Investigating Stroke Risk Factors: A Cross-Cultural Network Analysis

A thesis
submitted in partial fulfilment
of the requirements for the degree
of
Master of Science (Research) in Psychology

at
The University of Waikato
by
Kelly Jane Olsen

THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

January 2026

Abstract

Stroke is the second leading cause of death and a major source of disability, with well-documented, complex risk factors. While stroke risk factors have been studied in their respective fields, research on how biological, psychological, and social factors interact remains limited, and no studies have explored whether patterns vary across ethnic groups. This study used network analysis to explore the unique associations and directional probabilities among known stroke risk factors across different ethnic groups. Data from 15,460 participants across 134 countries were collected via the Stroke Riskometer mobile application, classified into six ethnic groups (White/European, African, Asian, Indian, Latin American/Hispanic, Other). Bayesian Gaussian Graphical Models (BGGM) showed that 78-84% of associations were stable across ethnicities, indicating that most stroke risk relationships are consistent across cultures, with some unique differences. Associations between cardiovascular and non-modifiable risk factors of age, TBI and stroke history were identified cross-culturally, consistent with literature and confirming the viability of the networks. The Directed Acyclic Graph (DAG) revealed probable risk pathways stemming from non-modifiable risk factors such as age, stroke history, parental stroke, TBI, and sex, leading to psychological distress, with cardiovascular and lifestyle factors acting as powerful mediators, and resulting in poor memory and diabetes identified as terminal outcomes. These findings suggest that stroke risk develops through multiple biopsychosocial pathways. While non-modifiable risk factors have substantial upstream influence, the DAG identified modifiable factors, such as psychological distress, cardiovascular conditions, and lifestyle factors, that can be targeted for prevention. Stroke prevention should be multifaceted, reflecting the network's dynamic interconnections and hierarchical nature. Overall, these results highlight a diverse but largely consistent cross-cultural network of stroke risk factors.

Keywords: Stroke risk, biopsychosocial model, psychological distress, Network Analysis, Directed Acyclic Graph Analysis

Acknowledgements

Over the past two years of conducting and writing my master's thesis, I have received tremendous support and encouragement. This thesis is dedicated to my beautiful grandmother, Marie McRae, who sadly passed away in the final stages of writing. She was a clever writer and academic who always supported me, checked on me, and motivated me. She will be sorely missed and always loved.

I would like to thank my supervisor, Dr Oleg Medvedev, who has taught me new statistical techniques, provided feedback and been optimistic about my abilities. Your expertise and knowledge are why I was able to get this thesis completed, thank you. A big thank you to PhD student Rebecca Chalmers, who provided writing retreats and invaluable advice, always lending an eye or an ear when I needed it. What a wonderful friend to have!

This research would not have been possible without the participants worldwide who have used the Stroke Riskometer app. Thank you for your contribution to the research into stroke prevention.

My sincerest gratitude to the research team, led by Dr Valerie Feigin, who developed the app. Dr Feigen and the team at The National Institute for Stroke and Applied Neurosciences at Auckland University of Technology allowed me to use the data they have collected – I am eternally grateful and hope that my findings are helpful in your quest to improve stroke prevention.

I would not have been able to complete this thesis without the unconditional love and support from my partner, George and our best mate, Rudie the Vizsla, who always offer a hug and make me laugh when I need it most. Finally, my wonderful parents, siblings, nephews, niece, and friends who always cheer me on, listen to me on my bad days, and celebrate me on the good ones. Thank you for always supporting me and believing in me.

Table of Contents

Investigating Risk Factors of Stroke: A Cross-Cultural Network Analysis

Stroke is the second leading cause of death and a major contributor to adult disabilities worldwide (Feigin et al., 2023). The Global Burden of Stroke research indicates that over six million people died from a stroke in 2020 alone, and this figure is expected to increase by 50% by 2050 (Feigin et al., 2023). Many stroke survivors are left with significant cognitive, psychological, and physical impairments that require long-term care (Feigin et al., 2023). The economic burden is substantial, covering both direct healthcare and rehabilitation costs, as well as indirect costs like productivity losses (Rajsic et al., 2019). It is estimated that around 34% of total healthcare expenses are spent on stroke, including primary care, post-stroke rehabilitation, and ongoing follow-up with various specialists (Rochmah et al., 2021).

Adding to this, the psychological and physical effects of having a stroke are well documented, including long recovery periods, speech and mobility difficulties, as well as the onset of stress, anxiety, depression, and sometimes PTSD. Stroke-induced cognitive decline also negatively impacts the individual and their families (Barker-Collo et al., 2024; Devereux & Berns, 2023). Alarmingly, the incidence of stroke in individuals under 55 is increasing globally, with 60% of strokes affecting those under 75 and stroke incidence in 20- to 64-year-olds rising by 25% (Feigin et al., 2023; Parmar et al., 2015; Rochmah et al., 2021). The rise in younger people experiencing stroke is reportedly linked to increases in cardiovascular risk factors, diabetes, and obesity (Parmar et al., 2015; Rochmah et al., 2021). Additionally, over 86% of the stroke burden is impacting low- and middle-income countries, and the gap continues to widen between wealthier and poorer nations (Feigin et al., 2023).

There has been improved control of communicable diseases, advancements in healthcare, and people are living longer, with the global average life expectancy is expected to rise from 73.6 years in 2022 to 78.2 years by 2050 (Wise, 2024). This indicates that unless better surveillance, prevention, care, and rehabilitation measures are identified and put into practice, an increase in stroke-related death, disability, and economic burden is likely (Feigin et al., 2023; Parmar et al., 2015). Gaining a clearer understanding of how the major risk factors for stroke interact would enable a more precise assessment of individual stroke risk and be valuable in developing prevention strategies to lower that risk.

Stroke Types

A stroke is the rapid onset of disturbance in brain function, which is identified by objective imaging and pathological evidence of brain, spinal cord, or retinal cell death, with a vascular aetiology, presenting with or without clinical symptoms (Saini & Gurvendra, 2022). Simplified, a stroke is a loss of blood flow caused by a blockage or rupture of an artery in the brain, resulting in a lack of oxygen and the rapid death of brain, spinal cord, or retinal cells (Saini & Gurvendra, 2022). Clinical signs of stroke can impact consciousness, motor abilities, vision, perception, language, and cognitive function. This frequently manifests as hemispheric numbness and weakness in the face and limbs, confusion, speech changes or difficulty understanding speech, visual disturbance, dizziness, loss of coordination and balance or a severe headache (Saini & Gurvendra, 2022).

Strokes are generally divided into two main types. Ischemic strokes are the most common, accounting for around 80% of all reported strokes. These occur when a blood clot blocks or slows blood flow, depriving the brain, spinal cord, or retina of oxygen for more than 24 hours. A Transient Ischemic Attack (TIA), or mini stroke, is a shorter episode of similar symptoms that lasts for minutes to hours. Haemorrhagic strokes are the second type, with two subtypes: Subarachnoid haemorrhage (SAH) and Intracerebral haemorrhage (ICH). SAHs occur in the region around the brain and are caused by a vascular abnormality, a rupture to cerebral blood vessels, or an aneurysm. They make up approximately 5% of all strokes. ICHs happen when a weakened blood vessel in the brain ruptures, causing blood to leak and intracranial pressure to increase, which damages brain cells. ICHs account for about 10% of all strokes (Saini & Gurvendra, 2022).

Risk Factors of Stroke

A stroke risk factor increases the likelihood of stroke. Risk factors can be modifiable, meaning they can be reduced or controlled through lifestyle changes or medication, or non-modifiable, meaning they are beyond an individual's control and can progress at different rates (Boehme et al., 2017; Saini & Gurvendra, 2022). More emphasis is on modifiable risk factors, where individual or community action can reduce stroke occurrence, thereby reducing the burden of stroke at both personal and societal levels (Boehme et al., 2017).

Modifiable Risk Factors

Modifiable risk factors include psychological distress, some cognitive impairments or poor memory, lifestyle risk factors like BMI, alcohol intake, diet, exercise and smoking, as well as cardiovascular risk factors like high blood pressure, heart disease, enlarged heart and atrial fibrillation. Physical risk factors, such as diabetes, can also often be modified (Nindrea & Hasanuddin, 2023; Saini & Gurvendra, 2022). With adequate control of these risk factors, strokes could be prevented or delayed (Senff et al., 2025).

Psychological Distress. Psychological distress includes various mental health factors such as depression, work or personal stress, adverse life events, and locus of control (Barry et al., 2020; O'Donnell et al., 2010; Saini & Gurvendra, 2022). Research shows that females experience higher rates of distress than males, and distress levels tend to decrease with age, a pattern that is consistent cross-culturally, but impacts minority groups to a higher degree (Charles et al., 2023; Mirzaei-Alavijeh et al., 2025; Paradise et al., 2011; Tran et al., 2015). Systematic reviews have confirmed that psychological distress is strongly linked to adverse health outcomes like stroke, diabetes, immune dysregulation, and cancer (Barry et al., 2020). Additionally, psychological distress has been connected to more subjective memory complaints and significantly higher smoking rates compared to non-distressed individuals, possibly as a coping mechanism for distress (Moran, 2016; Paradise et al., 2011). A longitudinal study in a US sample by Sol et al. (2020) examined whether psychological factors contribute to ethnic differences in poor memory over time, considering baseline memory performance in White, Black and Hispanic older adults. Their findings identified that psychological distress, particularly depressive symptoms, was associated with poorer memory at baseline and an accelerated decline. However, ethnicity did not moderate these associations, indicating that psychological distress and memory decline operated similarly across ethnicities (Sol et al., 2020).

The INTERSTROKE study is a large, international case-control investigation designed to identify potential modifiable risk factors of first-ever stroke, utilising around 27,000 participants (O'Donnell et al., 2010). This study found that individuals who experienced psychological distress were more likely to have had a first stroke (O'Donnell et al., 2010). Moreover, the more severely depressed the person was, the greater their risk of stroke. This pattern was consistent across all countries and income levels (O'Donnell et al., 2010). These results have been supported by further cohort studies and meta-analyses, which suggest that depression and depressive symptoms increase the risk of stroke as well as the likelihood of developing heart disease (Dong et al., 2012; Harshfield et al., 2020; Pan et al., 2011).

More studies have utilised the INTERSTROKE data to explore further the relationship between stress and the occurrence of first-ever stroke (Reddin et al., 2022), as well as the different aspects of psychological distress and stroke (Henderson et al., 2013; O'Donnell et al., 2010). A study by Reddin et al. (2022) found that stress significantly increased the likelihood of a stroke, and individuals experiencing multiple life stressors had an even stronger association with first-time stroke occurrence. This effect was observed in both ischemic and haemorrhagic strokes, with a notably stronger association for haemorrhagic stroke (Reddin et al., 2022). Interestingly, people with a higher internal locus of control, which is the belief that they have control over their lives and outcomes through their actions, were less likely to have a stroke and exhibited a weaker connection between stress and stroke. These findings were consistent across different socioeconomic groups (Reddin et al., 2022). As well, Henderson et al. (2013) identified that depression, perceived stress, external locus of control (which is the belief that individuals' circumstances are outside of their own influence), and stressful life events were all independently linked to a greater risk of fatal and non-fatal stroke. Furthermore, they observed a dose-response relationship, whereby individuals with more indicators of distress faced gradually increasing stroke risk, with findings consistent across various ethnicities and socioeconomic backgrounds (Henderson et al., 2013).

Based on several meta-analytic and cohort studies, psychological distress when assessing stroke risk is a significant factor and has been identified as one of the ten simple risk factors that are linked to 90% of stroke risk (O'Donnell et al., 2010).

Poor Memory. While dementia, a non-modifiable risk factor, and stroke have a well-established relationship, subjective memory complaints not linked to dementia have also been recognised as a risk factor for stroke (Sajjad et al., 2015). A study by Sajjad et al. (2015) examined the link between subjective memory complaints and stroke in approximately 10,000 participants aged 55 and over. They excluded individuals who, at baseline, had already experienced a stroke, were diagnosed with dementia, or had both conditions. Participants were asked about memory issues and completed the Mini-Mental State Examination (MMSE), a neurocognitive test used to assess cognitive decline (Sajjad et al., 2015). The data included follow-up measures and tracked health changes, such as the occurrence of a first stroke or a diagnosis of dementia, from 1990 to 2012. They found that individuals reporting subjective memory complaints faced a higher risk of experiencing their first stroke compared to those without such complaints. They also observed that those with higher levels of education were particularly vulnerable, likely because highly educated individuals are more aware of subtle changes in their cognitive performance than those with lower levels of education (Sajjad et al., 2015). The authors suggest that subjective memory complaints as a risk factor for stroke are likely explained by vascular changes, with subjective memory serving as a marker for injuries in the brain's vascular tissue (Sajjad et al., 2015). The authors recommend that stroke risk researchers and clinicians incorporate a simple, self-rated question about memory problems to identify potential vascular risk associated with memory complaints (Sajjad et al., 2015).

Lifestyle Risk Factors. Several well-known lifestyle risk factors increase the likelihood of stroke (Boden-Albala & Sacco, 2000). These include a person's diet, exercise, Body Mass Index (BMI), alcohol consumption, and cigarette smoking (Boden-Albala & Sacco, 2000; Boehme et al., 2017; Saini & Gurvendra, 2022). The INTERSTROKE study found that lifestyle factors such as smoking, BMI, diet, and exercise make up four of the five variables responsible for 80% of global stroke risk, emphasising the significant impact lifestyle modifications can have on stroke risk (O'Donnell et al., 2010).

Studies have shown that diet plays a crucial role in preventing stroke, as well as affecting other known stroke risk factors such as physical activity levels, diabetes, BMI, and hypertension (Boden-Albala & Sacco, 2000; Guo et al., 2022). A recent umbrella review of 122 meta-analyses aimed at synthesising evidence on the link between dietary factors and stroke risk found that a diet rich in fruits, vegetables, whole grains, and dietary fibre offers significant protection against stroke (Guo et al., 2022). The same review also revealed that high consumption of red meat and sugars increases stroke risk (Guo et al., 2022). Additionally, studies have established a connection between high salt intake and elevated blood pressure (Saini & Gurvendra, 2022). Another study by Aigner et al. (2018) examined diet quality in a multiethnic cohort. They identified that lower diet quality was consistently associated with higher stroke death rate, and that by improving diet, 7.9% of stroke deaths could be prevented (Aigner et al., 2018). However, they noted that diet quality seems to contribute unequally to stroke risk across groups, with little to no association in Hispanic groups, and significantly higher in African American and White participants (Aigner et al., 2018). Another study from America considered how diet quality varies by age, sex, ethnicity, and socioeconomic factors and found that Hispanic people had higher overall diet quality than other ethnicities (Hiza et al., 2013). Diet quality followed a U-shaped pattern across the lifespan, with children and older adults eating higher quality diets than middle-aged adults (Hiza et al., 2013). Furthermore, a study from India examined diet patterns of patients with first-ever strokes (Durga & Manorenj, 2019). They found that, while Indian diets generally consist of a lot of legumes, grains, and vegetables, those who had strokes had diets characterised by high refined carbohydrate intake, low fruit and leafy greens, excessive salt, and higher meat and alcohol consumption (Durga & Manorenj, 2019).

Physical inactivity is well-established as being linked to many adverse health outcomes, including stroke, and it is known that physically active people have a lower risk of stroke and death after stroke (Saini & Gurvendra, 2022). Similar to diet, physical activity influences other known stroke risk factors such as lower BMI, reduced risk of premature death, hypertension, and heart disease (Boden-Albala & Sacco, 2000; Yu et al., 2025). Furthermore, regular moderate-to-high-intensity exercise has a protective effect against other risk factors, such as dementia and mild cognitive impairment, by slowing cognitive deterioration (Law et al., 2020). Regular exercise, even walking, is associated with a decreased risk of all strokes and other stroke risk factors like hypertension, making it an easily modifiable risk factor (O’Donnell et al., 2010).

A high BMI, or being obese, is linked to a greater risk of stroke, especially ischemic strokes (Saini & Gurvendra, 2022). Among all major stroke risk factors, obesity is seen as one of the most manageable, as it directly influences hypertension, diabetes, and heart diseases, which are other key stroke risk factors, and can be controlled through diet and exercise (Wang et al., 2022). A recent meta-analysis by Wang et al. (2022) found that having a BMI over 25, classified as overweight or obese, is associated with a higher risk of stroke, particularly ischemic stroke. Studies also show that central obesity (abdominal fat distribution) poses the greatest risk and may be a stronger indicator of increased stroke risk (Boden-Albala & Sacco, 2000).

Alcohol consumption and cigarette smoking are both well-established modifiable risk factors for stroke, often linked in the literature, both being physiologically and psychological addictive and use being shaped by cultural norms as well as (Boehme et al., 2017; Room, 2004). Interestingly, in some studies light to moderate alcohol intake of two drinks or fewer for men and one for women has shown protective effects against ischemic stroke but not haemorrhagic stroke (Boehme et al., 2017; Saini & Gurvendra, 2022). Heavy drinking is associated with hypertension, poor blood pressure control, and an increased risk of both ischemic and haemorrhagic stroke (Boehme et al., 2017; Smyth et al., 2023). Case-control studies using the INTERSTROKE data have found that there are substantial ethnic and regional differences in alcohol consumption and stroke risk (Smyth et al., 2023). An example is that while low intake is associated with reduced risk in Western ethnicities, it increases risk in Indian people (Smyth et al., 2023).

Cigarette smoking is a significant risk factor for all stroke types, with a well-established dose-effect relationship between pack:years and stroke risk, with smoking contributes to approximately 15% of stroke deaths annually (Boehme et al., 2017; Saini & Gurvendra, 2022). Those who quit smoking quickly reduce their stroke risk, with a notable decline in excess risk observed after two to four years post-cessation (Boehme et al., 2017). Studies from America have linked low education and neighbourhood disparity with higher smoking and lower cessation rates, impacting the African American, White and Latin/Hispanic populations (Nguyen-Grozavu et al., 2020; Ra et al., 2022). Furthermore, middle-aged people smoke more frequently, as well as men smoking at higher rates than women (Nguyen-Grozavu et al., 2020; Ra et al., 2022). The lowest cessation rates across education levels were among African American participants (Nguyen-Grozavu et al., 2020).

Diabetes. Diabetes Mellitus is a well-established risk factor for stroke, especially in ischemic cases, with evidence showing up to 33% of patients having diabetes (Mosenzon et al., 2023; Saini & Gurvendra, 2022). Studies have found that individuals with either type 1 or type 2 diabetes are 1.5 to 2 times more likely to experience a stroke than those without diabetes, and the risk increases further the longer diabetes remains unmanaged (Mosenzon et al., 2023). Some gender differences have been observed, with women who have diabetes facing a higher risk of stroke compared to men (Mosenzon et al., 2023; Saini & Gurvendra, 2022). Ethnic differences are also evident, for example diabetes is one of two most prevalent comorbidities for stroke risk and is often associated with hypertension (Ram et al., 2021). Diabetes contributes to stroke through vascular changes, including arterial plaque formation, microvascular dysfunction in the brain, and cardiac embolism (Mosenzon et al., 2023; Saini & Gurvendra, 2022). For those who have had a stroke, diabetics tend to have worse clinical outcomes, longer recoveries, larger areas of brain tissue affected, higher death rates, and a greater likelihood of stroke recurrence (Mosenzon et al., 2023).

While diabetes itself is not classified as a lifestyle risk factor, it can be influenced by lifestyle factors such as diet, exercise, changes in BMI, medication, and blood pressure management. As a result, diabetes is a highly modifiable risk factor for stroke that can be managed through behavioural, pharmacological, and lifestyle interventions (Mosenzon et al., 2023). People with pre-diabetes also face a higher risk of stroke, making early intervention crucial to reduce the likelihood of developing both stroke and diagnostically significant diabetes (Lee et al., 2012; Mosenzon et al., 2023).

Cardiovascular Risk Factors. Several modifiable cardiovascular conditions increase the risk of having a stroke (Saini & Gurvendra, 2022). Hypertension, or high blood pressure (HBP), has been identified in numerous studies as the most critical risk factor for all stroke types, with strong, linear relationships between HBP and stroke risk identified (Boehme et al., 2017). HBP is one of the five risk factors that account for approximately 80% of stroke risk (Boehme et al., 2017; O'Donnell et al., 2010; Saini & Gurvendra, 2022). It is important because of its modifiable nature, with research identifying a 2mmHg drop in systolic blood pressure as linked to a 25% reduction in stroke risk, and a reduction in diastolic blood pressure as linked to a 50% reduction in stroke risk (Saini & Gurvendra, 2022). Hypertension is connected to other risk factors such as psychological distress, diabetes, poor diet, alcohol consumption, high BMI or a sedentary lifestyle; therefore, behavioural adjustments to lifestyle choices can help lower blood pressure and reduce risk (Boehme et al., 2017; Kuwabara et al., 2018; Landi et al., 2018; Mosenzon et al., 2023; Saini & Gurvendra, 2022).

There is a protective factor associated with taking blood pressure medication; conversely, not taking blood pressure medication or taking it improperly can increase or worsen the risk of having a stroke – therefore, taking blood pressure medication is an important aspect of managing stroke risk (Boehme et al., 2017; Mosenzon et al., 2023; Saini & Gurvendra, 2022). Studies have shown that after an initial stroke, starting blood pressure control medication can decrease the incidence of recurrent strokes by 30 to 40%, further confirming the importance of blood pressure management (Nindrea & Hasanuddin, 2023).

Cardiovascular diseases, such as heart disease, atrial fibrillation, and left ventricular hypertrophy (an enlarged heart), are well-established risk factors for stroke. These risks arise from vascular changes, weakening of the heart and its function, the potential for clot formation or embolism, irregular heartbeat, narrowing of arteries, and plaque buildup (Boehme et al., 2017; Wolf et al., 1991; Xu et al., 2020). Heart disease includes conditions like ischaemic heart disease, heart failure, valve disorders, and artery diseases—all of which are known to increase the risk of all stroke types (Boehme et al., 2017; Saini & Gurvendra, 2022; Wolf et al., 1991). Research has found that first-time stroke survivors with a prior diagnosis of heart disease, especially ischaemic heart disease, face a higher risk of recurrent stroke (Nindrea & Hasanuddin, 2023).

Although atrial fibrillation is a type of cardiac disease, it has been recognised as an independent and significant risk factor for ischaemic, embolic, and non-embolic strokes, as well as mortality. Those with atrial fibrillation are five times more likely to experience a stroke and have a 2.2 times greater risk of recurrent stroke (Nindrea & Hasanuddin, 2023; Saini & Gurvendra, 2022).

Furthermore, while an enlarged heart has been identified as an independent risk factor for stroke, the underlying reasons remain somewhat unclear (Xu et al., 2020). A recent meta-analysis by Xu et al. (2020) aimed to better understand this risk factor and its implications. They found that an enlarged heart increases the risk of stroke, and notably, for each one centimetre increase in the left atrial diameter, the risk of stroke rises by 25% (Xu et al., 2020). The researchers hypothesise that an enlarged heart predisposes the development of atrial fibrillation and acts as a marker for hypertension and other heart diseases, all of which are highly associated with stroke risk. They also hypothesise that an enlarged heart promotes blood stasis, which heightens the chance of embolism (Xu et al., 2020).

Non-Modifiable Risk Factors

Non-modifiable risk factors include demographic details such as a person's age, gender, ethnicity, genetics, and a history of stroke or heart attack (Saini & Gurvendra, 2022).

Some neuropsychological factors are also non-modifiable, such as dementia, along with physical factors like Traumatic Brain Injury (Behl et al., 2024; Qu et al., 2022)

Age. Strokes in people aged 18 to 55 are less common and are considered early strokes, making up about 10-15% of all strokes (Potter et al., 2022). The risk of stroke rises with age, and it doubles every decade after 55, which means older people face a higher stroke risk (Boehme et al., 2017; Feigin et al., 2023; Saini & Gurvendra, 2022). A recent systematic review and meta-analysis looked at risk factors for recurrent stroke and found that, among non-modifiable factors, older age was the most significant contributing factor (Nindrea & Hasanuddin, 2023). The evidence shows most recurrent strokes happen in those aged 65 or over (Nindrea & Hasanuddin, 2023). As people get older, blood vessels develop plaque and become stiffer because of natural degeneration, which further raises the risk of stroke (Nindrea & Hasanuddin, 2023).

Sex. The Global Burden of Disease study examined sex disparities and trends in stroke incidence from 1990 to 2021 (Ashraf et al., 2025). The researchers found that, although there has been a decline in stroke incidence during this period, males still have slightly higher rates than females (Ashraf et al., 2025). Other studies also show that males experience a stroke rate 1.25 times higher than women, but since women generally live longer, more women die of stroke each year (Saini & Gurvendra, 2022). However, a recent review on sex and stroke risk factors found that women are at increased risk because they have higher rates of atrial fibrillation, hypertension, and obesity compared to men (Hanna et al., 2024). The review also indicated that women face greater risks of diabetic complications than their male counterparts in terms of stroke and other cardiovascular concerns (Hanna et al., 2024). Additionally, there are sex-specific factors that raise a woman’s risk of stroke, such as pregnancy, menarche and menopause, as well as the use of hormone replacement therapy and the oral contraceptive pill (Hanna et al., 2024). Other differences influencing stroke risk include the higher prevalence of smoking among men and sex-specific low testosterone levels in men (Hanna et al., 2024).

Heritability of Stroke and Stroke History. A family history of stroke raises the risk of stroke by 30%, with other non-modifiable factors like age and sex also impacting how heritable stroke may be (Boehme et al., 2017). Premature stroke patients are more likely to have a first-degree relative who has had a stroke, women are more likely to have a parent who has had a stroke, and children born to individuals who have had a stroke by age 65 are twice as likely to experience an ischemic stroke (Boehme et al., 2017; Potter et al., 2022). Additionally, experiencing a first stroke increases the chance of having subsequent strokes, especially if lifestyle and medical health changes to lower the risk are not made (Boehme et al., 2017). One recent observational study in Australia followed 6000 first-ever stroke survivors for two years and found that 17% of these individuals had either a recurrent stroke or a major cardiovascular event within that period, with most occurring within the first year (Dharan et al., 2023).

