From Association to Causation: Integrated Causal Inference and Precision Intervention Pathways for Mental Health Risks in Construction Workers

Authors: Zhicheng Ling, Liu Yuhan, Wang Juan, Huanglong, Yellow Dragon

Date: 2025-11-03T00:00:00+00:00

Abstract

Construction workers face severe mental health challenges, with incidence rates of anxiety, depression, stress, and occupational burnout significantly higher than in other industries. However, existing research is largely confined to a "problem-oriented" approach, overly reliant on cross-sectional data and traditional statistical methods, which struggle to effectively handle high-dimensional, nonlinear data and neglect individual heterogeneity, resulting in weak causal inference. This study aims to overcome these limitations by constructing an innovative analytical framework, providing stronger scientific evidence for mental health interventions among construction workers.

This study proposes a "data-driven-causal inference-heterogeneity identification-robust evaluation" four-in-one analytical paradigm. The study randomly selected 1,000 employees from the Fourth Engineering Division of the Second Aviation Engineering Bureau of China Communications Construction Company, ultimately obtaining 912 valid questionnaires. In data analysis, elastic net regression was first employed for high-dimensional variable reduction and feature screening; second, Double Machine Learning (DML) combined with XGBoost was introduced to estimate causal effects and address confounding bias; next, Latent Profile Analysis (LPA) was utilized to identify heterogeneous subgroups of mental health; finally, the robustness of causal inference was evaluated through negative control outcome (NCO) and E-value analysis. All statistical analyses were completed using R 4.5.1 software and relevant core analysis packages, with $P<0.05$ set as the threshold for statistical significance.

The study identified factors such as smoking, gender, sense of work meaning, emotional self-regulation, workplace ostracism, emotional exhaustion, alcohol dependence, and cynicism as having significant causal effects on depression, anxiety, and stress. Latent Profile Analysis classified construction workers into three groups: high-risk (5.7%), medium-risk (32.6%), and mentally healthy (61.7%), with the high-risk group exhibiting significantly higher levels of anxiety, depression, and stress. E-value analysis results demonstrated that the causal effect models possess moderate to high robustness (E-values ranging from 1.21 to 1.78).

This study successfully constructed and applied an integrated analytical framework, substantially enhancing the strength of evidence for causal inference from cross-sectional observational data. The findings not only reveal key risk and protective factors influencing construction workers' mental health but also identify heterogeneous groups with varying mental health risks. These discoveries provide direct scientific evidence for developing precise and efficient mental health intervention strategies, particularly emphasizing the importance of targeting the "emotional exhaustion-cynicism" core pathway and enhancing protective factors such as "emotional self-regulation" and "sense of work meaning," holding significant practical guidance for improving construction workers' mental health.

Full Text

From Association to Causation: An Integrated Causal Inference and Precision Intervention Pathway for Mental Health Risks Among Construction Workers–An Empirical Study Based on a Four-in-One Framework and Cross-Sectional Data

Zhicheng Ling[1], Yuhan Liu[1], Juan Wang[2], Long Huang[1]#

[1]Affiliation 1
[2]Nantong University Xinglin College, Qidong, Jiangsu

#Corresponding Author

Background

Construction workers face severe mental health challenges, with rates of anxiety, depression, stress, and occupational burnout far exceeding those of other industries. However, existing research remains largely "problem-oriented,"overly reliant on cross-sectional data and traditional statistical methods that struggle to handle high-dimensional, nonlinear data while neglecting individual heterogeneity. These limitations result in weak causal inference. This study aims to overcome these constraints by constructing an innovative analytical framework that provides stronger scientific evidence for mental health interventions among construction workers.

Methods

This study proposes a four-in-one analytical paradigm: "data-driven causal inference, heterogeneity identification, and robust evaluation." We randomly sampled 1,000 employees from the Fourth Engineering Office of the Second Aviation Engineering Bureau of China Communications Construction Company, ultimately obtaining 912 valid questionnaires. For data analysis, we first employed elastic net regression for high-dimensional variable reduction and feature screening. Second, we introduced Double Machine Learning (DML) combined with XGBoost for causal effect estimation to address confounding bias. Next, we used Latent Profile Analysis (LPA) to identify heterogeneous subgroups of mental health. Finally, we evaluated the robustness of causal inference through Negative Control Outcome (NCO) and E-value analysis. All statistical analyses were completed using R 4.5.1 software and relevant core analysis packages, with $P < 0.05$ set as the threshold for statistical significance.

