ISSN 1881-7815 Online ISSN 1881-7823

BST BioScience Trends

Volume 18, Number 1
February, 2024

IRCA
BSSA
IACMHR

www.biosciencetrends.com

BST
BioScience Trends

ISSN: 1881-7815
Online ISSN: 1881-7823
CODEN: BTIRCZ
Issues/Year: 6
Language: English
Publisher: IACMHR Co., Ltd.

BioScience Trends is one of a series of peer-reviewed journals of the International Research and Cooperation Association for Bio & Socio-Sciences Advancement (IRCA-BSSA) Group. It is published bimonthly by the International Advancement Center for Medicine & Health Research Co., Ltd. (IACMHR Co., Ltd.) and supported by the IRCA-BSSA.

BioScience Trends devotes to publishing the latest and most exciting advances in scientific research. Articles cover fields of life science such as biochemistry, molecular biology, clinical research, public health, medical care system, and social science in order to encourage cooperation and exchange among scientists and clinical researchers.

BioScience Trends publishes Original Articles, Brief Reports, Reviews, Policy Forum articles, Communications, Editorials, News, and Letters on all aspects of the field of life science. All contributions should seek to promote international collaboration.

Editorial Board

Editorial and Head Office

Pearl City Koishikawa 603,
2-4-5 Kasuga, Bunkyo-ku, Tokyo 112-0003, Japan
E-mail: office@biosciencetrends.com

BioScience Trends

Editorial and Head Office

Pearl City Koishikawa 603, 2-4-5 Kasuga, Bunkyo-ku,
Tokyo 112-0003, Japan

E-mail: office@biosciencetrends.com
URL: www.biosciencetrends.com

Editorial Board Members

CONTENTS

Volume 18, Number 1, 2024

Review

Original Article

Letter to the Editor

Review

DOI: 10.5582/bst.2023.01027

Comparison of diagnosis-related groups (DRG)-based hospital payment system design and implementation strategies in different countries: The case of ischemic stroke

Yuan Liu[1,2], Gang Wang[1], Tian-Ge Qin[3], Susumu Kobayashi[2], Takashi Karako[2,4], Peipei Song[2,4,*]

[1] Statistics Center, Shanghai Chest Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China;
[2] Center for Clinical Sciences, National Center for Global Health and Medicine, Tokyo, Japan;
[3] Anqing Medical College, Anqing, Anhui, China;
[4] National College of Nursing, Japan, Tokyo, Japan.

SUMMARY Diagnosis-related groups (DRG) based hospital payment systems are gradually becoming the main mechanism for reimbursement of acute inpatient care. We reviewed the existing literature to ascertain the global use of DRG-based hospital payment systems, compared the similarities and differences of original DRG versions in ten countries, and used ischemic stroke as an example to ascertain the design and implementation strategies for various DRG systems. The current challenges with and direction for the development of DRG-based hospital payment systems are also analyzed. We found that the DRG systems vary greatly in countries in terms of their purpose, grouping, coding, and payment mechanisms although based on the same classification concept and that they have tended to develop differently in countries with different income classifications. In high-income countries, DRG-based hospital payment systems have gradually begun to weaken as a mainstream payment method, while in middle-income countries DRG-based hospital payment systems have attracted increasing attention and increased use. The example of ischemic stroke provides suggestions for mutual promotion of DRG-based hospital payment systems and disease management. How to determine the level of DRG payment incentives and improve system flexibility, balance payment goals and disease management goals, and integrate development with other payment methods are areas for future research on DRG-based hospital payment systems.