Dementia. Dementia and stroke share several non-modifiable risk factors, such as age, sex, genetics, and ethnicity, as well as many modifiable ones, including smoking, poor diet and exercise, alcohol consumption, diabetes, hypertension, obesity, and other cardiovascular risks (Behl et al., 2024; Gardener et al., 2015; Tack et al., 2025). These risk factors increase the chance of either or both conditions developing due to potential organic changes in brain tissue, linked to vascular dysfunctions, inflammation, or the presence of disease-specific proteins (Behl et al., 2024).

Dementia is a gradual and irreversible neurodegenerative disease that involves a decline in cognitive function. Depending on the type of dementia, progression can vary from rapid decline to step-wise changes, during which an individual may remain stable for an extended period (Behl et al., 2024; Warren et al., 2022). Because of its progressive and irreversible nature, it is regarded as a non-modifiable risk factor for stroke. Symptoms and warning signs include progressive memory loss, changes in executive function, language, thinking, and personality, leading to an inability to perform routine activities (Behl et al., 2024; Warren et al., 2022). While the causes of dementia are still being researched, organic changes to brain tissue and blood vessels are key factors in the development of plaques and tangles in neurons, resulting in brain atrophy, neurotoxicity, and synaptic loss (Behl et al., 2024; Warren et al., 2022). Most studies of dementia have been conducted on European participants; however, a study based in the United Kingdom investigated ethnic differences and found that similar risk factors for stroke were also risk factors for dementia (Mukadam et al., 2023). Furthermore, researchers found that known risk factors bring about a higher risk of dementia for Black and South Asian people, most notably in cardiovascular risk factors like hypertension, obesity, diabetes as well as age and being female (Islam et al., 2024; Mukadam et al., 2023). They also identified that dementia diagnosis rates are 22% higher among Black people in the UK when compared to their White counterparts, and that Black and South Asian dementia patients die sooner after diagnosis and younger (Mukadam et al., 2023).

Although the literature is limited, dementia, especially vascular dementia and Alzheimer's disease (AD), can increase the risk of stroke occurrence (Behl et al., 2024; Gardener et al., 2015). Researchers suggest that early cognitive impairment, whether due to AD or vascular dementia, may indicate organic vascular changes in the brain that raise the likelihood of strokes, with dementia-related brain changes potentially serving as a precursor to stroke (Gardener et al., 2015). Additionally, individuals with AD are more vulnerable to both ischemic and haemorrhagic strokes compared to those without AD (Behl et al., 2024). Preexisting dementia is also linked to worse outcomes after a stroke, including more severe and rapid cognitive decline, as well as higher rates of disability and death (Behl et al., 2024). There is substantial evidence that stroke can lead to post-stroke dementia within the first 5-10 years following the event (Behl et al., 2024). Consequently, a cyclical relationship exists between the two conditions.

Traumatic Brain Injury. Like stroke, TBI is a significant cause of mortality as well as disability, and the impacts of TBI can be lifelong, with individuals suffering from long-term physical, cognitive, and psychological concerns (Feigin et al., 2010; Qu et al., 2022). Furthermore, both are relatively common and impose economic and social burdens on individuals and society (Qu et al., 2022). While the available literature is limited, studies have suggested a clear link between having a TBI and then experiencing a subsequent first stroke, hypothesising that the vascular damage caused by TBI likely leads to blood vessel disruption and subsequent stroke (Esterov et al., 2023; Izzy et al., 2023; Qu et al., 2022; Sperl et al., 2022). Furthermore, similar to dementia, TBI also share many of the same risk factors, like vascular abnormality and has an increased risk of cardiovascular disease post-TBI (Izzy et al., 2023). Generally, TBI impact younger populations between 15 and 24 years old, often involving alcohol consumption and many mild TBI are not reported (Feigin et al., 2010). A recent meta-analysis and systematic review examined the associations between TBI and subsequent stroke, determined whether TBI increases stroke risk, and summarised the broader literature (Qu et al., 2022). The authors found that in over 2.2 million patients across six studies, TBI was indeed associated with an increased first incidence of both ischaemic and haemorrhagic stroke, with particular emphasis on haemorrhagic stroke (Qu et al., 2022). The authors highlight that patients with TBI need to be closely monitored and offered early and effective treatments for their brain injury, as well as ensuring that patients and their carers are aware of developing stroke symptoms and the importance of early intervention (Qu et al., 2022). There is also evidence of sympathetic nervous system changes after TBI, which can result in atrial fibrillation, regardless of the severity of TBI, and identifies TBI as a risk factor for cardiac arrythmias (Stewart et al., 2025).

Ethnicity. The differences in stroke risk factors and prevalence among various ethnic groups are well documented in the literature, with minority groups, indigenous communities, and low socioeconomic populations experiencing higher rates of stroke and its risk factors than European and white populations (Balabanski et al., 2024; Chen et al., 2014; Hajat et al., 2004; Howard et al., 2019; G. Howard, 2001; V. Howard, 2013; Potter et al., 2022; Saini & Gurvendra, 2022; Sur et al., 2024). These differences are believed to be linked to higher rates of modifiable risk factors in non-European populations compared to Europeans (Avezum et al., 2015; Thompson et al., 2022). For instance, studies have shown that African American and Latin cohorts have a greater prevalence of modifiable risk factors such as hypertension and diabetes (Avezum et al., 2015). Other studies have shown that major stroke risk factors, like cardiac conditions and lifestyle factors, are shared in Western and Asian populations; their effects differ (Chen et al., 2014). For example, smoking is more influential in Western cohorts, while high BMI, alcohol consumption and blood pressure are stronger in Asian cohorts (Chen et al., 2014). Research indicates that Black and African American individuals are between two and five times more likely to have a stroke and are at greater risk of experiencing a premature stroke, between the ages of 35 and 44, than white people (Potter et al., 2022).

Kittner et al. (2021) examined how stroke risk factors differ by ethnicity within an American population, finding that white participants had higher genetic risk factors, while smoking and socioeconomic issues, such as limited access to healthcare and insurance, were notably more prevalent among Black and Latin participants. A review by Sur et al. (2024) agreed that minority populations, including Black, Latin, and low-income groups, tend to have worse stroke outcomes, poorer access to care, and struggle more with controlling modifiable risk factors, all of which are heavily influenced by social determinants of health. These include poverty, lower levels of education, limited access to insurance, and racism (Sur et al., 2024). Specific disparities noted include higher rates of diabetes and more common, harder-to-control hypertension among Black and Latin individuals (Sur et al., 2024). They also highlighted that modern diabetes medications are under-prescribed to Black, Latin, and Asian populations compared to white populations, and that hypertension accounts for 37% of stroke risk in Black individuals versus 25% in white individuals.

Utilising Global Burden of Disease data over three decades, a study by Goyal et al. (2025) examined trends in incidence, prevalence, mortality, disability and attributable risk factors of stroke in the Indian population. They identified a marked increase in incidence, prevalence and disability despite a decrease in mortality, and significantly more individuals living with post-stroke impairments (Goyal et al., 2025). Cardiovascular and metabolic risk factors were the most dominant contributors to stroke in this population, as well as lifestyle risk factors of diet, alcohol and smoking (Goyal et al., 2025).

A recent study aimed to understand the prevalence of strokes in indigenous populations within countries with a high human development index, as stroke data is scarce in these communities, but health disparities are well-documented (Balabanski et al., 2024). Their review encompassed seven highly developed nations, including New Zealand, Australia, the USA, Canada, Norway, Sweden, and Singapore, and found that the incidence of stroke was generally higher among indigenous groups compared to non-indigenous populations, with the highest rates among Aboriginal and Torres Strait Islander Australians (Balabanski et al., 2024). Another key finding was that stroke rates were decreasing in non-indigenous populations, but not consistently in indigenous groups (Balabanski et al., 2024).

Feigin et al. (2025) studied stroke incidence, death, and disability outcomes in a multi-ethnic population in Auckland, New Zealand, across five time periods from 1981 to 2022 to understand trends over time. They found ethnic disparities consistent with previous research on indigenous populations, with younger stroke onset and 1.5 to 2 times higher stroke incidence in Māori and Pacific peoples compared to Pākehā (Feigin et al., 2025). Their results also showed that although stroke outcomes have significantly improved over the past 40 years, ethnic minorities still carry a disproportionate burden, emphasising the need for better, more targeted prevention strategies (Feigin et al., 2025). Other studies have similarly noted that Māori, Pacific peoples, and Asians have more risk factors, less access to acute stroke care facilities and are less likely to receive a high standard of care than their Pākehā counterparts (Thompson et al., 2022). Additional findings indicated that Māori experienced higher death rates one year after stroke (Thompson et al., 2022).

Based on the current literature review, the distinctive relationships between stroke risk factors and ethnicities have not been thoroughly explored, which this study aims to address.

The Biopsychosocial Model

The Biopsychosocial model was introduced by George Engel in 1977 as a response to his critiques of the biomedical model, which he considered reductionist and insufficient for understanding the whole process of disease and patient care (Engel, 2012). He saw the biomedical model as problematic because it focused solely on biological factors while overlooking psychosocial aspects, such as mental illness, which can provide valuable insights (Engel, 2012). He proposed the need for a model that combines biological (e.g., sex, disease processes), psychological (e.g., distress, depression, anxiety), and social (e.g., smoking, alcohol use, diet, ethnicity, poverty) dimensions of health to guide personalised diagnosis and treatment of individual concerns (Engel, 2012). This enables clinicians to develop a treatment plan that addresses biological factors as well as emotional, behavioural, and social challenges faced by the patient (Engel, 2012). Critics have raised concerns about applying this model in practice due to time pressures on clinicians, the complexity of integrating different aspects for each patient, and the fact that some clients may not require such in-depth consideration (Weston, 2005). Nonetheless, clinicians also recognise the importance of a balanced approach, utilising a patient-centred model that makes a biopsychosocial approach practical and adaptable (Weston, 2005).

A recent study by Hoorelbeke et al. (2025) explored how biopsychosocial factors interrelate and uniquely influence postpartum depression through network analysis, similar to the approach of the current study. They discovered that pre-pregnancy health (biological), stress exposure (psychological), and the intentionality of pregnancy (social) were directly associated with postpartum depression and concluded that this type of research supports targeted prevention and early detection for at-risk mothers (Hoorelbeke et al., 2025).

While this model has primarily been utilised in mental health and psychiatry, its popularity is increasing in other medical areas (Saxena et al., 2022). Saxena et al. (2022) recommended incorporating the biopsychosocial model into neurology. They provided case examples involving post-stroke epilepsy, migraine, post-concussion management, and functional movement disorders (Saxena et al., 2022). They highlighted the advantages of addressing psychosocial stress, financial concerns, insomnia, and the use of therapy to encourage lifestyle modifications related to alcohol consumption, smoking, diet, and exercise, which improved recovery and compliance with medication regimes (Saxena et al., 2022). Moreover, the biopsychosocial model has been referenced in literature to predict heart attacks (Tomás et al., 2023). Researchers analysed data from roughly 45,000 adults to explore how biological, psychological, and social factors collectively influence the risk of a heart attack (Tomás et al., 2023). They found that psychosocial factors like frailty and low psychological wellbeing, such as depression, along with cardiovascular risk factors, significantly affect the likelihood of heart attacks (Tomás et al., 2023). The authors concluded that adopting a holistic approach is vital in predicting and potentially preventing heart attacks and supported the view that prevention strategies should include targeting psychosocial wellbeing (Tomás et al., 2023).

The Biopsychosocial model posits that people's health should be viewed holistically, within their context, to support accurate diagnoses and treatment plans (Engel, 2012). It emphasises that understanding the interaction between risk factors for different individuals and groups is crucial (Engel, 2012). Sur et al. (2024) found that minority ethnic groups face worse social determinants of health, such as poverty, lower education levels, and limited insurance coverage, which deepen disparities and raise the risk of stroke. They highlight the need to tailor stroke prevention strategies for ethnic minority groups while recognising their social context (Sur et al., 2024). Concerning stroke, the biopsychosocial model has been applied in aftercare research; however, it has not been utilised to understand how to reduce first-time stroke occurrences (Saxena et al., 2022). Thus, the current study aims to adopt a holistic biopsychosocial approach to researching stroke risk among different ethnicities by examining the unique and complex interactions between well-known stroke risk factors across groups. It is important to acknowledge the differences that may exist among ethnic communities to provide appropriate care.

Measuring Stroke Risk

In a recent review of the global burden of stroke, Feigin et al. (2017a) emphasise that the key to reducing the prevalence of strokes worldwide lies in primary prevention. To prevent strokes from happening, stroke risk factors need to be accurately assessed; however, traditional primary prevention methods are underused and have not been effective enough (Feigin et al., 2017b, 2017a). Since the burden of stroke is greater in low- and middle-income countries, stroke risk measurement must be easily accessible and affordable across the globe (Feigin et al., 2023; Medvedev et al., 2021)

Over 4.8 billion people worldwide use modern technology devices and their applications, giving consumers better access to a wide range of products (Feigin & Norrving, 2014). Included in this are health-related apps, which, when developed by specialists, can offer personalised assessments and new ways to improve positive health outcomes for individuals and lessen the global burden of noncommunicable diseases, including strokes (Medvedev et al., 2021). This innovative approach has been used at the National Institute for Stroke and Applied Neurosciences (NISAN) at Auckland University of Technology, where a free, globally accessible mobile app called the Stroke Riskometer has been developed (Feigin & Norrving, 2014; Medvedev et al., 2021). The Stroke Riskometer was adapted from the Framingham Stroke Risk Score, an algorithmic tool used to estimate a person's 10-year stroke risk and was improved by adding seven additional known risk factors (Feigin & Norrving, 2014; Medvedev et al., 2021; O'Donnell et al., 2010; Wolf et al., 1991). These include lifestyle and psychological factors, such as diet, exercise, alcohol use, stress, ethnicity, and family or personal history of strokes, which have been previously described (Feigin & Norrving, 2014; O'Donnell et al., 2010). The app calculates absolute five- and 10-year risks and, importantly, provides a relative risk comparison for other users with similar demographic characteristics (Feigin & Norrving, 2014). This approach is effective in increasing motivation for users to understand their risk factors and explore ways to reduce them (Medvedev et al., 2021).

To assess its validity, Parmar et al. (2015) conducted validation studies using classical test theory, comparing the predictive validity of the Stroke Riskometer with two well-established measures: the QStroke and the Framingham Stroke Risk Score (Hippisley-Cox et al., 2013; Wolf et al., 1991). The study concluded that the Stroke Riskometer is as accurate as the measures it was compared against in predicting stroke occurrence (Parmar et al., 2015). Since the Stroke Riskometer was developed in New Zealand but has a global reach, it required cross-cultural validation (Medvedev et al., 2021). G-Theory, an advanced statistical technique, was utilised to examine the reliability and generalisability of stroke risk scores across cultural contexts in different samples (Medvedev et al., 2021). The study by Medvedev et al. (2021) found that the Stroke Riskometer demonstrated strong reliability and generalisability across countries and that, in its current format, it has optimal reliability for measuring stroke risk. This meant that no modifications of risk factor items were necessary, and the Stroke Riskometer can be used cross-culturally (Medvedev et al., 2021).

According to this literature, collecting data through an app like The Stroke Riskometer is useful for further research because it enables researchers to gather and analyse large datasets worldwide to develop more affordable and effective prediction and prevention strategies aimed at reducing stroke burden and prevalence (Feigin et al., 2023; Medvedev et al., 2021; Parmar et al., 2015).

Network Analysis

Network analysis is an advanced statistical technique that offers a visual depiction of the relationships between variables within a network, using statistical measures to capture these connections (Borsboom et al., 2021). It is a novel approach in psychological and health sciences, aligning well with the biopsychosocial model because of its ability to interpret complex and unique interactions among different factors, viewing them as a system of interconnected elements rather than as a result of a single condition (Chalmers et al., 2022). This approach allows biological, psychological, demographic, behavioural, and environmental factors to be represented and modelled as “nodes” within a network, with the strongest associations indicated by “edges” (lines connecting nodes), while also considering the influence of all other nodes in the network (Borsboom, 2017; Chalmers et al., 2022). This methodology has become vital in contemporary health research, as its findings can guide and inform clinical practice, health training, and targeted treatment strategies by addressing all factors in a holistic manner (Hevey, 2018; Thomann et al., 2023)

In psychological and health science research, several studies have explored various health issues from a biopsychosocial perspective using network analysis, as previously mentioned (Geenen & Dures, 2019; Gevers-Montoro et al., 2023; Hoorelbeke et al., 2025; Thomann et al., 2023; Tomás et al., 2023). A study by Chalmers et al. (2024) employed cross-cultural network analysis to examine links between protective and risk psychological factors affecting immune function in New Zealand, India, and Italy. Among other findings, they found that the psychological risk factor of anxiety was consistently associated with poor immunity (Chalmers et al., 2024). By using network analysis, the researchers could explore the complex relationships between these factors across cultures, which might not be possible with other methods. Additionally, Roemer et al. (2024) utilised BGGM analysis to investigate how the Big Five personality traits relate uniquely to the five facets of mindfulness across cultures. They developed specific hypotheses from the exploratory network and tested them in the confirmatory and cross-cultural networks, identifying four consistent patterns across all countries. Their results demonstrate that network analysis can reveal unique relationships among variables and highlight cross-cultural similarities and differences. Furthermore, their study shows how BGGM networks can clarify complex psychological structures beyond what traditional correlations and regressions can reveal (Roemer et al., 2024).

In a network, factors are represented as "nodes," while the lines connecting them are called “edges” (Borsboom et al., 2021; Chalmers et al., 2022). The colour of the edges indicates a positive or negative connection, while the strength of the connection is shown by the thickness of the edge, known as the “weight edge” (Borsboom et al., 2021). The weight edge reflects the statistical coefficient between two nodes, revealing the nature and strength of their association, typically ranging from -1 to +1 after accounting for all other factors in the network (Borsboom et al., 2021).

Network analysis models can be either non-directional or directional (Chalmers et al., 2022). A non-directional network, such as a BGGM, indicates the factors associated with each other and their strength but does not imply a cause-and-effect relationship (Chalmers et al., 2022). These models are commonly used in psychological research, especially in cross-sectional studies (Chalmers et al., 2022). Meanwhile, directional networks, such as a DAG, show which factor influences another and are useful for researchers trying to understand possible cause-and-effect relationships (Chalmers et al., 2022).

Bayesian Gaussian Graphical Models (BGGM) are a form of non-directional network analysis offering various benefits over similar methods (Williams, 2019). BGGM employs posterior distributions and Credible Intervals (CIs) for each partial correlation to estimate parameters (Williams, 2019). One key advantage of the BGGM model is its suitability for use with mixed data, regardless of distribution (Williams, 2019). Additionally, by using the "explore" and "estimate" functions of BGGM, researchers can calculate pairwise partial correlations while controlling for all other factors within a network. BGGM matrices only keep edges whose 95% credible intervals do not include zero, indicating statistically credible associations (Chalmers et al., 2022; Williams & Mulder, 2020)

BGGM uses predictability estimates for each node in the network to understand its explanatory power, that is, how much of a node's variance can be explained by other connected nodes in the network (Haslbeck & Waldorp, 2018). Predictability helps differentiate factors that are controlled by other factors within the network from those that operate independently or are affected by unmeasured factors outside the network (Haslbeck & Waldorp, 2018). This helps determine which factors might be useful targets for interventions to produce multiple health benefits through network relationships (Chalmers et al., 2022; Haslbeck & Waldorp, 2018). A predictability score ranges from 0 to 1, where 0 indicates the node cannot be predicted at all by its connected nodes, and scores close to 1 mean the node's value is highly explained by its connections (Haslbeck & Waldorp, 2018). Another advantage of predictability is that it is reported on an absolute scale, making it easier to interpret than centrality measures used in other forms of network analysis (Haslbeck & Waldorp, 2018).

In network analyses, BGGM is regarded as the gold standard in psychological research and tends to offer several advantages over its counterparts, such as LASSO networks (Epskamp & Fried, 2018; Williams, 2019). LASSO creates a sparse network by forcing weak edges to zero, retaining only the strongest connections, and is especially suitable for datasets where the number of nodes is large relative to the sample size, as it helps avoid overfitting (Epskamp & Fried, 2018; Williams, 2019). Instead of providing a single value for each connection, BGGM generates a range of possible values along with their probabilities. This enables researchers to evaluate uncertainty, as each edge has a credible interval indicating the precision or uncertainty of an estimate, avoiding a false sense of certainty (Borsboom et al., 2021; Epskamp & Fried, 2018). Additionally, BGGM utilises predictability over centrality (Epskamp & Fried, 2018). While centrality measures analyse the position of variables within the network, predictability assesses how well variables can be explained by their connections, offering more practical insights for intervention planning (Borsboom et al., 2021; Epskamp & Fried, 2018).

Directed Acyclic Graphs (DAGs) are directional network analyses that identify how one factor influences another within a network (Chalmers et al., 2022). DAGS differ from most directional networks because of their “acyclic” nature, meaning the network structure prevents feedback loops or cycles between factors (Chalmers et al., 2022; Heeren et al., 2021). This creates a probabilistic cause-and-effect relationship between factors, and although data must be longitudinal to imply causality, DAG networks offer hypotheses about causal relationships (Heeren et al., 2021).

DAG analyses are commonly used in psychological research as they account for uncertainty, optimise model fit, and improve reliability through bootstrapping procedures

(Heeren et al., 2021). R software uses hill-climbing algorithms to iteratively search for the best-fitting network structure by adding, removing, and reversing edges based on the Bayesian Information Criterion (BIC), providing a target goodness-of-fit score and determining which edges are present (Heeren et al., 2021). This process is stabilised through averaging across bootstrapped networks with 5,000-10,000 iterations, then only retaining edges that appear in at least 95% of samples, thereby reducing the risk of false-positive relationships (Chalmers et al., 2022, 2024). The BIC is calculated for each directional edge and signifies the overall importance within the model. The smaller the BIC value, the greater the model fit, identifying a more substantial contribution to the overall network structure. The direction of the majority percentage determines the edge's direction; for example, an edge from factor A to factor B in at least 60% of network resampling indicates the direction of probabilistic dependence and adds confidence (Chalmers et al., 2022; Heeren et al., 2020). While causation cannot be guaranteed, the Bayesian DAG offers a highly robust and interpretable model suitable for complex datasets where causal structures are unknown (Chalmers et al., 2022).

Based on the current literature review, no studies using network analysis have been conducted to explore the unique interrelationships of stroke risk factors across different cultures, nor has DAG analysis been utilised to examine potential causal relationships between factors within a holistic biopsychosocial framework.

Current Study

The burden of stroke is multifaceted, indicating significant personal, societal and financial impacts. While stroke risk factors have been identified and researched in their respective fields, research into the complex, interactive relationships between various biopsychosocial factors, including psychological, lifestyle, demographic, and physical risk factors, is almost nonexistent. No studies have examined whether unique patterns exist across ethnic groups that could inform population-based prevention strategies. Adopting a holistic approach, such as the biopsychosocial model, is essential to understanding how all risk factors influencing stroke occurrence interact and influence each other. Using network analysis, this study aimed to explore the unique associations among known stroke risk factors across different ethnic groups and to estimate probabilities of directional connections among these factors.

Method

Participants and Procedure

The current study utilised data from previous research, with approval from the Auckland University of Technology Ethics Committee (AUTEC Ref.#19/236) (Medvedev et al., 2021). An ethics research proposal was submitted to the University of Waikato ALPSS Human Research Ethics Committee, and the application was approved under the number FS2025-02 (Appendix A). All participants provided informed consent for their information to be used for research, and the data were anonymised. The current study included 15,460 participants from 134 countries, all of whom completed the Stroke Riskometer Questionnaire through the mobile app. To have adequate sample sizes for Network Analysis and to understand unique relationships in different groups, participants were categorised by their disclosed ethnicity. Table 1 demonstrates the demographic information of the participants. There were six ethnic categories: White/European (n = 10,995), African (n = 953), Asian (n = 1390), Indian (n = 814), Latin American/Hispanic (n = 540) and other ethnicities (n = 768). Participants ranged in age from 20 to 90 (M = 46.11, SD = 14.63). There were 7,912 (51.2%) male participants and 7,548 (48.8%) female participants.

Table 1

Table of Number, Mean Age, and Sex in Each Ethnic Group Used for Network Analysis

Ethnicity	*n*	%	Age *(M/SD)*	Male (%)
White/European	10995	71.1	47.97 / 14.27	5346 (48.6)
African	953	6.2	48.96 / 15.46	429 (45.0)
Asian	1390	9	38.70 / 12.93	857 (61.5)
Indian	814	5.3	40.29 / 14.81	654 (80.3)
Latin American/ Hispanic	540	3.5	40.87 / 13.08	306 (56.7)
Other	768	5	39.29 / 13.51	320 (41.7)
Total	15460	100	46.11 /14.63	7912 (51.2)

Measure

Data for this study were collected using the Stroke Riskometer mobile app-based questionnaire, which contained 21 assessment items (Parmar et al., 2015). These items can be categorised by Lifestyle Factors (Smoking, Alcohol use, Diet, Exercise and BMI), Psychological Health Factors (Distress, Dementia, and Memory Deficit), Physical Health Factors (Heart Disease, Enlarged Heart, Atrial Fibrillation, Blood Pressure, Blood Pressure Medication, Diabetes, Traumatic Brain Injury and Stroke History), and Demographic Factors (Parental Stroke, Age, Sex and Ethnicity). The demographic items are considered non-modifiable factors, including age, sex, ethnicity, family history and genetic factors. Items use categorical response options of two (0-1) (Item 4: “Have you experienced significant emotional stress or depression in the past year?” – yes, no) or three (0-2) (Item 8: Have you ever been told by a doctor that you have diabetes?” yes over 12 months ago, yes less than 12 months ago, no), and some utilise a cut-off point. An example is item 12, “What is your systolic blood pressure?”, where participants with a blood pressure of less than 120 are recoded to 0, and those with a blood pressure higher than 120 are recoded to 1. The Stroke

Riskometer App then computes five- and ten-year total stroke risk estimates based on the user's data.

