BPR MCS | Bulletin of Pioneering Researches of Medical and Clinical Science

Bulletin of Pioneering Researches of Medical and Clinical Science

Available online: https://bprmcs.com
2021 | Volume 1 | Issue 1 | Page: 82-93

Design and Validation of a Model to Forecast Depression Risk Among Cancer Survivors Using NHANES Data from 2005–2018

John M. Randall[1]*, Emily S. Carter[1], Ahmed S. Farouk[1]

[1]Department of Medical and Clinical Sciences, Harvard Medical School, Boston, United States.

Abstract

Depression represents a major concern for cancer survivors, and early detection of depressive symptomatology is essential for timely intervention. This work set out to construct and validate a model capable of estimating depression risk in cancer survivors. A total of 2,279 cancer survivors from the National Health and Nutrition Examination Survey (NHANES) were analyzed. Participants were randomly divided into training and validation cohorts at a 7:3 ratio. Independent determinants of depression—defined as a PHQ-9 score ≥ 10—were identified using least absolute shrinkage and selection operator regression followed by multivariable logistic regression, and these factors were incorporated into a nomogram. Model discrimination was evaluated through receiver operating characteristic analysis, and its reliability was examined using calibration plots, the Hosmer–Lemeshow goodness-of-fit test, and decision curve analysis. Seven predictors were ultimately retained: age, education, poverty-to-income ratio, smoking behavior, congestive heart failure, sleep disorders, and number of cancers. A nomogram based on these variables was created. The area under the curve in the training cohort reached 0.802 (95% CI: 0.767–0.836), while the validation cohort yielded 0.794 (95% CI: 0.740–0.849). Bootstrapped internal validation generated an optimism-adjusted AUC of 0.812 (95% CI: 0.784–0.840). Calibration assessments confirmed strong agreement between predicted and observed risks, and decision curve analysis demonstrated meaningful clinical benefit. Collectively, this study presents a nomogram capable of estimating depression risk among cancer survivors, offering a potentially valuable tool for recognizing individuals who may require further psychological evaluation.

Keywords: Predictive model, Cancer survivors, Nomogram, Depression

Corresponding author: John M. Randall
E-mail: john.randall@gmail.com

How to Cite This Article: Randall JM, Carter ES, Farouk AS. Design and Validation of a Model to Forecast Depression Risk Among Cancer Survivors Using NHANES Data from 2005–2018. Bull Pioneer Res Med Clin Sci. 2021;1(1):82-93. https://doi.org/10.51847/8iEuovDHfT

Introduction

With continual progress in cancer screening and therapeutic modalities, the number of cancer survivors has risen substantially over recent decades [1]. Although extended survival is increasingly achievable, individuals who have lived through cancer frequently continue to experience a wide array of adverse consequences. Persistent physical symptoms, treatment-related complications, and disruptions to daily functioning—including work, interpersonal relationships, and lifestyle habits—remain common after the completion of therapy [2]. As a result, depression has emerged as a pervasive concern within the cancer survivor population [3]. Compared with adults in the general community, people with cancer are estimated to exhibit nearly double the

prevalence of major depressive disorder [4]. A meta-analysis reported that in non-palliative contexts, approximately 16.3% of cancer patients experience depression (95% CI: 13.4–19.5%) [5]. Importantly, depressive symptoms may arise at any point along the survivorship continuum [6].

The coexistence of depression and cancer imposes wide-ranging negative consequences on survivors' health outcomes and overall functioning. Depressive symptoms diminish quality of life, hinder adherence to treatment plans and follow-up care, and may even accelerate disease progression [7]. Depression is also associated with elevated inflammatory activity and impaired immune responses [8]. Moreover, depression substantially heightens the risk of suicidal ideation [9]. Over time, it has been recognized as an independent predictor of both cancer-specific and all-cause mortality [10]. Given these far-reaching implications, timely recognition of depression in cancer survivors is essential.

However, depressive symptoms in cancer survivors are often overlooked [11]. Emotional distress may be dismissed as an expected reaction to diagnosis or masked by somatic manifestations such as fatigue. Additionally, the determinants of depression in this population are multifaceted, encompassing sociodemographic factors like age, sex, race, educational level, and marital status [12], as well as behavioral characteristics, comorbid physical illnesses, and cancer-related features such as tumor type [13]. Thus, a practical and accurate tool that can help clinicians anticipate depression risk among cancer survivors is of substantial value.

A nomogram provides a visual mechanism for transforming complex statistical models into intuitive prognostic tools, enabling clinicians to estimate individualized risk and support decision-making [14]. Although machine-learning models may offer improved predictive performance, their limited interpretability often restricts clinical implementation. Nomograms, by contrast, supply transparent and user-friendly representations of predictive relationships. Previous efforts have produced nomograms aimed at forecasting depression among individuals with colorectal [15], breast [16, 17], and lung cancers [18]. These tools, however, primarily target single cancer types and frequently rely on inpatient populations, potentially limiting applicability to broader groups of cancer survivors living in the community. Research addressing depression risk across diverse cancer types remains scarce. A recent investigation by Zuo and Yang [19] introduced a dynamic nomogram constructed from four factors using NHANES data, but that model did not integrate several clinically meaningful variables, including comorbid conditions and cancer-related attributes such as the number of tumors.

In line with earlier findings, we posited that a concise set of readily accessible demographic, clinical, and lifestyle indicators could be synthesized into an effective model for detecting depression risk among cancer survivors. Drawing on NHANES—a nationally representative, large-scale cross-sectional survey conducted in the United States—we sought to broaden the range of evaluated factors and develop a nomogram capable of estimating depression risk in adult cancer survivors. The primary aim of this study was to construct a more comprehensive predictive tool to facilitate the identification and appropriate management of survivors at heightened risk for depression.

Materials and Methods

Study design and data source

This study employed a secondary analysis of data drawn from the cross-sectional National Health and Nutrition Examination Survey (NHANES). NHANES is a nationwide program in the United States conducted in biennial cycles, designed to generate health and nutritional information that reflects the demographic composition of the national population. For this investigation, the combined NHANES cycles from 2005 to 2018—spanning seven 2-year waves—served as the analytical foundation. NHANES is publicly accessible, and all procedures were approved by the Ethics Review Committee of the National Center for Health Statistics (NCHS). Written informed consent was obtained from all study participants, and all analyses adhered to the ethical standards outlined in the Declaration of Helsinki.

