Journal of Public Health Sciences
Vol. 5, No. 02, pp. 315-327
journal.iistr.org/index.php/JPHS
DOI: 10.56741/IISTR.jphs.001951

Predictive Statistical Modeling of Unemployment-Related Lack of Health Insurance Coverage among US Adults

[1]Emmanuel Thompson*, [1]Mohamed Abu Sheha, [1]Seidu Sofo
Corresponding Author: *ethompson@semo.edu
[1] Southeast Missouri State University, Missouri, USA

ARTICLE INFO

Article history
Received 29 December 2025
Revised 20 February 2026
Accepted 22 March 2026

Keywords
Binary Logistics
Health Insurance
Predictive Modeling
Risk Factors
Unemployment

ABSTRACT

The study aimed to develop an accurate predictive model to identify individuals without health insurance due to unemployment. The study participants comprised 2,496 individuals who lacked general health insurance coverage in the United States for the period 2020-2023. Microdata records from the National Health Interview Survey available in the Integrated Public Use Microdata Series (IPUMS) served as the data source. The response variable was a binary variable indicating whether an individual had no health insurance (uninsured) due to unemployment. Statistical analyses involved descriptive measures: frequencies, percentages, and quartiles, and inferential tests: Pearson's chi-squared test, Wilcoxon rank-sum test, and Goodman and Kruskal's measures. For predictive analysis, data were partitioned into training and testing sets, with class imbalance addressed using the Synthetic Minority Over-sampling Technique (SMOTE). Binary logistic regression and probit models were applied, and model performance was assessed using accuracy, precision, recall, and F1 score. Most participants reported being in good, very good, or excellent health. The demographic and health-related factors in this study had limited standalone predictive values. However, Probit and Logit models identified obesity, age, lower education, and lack of healthcare access as critical risk factors. Furthermore, the prediction performance metrics for both Logit and Probit models were identical.

This is an open-access article under the CC–BY-SA license.

Introduction

Lack of health insurance significantly impacts overall health and well-being, especially among unemployed individuals [1[,]2]. Individuals often lose their health insurance coverage when they become unemployed, making unemployment a critical factor in the prevalence of uninsured adults [3]-[5]. Individuals without health insurance due to unemployment pose a challenge to the health sector. According to Claxton et al. [6], loss of a job does not only mean loss of income but also loss of employer-related health insurance and a higher out-of-pocket healthcare cost [7]. The loss of income implies that the uninsured would also pay the full price for their medications [8].

Several factors account for the loss of health insurance due to unemployment. The most common factors include the changing nature of work, the changing nature of the labor market, and employers' neglect of health outcomes [9]. A study of middle-aged adults in the United States by Kwon et al. [10] identified unstable employment and unskilled manual jobs as risk factors for loss of insurance coverage. Additionally, individuals with poor health and incomes below the federal poverty level were also at risk of being uninsured [10].

Pratap et al. [9] characterized unemployment not only as an economic issue, but also as a public health problem. The rise in opioid overdose, for example, can be attributed in part to waning economic opportunities in the United States [11]. Loss of insurance coverage due to unemployment has had negative effects on health indicators [12], including mental health [13],[14]. For instance, [13] reported that unemployed individuals were more in danger of committing suicide than their insured counterparts. According to Nguyen et al. [7], an unemployment insurance claim has a statistically significant negative effect on prescription drug utilization. Unemployment-related reductions in prescription refills are highest in states without expanded Medicaid services—their out-of-pocket costs skyrocketed by 168%. In a study of North Carolina in the United States, Shafer et al. [15] reported that 15% of unemployed individuals gained Medicaid coverage statewide, adding that the percentage was higher in socially vulnerable counties.

Researchers have suggested preventive measures as key to solving the public health crisis caused by loss of insurance coverage due to unemployment [9]. Stroud et al. [16] noted that while young adults are covered by provisions of the Affordable Care Act, little attention is paid to them in the healthcare system. Therefore, they recommended investing in the health and well-being of young adults [16]. Unemployment benefits are used as a stopgap to help reduce the harmful effects of unemployment on depression [17]. Cylus et al. [18] reported that a 63% increase in benefits mitigated the effects of unemployment on health, concluding that unemployment benefits have the potential to enhance health. Another preventive measure suggested to mitigate the negative effects of loss of insurance coverage is the provision of rapid

housing. There is evidence showing that the provision of rapid housing positively impacted short-term health outcomes, such as hospitalizations and shorter periods of hospitalization [19]. A critical factor in the effectiveness of preventive measures, the early identification of individuals who lack health insurance due to unemployment, is an area that has been given little attention. Therefore, there is a need to develop predictive models for the identification of such individuals.

