Assessment of Dietary and Nutritional Risk Factors Associated with Cancer Disease Through Factor Analysis

Sujatha Inginshetty[1*], Shweta. S[2]
[1]Department of Statistics, Gulbarga University, Kalaburagi-585106, India.
(ORCiD: https://orcid.org/0009-0009-5971-4255)
[2]Department of Statistics, Gulbarga University, Kalaburagi-585106, India.
(ORCiD: https://orcid.org/0000-0001-5152-2771)

ABSTRACT:

Cancer is one of the most serious non-communicable diseases affecting populations worldwide, and its rising incidence in India has become a major public health concern. Among various risk factors, unhealthy dietary and nutritional habits are now increasingly being recognized as critical contributors to cancer development. Rapid urbanization, lifestyle transitions, and lack of awareness regarding balanced nutrition have led to a growing prevalence of diet-related cancers. The present study aims to examine the association between dietary and nutritional patterns and the risk of cancer among cancer patients in Kalaburagi city. A total of 500 cancer patients from two prominent cancer care hospitals were selected for this study. Data were collected using a structured questionnaire that captured information on sociodemographic characteristics, dietary habits, and nutrition-related behaviours. To identify the most influential dietary risk patterns, a factor analysis approach was employed. This multivariate statistical technique helps in reducing a large set of interrelated variables into key underlying dietary factors influencing cancer occurrence. The study seeks to provide evidence-based insights into how specific nutrition-related behaviours may be linked to cancer risk, and to inform public health strategies and interventions aimed at dietary improvement and cancer prevention in Kalaburagi city.

Keywords: Cancer, Dietary patterns, Nutritional risk factors, Factor analysis.

1. INTRODUCTION

Cancer remains one of the most critical public health challenges worldwide, accounting for a growing share of global morbidity and mortality. The increasing burden in developing nations like India is largely attributed to rapid urbanization, lifestyle changes, and dietary transitions from traditional plant-based diets to energy-dense, nutrient-poor foods (Sinha et al., 2003). Dietary habits are among the most modifiable determinants of cancer, influencing both its initiation and progression. Over the past few decades, the globalization of food systems has resulted in a gradual shift towards Western dietary patterns characterized by higher consumption of red and processed meat, refined grains, sugary beverages, and saturated fats (Shridhar et al., 2018).

*Corresponding Author Email: sujathainginshetty22@gmail.com
Published: 03/12/2025
DOI: https://doi.org/10.70558/IJST.2025.v2.i4.241140

Several Indian studies have emphasized the growing link between poor dietary practices and increased non-communicable disease burden. In Karnataka, recent research revealed that frequent consumption of fried snacks, pickles, fast food, and sweetened beverages was associated with higher dietary risk scores and metabolic imbalances among rural and urban populations (Shalini et al., 2024). Similarly, nutritional surveys in Kalaburagi district identified persistent challenges such as undernutrition, irregular meal patterns, and micronutrient deficiencies, particularly among women and low-income groups. These findings highlight the double burden of malnutrition where undernutrition coexists with rising overweight and obesity, thus increasing susceptibility to cancer and other chronic diseases.

Cultural dietary practices, including vegetarianism, have drawn attention for their potential protective role against cancer. However, research indicates mixed outcomes: while some studies associate plant-based diets with reduced oxidative stress and cancer risk, others report no significant association between lifelong vegetarianism and specific cancers such as breast cancer (Gathani et al., 2017). On the other hand, diets rich in animal fats, processed meats, and refined carbohydrates have been linked with gastrointestinal, colorectal, and breast cancers, whereas high-fiber diets rich in fruits, vegetables, and whole grains have shown protective effects (Shivanna & Urooj, 2016).