Participants

Cancer status was determined using self-reported responses from the NHANES "Medical Conditions Questionnaire." Participants were asked, "Have you ever been told by a doctor or other health professional that you had cancer or a malignancy of any kind?" Individuals responding "yes" were classified as cancer survivors. Eligible participants were required to be at least 20 years old and have a documented history of cancer or malignancy. Exclusion criteria consisted of age under 20 years, absence of a cancer diagnosis, and incomplete information regarding PHQ-9 scores or other essential covariates. From an initial pool of 70,190 survey entries, 3782 adults reported having cancer. After removing cases with incomplete or invalid data, the final analytic sample comprised 2279 cancer survivors.

Assessment of depression

Depression severity was evaluated using the Patient Health Questionnaire-9 (PHQ-9), a validated instrument composed of nine symptom-based items. The

questionnaire assesses domains such as anhedonia, low mood, sleep disturbances, fatigue, appetite change, negative self-perception, impaired concentration, psychomotor changes, and suicidal ideation. Each item is rated from 0 (not at all) to 3 (nearly every day), producing a composite score ranging from 0 to 27. Higher totals indicate greater depressive severity. Consistent with prior research, PHQ-9 scores ≥10 were treated as indicative of clinically relevant depression [20].

Data selection and interpretation of predictors

The analysis incorporated a broad set of sociodemographic, behavioral, and clinical variables previously linked to depression and cancer outcomes. Demographic indicators included age, sex, racial/ethnic classification, educational level, and marital status. Age was grouped into 20–44, 45–64, and ≥65 years, aligning with standard NCHS life-stage categories [21]. Socioeconomic position was measured using the poverty-to-income ratio (PIR), which was divided into <1.3, 1.3–3.5, and >3.5, representing decreasing economic security [22]. Lifestyle measures consisted of smoking behavior, alcohol consumption, sleep duration, and body mass index. Comorbid conditions were identified from self-reports on the "Medical Conditions Questionnaire," with affirmative responses indicating the presence of each illness.

Cancer-related information was also obtained from self-report items. The number of cancers an individual had experienced (one, two, or multiple) and the reported cancer type were recorded. The initial cancer reported was treated as the primary cancer and classified according to anatomical system categories [23]: non-melanoma skin, genitourinary, breast, gynecologic, digestive/gastrointestinal, skin unspecified, and melanoma. Less common cancers were aggregated into an "other" group. Information on age at diagnosis and duration since diagnosis was computed using the time interval between the reported cancer onset and the examination date. Additionally, laboratory variables relevant to routine monitoring in cancer survivors—such as blood cell counts, liver function markers, renal indices, and lipid profiles—were extracted.

Statistical analysis

Survey weighting procedures were applied to ensure national representativeness. NHANES stratification (SDMVSTRA), clustering (SDMVPSU), and MEC examination weights (WTMEC2YR) were incorporated according to analytic guidelines. Continuous variables were summarized as weighted means ± standard deviations and compared using weighted independent-samples t-tests. Categorical variables were expressed as weighted frequencies or percentages, with differences assessed using the Pearson chi-square test.

Participants were randomly assigned to training and validation sets in a 7:3 ratio for model development. Variable reduction was performed using least absolute shrinkage and selection operator (LASSO) regression, after which weighted multivariable logistic regression was used to determine independent predictors of depression. A nomogram for estimating depression risk among cancer survivors was constructed using the final predictors. Robustness was evaluated through sensitivity analyses involving multiple imputation (m = 20) using the mice R package. For each imputed dataset, LASSO and multivariable logistic regression were repeated, and results were combined using Rubin's rules.

Discriminative performance in the training and validation sets was quantified using the area under the receiver operating characteristic curve (AUC). Internal validation relied on 1000-sample bootstrapping to derive optimism-corrected performance metrics. Calibration was assessed with calibration plots comparing predicted and observed probabilities; proximity to the ideal 45-degree reference line indicated accurate prediction. The Hosmer–Lemeshow goodness-of-fit test provided additional evaluation of calibration. Decision curve analysis (DCA) was conducted to examine the potential clinical utility of the model across a range of threshold probabilities. Statistical significance was defined as $P < .05$. All analyses were conducted with R version 4.2.1 (R Foundation for Statistical Computing, Vienna, Austria).

Results and Discussion

Baseline characteristics

The final sample included 2279 cancer survivors, among whom 241 individuals (10.6%) met the PHQ-9 criteria for depression and 2038 (89.4%) did not, yielding a weighted depression prevalence of 9.0%. Compared with their non-depressed counterparts, cancer survivors with depression tended to be middle-aged, predominantly female, have lower educational attainment, and exhibit reduced PIR values. Depressed individuals also showed higher frequencies of coronary heart disease, stroke, asthma, arthritis, congestive heart failure, angina, and chronic bronchitis. Additionally, shorter sleep duration, elevated rates of sleep disorders, and higher smoking prevalence were observed. Significant differences between groups were likewise identified for race, age at cancer diagnosis, cancer type, HDL levels, blood urea nitrogen, and triglycerides. Weighted descriptive statistics are provided in **Table 1**.

The dataset was randomly partitioned into a training set (n = 1594) and a validation set (n = 685) in a 7:3 ratio. Comparative analyses demonstrated that the two subsets

did not differ significantly across baseline variables (P > .05), supporting their suitability for model construction and evaluation.

Table 1. Characteristics of cancer survivors with and without depression included in the study