A. Rationale for the Study

The rationale for this study was to develop an accurate predictive model to identify individuals without health insurance due to unemployment. This focus is essential and would be beneficial to insurers, collection agencies, and policymakers for effective policy interventions. Predictors will help those agencies in terms of planning, decision making, and monitoring and evaluation of the target population, individuals without health insurance due to unemployment.

B. Objectives of the Study

The objectives of the study were threefold:

- To quantify the strength and direction of pairwise association among risk (demographic and health-related) factors.
- To identify significant risk (demographic and health-related) factors that contribute to the outcome.
- To develop and evaluate the performance of predictive models based on the identified significant risk factors.

Material and Methods

A. Participants

The study sample comprised 2,496 individuals in the United States who lacked general health insurance coverage between 2020 and 2023. This sample size was obtained following a structured data cleaning process. After restricting the dataset to the variables of interest and removing duplicate records, 3,566 unique observations remained. Next, 314 observations with missing information were excluded, leaving 3,252 observations (91.2%). The missing data were minimal and did not show any clear pattern related to the study outcomes or key predictors. Given the small proportion of missingness, a complete-case approach was adopted to maintain consistency across analyses without meaningfully affecting the sample size. We then examined unusually extreme values to ensure accuracy in the dataset. Rather than removing all outliers automatically, each case was reviewed carefully. Only those observations that appeared to reflect likely data-entry errors were excluded. This step resulted in the

removal of 756 additional observations, producing a final sample of 2,496 individuals, representing approximately 70.0% of the post-selection dataset.

B. Data and Variables

The study utilized microdata records from the National Health Interview Survey available in the Integrated Public Use Microdata Series (IPUMS) [20]. IPUMS is a widely recognized repository of harmonized census and survey data from around the world integrated across time and space. The response variable was HINOUNEMPR, a binary variable indicating whether an individual had no health insurance (uninsured) due to unemployment (1 = Yes, 0 = No). The predictor variables included a range of risk factors related to demographic, socioeconomic, and health-related characteristics. These variables were selected based on their relevance to health insurance coverage (Table 1).

Table 1. Variable Definition

Code	Definition	Type
REGION	Region where the individual resides	Categorical
AGE	Age in years	Continuous
SEX	Sex of the individual	Categorical
MARSTCUR	Current marital status	Categorical
RACENEW	Race/ethnicity	Categorical
EDUC	Education level	Categorical
HEALTH	Self-reported health	Categorical
BMICAT	BMI category	Categorical
USUALPL	Usual place for care	Categorical
HINOUNEMPR	Reason for no insurance: Unemployment	Categorical
SMOKFREQNOW	Current smoking frequency	Categorical

C. Statistical Analyses

Descriptive statistics, including frequencies, percentages, and quartiles, were computed to provide an overview of the demographic, socioeconomic, and health-related characteristics of the study participants. These statistics outlined the general profile of the study population, revealing both the distribution and variability in key variables. Several inferential statistical tests were conducted to examine associations between the predictor and response variables. Pearson's Chi-Squared Test was used to assess associations between categorical variables (REGION, SEX, MARSTCUR, RACENEW, EDUC, BMICAT, USUALPL, and SMOKFREONOW) and the response variable. The Wilcoxon Rank Sum Test was applied to compare the distribution of AGE between uninsured individuals due to unemployment and those who were not, serving as a nonparametric alternative to the t-test for comparing two independent samples. In addition, Goodman and Kruskal's Tau was used to evaluate pairwise associations among the categorical predictors, where a value of zero signifies no association and a value of one indicates perfect association.

Before modeling, the cleaned data set was divided into a training set (approximately 70% of the data) and a testing set (approximately 30%). This partition allowed the model to be

trained in one subset and evaluated on an independent subset, thereby reducing the risk of overfitting. The training data exhibited class imbalance, with a disproportionately higher number of insured individuals compared to those uninsured. To address this, the Synthetic Minority Over-sampling Technique (SMOTE) [21] was employed to generate synthetic samples of the minority class, resulting in a more balanced dataset and improved model performance. Binary Logistic Regression and Probit models (classification methods), both incorporating SMOTE, were then applied to model the probability of being uninsured due to unemployment based on the predictor variables. All significance tests were conducted at the 0.05 level to ensure statistical robustness and to minimize the likelihood of Type I errors.

