Bayesian Reasoning and Learning

Ning Wei Zhou (852757814)

Open University of the Netherlands

1 Introduction

Depression affects emotional, cognitive, and social well-being and manifests as prolonged low mood, diminished interest, hopelessness, fatigue, and altered sleep or appetite. Globally, around 280 million people suffer from depression World Health Organization [2023].

Conventional treatments include psychotherapy and pharmacological interventions Institute for Quality and Efficiency in Health Care [2024]. When these prove insufficient, alternative approaches like AI-driven mental healthcare offer significant promise by enabling early detection, personalized interventions, and virtual care. AI chatbots, in particular, enhance patient engagement and adherence Olawade et al. [2024], Liu et al. [2022].

Recent studies identify four dimensions of depression risk factors—socioeconomic status, physical attributes, genetic predispositions, and adverse childhood experiences Dagnino et al. [2020]. This study uses Bayesian network modeling to explore the interplay among these factors and depression onset, extending the framework with interventional nodes for both AI-based and traditional treatments to assess their moderating effects.

2 Domain

Depression stems from multiple interacting factors. For clarity, this study limits predictors to a select few. In the sociodemographic domain, only education and income are considered Dagnino et al. [2020]. In the physical domain, the focus is on physical activity and obesity status—highlighting obesity's impact over traditional BMI measurements (World Health Organization [2021b,a], Rutherford et al. [2022]). Genetic predisposition is modeled multiplicatively, reflecting that a family history can more than double depression risk (Dunn et al. [2015], Smith [2014]). Adverse childhood experiences (ACE) is incorporated using approximate estimates to maintain parsimony (Martins-Monteverde et al. [2019], Swedo et al. [2023]). Finally, nodes for AI-driven interventions and traditional treatments act as dampening factors in mitigating depression risk. This framework motivates the following research questions:

Research Question 1: How do sociodemographic factors, physical attributes, genetic influences, and adverse childhood experiences interact within a Bayesian network to predict depressive outcomes?

Research Question 2: How do traditional support methods and AI-enabled interventions interact to influence depression risk trajectories and outcomes within a Bayesian network framework?

3 Model

We use Python and the PyAgrum framework pyAgrum 2.1.1 Documentation [2025] to build our Bayesian network. Each variable represents a distinct determinant of depression, allowing us to examine their interrelationships. Table 1 lists the key variables in our simplified model. Nodes and arcs are then added using PyAgrum, as illustrated in the following code example.

```
# Create the combined Bayesian network for the depression model
bn = gum.BayesNet("CombinedDepressionModel")
```

```python
edu = bn.add(gum.LabelizedVariable("Education", "Education␣Level", ["Low", "Medium",
    ↪  "High"]))
inc = bn.add(gum.LabelizedVariable("Income", "Income␣Level", ["Low", "Medium", "High
    ↪ "]))

socio = bn.add(gum.LabelizedVariable("SocioStatus", "Overall␣Socioeconomic␣Status",
    ↪ ["High", "Low"]))

bn.addArc(inc, socio)
bn.addArc(edu, socio)
```

Variable	Domain	Description
Education	{Low, Medium, High}	Low corresponds to primary education, VMBO, the lower secondary phase of Havo/VWO, and MBO1. Medium denotes the upper secondary phase of Havo/VWO, MBO2, MBO3, and MBO4. High corresponds to HBO/WO. Data are obtained from Centraal Bureau voor de Statistiek [2022].
Income	{Low, Medium, High}	Categorizes household income in the USA, with data sourced from Pew Research Center [2020].
Physical Activity	{Inactive, Active}	Reflects the global percentage of the population that is physically inactive [World Health Organization, 2021a].
Obesity	{Not Obese, Obese}	Represents the proportion of the global population classified as obese [World Health Organization, 2021b].
Socioeconomic Status (SS)	{High, Low}	Synthesizes key indicators of High or Low economic status. Outcomes computed through a Noisy-OR.
Physical Attribute (PA)	{Poor, Good}	Synthesizes key indicators of physical health. Data is estimated heuristically for educational purposes
Adverse Childhood Experience (ACE)	{Absent, Present}	Indicates whether adverse childhood experiences were encountered. Data are retrieved from Swedo et al. [2023].
Genetic Predisposition (GP)	{Absent, Present}	Denotes the presence or absence of a genetic risk factor for depression. Data are retrieved from Kobayashi et al. [2024].
Depression	{No, Yes}	Binary outcome representing the absence or presence of depression. Outcomes computed through an Adapted Leaky Noisy-OR.
Artificial Intelligence	{No, Yes}	A binary variable denoting whether artificial intelligence has been implemented as an intervention to mitigate depression. The impact is estimated heuristically for educational purposes.
Traditional Methods	{No, Yes}	A binary variable denoting whether conventional treatment modalitieshave been employed to manage depression. The impact is estimated heuristically for educational purposes.