Variables	Without Depression*	With Depression*	Overall*	P-value†
Total	N = 2,038 (91.0%)	N = 241 (9.0%)	N = 2,279 (100.0%)	
Age (years)				**<.001**
20–44	184 (11.5%)	45 (20.0%)	229 (12.3%)	
45–64	600 (36.8%)	107 (53.1%)	707 (38.2%)	
≥65	1,254 (51.7%)	89 (27.0%)	1,343 (49.5%)	
Gender				**.031**
Male	989 (44.1%)	80 (33.6%)	1,069 (43.1%)	
Female	1,049 (55.9%)	161 (66.4%)	1,210 (56.9%)	
Race/Ethnicity				**.001**
Non-Hispanic White	1,482 (88.8%)	152 (80.0%)	1,634 (88.0%)	
Non-Hispanic Black	270 (4.5%)	29 (6.1%)	299 (4.6%)	
Mexican American	100 (1.7%)	32 (4.8%)	132 (2.0%)	
Other Hispanic	117 (2.2%)	14 (2.1%)	131 (2.1%)	
Other (including multi-racial)	69 (2.9%)	14 (7.0%)	83 (3.3%)	
Education level				**<.001**
Less than 9th grade	156 (3.8%)	43 (9.3%)	199 (4.3%)	
9–11th grade	219 (7.3%)	39 (12.1%)	258 (7.7%)	
High school graduate/GED	458 (20.3%)	55 (26.4%)	513 (20.8%)	
Some college or associate degree	611 (32.0%)	74 (37.7%)	685 (32.5%)	
College graduate or higher	594 (36.7%)	30 (14.5%)	624 (34.7%)	
Poverty Income Ratio (PIR)				**<.001**
<1.3	411 (12.3%)	122 (37.6%)	533 (14.6%)	
1.3–3.5	841 (36.2%)	86 (40.4%)	927 (36.6%)	
≥3.5	786 (51.5%)	33 (22.1%)	819 (48.9%)	
Marital status				**<.001**
Married	1,225 (64.0%)	98 (48.8%)	1,323 (62.6%)	
Widowed	337 (13.4%)	41 (13.1%)	378 (13.3%)	
Divorced	241 (11.2%)	56 (21.3%)	297 (12.1%)	
Separated	55 (2.2%)	16 (6.5%)	71 (2.6%)	
Never married	112 (5.6%)	23 (7.2%)	135 (5.7%)	
Living with partner	68 (3.7%)	7 (3.2%)	75 (3.6%)	
Body Mass Index (BMI)				.10
Underweight (<18.5)	33 (1.7%)	3 (0.9%)	36 (1.7%)	
Normal (18.5–24.9)	534 (28.2%)	61 (24.3%)	595 (27.8%)	
Overweight (25.0–29.9)	735 (35.3%)	68 (30.2%)	803 (34.9%)	
Obese (≥30.0)	736 (34.8%)	109 (44.6%)	845 (35.6%)	
Comorbid conditions				

Diabetes	358 (14.5%) Yes	60 (16.5%) Yes	418 (14.6%) Yes	.4
Coronary heart disease	168 (6.6%) Yes	32 (11.9%) Yes	200 (7.1%) Yes	**.009**
Stroke	157 (5.3%) Yes	36 (11.8%) Yes	193 (5.9%) Yes	**<.001**
Asthma	311 (16.4%) Yes	61 (24.9%) Yes	372 (17.2%) Yes	**.029**
Arthritis	1,021 (48.0%) Yes	156 (62.0%) Yes	1,177 (49.3%) Yes	**.004**
Congestive heart failure	121 (4.9%) Yes	32 (11.1%) Yes	153 (5.5%) Yes	**<.001**
Angina	102 (3.8%) Yes	22 (9.0%) Yes	124 (4.3%) Yes	**.003**
Heart attack	174 (6.5%) Yes	31 (9.6%) Yes	205 (6.8%) Yes	.073
Chronic bronchitis	169 (9.2%) Yes	44 (17.7%) Yes	213 (10.0%) Yes	**.001**
Thyroid problems	390 (19.7%) Yes	52 (22.0%) Yes	442 (19.9%) Yes	.5
Sleep duration				**.001**
≤7 hours	1,129 (53.2%)	159 (67.2%)	1,288 (54.4%)	
7–9 hours	806 (42.6%)	67 (27.6%)	873 (41.2%)	
>9 hours	103 (4.2%)	15 (5.2%)	118 (4.3%)	
Sleep disorder	661 (35.3%) Yes	155 (70.7%) Yes	816 (38.5%) Yes	**<.001**
Smoking status				**<.001**
Current smoker	270 (13.8%)	85 (38.0%)	355 (15.9%)	
Former smoker	823 (39.0%)	79 (32.1%)	902 (38.4%)	
Never smoker	945 (47.3%)	77 (29.9%)	1,022 (45.7%)	
Alcohol consumption				.8
Current drinker	1,460 (76.3%)	168 (76.8%)	1,628 (76.3%)	
Former drinker	323 (13.4%)	43 (14.4%)	366 (13.5%)	
Never/lifetime abstainer	255 (10.3%)	30 (8.9%)	285 (10.1%)	
Hypertension	1,153 (50.6%) Yes	142 (53.4%) Yes	1,295 (50.9%) Yes	.5
Age at cancer diagnosis (years), mean (SD)	51.12 (16.87)	44.10 (16.49)	50.49 (16.96)	**<.001**
Years since cancer diagnosis, mean (SD)	11.69 (11.46)	11.92 (10.77)	11.71 (11.40)	.8
Number of cancers				.2
1	1,836 (89.6%)	212 (87.0%)	2,048 (89.4%)	
2	186 (9.4%)	23 (10.6%)	209 (9.5%)	
≥3	16 (1.0%)	6 (2.5%)	22 (1.1%)	
Cancer type				**<.001**
Non-melanoma skin	360 (23.4%)	27 (16.9%)	387 (22.8%)	
Genitourinary	357 (11.5%)	24 (8.4%)	381 (11.2%)	
Breast	313 (14.0%)	38 (13.9%)	351 (14.0%)	
Gynecological	244 (12.1%)	65 (27.4%)	309 (13.4%)	
Digestive/gastrointestinal	198 (7.7%)	32 (8.6%)	230 (7.8%)	
Skin (unknown type)	159 (9.3%)	17 (7.0%)	176 (9.1%)	
Melanoma	130 (7.8%)	11 (5.8%)	141 (7.6%)	
Other	277 (14.4%)	27 (12.0%)	304 (14.1%)	
Laboratory measures, mean (SD)				
Total cholesterol (mmol/L)	5.08 (1.15)	5.19 (1.23)	5.09 (1.15)	.3
HDL cholesterol (mmol/L)	1.44 (0.46)	1.31 (0.43)	1.42 (0.46)	**<.001**
Hemoglobin A1c (%)	5.75 (0.81)	5.88 (1.14)	5.76 (0.84)	.13