To determine the most appropriate classification method, model performance was evaluated using the testing set. A confusion matrix was constructed, and key evaluation metrics: accuracy, precision, recall, and F1 score were calculated to quantify and compare the predictive performance of the classification methods. All statistical analyses were performed using R version 4.3.0 and RStudio version 2023.06.1+524.

Results and Discussions

Table 2 shows a univariate analysis of demographic factors for a sample size of N =2,496. Most of the participants were from the South (54.21%), followed by the North Central Midwest (20.91%), West (16.87%), and Northeast (8.01%). This indicates a significant regional concentration in the South. The median age of the participants was 43 years, with an interquartile range (IQR) of 33 to 55 years, which suggests a relatively wide age distribution. There was a higher proportion of males (63.70%) compared to females (36.30%). A significant portion of the participants were under the "Other" category (41.15%), which included divorced, separated, or widowed individuals. Never-married individuals made up 31.33%, and married individuals accounted for 27.52%. The participants were 81.01% White, 12.06% Black, and 6.93% other races. Additionally, their educational backgrounds varied, with the largest group having a high school education (40.71%). The others were higher education (25.20%), some college (17.75%), and less than a high school education (16.35%).

Table 2. Summary Statistics of Demographic Factors

Characteristic	$\boldsymbol{N = 2,496}$[1]
REGION	
Northeast	200 (8.01%)
North Central Mid-West	522 (20.91%)
South	1,353 (54.21%)
West	421 (16.87%)
AGE	43 (33, 55)
SEX	
Female	906 (36.30%)
Male	1,590 (63.70%)
MARSTCUR	

Characteristic	$N = 2,496$[1]
Never Married	782 (31.33%)
Married	687 (27.52%)
Other	1,027 (41.15%)
RACENEW	
Black	301 (12.06%)
White	2,022 (81.01%)
Other	173 (6.93%)
EDUC	
Less Than High School	408 (16.35%)
High School	1,016 (40.71%)
Some College	443 (17.75%)
Higher Education	629 (25.20%)

Table 3 provides a univariate analysis of health-related factors. The majority of participants reported their health status as "Good" (34.46%) or "Very Good" (31.05%). A smaller proportion reported "Excellent" health (17.63%), while 16.87% reported "Fair/Poor" health. This indicates that most participants perceived their health as good or very good, with few reporting less favorable health status. The distribution across BMI categories is relatively even, with 33.97% classified as "normal weight", 33.73% as "overweight", and 32.29% as "Obese". Most participants (61.66%) had USUALPL, while 38.34% did not. This could have implications for access to and continuity of care. Also, only 4.73% of participants had a history of unemployment, with the vast majority (95.27%) reporting no such history. Smoking habits were divided, with 44.67% reporting smoking every day, 43.07% not smoking at all, and 12.26% smoking some days. This highlights a significant portion of daily smokers, which could be a critical factor in health outcomes.

Table 3. Summary Statistics of Health-Related Factors

Characteristic	$N = 2,496$[1]
HEALTH	
Fair Poor	421 (16.87%)
Excellent	440 (17.63%)
Very Good	775 (31.05%)
Good	860 (34.46%)
BMICAT	
Normal Weight	848 (33.97%)
Overweight	842 (33.73%)
Obese	806 (32.29%)
USUALPL	
1	1,539 (61.66%)
0	957 (38.34%)
HINOUNEMPR	
1	118 (4.73%)
0	2,378 (95.27%)
SMOKFREONOW	
Not at all	1,075 (43.07%)
Some Days	306 (12.26%)
Every Day	1,115 (44.67%)

[1] n (%)

Table 4 provides the results of the Pearson's chi-square and the Wilcoxon rank-sum tests. For Pearson's chi-square, REGION, SEX, MARSTCUR, RACENEW, HEALTH, BMICAT, USUALPL, and SMOKFREONOW were each not statistically associated with HINOUNEMPR. Additionally, AGE, based on the Wilcoxon rank-sum test, was found not to be statistically significant between the two groups HINOUNEMPR. All tests were performed at the 0.05 level of significance. Further multivariate analysis could help better understand the relationships and potential confounding factors.

