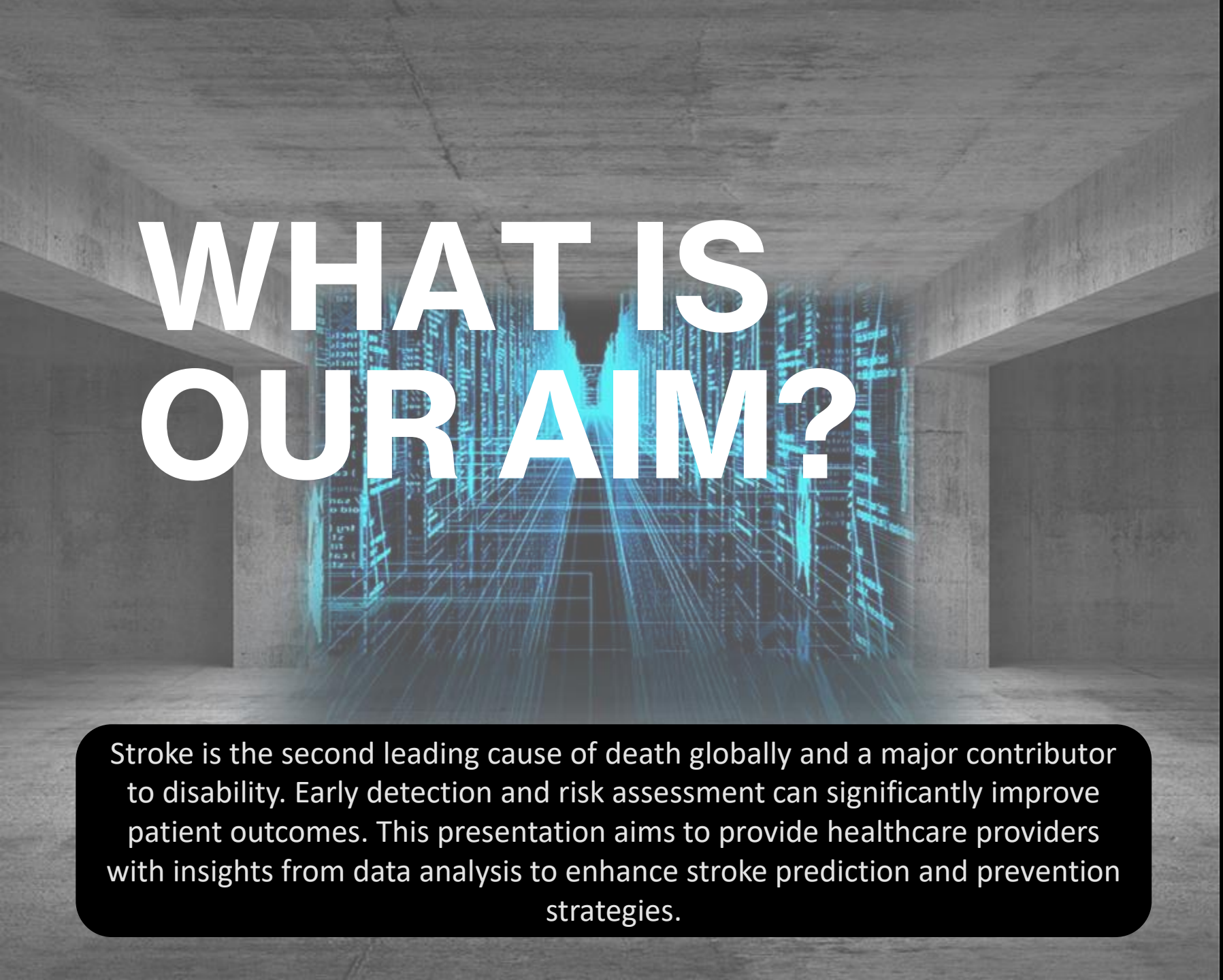
STROKE PREDICTIONS USING MACHINE LEARNING