Data Analyses

Before formal analysis, 465 missing values were identified in the dataset (0.12% of the entire dataset). Given the size of the dataset and the mix of categorical and numerical variables, a tailored, cautious imputation was considered the most appropriate statistical method for handling missing data. Numeric variables, such as height, weight, BMI, age, and blood pressure, used the column median for imputation. In contrast, categorical variables (including binary variables) used the column's mode. The imputation was performed using the SimpleImputer from Scikit-Learn (Pedregosa et al., 2011), which carries out univariate imputation to replace missing data based on the predicted mode and median values. Descriptive statistics were calculated in IBM SPSS v.29.

Network analyses were conducted using R Studio software (Version 2023.12, R Core Team, 2023). To examine the unique associations between various known stroke risk factors across different ethnic groups, Bayesian Gaussian Graphical Models (BGGMs) were utilised. To assess the overall potential causal links among stroke risk factors, Directed Acyclic Graphs (DAGs) were employed. The BGGM approach offers a detailed analysis of the unique relationship between each pair of variables while considering all other variables in the network. It is also suitable for large datasets with multiple variables and visually depicts relationships with nodes representing variables and edges indicating significant connections (Chalmers et al., 2022). Blue edges represent positive relationships, while red edges indicate negative ones. Thicker, darker lines signify stronger relationships, and thinner, lighter lines represent weaker ones. The DAG approach estimates directional probabilities between nodes based on probabilistic dependency, with a detailed explanation provided for this specific analysis below (Chalmers et al., 2022; Heeren et al., 2020).

The models used the following well-known stroke risk factors: Lifestyle (smoking, alcohol use, diet, exercise, and BMI), Psychological Health (distress, dementia, poor memory), Physical Health (heart disease, enlarged heart, atrial fibrillation, high blood pressure, blood pressure medication, diabetes, traumatic brain injury, stroke history), and Demographic factors (parental stroke, age, and sex). These were included as nodes in the models (Feigin, 2013). All analysis code completed in R Studio is available in Appendix B.

BGGM Network Analyses

To evaluate whether the relationships between stroke risk factors were stable, exploratory (labelled ensample) and confirmatory (labelled cnsample) network analyses were performed using the largest ethnic group (White/European), which was divided into two equal parts. Exploratory network analysis was conducted on the first half of the sample (n = 5497) to investigate network connections among stroke risk factors and to develop hypotheses in a format similar to that of the Roemer et al. (2024) study. After identifying patterns, they were tested using confirmatory network analysis on the second half of the sample (n = 5498). Only confirmed node relationships present in both exploratory and confirmatory networks were considered genuine. To determine significant differences between the exploratory and confirmatory networks, 5000 posterior estimates with 95% credible intervals (CI) were utilised to compare mean differences for each node-to-node connection. Conducting exploratory and confirmatory analyses with random split data helps prevent issues related to post-model-selection inference (Faraway, 2016).

Cross-cultural networks for African, Asian, Indian, Hispanic, and Other ethnic groups were estimated and compared to the 'ensample' to determine if the established relationships observed in the White/European samples were replicated in other ethnicities. The *explore* function in the BGGM package in RStudio estimated partial correlations to identify the unique association between each node pair in each network. Only partial correlations with

95% CIs that did not contain zero were considered significant and included in the matrix. Separate networks were plotted for each ethnicity using the *qgraph* function in R. The layout was averaged across the seven networks (ensample, cnsample, African, Asian, Indian, Hispanic, and Other) using the Fruchterman-Reingold algorithm (Fruchterman & Reingold, 1991) to facilitate visual comparison. All comparisons between ethnicities were based on 5000 posterior samples to ensure robustness. Any 95% CIs that included zero indicated no significant difference between the two networks.

Predictability

The *predictability* function in R was used to evaluate the predictability of each node in the networks through the Bayesian predictability estimate R^2 value. Bayesian R^2 estimates the proportion of variance in a node that can be explained by the variance in all other nodes across the network (Chalmers et al., 2022; Haslbeck & Waldorp, 2018). Predictability is essential for understanding how interconnected a node is and how much influence one node has on the others in a network. It also helps in designing interventions (Chalmers et al., 2024).

Predictability analysis shows which risk factors are most closely linked within the network and which might be influenced by factors outside the measured variables. For interpretation, predictability values were categorised as high ($R^2 \geq 0.40$), moderate ($0.20 \leq R^2 < 0.40$), and low ($R^2 < 0.20$). Risk factors with high predictability are seen as heavily influenced by other measured variables, suggesting potential intervention points where targeting connected factors could have a notable effect. Conversely, nodes with low predictability indicate risk factors likely affected by unmeasured variables outside the network.

Directed Acyclic Graph (DAG) Analysis

A DAG was constructed for the entire sample to visually display the direction of relationships between stroke risk factors. Combining the DAG findings with nondirectional BGGM networks provides a more comprehensive interpretation of the connections between variables (Kokou-Kpolou et al., 2023). Similar to the BGGM analyses, the DAG includes nodes representing each variable as well as directed arrows and edges indicating possible causal pathways between predictor (ancestor) and outcome (descendant) nodes (Rohrer, 2018).

The hill-climbing algorithm in the bnlearn package in R was used to calculate the best-fitting model for the network. This algorithm iteratively adds, removes, and reverses edges until a target goodness-of-fit score, specifically the Bayesian Information Criterion (BIC), is achieved (Scutari & Nagarajan, 2013). The stability of the DAG was evaluated using 5000 bootstrapped samples to generate the final network, which was then visually plotted with the Rgraphviz package (Hansen et al., 2024; Scutari & Silander, 2024). Edges that exceeded the significance threshold established through bootstrapping were retained, highlighting the most stable and probable connections (Scutari & Nagarajan, 2013). In the DAG, arrows indicate the likely direction of prediction, with the predictor (ancestor) variable at the start and the outcome (descendant) variable at the end. The thickness of edges and arrows reflects the strength of the prediction (Kokou-Kpolou et al., 2022). Both strength and direction have a maximum of 1. A strength of 1 means that the edge between two variables was recovered in 100% of the 5000 bootstrapped samples. A direction of 1 means that every time the edge appeared, it appeared in the same direction, i.e. A → B. Age, sex, TBI, history of stroke, and parental stroke were blacklisted because they are non-modifiable variables that are unlikely or impossible to change. Therefore, they could not serve as outcome or descendant variables in the DAG.

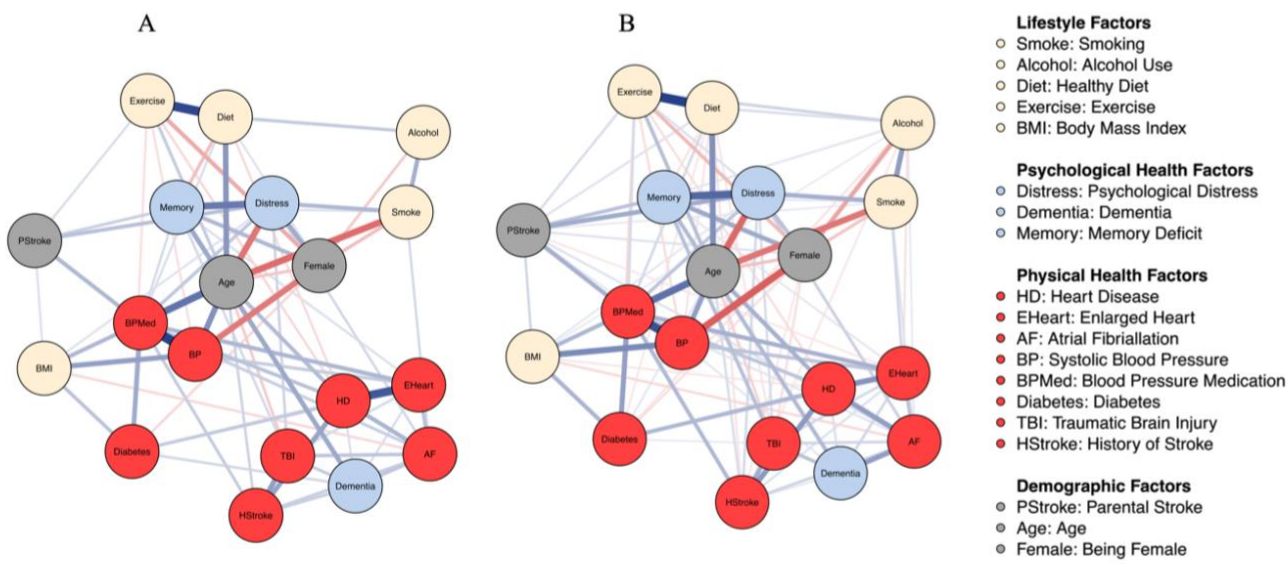
Results

Initial Exploration of Hypotheses

A random half of the White/European cohort (n = 5497) was used to estimate an exploratory network, presented in Figure 1A. Due to the large number of nodes in the network, only r-values greater than 0.100 are described. To guide the confirmatory and cross-cultural analyses, a set of hypotheses was developed based on patterns observed in the exploratory European network. These hypotheses were structured into three clusters: psychological, physical health, and lifestyle.

Figure 1

Exploratory (A) and Confirmatory (B) Network Analyses of European Ethnic Groups for Stroke Risk Factors



Psychological Cluster

Based on the exploratory network, psychological distress was positively associated with poor memory (r = 0.186) and poor memory was positively associated with dementia (r = 0.120). Distress was also negatively associated with age (r = -0.168) and positively associated with being female (r = 0.116). Based on these findings, it was hypothesised that

(H1A) psychological distress will show a positive association with poor memory, (H1B) poor memory will show a positive association with dementia, (H1C) psychological distress will be negatively associated with age, and (H1D) psychological distress will be positively associated with being female.

Physical Health Cluster

A clustering among cardiovascular and metabolic risk factors was identified. High blood pressure showed positive associations with blood pressure medication use (r = 0.267), high BMI (r = 0.133), and age (r = 0.153) and was less common among females (r = -0.163). Use of blood pressure medication also showed positive associations with an enlarged heart (r = 0.101), diabetes (r = 0.118), and age (r = 0.194). Additionally, heart disease, an enlarged heart and atrial fibrillation formed a connected cardiovascular grouping (r = 0.084 – 0.234). A history of having a stroke was associated with heart disease (r – 0.126) and TBI (r – 0.134). Therefore, it is hypothesised that (H2A) high blood pressure will be positively associated with taking blood pressure medication, (H2B) high blood pressure will be positively associated with higher BMI, (H2C) older age, and (H2D) will be less common among females. (H2E) Taking blood pressure medication will be positively associated with diabetes and (H2F) older age. (H2G) A history of stroke will be positively associated with heart disease, and (H2H) A history of stroke will be positively associated with traumatic brain injury. (H2I) Heart disease and atrial fibrillation will be positively associated, as well as (H2J) heart disease and an enlarged heart.

Lifestyle Cluster

Among lifestyle factors, diet and exercise were strongly and positively associated (r = 0.257), as were diet and age (r = 0.144). Smoking was associated with alcohol consumption (r = 0.109) and age (r = -0.178). Exercise was negatively associated with being female (r = -0.102). Based on this, it is hypothesised that (H3A) exercise and a healthy diet will be positively associated, (H3B) smoking and alcohol consumption will be positively associated, (H3C) smoking will be negatively associated with age, (H3D) a healthier diet will be positively associated with age, and (H3E) exercise will be negatively associated with being female.

Confirmatory Test of Hypotheses in White/European Sample

Of all associations found in the exploratory group, only 10 node pairs had credible intervals that did not include zero and could not be confirmed as reliable or stable in the confirmatory group. These included BMI and memory, smoking and enlarged heart, heart disease and enlarged heart, dementia and atrial fibrillation, memory and atrial fibrillation, dementia and parental stroke, heart disease and diabetes, diabetes and TBI, blood pressure and history of stroke, and alcohol and being female.

The hypotheses mentioned above were tested in the confirmatory network analysis (Figure 1B). The results showed that all hypotheses were supported. In respect to Hypothesis 2J, a significant difference in credible intervals between heart disease and an enlarged heart was identified, CL [0.05, 0.13]. However, on visual inspection and in terms of *r*-value, a moderate to strong relationship was observed in both the exploratory and confirmatory networks, and so the hypothesis was tested in other cohorts.

Cross-Cultural Test of Hypotheses and Unique Findings

Hypotheses H1A–H3E were tested across the African, Asian, Indian, Hispanic/Latin, and Other cohorts. The estimated networks for each ethnicity are presented in Figures 2-4, with averaged network layouts. Visual graph inspection, association magnitude, and credible intervals were used to test the hypotheses. Of all hypotheses, H2A, H2C, H2F, H2H, H2I and H2J were supported across all cohorts. (H2C), High blood pressure will be positively associated with older age, and this association was stronger in magnitude in the African cohort CL [-0.26, -0.13]. (H2H), A history of stroke will be positively associated with TBI, was stronger in magnitude in the African CL [-0.19, -0.06] and Indian cohort CL [-0.22, -0.08]. (H2I), Heart disease and atrial fibrillation will be positively associated, was stronger in magnitude in the Indian cohort CL [-0.24, -0.10] and Asian cohort CL [-0.24, -0.13].

Across the BGGM networks, there was cross-cultural variation in whether the hypotheses were supported. However, when comparing the exploratory groups with other ethnicities, 77% to 84% of associations were found to be stable. Below are the unique findings for each ethnic group.

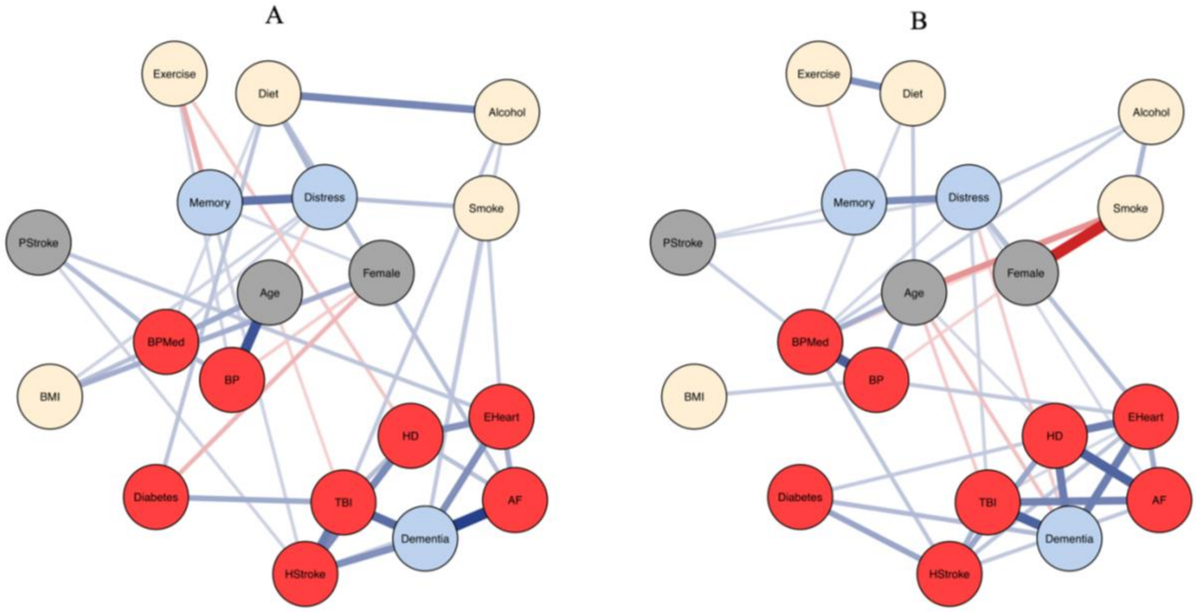
African Cohort

Unique hypotheses supported in the African cohort include (H1A) psychological distress will show a positive association with poor memory (r = 0.215), with a significantly stronger magnitude than the exploratory group CL [-0.16, -0.03], and (H1C) psychological distress will be negatively associated with age (r = -0.090). Physical hypotheses supported include (H2D) high blood pressure will be less common among females (r = -0.091), and (H2G) a history of stroke will be positively associated with heart disease (r = 0.243), with a significantly stronger magnitude than the exploratory group CL [-0.18, -0.04]. No lifestyle hypotheses were supported in the African cohort. H1B, H1D, H2B, H2E and H3A-E were not supported.

The strongest associations of confirmed nodes in the African network include: Age and high blood pressure, TBI and dementia, memory and distress, history of stroke and TBI, diet and alcohol use and heart disease and history of stroke (r = 0.243 to 0.355).

Figure 2

African Ethnicity (A) and Asian Ethnicity (B) Network Analyses for Stroke Risk Factors



Other significantly different node-to-node results of note include diabetes and being female, BMI and being female, dementia and history of stroke, heart disease and TBI, dementia and TBI, alcohol use and TBI, diet and diabetes, enlarged heart and parental stroke, enlarged heart and atrial fibrillation, dementia and atrial fibrillation, diet and atrial fibrillation, dementia and enlarged heart, exercise and memory, alcohol use and dementia, smoking and dementia, diet and distress, alcohol use and diet.

Asian Cohort

Unique hypotheses supported in the Asian cohort include (H1A) psychological distress will show a positive association with poor memory (r = 0.215), (H1D) psychological distress will be positively associated with being female (r = 0.093). Physical hypotheses supported include (H2B) high blood pressure will be positively associated with higher BMI (r = 0.101) and (H2G) a history of stroke will be positively associated with heart disease r = 0.170). Lifestyle hypotheses supported include (H3A) exercise and a healthy diet will be positively associated (*r* = 0.211), (H3B) smoking and alcohol consumption will be positively associated, (*r* = 0.137), (H3C) smoking will be negatively associated with age, (*r* = -0.181), (H3D) a healthier diet will be positively associated with age (*r* = 0.100). H1B, H1C, H2E and H3E were not supported.

The strongest associations of confirmed nodes in the Asian cohort include: Being male and smoking, high blood pressure and taking blood pressure medication, TBI and dementia, heart disease and atrial fibrillation, dementia and heart disease and dementia and enlarged heart (*r* = -0.362 to 0.243).

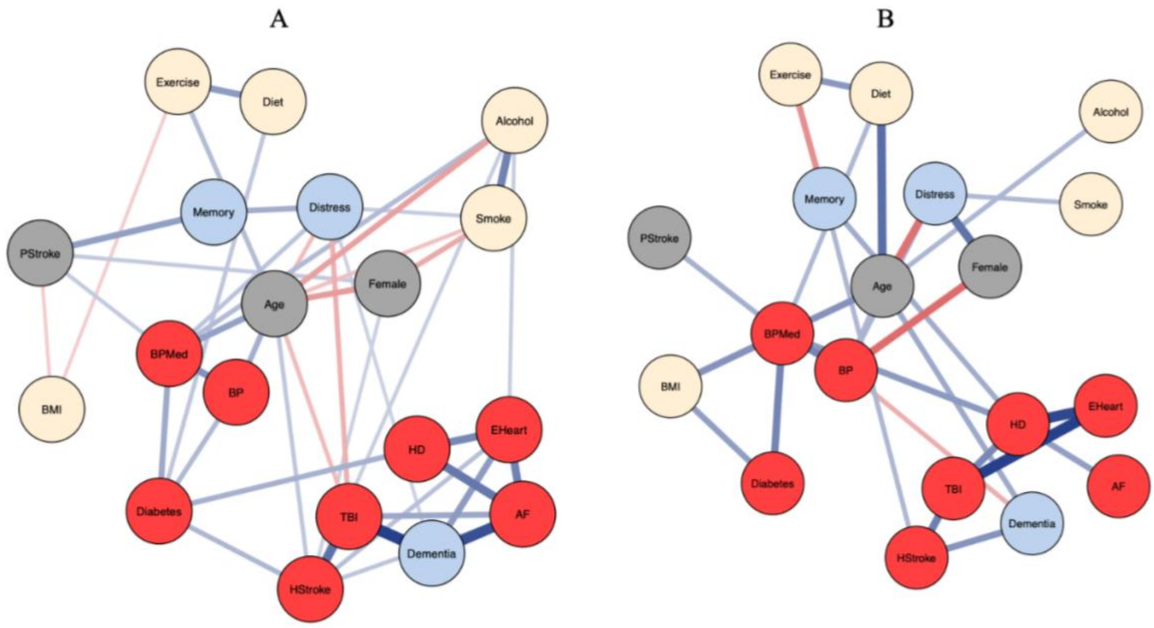
Other significantly different node-to-node results of note include dementia and age, enlarged heart and history of stroke, atrial fibrillation and TBI, dementia and diabetes, enlarged heart and atrial fibrillation, and distress and enlarged heart.

Indian Cohort

Unique hypotheses supported in the Indian cohort include (H1A) Psychological distress will show a positive association with poor memory (*r* = 0.168), and (H1C) Psychological distress will be negatively associated with age (*r* = -0.118). Physical hypotheses supported include (H2E) taking blood pressure medication was positively associated with diabetes (*r* = 0.176), with a significantly stronger magnitude than the exploratory group CL [-0.22, -0.07]. Unique lifestyle hypotheses supported in the Indian cohort include (H3A) exercise and a healthy diet will be positively associated (*r* =0.205), (H3B) smoking and alcohol consumption will be positively associated (*r* = 0.253), with a significantly stronger magnitude than the exploratory group CL [-0.21, -0.07], and (H3C) smoking will be negatively associated with age (*r* = -0.111). H1B, H1D, H2B, H2D, H2G, H3D and H3E were not supported.

Figure 3

Indian Ethnicity (A) and Hispanic/Latin Ethnicity (B) Network Analyses for Stroke Risk Factors



The strongest associations of confirmed nodes in the Indian cohort include: TBI and dementia, TBI and a history of stroke, heart disease and atrial fibrillation, atrial fibrillation and an enlarged heart, alcohol use and smoking and high blood pressure and blood pressure medication (r = 0.221 to 0.388)

Other significantly different node-to-node results of note include age and being female, parental stroke and being female, alcohol use and age, diabetes and history of stroke, an enlarged heart and history of stroke, alcohol use and history of stroke, atrial fibrillation and TBI, dementia and TBI, distress and TBI, diet and diabetes, memory and parental stroke, alcohol use and taking blood pressure medication, dementia and atrial fibrillation, dementia and enlarged heart, alcohol use and enlarged heart.

Latin/Hispanic Cohort

Unique hypotheses supported in the Latin/Hispanic cohort include (H1C) Psychological distress will be negatively associated with age (r = -0.194), and (H1D)

Psychological distress will be positively associated with being female (*r* = 0.219), which was significantly stronger in magnitude than the exploratory group CL [-0.19, -0.02]. Physical hypotheses supported include (H2D), high blood pressure was less common among females (*r* = -0.199), and (H2E), taking blood pressure medication will be positively associated with diabetes (*r* = 0.190). Lifestyle hypotheses supported include (H3A) exercise and a healthy diet will be positively associated (r = 0.161), but with a significantly weaker magnitude than the exploratory group CL [0.01, 0.18]. (H3D) A healthier diet will be positively associated with age, was also supported (*r* = 0.226). H1A, H1B, H2B, H2G, H3B, H3C and H3E were not supported.

The strongest associations of confirmed nodes in the Latin/Hispanic cohort include: An enlarged heart and TBI, diet and age, distress and being female, high blood pressure and being male, TBI and history of stroke and distress and age (*r* = -0.194 to 0.296).

Other significantly different node-to-node results of note in the Latin/Hispanic cohort include dementia and age, alcohol use and age, smoking and age, poor memory and a history of stroke, dementia and a history of stroke, an enlarged heart and TBI, heart disease and TBI, dementia and high blood pressure, poor memory and high blood pressure, poor memory and heart disease, exercise and memory.

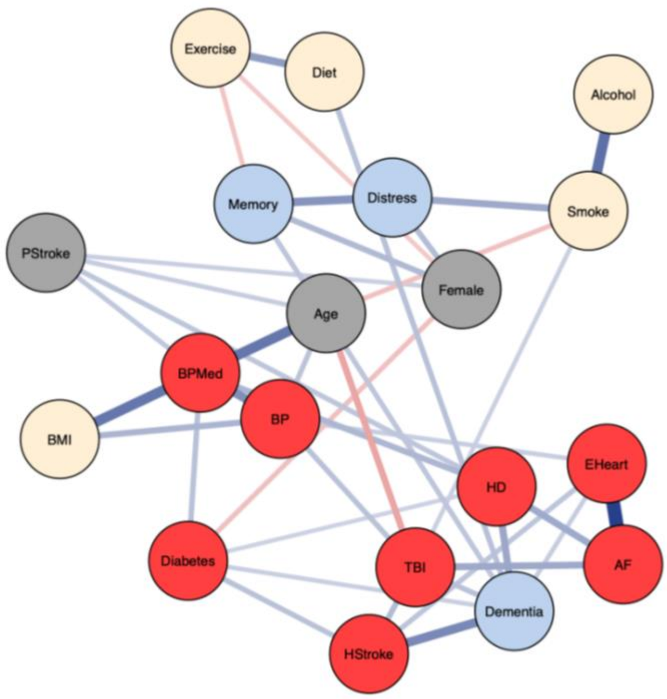
Other Cohort

Psychological hypotheses supported in the Other ethnicity cohort include (H1A) Psychological distress will show a positive association with poor memory (*r* = 0.225), (H1B) Poor memory will show a positive association with dementia (*r* = 0.122) and (H1D) Psychological distress will be positively associated with being female (*r* = 0.140). Physical hypotheses supported included (H2B) high blood pressure was positively associated with higher BMI (*r* = 0.157) and (H2E) taking blood pressure medication was positively associated with diabetes (*r* = 0.131). Supported lifestyle hypotheses included (H3A) exercise and a healthy diet were positively associated (r = 0.205), (H3B) smoking and alcohol consumption were positively associated (r = 0.303) and were significantly stronger in magnitude than the exploratory group CL [-0.26, -0.12]. (H3C) smoking was negatively associated with age (r = -0.116), and (H3E) exercise was negatively associated with being female (r = -0.101). H1C, H2D, H2G and H3D were not supported.