Another crucial factor in cancer prevention is public awareness of the diet–disease relationship. Unfortunately, inadequate knowledge and misconceptions about healthy eating remain widespread in Indian communities (Elangovan et al., 2016). The Indian Cancer Society estimates that nearly one-third of cancer cases in India could be prevented through appropriate diet, regular physical activity, and reduced tobacco and alcohol use. Despite this, localized research on dietary and nutritional patterns linked to cancer remains limited, particularly in semi-urban areas such as Kalaburagi, which is currently undergoing major nutritional and lifestyle transitions (Anandi et al., 2018).

By analyzing the dietary habits and nutritional intakes of cancer patients, this study aims to identify prevalent dietary patterns and behaviors associated with cancer risk. The results are expected to provide valuable insights into the dietary and nutritional determinants of cancer patients in Kalaburagi city and contribute to the formulation of targeted nutritional interventions and cancer prevention strategies.

2. MATERIALS AND METHODS

The present study was carried out to examine the relationship between dietary and nutritional patterns and cancer risk among patients in Kalaburagi city, Karnataka**.** A cross-sectional study design was adopted, and a total of 500 cancer patients were selected from two major cancer care hospitals in the Kalaburagi city.

Data were collected using a structured and pre-tested questionnaire developed specifically to study the dietary and nutritional risk factors among cancer patients. The questionnaire included information on sociodemographic factors, dietary habits, nutritional behaviors, lifestyle factors, anthropometric factors, and biochemical factors.

Although the questionnaire covered all these aspects, the present study specifically focused on analysing the dietary and nutritional risk factors associated with cancer. Data collection was carried out through personal interviews after obtaining informed consent from all participants.

To identify major dietary and nutritional patterns, Exploratory Factor Analysis (EFA) was employed using the following steps:

1. Testing Data Suitability:

The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and Bartlett's Test of Sphericity were performed to verify the suitability of data for factor analysis.

2. Factor Extraction:

Principal Component Analysis (PCA) was used to extract the main dietary and nutritional components that contribute to cancer risk.

3. Factor Rotation:

The Varimax rotation method was applied to obtain a clear and interpretable factor structure.

4. Interpretation of Factors:

The Component Matrix and Rotated Component Matrix were examined to identify key dietary and nutritional factors. Each extracted factor was named based on the variables with the highest loadings, representing dominant dietary and nutritional patterns among cancer patients.

The identified factors provide important insights into the underlying dietary and nutritional risk behaviours associated with cancer in the Kalaburagi population.

3. RESULTS AND DISCUSSION

In the analysis of dietary and nutritional risk factors among cancer patients, 18 dietary pattern categories were formulated based on patient responses. Similar food items and eating behaviours were grouped together, and these categories were later analysed using PCA to derive meaningful factors. This approach helped in identifying how specific dietary habits may be associated with cancer risk.

3.1 KMO and Bartlett's Test

Table 1:KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		**0.733**
Bartlett's Test of Sphericity	Approx. Chi-Square	1593.299
	Degrees of freedom	153
	Significance	**0.000**

The KMO statistic assesses whether the dataset is appropriate for factor analysis. A value above 0.5 is generally regarded as adequate, and values exceeding 0.7 reflect good adequacy. In the present study, the obtained KMO value of 0.733 indicates that the sample meets the required conditions for conducting factor analysis.

Similarly, Bartlett's Test of Sphericity evaluates whether the correlation matrix differs significantly from an identity matrix. The test yielded a p-value below 0.05, confirming that sufficient correlations exist among the variables. This supports the use of factor analysis for the dataset.

3.2 The Initial Factor Analysis Solution (Extraction Method: PCA)

In factor analysis using Principal Component Analysis (PCA), communalities represent the proportion of a variable's variance explained by all the extracted factors. Initially, the communality values are set to 1, meaning each variable is fully represented. After extraction, the final communalities indicate how much of each variable's variance is accounted for by the principal components. Variables with higher extraction values are better represented in the factor solution.