Results

The study identified smoking, gender, sense of work meaning, emotional self-regulation, workplace exclusion, emotional exhaustion, alcohol dependence, and cynicism as factors with significant causal effects on depression, anxiety, and stress. Latent Profile Analysis classified construction workers into three groups: high-risk (5.7%), medium-risk (32.6%), and mentally healthy (61.7%). The high-risk group exhibited significantly higher levels of anxiety, depression, and stress.

Conclusions and Significance

This study successfully constructed and applied an integrated analytical framework that substantially enhances the strength of evidence for causal inference from cross-sectional observational data. The findings not only reveal key risk and protective factors affecting construction workers' mental health but also identify heterogeneous groups with different mental health risk profiles. These results provide direct scientific evidence for developing precise and effective mental health intervention strategies, particularly emphasizing the importance of targeting the "emotional exhaustion-cynicism" core pathway and enhancing protective factors such as "emotional self-regulation" and "sense of work meaning." These findings offer important practical guidance for improving construction workers' mental health.

Introduction

The construction industry, as the cornerstone of global economic and social development, shapes the contours of modern civilization. Yet behind this civilization lies a vast group often overlooked by society—construction workers. They are contributors to human civilization's progress, yet they labor long-term in extremely harsh physical and psychological environments, enduring unimaginable pressure. In recent years, as societal attention to mental health issues has generally increased, the psychological well-being of construction workers as a special group has gradually moved from the margins into the spotlight, becoming a public health and social issue that cannot be ignored. Existing research reveals a troubling picture: construction workers' mental health status is generally concerning, with significantly higher rates of anxiety, depression, stress, and resulting occupational burnout and substance abuse (particularly alcohol dependence) compared to many other industries. According to the *2023 Blue Book on Mental Health in the Construction Industry*, as many as 38% of construction workers experience anxiety, 22% show depressive tendencies, and the occupational burnout rate reaches a staggering 45% (曾智 et al., 2017). Behind these cold numbers lies immense suffering at the individual level across physiological, psychological, and social functioning dimensions, with profound impacts on personal quality of life, family harmony, and even safety production across the entire industry. Mental health problems not only severely impair workers' cognitive function, emotional regulation capacity, and interpersonal abilities, increasing their risk of various physical and mental diseases, but also directly lead to reduced concentration and increased operational errors, becoming a critical trigger for safety accidents (王伯军, 2004). Related research indicates that over 80% of safety accidents in the construction industry are directly related to workers' unsafe behaviors, which in turn are closely linked to psychological exhaustion and occupational burnout caused by long-term harsh working conditions and high-intensity labor. Therefore, deeply exploring the internal mechanisms of construction workers' mental health problems, key risk factors, and their complex interactions is not only a humanitarian concern but also an urgent need to ensure the industry' s sustainable development and maintain public safety (陈秋珠, 2000).

Although the severity of mental health issues among construction workers has become increasingly prominent and relevant academic research has made some progress, the entire field remains in a relatively preliminary exploration stage, with its knowledge system and evidence base far from solid and mature. While existing research reveals the universality of the problem, it also exposes many profound limitations that collectively constitute a chronic weakness in the current evidence chain, making it exceptionally fragile when facing complex real-world problems (Samuel et al., 2022). First, in terms of research content, the vast majority of studies adopt a "problem-oriented" perspective, overly focusing on measuring and describing negative psychological states such as anxiety, depression, and stress, while paying little attention to protective psychological factors like work well-being, social support, and traditional lifestyle habits. This "disease model" -centered research approach, while helpful for identifying high-risk individuals, fails to comprehensively depict the complete picture of workers' mental health and, more importantly, cannot effectively guide intervention development from a positive promotion perspective. Second, in terms of research methods, cross-sectional designs dominate, which is inherently inadequate for causal effect interpretation. More critically, existing research generally employs traditional analytical paradigms, relying mostly on correlation analysis (such as linear regression) and simple mediation/moderation effect tests. These methods are inadequate when dealing with the high-dimensional, nonlinear, and multicollinear complex data prevalent in construction worker populations (Chernozhukov et al., 2018). They not only struggle to effectively address model specification bias but also cannot accurately estimate the causal effects of key risk factors while strictly controlling for confounding factors. Additionally, traditional methods tend to treat the research population as a homogeneous whole, using "average effects" to mask substantial heterogeneity among individuals. In reality, construction worker populations exhibit enormous internal differences that likely lead to vastly different psychological reactions when facing the same stressors. Ignoring this heterogeneity not only limits research depth but also renders interventions developed based on "average effects" minimally effective in practice (Samuel et al., 2025). Finally, regarding the robustness of research conclusions, especially sensitivity to unobserved confounding factors, existing literature rarely includes systematic evaluation and discussion. This leaves many seemingly significant "findings" permanently at the association level, with potentially fragile reliability. Once certain unmeasured key variables exist, the entire conclusion could be undermined. These deep-rooted limitations collectively render the evidence chain in current construction worker mental health research exceptionally fragile, unable to support scientific, precise, and effective intervention strategies (Fagbenro et al., 2024).