The strongest associations of confirmed nodes in the Other ethnicity cohort include: Atrial fibrillation and an enlarged heart, alcohol use and smoking, age and BMI, history of stroke and dementia, BMI and using blood pressure medication and poor memory and distress (r = 0.225 to 0.405).

Figure 4

Other Ethnicity Network Analyses for Stroke Risk Factors



Other significantly different node-to-node results of note in the Other cohort include high blood pressure and being female, TBI and age, parental stroke and age, BMI and age, diabetes and a history of stroke, an enlarged heart and history of stroke, dementia and history of stroke, high blood pressure and TBI, atrial fibrillation and TBI, heart disease and parental stroke, enlarged heart and taking blood pressure medication, an enlarged heart and atrial fibrillation, dementia and heart disease, exercise and poor memory, diet and dementia, smoking and distress, smoking and alcohol use.

Predictability

Predictability analyses were conducted to assess which variables were most strongly predicted by other network variables. Figures 5-8 show predictability for all cohorts.

Across the exploratory and confirmatory networks, blood pressure medication, age, and high blood pressure show the highest predictability, with R^2 values of 0.25–0.30. In contrast, alcohol use, parental stroke, and dementia had the lowest predictability, making them less predictable by other variables in the network, meaning they may be influenced more broadly, with less direct connection within the network or being influenced by factors outside those measured.

In the African network, dementia, atrial fibrillation, and traumatic brain injury showed the highest predictability ($R^2 \approx 0.55 – 0.63$). Being female, exercise, and a high BMI had the lowest predictability.

In the Asian network, heart disease, dementia, and an enlarged heart had the highest predictability ($R^2 \approx 0.46 – 0.50$), with high BMI, exercise, and parental stroke showing the lowest predictability.

In the Indian group, atrial fibrillation, dementia and traumatic brain injury showed the highest predictability ($R^2 \approx 0.76 – 0.78$). Poor memory, exercise, and high BMI were the lowest predictive factors in the Indian cohort.

In the Latin/Hispanic group, age, taking blood pressure medication and having heart disease showed the highest predictability ($R^2 \approx 0.31$), while atrial fibrillation, smoking and alcohol had the lowest predictability.

In the Other ethnic group, atrial fibrillation, an enlarged heart and dementia emerged with the highest predictability ($R^2 \approx 0.36 - 0.40$). Being female, exercise and diet were the least predictable.

Figure 5

Predictability of Nodes for Exploratory (A) and Confirmatory (B) Networks for White/European Ethnic Group

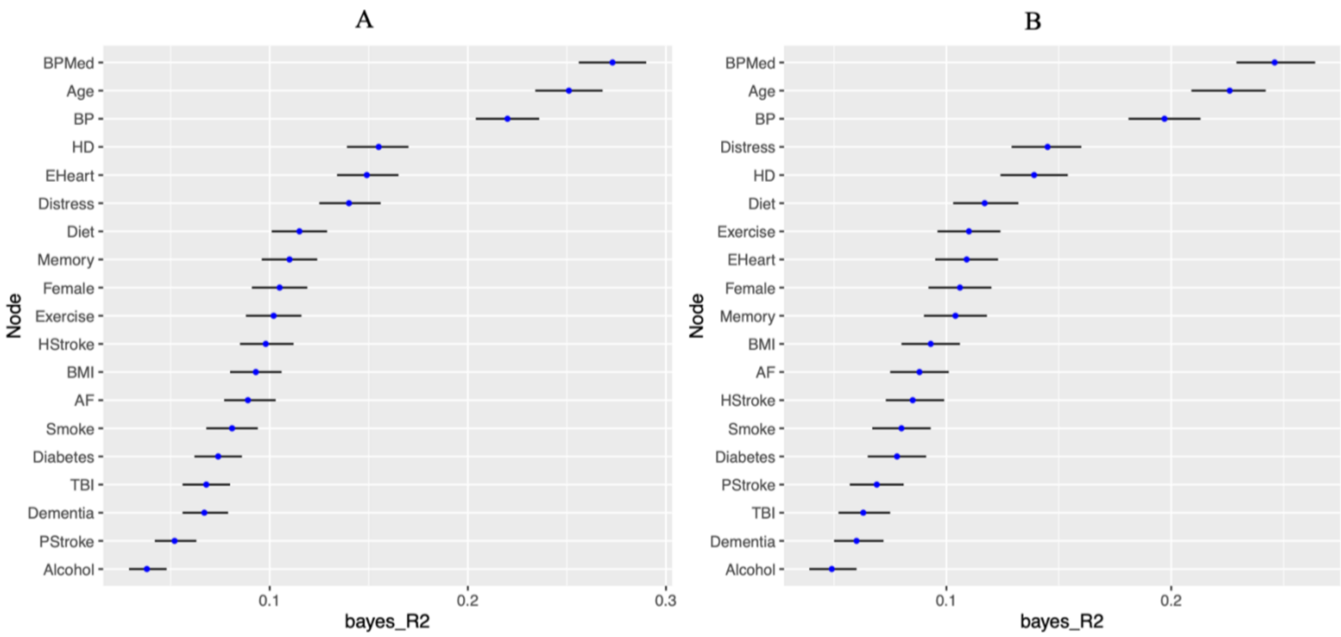


Figure 6

Predictability of Nodes for African (A) and Asian (B) Networks

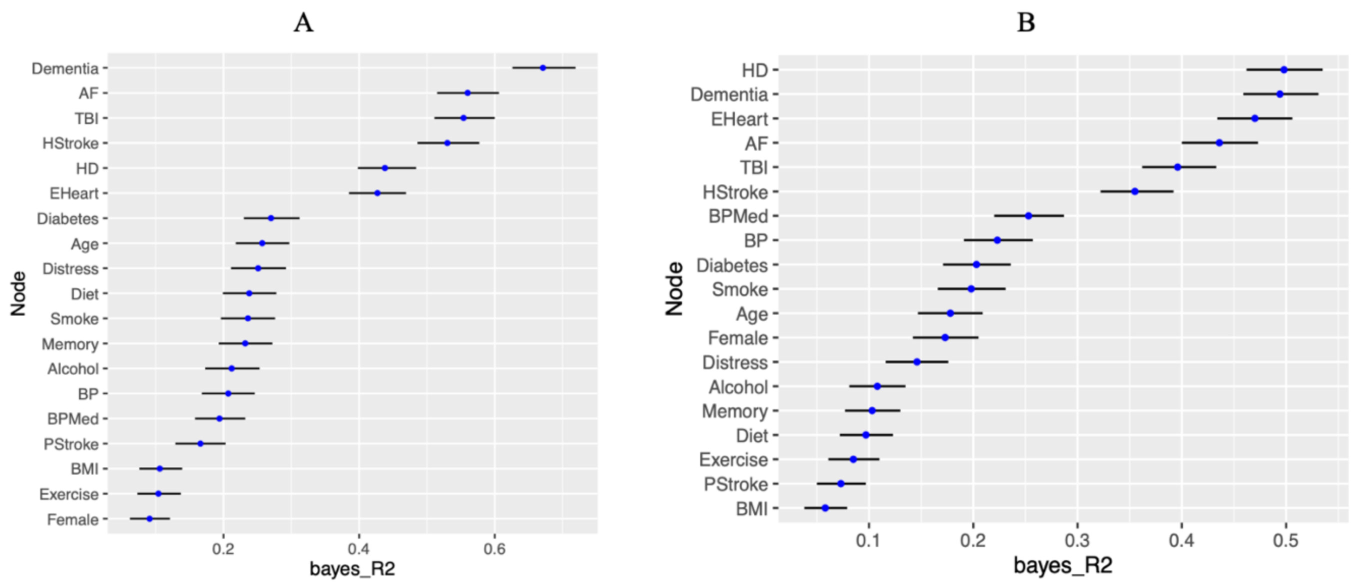


Figure 7

Predictability of Nodes for Indian (A) and Latin/Hispanic (B) Networks

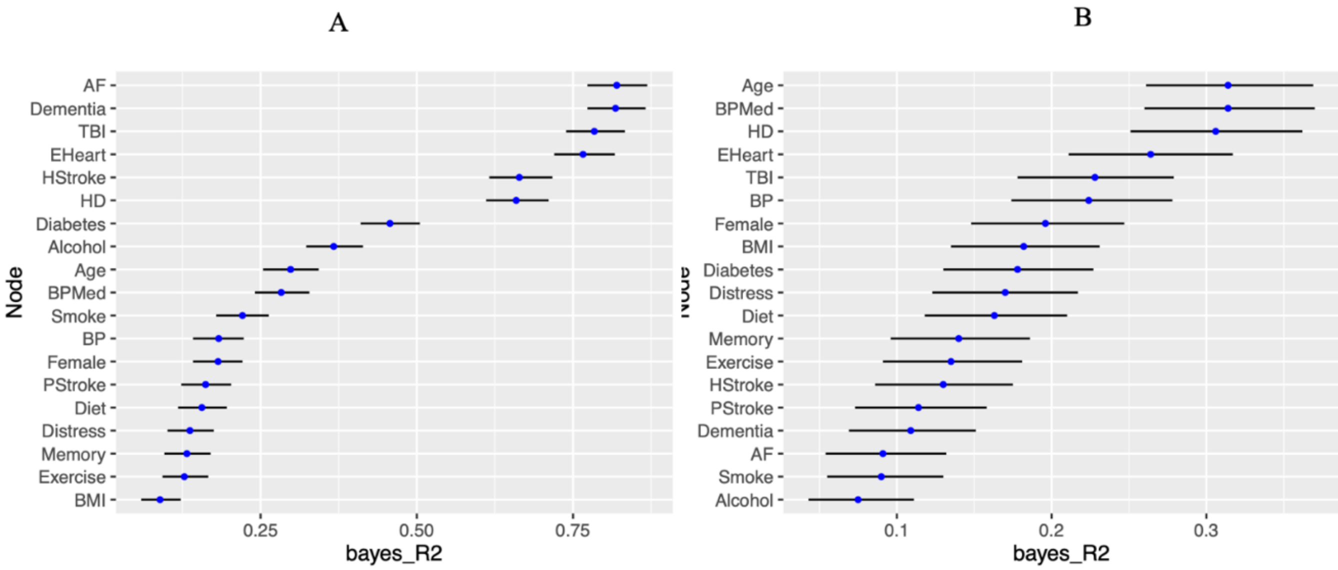
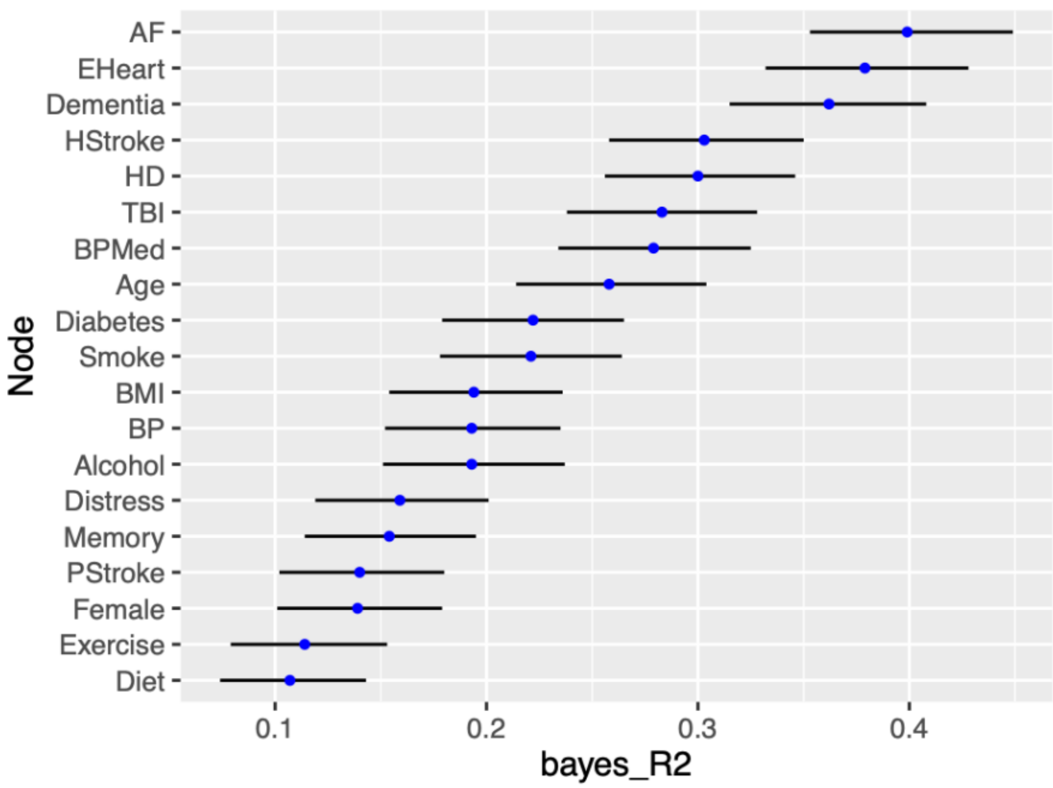


Figure 8

Predictability of Nodes for Other Ethnicity Network

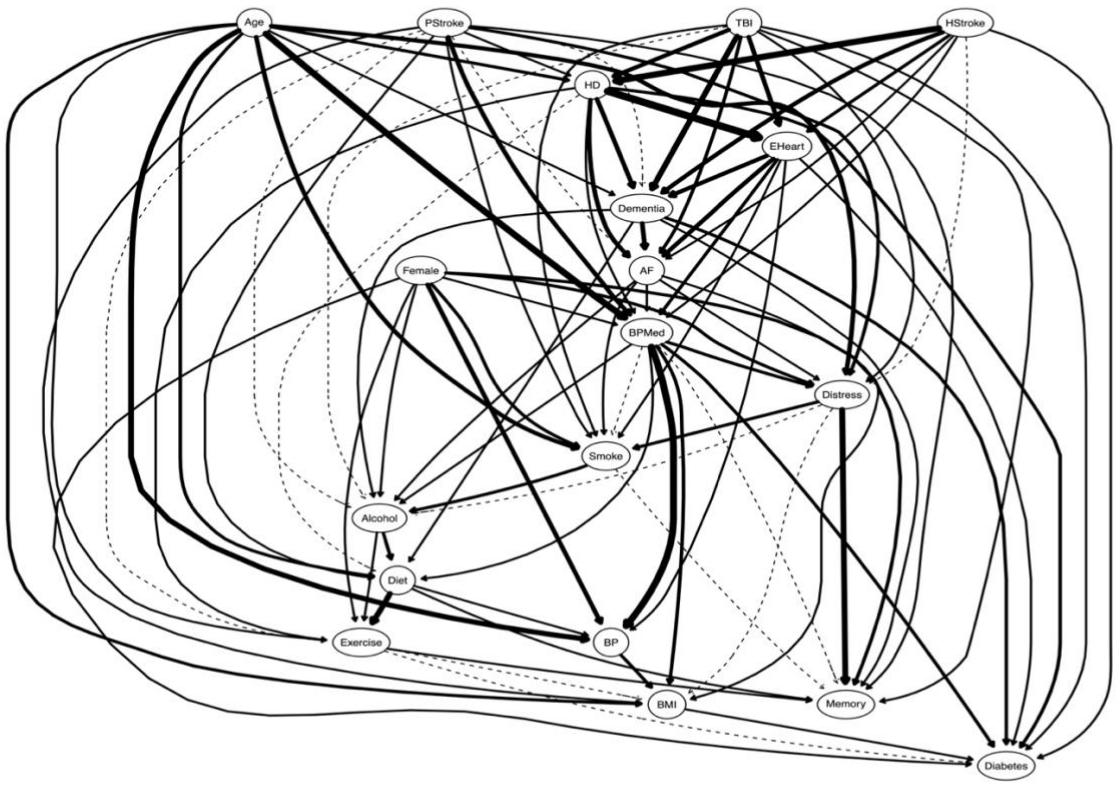


Directed Acyclic Graph (DAG) Analysis

Figure 9 illustrates the Directed Acyclic Graph (DAG) analysis based on 5,000 bootstrapped samples to assess edge stability and directional probability. High strength and directional probability were determined to be 0.85 or above, indicating strong, robust directional associations. A result of 1.00 means 100% probability for either strength or direction, and if both are 1.00 or close, it suggests a high degree of certainty in the direction and influence between variables. The DAG extends BGGM by revealing possible directional relationships among stroke risk factors, thereby discerning how one factor directly affects another. BIC and bootstrapping procedures were both used to assess edge stability and directional probability.

Figure 9

A Directed Acyclic Graph (DAG) of variables associated with Stroke Risk



Note. The thickness and direction of arrows between variables depict the strength (representing the magnitude of association) and directionality (representing the influence of one variable on another) of relationships. This is learned through the bootstrapping procedure.

Several edges had extremely large negative BIC scores, indicating their importance to the model. This included heart disease and an enlarged heart (ΔBIC = -742.86), age and using blood pressure medication (ΔBIC = -546.51), blood pressure medication and high blood pressure (ΔBIC = -542.51), TBI and dementia (ΔBIC = -305.70), and an enlarged heart and dementia (ΔBIC = -252.58).

Several non-modifiable ancestor nodes were identified in the network, including age, parental stroke, TBI, a history of stroke and being female, which created a cascading impact on other factors, with no factors impacting them.

An individual's age is strongly predictive of lifestyle factors, including smoking, a healthy diet, exercise, and BMI, all of which have a strength and directionality of 1.00. Aging also had a psychological pathway present, predicting psychological distress (strength = 1.00, directionality = 1.00) and dementia (strength = 0.99, directionality = 1.00), as well as a cardiovascular pathway predicting heart disease, high blood pressure, and taking blood pressure medication (strength = 1.00, directionality = 1.00).

Having a history of stroke predicted the psychological factor of dementia, as well as cardiovascular factors of heart disease, an enlarged heart, atrial fibrillation, and the physical health factor of diabetes, with strength and directionality ranging from 0.99 to 1.00, indicating a high level of certainty of influence.

The familial factor of a parental stroke predicted lifestyle factors, including smoking, diet, and BMI, with a strength ranging from 0.92 to 1.00 and directionality of 1.00. Parental stroke also predicted psychological distress and memory, as well as the cardiovascular factor of taking blood pressure medication, with a strength and directionality of 1.00.

Having a traumatic brain injury predicted the lifestyle factor of smoking as well as psychological factors of distress, dementia and poor memory (strength ranging from 0.86 to 1.00, and directionality = 1.00). TBI also predicted cardiovascular factors of heart disease, enlarged heart, and atrial fibrillation, as well as the physical health factor of diabetes (strength ranged from 0.88 to 1.00, directionality = 1.00).

Being female presented cascading pathways, predicting lifestyle factors of smoking, alcohol, and exercise, a psychological pathway to distress and memory, as well as a cardiovascular and physical health pathway, predicting high blood pressure and diabetes, all with a strength and directionality of 1.00.

Critical psychological pathways emerged. Psychological distress strongly predicted poor memory (strength = 1.00, directionality = 0.92) and moderately predicted smoking (strength = 1.00, directionality = 0.73). Dementia strongly predicted alcohol use (strength = 1.00, directionality = 0.96) and a healthy diet (strength = 0.98, directionality = 0.87). Dementia also strongly predicted poor memory (strength = 1, directionality = 0.99) and diabetes (strength = 1.00, directionality = 0.99).

A healthy diet predicted poor memory (strength = 0.91, directionality = 0.88) and high blood pressure (strength = 0.99, directionality = 0.74). Exercise was predictive of poor memory (strength = 1.00, directionality = 0.83). A high BMI was a strong predictor of diabetes (strength = 1.00, directionality = 0.69).

Heart disease strongly predicted diabetes (strength = 1.00, directionality = 0.97). An enlarged heart also strongly predicted diabetes (strength = 1.00, directionality = 0.96). Atrial fibrillation was a strong predictor of high BMI (strength = 0.96, directionality = 0.97) and a predictor of psychological distress (strength =1.00, directionality 0.82). Taking blood pressure medication was a predictor of a high BMI (strength = 1.00, directionality = 0.83), psychological distress (strength = 1.00, directionality = 0.87), and of diabetes (strength = 1.00, directionality = 0.98).

Some edges showed high strength but uncertain directionality, suggesting either ambiguous or reciprocal relationships. For example, an enlarged heart was a strong predictor of dementia (strength = 1.00), but the direction was 0.61, indicating a possible reciprocal relationship. This also had a strong BIC (-252.58), indicating a significant edge in the network. Similarly, heart disease strongly predicted an enlarged heart, with directionality in only 0.61 of iterations and with a BIC of -742.86. Another example is the bidirectional association between smoking and alcohol use, which had a strength of 1.00 but a direction of 0.55, indicating an almost even split and implying a non-causal or cyclical relationship between these two risk factors.

Two critical terminal factors were identified in the network, showing that many factors predict them, as described above, but they have little to no predictive input on other factors measured within this network. This includes the physical health risk factor of diabetes and the psychological risk factor of memory. This identifies stroke risk factors that, with increased therapeutic input, may decrease other factors higher up in the network.

Discussion

This study used BGGM and DAG network analysis to explore the unique associations and directional probabilities among known stroke risk factors across different ethnic groups. The BGGM revealed that 77-84% of all associations were stable across ethnic groups, suggesting that the majority of associations between stroke risk factors are universal and generalisable. This is beneficial for creating efficient screening and prevention strategies that are applicable to multiple ethnicities, which is the primary novelty of this study.

Robust associations between cardiovascular risk and non-modifiable risk factors, including age, stroke history, and TBI, that were evident cross-culturally, consistent with the support for hypotheses 2A, 2C, 2F, 2H, 2I, and 2J. This indicates that associations exist universally between high blood pressure and the use of blood pressure medication, high blood pressure and older age, and the use of blood pressure medication and older age, as expected based on the literature, and confirms the viability of the networks. These findings, coupled with the DAG results discussed below, support research suggesting that age is highly predictive of cardiovascular conditions and increases stroke risk (Boehme et al., 2017; Nindrea & Hasanuddin, 2023; O’Donnell et al., 2010; Saini & Gurvendra, 2022; Wolf et al., 1991). A strong cardiovascular disease pattern was noted cross-culturally between heart disease, an enlarged heart and atrial fibrillation, all being associated and confirmed in the DAG, with strong, reciprocal relationships being found. However, heart disease was slightly more predictive of the other two conditions.

Some associations were stronger in certain ethnicities, indicating greater importance for those cohorts. For example, the association between a history of stroke and TBI was much stronger in the African and Indian populations, while the association between heart disease and atrial fibrillation was much stronger in the Indian and Asian cohorts.

The findings of the DAG suggest that stroke risk originates from non-modifiable risk factors of age, sex, parental stroke, history of stroke and TBI, acting as ancestor nodes in the network, cascading through mediating nodes and ending in two terminal nodes. The ancestor nodes have no incoming edges, but have significant downstream impacts on mediating nodes, including cardiovascular risk factors (high blood pressure, blood pressure medication, atrial fibrillation, heart disease, enlarged heart), lifestyle factors (diet, exercise, smoking, alcohol consumption) and neuropsychological factors (psychological distress, dementia).

Mediating nodes were modifiable, offering key intervention points through behavioural, lifestyle, and medical interventions, highlighting the importance of a biopsychosocial approach to stroke risk prevention and management. Mediating nodes also offer protective factors. For example, diet and exercise had a strong reciprocal relationship, while exercise was predictive of BMI, which strongly predicted diabetes. This identified exercise as protective, in that more exercise likely improves diet, reduces BMI and reduces the likelihood of metabolic disease. Heart disease and an enlarged heart were also highly predictive of diabetes, identifying the increased risk associated with prolonged cardiovascular dysfunction resulting in diabetes. Blood pressure medication use was highly predictive of alcohol consumption, identifying points of behavioural intervention as cardiovascular risk increases with alcohol use. Psychological distress is predictive of smoking, poor memory and alcohol consumption; thus, improving individuals' psychological well-being may provide them with enhanced coping strategies that can reduce stroke risk through behavioural change and reduce cognitive decline.

Both ancestor and mediating nodes converge on diabetes and poor memory, two critical terminal nodes that have many incoming edges but no outgoing ones. These critical terminal nodes reveal the consequences of accumulated risk and the importance of multifocal, targeted intervention and that diabetes and poor memory may be manifestations of a multitude of other risk factors.

The DAG findings suggest that to reduce stroke risk, it is crucial to focus on modifiable risk factors and to take a holistic, biopsychosocial approach to intervention, whilst understanding how non-modifiable factors play a significant role in stroke risk. These factors cannot be changed, however, relationships between modifiable risk factors can be protective in reducing risk.

BGGM Findings

The association between high blood pressure and age (H2C) was much stronger in the African cohort, suggesting that, while this association is common globally, high blood pressure and aging are a more significant risk factor pairing among African people. This is consistent with the literature identifying undertreatment of hypertension in minority communities, particularly among Black communities (Mukadam et al., 2023; Sur et al., 2024). The association between a history of stroke and TBI (H2H) was much stronger in the African and Indian populations, possibly indicating a higher prevalence of these risk factors or better recording in these cohorts. While the literature on stroke following TBI is limited to stroke occurring after TBI or stroke that causes a brain injury, the findings in this study support this association (Esterov et al., 2023; Feigin et al., 2010; Qu et al., 2022; Sperl et al., 2022). In the current study's data, it is not known whether those with a TBI and stroke history experienced the TBI or the stroke first. Finally, the association between heart disease and atrial fibrillation (H2I) was much stronger in the Indian and Asian cohorts, identifying a co-occurring and elevated cardiovascular risk, highlighting the importance of cardiac monitoring in these populations (Chen et al., 2014; Goyal et al., 2025). Below, all other hypotheses are discussed.

