening tool used	No. of variables studied	Main body region/ Prevalence of WRMSDs	Outcome
Sanchez-Guillen et al.[1] (2024) Spain	651	Surgeons (healthcare sector)	NMQ	59	Cervical, lumbar followed by shoulder Prevalence: 90.6%	**Predictors** (not based on ranking): - Age - Sex - BMI - Height - Total surgeries per year - Perform high frequency of surgery (3–5 per week) - Duration of the surgery - Type of surgery performed - Minor surgery procedures per year - Performing endoscopic surgery procedure - Instrumental bad ergonomics - Inappropriate location of the laparoscopic screen - Long interventions without breaks Machine learning algorithm used: **RF** and **gradient boosting** **Division of data:** not mentioned **Best ML algorithm:** not mentioned Model performance according to site of WRMSDs: (body trunk[a]/hand/knees/feet and ankle) - Accuracy: 0.738/0.667/0.600/0.786 - Sensitivity: 0.469/0.423/0.939/1.000 - Specificity: 0.856/0.893/0.091/0.061 - PPV: 0.590/0.786/0.608/0.783 - NPV: 0.785/0.625/0.500/1.000 AUC-ROC: not mentioned

Luo et al.[2] (2024) China	617	Healthcare professionals (healthcare sector)	Modified NMQ	21	Study only focuses on neck and shoulders MSD Prevalence: 44.89%	**Predictors** (not based on ranking): **<u>Neck MSD</u>** - Feel physically tired after work - Frequent interaction with patients or the public - Turn body around during work - Frequently lean your neck forward or maintain this posture for extended periods during work - Frequently lean neck backward or maintain this posture for extended periods during work - Frequently bend wrists or maintain this posture for extended periods during work - Frequently twist wrists and maintain this posture for extended periods during work - Frequently repeat the same movements with upper arms and fingers multiple times per minute during work - Frequently repeat the same movements with head multiple times per minute during work - Work in uncomfortable postures - Spend long periods sitting during work - Maintain the same posture for extended periods during work **<u>Shoulder MSD</u>** - Feel physically tired after work - Frequently lean your neck forward or maintain this posture for extended periods during work - Frequently lean neck backward or maintain this posture for extended periods during work - Frequently bend wrists or maintain this posture for extended periods during work - Frequently twist wrists and maintain this posture for extended periods during work - Frequently repeat the same movements with upper arms and fingers multiple times per minute during work

- Spend long periods sitting during work
- Maintain the same posture for extended periods during work
- Self-assessment of health status
- Taken sick leave in the past year due to illness
- Having Work-related stress
- Having chronic diseases
- Feel rest periods are not sufficient
- Frequently make large turns of body during work
- Frequently maintain a slightly bent posture for extended periods during work
- Frequently maintain a slightly turned posture for extended periods during work

Machine learning algorithm used:

- Tree-based models: DT, RF, and XGBoost models (Xgboost)
- Single hidden-layer neural network models (MLP)
- Elastic net models (ENet)
- SVM

Division of data: training 75%, testing 25%
Neck best ML algorithm (Training): SVM[b]
Shoulder best ML algorithm (Training): MLP[b]
Neck best ML algorithm (Testing): SVM[b]
Shoulder best ML algorithm (Testing): MLP[b]
Model performance:
Neck MSD
Training

- Best accuracy: SVM & Enet (68.4%)
- Best sensitivity: SVM (86.5%) & XgBoost (85.5%)
- Best specificity: RF (66.7%)
- Highest AUC: RF (77.2%)

Testing

						- Best accuracy: RF (60.6%) - Best sensitivity: Xgboost (78.6%) - Best specificity: RF (61.2%) - Highest AUC: RF (63%) **MAE** 1) SVM: 0.9165 2) MLP: 0.9850 3) RF: 0.9855 **RSME** 1) SVM: 1.0385 2) MLP: 1.0940 3) RF: 1.1045 **<u>Shoulder MSD</u>** **Training** - Best accuracy: SVM (78.1%) - Best sensitivity: SVM (80.2%) - Best specificity: SVM (76.8%) - Highest AUC: SVM (86.6%) **Testing** - Best accuracy: Enet (67.7%) - Best sensitivity: RF (72.1%) - Best specificity: Enet (79.9%) - Highest AUC: XgBoost (73.4%) **MAE** 1) MLP: 0.946 2) XGBoost: 0.974 3) SVM: 1.001
Rmadi et al.[3] (2024) Tunisia	145	Sewing machine operators (manufacturing sector)	NMQ	Not mentioned	Lower back, upper back, followed by hand/wrist Prevalence: 89%	**Predictors** (not based on ranking): - History of MSDs - Job seniority greater than 6.5 years - Age greater than 25 years - Difficulty to keep up with production rhythm - Female sex **Machine learning algorithm used:** CART (Classification and Regression Tree - graphical)