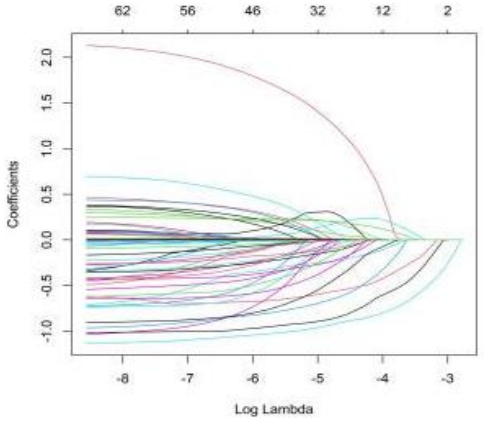
Albumin (g/L)	5.56 (2.32)	4.90 (2.40)	5.50 (2.34)	**.001**
Creatinine (μmol/L)	42.30 (3.17)	41.97 (3.64)	42.27 (3.21)	.3
Blood urea nitrogen (mg/dL)	83.76 (38. 24)	81.21 (37.82)	83.53 (38.20)	.4
Aspartate aminotransferase (U/L)	25.56 (11.38)	26.11 (16.00)	25.61 (11.86)	.6
Alanine aminotransferase (U/L)	23.73 (14.92)	25.80 (22.12)	23.91 (15.71)	.15
Triglycerides (mmol/L)	1.81 (1.76)	2.11 (1.53)	1.84 (1.74)	.053

*n (unweighted) (% (weighted)); mean (SD) or frequency (percentage).
†Pearson's X2: Rao & Scott adjustment; design-based t-test.

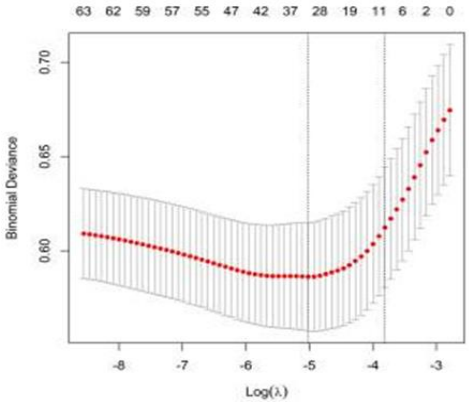
Predictor screening and nomogram construction

To identify the variables most relevant for model development, predictor reduction was carried out through LASSO regression using a 10-fold cross-validation procedure. The regularization parameter was selected according to the one-standard-error criterion (lambda.1se), ensuring a parsimonious and stable model. The initial pool consisted of 62 potential predictors spanning demographic features, socioeconomic indicators, chronic disease history, cancer-specific measures, and biochemical laboratory values. Following penalization, LASSO retained 11 variables with non-zero coefficients **(Figure 1)**. No additional subjective removal of borderline predictors was undertaken after the automated selection.

Subsequently, weighted multivariable logistic regression was used to validate the LASSO-derived variables. This analysis identified seven predictors with statistically significant independent associations with depression risk **(Table 2)**. These included being in the middle-age category, having lower educational attainment, reduced PIR, a history of congestive heart failure, the presence of sleep disorders, current smoking behavior, and having multiple cancer sites.



a)



b)

Figure 1. (a) Trajectory of coefficient estimates generated through the LASSO regression procedure. (b) Cross-validation plot illustrating the tuning of the LASSO penalty parameter.

Table 2. Results of multivariate logistic regression for variables identified by LASSO.

Variables	*P*-value	95% CI	OR
Age	**<.001**		
45–64		—	—
20–44		0.25, 1.07	0.52
≥65		0.23, 0.64	0.39
Education	**.013**		
<9th grade		—	—
9–11th grade		0.18, 1.06	0.43

High School graduate or equivalent		0.26, 1.50	0.62
Some college or AA degree		0.28, 1.54	0.65
College graduate or above		0.10, 0.67	0.26
PIR	**<.001**		
<1.3		—	—
1.3–3.5		0.32, 0.96	0.55
≥3.5		0.15, 0.51	0.27
Congestive heart failure	**<.001**		
Yes		—	—
No		0.15, 0.60	0.29
Sleep disorder	**<.001**		
Yes		—	—
No		0.15, 0.38	0.24
Smoking status	**.012**		
Current		—	—
Former		0.20, 0.85	0.42
Never		0.22, 0.77	0.41
Number of cancer	**.035**		
1		—	—
2		0.75, 3.61	1.65
Multiple		1.33, 34.3	6.76
Age at cancer diagnosis	.3	0.97, 1.01	0.99
Cancer type	.5		
Skin (non-melanoma)		—	—
Genitourinary		0.56, 3.30	1.36
Breast		0.43, 2.63	1.07
Gynecological		0.49, 2.89	1.19
Digestive/Gastrointestinal		0.38, 1.71	0.81
Skin (unknown kind)		0.38, 2.50	0.97
Melanoma		0.22, 2.66	0.77
Other		0.24, 1.71	0.64
Hemoglobin A1c	.2	0.92, 1.47	1.16
Aspartate aminotransferase (AST)	.10	1.00, 1.02	1.01

CI: confidence interval, LASSO: least absolute shrinkage and selection operator, OR: odds ratio, PIR: poverty-to-income ratio.

Findings from the multiple imputation sensitivity assessment further reinforced the stability of the complete-case results. The same 7 baseline predictors continued to show statistically meaningful associations, and both their estimated effects and corresponding confidence intervals remained virtually unchanged. The analysis did not reveal any new significant predictors. Using these 7 established predictors, we developed a nomogram that visually translates the model into an individualized risk-estimation tool for depression among cancer survivors, illustrated in **Figure 2**. In this diagram, each predictor contributes a defined number of points on the upper scoring scale; these individual point contributions are then summed. The cumulative total is matched to the total-points line, which in turn yields the estimated probability of depression on the final risk scale. To illustrate its application, consider a 70-year-old individual with primary-school education, a PIR of 1.0, no congestive heart failure, presence of a sleep disorder, no history of smoking, and a single cancer site. Their combined score would reach 202 points, corresponding to an estimated depression probability of roughly 48%.