Psychological Hypotheses

H1A Psychological Distress will Show a Positive Association with Poor Memory. This hypothesis was supported across all networks, except for the Latin/Hispanic population, identifying a cross-cultural pattern consistent with the literature associating high levels of distress with poor memory outcomes (Moran, 2016; Paradise et al., 2011). In contrast to previous longitudinal studies conducted in US cohorts by Sol et al. (2020), the current study identified that this association was significantly stronger among African participants, and no association was observed in the Latin/Hispanic cohort. This identifies the importance of psychological intervention in African communities and a gap in access to psychological care. Importantly, Sol et al. (2020) reported that psychological symptoms predicted poorer baseline memory and faster memory decline across all racial and ethnic groups, after controlling for ethnicity, findings that are somewhat in keeping with the current findings. The divergence in the Latin/Hispanic findings suggests that the relationship between psychological distress and poor memory may be more context-specific than studies have identified.

H1B Poor Memory will Show a Positive Association with Dementia. Interestingly, this hypothesis was only supported in the White/European cohort and the Other ethnicity cohort, which is not in keeping with what is known about the process of dementia, with the key symptoms being memory complaints (Behl et al., 2024; Gardener et al., 2015; Tack et al., 2025). This may be due to differences in reporting and measuring symptoms of dementia, or an inability for those with dementia to identify their subjective memory complaints accurately. Furthermore, prior literature suggests that subjective memory complaints should be documented as a precursor for a potential dementia diagnosis, due to the potential for vascular changes to the brain (Warren et al., 2022). The current findings may indicate that those with poor memory should be monitored for other vascular changes that are symptoms of both stroke and dementia, as a dementia diagnosis may not have yet occurred.

H1C Psychological Distress will be Negatively Associated with Age. Psychological distress decreased as people aged in the White/European, African, Indian and Latin/Hispanic cohorts, which is consistent with existing literature showing declining distress and negative affect from middle age onwards (Charles et al., 2023). In contrast, this association was not present in the Asian and Other ethnicity cohorts. No association, positive or negative, was reported in the Asian cohort, which diverges from previous literature (Charles et al., 2023). However, it is noted in some literature that psychological distress differs across ethnicity and is underreported in certain ethnicities, including Asian communities, which may be an explanation for the current study's results (Tran et al., 2015).

H1D Psychological Distress will be Positively Associated with Being Female. Psychological distress was associated with being female in the White/European, Asian, and Hispanic cohorts, but not the African, Indian or Other ethnicity cohorts, which is partially in keeping with the literature identifying that females have higher levels of psychological distress than males (Mirzaei-Alavijeh et al., 2025). Previous findings by Mirzaei-Alavijeh et al. (2025) showed that psychological distress was more pronounced in females, and those with lower education and socioeconomic status were more strongly associated with profound distress. This association was significantly stronger in the Latin/Hispanic cohort, which may reflect the cultural or socioeconomic position of the current participants (Mirzaei-Alavijeh et al., 2025). However, the current study's findings add additional information showing that cultural differences may play a role, identifying that there is no association between sex and psychological distress in African, Indian or Other ethnicity cohorts in the current data set.

Physical Health Hypotheses

H2B High Blood Pressure was Positively Associated with Higher BMI. High blood pressure and high BMI were positively associated in the White/European, Asian, and Other ethnicity cohorts, which is consistent with prior literature that has found that BMI is a strong and independent predictor of high blood pressure, and that Asian populations may be more metabolically sensitive to smaller BMI increases (Kuwabara et al., 2018; Landi et al., 2018). These findings suggest that keeping body weight lower reduces the chances of high blood pressure, and may reduce stroke risk prevalence, given hypertension's significant role in stroke risk (Boehme et al., 2017; O'Donnell et al., 2010). The African, Indian, and Latin/Hispanic cohorts had no identified association between high blood pressure and BMI. This could be due to lower reporting or access to identification of hypertension or cultural and lifestyle differences (Sur et al., 2024). There are also criticisms of using BMI as a measure in some research areas, and, as such, other measures, such as hip-to-waist ratio, may be used instead (Boehme et al., 2017).

H2E Taking Blood Pressure Medication was Positively Associated with Diabetes. This hypothesis was supported in the White/European, Indian, Latin/Hispanic and Other ethnicity cohorts. This is a unique finding, where literature support is incidental rather than direct. People who have high blood pressure are often at a higher risk of diabetes due to the vascular and lifestyle nature of the disease, and vice versa (Saini & Gurvendra, 2022). Therefore, taking blood pressure medication is more likely among those with diabetes (Saini

& Gurvendra, 2022). This association was significantly stronger in the Indian cohort, suggesting a higher risk in this population. No association was identified in the African and Asian cohorts. This could be due to access problems to adequate health care, as it is reported that African-Americans and other minority groups are regularly undermedicated in terms of hypertension medication (Sur et al., 2024).

H2G A History of Stroke will be Positively Associated with Heart Disease. This hypothesis was supported in the White/European, African and Asian cohorts, indicating that in these populations, heart disease is a crucial cardiovascular risk factor when someone has had a stroke, requiring monitoring by medical professionals (Nindrea & Hasanuddin, 2023). This supports previous findings by Nindrea & Hasnuddin (2023), who identified that first-time stroke survivors with a prior heart disease diagnosis are at increased risk of recurrent stroke. The African cohort had a significantly stronger association; therefore, this identified a very important cardiac risk factor for the African population. Indian, Latin/Hispanic and Other ethnicity cohorts did not have this association; however, they had associations between other heart conditions, like atrial fibrillation or an enlarged heart and a history of stroke. This identifies unique cardiovascular concerns that may be more strongly associated with having a stroke in certain ethnicities. While it is well established that cardiac concerns are a significant risk factor for strokes, this suggests that different conditions may impact populations to a greater or smaller extent in terms of stroke recurrence risk.

Lifestyle Hypotheses

H3A Exercise and a Healthy Diet will be Positively Associated. This hypothesis was supported across all networks, except for the African population, identifying a cross-cultural pattern consistent with the literature associating a healthy diet with more exercise (Wang et al., 2022; Yu et al., 2025). It was particularly strong in the Latin/Hispanic cohort, which is in keeping with the literature that Latinos have a generally healthier diet than other ethnicities, and nutritious diet and exercise are positively related to one another (Hiza et al., 2013; Wang et al., 2022; Yu et al., 2025). The current results add to those of Yu et al. (2025): both a healthy diet and regular physical activity act as protective factors in reducing stroke risk, and the network analysis identified strong associations between these factors. If the DAG results are added to this picture, it shows that a healthy diet is more likely to precede more physical exercise. This means that those who eat better are more likely to exercise, which would be important in intervention planning for a staged approach, focusing on improving an individual's diet before increasing exercise.

H3B Smoking and Alcohol Consumption will be Positively Associated. Consistent with prior literature, smoking and drinking were positively associated as complementary behaviours in the White/European, Asian, Indian, and Other ethnicity cohorts (Room, 2004). Both are known to increase stroke risk significantly as well as increase other risk factors, like hypertension (Boehme et al., 2017; Room, 2004). Both substances share physiological, psychological and sociocultural mechanisms of dependence, making cessation challenging for the individual (Room, 2004). This finding was significantly stronger in the Indian cohort, identifying possible cultural norms for this population. This provides support for focused intervention and education to adapt cultural norms around substance use and for people to better understand health risks associated with these behaviours (Room, 2004). When considering the results of the DAG, smoking is predictive of alcohol consumption; therefore, if an individual engages in smoking cessation, they may also reduce alcohol consumption organically, which is contrary to prior literature that suggests that alcohol is a cue for smoking (Room, 2004). No association was found between alcohol consumption and smoking in the African or Latin/Hispanic cohorts, indicating potential cultural, lifestyle and social norm differences.

H3C Smoking will be Negatively Associated with Age. Less smoking was associated with getting older in the White/European, Asian, Indian, and Other ethnicity cohorts, but not in the African or Latin/Hispanic cohorts. Current literature and studies from Smoke Free agencies suggest that middle age is when people smoke most, and that smoking cessation occurs after this time, which is in keeping with our results (Nguyen-Grozavu et al., 2020). Interestingly, this finding was not observed in the African or Latin/Hispanic cohorts, which may be due to education or neighbourhood disparities, as mentioned in prior literature, which suggests that cessation rates decreased with lower education and higher poverty rates; however, this is an assumption (Nguyen-Grozavu et al., 2020). This does provide crucial information as to where smoking cessation campaigns may be most needed, particularly given that smoking contributed to approximately 15% of stroke deaths each year (Boehme et al., 2017).

H3D A Healthier Diet will be Positively Associated with Age. Aging and a healthier diet were supported in the White/European, Asian, and Latin/Hispanic cohorts, which aligns with prior literature identifying that diet quality follows a U-shape across the lifespan, with 65+ year olds eating a healthier diet with more fruits, vegetables, and legumes than those in middle age (Aigner et al., 2018). This likely identifies a protective factor and a focus area for prevention to improve the diets of younger individuals and reduce the stroke mortality risk associated with low diet quality (Aigner et al., 2018). This association was not identified in the African, Indian, or Other ethnicity cohorts, suggesting that diets may be more consistent across the lifespan in these populations.

H3E Exercise will be Negatively Associated with Being Female. This hypothesis was not supported in any groups, other than White/European and Other ethnicity cohorts. This indicates that, among African, Asian, Indian, and Latin/Hispanic ethnicities, there was no association, positive or negative, between being female and exercising. This variation is likely due to cultural norms or limited exposure to exercise for White/European women. This suggests that White/European women should be exercising more frequently to lower their BMI, hypertension, heart disease risk and overall stroke and early death risk and protocols of as little as increasing step count are associated with decreased risk (Boden-Albala & Sacco, 2000; Saini & Gurvendra, 2022).

Cross-Cultural Differences

Based on an extensive literature review, this study is the first to apply BGGM network analysis to examine stroke risk factor associations by ethnicity; as such, it is difficult to interpret all findings across cohorts, as they reflect unique associations.

African Cohort. The African cohort had the most variation compared to the White/European cohort, with 39 significantly different node pairings, meaning 77.19% of associations in the network were consistent across the groups. This means that, of all associations found, over three-quarters are consistent, which is helpful for universal prevention and treatment planning.

Unique and strong associations between biopsychosocial stroke risk factors were identified in the African cohort. The use of alcohol was strongly associated with diet, having a traumatic brain injury, and having a dementia diagnosis. This identifies an area of importance for modifiable behaviour change in this cohort, as prior studies have determined that African populations have a higher risk of stroke if they are heavy alcohol users compared to their White counterparts; therefore, reducing alcohol consumption would likely be a protective factor (Kittner et al., 2021).

Being female was associated with high BMI, indicating a unique sex-related modifiable risk factor that could reduce stroke risk for African women (Wang et al., 2022). A diet with fruits and vegetables was associated with having diabetes, atrial fibrillation and being more psychologically distressed. This is a unique finding and may suggest that there is a higher prevalence of health conditions like diabetes and atrial fibrillation in African people, which other studies have identified (Kittner et al., 2021). Furthermore, it is unknown whether unhealthy food choices are also being made, as the survey question focuses only on fruit and vegetable servings and does not assess other important dietary factors, such as salt, sugar, and animal product intake (Guo et al., 2022).

Those with poor memory exercise less, which was not identified in the exploratory group. However, literature suggests that exercise has been shown to slow the progression of cognitive impairment for those with dementia or subjective memory complaints (Law et al., 2020). As such, people with cognitive impairment should continue exercising to slow this progression and reduce health risks (Law et al., 2020). Together, these findings suggest that there are unique lifestyle factors important to consider in prevention planning for the African population, encompassing behavioural changes, physical health conditions, and psychological factors, supporting the importance of a biopsychosocial model in stroke prevention (Engel, 2012).

Psychological distress was positively associated with poor memory and higher rates of smoking at a much larger magnitude than in the White/European cohort. This is consistent with the existing literature that suggests that psychological distress is associated with poorer memory, as well as using smoking as a common coping strategy for dealing with distressing situations (Boehme et al., 2017; Room, 2004). Behavioural interventions through smoking cessation, as well as distress management tools, would likely benefit African people who are at higher risk of stroke and act as a protective factor.

Dementia was the most predictive node in the African network. It was strongly associated with an enlarged heart, atrial fibrillation, TBI, and a stroke history, suggesting a robust vascular picture for African people, which is in keeping with prior literature (Mukadam et al., 2023). This is likely because stroke and dementia share common risk factors, and prior literature suggests that cardiovascular conditions affect Black people at higher rates than White people (Mukadam et al., 2023; Sur et al., 2024). Those with dementia also have higher smoking rates, which may be a coping strategy. It is well-established that smoking is a modifiable risk factor for both stroke and dementia, and it can be assumed that a person who is smoking and has a dementia diagnosis has a higher chance of stroke occurrence (Mukadam et al., 2023; Tack et al., 2025). This provides a vital area for behavioural modification to reduce risk through smoking cessation (Mukadam et al., 2023; Senff et al., 2025).

Diabetes was associated with being male in the African group, contrary to previous findings, which did not identify sex differences between diabetes as a risk factor for Black individuals in America; however, globally, this could be different (Howard et al., 2019). Moderate associations were found between heart disease and TBI, an enlarged heart and atrial fibrillation and parental stroke, which were not confirmed in the White/European groups.

Overall, the cardiovascular risk in the African cohort is substantial, supported by the high predictability of dementia and atrial fibrillation. This is in keeping with the literature, which states that controlling cardiovascular risk factors is essential for all people; however, it is particularly crucial in African and Black populations, as their risk for stroke is three times higher when cardiovascular conditions, such as uncontrolled hypertension, occur (Howard et al., 2019). In terms of cohort-based interventions, this suggests closer monitoring of cardiovascular conditions for African communities, as well as better treatment for non-modifiable conditions like TBI and dementia, which generally have poor outcomes for African patients (Mukadam et al., 2023). Furthermore, behavioural interventions to help those at risk of stroke decrease alcohol use and cigarette consumption would be beneficial for this cohort, as well as psychological help for managing psychological distress in ways that improve their health and well-being.

Asian Cohort. The Asian cohort had 29 significantly different node pairings when compared to the White/European cohort, meaning 83.04% of associations in the network were consistent across groups, indicating a high degree of similarity in stroke risk factors, which is beneficial for the generalisability of stroke risk prevention.

Unique and strong associations between biopsychosocial stroke risk factors were identified in the Asian cohort. Being male was strongly associated with smoking, which is in keeping with existing literature regarding Asian men smoking at higher rates than women, however, recent literature in this area is scarce (Ra et al., 2022). This suggests sex-specific intervention for the Asian cohort, likely requiring further smoking cessation education and risks.

Dementia was associated with TBI, diabetes, being younger, heart disease and an enlarged heart. It was one of the risk factors with the highest predictability in the Asian cohort, similar to the African cohort, where a robust vascular picture is present. This is consistent with prior studies identifying an increased risk of dementia in Asian populations and a growing burden of dementia diagnoses in Asia (Islam et al., 2024; Mukadam et al., 2023). The association between dementia and younger age is unique and concerning, as dementia generally affects older people more frequently, which has been consistent in Asian populations in the literature (Islam et al., 2024). The association between dementia and TBI was very strong, and provides a possible support for brain injury increasing the chances of dementia (Mukadam et al., 2023).

Psychological distress was associated with an enlarged heart, which is an incidental finding that has not been established in the literature, although psychological distress has been correlated to other cardiac conditions, like high blood pressure (Boehme et al., 2017).

A much stronger cardiac picture was present in the Asian cohort; heart disease, which had the highest predictability, was associated with an enlarged heart, atrial fibrillation, diabetes and a history of stroke. Atrial fibrillation was associated with TBI and an enlarged heart. An enlarged heart was associated with a history of stroke. This is consistent with prior knowledge, identifying that cardiac conditions were associated with significantly elevated stroke risk in Asian populations (Chen et al., 2014).

The findings from the Asian cohort suggest that cardiovascular conditions are central components of stroke risk in Asian populations, requiring clinical screening and management at population-based levels, which is consistent with the literature (Chen et al., 2014; Islam et al., 2024; Mukadam et al., 2023). However, some lifestyle and sex-specific behavioural changes could be implemented to improve smoking rates in Asian men.

Indian Cohort. The Indian cohort had 38 significantly different node pairings when compared to the White/European cohort, meaning 77.78% of associations in the network were consistent across groups, indicating a high degree of similarity in stroke risk factors.

Unique and strong associations between biopsychosocial stroke risk factors were identified in the Indian cohort. Of all the cohorts, the Indian group had the most node-to-node associations with alcohol use, exhibiting a possible cultural norm (Smyth et al., 2023). Alcohol use was strongly associated with smoking, having a stroke history, having an enlarged heart, taking blood pressure medication and with younger age. Previous studies have shown that heavy episodic drinking was common in India, but that Indian participants drank less than other ethnicities generally, but their risk was increased (Smyth et al., 2023). However, they also identified that spirits and arrack were associated with increased stroke risk, which is most consumed in India (Smyth et al., 2023). This identifies a risk factor that requires specific monitoring for this group, appropriate psychoeducation on the impacts of alcohol consumption, as well as assistance with behavioural modifications to lessen stroke risk.

TBI had several strong associations and was one of the risk factors with the highest predictability. TBI was associated with dementia, atrial fibrillation, being younger and less psychological distress, which is a unique finding. Whilst these findings are not established in the literature specifically for the Indian population, it is consistent with the broader literature, finding that TBI impacts younger people and is linked to cognitive changes and potentially dementia (Feigin et al., 2010). The association of TBI and atrial fibrillation is also supported in that individuals with TBI have an elevated risk of arrhythmias due to increased activation of the sympathetic nervous system, identifying essential points of monitoring for individuals with TBI in terms of stroke risk and cardiac risk, particularly in Indian patients, given the current findings (Stewart et al., 2025).

Similar to other groups, a cardiac profile was identified. Heart disease was associated with atrial fibrillation and an enlarged heart. Atrial fibrillation, which had the highest predictability, was associated with an enlarged heart, stroke history and dementia, and an enlarged heart was associated with a history of stroke and dementia as well, suggesting significant vascular involvement in stroke risk in Indian people, consistent with the literature (Goyal et al., 2025).

Diabetes was associated with a history of stroke and diet. This is consistent with the literature, which identifies diabetes as one of the two most commonly reported predisposing risks for stroke among Indian people (Goyal et al., 2025; Ram et al., 2021). Furthermore, Indian diets consist of lentils, grains, and vegetables; however, those who have had strokes in India primarily consume white rice and dal, with fewer leafy green vegetables and fruits (Durga & Manorenj, 2019). Similarly to the African cohort, it is difficult to interpret the relationship, as the Stroke Riskometer only measures fruit and vegetable intake and does not include other food sources commonly linked to diabetes or stroke risk.

Parental stroke was associated with being female and poor memory, another unique finding. This may mean that females in this cohort are more likely to have parents who have had a stroke than their male counterparts, and those with parents who have had a stroke tend to have poor memory. This could be genetically influenced, but also environmentally, with lifestyle behaviours such as smoking, drinking, poor diet, and low exercise being modelled (Linghu et al., 2025).

Being male is associated with older age in this cohort, which may indicate that females are utilising the Stroke Riskometer at a younger age, or that older females have less of a need to consider their stroke risk in this cohort due to men having a higher prevalence of stroke (Saini & Gurvendra, 2022).

Based on the BGGM results, the most important avenues for stroke prevention seem to be the prevention, accurate registration, and monitoring of all types of TBI, given their relationships with many other significant risk factors and the high predictability of TBI in this cohort (Feigin et al., 2010). Lifestyle and behavioural modifications to alcohol use are likely to reduce stroke risk, which could be done at an individual or population-based level through policy and psychological support to quit drinking (Smyth et al., 2023). Finally, adequate management of both metabolic and cardiovascular concerns such as diabetes, heart disease, atrial fibrillation and hypertension by physicians would benefit Indian patients in reducing stroke risk occurrence (Durga & Manorenj, 2019; Goyal et al., 2025).

Latin/Hispanic Cohort. The Latin/Hispanic cohort had the fewest significantly different node pairings (28) compared with the White/European cohort, meaning 83.63% of associations in the network were consistent across groups, indicating a high degree of similarity in stroke risk factors.

Unique and strong associations between biopsychosocial stroke risk factors were identified in the Latin/Hispanic cohort. Age is associated with dementia, high BMI, and high alcohol use, and it was the risk factor with the highest predictability. The association between older age and dementia diagnosis is consistent with what is known about dementia diagnosis and progression (Behl et al., 2024). The association between aging and having a higher BMI is also consistent with the literature, in that older Latinos tend to exercise less and have higher abdominal obesity, which increases the chances of diabetes in the older Latin population, which was another association found in this cohort: diabetes and high BMI (Avezum et al., 2015). The only association for alcohol consumption was with older age, and alcohol had the lowest predictability in this group. This suggests that either alcohol use is not a central modifiable risk factor in the Latin/Hispanic population or that it is connected to other risk factors that were not measured in this study.

Taking blood pressure medication was associated with diet, high BMI, heart disease, parental stroke and diabetes, and had one of the highest predictabilities for this cohort. Considering the existing knowledge of relationships between hypertension and lifestyle, environment and metabolic conditions (Boehme et al., 2017; Mosenzon et al., 2023; Nindrea & Hasanuddin, 2023), these associations are consistent with the literature. These findings support that this cohort must have access to adequate treatment and medication control of cardiovascular and metabolic conditions, as in the literature, they are often under-prescribed and with worse outcomes (Sur et al., 2024).

Poor memory was associated with less exercise, heart disease and a history of stroke. The findings between poor memory and less exercise were also identified in the African cohort, and support the notion that those with cognitive decline should continue exercising, as it is protective for slowing cognitive decline (Law et al., 2020). Subjective memory complaints and having a stroke history are consistent with the literature and likely identify vascular changes in the brain either before or after a first-event stroke (Sajjad et al., 2015).

Dementia was associated with a history of stroke, which is consistent with the literature, particularly as stroke and dementia share many similar risk factors (Behl et al., 2024; Gardener et al., 2015; Tack et al., 2025). Interestingly, dementia was associated with lower blood pressure in the Latin/Hispanic cohort, which is a unique finding that opposes the literature.

Psychological distress was associated with higher smoking rates, and high blood pressure, both of which are consistent with existing literature (Boehme et al., 2017; Room, 2004). This identifies the importance of smoking cessation psychoeducation, as well as support for distress tolerance and other mental health support, to reduce the risk of health concerns like stroke (Boehme et al., 2017; Room, 2004).

TBI was associated with cardiac conditions of an enlarged heart and heart disease, which has support in the literature, and supports that TBI, dementia and stroke share similar risk factors, highlighting the importance of TBI reduction and appropriate management when they do occur (Esterov et al., 2023; Izzy et al., 2023; Stewart et al., 2025).

This cohort had fewer associations with lifestyle factors, identifying that physical health factors, such as aging and the management of heart disease and hypertension, are the most central risk factors. While psychological and behavioural intervention may be beneficial, the priority should be on cardiovascular and metabolic conditions identification and appropriate management to reduce stroke risk for this cohort.

Other Ethnicity Cohort. The Other ethnicity cohort comprised individuals who either did not state their ethnicity in the application or were not large enough to form a robust network. For example, there was a very small percentage of Māori or Pacific Islander peoples in the entire cohort. While this type of analysis would be beneficial for these underserved Indigenous populations, it was not possible to perform network analysis with such small sample sizes (Balabanski et al., 2024; Feigin et al., 2025; Thompson et al., 2022). There were

30 significantly different node pairings compared to the White/European cohort, indicating that 82.46% of node-to-node associations were consistent.

Atrial fibrillation was associated with an enlarged heart, heart disease and TBI, and was the most predictive node in the network. An enlarged heart also had high predictability and was associated with a history of stroke and taking blood pressure medication. These findings are supported in the literature and underscore the importance of identifying, monitoring, and treating cardiac conditions to reduce stroke risk, given their established significant role in increasing stroke risk (Boehme et al., 2017; Wolf et al., 1991; Xu et al., 2020).

Dementia was associated with a history of stroke, heart disease, an enlarged heart, diabetes, TBI and diet. It was one of the risk factors with the highest predictability in the Other ethnicity cohort, similar to the African and Asian cohorts, in which a robust vascular picture is present along with shared risk factors (Law et al., 2020; Mukadam et al., 2023). The association between fruit and vegetable intake and dementia is unique to this group, and there is no known literature to support this finding, particularly given that ethnicity and location are unknown among this cohorts participants.

In terms of lifestyle factors, smoking was associated with psychological distress, similarly to the White/European, African and Latin/Hispanic cohorts (Boehme et al., 2017; Room, 2004). Less exercise was associated with poor memory, a finding observed in the Latin/Hispanic, Asian, and African cohorts (Law et al., 2020). Older age was associated with higher BMI, similar to the Latin/Hispanic cohort, supporting the literature that older people tend to exercise less, hold more weight and increase other known risk factors for stroke, like diabetes and hypertension (Nindrea & Hasanuddin, 2023; Wang et al., 2022). High BMI was associated with taking blood pressure medication, consistent with H2B, which indicates that people in this cohort are on blood pressure medication. This identifies the importance of lowering BMI in this cohort, to reduce the risk of the well-established risk factor of hypertension and subsequently, medication use (Landi et al., 2018).