Table 4. Descriptive and Inferential Analysis by HINOUNEMPR

Characteristic	**1 (*Yes*)** $N = 18$[1]	**0 (*No*)** $N = 2,378$	***p*-value**[2]
REGION			0.2469
• Northeast	8 (6.78%)	192 (8.07%)	
• North Central Midwest	31 (26.27%)	491 (20.65%)	
• South	55 (46.61%)	1,298 (54.58%)	
• West	24 (20.34%)	397 (16.69%)	
AGE	41 (33, 51)	43 (33, 53)	0.1568
SEX			0.3107
• Female	48 (40.68%)	858 (36.08%)	
• Male	70 (59.32%)	1,520 (63.92%)	
MARSTCUR			0.0808
• Never Married	48 (40.68%)	734 (30.87%)	
• Married	28 (23.73%)	659 (27.71%)	
• Other	42 (35.59%)	985 (41.42%)	
RACENEW			0.7622
• Black	15 (12.71%)	286 (12.03%)	
• White	93 (78.81%)	1,929 (81.12%)	
• Other	10 (8.47%)	163 (6.85%)	
EDUC			0.0077
• Less than High School	10 (8.47%)	398 (16.74%)	
• High School	41 (34.75%)	975 (41.00%)	
• Some College	25 (21.19%)	418 (17.58%)	
• Higher Education	42 (35.59%)	587 (24.68%)	
HEALTH			0.6066
• Fair/Poor	18 (15.25%)	403 (16.95%)	
• Excellent	25 (21.19%)	415 (17.45%)	
• Very Good	32 (27.12%)	743 (31.24%)	
• Good	43 (36.44%)	817 (34.36%)	
BMICAT			0.4929
• Normal Weight	37 (31.36%)	811 (34.10%)	
• Overweight	37 (31.36%)	805 (33.85%)	
• Obese	44 (37.29%)	762 (32.04%)	
USUALPL			0.1098
• Yes	81 (68.64%)	1,458 (61.31%)	
• No	37 (31.36%)	920 (38.69%)	
SMOKFREONOW			0.8758
• Not at all	53 (44.92%)	1,022 (42.98%)	
• Some Days	15 (12.71%)	291 (12.24%)	
• Every Day	50 (42.37%)	1,065 (44.79%)	

Fig. 1 and Fig. 2 represent Goodman and Kruskal's Tau, which measures the association between two categorical variables within the overall data and the oversampled training set, respectively. Zero indicates no association, and one indicates perfect association. The higher the value, the stronger the association. Each variable has multiple values, which correspond to different levels. For example, regions have K=4, meaning there are four levels (Northeast, North Central, Midwest, South, West). There were no associations among each pair of categorical risk factors based on the Goodman-Kruskal tau metric (Tau values ≤ 0.03). This implies that the risk factors examined (predictor variables) were not or were weakly associated with one another.

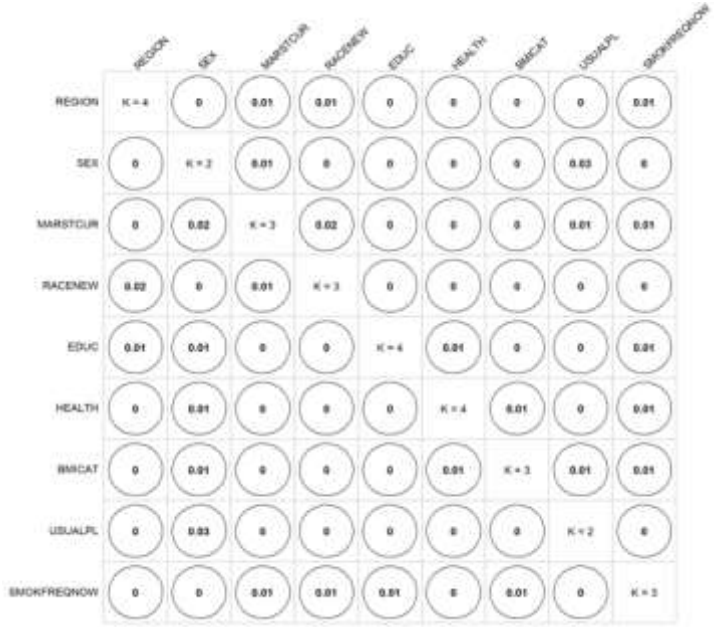


Fig. 1. Goodman & Kruskal's Tau Measure (Overall Data)