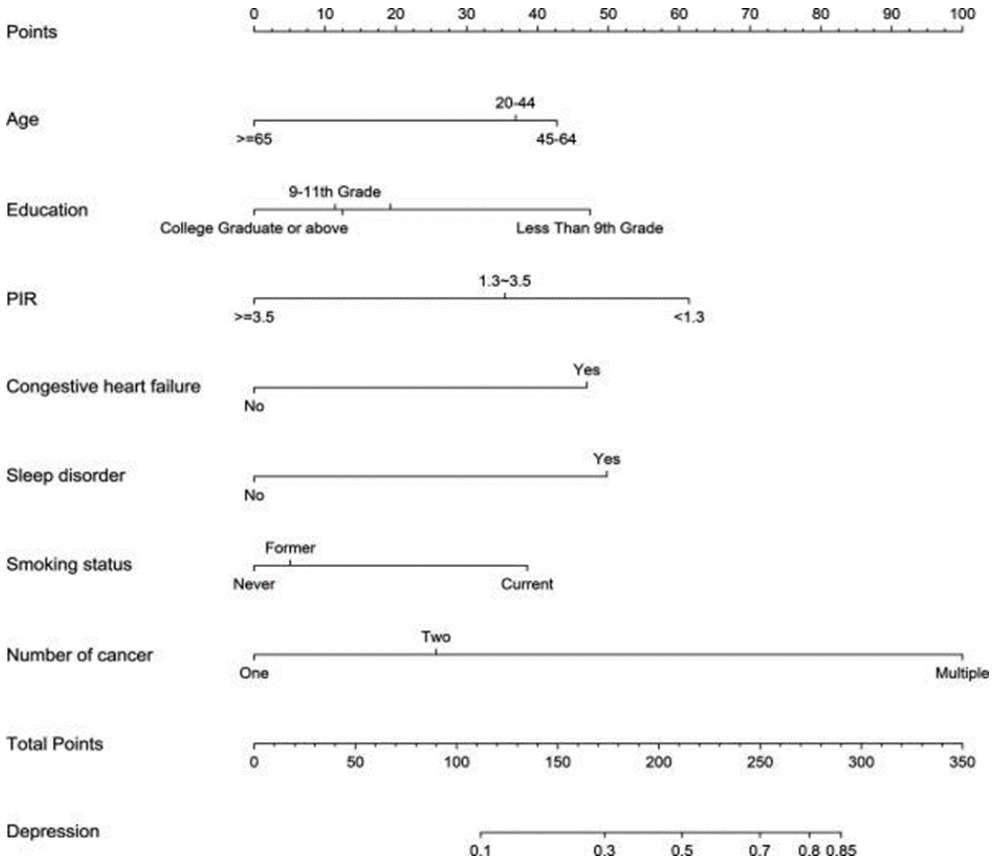
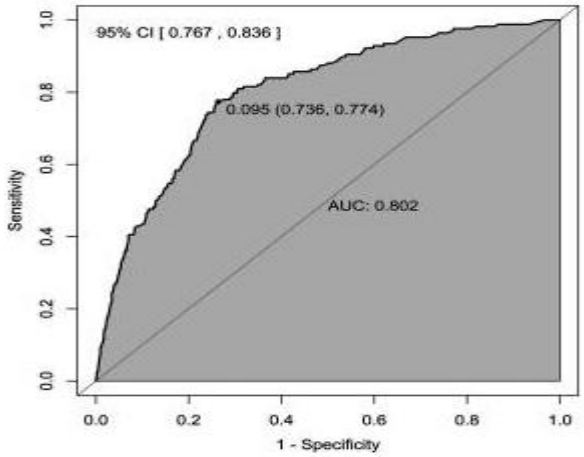


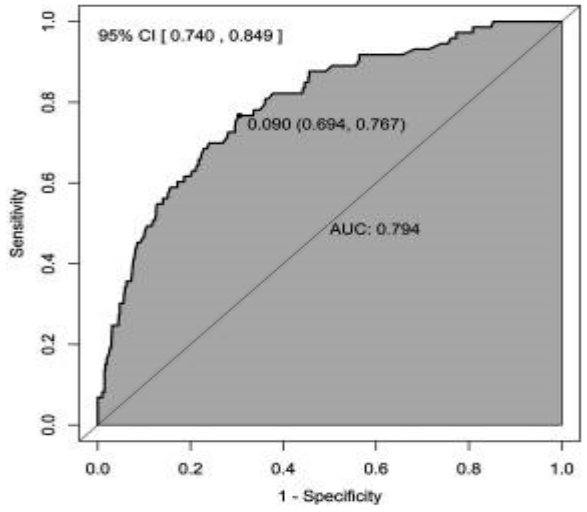
Figure 2. Nomogram for estimating depression risk in cancer survivors.

Predictive model validation

The accuracy of the constructed nomogram was evaluated using receiver operating characteristic (ROC) curves in both the training and validation datasets. The model displayed strong discriminative ability, with area under the curve (AUC) values of 0.802 (95 percent CI: 0.767–0.836) for the training cohort and 0.794 (95 percent CI: 0.740–0.849) for the validation cohort, reflecting reliable predictive performance **(Figure 3)**. To further assess the stability of the model, bootstrap resampling with 1000 iterations was conducted, resulting in an optimism-adjusted AUC of 0.812 (95 percent CI: 0.784–0.840).



a)



b)

Figure 3. ROC analysis for the training set (a) and validation set (b).

Prediction accuracy was further evaluated using calibration curves for both datasets. As shown in **Figure 4**, the predicted probabilities closely matched the observed outcomes, with curves aligning near the ideal 45-degree line. In the training cohort, the calibration slope was 1.00 with an intercept near 0, whereas in the validation cohort, the slope was 1.01 and the intercept −0.03. Additionally, the Hosmer–Lemeshow test supported the model's calibration, yielding nonsignificant P-values of 0.112 for

the training set and 0.475 for the validation set, confirming satisfactory agreement between predicted and actual outcomes.

