

Big data solutions for stroke research

A new comprehensive big data centre in China is set to **TRANSFORM KNOWLEDGE ABOUT STROKE**.

Stroke is a global challenge, affecting more than 13 million people each year. The majority of stroke patients are left with permanent disabilities that undermine their independence, and create a tremendous healthcare burden. Here, Yongjun Wang, chief neurologist and president of Beijing Tiantan Hospital, the founding director of the Centre of excellence for Omics Research (CORe), and deputy director of the China National Clinical Research Center for Neurological Diseases, shares the latest contributions from his team towards stroke research.

What are the major global impacts of stroke?

The World Stroke Organization predicts 14.5 million people will have a stroke in 2021 and 30% of survivors will suffer a recurrence with complex causes. In China, stroke is the leading factor of death, and the average age for patients is about 62, 5-10 years younger than that in western countries.

We lack knowledge about stroke. It's vital that we understand this disease in multiple dimensions, including genetic cues, epigenetic regulations, protein integrations, and even

environmental and socio-economic factors.

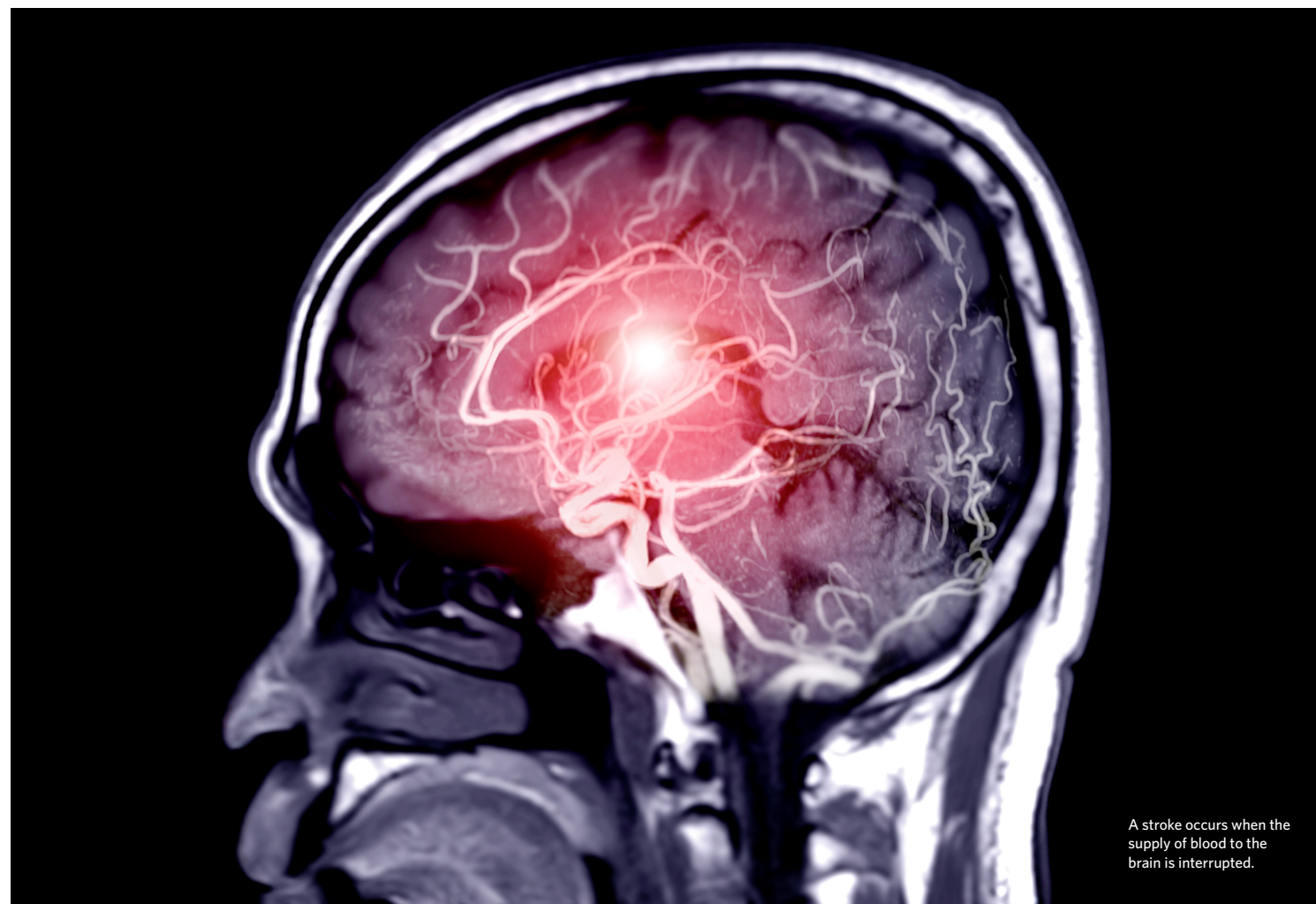
What is omics research in stroke?