The Four-in-One Analytical Framework

Faced with the widespread fragility of evidence chains in construction worker mental health research, the fundamental solution lies in profound methodological innovation. To achieve the leap from "describing phenomena" to "revealing causality" and then to "precision intervention," a more powerful, rigorous, and comprehensive analytical framework must be constructed. To this end, we propose and systematically elaborate an integrated analytical paradigm of "data-driven causal inference, robust evaluation, and heterogeneity identification." Each component of this paradigm addresses a specific weakness in current research, and their integrated application will collectively advance the entire field toward higher-level scientific evidence. Simultaneously, through causal effect estimation and in-depth heterogeneity research, this study will further supplement and refine theoretical research and preliminary guidance for precision intervention in construction worker mental health.

2.1 Research Process

Based on collaboration with the Fourth Engineering Office of the Second Aviation Engineering Bureau of China Communications Construction Company, this study conducted random sampling of 1,000 regular employees (50% of the total company workforce) through an online questionnaire system from April to

May 2023. After data cleaning (excluding incomplete, anomalous, and invalid responses), 912 valid questionnaires were retained (valid response rate: 91.2%), including 790 males and 122 females. To ensure ethical standards and data quality, we conducted thorough briefings before the survey, used covert backend sampling, and provided secondary explanations to selected employees to confirm their informed consent and accurate understanding of the questionnaire.

In terms of analytical strategy, this study constructed and implemented an integrated analytical paradigm (Figure 1 [Figure 1: see original paper]). The paradigm begins with data-driven approaches: using elastic net regression for feature screening to prevent model specification bias; then shifts to causal inference: applying double machine learning models to estimate causal effects of key factors on depression, anxiety, and stress after controlling for high-dimensional confounding; subsequently conducts robustness evaluation: introducing negative controls screened by network centrality indicators and E-value analysis to quantify potential impacts of unmeasured confounding; and finally achieves heterogeneity identification: through latent profile analysis, identifying high-risk subgroups beneath average effects and characterizing their unique risk-protective factor profiles.

Figure 1. Research Flowchart

2.2 Algorithm Selection

To avoid model specification bias and achieve data-driven feature screening, this study employed elastic net regression for high-dimensional variable reduction. By combining L1 and L2 regularization, this method simultaneously performs variable selection and coefficient shrinkage, yielding a refined and robust feature set that lays the foundation for subsequent causal inference.

To establish more causal interpretations in cross-sectional data, this study introduced the cutting-edge econometric double machine learning model. DML aims to address the core challenge in observational research—confounding bias. Its fundamental idea is: using machine learning models to separately capture the portions of treatment variables (risk factors) and outcome variables (mental health) that can be explained by confounding variables, "stripping" them from the original variables, and then regressing the remaining relatively exogenous variation to obtain more robust causal effect estimates. This approach allows us to flexibly capture complex relationships between variables using algorithms like XGBoost without prespecifying linear forms, providing an ideal tool for handling the high-dimensional, nonlinear data in construction worker research.

To reveal potential heterogeneous subgroups behind overall average effects, this study adopted latent profile analysis. This model identifies mutually exclusive latent classes based on individuals' response patterns on mental health indicators using probabilistic models, thereby segmenting the population and providing targeted evidence for precision intervention. Unlike traditional clustering methods, LPA operates within a probabilistic modeling framework, identifying latent class structures through maximum likelihood estimation and assigning each individual a posterior probability of class membership, thus providing statistically testable classification results.

