

Stroke Risk Factors

- Factors associated with oral anticoagulant non-use at first ischemic stroke in atrial fibrillation: A nationwide study
- Prestroke sleep and stroke: a narrative review
- The future of biomarkers for vascular contributions to cognitive impairment and dementia (VCID): proceedings of the 2025 annual workshop of the Albert research institute for white matter and cognition
- Early life adversity increases risk for chronic post-traumatic pain, data from humans and rodents
- Health Behavior Decision-Making of Stroke High-Risk Population: A Latent Profile Analysis
- Impact of Depressive Symptom Severity on Stroke Risk in a US Cohort of People with HIV
- The relationship between occupation and specific forms of idiopathic adult-onset dystonia
- Novel Insights into the Mechanism and Treatment of Diabetes-Related Brain Complications: Focusing on the Blood-Brain Barrier Impairment

Stroke **risk factors** are conditions or behaviors that increase the likelihood of having a stroke. They are usually classified into **modifiable** and **non-modifiable** categories.

Non-modifiable risk factors

- **Age** – risk increases significantly after age 55
- **Sex** – men have a higher risk, but women have worse outcomes
- **Family history** – genetic predisposition
- **Race/Ethnicity** – higher risk in African American, Hispanic, and South Asian populations
- **Previous stroke or TIA (transient ischemic attack)**

Modifiable risk factors

- **Hypertension** – the most important modifiable factor
- **Diabetes mellitus**
- **Dyslipidemia** – high LDL, low HDL
- **Smoking**
- **Obesity and sedentary lifestyle**
- **Atrial fibrillation and other cardiac arrhythmias**
- **Excessive alcohol intake**
- **Obstructive sleep apnea**
- **Poor diet** – low in fruits/vegetables, high in saturated fats or sodium
- **Chronic stress and depression**

Emerging biochemical markers (from recent studies)

- Elevated **glucose** and **triglycerides**
- High **alkaline phosphatase**
- Low **serum albumin**
- High **neutrophil percentage**
- Altered **lymphocyte percentage**

- **Low socioeconomic status** and **limited access to healthcare**

Understanding and addressing modifiable risk factors is essential for the primary and secondary prevention of stroke.

Retrospective observational studies with predictive modeling

In a Retrospective observational study with predictive modeling Wu et al. ¹⁾ wish to identify key clinical, biochemical, and socioeconomic risk factors for stroke and post-stroke depression, and to develop a reliable and explainable [prediction model](#) and scoring tool based on population-level data, in order to improve strategies for primary [stroke prevention](#).

This study, while superficially appealing due to its use of “modern” AI methods [Shapley Additive Explanations](#) (SHAP), ultimately collapses under the weight of [methodological overreach](#) and [interpretative overconfidence](#).

□ 1. Rhetorical Inflation in Purpose

The authors wish to “provide novel prevention strategies” for stroke and post-stroke depression using a retrospective cross-sectional dataset. This is a textbook case of [rhetorical inflation](#)—confusing statistical association with clinical innovation. National Health and Nutrition Examination Survey (NHANES) data are observational, not designed for [causal inference](#) or dynamic modeling. Yet, the conclusions sound as if they were derived from a [randomized trial](#) or [prospective cohort](#).

□ 2. Overfitting and Illusion of Precision

The model boasts an AUC of 82%. However:

No [external validation](#) cohort is presented.

[Internal validation](#) alone, especially on a subset of ~4,000 out of ~49,000 participants, is insufficient and prone to [overestimation](#).

SHAP values do not inherently confer [reliability](#) or [causality](#)—they only explain how a given model behaves, not whether it makes sense clinically.

Using SHAP on weakly curated variables from a non-stroke-focused dataset is like explaining how a broken compass points north.

□ 3. Biochemical Noise Masquerading as Insight

Variables such as [alkaline phosphatase](#), [albumin](#), [neutrophils](#), and [lymphocyte](#) percentage are included in the model, yet:

No mechanistic [rationale](#) is given.

No [stratification](#) by stroke subtype (ischemic vs. hemorrhagic) is made.

Their clinical [interpretability](#) is poor, turning the resulting “risk score” into a black box with false [credibility](#).

This is conceptual [ambiguity](#) in action: quantitative signals presented as meaningful without context, causality, or physiological grounding.

□ 4. False Equivalence Between Predictive and Preventive Value

The authors conflate model accuracy with [clinical utility](#). A scoring tool built on retrospective correlations does not equal a preventive instrument. [Prevention](#) requires prospective testing, behavioral change modeling, and context-aware implementation. None of that is attempted here.

Instead, they offer a DIY self-assessment tool from a population-level dataset — a tempting but misleading notion, potentially causing false [reassurance](#) or anxiety.

□ 5. Neglect of Stroke Complexity

Stroke is heterogeneous. It cannot be reduced to a one-size-fits-all model using NHANES data, which lacks imaging, timing, vascular anatomy, or medication data. No mention is made of key confounders such as:

Atrial fibrillation

Hypertension treatment

Prior antithrombotic use

This reflects sample simplification fallacy—where convenience trumps clinical complexity.

□ 6. No Contribution to Post-Stroke Depression Understanding

Despite including it in the title, the “post-stroke depression” angle is entirely speculative. The authors provide no validated psychiatric metrics, no follow-up data, and no proper modeling of mental health outcomes. It's a keyword, not a conclusion.

△ Conclusion: This study commits multiple scientific sins:

Rhetorical overreach (promising prevention from cross-sectional data)

Overfitted modeling with uncritical use of SHAP

Biochemical cherry-picking without clinical plausibility

False equivalence between correlation and clinical actionability

It is dressed as data science, but fundamentally lacks the anatomical, temporal, and pathophysiological depth required for meaningful stroke research.

A [flashy model](#) on shallow ground — predictive numerology disguised as [precision medicine](#).

1)

Wu B, Yu W, Zhang G, Jiang H, Chen Y, Wu N. Mining the risk factors for stroke occurrence and dietary protective factors based on the NHANES database: Analysis using SHAP. J Affect Disord. 2025 Jun 12:119671. doi: 10.1016/j.jad.2025.119671. Epub ahead of print. PMID: 40516626.

From:

<https://neurosurgerywiki.com/wiki/> - **Neurosurgery Wiki**

Permanent link:

https://neurosurgerywiki.com/wiki/doku.php?id=stroke_risk_factors

Last update: **2025/06/15 11:18**