Table 1: Key variables in the simplified depression model.

Determinants of depression are grouped into four domains. The SocioStatus variable is computed using a Noisy-OR gate and then combined with other factors via an adapted leaky Noisy-OR mechanism to yield the final outcomes. Figure 1 shows the resulting structure, and the code used to generate this graph is provided below.

```python
node_colors = {bn.variable(i).name(): 1 for i in range(bn.size())}

node_colors["Artificial␣Intelligence"] = 0.1
node_colors["Traditional␣Methods"] = 0.1
node_colors["SocioStatus"] = 0.2
node_colors["Depression"] = 0.3

graph = gnb.show(bn, size="11", nodeColor=node_colors)
```

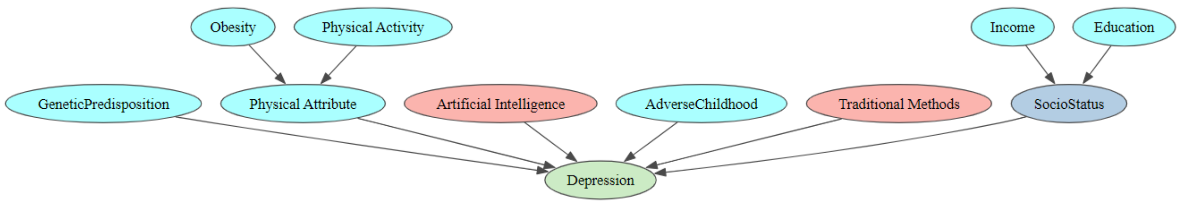

Figure 1: Bayesian network: red nodes are inactive by default, dark blue nodes use a standard Noisy-OR, and the green node uses a Adapted Leaky Noisy-OR.

3.1 Noisy-OR

In the Noisy-OR model, the probability that a child variable X remains inactive (x^0) given its parents U is computed by

$$P(x^0 \mid U) = \prod_{u_i \in U} (1 - \lambda_i)\,,$$

where λ_i is the probability that the i-th parent (here, SocioStatus) activates X. Thus, the probability that X is active (x^1) is

$$P(x^1 \mid U) = 1 - \prod_{u_i \in U} (1 - \lambda_i)\,.$$

Table 2 lists the probabilities used for the SocioStatus variable.

Parameter	Key	State/Interpretation	Value
	0	Low	0.8
λ_{Income}	1	Medium	0.5
	2	High	0.2
	0	Low	0.7
$\lambda_{Education}$	1	Medium	0.4
	2	High	0.1

Table 2: Parameters used in the Noisy-OR for SocioStatus

Code used to compute SocioStatus:

```
cpt_socio = bn.cpt(socio)

r_inc = {"Low": 0.8, "Medium": 0.5, "High": 0.2}
r_edu = {"Low": 0.7, "Medium": 0.4, "High": 0.1}

income_states = bn.variable(inc).labels() # e.g., ["Low", "Medium", "High"]
edu_states = bn.variable(edu).labels() # e.g., ["Low", "Medium", "High"]

for inc_state, edu_state in itertools.product(income_states, edu_states):
    factor_inc = 1 - r_inc[inc_state]
    factor_edu = 1 - r_edu[edu_state]

    product = factor_inc * factor_edu

    p_low = 1 - product
    p_high = product # so that p_high + p_low = 1

    cpt_socio[{
      bn.variable(inc).name(): inc_state,
      bn.variable(edu).name(): edu_state,
```

```
}] = [p_high, p_low]
```

3.2 Adapted Leaky Noisy-OR

When nodes for genetic predisposition, traditional methods, and AI-driven interventions are active, the model adjusts by computing $P(x^1 \mid U)$ with depression-specific parameters. We define the adjusted probability as

$$P(\tilde{x}^1 \mid U) = \min\left\{1, \prod_{j \in M} r_j \Big(P(x^1 \mid U)\Big)\right\}. \tag{1}$$

In this equation, λ_i (see Table 3) represents the direct influences, and the r_j factors (Table 4) serve as multiplicative scaling modifiers. A latent baseline risk, λ_{Leak}, accounts for depression risk from factors not explicitly modeled. All risk configurations are evaluated, with outcomes detailed in Tables 5–8 in the Probabilities subsection.