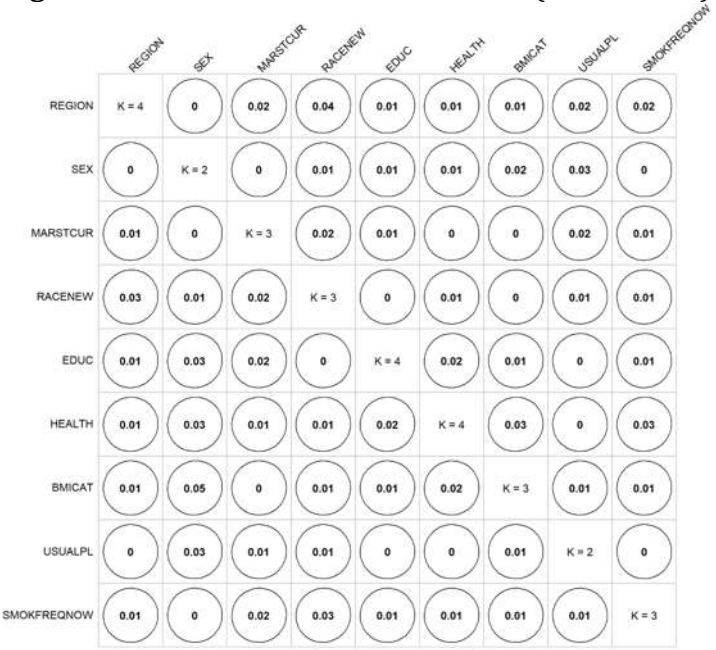


Fig. 2. Goodman & Kruskal's Tau Measure (Oversampled Training Set)

Fig. 3 shows the p-values for the risk factors in both the Logit (Logistic regression) and the Probit models based on the oversampled training set. Apart from RACENEW and HEALTH (Good and Very Good), other risk factors (REGION, SEX, MARSTCUR, EDUC, BMICAT, USUALPL, SMOKEFREQNOW) were identified as statistically significant at the 0.05 level by both the Logit and the Probit models.

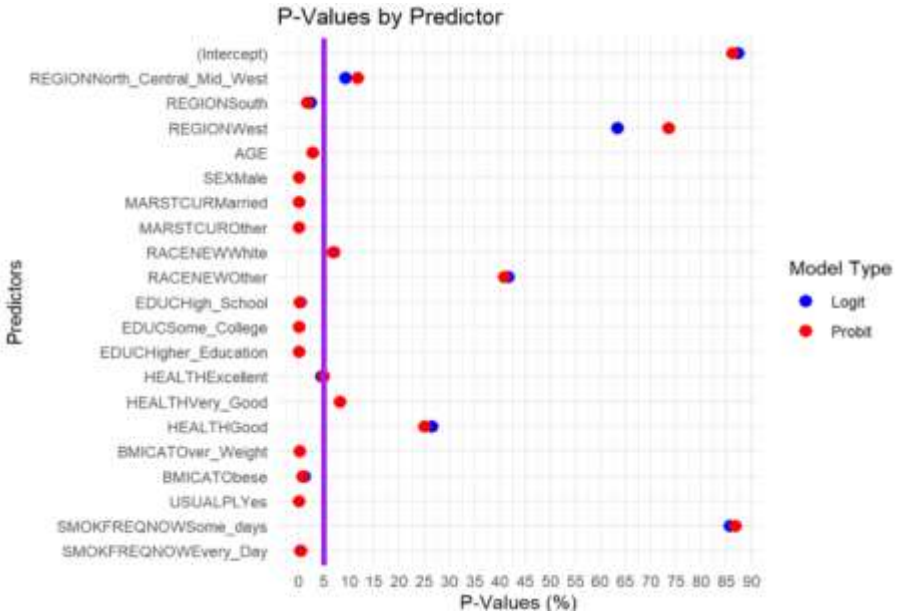


Fig. 3. *P*-values: Logit & Probit Models (Oversampled Training Dataset)

Fig. 4 shows the estimated odds ratios for various risk factors in both the Logit and the Probit models. An odds ratio (OR) greater than 1 indicates that the risk factor is associated with higher odds of the outcome, while an OR less than 1 means lower odds. If the OR is 1, there's no effect. The Logit and Probit models identified obesity, age, lower education, and lack of healthcare access as critical (statistically significant) risk factors.