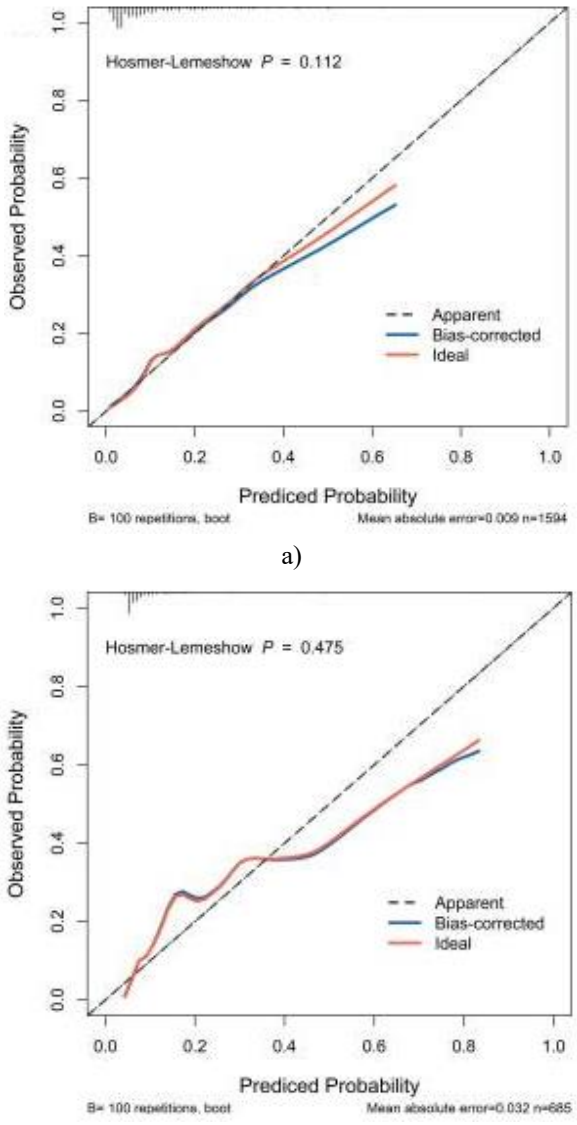


Figure 4. Calibration curves depicting the nomogram's predictive performance in the training set (a) and validation set (b).

The clinical relevance of the nomogram was examined through decision curve analysis, which evaluates the net benefit of using the model across varying risk thresholds. In these plots, the x-axis represents different probability thresholds, while the y-axis shows the corresponding net benefit, facilitating a comparison of the model's added value relative to alternative strategies. **Figure 5** demonstrates that the nomogram consistently provides a positive net benefit in both the training and validation cohorts, indicating that it could meaningfully inform clinical decisions and support risk-guided management in cancer survivors.

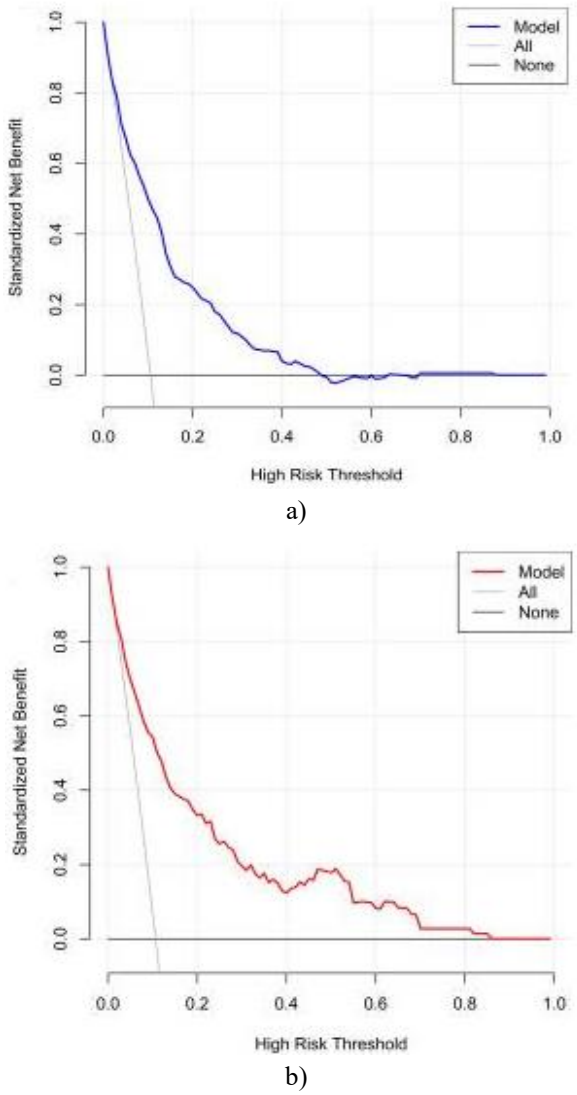


Figure 5. Decision curve analysis (DCA) for the training set (a) and validation set (b).

Using nationally representative NHANES data from 2005–2018, this study developed and validated a predictive nomogram to assess depression risk among cancer survivors. Multivariable analysis revealed seven significant predictors: middle-aged status, lower educational attainment, reduced poverty-income ratio, current smoking, presence of congestive heart failure, sleep disturbances, and multiple cancer sites. The nomogram showed strong discriminative ability and practical clinical utility through extensive validation.

Our findings indicate that middle-aged cancer survivors have a higher likelihood of depression, consistent with prior studies. Compared to older adults, those diagnosed during working age may face greater disruptions to employment and social life, leading to elevated depression rates [24, 25]. The study also identified low income as a robust predictor of depression among cancer survivors, aligning with previous research suggesting that limited financial resources exacerbate medical and living expenses, potentially restricting access to timely and high-

quality care and increasing psychological stress [26, 27]. Furthermore, lower educational attainment emerged as a significant predictor of depression, with prior research on prostate cancer patients showing a 1.86-fold higher risk of depression among individuals with less education [28]. Limited education may hinder understanding of complex medical information and reduce the ability to adopt effective coping strategies necessary for managing cancer-related emotional challenges [29].

Comorbid chronic conditions appear to further elevate depression risk in cancer survivors. Literature indicates that patients with chronic diseases have up to a 1.7-fold higher risk of depression than those without comorbidities [12]. In this study, congestive heart failure was identified as a notable predictor of depression. Previous evidence shows that up to 30% of patients with heart failure experience depressive symptoms [30], and this condition is commonly observed following cancer treatment, particularly in those receiving combined radiotherapy and chemotherapy [31, 32]. The mechanisms linking heart failure and depression in cancer survivors are not fully understood, but may involve reduced physical functioning and quality of life, which negatively impact mood [33], along with shared biological pathways such as neuroendocrine dysregulation [34] and elevated inflammatory markers (e.g., C-reactive protein and interleukin-6) [35, 36]. However, these explanations remain speculative, highlighting the need for further research to clarify how cardiac function affects mood in this population.