Parameter	Key	State/Interpretation	Value
λ_{SS}	0	Favorable (High SES)	0.005
	1	Unfavorable (Low SES)	0.05
λ_{PA}	0	Favorable (Good health)	0.005
	1	Unfavorable (Poor health)	0.03
λ_{ACE}	0	Absent	0.005
	1	Present	0.03
λ_{Leak} – baseline risk (always present)			0.01

Table 3: Parameters triggering the Adapted Noisy-OR for Depression Outcome

Intervention	Condition	r_j
Genetic Predisposition	Present (doubles risk)	2
Artificial Intelligence	Present (reduces risk)	0.80
Traditional Methods	Present (reduces risk)	0.90
Combined Intervention	AI & Traditional Methods Present	0.70

Table 4: Effective Risk Multipliers

Code for computing the Depression node:

```
ss_params = {0: 0.005, 1: 0.05}
ph_params = {0: 0.005, 1: 0.03}
ace_params = {0: 0.005, 1: 0.03}
gp_multiplier = {0: 1.0, 1: 2}
chatbot = {0: 1.0, 1: 0.80}
traditional = {0: 1.0, 1: 0.9}
leak = 0.01
cpt_depress = bn.cpt(depress)
cpt_values = [] # This flat list will accumulate the CPT probabilities for all
    ↪ combinations
for th, cb, ace_state, gp_state, ph, ss in itertools.product(range(2), range(2),
    ↪ range(2), range(2), range(2), range(2)):
    prod_failure = (1 - ss_params[ss]) * (1 - ph_params[ph]) * (1 - ace_params[
        ↪ ace_state]) * (1 - leak)
    p_baseline = 1 - prod_failure
    if gp_state == 1:
        p_dep = min(1.0, p_baseline * gp_multiplier[1])
    else:
```

```
        p_dep = p_baseline
    if cb == 1 and th == 1:
        p_dep = p_dep * 0.70
    elif th == 1:
        p_dep = p_dep * traditional[1]
    elif cb == 1:
        p_dep = p_dep * chatbot[1]
    cpt_values.extend([1 - p_dep, p_dep])
cpt_depress.fillWith(cpt_values)
```

4 Results

4.1 Probabilities

Data from the sources listed in Table 1 were cleaned and processed to derive the corresponding probability estimates. Tables 5 to 8 summarize depression outcomes under nonintervention and intervention scenarios. Due to the wide scope of the Depression variable, its representation was split into several smaller, manageable tables. Figures 2 through 9 depict the probability distributions for key variables including education, income, obesity, genetic predisposition, physical activity, physical attribute, adverse childhood experiences, and socioeconomic status. The code for generating these visualizations is provided below.

```
# Iterate over all nodes in the BN and show their CPTs:
for node in bn.nodes():
    var_name = bn.variable(node).name()
    print(f"Statistics␣for␣{var_name}:")
    gnb.showPotential(bn.cpt(node))
```

Adverse Childhood	**Genetic Predisposition**	**Physical Attribute**	**SocioStatus**	**Depression (No)**	**Depression (Yes)**
Absent	Absent	Poor	High	0.9752	0.0248
Absent	Absent	Poor	Low	0.9311	0.0689
Absent	Absent	Good	High	0.9507	0.0493
Absent	Absent	Good	Low	0.9077	0.0923
Absent	Present	Poor	High	0.9504	0.0496
Absent	Present	Poor	Low	0.8622	0.1378
Absent	Present	Good	High	0.9014	0.0986
Absent	Present	Good	Low	0.8154	0.1846
Present	Absent	Poor	High	0.9507	0.0493
Present	Absent	Poor	Low	0.9077	0.0923
Present	Absent	Good	High	0.9268	0.0732
Present	Absent	Good	Low	0.8849	0.1151
Present	Present	Poor	High	0.9014	0.0986
Present	Present	Poor	Low	0.8154	0.1846
Present	Present	Good	High	0.8537	0.1463
Present	Present	Good	Low	0.7698	0.2302