A Data-Driven Approach

By Ismaa-eel Hendricks

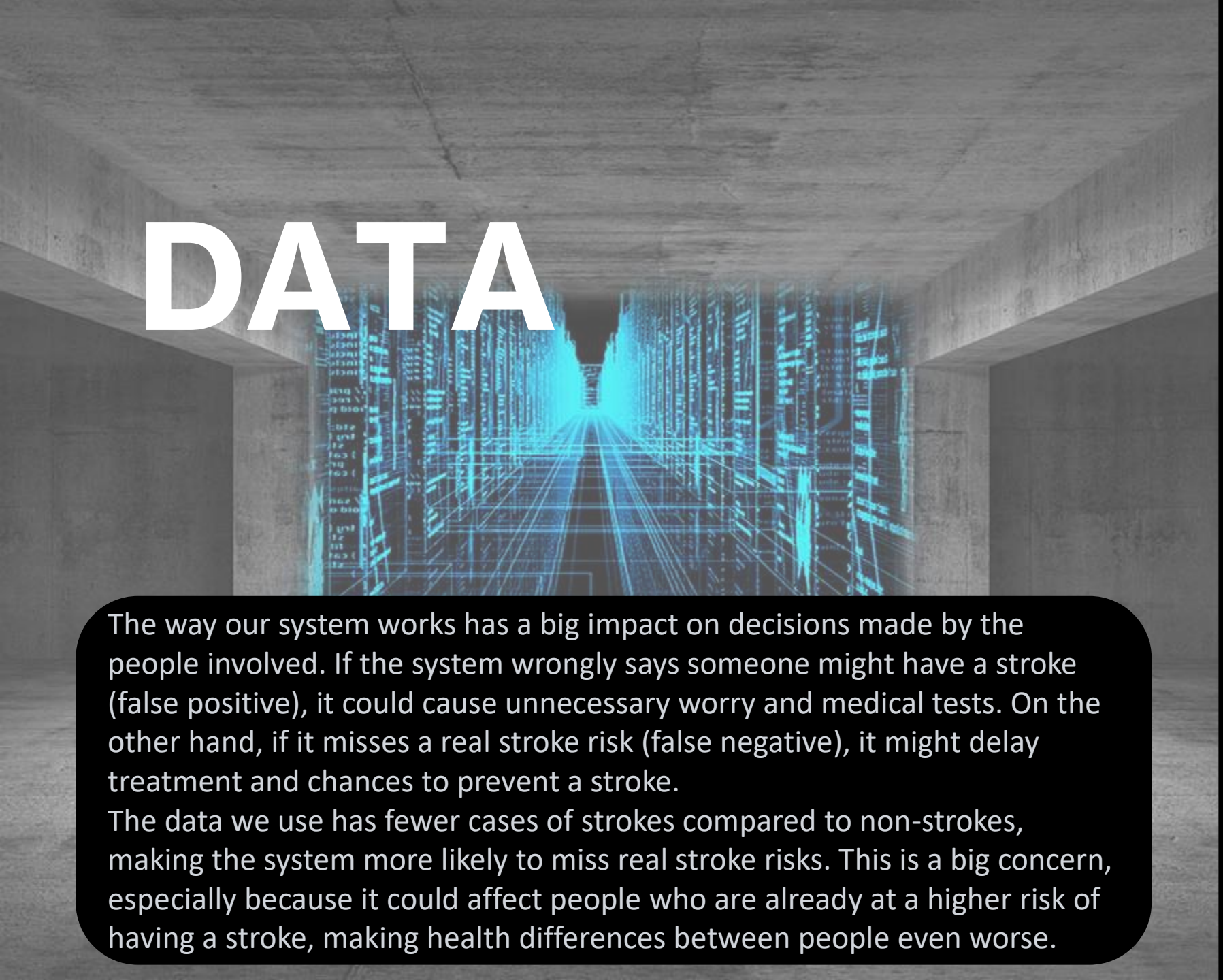


WHAT IS OUR AIM?

Stroke is the second leading cause of death globally and a major contributor to disability. Early detection and risk assessment can significantly improve patient outcomes. This presentation aims to provide healthcare providers with insights from data analysis to enhance stroke prediction and prevention strategies.

Our aim is to identify the leading lifestyle factors that contribute towards strokes by analyzing a set of lifestyle factors collected from a sample of individuals.