Table 2: Communalities

	Initial	Extraction
Consumption of sugary beverages	1.000	0.492
Eat processed good high in sugar &fat	1.000	**0.634**
Consumption of fruits & vegetables	1.000	**0.587**
Consumption of high calories foods or deserts	1.000	**0.543**
Eat variety of colourful fruits &vegetables	1.000	0.462
Consumption of sugary snacks	1.000	**0.610**
Consumption of homemade meals	1.000	**0.527**
Consumption of processed or packaged food	1.000	**0.603**
Including whole grains in meals	1.000	**0.657**
Consumption of dairy products	1.000	0.455
Eat late at night or before bed	1.000	0.405
Eating red or processed meat	1.000	**0.687**
Consumption of low-carbohydrate foods	1.000	0.394
Include nuts and seeds in diet	1.000	0.443
Read labels before purchasing food products	1.000	**0.582**
Follows specific dietary plan	1.000	**0.760**

Dietary supplements'	1.000	**0.734**
Consumption of high-caffeine beverages	1.000	**0.547**
Extraction Method: Principal Component Analysis.		

Table 2 shows the communalities for the dietary and nutritional variables, indicating how much of the variance in each variable is explained by the extracted factors. A communality value above 0.5 is generally considered acceptable for inclusion in factor analysis.

In this study, most variables had communalities above 0.5, showing that a substantial proportion of their variance was explained by the extracted factors. For example, following a specific dietary plan (0.760), use of dietary supplements (0.734), eating red or processed meat (0.687), including whole grains in meals (0.657), consuming processed foods high in sugar and fat (0.634), consuming processed or packaged food (0.603), and consuming sugary snacks (0.610) all had high communalities. Other variables such as consumption of fruits and vegetables (0.587), reading labels before purchasing food products (0.582), consumption of high-caffeine beverages (0.547), consumption of homemade meals (0.527), and consumption of high-calorie foods or desserts (0.543) also met the threshold for further analysis.

A few variables had communalities below 0.5, which indicates that a smaller proportion of their variance was explained by the extracted factors. The majority of variables showed adequate communalities, confirming that they are suitable for further analysis using PCA.

3.3 Total Variance Explained

Table 3 presents the eigenvalues and the percentage of variance explained by each extracted factor. Factors with eigenvalues greater than 1 are considered the most important and influential. The table is divided into three sections: Initial Eigenvalues, Extraction Sums of Squared Loadings, and Rotation Sums of Squared Loadings. These values help determine the number of factors to retain and show how much variance in the data is explained by each factor.

Table 3: Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.078	17.099	17.099	3.078	17.099	17.099	2.643	14.685	14.685
2	2.386	13.254	30.353	2.386	13.254	30.353	2.428	13.490	28.175
3	1.295	7.196	37.548	1.295	7.196	37.548	1.307	7.261	35.435
4	1.241	6.894	44.443	1.241	6.894	44.443	1.290	7.168	42.603
5	1.117	6.205	50.648	1.117	6.205	50.648	1.252	6.954	49.557

6	1.004	5.578	56.226	1.004	5.578	56.226	1.200	6.668	56.226
7	0.934	5.191	61.416						
8	0.878	4.878	66.295						
9	0.834	4.633	70.928						
10	0.810	4.498	75.426						
11	0.776	4.309	79.735						
12	0.713	3.959	83.694						
13	0.647	3.596	87.290						
14	0.606	3.365	90.655						
15	0.553	3.070	93.725						
16	0.499	2.772	96.497						
17	0.414	2.300	98.797						
18	0.216	1.203	100.000						
Extraction Method: Principal Component Analysis.									

A total of six components were extracted, as their eigenvalues were greater than 1, which is the standard criterion for factor retention.

The initial eigenvalues represent the variance of each component before extraction. Because PCA was conducted on standardized variables (correlation matrix), each variable had a variance of 1, and the total variance equaled the number of variables (18). After extraction, the variance explained by each component was slightly reduced, as shown in the Extraction Sums of Squared Loadings.