Keywords diagnosis-related groups, payment system, hospitals, stroke

1. Introduction

Global health expenditures are increasing. Predictions indicate that spending will increase from $7.9 trillion in 2017 to $11.0 trillion in 2030 (*1*). Hospital expenses account for one of the largest shares of total healthcare expenses in all countries (*2*). Countries are seeking innovations in the methods of paying for hospital care to better allocate healthcare resources, improve hospital efficiency, and control the growth of healthcare costs. In 1983, a diagnosis-related groups (DRG)-based hospital payment system was first introduced as a new prospective case-based reimbursement system for medical care in the United States (U.S.). Since then, a range of DRG-based hospital payment systems have been widely used in inpatient care worldwide in an effort to reduce healthcare costs, such as in Europe and rapidly developing countries in Asia and sub-Saharan Africa (*3*). DRG-based hospital payment systems are gradually becoming the main mechanism for reimbursement of acute inpatient care.

DRG-based hospital payment systems are a form of activity-based funding used to classify hospital care according to the care provided. The basic idea of a DRG-based hospital payment system is that all patients treated by a hospital are classified into a limited number of DRGs, which are supposed to be clinically meaningful and relatively homogenous in their patterns of resource consumption (*4*). Each DRG is associated with a specific cost weight or tariff, and hospitals using a DRG-based hospital payment system either receive a DRG-based case payment or a DRG-based budget allocation. Classifying patients into groups with similar levels of resource use would standardize the case-mix of patients and allow valid comparisons of hospital efficiency and output-based payment. Studies have shown that DRG-

based hospital payment systems largely increased transparency, efficiency, and the quality of hospitals in many countries (*5*).

DRG-based hospital payment systems integrate a wide range of patient information that helps to describe and understand the patient, resulting in care that optimizes patients' needs and goals (*6*). Studies have revealed differences in the ability of DRG-based hospital payment systems to explain variance in the costs and length of stay (LoS) across countries (*7*). A comparison of patient characteristics in classification systems by DRG in different countries can improve the performance of DRG classification and patient control strategies can benefit. With changes in people's lifestyles and global aging, stroke has become the second highest cause of death globally and a leading cause of disability (*8*). According to the Global Burden of Disease estimates, there were around 12.2 million incident cases of stroke, 143 million disability-adjusted life-years lost, and 6.6 million deaths globally in 2019 (*9*). The disease burden of stroke varies widely geographically and economically, with almost 90% of all deaths and disability from stroke occurring in lower-income and middle-income countries, particularly in sub-Saharan Africa and Asia (*10*). Ischemic stroke is the most common type of stroke, causing severe disability to the patient and placing a heavy burden on families and counties. Novel strategies for the prevention and management of stroke are needed in countries around the world. In this article, ischemic stroke is used as an example to summarize the design of existing DRG-based hospital payment systems and experience with their implementation to provide a reference for policymakers in countries concerned about DRG-based payment systems and to provide suggestions for stroke management strategies.

2. Search strategy

We conducted a search of the literature published from January 1983 to December 2023. We started by searching for English-language publications indexed in PubMed with "diagnosis-related groups", "DRG", "diagnosis related group", or "case-mix" in the title, keywords, or abstract. We also searched Google for the same keywords to identify grey literature, books, government reports, *etc*. Following the literature search, identified publications were reviewed and a list of countries with DRG-based payment systems was created. Certain countries that use DRG just for patient classification and not for hospital payments were excluded from the list. Once a list of countries was created, we performed a second literature search with no language restrictions in PubMed and Google that focused on countries on the list, using the name of each country combined with the same keywords mentioned above. In this way we further validated the list of countries while obtaining detailed information on each country's system design and implementation strategy. As we further explored the development of DRG-based hospital payment systems in countries with different income classifications, we used the World Bank's country income classification of 2022 (*11*).

Ischemic stroke was identified as cases with a principal diagnosis coded for cerebral infarction (I63) using the International Classification of Diseases (ICD), 10th edition. Classification variables and grouping algorithms for ischemic stroke cases were retrieved from the newest national DRG systems (*12-15*) and detailed comparisons were made to ascertain similarities and differences in DRG system design across countries.