Sleep disturbances were strongly associated with depression, with cancer survivors without sleep problems showing a markedly lower risk. These results are consistent with prior research, including a systematic review and meta-analysis reporting that 57.4% of cancer patients experience impaired sleep quality (95% CI: 53.3–61.6%) [37], and meta-regression analyses indicate a positive correlation between poor sleep and comorbid depression [37]. Sleep problems can cause fatigue, cognitive deficits, and disruptions in metabolic and neuroendocrine function, which can diminish quality of life and exacerbate depressive symptoms [38, 39]. Additionally, consistent with existing literature [40], this study found that never-smokers and former smokers had lower depression risk compared to current smokers, potentially due to the pro-inflammatory effects of smoking that may trigger neuroimmune changes contributing to depression [41].

In this study, we also investigated cancer-specific factors associated with depression, identifying the number of tumors as a significant predictor: cancer survivors with multiple tumor sites had a markedly higher risk of depression than those with a single site (OR = 7.51, 95% CI: 1.66–33.90). A matched-cohort study in Japan reported similar findings, showing the highest depression risk in patients with multiple cancers compared to cancer-free individuals [42]. Multiple tumors often reflect advanced disease stages or poorer overall health, and prior studies have linked advanced cancer stages to higher depression rates across various cancer types [12, 13]. The presence of multiple tumors can increase disease burden, necessitate more complex treatments, and cause greater somatic discomfort, while potentially triggering inflammation and hormonal imbalances, all of which may collectively heighten depression risk.

Several prior studies have proposed predictive tools for depression in cancer patients. Consistent with Zuo and Yang (2025) [19], our study identified lower poverty-income ratio (PIR) and sleep disturbances as key predictors, while also including additional clinically relevant factors—congestive heart failure, smoking status, and multiple cancer sites—that were not part of their model, enhancing clinical utility and risk stratification. Other studies on colorectal and breast cancer patients similarly identified lower income and multiple comorbidities as depression predictors [15, 16]. However, inpatient populations undergoing active treatment may have mental health influenced more heavily by treatment-related factors such as pain, postoperative complications, and adjunctive therapies [15, 43]. Previous NHANES-based models for depressive symptoms included over 20 variables, limiting practical use and lacking validation for calibration and clinical efficacy [44]. In contrast, our approach employed LASSO regression to select variables before multivariable logistic regression, reducing overfitting and multicollinearity while simplifying the model. The resulting nomogram comprises seven easily obtainable clinical factors, and its validity and reliability were confirmed using calibration curves and DCA analysis.

Early identification of depression among cancer survivors is crucial for timely intervention and improved prognosis. In low-income settings, depressive symptoms are often overlooked, as patients are more likely to present with advanced disease, limited treatment options, and poorer outcomes [27]. Additionally, primary care settings may lack sufficient psychiatric resources, hindering depression screening. The nomogram developed in this study enables clinicians to quantify depression risk based on seven key factors, facilitating early recognition of high-risk individuals and the delivery of targeted interventions. Special attention should be given to patients presenting with one or more of these risk factors.

Several limitations should be acknowledged. First, NHANES data on medical conditions and depression rely on self-report or screening tools such as the PHQ-9, rather than clinician diagnosis, which may introduce reporting or recall bias. NHANES excludes institutionalized or

severely ill individuals, potentially causing survivorship bias and underestimation of depression prevalence. Second, the cross-sectional design precludes causal inferences, and unmeasured confounders—such as psychosocial factors and treatment-specific variables—may influence the associations observed. Third, missing data were present due to the extensive covariate set, raising the possibility of selection bias, although multiple imputation sensitivity analyses yielded consistent results. Fourth, while internal validation was performed, the model lacks external validation, and its generalizability to non-U.S. populations, particularly in low- and middle-income countries, remains uncertain due to differences in healthcare systems, cultural factors, and survivorship care. Future research should evaluate the model in diverse populations and explore additional psychosocial and treatment-related factors to enhance predictive accuracy.

Conclusion

This study developed and validated a visual nomogram to predict depression risk among cancer survivors, incorporating seven easily obtainable predictors. The model demonstrated strong discrimination, calibration, and clinical applicability, offering a practical tool for healthcare providers to identify high-risk individuals and facilitate early intervention. Future prospective studies and external validation are needed to assess generalizability across populations and healthcare settings and to confirm clinical utility.

Acknowledgments: The authors gratefully acknowledge the National Health and Nutrition Examination Survey (NHANES) participants and research personnel for their invaluable contributions to data collection and study implementation.

Conflict of interest: None

Financial support: None

Ethics statement: None

References

1. Andrade LH, Alonso J, Mneimneh Z, Wells JE, Al-Hamzawi A, Borges G, et al. Barriers to mental health treatment: results from the WHO World Mental Health Surveys. Psychol Med. 2014;44(6):1303–17.
2. Miller KD, Nogueira L, Devasia T, Mariotto AB, Yabroff KR, Jemal A. Cancer treatment and survivorship statistics, 2022. CA Cancer J Clin. 2022;72(5):409–36.
3. Pitman A, Suleman S, Hyde N, Hodgkiss A. Depression and anxiety in patients with cancer. BMJ. 2018;361(1):k1415.
4. Vehling S, Mehnert-Theuerkauf A, Philipp R, Koch U, Härter M, Schulz H. Prevalence of mental disorders in patients with cancer compared to matched controls. Acta Oncol. 2022;61(1):7–13.
5. Mitchell AJ, Chan M, Bhatti H, Halton M, Grassi L, Johansen C. Prevalence of depression, anxiety, and adjustment disorder in oncological settings. Lancet Oncol. 2011;12(2):160–74.
6. Andersen BL, McElroy JP, Carbone DP, Phillips KM, Jones LW. Psychological symptom trajectories and non-small cell lung cancer survival. Psychosom Med. 2022;84(2):215–23.
7. Andersen BL, Lacchetti C, Ashing K, Barsevick AM, Beckjord E, Blanchard CM, et al. Management of anxiety and depression in adult survivors of cancer. J Clin Oncol. 2023;41(20):3426–53.
8. Nakhlband A, Farahzadi R, Saeedi N, Barzegar H, Montazersaheb S, Soofiyani SR. Bidirectional relations between anxiety, depression, and cancer. Curr Drug Targets. 2023;24(2):118–30.
9. Abdel-Rahman O. Depression and suicidal ideation among patients with cancer in the United States. JCO Oncol Pract. 2020;16(7):e601–9.
10. Wang YH, Li JQ, Shi JF, Que JY, Liu JJ, Lappin JM, et al. Depression and anxiety in relation to cancer incidence and mortality. Mol Psychiatry. 2020;25(7):1487–99.
11. Zhao L, Li X, Zhang Z, Song C, Wang Y, Wang J. Prevalence, correlates and recognition of depression in Chinese inpatients with cancer. Gen Hosp Psychiatry. 2014;36(5):477–82.
12. Riedl D, Schüßler G. Factors associated with and risk factors for depression in cancer patients. Transl Oncol. 2022;16(1):101328.
13. Ikhile D, Ford E, Glass D, Gremesty G, van Marwijk H. Risk factors associated with depression and anxiety in cancer patients. PLoS One. 2024;19(2):e0296892.
14. Yu X, Tian S, Wu L, Zheng H, Liu M, Wu W. Construction of a depression risk prediction model for type 2 diabetes mellitus patients. J Affect Disord. 2024;349(1):217–25.
15. Hu Z, Zhang H, Wang J, Liu X, Wu X, Zhang Y. Nomogram to predict postoperative anxiety and depression in colorectal cancer patients. Int J Gen Med. 2022;15(1):4881–95.
16. Mao Y, Li J, Shi R, Gao L, Xu A, Wang B. Nomogram risk prediction model for depressive symptoms in elderly breast cancer patients. Sci Rep. 2024;14(1):26433.