The first component accounted for the maximum variance (Eigenvalue = 3.078, 17.10%), followed by the second component (Eigenvalue = 2.386, 13.25%), third (Eigenvalue = 1.295, 7.20%), fourth (Eigenvalue = 1.241, 6.89%), fifth (Eigenvalue = 1.117, 6.21%), and sixth (Eigenvalue = 1.004, 5.58%). Together, these six components explained 56.23% of the total variance, as shown by the cumulative percentage.

The Rotation Sums of Squared Loadings optimized the factor structure, making the variance distribution more balanced across the six components. After rotation, the variance explained by each component was: 14.69%, 13.49%, 7.26%, 7.17%, 6.95%, and 6.67%, respectively. This rotation helps in better interpretation of the factors by distributing variance more evenly among them.

The PCA results indicate that the first six components represent the most important dietary and nutritional patterns influencing cancer risk, and they were retained for further analysis and interpretation.

3.4 Scree plot

A scree plot displays the eigenvalues of each component in factor analysis. Components with eigenvalues above 1 are considered significant, and the "elbow" point indicates where additional factors contribute little variance.

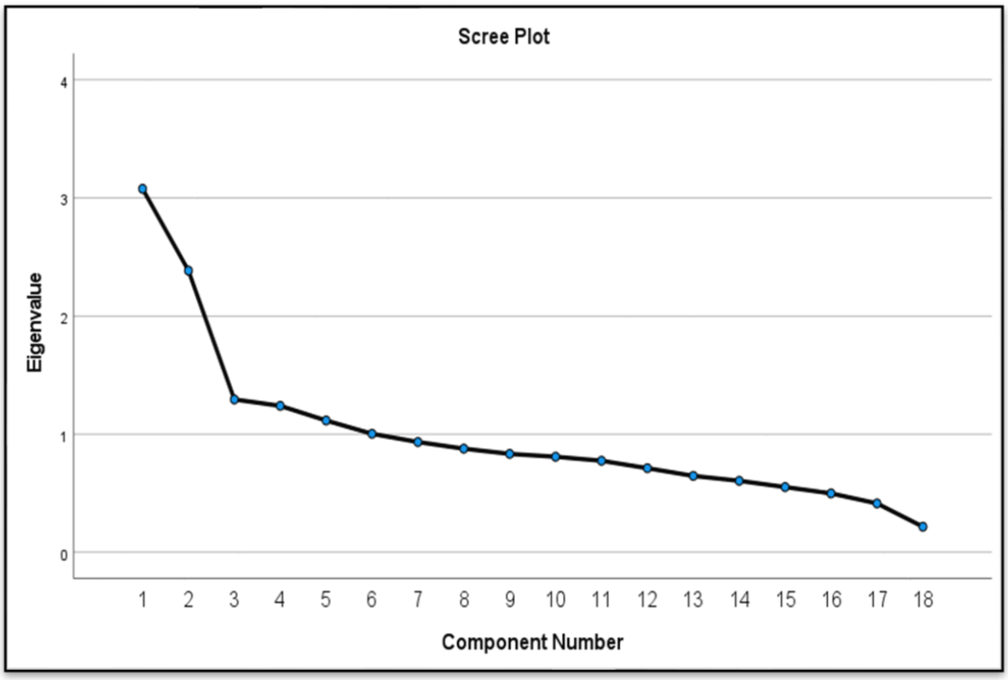

Figure 1: Scree plot

The scree plot shows eigenvalues plotted against component numbers. It indicates that the first six components—Dietary supplements intake, Following a dietary plan, Consumption of dairy products, Late-night eating habits, Eating fruits and vegetables, and Consumption of processed foods high in sugar and fat, explain the most variance. Components from the seventh onwards contribute much less to the total variation. Only components with eigenvalues greater than one were retained.