2.3 Sensitivity Analysis

Even when applying state-of-the-art causal inference methods, conclusions still rely on the strong assumption of "no unmeasured confounding." However, in social science research, unobserved or unmeasured confounding factors are unavoidable. To systematically evaluate bias that may arise from unmeasured confounding, this study introduced two complementary tools: negative control outcome analysis and E-value analysis.

Negative control variables refer to variables that theoretically have no causal association with either the exposure or outcome. This study innovatively quantified their screening process through two thresholds: first, calculating the comprehensive centrality of candidate variables in the network and selecting the top 70% of nodes; second, requiring their correlation coefficient with depression, anxiety, and stress to be below 0.2. Only variables simultaneously satisfying "high confounding potential, low causal association" could serve as negative controls. If the model detects significant "effects" on these variables, it suggests the presence of uncontrolled confounding.

The E-value provides a more direct quantitative standard, defined as the minimum strength of association that an unmeasured confounder would need to have with both the exposure and outcome to fully explain the observed effect. A larger E-value indicates that stronger unmeasured confounding would be required to overturn the conclusion, suggesting more robust results.

2.4 Scale Selection

This study employed a series of standardized scales to measure core variables (see Appendix Table 1). All scales demonstrated acceptable internal consistency reliability in this sample (Cronbach' s $\alpha = 0.773$–0.926). Mental health was assessed using the 21-item Depression Anxiety Stress Scale (DASS-21) to measure symptom severity. Sleep problems were quantified using the 7-item Insomnia Severity Index (ISI). Alcohol use behavior was evaluated through the 25-item Alcohol Dependence Scale. Emotional exhaustion, cynicism, and reduced professional efficacy were measured using the 15-item Maslach Burnout Inventory-General Survey (MBI-GS). Work-family conflict was assessed using 5 items from the Work-Family Conflict Scale. Individual psychological capital and resource pressure were evaluated using the 10-item General Self-Efficacy Scale and the Scarcity Mindset Scale. Workplace exclusion was measured using the Workplace Ostracism Scale.

2.5 Statistical Tools

This study used R 4.5.1 software for data analysis. The analytical approach followed the "data-driven causal inference, heterogeneity identification, and robust evaluation" framework. Continuous variables are presented as mean ± standard deviation (M ± SD), while categorical variables are presented as frequencies or percentages. Advanced statistical analyses were performed using core R packages including "glmnet," "DoubleML," "simex," "EValue," and "tidyLPA." In all statistical analyses, $P < 0.05$ was considered statistically significant.

3.1 Descriptive Statistics of Demographic Variables

This study included 912 participants. Demographic characteristic analysis showed that the sample was predominantly male (86.62%). Age distribution was relatively balanced, with 52.08% under 35 years old and 47.92% aged 35 and above. Work tenure showed diverse distribution: 28.73% had less than 5 years, 37.72% had 5–15 years, and 33.55% had more than 15 years. Regarding behavioral and health characteristics, 56.47% of participants had drinking habits, the workplace injury incidence rate was relatively low at 18.86%, COVID-19 infection rate was high at 81.47%, and the vast majority were migrant workers (92.87%). The sample overall accurately reflected the construction industry population dominated by young male migrant workers.

Table 1. Demographic Descriptive Analysis

3.2.1 Correlation Analysis

During the feature engineering stage, we plotted Pearson correlation heatmaps with various candidate features (covering individual traits, organizational contexts, etc.) to visually quantify the direction and strength of linear associations between variables using color gradients. This approach helped identify key features with high inter-variable correlations while examining feature collinearity, guiding feature selection and derived feature construction, thereby ensuring the effectiveness and parsimony of the feature set and laying a foundation for subsequent model interpretation and robustness validation.

Figure 2. Correlation Heatmap

3.2.2 Elastic Net Regression

This study used elastic net regression for embedded feature screening. After tuning via 10-fold cross-validation ($\lambda = 0.0153$, $R^2 = 0.593$), the model retained 27 non-zero coefficients, with coefficient values serving as measures of feature importance. As shown in Figure 3 [Figure 3: see original paper], emotional self-regulation, sense of work meaning, work well-being, alcohol consumption, and utilization of emotions ranked as the top five negative contributors; cynicism, emotional exhaustion, workplace exclusion, alcohol dependence, and organizational identification constituted the first tier of positive contributors. All non-zero coefficients were saved as the variable set for subsequent analysis.