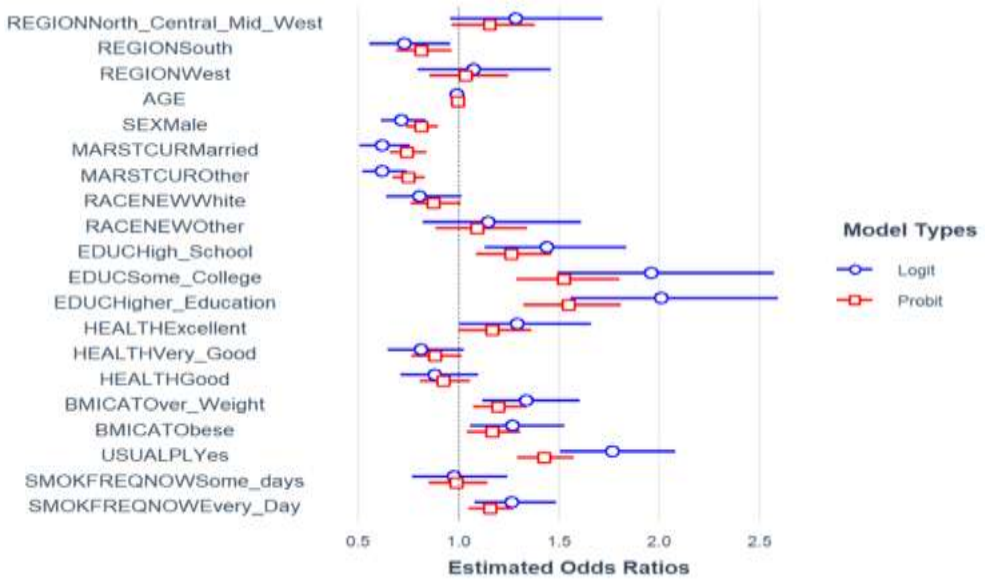


Fig. 4. Estimated Odds Ratios: Logit & Probit Models (Oversampled Training Dataset)

The prediction performance metrics for both Logit and Probit models were nearly identical, as shown in Table 5. Both models demonstrated high and comparable accuracy, with Probit marginally outperforming the Logit model by 0.14%.

Table 5. Out-of-Sample Prediction Performance Metrics

Metric	Logit (%)	Probit (%)
Accuracy	81.30	81.44
Recall	83.90	84.02
Precision	95.90	95.93
F1 Score	89.20	89.35

Discussion

The results of the current study showed that slightly over two-thirds of the participants were either overweight or obese. Research has identified obesity as a risk factor for exiting paid employment [22], as it is significantly associated with higher rates of unemployment and disability [23]. Another finding of the study was that lower education was a significant predictor of lack of health insurance due to unemployment. For instance, the prevalence of diabetes was found to be higher among individuals with less than a high school education than among those with a high school education [24].

The results identified that age is a major risk factor for lack of health insurance because of unemployment. This finding is supported by Virgolino et al. [13], who reported that men and young adults were the most at risk of being unemployed. Stroud et al. [16] recommended that young adults aged 18- 26 years be treated as a distinct subgroup in health planning and programming, as little attention is paid to this group once they graduate from pediatric to adult medical care.

Lack of healthcare access was found to be another critical risk factor in identifying individuals without healthcare insurance due to unemployment. This is consistent with Rinne et al. [25], who reported no employment status differences for the need for services from doctors and nurses. However, the unemployed had higher odds of having unmet healthcare needs when health-related factors were taken into consideration. Unemployment must be considered a public health issue, as limited access to healthcare could result in physical and mental health crises [26].

The demographic and health-related factors in the current study had limited standalone predictive value. That is, the factors in this were not statistically significant when assessed individually. However, when studied within the framework of logit and probit models, the researchers observed statistical significance as well as high predictive performance. Therefore, for accurate prediction of the demographic and health-related factors, practitioners should consider them jointly rather than in isolation. The predictive performances of the two models were similar. However, the probit model had a slightly higher predictive ability compared to the logit model. The implication is that the probit model had better potential than

the logit to assist in identifying individuals with no insurance due to unemployment. The contribution to existing modeling is for practitioners to use these variables (Demographic and health-related factors) in combination rather than in isolation. Another contribution made by this paper is that there was a slight difference in the predictive performances of the two models. This has the potential of guiding practitioners in making informed decisions as to which of the two models to use.

A major limitation of the study is that potential confounders were not accounted for. Some common confounding variables could be the type of employment lost, the length of unemployment, and whether the previous insurance was employer-related or through private insurance. In addition, the definition of studied variables, such as Region and Health Status, could potentially affect the results of the study. For example, there could be variability within regions that were not accounted for in the present study. Moreover, Health Status as a variable in this study was measured subjectively through self-reporting.