3.5 Component matrix

The component matrix presents the initial factor loadings for each variable before rotation. To facilitate clearer interpretation, the factors were rotated using the Varimax rotation with Kaiser Normalization, producing the rotated component matrix. Rotation simplifies the structure by maximizing high loadings and minimizing low loadings for each factor. In the matrix, variables with high factor loadings (greater than 0.5) are considered significant and are highlighted in bold for each component.

Table 4 : Component Matrix[a]						
	Factors					
	1	2	3	4	5	6

Consumption of sugary beverages	0.167	0.140	0.306	0.287	-0.517	-0.035
Eat processed good high in sugar &fat	0.396	**0.634**	0.074	0.166	-0.146	-0.143
Consumption of fruits & vegetables	0.168	-0.066	**0.605**	0.156	0.062	-0.401
Consumption of high calories foods or deserts	0.221	**0.657**	0.152	-0.076	-0.125	-0.136
Eat variety of colourful fruits &vegetables	0.402	-0.380	0.062	0.238	0.306	0.034
Consumption of sugary snacks	**0.577**	0.475	-0.141	-0.093	0.063	0.136
Consumption of homemade meals	-0.260	-0.188	0.602	-0.161	0.104	-0.156
Consumption of processed or packaged food	0.387	0.656	-0.083	0.080	0.057	-0.076
Including whole grains in meals	0.213	-0.186	0.303	0.403	-0.149	**0.548**
Consumption of dairy products	-0.388	0.219	0.300	0.025	0.049	0.404
Eat late at night or before bed	**0.526**	-0.108	-0.254	0.179	0.010	-0.141
Eating red or processed meat	-0.070	0.055	-0.094	0.695	0.405	-0.153
Consumption of low-carbohydrate foods	0.364	-0.217	-0.181	0.147	-0.251	0.312
Include nuts and seeds in diet	**0.550**	0.025	0.295	-0.004	0.093	0.210
Read labels before purchasing food products	0.330	0.267	0.187	-0.481	0.273	0.246
Follows specific dietary plan	**0.728**	-0.407	0.051	-0.204	0.136	-0.051
Dietary supplements	**0.721**	-0.431	0.026	-0.112	0.051	-0.114
Consumption of high-caffeine beverages	-0.180	0.354	0.036	0.132	**0.584**	0.172
Extraction Method: Principal Component Analysis.						
a. 6 components extracted.						

3.6 Sorted rotated factor loadings

The purpose of rotation is to reduce the number of factors that have high loadings for each variable, making the results easier to understand.

Table 5: Rotated Component Matrix[a]

	Factors					
	1	**2**	**3**	**4**	**5**	**6**
Dietary supplements	**0.812**	-0.029	0.138	0.158	0.086	-0.148
Follows specific dietary plan	**0.792**	-0.041	0.121	0.299	0.084	-0.139
Eat late at night or before bed	**0.564**	0.182	0.058	-0.116	-0.174	0.082
Eat variety of colourful fruits & vegetables	**0.523**	-0.165	0.244	0.080	0.091	0.294
Eat processed food high in sugar & fat	0.034	**0.790**	0.066	-0.040	0.028	0.035
Consumption of processed or packaged food	0.050	**0.739**	-0.052	0.117	-0.123	0.154
Consumption of high-calorie foods or desserts	-0.134	**0.701**	-0.073	0.102	0.100	-0.088
Consumption of sugary snacks	0.230	**0.610**	0.069	0.333	-0.261	0.018
Including whole grains in meals	0.026	-0.071	**0.803**	0.014	0.017	0.071
Read labels before purchasing food products	0.058	0.239	-0.013	**0.715**	0.031	-0.094
Consumption of fruits & vegetables	0.190	0.136	0.068	-0.090	**0.718**	0.070
Consumption of homemade meals	-0.169	-0.252	-0.031	0.121	**0.643**	-0.073
Eating red or processed meat	0.050	0.043	0.042	-0.322	0.032	**0.759**
Consumption of high-caffeine beverages	-0.288	0.096	-0.081	0.319	-0.012	**0.588**