Figure 3. Non-Zero Coefficients from Feature Engineering

3.3 Causal Effect Estimation

Using the Double Machine Learning framework with XGBoost as the base learner and selecting optimal parameters after tuning (Table 2), results showed (Figure 4 [Figure 4: see original paper]) that after stripping away high-dimensional confounding, eight independent variables–smoking status, gender, sense of work meaning, emotional self-regulation, workplace exclusion, emotional exhaustion, alcohol dependence, and cynicism–demonstrated relatively significant causal effect estimates on the depression-anxiety-stress model. Moreover, the overall average treatment effects across all pathways were relatively strong (Table 3).

Table 2. DoubleML Hyperparameter Configuration Table

Figure 4. DoubleML Causal Inference Results

Table 3. Specific Coefficients of Significant Variables from DoubleML Causal Inference

Variable	Average Treatment Effect (ATE)	95% Confidence Interval
Emotional Self-Regulation	[-0.201, -0.025]	
Sense of Work Meaning	[-0.146, -0.004]	
Smoking Status	[0.022, 0.042]	
Gender	[0.028, 0.207]	
Workplace Exclusion	[0.087, 0.239]	
Emotional Exhaustion	[0.098, 0.307]	
Alcohol Dependence	[0.048, 0.365]	
Cynicism	[0.141, 0.283]	

3.4.1 Covariate Selection for NCO Analysis

Network centrality analysis showed that "self-emotion evaluation" had standardized comprehensive centrality in the top 70% of all variable nodes, demonstrating its high correlation with multiple confounders (see Appendix Table 2). Moreover, prior data indicated no theoretical causal path to the main outcome and a correlation coefficient below 0.2, satisfying negative control assumptions. Therefore, it was uniquely included in the final covariate set to absorb residual confounding.

3.4.2 Sensitivity Analysis

Through Negative Control Outcome (NCO) analysis using "self-emotion evaluation" as the control variable (Table 4), NCO p-values for all variables were greater than 0.05, supporting model validity. Overall, this sensitivity analysis suggests that causal inference between depression-anxiety-stress and most psychological-behavioral variables possesses considerable reliability. In this study, E-values for each effect model ranged from 1.21 to 1.78, indicating moderate-to-high robustness of the effect models.

Table 4. Sensitivity Analysis Results

Variable	NCO P-value	Main Effect RR	e_{point}
Emotional Self-Regulation			
Sense of Work Meaning			

3.5.1 Model Selection

Using anxiety, depression, and stress subscale data from the DASS-21 as outcome variables, this study tested 2–6 profiles commonly used in latent profile analysis. Model selection was based on comprehensive evaluation of AIC, BIC, and entropy values (Table 5). Ultimately, the 3-profile solution showed the best fit. The clustering plot (Figure 5 [Figure 5: see original paper]) showed clear boundaries, indicating good discrimination and comprehensive classification in this clustering.

Table 5. Parameters for Tested Profiles

Figure 5. Cluster Profile Plot

3.5.2 Profile Characterization

Through between-group comparisons (Table 6), the study found that Class 1 (5.7%) exhibited a "high triad" pattern—mean anxiety, depression, and stress scores reached 22.1, 23.7, and 26.3 respectively, all significantly higher than the overall population, representing a high-risk group. Class 2 (32.6%) was the medium-risk group, with scores of approximately 10–15 across the three dimensions, falling in the borderline range requiring close attention and timely support. Class 3 (61.7%) had mean scores below 4 across all dimensions and could be considered the "mentally healthy" group. Through factor identification, the study found six high-risk factors for the high-risk group: alcohol dependence, cynicism, emotional exhaustion, work-family conflict, scarcity mindset, and workplace exclusion. Self-capability motivation, sense of work meaning, leadership support, and work well-being emerged as significant protective factors for the population (Figure 6 [Figure 6: see original paper]).

Table 6. Between-Profile Comparisons

Figure 6. Core Factors by Profile

Discussion

This study conducted an in-depth causal exploration and heterogeneity analysis of mental health risk factors among Chinese construction workers by constructing and applying a "four-in-one" analytical paradigm that integrates cutting-edge econometric and machine learning methods based on cross-sectional survey data. The findings not only validate and supplement core theories in occupational health but, more importantly, provide solid scientific evidence and clear target pathways for precision intervention.