Table 5: Depression probabilities in default case

Adverse Childhood	Genetic Predisposition	Physical Attribute	SocioStatus	Depression (No)	Depression (Yes)
Absent	Absent	Good	High	0.9802	0.0198
Absent	Absent	Good	Low	0.9449	0.0551
Absent	Absent	Poor	High	0.9606	0.0394
Absent	Absent	Poor	Low	0.9262	0.0738
Absent	Present	Good	High	0.9604	0.0396
Absent	Present	Good	Low	0.8898	0.1102
Absent	Present	Poor	High	0.9212	0.0788
Absent	Present	Poor	Low	0.8524	0.1476
Present	Absent	Good	High	0.9606	0.0394
Present	Absent	Good	Low	0.9262	0.0738
Present	Absent	Poor	High	0.9415	0.0585
Present	Absent	Poor	Low	0.9079	0.0921
Present	Present	Good	High	0.9212	0.0788
Present	Present	Good	Low	0.8524	0.1476
Present	Present	Poor	High	0.8829	0.1171
Present	Present	Poor	Low	0.8159	0.1841

Table 6: Depression probabilities with Artificial Intelligence Intervention

Adverse Childhood	Genetic Predisposition	Physical Attribute	SocioStatus	Depression (No)	Depression (Yes)
Absent	Absent	Good	High	0.9777	0.0223
Absent	Absent	Good	Low	0.9380	0.0620
Absent	Absent	Poor	High	0.9556	0.0444
Absent	Absent	Poor	Low	0.9170	0.0830
Absent	Present	Good	High	0.9554	0.0446
Absent	Present	Good	Low	0.8760	0.1240
Absent	Present	Poor	High	0.9113	0.0887
Absent	Present	Poor	Low	0.8339	0.1661
Present	Absent	Good	High	0.9556	0.0444
Present	Absent	Good	Low	0.9170	0.0830
Present	Absent	Poor	High	0.9342	0.0658
Present	Absent	Poor	Low	0.8964	0.1036
Present	Present	Good	High	0.9113	0.0887
Present	Present	Good	Low	0.8339	0.1661
Present	Present	Poor	High	0.8683	0.1317
Present	Present	Poor	Low	0.7928	0.2072

Table 7: Depression probabilities with Traditional Methods Intervention

Adverse Childhood	Genetic Predisposition	Physical Attribute	SocioStatus	Depression (No)	Depression (Yes)
Absent	Absent	Good	High	0.9827	0.0173
Absent	Absent	Good	Low	0.9518	0.0482
Absent	Absent	Poor	High	0.9655	0.0345
Absent	Absent	Poor	Low	0.9354	0.0646
Absent	Present	Good	High	0.9653	0.0347
Absent	Present	Good	Low	0.9036	0.0964
Absent	Present	Poor	High	0.9310	0.0690
Absent	Present	Poor	Low	0.8708	0.1292
Present	Absent	Good	High	0.9655	0.0345
Present	Absent	Good	Low	0.9354	0.0646
Present	Absent	Poor	High	0.9488	0.0512
Present	Absent	Poor	Low	0.9194	0.0806
Present	Present	Good	High	0.9310	0.0690
Present	Present	Good	Low	0.8708	0.1292
Present	Present	Poor	High	0.8976	0.1024
Present	Present	Poor	Low	0.8389	0.1611

Table 8: Depression probabilities with both intervention methods

Education		
Low	Medium	High
0.1926	0.3995	0.4123

Figure 2: Education

Income		
Low	Medium	High
0.0900	0.4300	0.4800

Figure 3: Income

Obesity	
Obese	Not Obese
0.4300	0.5700

Figure 4: Obesity

GeneticPredisposition	
Absent	Present
0.6000	0.4000

Figure 5: Genetic Predisposition

Physical Activity	
Inactive	Active
0.3100	0.6900

Figure 6: Physical Activity

		Physical Attribute	
Obesity	Physical Activity	Good	Poor
Obese	Inactive	0.9000	0.1000
	Active	0.8000	0.2000
Not Obese	Inactive	0.6000	0.4000
	Active	0.5000	0.5000