17. Mao Y, Shi RX, Gao LM, Li J, Xu A, Wang B. Nomogram-based risk prediction model for depressive symptoms in breast cancer patients. World J Clin Oncol. 2025;16(4):102208.
18. Xing Y, Zhao W, Duan C, Liu Y, Zhang Y, Li H. Visual model for predicting depression in lung cancer patients. J Clin Nurs. 2023;32(21–22):4614–25.
19. Zuo W, Yang X. Dynamic online nomogram for predicting depression risk in cancer patients. J Affect Disord. 2025;385(1):119402.
20. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med. 2001;16(9):606–13.
21. National Center for Health Statistics. Age adjustment – Health, United States. Centers for Disease Control and Prevention. 2023;(1):1–5.
22. Yao J, Chen X, Meng F, Cao H, Shu X. Nutritional and inflammatory status and depressive symptoms on mortality among cancer survivors. Brain Behav Immun. 2024;115(1):109–17.
23. Petrova D, Catena A, Rodríguez-Barranco M, Redondo-Sánchez D, Chacón-Cuberos R, Amiano P. Physical comorbidities and depression in adult cancer survivors. Cancers (Basel). 2021;13(13):3368.
24. Kuba K, Esser P, Mehnert A, Johansen C, Schwinn A, Koch U. Risk for depression and anxiety in hematologic cancer survivors. Health Psychol. 2019;38(3):187–95.
25. Götze H, Friedrich M, Taubenheim S, Dietz A, Lordick F, Mehnert A. Depression and anxiety 5 and 10 years after cancer diagnosis. Support Care Cancer. 2020;28(1):211–20.
26. Li S, He Y, Liu J, Wang J, He M, Zhang Y, et al. Socioeconomic status and cancer: umbrella review. Nat Commun. 2024;15(1):9993.
27. Walker ZJ, Xue S, Jones MP, Ravindran AV. Mental disorders in cancer patients in LMICs. JCO Glob Oncol. 2021;7(1):1233–50.
28. Friberg AS, Rask Moustsen I, Benzon Larsen S, Johansen C, Dalton SO. Educational level and depression after prostate cancer. Acta Oncol. 2019;58(5):722–9.
29. Lu W, Pikhart H, Peasey A, Kubinova R, Pitman A, Bobak M. Risk of depressive symptoms before and after cancer hospitalisation. J Affect Disord. 2020;276(1):76–83.
30. Celano CM, Villegas AC, Albanese AM, Gaggin HK, Huffman JC. Depression and anxiety in heart failure. Harv Rev Psychiatry. 2018;26(4):175–84.
31. Ahmad TA, Gopal DP, Chelala C, Dayem Ullah AZ, Taylor SJ. Multimorbidity in people living with and beyond cancer. Am J Cancer Res. 2023;13(8):4346–65.
32. Hooning MJ, Botma A, Aleman BM, Baaijens MH, Bartelink H, Klijn JG. Cardiovascular disease risk in breast cancer survivors. J Natl Cancer Inst. 2007;99(5):365–75.
33. Sbolli M, Fiuzat M, Cani D, O'Connor CM. Depression and heart failure. Eur J Heart Fail. 2020;22(11):2007–17.
34. Redwine LS, Wirtz PH, Hong S, Bosch JA, Ziegler MG, Greenberg B. Depression and leukocyte mobilization in heart failure. J Am Coll Cardiol. 2010;56(21):1720–7.
35. Xiong GL, Prybol K, Boyle SH, Dunlap ME, Sheps DS, Freedland KE. Inflammation markers and major depressive disorder in heart failure. Psychosom Med. 2015;77(7):808–15.
36. Mommersteeg PMC, Schoemaker RG, Naudé PJW, Eisel UL, Garrelds IM, Schalkwijk CG. Depression and inflammation predicting mortality in heart failure. Brain Behav Immun. 2016;57(1):144–50.
37. Chen MY, Zheng WY, Liu YF, Wang J, Liu Z, Li Y. Poor sleep quality in cancer patients. Gen Hosp Psychiatry. 2024;87(1):92–102.
38. Kisamore CO, Kisamore CA, Walker WH. Circadian rhythm disruption in cancer survivors. Cancer Med. 2024;13(6):e70353.
39. Lanza G, Mogavero MP, Salemi M, Ferri R. Sleep, immunity, and cancer. Cells. 2024;13(14):1246.
40. Yan G, Zhang Q, Yan Y, Wang Y, Liu X, Chen Y. Trends in comorbid depression in cancer patients. J Affect Disord. 2023;340(1):743–50.
41. McFarland DC, Riba M, Grassi L. Cancer-related inflammation and depression. Clin Pract Epidemiol Ment Health. 2021;17(1):287–94.
42. Akechi T, Mishiro I, Fujimoto S, Murase K. Risk of major depressive disorder in Japanese cancer patients. Psychooncology. 2020;29(10):1686–94.
43. Li RX, Li XL, Wu GJ, Zhang Y, Wang Y, Chen X. Risk factors for anxiety and depression after prostate cancer castration. World J Psychiatry. 2024;14(3):255–65.
44. Chen X, Ye C, Liu L, Li X. Factors associated with depressive symptoms among cancer patients. BMC Public Health. 2024;24(1):1443.