Conclusion

Most participants in the present study reported being in good, very good, or excellent health. The demographic and health-related factors in this study had limited standalone predictive values. However, Probit and Logit models identified obesity, age, lower education, and lack of healthcare access as critical risk factors for determining individuals without health insurance due to unemployment. Furthermore, the prediction performance metrics for both Logit and Probit models were identical.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] C. DeNavas-Walt, *Income, Poverty, and Health Insurance Coverage in the United States*. Washington, DC, USA: Diane Publishing, 2010.

[2] M. Majerol, V. Newkirk, and R. Garfield, "The uninsured: A primer," Kaiser Family Foundation Publication, no. 7451-10, 2015. [Online]. Available: Kaiser Family Foundation

[3] S. Dorn, "The relationship between unemployment and health insurance coverage: Findings from the Current Population Survey," Urban Institute, 2020. [Online]. Available: Urban Institute Report. Accessed: Nov. 25, 2025.

[4] Kaiser Family Foundation, "The uninsured and the difference health insurance makes," 2023. [Online]. Available: KFF Uninsured Resources. Accessed: Dec. 17, 2025.

[5] National Center for Health Statistics, "National Health Interview Survey," 2023. [Online]. Available: CDC NHIS. Accessed: May 9, 2025.

[6] G. Claxton, R. Mathew, D. Anthony, Y. Gregory, M. Daniel, and H. Whitmore, "Health benefits in 2019: Premiums inch higher, employers respond to federal policy," *Health Aff.*, vol. 38, no. 10, pp. 1752–1761, 2019, doi: 10.1377/hlthaff.2019.01026.

[7] A. Nguyen, A. Guttentag, D. Li, and J. V. Meijgaard, "The impact of job and insurance loss on prescription drug use: A panel data approach to quantifying the health consequences of unemployment during the COVID-19 pandemic," *Int. J. Health Serv.*, vol. 52, no. 3, pp. 312–322, 2022, doi: 10.1177/00207314221078749.

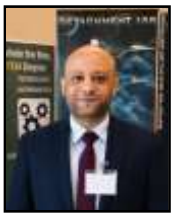
[8] T. Marsh and S. Kim, "Drug prices are growing faster than any other commodity or service," GoodRx Research, 2020. [Online]. Available: GoodRx Research Report