						Division of data: not mentioned **Model performance:** Not mentioned, however comparable to logistic regression based on ROC curve graphic. ROC value not mentioned.
Byeon[4] (2024) South Korea	6,885	Male office workers (multiple sectors)	Not mentioned from Korean Working Conditions Survey (KWCS)	67	Not mentioned Prevalence: 17%	**Predictors** (based on ranking): 1) Fatigue of painful postures 2) Repetitive hand/arm movement 3) Standing posture 4) Carrying heavy objects 5) Using the internet/email 6) Working with computers 7) Educational level 8) Weekly working hours 9) Number of absences in the past years 10) Noise **Machine learning algorithm used:** - RSTSVM - SVM - GBM **Division of data:** using 10-fold cross-validation **Best ML algorithm:** RSTSVM **Model performance:** Accuracy/precision/recall /F1 Score/AUC-ROC 1) RSTSVM 85.2/84.5/83.8/84.1/0.84 2) SVM 83.7/82.9/82.1/82.5/0.80 3) GBM 84.9/84.3/83.5/83.9/0.82
Hanumegowda and Gnanasekaran[5] (2022) India	370	Bus drivers (Transport Sector)	Modified NMQ	21	Lower back, hip/buttocks, followed by shoulder Prevalence: 78%	**Predictors (not based on ranking):** - Involvement in physical activities - Tobacco consumption - Frequent posture change - Egress/ingress - Exposure to vibration

						- On duty breaks - Seat adaptability issues - Tired at end of work - Sleeping in the bus (after duty) **Machine learning algorithm used:** - DT - RF - Naïve Bayes Classifier **Division of data:** 70% train 30% testing **Best ML algorithm from training:** DT and RF (no data on testing)[b] **Model performance on training dataset post 10-fold cross validation:** **1) DT** Post ten-fold cross-validation: Classification accuracy 100% MAE 0 RMSE 0 **2) RF** Post ten-fold cross-validation: Classification accuracy 100% MAE 0.0003 RMSE 0.0068 **3) Naïve Bayes** Post ten-fold cross-validation: Classification accuracy 93.28% MAE 0.0313 RMSE 38.16% *AUC-ROC value: not mentioned
Shaikh and Mandal[6] (2025) India	56	Shuttle car operators (Mining & Quarrying sector)	NMQ [a]	6	Non mentioned Prevalence: not mentioned	**Predictors (not based on ranking):** - Age - BMI - Experience - Posture - Frequency-weighted RMS acceleration (aw)

						- Vibration dose values **Machine learning algorithm used:** artificial neural networks **Division of data:** 0.96% training 0.14% testing **Model performance on testing dataset:** - Accuracy (0.975 ± 0.014) - Precision (0.805 ± 0.083) - Recall (1.000 ± 0.000) - F1 score (0.890 ± 0.053) - AUC-ROC: (0.996 ± 0.005)
Ali et al.[7] (2020) Bangladesh	593	Bank employees (financing sector)	Musculoskeletal subscale of subjective health complaints	11	Study only studies low back pain Prevalence: 36.6%	**Predictors (based on ranking):** 1) Length of employment (>10 years) 2) Age (higher age) 3) Prolong office hours (>9 hours) 4) Presence of chronic illness 5) Physical activity **Machine learning algorithm used:** RF **Division of data:** 70% training 30% testing **Model performance on testing:** Only mentioned accuracy: 0.77 (0.70–0.84) AUC-ROC: not mentioned
Raza et al.[8] (2024) India	88	Heavy vehicle drivers and office workers (transport sector)	Modified NMQ	9	low back pain (LBP), knee pain (KP), followed by neck pain (NP) Prevalence: 87.5% of the heavy vehicle drivers and 67.5% of the office workers experience at	**Predictors based on significance:** **LBP** Age – B coefficient 0.1310/AOR 0.049 (0.034, 0.228)/P value: 0.008 **Neck pain** Sleeping duration (hours) - B coefficient −0.6035/AOR 0.546 (−1.181, −0.026)/p = 0.041 *No significant predictors for knee pain **Machine learning algorithm used:** logistic regression **Division of data:** 70% training 30% testing Model performance: **Training** (LBP/KP/NP) - Accuracy: 0.69/0.63/0.68

					least one type of MSD.	- Precision: 0.69/0.55/0.54 - Recall: 0.71/0.66/0.77 - F1 Score: 0.69/0.59/0.63 **Testing** (LBP/KP/NP) - Accuracy: 0.63/0.56/0.59 - Precision: 0.73/0.47/0.4 - Recall: 0.54/0.64/0.45 - F1 Score: 0.62/0.54/0.42 **AUC-ROC:** - Training (LBP/KP/NP) - AUC-ROC: 0.72/0.68/0.74 - Testing (LBP/KP/NP) - AUC-ROC: 0.65/0.65/0.47
Khoshakhlagh et al.[9] (2024 Iran	2,339	Firefighters (service sectors)	NMQ	12	Not mention site of WRMDS Prevalence: 49.3%	**Predictors based on significance:** - High job stress and high PTSD each increased the probability of musculoskeletal symptoms by 34% - Combined, job stress and high PTSD increased the probability of musculoskeletal symptoms by 37% - Burnout and depression have association with MSD **Machine learning algorithm used**: Bayesian Network **Division of data:** 10-fold cross validation analysis **Model performance:** - Accuracy: 0.742 - Sensitivity: 0.887 - Specificity: 0.593 - AUC-ROC: 0.759
Kar et al.[10] (2023) India	246	Dumper operators (mining and quarrying sector)	NMQ	15	Not mention site of WRMDS Prevalence: not mentioned	**Predictors (based on ranking):** 1) Age 2) Awkward posture, 3) Experience in mines 4) Job demand