Omics research is an emerging field in precision medicine using big data science. The integration of omics data means that thousands of genes, proteins, RNAs, and metabolites can be studied simultaneously, revealing interaction networks at molecular levels.

We set a standard for stroke omics data, which should include both clinical observations (disease characteristics, treatment, progress, CT/MRI images, blood/urine biomarkers) and molecular information (DNA sequences, gene and protein expression profiles, epigenetic modifications, circulating metabolites). The new centre, CORe, is responsible for the collection, management and analysis of stroke omics data.

How can CORe help improve stroke outcomes?

We won't be able to solve the problem until we understand it. CORe's mission is to produce new knowledge of stroke by big-data technologies, incorporating high-throughput sequencing,



A stroke occurs when the supply of blood to the brain is interrupted.

mass-spectrometry and artificial intelligence computation. Only then, we will have sufficient big data facts to redirect basic research. We call this a bedside-to-bench strategy.

To have a systemic view of stroke, we have been collecting data from the Third China National Stroke Registry (CNSR-III), involving 15,166 patients from 201 sites that cover 22 provinces and 4 municipalities across China. For each patient, we collect omics data and integrate it in a big database, representing a complex map of stroke. So far, this consists of 6 million GB of data about stroke.

With this map, we can direct basic research more efficiently to determine various factors contributing to stroke and relapse. For example, it will tell us about how blood proteins contribute, and which genes are responsible for this association. This knowledge of disease-causing genes and progression will improve our precision therapy.

What are the major advantages of this omics research?

Omics data-driven drug discovery will be more efficient, precise, and cost-effective. Traditional research selects drug targets that often lack direct

clinical relevance, leading to results that are not reproducible. Omics research, by integrating big data, allows us to identify more reliable biological markers related to stroke progress, and thus facilitate drug target identification. We estimate that this approach will help to reduce the drug discovery timeframe from 15 to 7 years.

Has the CORe yielded any significant research results so far?

An important observation is that the urban disease pattern in China has been more similar to that in western countries, which disproves our old presumption

that East Asians are more susceptible to stroke.

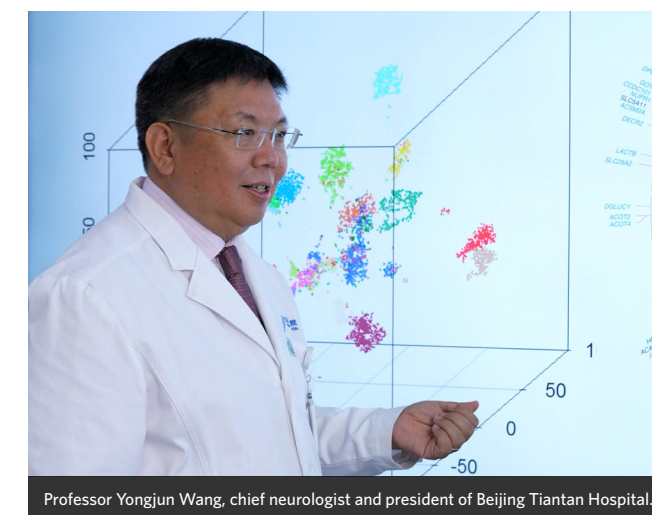
Additionally, CORe data has helped to optimize China's clinical guidelines for stroke treatment. In drug discovery, we have identified nine new targets and two are being tested in clinical trials.

What are your future expectations for the CORe?

CORe manages the world's most comprehensive stroke omics database, which is expected to lead global collaborations against stroke. The CORe database is maintained using international standards for easy exchange and knowledge



Wang's team at the Beijing Tiantan Hospital.



Professor Yongjun Wang, chief neurologist and president of Beijing Tiantan Hospital.

transfer. As a publicly funded institution, CORe is open to society for data sharing. So far, we have shared data with colleagues from Oxford University, Duke University, and Harvard University.

We encourage stroke researchers from all over the world to collaborate with us. We will launch the STROMICS (Stroke Omics) project and website to facilitate collaboration and data sharing. Our database is categorized based on projects with a user-friendly system, in which researchers need to submit their basic information, proposal, and sign the data use agreement

to complete a quick online application.

At the same time, we also hope that researchers will share their data and collaborate with CORe. We believe that sharing and collaborating will bring incredible innovation to stroke research. ■

 首都医科大学附属北京天坛医院
BEIJING TIAN TAN HOSPITAL, CAPITAL MEDICAL UNIVERSITY

 国家神经系统疾病临床医学研究中心
China National Clinical Research Center for Neurological Diseases

 CORe 卒中多组学创新中心
Center of excellence for Omics Research

Email: yongjunwang@nrcnd.org.cn
sicheng@nrcnd.org.cn