[9] P. Pratap, A. Dickson, M. Love, J. Zanoni, C. Donato, M. A. Flynn, and P. A. Schulte, "Public health impacts of underemployment and unemployment in the United States: Exploring perceptions, gaps and opportunities," *Int. J. Environ. Res. Public Health*, vol. 18, no. 19, p. 10021, 2021, doi: 10.3390/ijerph181910021.
[10] E. Kwon, S. Park, and T. D. McBride, "Effects of the Affordable Care Act on health insurance coverage among middle-aged adults," *Int. J. Health Serv.*, vol. 49, no. 4, pp. 712–732, 2019, doi: 10.1177/00207314198651.
[11] R. E. McGee and N. J. Thompson, "Unemployment and depression among emerging adults in 12 States, Behavioral Risk Factor Surveillance System 2010," *Prev. Chronic Dis.*, vol. 12, p. E38, 2015, doi: 10.5888/pcd12.140451.
[12] H. Wilson and D. Finch, "Unemployment and mental health: Why both require action for our COVID-19 recovery," The Health Foundation, 2021.
[13] A. Virgolino et al., "Lost in transition: A systematic review of the association between unemployment and mental health," *J. Ment. Health*, vol. 31, no. 3, pp. 432–444, 2022, doi: 10.1080/09638237.2021.2022615.
[14] K. Wang and B. D. Smith, "A county-level analysis of unemployment and mental health in the Southeastern US: Applying a social determinants of health framework," *Soc. Work Ment. Health*, vol. 20, no. 5, pp. 559–577, 2022, doi: 10.1080/15332985.2022.2038761.
[15] P. R. Shafer, D. M. Anderson, R. Whitaker, C. A. Wong, and B. Wright, "Association of unemployment with Medicaid enrollment by social vulnerability in North Carolina during COVID-19," *Health Aff.*, vol. 40, no. 9, pp. 1491–1500, 2021, doi: 10.1377/hlthaff.2021.00377.
[16] C. Stroud, L. R. Walker, M. Davis, and C. E. Irwin Jr., "Investing in the health and well-being of young adults," *J. Adolesc. Health*, vol. 56, no. 2, pp. 127–129, 2015, doi: 10.1016/j.jadohealth.2014.11.012.
[17] D. Maupin, J. O. Lee, A. Kleinman, W. J. Lee, and H. Jin, "The impact of unemployment benefits on the mental health of underemployed workers: Findings from the Understanding America Study," *medRxiv*, pp. 2025–10, 2025, doi: 10.1101/2025.10.16.25338149.
[18] J. Cylus, M. M. Glymour, and M. Avendano, "Health effects of unemployment benefit program generosity," *Am. J. Public Health*, vol. 105, no. 2, pp. 317–323, 2015, doi: 10.2105/AJPH.2014.302253.
[19] A. J. Baxter, E. J. Tweed, S. V. Katikireddi, and H. Thomson, "Effects of housing first approaches on health and well-being of adults who are homeless or at risk of homelessness: Systematic review and meta-analysis of randomised controlled trials," *J. Epidemiol. Community Health*, vol. 73, no. 5, pp. 379–387, 2019, doi: 10.1136/jech-2018-210981.
[20] L. A. Blewett, J. A. R. Drew, M. L. King, K. C. W. Williams, A. Chen, S. Richards, and M. Westberry, "IPUMS Health Surveys: National Health Interview Survey," Version 7.3, Minneapolis, MN, USA: IPUMS, 2023, doi: 10.18128/D070.V7.3.
[21] R. Vaughan, "Oversampling in health surveys: Why, when, and how?" *Am. J. Public Health*, vol. 107, no. 8, pp. 1214–1215, 2017, doi: 10.2105/AJPH.2017.303895.
[22] S. J. W. Robroek et al., "The contribution of overweight, obesity, and lack of physical activity to exit from paid employment: A meta-analysis," *Scand. J. Work Environ. Health*, vol. 39, no. 3, pp. 233–240, 2013. [Online]. Available: JSTOR Article
[23] Y. Le Strat et al., "The role of comorbidity in the association of obesity with unemployment and disability," *Ann. Epidemiol.*, vol. 45, pp. 61–68, 2020, doi: 10.1016/j.annepidem.2020.03.004.
[24] Centers for Disease Control and Prevention, *National Diabetes Statistics Report, 2020*. Atlanta, GA, USA: Centers for Disease Control and Prevention, 2020. [Online]. Available: CDC Diabetes Report
[25] H. Rinne et al., "Unmet need for healthcare services among the unemployed," *Eur. J. Public Health*, vol. 35, suppl. 4, p. ckaf161-1572, 2025, doi: 10.1093/eurpub/ckaf161.1572.
[26] B. Ilić, S. L. Fičko, A. M. Hošnjak, I. Kovačević, M. Smrekar, and B. Sedić, "Unemployment and access to health care," *Croat. Nurs. J.*, vol. 3, pp. 183–192, 2019, doi: 10.24141/2/3/2/6.

Authors



Emmanuel Thompson (PhD, AASG) is a professor of actuarial science and statistics in the Department of Mathematics at Southeast Missouri State University, USA. His academic preparation is in Statistics and Actuarial Science. His research interest is mainly inspired by applications, especially those from the financial, insurance, and health services industries. His research has been on actuarial risk and mortality modeling with its applications. Currently, his research interests are centered on the applications of modern statistical methods and data science techniques to high-dimensional, messy, and massive data sets. (email: ethompson@semo.edu).



Mohamed Abu Sheha (PhD) is an Assistant Professor of Statistics in the Department of Mathematics at Southeast Missouri State University. His academic background is in Statistics, with expertise in Statistical Learning, Causal Inference, and Bayesian Methods. He develops advanced analytical and AI-driven models to address critical challenges in areas like health science and climate. Currently, he focuses on modeling the cognitive and brain changes associated with aging and Alzheimer's disease. His broader research interests span data science applications across health, environmental science, and engineering. (email: mabusheha@semo.edu).



Seidu Sofo is a professor of physical education pedagogy in the Department of Allied Health, Kinesiology, and Sport Sciences at Southeast Missouri State University, USA. His academic preparation is in Physical Education Pedagogy concentration, with cognates in Health Education and Educational Research. His research interests are in early childhood/elementary (primary) physical education, physical activity, perceptual-motor development, and decolonizing pedagogy. (email: ssofo@semo.edu).