Figure 7: Physical Attribute

AdverseChildhood	
Absent	Present
0.3610	0.6390

Figure 8: Adverse Childhood Experience

		SocioStatus	
Education	Income	High	Low
Low	Low	0.0600	0.9400
	Medium	0.1500	0.8500
	High	0.2400	0.7600
Medium	Low	0.1200	0.8800
	Medium	0.3000	0.7000
	High	0.4800	0.5200
High	Low	0.1800	0.8200
	Medium	0.4500	0.5500
	High	0.7200	0.2800

Figure 9: Socioeconomic Status

4.2 Inference

The Bayesian network yields a baseline depression probability of 10.39% under default conditions. With the Artificial Intelligence intervention, this drops to 8.31%, while the Traditional Methods intervention lowers it to 9.35%. The combined intervention further reduces the probability to 7.27%. The following code computes and visualizes these outcomes:

```
gnb.showInference(bn, evs={})
gnb.showInference(bn, evs={"Artificial Intelligence": "Yes"})
gnb.showInference(bn, evs={"Traditional Methods": "Yes"})
gnb.showInference(bn, evs={
"Artificial Intelligence": "Yes",
"Traditional Methods": "Yes"
})
```

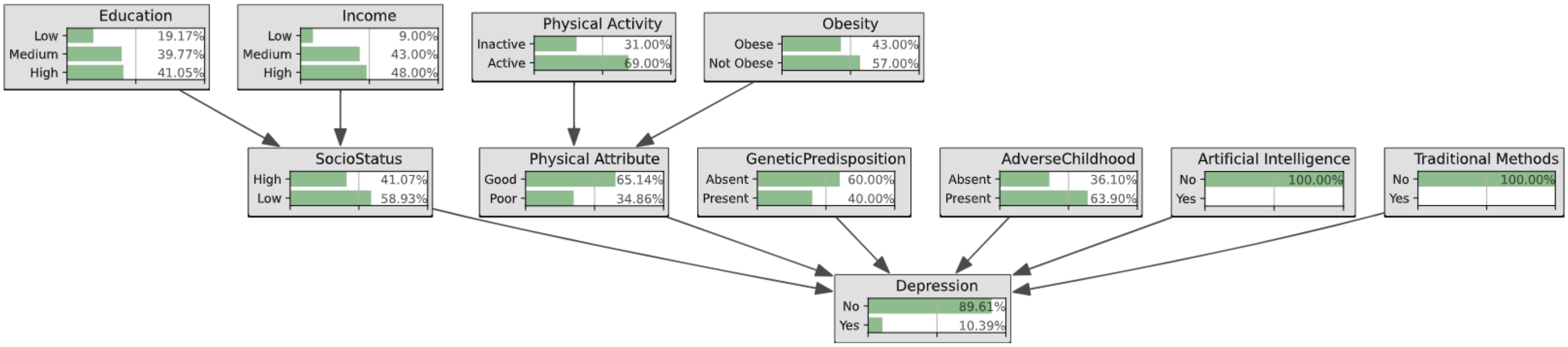

Figure 10: Default Outcome (10.39% probability)

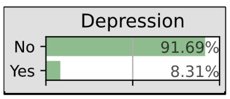

Figure 11: Artificial Intelligence Outcome (8.31% probability)

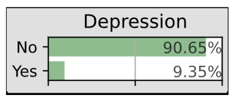

Figure 12: Traditional Method Outcome (9.35% probability)

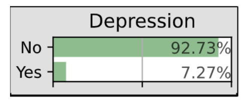

Figure 13: Combined Outcome (7.27% probability)

5 Discussion

For educational purposes, the model has been significantly simplified to highlight the essential mechanisms and interactions of risk factors within the Bayesian network, though it may not capture the full complexity of depression.

Data were drawn from diverse international sources, which may introduce discrepancies in risk factor measurements and interpretations due to variation in population characteristics and data collection methods.

Additionally, some parameter values—particularly the trigger factors for the Noisy-ORs—are based on estimates to provide a tractable example rather than definitive clinical or epidemiological values.

References

Centraal Bureau voor de Statistiek. Leeftijd, opleiding en inkomen in nederland [data set], 2022. Accessed: 08 May 2025. Available at: https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83687NED/table.

Paula Dagnino, María José Ugarte, Felipe Morales, Sofia González, Daniela Saralegui, and Johannes C. Ehrenthal. Risk factors for adult depression: Adverse childhood experiences and personality functioning. *Frontiers in Psychology*, 11, 2020. ISSN 1664-1078. doi: 10.3389/fpsyg.2020.594698. URL https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2020.594698.