Demographic factors identified that being female was associated with diabetes and more memory problems. Being younger was associated with more TBI, and those with TBI had higher blood pressure, which is consistent with the literature (Feigin et al., 2010). This supports the notion that accurate TBI assessment and management are important in minimising stroke risk, as well as the idea that TBI and stroke share common risk factors (Esterov et al., 2023; Izzy et al., 2023; Stewart et al., 2025). Having a parent who had a stroke was associated with being older, having heart disease and taking blood pressure medication, which could be attributed to either genetic or shared lifestyles and environmental learning (Linghu et al., 2025). Those with a familial history of stroke should be well-informed of the increased risk of stroke for them, as well as the common risk factors that increase the chances of stroke, so they can make adequate physical health lifestyle changes to reduce modifiable risk factors and their chances of stroke. Having a parent who had a stroke was also moderately associated with being female, which is an incidental finding with no known literature.

Summary of BGGM Results

Overall, the findings in the BGGM indicate variation across ethnic groups relative to the exploratory and confirmatory White/European group. As such, unique protocols for stroke prevention could be developed to target specific risk factor relationships, particularly those that are protective in nature, like improving diet and exercise, lowering BMI, improving metabolic and cardiovascular disease management, and decreasing smoking and psychological distress. This indicates the importance of considering stroke risk factors and their management in a holistic manner, such as that described by the Biopsychosocial model.

Universally, concentrating on the identification, management and treatment of hypertension, appropriate medication, cardiovascular disease, particularly in older age, and identification, management and treatment of traumatic brain injuries has been identified as crucial across ethnicities, but in particular, those that are often underserved like the African, Latin/Hispanic and Indigenous communities (Balabanski et al., 2024; Feigin et al., 2025; Mukadam et al., 2023; Thompson et al., 2022). The findings of the BGGM build upon prior knowledge and are the first of their kind to explore unique relationships between the well-established risk factors of stroke. It also enhances understanding of ethnic differences that prior research has identified, providing a basis for new exploration of unique pairings and, eventually, the development of new ways to monitor and treat stroke risk factors, which will also impact other known disorders that share risk factors. This includes both TBI and dementia, which are well-established in the literature to share many known risk factors with stroke.

Directed Acyclic Graph (DAG) Findings

The DAG was undertaken on the complete data set, as the majority of node-to-node associations confirmed in the White/European group were found in the other ethnic groups. The DAG showed hierarchical and cascading organisation of stroke risk factors beginning with non-modifiable demographic, family history, and neurological factors, with lifestyle, psychological and cardiovascular pathways forming and converging on two key terminal nodes, being diabetes and poor memory. Certain edges in the network exhibited extremely large BIC values, indicating that these relationships were structurally essential to the model, likely forming causal pathways.

Ancestor Nodes

Age, parental stroke, TBI, history of stroke, and sex were identified as ancestor nodes, with no incoming edges but a vast downstream influence, identifying that demographic, lifestyle, psychological and physical health risk factors are intertwined.

Age initiated the most cascading pathways that ended in key terminal outcomes of diabetes and poor memory. This supports prior literature indicating that, among non-modifiable risk factors, age is the most significant contributor, suggesting that age is a driving factor of physiological and behavioural changes that can lead to an increased risk of stroke (Boehme et al., 2017; Feigin et al., 2023; Linghu et al., 2025; Nindrea & Hasanuddin, 2023). Age had the greatest predictive strength and directionality for lifestyle behaviours, psychological outcomes, and cardiovascular factors. Literature suggests that age likely influences stroke through mechanisms like narrowing and hardening of arteries, the increased likelihood of chronic diseases like hypertension, cardiac conditions and diabetes, as well as changes to cognitive function and lifestyle factors as people get older (Linghu et al., 2025). These existing findings explain and support the directional cascade observed in this study's DAG network. The exceptionally large BIC value between age and blood pressure medication indicates that age-related cardiovascular management is central to the network, meaning it is central to stroke risk and cardiovascular dysregulation.

While age cannot be altered, many modifiable risk factors associated with aging can be improved through lifestyle and behavioural adaptations, thereby reducing stroke risk. These findings identify possible protective factors of aging. For example, age is a predictor of diet with perfect strength and direction in the DAG. When coupled with BGGM findings, indicating a positive association between age and diet, this suggests that as people age, they are more likely to eat more fruits and vegetables, which has been identified as a protective factor against stroke risk in the literature, and identifies a vital lifestyle modification target for prevention in younger populations (Guo et al., 2022). Diet predicts blood pressure medication use, which is predictive of diabetes and high blood pressure, indicating the downstream impact diet has on other health outcomes (Mosenzon et al., 2023).

A history of stroke was a strong ancestor node of dementia, heart disease, an enlarged heart, atrial fibrillation and diabetes, indicating a strong vascular change profile that is consistent with the findings of the BGGM and is supported in the literature (Nindrea & Hasanuddin, 2023). Parental stroke was a strong ancestor node predicting smoking, diet, high BMI, psychological distress and using blood pressure medication, which supports prior literature that while this can indicate a genetic contribution, it is strongly supportive of learnt lifestyle and environmental behaviours, identifying the importance of psychoeducation to reduce stroke risk in families (Linghu et al., 2025).

Traumatic Brain Injury emerged as a strong ancestor node, strongly predicting dementia, poor memory, psychological distress, cardiovascular disease, atrial fibrillation and diabetes. A very large BIC value was identified between TBI and dementia, indicating this is a central relationship in this network. This is an important finding, identifying possible causal links between TBI and the onset of dementia. This relationship has support in the literature and suggests that neurological injury may accelerate a vascular process contributing to cognitive decline as well as a vulnerability to cardiovascular and metabolic disease (Esterov et al., 2023; Feigin et al., 2010; Izzy et al., 2023; Qu et al., 2022; Stewart et al., 2025). Based on these findings, TBI appears to accelerate other known risk factors and provides robust support for improving the recording of even mild TBI occurrence. Furthermore, adequate psychological and physical rehabilitative support and ongoing monitoring are necessary for those who have experienced TBI to reduce stroke risk (Feigin et al., 2010).

Sex was a strong predictor of smoking and alcohol consumption, which, when coupled with BGGM results, suggests that men smoke and drink more than women, which likely increases their chances of stroke and identifies sex-specific interventions (Linghu et al., 2025). Furthermore, being male predicts having higher blood pressure and diabetes, which is supported in the literature. These findings have also been identified in a recent network analysis exploring sex differences in stroke risk factors, adding to the robustness of this study and those by Linghu et al. (2025). Being female is predictive of less exercise, poorer memory and being more psychologically distressed, also providing unique female sex-specific interventions for stroke prevention (Boden-Albala & Sacco, 2000; Mirzaei-Alavijeh et al., 2025). These findings were not identified in Linghu et al.'s (2025) study; however, poor memory and psychological distress were not used as risk factors in their research.

Mediating Nodes

Cardiovascular risk factors formed the core of the directional network, and the relationship between heart disease and an enlarged heart had the largest BIC value in the network, indicating a highly stable and essential relationship in stroke risk. The directionality was uncertain, suggesting a reciprocal pathological relationship rather than a causal relationship. This is consistent with and builds upon existing literature, as there are many known causes of both heart disease and an enlarged heart that perhaps overlap and increase stroke risk (Boehme et al., 2017; Xu et al., 2020).

A similar pattern was observed between high blood pressure and the use of blood pressure medication. While the dominant directionality indicates that high blood pressure precedes the use of antihypertensive medication, a reciprocal structure shows that this is not a cause-and-effect pathway. This reciprocal relationship is reflective of the complex dynamic of chronic disease management, in that using hypertensive medication functions less as an intervention that resolves the risk of hypertension, and more as an identifier of sustained cardiovascular dysregulation. This means that individuals who are prescribed these medications are likely those who have a history of persistent cardiovascular risk, rather than acting as a protective factor. The reciprocal relationship suggests that high blood pressure may persist even in the presence of antihypertensive medication, due to incomplete control, non-adherence, or underlying cardiovascular pathology, as indicated by downstream outcomes including atrial fibrillation, high BMI and diabetes (Boehme et al., 2017). This may suggest that people need to focus more on lifestyle changes in conjunction with hypertension medication, like diet, exercise, smoking and alcohol use changes, to reduce the impact of hypertension, rather than relying on medication alone, particularly as literature shows that hypertension is inadequately treated (Boehme et al., 2017). This finding supports the notion that a biopsychosocial approach to stroke risk management would likely benefit people by linking lifestyle behaviours, physical health conditions, and pharmacological interventions to develop an effective treatment plan that reduces the risk of chronic disease (Engel, 2012).

Dementia was an intermediate node, both predicting and being predicted by other nodes in the network. It was a strong predictor of poor memory, the most common symptom of dementia, providing evidence for this model (Gardener et al., 2015). It was also predictive of alcohol use, diet and diabetes, which suggests that cognitive decline may alter health behaviours, brain function and metabolic regulation, strengthening downstream risk accumulation (Gardener et al., 2015). Dementia showed strong incoming connections from heart disease and an enlarged heart. However, the directionality appeared reciprocal, suggesting a mutual relationship and supporting the notion that stroke risk factors share many similarities with those for dementia (Gardener et al., 2015; Mukadam et al., 2023; Senff et al., 2025).

Psychological distress was an intermediate node in the DAG network, having strong input from previously mentioned risk factors like cardiovascular disease, TBI, sex and age. It was also strongly predictive of smoking, poor memory, and diabetes, suggesting psychological processes may exacerbate risk pathways, particularly through unhealthy coping strategies (Boehme et al., 2017; Room, 2004).

Lifestyle factors acted as mediator nodes, passing upstream risk from demographic and psychological factors to metabolic factors. A high BMI was a strong predictor of diabetes, which is consistent with the literature (Kuwabara et al., 2018; Mosenzon et al., 2023). Smoking and alcohol consumption showed a bidirectional relationship in the network, suggesting a mutual relationship between the two nodes rather than a causal relationship, consistent with the BGGM and prior literature. Diet and exercise also showed a strong, reciprocal relationship, consistent with BGGM findings and prior literature (Aigner et al., 2018). This indicates that of these lifestyle behaviours, neither is more likely to precede the other.

Although the DAG primarily identified risk pathways, modifiable factors emerged as potential protective factors. Exercise emerged as a modifiable upstream factor to high BMI, which is highly predictive of diabetes. Therefore, exercise represents a possible point of lifestyle modification that could disrupt the progression of metabolic disease. Exercise has also been shown to have protective implications in the literature, reducing BMI, improving cardiovascular health and slowing the progression of memory complaints (Boden-Albala & Sacco, 2000; Law et al., 2020).

The position of smoking being downstream of psychological distress suggests that smoking cessation may be most effectively approached indirectly through psychological interventions to address distress. This may include developing better coping strategies, learning mindfulness regulation techniques, utilising pharmacological interventions, or engaging in talk therapy (Chalmers et al., 2024; Roemer et al., 2024). If people can develop psychological well-being, they may be less likely to utilise burdensome coping strategies like smoking. Furthermore, improving psychological well-being may have a protective effect on memory impairment.

Terminal Nodes

In the DAG, diabetes and poor memory served as terminal nodes that did not predict any other risk factors in the network but had strong predictive input from all other risk factors. These are important points in the directional network that identify the consequences of accumulated downstream risk from prolonged dysregulation of psychological, behavioural, environmental, neurological and cardiovascular risk factors, culminating in elevated metabolic and cognitive risk.

Overall, the DAG suggest that stroke risk emerges from interacting cascades of biopsychosocial factors. This includes demographic vulnerability, neurological injury, cardiovascular burden and psychological and memory complaints, rather than isolated, siloed risk factors. This model identifies important potential risk factors and intervention points to reduce stroke prevalence. However, to infer causation, longitudinal studies would be needed.

Theoretical Implications

There are several theoretical implications from this extensive exploratory study. The current research suggests that, consistent with a biopsychosocial model, cumulative stroke risk is shaped not by isolated variables but by a dynamic system of interconnected biological, psychological, and social factors that form a network of stroke risk (Engel, 2012). The findings add to existing knowledge that the biopsychosocial model is an informative way to examine biological conditions and to consider a broader scope when assessing risk factors and developing interventions and prevention strategies.

Many studies have reported differences in stroke risk factors across ethnicities (Feigin et al., 2025; Hajat et al., 2004; Howard et al., 2019; G. Howard, 2001; Rajsic et al., 2019). This study identified that while group differences do exist cross-culturally, relationships were largely stable across ethnic groups, indicating a broadly generalisable underlying structure to stroke risk. Furthermore, this study adds to a growing body of literature utilising network analysis as a robust statistical tool to understand how multiple factors interact with one another, with the DAG providing a basis for developing and testing hypotheses for future research in a longitudinal fashion, moving stroke risk factor understanding forward (Chalmers et al., 2024; Gevers-Montoro et al., 2023; Hoorelbeke et al., 2025; Roemer et al., 2024; Thomann et al., 2023).

Furthermore, the DAG demonstrated that stroke risk is organised hierarchically, with upstream non-modifiable risk factors shaping downstream modifiable behaviours and outcomes. In particular, age was a central node that initiated pathways across all domains and was identified as a strongly predictive node of stroke risk, which supports the literature (Linghu et al., 2025).

The Stroke Riskometer is an effective and accessible measurement tool, which has been validated cross-culturally (Medvedev et al., 2021). However, this study has identified possible modifications that may help improve data. The Stroke Riskometer only includes fruit and vegetable intake in its diet question, missing information around salt, animal product consumption and complex carbohydrates, and the literature on diets' impact on stroke risk is vast (Aigner et al., 2018; Boehme et al., 2017; Durga & Manorenj, 2019; Guo et al., 2022; Saini & Gurvendra, 2022). The findings from the current study identified some unknowns in the results due to limited information, suggesting that improvements in data collection may be possible. Furthermore, there is no information about socioeconomic position, another factor that would likely be beneficial, given that those in lower socioeconomic positions are at a higher risk of stroke (Balabanski et al., 2024; Kittner et al., 2021; Sur et al., 2024).

Together, the findings in this study extend biopsychosocial theory by demonstrating that stroke vulnerability is best understood as an interacting behavioural, psychological and physical health factors that increase susceptibility to stroke.

Practical Implications

Researchers in stroke have made significant strides in identifying and understanding stroke risk factors for decades to create accessible preventive tools and treatment methods (Boden-Albala & Sacco, 2000; Boehme et al., 2017; Feigin, 2013; Feigin & Norrving, 2014; Gardener et al., 2015; Hippisley-Cox et al., 2013; Medvedev et al., 2021; Parmar et al., 2015; Saini & Gurvendra, 2022; Sur et al., 2024; Wolf et al., 1991). This study adds to the literature by examining how risk factors relate to one another and identifying that altering one risk factor can significantly affect overall risk, particularly across biopsychosocial domains.

The findings in this study suggest that most risk factors are stable across ethnicities. Practically, this means that preventative strategies and therapeutic interventions may be generalisable, with some cultural tweaks required depending on ethnicity. Findings indicate that stroke prevention and screening may be most effective when designed around clustered pathways, rather than isolated risk factors. For example, cardiovascular risk management remains central, but the network structure suggests that medication alone is unlikely to disrupt risk trajectories without concurrent lifestyle behavioural changes (Boehme et al., 2017). As such, antihypertensive medication could be considered as a marker of sustained risk that requires ongoing management and behavioural intervention, rather than an endpoint (Boehme et al., 2017).

It supports the literature that cardiovascular conditions and their prevention are crucial globally, but also that some minority groups are more impacted to a larger extent (O'Donnell et al., 2010). This finding can help shape monitoring and treatment practices and suggests that more needs to be done to help minority groups access appropriate care (Avezum et al., 2015; Kuwabara et al., 2018; O'Donnell et al., 2010; Thompson et al., 2022).

Modifiable risk factors like diet, smoking, physical exercise and psychological distress represent important intervention points because they link to risk factors of high BMI, diabetes and poor memory (Henderson et al., 2013; Reddin et al., 2022). This is even more important given that much of stroke risk stems from non-modifiable risk factors that cannot be altered; it is important to target what can be modified. This suggests that, consistent with prior suggestions, modifying lifestyle factors through government policy and programmes, and individual behavioural intervention would be beneficial in reducing the risk of known terminal risk factors (Boehme et al., 2017; Linghu et al., 2025).

Psychological support for people with distress, particularly with other known risk factors, represents a critical intervention point. Particularly given its upstream position of smoking, suggesting that cessation efforts could be strengthened by integrating distress-focused interventions like coping skills, mindfulness and pharmacological and psychological treatment (Barry et al., 2020; Chalmers et al., 2024; Henderson et al., 2013; Mirzaei-Alavijeh et al., 2025; Paradise et al., 2011; Roemer et al., 2024). This will likely have a larger evidence base and prove more impactful than relying on behavioural change messaging alone.

Clinically, a history of TBI should prompt long-term cardiovascular and cognitive surveillance given its strong connections with dementia and downstream cardiovascular health outcomes. Improvements are likely necessary in the documentation of even mild TBI and in the adequate follow-up of TBI (Feigin et al., 2010).

Finally, although the overall network structure was broadly stable, there were differences in edge strength and configuration across ethnic groups. This supports culturally responsive prevention strategies, prioritising cohort-specific needs. For example, some cohorts require more behavioural change and intervention due to stronger associations with lifestyle risk factors, while others have significantly more pronounced cardiovascular or neuropsychological concerns. Overall, these findings support an integrated, culturally informed stroke-prevention model, grounded in a biopsychosocial model of care, rather than the biomedical model most nations use (Gevers-Montoro et al., 2023; Tomás et al., 2023; Weston, 2005).

Limitations and Future Directions

While this study presented novel findings using advanced statistical methodology and a large data sample, it also presents limitations and future opportunities for consideration. Data had already been collected, so some questions were missing important components. For example, the question about diet included only fruit and vegetable consumption, whereas the literature suggests that whole-grain, dietary fibre, sugar, salt, and animal product consumption alter stroke risk (Boden-Albala & Sacco, 2000; Guo et al., 2022). We cannot know how these factors would change associations between nodes, and it may be helpful to update the application questions to reflect these gaps. We also do not know if those participants who had stroke histories had other complaints before or after stroke, for example, the association of history of stroke and TBI.

Another limitation is that, due to the extensive, global data, it was not possible to classify by ethnicity and country, so the data was classified by ethnicity alone to ensure robust numbers for network analysis in each cohort. This does not account for participants' environments, access to medical care, and socioeconomic position, all of which are known to affect stroke risk (Hajat et al., 2004; Mirzaei-Alavijeh et al., 2025; Sur et al., 2024). Future research could conduct a similar study by ethnicity within a country. For example, the data set included very few Māori and Pacific peoples, although several studies in Aotearoa have examined stroke risk and treatment disparities (Balabanski et al., 2021; Feigin et al., 2025; Thompson et al., 2022). The current research format would provide important information on differences across ethnic groups within a country and could inform policy changes to optimise stroke prevention strategies at a practical level, which would require further research.

Finally, while DAG is a robust analysis and identifies possible causal pathways, it cannot confirm causation on exploratory data. Future studies may utilise longitudinal research, developing hypotheses to test based on the DAG results from the current study and tracking them over time.

Conclusion

This thesis aimed to explore the unique associations and directional probabilities among known stroke risk factors across ethnic groups from a Biopsychosocial model perspective, utilising BGGM and DAG network analysis. The BGGM findings indicated that most associations (78-84%) were stable across ethnicities, with some unique differences identified. At the core of stroke risk, consistent, stable associations between cardiovascular and non-modifiable risk factors of age, TBI and stroke history emerged across groups that had some variation in strength cross-culturally, highlighting the need for culturally responsive prioritisation of risk pathways. Expected associations confirmed the viability of the network models. The DAG extended these findings by revealing a hierarchical cascading network, in which upstream non-modifiable factors shape downstream lifestyle, psychological, cardiovascular, metabolic, and cognitive outcomes. This network converged on diabetes and poor memory as key terminal points that identify the consequences of prolonged risk dysregulation. These findings suggest that stroke risk develops through multiple biopsychosocial pathways, highlighting the importance of reducing modifiable risk factor influence, since most risk stems from non-modifiable factors. Together, these results identify practical intervention targets, such as integrated management of hypertension and cardiovascular disease, lifestyle behavioural modification and psychological support, and long-term TBI surveillance and management. This multifaceted approach to stroke risk prevention could reduce the cascading impact of layered stroke risk, reducing prevalence caused by non-modifiable factors.

It is emphasised that directional models in exploratory data remain probabilistic of causal links, and longitudinal research is needed to test the inferred pathways. It was also suggested that refining risk assessment by incorporating additional contextual information, including broader dietary markers and socioeconomic factors, would enable more precise, equitable and effective stroke prevention strategies across populations. The findings from the current study add to the extensive research on stroke risk prevention and management by providing novel insights into how biological, psychological, and social risk factors interact uniquely and dynamically.

References

Aigner, A., Becher, H., Jacobs, S., Wilkens, L. R., Boushey, C. J., Le Marchand, L., Haiman, C. A., & Maskarinec, G. (2018). Low diet quality and the risk of stroke mortality: The multiethnic cohort study. *European Journal of Clinical Nutrition*, *72*(7), 1035–1045. https://doi.org/10.1038/s41430-018-0103-4

Ashraf, H., Ashfaq, H., Dawood, M. H., Kan Changez, M. I., Butt, M., Ashraf, A., Ahmed, S., & Nadeem, Z. A. (2025). Sex disparities and trends in stroke incidence, prevalence, and mortality in the US and worldwide: Findings from the global burden of disease, 1990–2021. *Journal of Clinical Neuroscience*, *133*. https://doi.org/10.1016/j.jocn.2025.111029

Avezum, Á., Costa-Filho, F. F., Pieri, A., Martins, S. O., Marin-Neto São Paulo, J. A., Grande do Sul, R., & Preto, R. (2015). Stroke in Latin America burden of disease and opportunities for prevention. *Global Heart*, *10*(4), 323. https://doi.org/10.1016/j.gheart.2014

Balabanski, A. H., Dos Santos, A., Woods, J. A., Mutimer, C. A., Thrift, A. G., Kleinig, T. J., Suchy-Dicey, A. M., Siri, S. R. A., Boden-Albala, B., Krishnamurthi, R. V., Feigin, V. L., Buchwald, D., Ranta, A., Mienna, C. S., Zavaleta-Cortijo, C., Churilov, L., Burchill, L., Zion, D., Longstreth, W. T., … Katzenellenbogen, J. M. (2024). Incidence of Stroke in Indigenous Populations of Countries With a Very High Human Development Index. *Neurology*, *102*(5). https://doi.org/10.1212/wnl.0000000000209138

Balabanski, A. H., Dos Santos, A., Woods, J. A., Thrift, A. G., Kleinig, T. J., Suchy-Dicey, A., Siri, S. R., Boden-Albala, B., Krishnamurthi, R., Feigin, V. L., Buchwald, D., Ranta, A., Mienna, C. S., Zavaleta, C., Churilov, L., Burchill, L., Zion, D., Longstreth, W. T., Tirschwell, D. L., … Katzenellenbogen, J. M. (2021). The Incidence of Stroke in Indigenous Populations of Countries With a Very High Human Development Index: A

Systematic Review Protocol. *Frontiers in Neurology*, *12*. https://doi.org/10.3389/fneur.2021.661570

Barker-Collo, S., Krishnamurthi, R., Nair, B., Ranta, A., Douwes, J., & Feigin, V. (2024). Depression and anxiety at 1- and 12-months post ischemic stroke: methods for examining individual change over time. *Brain Impairment*, *25*(4). https://doi.org/10.1071/IB24025

Barry, V., Stout, M. E., Lynch, M. E., Mattis, S., Tran, D. Q., Antun, A., Ribeiro, M. J. A., Stein, S. F., & Kempton, C. L. (2020). The effect of psychological distress on health outcomes: A systematic review and meta-analysis of prospective studies. In *Journal of Health Psychology* (Vol. 25, Issue 2, pp. 227–239). SAGE Publications Ltd. https://doi.org/10.1177/1359105319842931

Behl, T., Kaur, I., Sehgal, A., Khandige, P. S., Imran, M., Gulati, M., Khalid Anwer, M., Elossaily, G. M., Ali, N., Wal, P., & Gasmi, A. (2024). The link between Alzheimer's disease and stroke: A detrimental synergism. *Ageing Research Reviews*, *99*, 1–3. https://doi.org/10.1016/j.arr.2024.102388

Boden-Albala, B., & Sacco, R. L. (2000). Lifestyle factors and stroke risk: Exercise, alcohol, diet, obesity, smoking, drug use, and stress. *Current Atherosclerosis Reports*, *2*, 160–166.