DATA

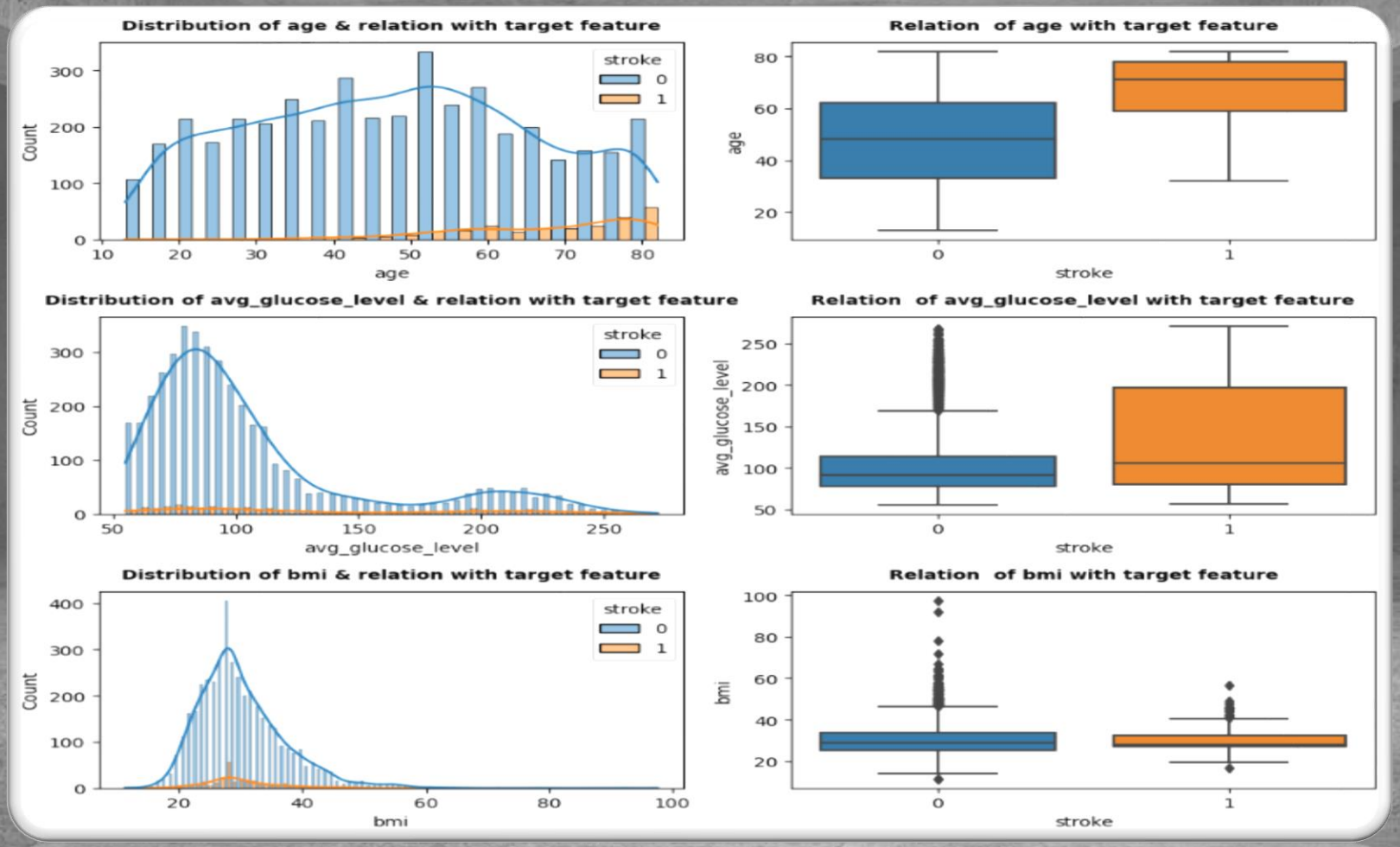


The way our system works has a big impact on decisions made by the people involved. If the system wrongly says someone might have a stroke (false positive), it could cause unnecessary worry and medical tests. On the other hand, if it misses a real stroke risk (false negative), it might delay treatment and chances to prevent a stroke.

The data we use has fewer cases of strokes compared to non-strokes, making the system more likely to miss real stroke risks. This is a big concern, especially because it could affect people who are already at a higher risk of having a stroke, making health differences between people even worse.

The Stroke Prediction Dataset from Kaggle comprises 5110 patient records, each characterized by 12 attributes, including demographic information, medical history, and lifestyle factors. This dataset provides a rich source of information for understanding stroke risk factors.