The core finding of this study—that emotional exhaustion and cynicism have significant causal effects on depression, anxiety, and stress (DASS)—provides strong empirical evidence for understanding the psychological consequences of occupational burnout, profoundly corroborating the Job Demands-Resources (JD-R) model and Conservation of Resources (COR) theory (Ling et al., 2025). The JD-R model posits that when job demands continuously deplete employees' resources without effective replenishment, emotional exhaustion results; continued resource loss further triggers defensive cynicism, an emotional detachment strategy adopted to preserve remaining psychological resources. Our causal estimates, after strictly controlling for confounding factors, quantify this pathological pathway from "resource depletion" to "psychological distress," confirming that emotional exhaustion is not only a core dimension of occupational burnout but also a key causal antecedent for inducing generalized negative emotions (Fernandez et al., 2000).

Moreover, this study transcends traditional research' s singular focus on risk factors by identifying the significant protective roles of sense of work meaning and emotional self-regulation through causal inference. This provides empirical support for COR theory, which emphasizes that individuals not only strive to protect existing resources but also actively invest in and acquire new resources (such as psychological capital and positive work experiences) to cope with stress and threats. Sense of work meaning, as a key internal resource, helps workers find value and belonging in harsh environments, thereby buffering the impact of external stressors (宋志方 et al., 2010). Emotional self-regulation, as a dynamic personal skill resource for managing emotional responses, can directly reduce the risk of anxiety and depression (张茂棠 et al., 2009). Therefore, this study not only validates the core mechanisms of JD-R and COR theories but also constructs a more comprehensive "demands-resources" balance model by integrating risk and protective factors, supplementing existing theories' insufficient exploration of protective factors' causal roles in occupational health contexts.

We found that high-risk factors identified by LPA highly overlapped with risk factors showing significant causal effects in the DML model. This convergence holds profound practical significance: it reveals that factors driving population-level average risk increases are precisely the problems concentrated among extreme high-risk individuals. This convergent evidence substantially enhances the credibility of these factors as core intervention targets. Conversely, protective factors identified by LPA also highly coincided with protective causal variables in DML, further confirming the critical role of enhancing these resources in building psychological resilience.

This overlap and interaction provides a direct blueprint for developing "precision-targeted" intervention strategies. For the 61.7% "mentally healthy group," intervention should focus on universal prevention, such as broadly enhancing sense of work meaning and emotional regulation skills through corporate culture promotion and team activities. For the 32.6% "medium-risk group," selective prevention should be implemented, providing mental health screening and stress management workshops to prevent sliding into high-risk status. For the critical 5.7% "high-risk group," indicated prevention and early treatment must be implemented, offering one-on-one counseling, Employee Assistance Program (EAP) services, and considering comprehensive interventions such as job adjustments (Mrazek & Haggerty, 1994). This risk-stratified intervention model can allocate limited public health resources with maximum efficiency to where they are most needed, achieving optimal intervention effects.

This study's most prominent contribution lies at the methodological level. Faced with the fundamental challenge of causal inference from cross-sectional data, our proposed "data-driven causal inference, robust evaluation, and heterogeneity identification" four-in-one framework provides a systematic, highly robust solution for handling complex observational data, whose advancement and universality warrant in-depth discussion.

In terms of advancement, this study' s integrated analytical paradigm makes core contributions through methodological synergistic innovation, systematically advancing data analysis from descriptive association to robust causal inference and precise population segmentation. First, data-driven elastic net screening avoids subjective model specification bias from the outset, laying an objective foundation for causal identification (Zou & Hastie, 2005). Subsequently, the double machine learning model controls high-dimensional confounding through flexible machine learning algorithms, enabling us to obtain near-causal effect estimates from cross-sectional data and significantly enhancing the theoretical value of research findings. More importantly, latent profile analysis lifts the veil on "average effects," identifying heterogeneous subgroups with distinct internal risk patterns, transforming "precision intervention" from concept into actionable practice. Finally, the built-in robustness evaluation framework goes beyond conventional analysis by actively challenging and quantifying the robustness of research conclusions to unmeasured confounding, greatly enhancing scientific credibility. Additionally, the quantitative indicator screening scheme for NCO covariate selection in this study demonstrates originality (Steger et al., 2012).