E. C. Dunn, R. C. Brown, Y. Dai, J. Rosand, N. R. Nugent, A. B. Amstadter, and J. W. Smoller. Genetic determinants of depression: Recent findings and future directions. *Harvard Review of Psychiatry*, 23(1):1–18, 2015. doi: 10.1097/HRP.0000000000000054.

Institute for Quality and Efficiency in Health Care. Depression: Learn more – how is depression treated? InformedHealth.org [Internet]. Cologne, Germany: Institute for Quality and Efficiency in Health Care (IQWiG); 2006-, 2024. [Updated 2024 Apr 15]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK279282/.

Nobuyuki Kobayashi, Kazuya Shimada, Azusa Ishii, Rui Osaka, Toshiko Nishiyama, Masahiro Shigeta, Hiroyuki Yanagisawa, Naomi Oka, and Kazuhiro Kondo. Identification of a strong genetic risk factor for major depressive disorder in the human virome. *iScience*, 27(3):109203, 2024. ISSN 2589-0042. doi: https://doi.org/10.1016/j.isci.2024.109203. URL https://www.sciencedirect.com/science/article/pii/S2589004224004243.

Hao Liu, Huaming Peng, Xingyu Song, Chenzi Xu, and Meng Zhang. Using ai chatbots to provide self-help depression interventions for university students: A randomized trial of effectiveness. *Internet Interventions*, 27:100495, 2022. ISSN 2214-7829. doi: https://doi.org/10.1016/j.invent.2022.100495. URL https://www.sciencedirect.com/science/article/pii/S2214782922000021.

C. M. S. Martins-Monteverde, C. V. W. Baes, E. Reisdorfer, T. Padovan, S. M. C. Tofoli, and M. F. Juruena. Relationship between depression and subtypes of early life stress in adult psychiatric patients. *Frontiers in Psychiatry*, 10:19, 2019. doi: 10.3389/fpsyt.2019.00019.

David B. Olawade, Ojima Z. Wada, Aderonke Odetayo, Aanuoluwapo Clement David-Olawade, Fiyinfoluwa Asaolu, and Judith Eberhardt. Enhancing mental health with artificial intelligence: Current trends and future prospects. *Journal of Medicine, Surgery, and Public Health*, 3:100099, 2024. ISSN 2949-916X. doi: https://doi.org/10.1016/j.glmedi.2024.100099. URL https://www.sciencedirect.com/science/article/pii/S2949916X24000525.

Pew Research Center. Trends in income and wealth inequality. https://www.pewresearch.org/social-trends/2020/01/09/trends-in-income-and-wealth-inequality/, 2020. Accessed: May 10, 2025.

pyAgrum 2.1.1 Documentation. *pyAgrum Documentation*. pyAgrum Project, 2025. URL https://pyagrum.readthedocs.io/en/2.1.1/. Version 2.1.1.

E.R. Rutherford, C. Vandelanotte, and J. et al. Chapman. Associations between depression, domain-specific physical activity, and bmi among us adults: Nhanes 2011-2014 cross-sectional data. *BMC Public Health*, 22:1618, 2022. doi: 10.1186/s12889-022-14037-4. URL https://doi.org/10.1186/s12889-022-14037-4.

Katana Smith. More than 1 in 5 people have a genetic predisposition to depression. https://neura.edu.au/news-media/media-releases/1-5-people-genetic-predisposition-depression, 2014. Media Release, 13 April 2014.

E. A. Swedo, M. V. Aslam, L. L. Dahlberg, et al. Prevalence of adverse childhood experiences among u.s. adults — behavioral risk factor surveillance system, 2011–2020. *MMWR Morb Mortal Wkly Rep*, 72:707–715, 2023. doi: 10.15585/mmwr.mm7226a2.

World Health Organization. Physical activity. https://www.who.int/news-room/fact-sheets/detail/physical-activity, 2021a. Accessed: May 13, 2025.

World Health Organization. Obesity and overweight, 2021b. Accessed: May 09, 2025 Available at: https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight.

World Health Organization. Depressive disorder (depression). https://www.who.int/news-room/fact-sheets/detail/depression, 2023. Fact sheet. Published: 31 March 2023. Accessed: May 24, 2025.