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 23 iterations

The rotated component matrix shows factor loadings for each variable on the six components, making interpretation easier. The first component includes dietary supplements, following a dietary plan, late-night eating, and eating a variety of colourful fruits and vegetables, reflecting a pattern of health-conscious or structured eating. The second component represents consumption of processed foods, high-calorie desserts, and sugary snacks, indicating a preference for unhealthy or high-fat foods. The third component highlights whole grain intake, while the fourth component reflects food label awareness. The fifth component includes fruits, vegetables, and homemade meals, suggesting a natural or home-based diet. The sixth component includes red or processed meat and high-caffeine beverages, representing processed food and stimulant consumption. Variables with high loadings in each component are considered significant.

3.7 Component transformation Matrix

The component transformation matrix again displays the correlation among the components prior to and after rotation

Table 6:Component transformation Matrix

Factors	1	2	3	4	5	6
1	0.799	0.461	0.285	0.240	-0.054	-0.082
2	-0.468	0.841	-0.154	0.149	-0.106	0.131
3	-0.179	0.055	0.373	0.215	0.879	-0.082
4	0.020	0.112	0.467	-0.623	0.009	0.618
5	0.132	-0.177	-0.285	0.521	0.103	0.767
6	-0.304	-0.185	0.676	0.464	-0.450	0.006

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.

3.8 Factors Loading

The variables with the highest factor loadings in each component are considered significant and are presented in the table. For each component, the variable with the strongest loading is listed first, followed by the next highest, and so on for all factors.

Table 7: Factors loading values

Factors	**Label**	**High loading variables**	**Factors loading value**
1		Dietary supplements	0.812

	Health-Conscious Dietary Practices	Follows specific dietary plan	0.792
		Eat late at night or before bed	0.564
		Eat variety of colourful fruits & vegetables	0.523
2	Processed and Sugary Food Consumption	Eat processed food high in sugar & fat	0.790
		Consumption of processed or packaged food	0.739
		Consumption of high-calorie foods or desserts	0.701
		Consumption of sugary snacks	0.610
3	Whole Grains and Balanced Nutrition	Including whole grains in meals	0.803
4	Food Awareness and Label Reading	Read labels before purchasing food products	0.715
5	Natural & Homemade Eating Patterns	Consumption of fruits & vegetables	0.718
		Consumption of homemade meals	0.643
6	Meat and Stimulant Consumption	Eating red or processed meat	0.759
		Consumption of high-caffeine beverages	0.588

The rotated component matrix identified six key dietary patterns among the respondents:

Health-Conscious Dietary Practices (Component 1): Measured by dietary supplements (0.812), following a dietary plan (0.792), late-night eating (0.564), and eating a variety of colourful fruits and vegetables (0.523). These variables reflect structured and health-oriented eating habits.

Processed and Sugary Food Consumption (Component 2): Includes processed foods high in sugar and fat (0.790), packaged foods (0.739), high-calorie foods or desserts (0.701), and sugary snacks (0.610), representing frequent intake of unhealthy foods.

Whole Grains and Balanced Nutrition (Component 3): Measured by including whole grains in meals (0.803), indicating an emphasis on balanced dietary intake.

Food Awareness and Label Reading (Component 4): Measured by reading food labels before purchase (0.715), reflecting conscious food choices.

Natural and Homemade Eating Patterns (Component 5): Includes consumption of fruits and vegetables (0.718) and homemade meals (0.643), highlighting preference for natural, home-prepared foods.

Meat and Stimulant Consumption (Component 6): Includes eating red or processed meat (0.759) and high-caffeine beverages (0.588), indicating intake of animal-based and stimulant foods.

These six components capture the main dimensions of dietary behavior, ranging from health-conscious habits to processed food consumption and stimulant intake.