In terms of universality, this framework is not limited to construction workers or mental health domains. Any field involving high-dimensional data, complex interactions, and difficulties in conducting randomized controlled trials—such as educational economics (evaluating education policy effects), sociology (studying causes of social inequality), or epidemiology (exploring chronic disease risk factors)—can 借鉴 this paradigm to mine higher-grade, more causal evidence from existing observational data, saving research costs, improving pilot study quality, and breaking certain research barriers, thereby promoting scientific decision-making and precision intervention in related fields.

Limitations and Future Directions

Despite significant methodological and applied advances, this study has several limitations that point to future research directions.

First, the inherent limitation of cross-sectional design is this study' s fundamental constraint. Although methods like DML substantially strengthen causal inference, they cannot replace true temporal sequences for establishing causal directionality. Future research should establish prospective cohorts with at least two waves of follow-up surveys to validate the temporal correctness of the causal pathways discovered in this study.

Second, all data originate from self-reports, which may be subject to common method bias and recall bias, particularly regarding sensitive issues like alcohol use. Future research could incorporate objective indicators, such as monitoring physiological stress markers (e.g., heart rate variability, cortisol levels) through wearable devices, to provide more comprehensive and objective health assessments.

Finally, intervention development and effectiveness evaluation represent the ultimate goal. While this study identifies intervention targets, the next step is to design specific intervention programs based on these targets and evaluate their cost-effectiveness through randomized controlled trials (RCTs) or quasi-experimental designs, thereby truly transforming "strong action evidence" into practical outcomes that improve construction workers' psychological well-being.

Author Contributions

Zhicheng Ling: conceptualization, methodology, data analysis, visualization, and original draft writing.
Yuhan Liu: data analysis and original draft writing.
Juan Wang: data analysis and original draft writing.
Long Huang: data collection, conceptualization, methodology, final manuscript review, and funding acquisition.

Ethics Statement

This study adheres to the Declaration of Helsinki and has been approved by the Ethics Review Committee of Wannan Medical College. All participants provided informed consent after thorough explanation. The research process strictly followed confidentiality principles, with data anonymized to ensure participants' privacy and rights were protected.

Funding

This research was supported by the following grants: Anhui Provincial Department of Education University Research Project (Grant No. 2023AH051720), Anhui Provincial Department of Education Young and Middle-aged Teacher Training Program (Project No. JNFX2024040), Anhui Provincial Education Research Project (JK24104), Anhui Provincial Education Information Technology Research Project (AH2024072), Anhui Provincial New Era Talent Training Project (Graduate Education) (2022lhpysfjd065, 2024lhpysfjd067, 2024zyxwjxalk199), and Wannan Medical College Doctoral Research Startup Project (WYRCQD2022030).

References

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21(1), C1–C68. https://doi.org/10.1111/ectj.12097

Fagbenro, R. K., Sunindijo, R. Y., Illankoon, C., & Frimpong, S. (2024). Importance of Prefabrication to Easing Construction Workers' Experience of Mental Health Stressors. *International Journal of Environmental Research and Public Health*, 21(9), 1218–1218. https://doi.org/10.3390/IJERPH21091218

Fernandez, R. M., Castilla, E. J., & Moore, P. (2000). Social Capital at Work: Networks and Employment at a Phone Center. *American Journal of Sociology*, 105(5), 1288–1356. https://doi.org/10.1086/210432

Ling, Z., Xu, Y., Tao, M., Zhang, B., Zhang, M., Zhang, Z., Da, X., Liu, X., & Huang, L. (2025). Construction workers' depression, anxiety, stress, and risk factors in China: A cross-sectional study. *Journal of Global Health*, https://doi.org/10.7189/jogh.15.04167

Samuel, F., Yosia, S. R., Changxin, W. C., & Frimpong, B. E. (2022). Domains of Psychosocial Risk Factors Affecting Young Construction Workers: A Systematic Review. *Buildings*, 12(3), https://doi.org/10.3390/BUILDINGS12030335

Samuel, F., Yosia, S. R., Changxin, W. C., Frimpong, B. E., Ayirebi, D., & Kolawole, F. R. (2025). A scoping review of research on mental health conditions among young construction workers. *Construction Innovation*, 25(1), https://doi.org/10.1108/CI-06-2023-0133