INSIGHTS

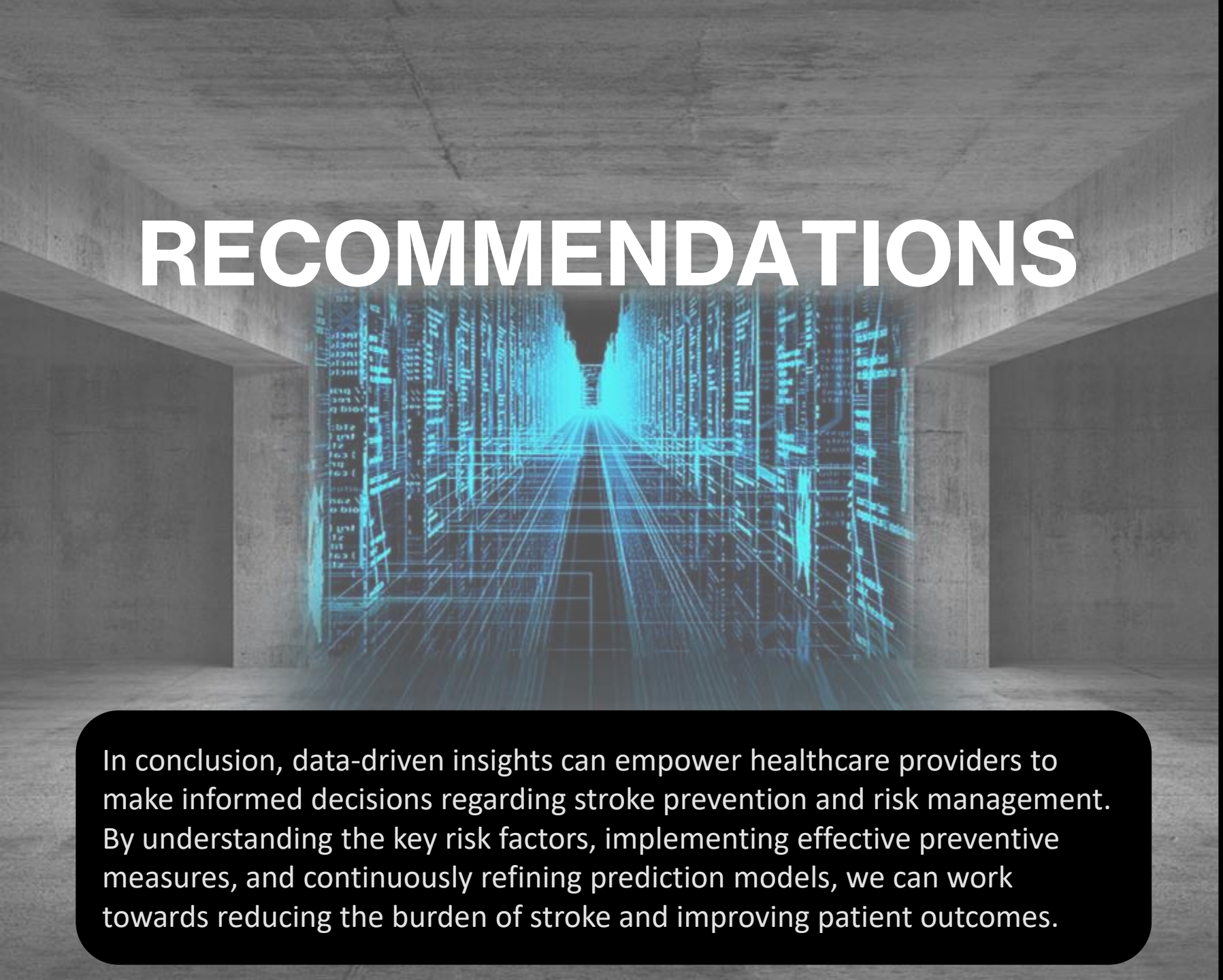


The first graph shows the distribution of age with the target feature. We can see that the majority of people who have had a stroke are over the age of 60. There are also a few people who have had a stroke under the age of 60, but these are much less common.

The second graph shows the distribution of glucose level with the target feature. We can see that people with higher glucose levels are more likely to have had a stroke. This is because high glucose levels can damage the blood vessels, which can lead to stroke.

The third graph shows the distribution of BMI with the target feature. We can see that people with higher BMIs are more likely to have had a stroke. This is because obesity is a risk factor for many chronic diseases, including stroke.

RECOMMENDATIONS



In conclusion, data-driven insights can empower healthcare providers to make informed decisions regarding stroke prevention and risk management. By understanding the key risk factors, implementing effective preventive measures, and continuously refining prediction models, we can work towards reducing the burden of stroke and improving patient outcomes.

1.Prioritize preventive measures, such as hypertension management and smoking cessation, for high-risk individuals.

2.Regularly monitor risk factors and provide personalized healthcare interventions to reduce stroke risk.

3.Implement data-driven risk assessment tools to identify individuals at high risk of stroke.

4.Conduct further research to improve the accuracy and generalizability of stroke prediction models, considering class imbalance and data quality.

5.Enhance healthcare provider education and training on stroke risk assessment and prevention strategies.