Boehme, A. K., Esenwa, C., & Elkind, M. S. V. (2017). Stroke risk factors, genetics and prevention. *Circulation Research*, *120*(3), 472–495. https://doi.org/10.1161/CIRCRESAHA.116.308398

Borsboom, D., Deserno, M. K., Rhemtulla, M., Epskamp, S., Fried, E. I., McNally, R. J., Robinaugh, D. J., Perugini, M., Dalege, J., Costantini, G., Isvoranu, A. M., Wysocki, A. C., van Borkulo, C. D., van Bork, R., & Waldorp, L. J. (2021). Network analysis of multivariate data in psychological science. *Nature Reviews Methods Primers*, *1*(1). https://doi.org/10.1038/s43586-021-00055-w

Chalmers, R. A., Cervin, M., Choo, C., Sutton, A., Barcaccia, B., Couyoumdjian, A., Pallini, S., Iqbal, N., Reid, V., Singh, N. N., & Medvedev, O. N. (2024). Psychological risk and protective factors of immunity: a cross-cultural network analysis. *Current Psychology*. https://doi.org/10.1007/s12144-024-07178-9

Chalmers, R. A., Cervin, M., & Medvedev, O. N. (2022). Network Analysis. In *Handbook of Assessment in Mindfulness Research* (pp. 1–16). Springer International Publishing. https://doi.org/10.1007/978-3-030-77644-2_70-1

Charles, S. T., Rush, J., Piazza, J. R., Cerino, E. S., Mogle, J., & Almeida, D. M. (2023). Growing old and being old: Emotional well-being across adulthood. *Journal of Personality and Social Psychology*, *125*(2), 455–469. https://doi.org/10.1037/pspp0000453

Chen, X., Zhou, L., Zhang, Y., Yi, D., Liu, L., Rao, W., Wu, Y., Ma, D., Liu, X., Zhou, X. H. A., Lin, H., Cheng, D., & Yi, D. (2014). Risk factors of stroke in western and asian countries: A systematic review and meta-analysis of prospective cohort studies. *BMC Public Health*, *14*(1). https://doi.org/10.1186/1471-2458-14-776

Devereux, N., & Berns, A. M. (2023). Evaluation and treatment of psychological effects of stroke. *Delaware Journal of Public Health*, *9*(3). https://doi.org/10.32481/DJPH.2023.08.011

Dharan, A. S., Dalli, L. L., Olaiya, M. T., Cadilhac, D. A., Nedkoff, L., Kim, J., Andrew, N. E., Sundararajan, V., Thrift, A. G., Faux, S. G., Grimley, R., Kilkenny, M. F., & Kuhn, L. (2023). Risk factors associated with major adverse cardiovascular events after ischemic stroke: A linked registry study. *Neuroepidemiology*, *58*(2), 134–142. https://doi.org/10.1159/000535872

Dong, J.-Y., Zhang, Y.-H., Tong, J., Li, ;, & Qin, -Qiang. (2012). Depression and risk of stroke: A meta-analysis of prospective studies. *Stroke*, *43*, 32–37. https://doi.org/10.1161/STROKEAHA.111.630871/-/DC1

Durga, A., & Manorenj, S. (2019). Dietary pattern in adult patients with acute stroke in South India: A Case-Control Study from a Tertiary Care Center in Hyderabad. *Journal of Neurosciences in Rural Practice*, *10*(2), 283–293. https://doi.org/10.4103/jnrp.jnrp_237_18

Engel, G. L. (2012). The need for a new medical model: A challenge for biomedicine. *Psychodynamic Psychiatry*, *40*(3), 377–396.

Epskamp, S., & Fried, E. I. (2018). A tutorial on regularized partial correlation networks. *Psychological Methods*, *23*(4), 617–634. https://doi.org/10.1037/met0000167

Esterov, D., Sperl, M. A., Hines, E. A., Kinzelman Vesely, E. A., & Brown, A. W. (2023). Association between traumatic brain injury and increased risk of stroke: A systematic review and meta-analysis. *Journal of Head Trauma Rehabilitation*, *38*(1), E44–E55. https://doi.org/10.1097/HTR.0000000000000785

Faraway, J. J. (2016). Does data splitting improve prediction? *Statistics and Computing*, *26*(1–2), 49–60. https://doi.org/10.1007/s11222-014-9522-9

Feigin, V. L. (2013). *Stroke Riskometer App*. http://www.nisan.aut.ac.nz/stroke-riskometer.

Feigin, V. L., Barker-Collo, S., Krishnamurthi, R., Theadom, A., & Starkey, N. (2010). Epidemiology of ischaemic stroke and traumatic brain injury. *Best Practice and Research: Clinical Anaesthesiology*, *24*(4), 485–494. https://doi.org/10.1016/j.bpa.2010.10.006

Feigin, V. L., Krishnamurthi, R., Nair, B., Rautalin, I., Parag, V., Anderson, C. S., Arroll, B., Barber, P. A., Barker-Collo, S., Bennett, D., Brown, P., Cadilhac, D. A., Douwes, J., Exeter, D., Ranta, A., Ratnasabapathy, Y., Swain, A., Tautolo, E.-S., Ao, B. Te, …

Tunnage, B. (2025). Trends in stroke incidence, death, and disability outcomes in a multi-ethnic population: Auckland regional community stroke studies (1981-2022). *The Lancet Regional Health - Western Pacific*, *56*, 1–18. www.thelancet.com

Feigin, V. L., & Norrving, B. (2014). A new paradigm for primary prevention strategy in people with elevated risk of stroke. *International Journal of Stroke*, *9*(5), 624–626. https://doi.org/10.1111/ijs.12300

Feigin, V. L., Norrving, B., & Mensah, G. A. (2017a). Global Burden of Stroke. *Circulation Research*, *120*(3), 439–448. https://doi.org/10.1161/CIRCRESAHA.116.308413

Feigin, V. L., Norrving, B., & Mensah, G. A. (2017b). Primary prevention of cardiovascular disease through population-wide motivational strategies: insights from using smartphones in stroke prevention. *BMJ Global Health*, *2*. https://doi.org/10.1136/bmjgh-2017

Feigin, V. L., Owolabi, M. O., Feigin, V. L., Abd-Allah, F., Akinyemi, R. O., Bhattacharjee, N. V., Brainin, M., Cao, J., Caso, V., Dalton, B., Davis, A., Dempsey, R., Duprey, J., Feng, W., Ford, G. A., Gall, S., Gandhi, D., Good, D. C., Hachinski, V., … Zagożdżon, P. (2023). Pragmatic solutions to reduce the global burden of stroke: a World Stroke Organization–Lancet Neurology Commission. *The Lancet Neurology*, *22*(12), 1160–1206. https://doi.org/10.1016/S1474-4422(23)00277-6

Fruchterman, T. M. J., & Reingold, E. M. (1991). Graph Drawing by Force-Directed Placement. *Software: Practice and Experience*, 1–44. https://doi.org/10.1002/spe.4380211102

Gardener, H., Wright, C. B., Rundek, T., & Sacco, R. L. (2015). Brain health and shared risk factors for dementia and stroke. *Nature Reviews Neurology*, *11*(11), 651–657. https://doi.org/10.1038/nrneurol.2015.195

Geenen, R., & Dures, E. (2019). A biopsychosocial network model of fatigue in rheumatoid arthritis: A systematic review. *Rheumatology (United Kingdom)*, *58*, 10-V21. https://doi.org/10.1093/rheumatology/kez403

Gevers-Montoro, C., Liew, B. X. W., Deldar, Z., Conesa-Buendia, F. M., Ortega-De Mues, A., Falla, D., & Khatibi, A. (2023). A network analysis on biopsychosocial factors and pain-related outcomes assessed during a COVID-19 lockdown. *Scientific Reports*, *13*(1). https://doi.org/10.1038/s41598-023-31054-4

Goyal, A. D., Gajjar, A. A., Muhammad, N., Wu, A. Q., Polavarapu, H., Tang, O., Salem, M. M., Paliwoda, E. D., Gupta, N., Doad, J., Jankowitz, B. T., Srinivasan, V. M., & Burkhardt, J. K. (2025). A Three-Decade Analysis of Ischemic Stroke in India: Mortality, Morbidity, and Risk Factors Using the Global Burden of Diseases Study from 1990 to 2019. *Journal of Clinical Medicine*, *14*(21). https://doi.org/10.3390/jcm14217807

Guo, N., Zhu, Y., Tian, D., Zhao, Y., Zhang, C., Mu, C., Han, C., Zhu, R., & Liu, X. (2022). Role of diet in stroke incidence: an umbrella review of meta-analyses of prospective observational studies. *BMC Medicine*, *20*(1). https://doi.org/10.1186/s12916-022-02381-6

Hajat, C., Tilling, K., Stewart, J. A., Lemic-Stojcevic, N., & Wolfe, C. D. A. (2004). Ethnic differences in risk factors for ischemic stroke: A European case-control study. *Stroke*, *35*(7), 1562–1567. https://doi.org/10.1161/01.STR.0000131903.04708.b8

Hanna, M., Wabnitz, A., & Grewal, P. (2024). Sex and stroke risk factors: A review of differences and impact. *Journal of Stroke and Cerebrovascular Diseases*, *33*(4). https://doi.org/10.1016/j.jstrokecerebrovasdis.2024.107624

Hansen, K. D., Gentry, J., Long, L., Gentleman, R., Falcon, S., Hahne, F., & Sarkar, D. (2024). *Package 'Rgraphviz": Provides plotting capabilities for R graph objects* (2.51.0). https://git.bioconductor.org/packages/Rgraphviz

Harshfield, E. L., Pennells, L., Schwartz, J. E., Willeit, P., Kaptoge, S., Bell, S., Shaffer, J. A., Bolton, T., Spackman, S., Wassertheil-Smoller, S., Kee, F., Amouyel, P., Shea, S. J., Kuller, L. H., Kauhanen, J., Van Zutphen, E. M., Blazer, D. G., Krumholz, H., Nietert, P. J., … Davidson, K. W. (2020). Association between depressive symptoms and incident cardiovascular diseases. *JAMA - Journal of the American Medical Association*, *324*(23), 2396–2405. https://doi.org/10.1001/jama.2020.23068

Haslbeck, J. M. B., & Waldorp, L. J. (2018). How well do network models predict observations? On the importance of predictability in network models. *Behavior Research Methods*, *50*(2), 853–861. https://doi.org/10.3758/s13428-017-0910-x

Heeren, A., Lannoy, S., Coussement, C., Hoebeke, Y., Verschuren, A., Blanchard, M. A., Chakroun-Baggioni, N., Philippot, P., & Gierski, F. (2021). A network approach to the five-facet model of mindfulness. *Scientific Reports*, *11*(1). https://doi.org/10.1038/s41598-021-94151-2

Henderson, K. M., Clark, C. J., Lewis, T. T., Aggarwal, N. T., Beck, T., Guo, H., Lunos, S., Brearley, A., Mendes De Leon, C. F., Evans, D. A., & Everson-Rose, S. A. (2013). Psychosocial distress and stroke risk in older adults. *Stroke*, *44*(2), 367–372. https://doi.org/10.1161/STROKEAHA.112.679159

Hevey, D. (2018). Network analysis: a brief overview and tutorial. *Health Psychology and Behavioral Medicine*, *6*(1), 301–328. https://doi.org/10.1080/21642850.2018.1521283

Hippisley-Cox, J., Coupland, C., & Brindle, P. (2013). Derivation and validation of QStroke score for predicting risk of ischaemic stroke in primary care and comparison with other risk scores: A prospective open cohort study. *BMJ (Online)*, *346*(7910). https://doi.org/10.1136/bmj.f2573

Hiza, H. A. B., Casavale, K. O., Guenther, P. M., & Davis, C. A. (2013). Diet quality of Americans differs by age, sex, race/ethnicity, income, and education level. *Journal of the Academy of Nutrition and Dietetics*, *113*(2), 297–306. https://doi.org/10.1016/j.jand.2012.08.011

Hoorelbeke, K., Fried, E. I., & Koster, E. H. W. (2025). A comprehensive network analysis of biopsychosocial factors associated with postpartum depression. *Journal of Affective Disorders*, *390*. https://doi.org/10.1016/j.jad.2025.119808

Howard, G. (2001). Ethnic Disparities in Stroke: The Scope of the Problem. *Ethnicity & Disease*, *11*(4), 761–768.

Howard, Madsen, T. E., Kleindorfer, D. O., Judd, S. E., Rhodes, J. D., Soliman, E. Z., Kissela, B. M., Safford, M. M., Moy, C. S., McClure, L. A., Howard, G., & Cushman, M. (2019). Sex and race differences in the association of incident ischemic stroke with risk factors. *JAMA Neurology*, *76*(2), 179–186. https://doi.org/10.1001/jamaneurol.2018.3862

Howard, V. (2013). Reasons underlying racial differences in stroke incidence and mortality. *Stroke*, *44*(SUPPL. 1). https://doi.org/10.1161/STROKEAHA.111.000691

Islam, B., Li, T., Xu, M., Yang, D., Lv, H., Gassara, G., Ibrahim, T. I., Radman, B. A., & Wang, J. (2024). Emerging trends in cognitive impairment and dementia among older populations in Asia: A systematic review. *Journal of Global Health*, *14*, 1–23. https://doi.org/10.7189/jogh.14.04233

Izzy, S., Grashow, R., Radmanesh, F., Chen, P., Taylor, H., Formisano, R., Wilson, F., Wasfy, M., Baggish, A., & Zafonte, R. (2023). Long-term risk of cardiovascular disease after traumatic brain injury: screening and prevention. In *The Lancet Neurology* (Vol.

22, Issue 10, pp. 959–970). Elsevier Ltd. https://doi.org/10.1016/S1474-4422(23)00241-7

Kittner, S. J., Sekar, P., Comeau, M. E., Anderson, C. D., Parikh, G. Y., Tavarez, T., Flaherty, M. L., Testai, F. D., Frankel, M. R., James, M. L., Sung, G., Elkind, M. S. V., Worrall, B. B., Kidwell, C. S., Gonzales, N. R., Koch, S., Hall, C. E., Birnbaum, L., Mayson, D., … Woo, D. (2021). Ethnic and racial variation in intracerebral hemorrhage risk factors and risk factor burden. *JAMA Network Open*, *4*(8), E2121921. https://doi.org/10.1001/jamanetworkopen.2021.21921

Kokou-Kpolou, C. K., Park, S., Bet, Q., Iorfa, S. K., Chinweuba, D. C., & Chukwuorji, J. B. C. (2023). Towards a more comprehensive understanding of depressive symptoms among young adults using Gaussian graphical and directed acyclic graph models. *Current Psychology*, *42*(35), 31579–31589. https://doi.org/10.1007/s12144-022-04192-7

Kuwabara, M., Kuwabara, R., Niwa, K., Hisatome, I., Smits, G., Roncal-Jimenez, C. A., Maclean, P. S., Yracheta, J. M., Ohno, M., Lanaspa, M. A., Johnson, R. J., & Jalal, D. I. (2018). Different risk for hypertension, diabetes, dyslipidemia, and hyperuricemia according to level of body mass index in Japanese and American subjects. *Nutrients*, *10*(8). https://doi.org/10.3390/nu10081011

Landi, F., Calvani, R., Picca, A., Tosato, M., Martone, A. M., Ortolani, E., Sisto, A., D'angelo, E., Serafini, E., Desideri, G., Fuga, M. T., & Marzetti, E. (2018). Body mass index is strongly associated with hypertension: Results from the longevity check-up 7+ study. *Nutrients*, *10*(12). https://doi.org/10.3390/nu10121976

Law, C. K., Lam, F. M., Chung, R. C., & Pang, M. Y. (2020). Physical exercise attenuates cognitive decline and reduces behavioural problems in people with mild cognitive impairment and dementia: a systematic review. *Journal of Physiotherapy*, *66*(1), 9–18. https://doi.org/10.1016/j.jphys.2019.11.014

Lee, M., Saver, J. L., Hong, K. S., Song, S., Chang, K. H., & Ovbiagele, B. (2012). Effect of pre-diabetes on future risk of stroke: Meta-analysis. *BMJ (Online)*, *344*(7861). https://doi.org/10.1136/bmj.e3564

Linghu, L., Huang, Y., Qiu, L., Wang, X., Zhang, J., Ma, L., Li, chenglian, & Wang, L. (2025). Exploring stroke risk factors in different genders using Bayesian networks: a cross-sectional study involving a population of 134,382. *BMC Public Health*, *25*(1). https://doi.org/10.1186/s12889-025-23946-z

Medvedev, O., Truong, Q., Merkin, A., Borotkanics, R., Krishnamurthi, R., & Feigin, V. (2021). Cross-cultural validation of the stroke riskometer using generalizability theory. *Scientific Reports*, *11*(1). https://doi.org/10.1038/s41598-021-98591-8

Mirzaei-Alavijeh, M., Najafi, F., Veis, R. G., Moradinazar, M., Farasati, F., & Jalilian, F. (2025). Psychological distress and its association with gender, socioeconomic status, education and health conditions. *Scientific Reports*, *15*(1). https://doi.org/10.1038/s41598-025-14848-6

Moran, T. (2016). Supplemental material for anxiety and working memory capacity: A meta-analysis and narrative review. *Psychological Bulletin*, *142*(8), 831–864. https://doi.org/10.1037/bul0000051.supp

Mosenzon, O., Cheng, A. Y. Y., Rabinstein, A. A., & Sacco, S. (2023). Diabetes and stroke: What are the connections? *Journal of Stroke*, *25*(1), 26–38. https://doi.org/10.5853/jos.2022.02306

Mukadam, N., Marston, L., Lewis, G., Mathur, R., Lowther, E., Rait, G., & Livingston, G. (2023). South Asian, Black and White ethnicity and the effect of potentially modifiable risk factors for dementia: A study in English electronic health records. *PLoS ONE*, *18*(10 October). https://doi.org/10.1371/journal.pone.0289893

Nguyen-Grozavu, F. T., Pierce, J. P., Sakuma, K. L. K., Leas, E. C., McMenamin, S. B., Kealey, S., Benmarhnia, T., Emery, S. L., White, M. M., Fagan, P., & Trinidad, D. R. (2020). Widening disparities in cigarette smoking by race/ethnicity across education level in the United States. *Preventive Medicine*, *139*. https://doi.org/10.1016/j.ypmed.2020.106220

Nindrea, R. D., & Hasanuddin, A. (2023). Non-modifiable and modifiable factors contributing to recurrent stroke: A systematic review and meta-analysis. *Clinical Epidemiology and Global Health*, *20*. https://doi.org/10.1016/j.cegh.2023.101240

O'Donnell, M. J., Zhang, H., Wang, X., Xavier, D., Liu, L., Zhang, H., Lim Chin, S., Rao-Melacini, P., Rangarajan, S., qul Islam, S., Pais, P., McQueen, M. J., Mondo, C., Damasceno, A., Lopez-Jaramillo, P., Hankey, G. J., Dans, A. L., Yusoff, K., Truelsen, T., … Yusuf, S. (2010). Risk factors for ischaemic and intracerebral haemorrhagic stroke in 22 countries (the INTERSTROKE study): a case-control study. *The Lancet*, *376*, 112–123. https://doi.org/10.1016/S0140

Pan, A., Sun, Q., Okereke, O. I., Rexrode, K. M., & Hu, F. B. (2011). Depression and risk of stroke morbidity and mortality: A meta-analysis and systematic review. *JAMA*, *306*(11), 1241–1249. https://doi.org/10.1001/jama.2011.1282

Paradise, M. B., Glozier, N. S., Naismith, S. L., Davenport, T. A., & Hickie, I. B. (2011). Subjective memory complaints, vascular risk factors and psychological distress in the middle-aged: A cross-sectional study. *BMC Psychiatry*, *11*. https://doi.org/10.1186/1471-244X-11-108

Parmar, P., Krishnamurthi, R., Ikram, M. A., Hofman, A., Mirza, S. S., Varakin, Y., Kravchenko, M., Piradov, M., Thrift, A. G., Norrving, B., Wang, W., Mandal, D. K., Barker-Collo, S., Sahathevan, R., Davis, S., Saposnik, G., Kivipelto, M., Sindi, S., Bornstein, N. M., … Feigin, V. L. (2015). The Stroke RiskometerTM App: Validation of a data collection tool and stroke risk predictor. *International Journal of Stroke*, *10*(2), 231–244. https://doi.org/10.1111/ijs.12411

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830. http://scikit-learn.sourceforge.net.

Potter, T. B. H., Tannous, J., & Vahidy, F. S. (2022). A Contemporary review of epidemiology, risk factors, etiology, and outcomes of premature stroke. *Current Atherosclerosis Reports*, *24*(12), 939–948. https://doi.org/10.1007/s11883-022-01067-x

Qu, D., Li, W., Zhang, S., Li, R., Wang, H., & Chen, B. (2022). Traumatic brain injury is associated with both hemorrhagic stroke and ischemic stroke: A systematic review and meta-analysis. *Frontiers in Neuroscience*, *16*. https://doi.org/10.3389/fnins.2022.814684

Ra, C. K., Pehlivan, N., Kim, H., Sussman, S., Unger, J. B., & Businelle, M. S. (2022). Smoking prevalence among Asian Americans: Associations with education, acculturation, and gender. *Preventive Medicine Reports*, *30*. https://doi.org/10.1016/j.pmedr.2022.102035

Rajsic, S., Gothe, H., Borba, H. H., Sroczynski, G., Vujicic, J., Toell, T., & Siebert, U. (2019). Economic burden of stroke: a systematic review on post-stroke care. *European Journal of Health Economics*, *20*(1), 107–134. https://doi.org/10.1007/s10198-018-0984-0

Ram, C. V. S., Kumar, S., Renjen, P. N., Kumar, G. P., Swaminathan, J., Reddy, C. R., Kondati, S., Sharma, M., Selvan, V. L. A., Sundaram, M., Vasudevan, A., & Lackland, D. (2021). Risk factors predisposing to acute stroke in India: a prospective study.

Journal of Hypertension, *39*(11), 2183–2189. https://doi.org/10.1097/HJH.0000000000002915

Reddin, C., Murphy, R., Hankey, G. J., Judge, C., Xavier, D., Rosengren, A., Ferguson, J., Alvarez-Iglesias, A., Oveisgharan, S., Iversen, H. K., Lanas, F., Al-Hussein, F., Członkowska, A., Oguz, A., McDermott, C., Pogosova, N., Málaga, G., Langhorne, P., Wang, X., … O'Donnell, M. (2022). Association of psychosocial stress with risk of acute stroke. *JAMA Network Open*, *5*(12), E2244836. https://doi.org/10.1001/jamanetworkopen.2022.44836

Rochmah, T. N., Rahmawati, I. T., Dahlui, M., Budiarto, W., & Bilqis, N. (2021). Economic burden of stroke disease: A systematic review. *International Journal of Environmental Research and Public Health*, *18*(14). https://doi.org/10.3390/ijerph18147552

Roemer, A., Cervin, M., Medvedeva, A., Bravo, A. J., & Medvedev, O. N. (2024). Big five of mindfulness and personality: Cross-cultural network analysis. *Mindfulness*, *15*(1), 37–47. https://doi.org/10.1007/s12671-023-02293-3

Room, R. (2004). Smoking and drinking as complementary behaviours. *Biomedicine and Pharmacotherapy*, *58*(2), 111–115. https://doi.org/10.1016/j.biopha.2003.12.003

Saini, N. R., & Gurvendra, A. L. (2022). Stroke-related risk factors: A review. *Asian Pacific Journal of Health Sciences*, *9*(3), 102–107. https://doi.org/10.21276/apjhs.2022.9.3.21

Sajjad, A., Mirza, S. S., Portegies, M. L. P., Bos, M. J., Hofman, A., Koudstaal, P. J., Tiemeier, H., & Arfan Ikram, M. (2015). Subjective memory complaints and the risk of stroke. *Stroke*, *46*(1), 170–175. https://doi.org/10.1161/STROKEAHA.114.006616

Saxena, A., Paredes-Echeverri, S., Michaelis, R., Popkirov, S., & Perez, D. L. (2022). Using the Biopsychosocial Model to Guide Patient-Centered Neurological Treatments. *Seminars in Neurology*, *42*(2), 80–87. https://doi.org/10.1055/s-0041-1742145

Scutari, M., & Nagarajan, R. (2013). Identifying significant edges in graphical models of molecular networks. *Artificial Intelligence in Medicine*, *57*(3), 207–217. https://doi.org/10.1016/j.artmed.2012.12.006

Scutari, M., & Silander, T. (2024). *Package "bnlearn": Bayesian Network Structure Learning, Parameter Learning and Inference* (5.0.1). https://www.bnlearn.com/

Senff, J., Tack, R. W. P., Mallick, A., Gutierrez-Martinez, L., Duskin, J., Kimball, T. N., Tan, B. Y. Q., Chemali, Z. N., Newhouse, A., Kourkoulis, C., Rivier, C., Falcone, G. J., Sheth, K. N., Lazar, R. M., Ibrahim, S., Pikula, A., Tanzi, R. E., Fricchione, G. L., Brouwers, H. B., … Singh, S. D. (2025). Modifiable risk factors for stroke, dementia and late-life depression: a systematic review and DALY-weighted risk factors for a composite outcome. *Journal of Neurology, Neurosurgery, and Psychiatry*. https://doi.org/10.1136/jnnp-2024-334925

Smyth, A., O'Donnell, M., Rangarajan, S., Hankey, G. J., Oveisgharan, S., Canavan, M., McDermott, C., Xavier, D., Zhang, H., Damasceno, A., Avezum, A., Pogosova, N., Oguz, A., Ryglewicz, D., Iversen, H. K., Lanas, F., Rosengren, A., Yusuf, S., & Langhorne, P. (2023). Alcohol intake as a risk factor for acute stroke: The INTERSTROKE study. *Neurology*, *100*(2), E142–E153. https://doi.org/10.1212/WNL.0000000000201388

Sol, K., Zaheed, A. B., Kraal, A. Z., Sharifian, N., Arce Rentería, M., & Zahodne, L. B. (2020). Psychological predictors of memory decline in a racially and ethnically diverse longitudinal sample of older adults in the United States. *International Journal of Geriatric Psychiatry*, *35*(2), 204–212. https://doi.org/10.1002/gps.5236

Sperl, M. A., Esterov, D., Ransom, J. E., Mielke, M. M., Witkowski, J. E., & Brown, A. W. (2022). Long-Term Risk of Stroke after Traumatic Brain Injury: A Population-Based

Medical Record Review Study. *Neuroepidemiology*, *56*(4), 283–290. https://doi.org/10.1159/000525111

Stewart, I. J., Howard, J. T., Amuan, M. E., Kennedy, E., Balke, J. E., Poltavskiy, E., Walker, L. E., Haigney, M., & Pugh, M. J. (2025). Traumatic brain injury is associated with the subsequent risk of atrial fibrillation or atrial flutter. *Heart Rhythm*, *22*(3), 661–667. https://doi.org/10.1016/j.hrthm.2024.09.019

Sur, N. B., Kozberg, M., Desvigne-Nickens, P., Silversides, C., & Bushnell, C. (2024). Improving stroke risk factor management focusing on health disparities and knowledge gaps. *Stroke*, *55*(1), 248–258. https://doi.org/10.1161/STROKEAHA.122.040449

Tack, R. W. P., Senff, J. R., Kimball, T. N., Prapiadou, S., Tan, B. Y. Q., Yechoor, N., Rosand, J., Singh, S. D., & Anderson, C. D. (2025). Reliability and validity of self-reported risk factors for stroke and dementia. *Journal of the American Heart Association*, *14*(7). https://doi.org/10.1161/JAHA.124.038730

Thomann, A. K., Knödler, L. L., Karthikeyan, S., Atanasova, K., Bernstein, C. N., Ebert, M. P., Lis, S., & Reindl, W. (2023). The interplay of biopsychosocial factors and quality of life in inflammatory bowel diseases: A network analysis. *Journal of Clinical Gastroenterology*, *57*(1), 57–65. https://doi.org/10.1097/MCG.0000000000001625

Thompson, S. G., Barber, P. A., Gommans, J. H., Cadilhac, D. A., Davis, A., Fink, J. N., Harwood, M., Levack, W., Mcnaughton, H., Feigin, V. L., Abernethy, V., Girvan, J., Denison, H., Corbin, M., Wilson, A., Douwes, J., & Ranta, A. (2022). The impact of ethnicity on stroke care access and patient outcomes: a New Zealand nationwide observational study. *The Lancet Regional Health - Western Pacific*. https://doi.org/10.1016/j

Tomás, J. M., Oliver, A., Torres, Z., Parker, J., Marques-Sule, E., & Sentandreu-Mañó, T. (2023). A biopsychosocial model predicting myocardial infarction. *Journal of Clinical Medicine*, *12*(17). https://doi.org/10.3390/jcm12175715

Tran, T. V., DeMarco, R., Chan, K., & Nguyen, T. N. (2015). Race, age and serious psychological distress in the USA. *International Journal of Culture and Mental Health*, *8*(2), 162–175. https://doi.org/10.1080/17542863.2014.913644

Wang, X., Huang, Y., Chen, Y., Yang, T., Su, W., Chen, X., Yan, F., Han, L., & Ma, Y. (2022). The relationship between body mass index and stroke: A systemic review and meta-analysis. *Journal of Neurology*, *269*(12), 6279–6289. https://doi.org/10.1007/s00415-022-11318-1

Warren, S. L., Reid, E., Whitfield, P., & Moustafa, A. A. (2022). Subjective memory complaints as a predictor of mild cognitive impairment and Alzheimer's disease. *Discover Psychology*, *2*(1). https://doi.org/10.1007/s44202-022-00031-9

Weston, W. W. (2005). Patient-centered medicine: A guide to the biopsychosocial model. *Families, Systems and Health*, *23*(4), 387–392. https://doi.org/10.1037/1091-7527.23.4.387

Williams, D. (2019). *Bayesian estimation for gaussian graphical models: Structure learning, predictability, and network comparisons*. 1–7.