3. Overview of the global use of DRG-based hospital payment systems

Internationally, a total of 49 countries have introduced DRG-based hospital payment systems as of 2023, in addition to 13 countries that are piloting or exploring the use of DRG-base hospital payment systems. Based on the annual incremental development rate of countries using DRG-base hospital payment systems, the period from 1983 to 2023 can be divided into three stages: birth, slow development (0.9 new countries per year), and rapid development (2.2 new countries per year) (Figure 1A). The U.S. was the world's first country to use a DRG-based hospital payment system in 1983. DRG-based hospital payment systems then entered a phase of slow development in the 20 years from 1984 to 2003, predominantly in European countries. In the two decades since 2004, the use of DRG-based payment systems has entered a phase of rapid development, and the systems are tending to spread globally. DRG-based payment systems have gradually become the principal means of reimbursing hospitals for acute inpatient care in most high-income countries (*5*). Across the Asian and Pacific region, increasing attention is now being paid to the use of DRG-based hospital payment systems as the basis for hospital funding arrangements (*16*). There are already 15 middle-income countries that have introduced DRG-based hospital payment systems, such as China, Malaysia and Thailand, and 12 middle-income countries are piloting or exploring the use of DRG-based payment, such as Vietnam and the Philippines (Figure 1).

4. Comparison of national versions of DRG-based hospital payment systems

There were two ways for countries to introduce DRG-based hospital payment systems, importing one of the already-existing DRG systems from abroad or developing a new DRG system (*17*). The former option requires a well-developed health administration and information system. The latter requires strong team support to meet the context of a particular country's needs. In this paper, we selected countries that have developed unique DRG-based hospital payment systems, including countries that