Steger, M. F., Dik, B. J., & Duffy, R. D. (2012). Measuring Meaningful Work: The Work and Meaning Inventory (WAMI). *Journal of Career Assessment*, 20(3), 322–337. https://doi.org/10.1177/1069072711436160

Zou, H., & Hastie, T. (2005). Regularization and Variable Selection Via the Elastic Net. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 67(2), 301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x

宋志方, 鹿德智, 甄惠君, 贾晓民, & 张美荣. (2010). 煤矿井下工人心理健康水平研究. 中国健康心理学杂志, 18(01), https://doi.org/10.13342/j.cnki.cjhp.2010.01.037

张茂棠, 肖全华, 梁宁, & 邝云洪. (2009). 接触正己烷印刷厂工人心理健康状况调查. 中国职业医学, 36(03), 222-224+226. https://doi.org/10.20001/j.issn.2095-2619.2009.03.016

曾智, 陈雯, 石景容, & 凌莉. (2017). 企业工人社会心理因素对心理健康影响结构方程模型分析. 中国职业医学, 44(02), 188-192+197.

王伯军. (2004). 国有企业下岗职工心理健康对照研究. 中国健康心理学杂志, 04, 301-302+274. https://doi.org/10.13342/j.cnki.cjhp.2004.04.027

陈秋珠郭文斌. (2000). 下岗职工心理健康状况调查. 健康心理学杂志, 04, 465-468. https://doi.org/10.13342/j.cnki.cjhp.2000.04.053

Appendix

Appendix Table 1. Detailed Scale Information

Scale Name (Abbreviation)	Number of Items	Scoring Rules	Total Score Range	Cutoff/Clinical Grading	Cronbach's α
Depression Anxiety Stress Scale (DASS-21)	21	4-point (0-3)	0-126 (subscale scores summed $\times$2)	Depression: 0-9 (normal), 10-13 (mild), 14-20 (moderate), 21-27 (severe), 28+ (extremely severe)Anxiety: 0-7 (normal), 8-9 (mild), 10-14 (moderate), 15-19 (severe), 20+ (extremely severe)Stress: 0-14 (normal), 15-18 (mild), 19-25 (moderate), 26-33 (severe), 34+ (extremely severe)	0.829 (depression)0.875 (anxiety)0.922 (stress)
Insomnia Severity Index (ISI)	7	7-point frequency scale (0-6)	0-28	0-7 (no insomnia), 8-14 (mild), 15-21 (moderate), 22-28 (severe)	
Alcohol Dependence Scale	25			0 (none), 1-13 (mild), 14-21 (moderate), 22-30 (severe), 31+ (extremely severe)	

Scale Name (Abbreviation)	Number of Items	Scoring Rules	Total Score Range	Cutoff/Clinical Grading	Cronbach's α
Maslach Burnout Inventory-General Survey (MBI-GS)	15	5-point (1-5)	Emotional Exhaustion: 0-30Cynicism: 0-24Reduced Efficacy: 0-36	Higher scores indicate greater burnout severity per manual standards	
Work-Family Conflict Scale	5	5-point (1-5)		Higher scores indicate greater work-family conflict	
General Self-Efficacy Scale (GSES)	10	4-point (1-4)		Higher scores indicate greater general self-efficacy	
Scarcity Mindset Scale		7-point (1-7)		Higher scores indicate greater perceived resource scarcity and psychological stress	
Workplace Ostracism Scale		5-point frequency scale (1-5)		Higher scores indicate greater perceived workplace exclusion	
Work Well-Being Sense of Work Meaning Self-Capability Motivation Work Self-Efficacy					

Scale Name (Abbreviation)	Number of Items	Scoring Rules	Total Score Range	Cutoff/Clinical Grading	Cronbach's α
Career Centrality					
Proactive Personality					
Self-Emotion Evaluation					
Work-Family Conflict					
Self-Capability Proof					
Emotional Self-Regulation					
Childhood SSS					
Others' Emotion Evaluation					
Pro-Organizational Unethical Behavior					
Workplace Injury Frequency					

Appendix Table 2. Standardized Network Centrality Node Coefficients

Variable	Betweenness Centrality	Closeness Centrality	Eigenvector Centrality	Comprehensive Centrality

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv —Machine translation. Verify with original.