	5) Alcohol consumption 6) Smoking cigarettes 7) Work design 8) Marriage status **Machine learning algorithm used:** - Logistic regression (LR) - SVM - DT - GBM - RF Best ML algorithm from training: **RF** Best ML algorithm from testing: **RF** **Division of data**: 80% training 20% testing **Model performance:** **Training** (LR/SVM/DT/GBM/RF) - Accuracy: 0.81/0.79/0.81/0.70/0.79 - Precision: 0.76/0.87/0.83/0.84/0.82 - Recall: 0.72/0.66/0.91/0.81/0.91 - F1 score: 0.79/0.68/0.88/0.83/0.88 **Testing** (LR/SVM/DT/GBM/RF) - Accuracy: 0.64/0.64/0.63/0.61/0.71 - Precision: 0.69/0.76/0.72/0.77/0.75 - Recall: 0.66/0.54/0.77/0.72/0.78 - F1 score: 0.68/0.63/0.74/0.75/0.76 **AUC-ROC:** RF (0.82), GBM (0.79), DT (0.76), SVM (0.73), and LR (0.69)

NMQ: Nordic Musculoskeletal Questionnaire; BMI: body mass index; RF: random forest; ML: machine learning; WRMSD: work-related musculoskeletal disorder; PPV: positive predictive value; NPV: negative predictive value; AUC: area under the curve; ROC, receiver operating characteristic; MSD: musculoskeletal disorder; DT: decision tree; MLP: multilayer perceptron; DT: decision tree; SVM: support vector machine; MAE: mean absolute error; RMSE: root mean square error; CART: classification and regression tree; RSTSVM: robust and sparse twin support vector machine; GBM: gradient boosting machine; AOR: adjusted odds ratio; PTSD: post-traumatic stress disorder.

[a]Body trunk = neck + shoulder + dorsal + lumbar region; [b]According to MAE and RSME.

REFERENCES

1. Sanchez-Guillen L, Lozano-Quijada C, Soler-Silva A, Hernandez-Sanchez S, Barber X, Toledo-Marhuenda JV, et al. A calculator for musculoskeletal injuries prediction in surgeons: a machine learning approach. Surg Endosc 2024;38(11):6577-85.

2. Luo N, Xu X, Jiang B, Zhang Z, Huang J, Zhang X, et al. Explainable machine learning framework to predict the risk of work-related neck and shoulder musculoskeletal disorders among healthcare professionals. Front Public Health 2024;12:1414209.

3. Rmadi N, Sellami I, Feki A, Hammami KJ, Masmoudi ML, Hajjaji M. Exploring multisite musculoskeletal symptoms among sewing machine operators in a tunisian leather and footwear industry using decision tree models. Clin Epidemiol Glob Health 2024;27:101575.

4. Byeon H. Predicting occupational musculoskeletal disorders in South Korean male office workers using a robust and sparse twin support vector machine. J Mens Health 2024;20:41-9.

5. Hanumegowda PK, Gnanasekaran S. Prediction of work-related risk factors among bus drivers using machine learning. Int J Environ Res Public Health 2022;19(22):15179.

6. Shaikh AM, Mandal BB. Predictive modeling of work-related musculoskeletal disorders among shuttle car operators using artificial neural network. J Mines Met Fuels 2025;73:65-74.

7. Ali M, Ahsan GU, Hossain A. Prevalence and associated occupational factors of low back pain among the bank employees in Dhaka City. J Occup Health 2020;62(1):e12131.

8. Raza M, Bhushan RK, Khan AA, Ali AM, Khamaj A, Alam MM. Prevalence of musculoskeletal disorders in heavy vehicle drivers and office workers: a comparative analysis using a machine learning approach. Healthcare (Basel) 2024;12(24):2560.

9. Khoshakhlagh AH, Sulaie SA, Yazdanirad S, Orr RM, Laal F. Relationships between job stress, post-traumatic stress and musculoskeletal symptoms in firefighters and the role of job burnout and depression mediators: a bayesian network model. BMC Public Health 2024;24(1):468.

10. Kar MB, Aruna M, Kunar BM. Risk factors associated with work-related musculoskeletal disorders among dumper operators: a machine learning approach. Clin Epidemiol Glob Health 2023;24:101438. https://doi.org/https://doi.org/10.1016/j.cegh.2023.101438