Williams, & Mulder. (2020). BGGM: Bayesian gaussian graphical models in r. *Journal of Open Source Software*, *5*(51), 1–4. https://doi.org/10.21105/joss.02111

Wise, J. (2024). Global life expectancy to increase by almost five years by 2050, study predicts. *BMJ*. https://doi.org/10.1136/bmj.q1126

Wolf, P. A., D'Agostino, R. B., Belanger, A. J., & Kannel, W. B. (1991). Probability of stroke: A risk profile from the Framingham study. *Stroke*, *22*(3). http://ahajournals.org

Xu, Y., Zhao, L., Zhang, L., Han, Y., Wang, P., & Yu, S. (2020). Left atrial enlargement and the risk of stroke: A meta-analysis of prospective cohort studies. *Frontiers in Neurology*, *11*. https://doi.org/10.3389/fneur.2020.00026

Yu, J., Li, J., Yu, H., & Zhou, Q. (2025). Influence of lifestyle on stroke risk among adults over 40 years in northern China: A retrospective case-control study. *Medicine*, *104*(45), e45707. https://doi.org/10.1097/MD.0000000000045707

Appendix A

Ethics Approval

Te Wānanga o Ngā Kete | ***Division of Arts, Law, Psychology & Social Sciences***

The University of Waikato
Private Bag 3105
Hamilton 3240
New Zealand

School of Psychological and Social Sciences
Psychology Programme
Dr Amy Bird
Tel: +64 7 838 4094
Email: amy.bird@waikato.ac.nz
www.waikato.ac.nz

THE UNIVERSITY OF WAIKATO
Te Whare Wānanga o Waikato

Kelly Olsen

Dr Oleg Medvedev

School of Psychological and Social Sciences
Psychology Programme

17 February 2025

Dear Kelly

Re: **FS2025-03: Investigating Associations between Risk Factors of Stroke: A Cross-Cultural Network Analysis**

Thank you for submitting an application to the ALPSS Human Research Ethics Committee. We have reviewed your application and the Committee is pleased to offer formal approval for your research activities.

We encourage you to contact the committee should issues arise during your data collection, or should you wish to add further research activities or make changes to your project as it unfolds. We wish you all the best with your research. Thank-you for engaging with the process of Ethical Review.

Kind regards,

Dr Amy Bird, Convenor
Division of Arts, Law, Psychology & Social Sciences Human Research Ethics

Appendix B

R-Software Code for BGGM and DAG Analyses

```
library(BGGM)
library(readxl)
library(qgraph)

setwd("/Users/kellyjolsen/Desktop/MASTERS 2024/r analysis")

df <- read_excel("stroke.xlsx")

#### #### #### #### #### #### #### #### #### #### ####
##### Explore in White/European data
#### #### #### #### #### #### #### #### #### #### ####

### Explore structure in White/European data
stroke <- subset(df, Ethnicity==1)

#----------- Split sample
## 50/50%
smp_size <- floor(0.5 * nrow(stroke))

## set the seed to make partition reproducible
set.seed(4321)

half <- sample(seq_len(nrow(stroke)), size = smp_size)

ensample <- stroke[half, ] # 5497 obs
cnsample <- stroke[-half, ] # 5498 obs

#----------- Explore

fitstroke <- explore(ensample[1:19], type = "continuous", prior_sd = 0.25, iter = 5000,
                 progress = TRUE, impute = F, seed = 1)
```

```
selstroke <- BGGM::select(fitstroke)

names1<-c("Smoke", "Alcohol", "Diet", "Exercise", "BMI",
        "Distress", "Dementia", "Memory", "HD", "EHeart", "AF", "BP", "BPMed",
"PStroke", "Diabetes", "TBI", "HStroke", "Age", "Female")

names2<-c("Smoking", "Alcohol Use", "Healthy Diet", "Exercise", "Body Mass Index",
        "Psychological Distress", "Dementia", "Memory Deficit", "Heart Disease", "Enlarged
Heart", "Atrial Fibriallation", "Systolic Blood Pressure", "Blood Pressure Medication",
"Parental Stroke", "Diabetes", "Traumatic Brain Injury", "History of Stroke", "Age", "Being
Female")

col<-c("papayawhip", "lightsteelblue2", "brown1", "gray65")

gr<-list(c(1,2,3,4,5),c(6,7,8),c(9,10,11,12,13,15,16,17), c(14,18,19))

names(gr)<-c("Lifestyle Factors", "Psychological Health Factors", "Physical Health Factors",
"Demographic Factors")

pdf("enw_stroke.pdf",width = 16, height = 12)
qgraph(selstroke$pcor_mat_zero,layout="spring",
      posCol="royalblue4", negCol="firebrick3", palette="pastel",
      groups=gr,color=col,esize=10,vsize=8,
      labels=names1, nodeNames = names2, label.scale.equal=T)
dev.off()

#### #### #### #### #### #### #### #### #### #### ####
#### Run confirmatory tests with new White/European data
#### #### #### #### #### #### #### #### #### #### ####

fitstroke2 <- estimate(cnsample[1:19], type = "continuous", prior_sd = 0.25, iter = 5000,
                  progress = TRUE, impute = F, seed = 1)
selstroke2 <- BGGM::select(fitstroke2)
```

```
## Print network using exploratory layout to increase interpretability

Lay<-qgraph(selstroke$pcor_mat_zero, layout = "spring", DoNotPlot = TRUE)
Lay      <- qgraph::averageLayout(Lay)

pdf("enw_stroke2.pdf",width = 16, height = 12)
qgraph(selstroke2$pcor_adj,layout=Lay,
     posCol="royalblue4", negCol="firebrick3", palette="pastel",
     groups=gr,color=col,esize=10,vsize=8,
     labels=names1, nodeNames = names2, label.scale.equal=T)
dev.off()

# Extracting the partial correlation matrix for ensample
pcor_matrix <- selstroke$pcor_mat_zero

# View the partial correlation matrix
print(pcor_matrix)

# Extracting the partial correlation matrix for cnsample
# Assuming the initial setup and the analysis for ensample is correct and remains unchanged.

#### Confirmatory/Second Sample Analysis for cnsample ####
# Use explore() for consistency, if that's your intent

# Adjusting the code to use explore() for cnsample, similar to ensample
fitstroke2 <- explore(cnsample[1:19], type = "continuous", prior_sd = 0.25, iter = 5000,
                progress = TRUE, impute = F, seed = 1)
selstroke2 <- BGGM::select(fitstroke2)

# Extracting the partial correlation matrix for cnsample, similar to ensample
pcor_matrix2 <- selstroke2$pcor_mat_zero  # Ensure this is the correct output you're
expecting
```

```
# View the partial correlation matrix for cnsample
print(pcor_matrix2)

#----------- Predictability Plots -----------

r2 <- predictability(fitstroke)
plot(r2)

r2c <- predictability(fitstroke2)
plot(r2c)

#### Comparing Links Between Two Subsamples ####

# Extracting the partial correlation matrix for ensample
library(BGGM)
library(readxl)
library(qgraph)

# Assuming you've set the working directory and loaded the data as per your initial script.
df <- read_excel("stroke.xlsx")

# Subsetting and sampling as per your initial script
stroke <- subset(df, Ethnicity == 1)
smp_size <- floor(0.5 * nrow(stroke))
set.seed(4321)
half <- sample(seq_len(nrow(stroke)), size = smp_size)
ensample <- stroke[half, ] # 5497 obs
cnsample <- stroke[-half, ] # 5498 obs

# Your analysis for ensample and cnsample here...
# Assuming this part includes fitting models and was performed correctly.

# Compare networks
```

```
fitcompare <- ggm_compare_estimate(as.matrix(ensample[1:19]), as.matrix(cnsample[1:19]),
type = "continuous", prior_sd = 0.25, iter = 5000, progress = TRUE, seed = 1)
# Plotting the differences - you mentioned this results in a mess, so we'll replace this with
table generation.
# pdf("diff_en_cn_samples_final.pdf")
# plot(summary(fitcompare))
# dev.off()

# Instead of plotting, we extract the summary into a more readable format
summary_fitcompare <- summary(fitcompare)

summary_fitcompare <- summary(fitcompare)
print(summary_fitcompare)
str(summary_fitcompare)

#### #### #### #### #### #### #### #### #### #### ####
### Explore in African ethnicity data
#### #### #### #### #### #### #### #### #### #### ####
stroke_african <- subset(df, Ethnicity == 2)

#----------- Explore
fitstroke_african <- explore(stroke_african[1:19], type = "continuous", prior_sd = 0.25, iter =
5000, progress = TRUE, impute = F, seed = 1)
selstroke_african <- BGGM::select(fitstroke_african)

## Print network to increase interpretability
Lay <- qgraph(selstroke_african$pcor_mat_zero, layout = "spring", DoNotPlot = TRUE)
Lay <- qgraph::averageLayout(Lay)

pdf("nw_African.pdf", width = 16, height = 12)
qgraph(selstroke2_african$pcor_adj, layout = Lay,
    posCol = "royalblue4", negCol = "firebrick3", palette = "pastel",
    groups = gr, color = col, esize = 10, vsize = 8,
    labels = names1, nodeNames = names2, label.scale.equal = TRUE)
```

```
dev.off()

# Extracting the partial correlation matrix for total African ethnicity data
pcor_matrix_african <- selstroke_african$pcor_mat_zero

# View the partial correlation matrix for total African ethnicity data
print(pcor_matrix_african)

# Predictability Plots for African ethnicity data
r2_african <- predictability(fitstroke_african)
plot(r2_african)

#### #### #### #### #### #### #### #### #### #### ####
### Explore in Asian ethnicity data
#### #### #### #### #### #### #### #### #### #### ####
stroke_asian <- subset(df, Ethnicity == 3)

#----------- Explore
fitstroke_asian <- explore(stroke_asian[1:19], type = "continuous", prior_sd = 0.25, iter =
5000, progress = TRUE, impute = F, seed = 1)

selstroke_asian <- BGGM::select(fitstroke_asian)

## Print network to increase interpretability
Lay <- qgraph(selstroke_asian$pcor_mat_zero, layout = "spring", DoNotPlot = TRUE)
Lay <- qgraph::averageLayout(Lay)

pdf("nw_Asian.pdf", width = 16, height = 12)
qgraph(selstroke_asian$pcor_mat_zero,layout="spring",
    posCol="royalblue4", negCol="firebrick3", palette="pastel",
    groups=gr,color=col,esize=10,vsize=8,
    labels=names1, nodeNames = names2, label.scale.equal=T)
dev.off()
```

```
# Extracting the partial correlation matrix for total Asian ethnicity data
pcor_matrix_asian <- selstroke_asian$pcor_mat_zero
# View the partial correlation matrix for total Asian ethnicity data
print(pcor_matrix_asian)

# Predictability Plots for Asian ethnicity data
r2_asian <- predictability(fitstroke_asian)
plot(r2_asian)

#### #### #### #### #### #### #### #### #### #### ####
### Explore in Indian ethnicity data
#### #### #### #### #### #### #### #### #### #### ####
stroke_indian <- subset(df, Ethnicity == 4)

#----------- Explore
fitstroke_indian <- explore(stroke_indian[1:19], type = "continuous", prior_sd = 0.25, iter =
5000, progress = TRUE, impute = F, seed = 1)

selstroke_indian <- BGGM::select(fitstroke_indian)

## Print network to increase interpretability
Lay <- qgraph(selstroke_indian$pcor_mat_zero, layout = "spring", DoNotPlot = TRUE)
Lay <- qgraph::averageLayout(Lay)

pdf("nw_Indian.pdf", width = 16, height = 12)
qgraph(selstroke_indian$pcor_mat_zero,layout="spring",
      posCol="royalblue4", negCol="firebrick3", palette="pastel",
      groups=gr,color=col,esize=10,vsize=8,
      labels=names1, nodeNames = names2, label.scale.equal=T)
dev.off()

# Extracting the partial correlation matrix for total Indian ethnicity data
pcor_matrix_indian <- selstroke_indian$pcor_mat_zero
```

```
# View the partial correlation matrix for total Indian ethnicity data
print(pcor_matrix_indian)
# Predictability Plots for Indian ethnicity data
r2_indian <- predictability(fitstroke_indian)
plot(r2_indian)

#### #### #### #### #### #### #### #### #### #### ####
### Explore in Hispanic ethnicity data
#### #### #### #### #### #### #### #### #### #### ####
stroke_hispanic <- subset(df, Ethnicity == 5)

#----------- Explore
fitstroke_hispanic <- explore(stroke_hispanic[1:19], type = "continuous", prior_sd = 0.25, iter
= 5000, progress = TRUE, impute = F, seed = 1)

selstroke_hispanic <- BGGM::select(fitstroke_hispanic)

## Print network to increase interpretability
Lay <- qgraph(selstroke_hispanic$pcor_mat_zero, layout = "spring", DoNotPlot = TRUE)
Lay <- qgraph::averageLayout(Lay)

pdf("nw_Hispanic.pdf", width = 16, height = 12)
qgraph(selstroke_hispanic$pcor_mat_zero,layout="spring",
      posCol="royalblue4", negCol="firebrick3", palette="pastel",
      groups=gr,color=col,esize=10,vsize=8,
      labels=names1, nodeNames = names2, label.scale.equal=T)
dev.off()

# Extracting the partial correlation matrix for total Hispanic ethnicity data
pcor_matrix_hispanic <- selstroke_hispanic$pcor_mat_zero

# View the partial correlation matrix for total Hispanic ethnicity data
print(pcor_matrix_hispanic)
```

```
# Predictability Plots for Hispanic ethnicity data
r2_hispanic <- predictability(fitstroke_hispanic)
plot(r2_hispanic)

#### #### #### #### #### #### #### #### #### #### ####
### Explore in Other ethnicity data
#### #### #### #### #### #### #### #### #### #### ####
stroke_other <- subset(df, Ethnicity == 6)

#----------- Explore
fitstroke_other <- explore(stroke_other[1:19], type = "continuous", prior_sd = 0.25, iter =
5000, progress = TRUE, impute = F, seed = 1)

selstroke_other <- BGGM::select(fitstroke_other)

## Print network to increase interpretability
Lay <- qgraph(selstroke_other$pcor_mat_zero, layout = "spring", DoNotPlot = TRUE)
Lay <- qgraph::averageLayout(Lay)

pdf("nw_Other.pdf", width = 16, height = 12)
qgraph(selstroke_other$pcor_mat_zero,layout="spring",
    posCol="royalblue4", negCol="firebrick3", palette="pastel",
    groups=gr,color=col,esize=10,vsize=8,
    labels=names1, nodeNames = names2, label.scale.equal=T)
dev.off()

# Extracting the partial correlation matrix for total Other ethnicity data
pcor_matrix_other <- selstroke_other$pcor_mat_zero

# View the partial correlation matrix for total Other ethnicity data
print(pcor_matrix_other)

# Predictability Plots for Other ethnicity data
r2_other <- predictability(fitstroke_other)
```

```
plot(r2_other)

##################################################
#### Print averaged network for other ethnicities
##################################################

# Compute layouts for each ethnicity's network graph
layout_g1 <- qgraph(selstroke$pcor_mat_zero, layout = "spring", DoNotPlot = TRUE)
layout_g2 <- qgraph(selstroke2$pcor_adj, layout = "spring", DoNotPlot = TRUE)
layout_g3 <- qgraph(selstroke_african$pcor_mat_zero, layout = "spring", DoNotPlot =
TRUE)
layout_g4 <- qgraph(selstroke_asian$pcor_mat_zero, layout = "spring", DoNotPlot = TRUE)
layout_g5 <- qgraph(selstroke_indian$pcor_mat_zero, layout = "spring", DoNotPlot =
TRUE)
layout_g6 <- qgraph(selstroke_hispanic$pcor_mat_zero, layout = "spring", DoNotPlot =
TRUE)
layout_g7 <- qgraph(selstroke_other$pcor_mat_zero, layout = "spring", DoNotPlot = TRUE)

# Combine layouts into a list
layout_list <- list(layout_g1, layout_g2, layout_g3, layout_g4, layout_g5, layout_g6,
layout_g7)

# Compute average layout
averaged_layout <- qgraph::averageLayout(layout_list)

# Plot each ethnicity's network graph using the averaged layout
pdf("network_plots.pdf", width = 16, height = 12)
for (i in 1:length(layout_list)) {
  graph_layout <- layout_list[[i]]
  ethnicity_data <- list(selstroke$pcor_mat_zero, selstroke2$pcor_adj,
selstroke_african$pcor_mat_zero,
selstroke_asian$pcor_mat_zero, selstroke_indian$pcor_mat_zero,
selstroke_hispanic$pcor_mat_zero, selstroke_other$pcor_mat_zero)
  ethnicity_graph <- ethnicity_data[[i]]
```

```
  qgraph(ethnicity_graph, layout = averaged_layout, posCol = "royalblue4", negCol =
"firebrick3",
        palette = "pastel", groups = gr, color = col, esize = 10, vsize = 8,
        labels = names1, nodeNames = names2, label.scale.equal = TRUE)}
dev.off()

##################################
# Compare Euro Exploratory with African
##################################

fitcompare <- ggm_compare_estimate(as.matrix(ensample[1:19]), as.matrix(stroke_african
[1:19]), type = "continuous", prior_sd = 0.25, iter = 5000, progress = TRUE, seed = 1)

# Plotting the differences - table generation.
# pdf("diff_en_african_samples_final.pdf")
# plot(summary(fitcompare))
# dev.off()

# Instead of plotting, we extract the summary into a more readable format
summary_fitcompare <- summary(fitcompare)

summary_fitcompare <- summary(fitcompare)
print(summary_fitcompare)
str(summary_fitcompare)

##################################
# Compare Euro Exploratory with Asian
##################################

fitcompare <- ggm_compare_estimate(as.matrix(ensample[1:19]), as.matrix(stroke_asian
[1:19]), type = "continuous", prior_sd = 0.25, iter = 5000, progress = TRUE, seed = 1)

# Plotting the differences - table generation.
```

```
# pdf("diff_en_asian_samples_final.pdf")
# plot(summary(fitcompare))
# dev.off()

# Instead of plotting, we extract the summary into a more readable format
summary_fitcompare <- summary(fitcompare)

summary_fitcompare <- summary(fitcompare)
print(summary_fitcompare)
str(summary_fitcompare)

##################################
# Compare Euro Exploratory with Indian
##################################

fitcompare <- ggm_compare_estimate(as.matrix(ensample[1:19]), as.matrix(stroke_indian
[1:19]), type = "continuous", prior_sd = 0.25, iter = 5000, progress = TRUE, seed = 1)

# Plotting the differences - table generation.
# pdf("diff_en_indian_samples_final.pdf")
# plot(summary(fitcompare))
# dev.off()

# Instead of plotting, we extract the summary into a more readable format
summary_fitcompare <- summary(fitcompare)

summary_fitcompare <- summary(fitcompare)
print(summary_fitcompare)
str(summary_fitcompare)

##################################
# Compare Euro Exploratory with Hispanic
##################################
```

```
fitcompare <- ggm_compare_estimate(as.matrix(ensample[1:19]), as.matrix(stroke_hispanic
[1:19]), type = "continuous", prior_sd = 0.25, iter = 5000, progress = TRUE, seed = 1)
# Plotting the differences - table generation.
# pdf("diff_en_hispanic_samples_final.pdf")
# plot(summary(fitcompare))
# dev.off()

# Instead of plotting, we extract the summary into a more readable format
summary_fitcompare <- summary(fitcompare)
print(summary_fitcompare)
str(summary_fitcompare)

##################################
# Compare Euro Exploratory with Other
##################################

fitcompare <- ggm_compare_estimate(as.matrix(ensample[1:19]), as.matrix(stroke_other
[1:19]), type = "continuous", prior_sd = 0.25, iter = 5000, progress = TRUE, seed = 1)

# Plotting the differences - table generation.
# pdf("diff_en_other_samples_final.pdf")
# plot(summary(fitcompare))
# dev.off()

# Instead of plotting, we extract the summary into a more readable format
summary_fitcompare <- summary(fitcompare)
summary_fitcompare <- summary(fitcompare)
print(summary_fitcompare)
str(summary_fitcompare)

 # Conducting robust DAG with Iterations

library(BGGM)
library(qgraph)
```

```
library(networktools)
library(readxl)
library(igraph)
library(polycor)
library(ggplot2)
library(Rgraphviz)

####import data set stroke
library(readxl)
stroke <- read_excel("strokeDAG.xlsx")
View(strokeDAG)

# rename stroke into df
df <- strokeDAG[1:16]

###define variables (number in square brackets is a number of column containing a variable starting from the left)
Smoke=strokeDAG[,1]
Alcohol=strokeDAG[,2]
Diet=strokeDAG[,3]
Exercise=strokeDAG[,4]
BMI=strokeDAG[,5]
Distress=strokeDAG[,6]
Dementia=strokeDAG[,7]
Memory=strokeDAG[,8]
HD=strokeDAG[,9]
EHeart=strokeDAG[,10]
AF=strokeDAG[,11]
BP=strokeDAG[,12]
BPMed=strokeDAG[,13]
Diabetes=strokeDAG[,14]
TBI=strokeDAG[,15]
HStroke=strokeDAG[,16]
```

```
######## remove any rows with n/a
DataFacets<- na.omit(df)

## DAG
# Load necessary libraries
library(bnlearn)
library(readxl)

# Load your dataset
strokeDAG <- read_excel (Users/kellyjolsen/Desktop/MASTERS 2024/r analysis")

# Keep only needed variables
df <- strokeDAG[, c( "Smoke", "Alcohol", "Diet", "Exercise", "BMI", "Distress",
"Dementia", "Memory", "HD", "EHeart", "AF", "BP", "BPMed", "Diabetes", "TBI",
"HStroke", "Age", "Female", "PStroke")]

# Remove missing values
DataFacets <- na.omit(df)

# Create a blacklist: these variables cannot be outcomes
non_outcomes <- c("Age", "Female", "PStroke", "TBI", "HStroke")
all_vars <- colnames(DataFacets)
blacklist <- do.call(rbind, lapply(non_outcomes, function(target) {data.frame(from =
setdiff(all_vars, target), to = target)}))

# Run bootstrapped DAG learning with Hill-Climbing algorithm
set.seed(123)
bootnet <- boot.strength(data = DataFacets, R = 5000, algorithm = "hc", algorithm.args = list(
 restart = 5, perturb = 10, blacklist = blacklist), cpdag = TRUE, debug = TRUE)

# Save the bootnet result
write.csv(bootnet, "C:/Users/kellyjolsen/Desktop/bootnet_result.csv", row.names = FALSE)
```

```
# Generate the averaged network
avgnet1 <- averaged.network(bootnet)

# Evaluate arc strengths
astr1 <- arc.strength(avgnet1, data = DataFacets, criterion = "bic-g")
write.csv(astr1, "C:/Users/kellyjolsen/Desktop/bic_result.csv", row.names = FALSE)

# Determine optimal threshold (e.g., 85th percentile)
thresh <- quantile(astr1$strength, 0.85)

# Plot the network using strength.plot
pdf("C:/Users/olegm/OneDrive/Desktop/meaningful_dag_plot.pdf", width = 10, height = 8)
strength.plot(avgnet1, strength = astr1, shape = "ellipse", threshold = thresh)
dev.off()
```