4. CONCLUSIONS

The present study provides strong evidence that dietary and nutritional patterns play a critical role in influencing cancer risk among patients in Kalaburagi city. Using a factor analysis approach, distinct dietary patterns were identified some of which were significantly associated with an increased risk of cancer. Unhealthy dietary behaviors, such as frequent consumption of high-fat processed foods, red meat, sugary beverages, and irregular meal timing, were among the major risk factors. Conversely, dietary habits that included regular consumption of fruits, vegetables, whole grains, and adequate water intake appeared to offer a protective effect.

These findings underline the importance of nutritional awareness and lifestyle modification in cancer prevention and management. Many participants demonstrated limited knowledge about balanced diets and healthy food practices, highlighting the need for widespread health education programs focused on nutrition and disease prevention. Public health policies should integrate diet-based counselling and cancer screening initiatives, especially in semi-urban and urban areas like Kalaburagi, where dietary transitions are occurring rapidly due to modernization. Moreover, this study sets the groundwork for future research to explore causal relationships using longitudinal data and biochemical assessments. Collaboration between nutritionists, oncologists, and public health experts is essential to translate these findings into effective interventions. Incorporating nutritional screening in routine cancer care and community outreach programs can significantly contribute to reducing cancer burden in the region.

Promoting healthy dietary patterns, raising awareness, and implementing preventive strategies at the grassroots level are vital steps toward reducing diet-related cancer risks and improving the overall well-being of the population in Kalaburagi city.

REFERENCES

1. Anandi, B. S., Reddy, S. B., & Indupalli, A. S. (2018). Assessment of nutritional status among under-five children in a rural area of Kalaburagi district. *National Journal of Community Medicine, 9*(8), 599–604.

2. Elangovan, A., Periasamy, A., Perumal, V., Ranganathan, D., Kothanda, V. R., Govindasamy, A., & Govindasamy, S. (2016). Awareness and Perception About Cancer Among the Public in Chennai, India. *Journal of Global Oncology*, *3*(5), 469–479. https://doi.org/10.1200/JGO.2016.006502

3. Gathani, T., Barnes, I., Ali, R., Arumugham, R., Chacko, R., Digumarti, R., Jivarajani, P., Kannan, R., Loknatha, D., Malhotra, H., Mathew, B. S., & INDOX Cancer Research Network Collaborators. (2017). Lifelong vegetarianism and breast cancer risk: A large multicentre case–control study in India. *BMC Women's Health, 17*(1), Article 6. https://doi.org/10.1186/s12905-016-0357-8

4. Shalini, S., Rajaram, D., Nooyi, S. C., Shivaraj, N. S., & Pruthvish, S. (2024). Assessment of dietary patterns and its risk factors among adults of a rural population of Chikkaballapur District, Karnataka. *Journal of Nutrition Research, 12*(2), 92–97. https://doi.org/10.55289/jnutres/v12i2.31.

5. Shivanna, L., & Urooj, A. (2016). A review on dietary and non-dietary risk factors associated with gastrointestinal cancer. *Journal of Gastrointestinal Cancer, 47*(3), 247–254. https://doi.org/10.1007/s12029-016-9845-1

6. Shridhar, K., Satija, A., Dhillon, P. K., Agrawal, S., Gupta, R., Bowen, L., Kinra, S., Bharathi, A. V., Prabhakaran, D., Reddy, K. S., & Ebrahim, S. (2018). Association between empirically derived dietary patterns with blood lipids, fasting blood glucose and blood pressure in adults – The India Migration Study. *Nutrition Journal, 17*, Article 15. https://doi.org/10.1186/s12937-018-0327-0

7. Sinha, R., Cross, A. J., Graubard, B. I., Leitzmann, M. F., & Schatzkin, A. (2009). Meat intake and mortality: A prospective study of over half a million people. *Archives of Internal Medicine, 169*(6), 562–571. https://doi.org/10.1001/archinternmed.2009.6