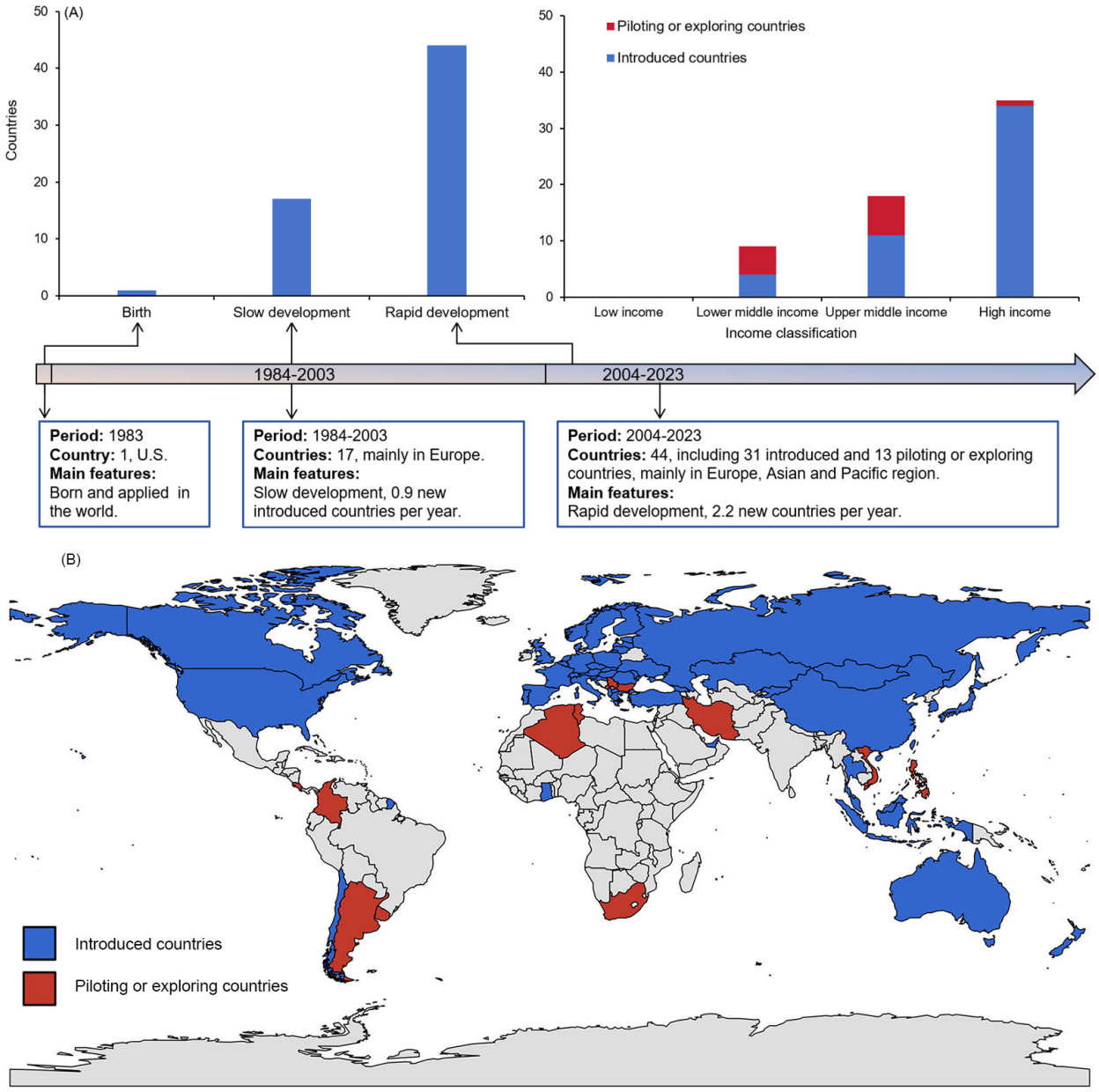

Figure 1. Global use of DRG-based hospital payment systems. (A) The three stages of development of DRG-based hospital payment systems and their distribution in countries with different income classifications. **(B)** Geographic distribution of countries using DRG-based hospital payment systems. *Abbreviation*: U.S., United States.

pioneered DRG-based hospital payment systems (the U.S. and Australia), countries that have long used DRG-based hospital payment systems (England, Sweden, Germany, the Netherlands, Japan, and Thailand), and countries where DRG-based payment systems were recently introduced (the Republic of Korea and China). The similarities and differences of DRG-based hospital payment systems in these countries were compared (Table 1). Learning from different countries' approaches and experiences is important for the development, use, and evolution of DRG-based hospital payment systems around the world.

4.1. Number of groups

The number of groups in DRG-based hospital payment systems mainly ranged from about 600 to over 2,000. Almost all countries, with the exception of the Netherlands, started with a relatively small number of groups when the DRG-based hospital payment system was initially introduced. As the payment system gradually matured, the number of DRG groups tended to increase. In the Republic of Korea, more than 90% of hospitals are private (*18*), and due to strong opposition from private hospitals payment has long been mandated for only 78 groups in the Korea-DRG (K-DRG) (out of 1,880 groups) covering seven conditions involving relatively simple surgery (*e.g.* cataract surgery and appendectomy), while other care is still paid for on a

Table 1. Comparison of different versions of DRG-based hospital payment systems

Country (Year introduced)	Patient classification system	Diagnosis/ procedure codes	No. of groups in 2023	No. of groups initially	Frequency of revision	Previous payment system	Original purposes	Classification variables*	Payment characteristics	*Ref.*
U.S. (1983)	MS-DRG	ICD-10-CM ICD-10-PCS	766	470	Annual	FFS	Improving efficiency and controlling medical costs	Age, sex, and discharge status of the patient 5 Severity levels	The per-discharge payment amount is based on the average national resource use for treating patients under similar circumstances, not including fees for physicians and some education & research.	(*13,22, 24,25*)
Sweden (1992)	NordDRG	ICD-10-SE KVÅ	1463	722	Annual	Global budget	Aiding transparency and management of hospital care	Age, discharge status and LoS 3 severity levels	Sweden has compiled national weights for DRG payments, but each region has a high degree of local autonomy and the usage of DRG payments varies greatly across regions.	(*24,26, 27*)
Australia (1993)	AR-DRG	ICD-10-AM ACHI	789	527	Irregular	Historic budgets	Improved efficiency, equity of resource allocation and healthcare quality	Age, sex, mode of separation, LoS, birth weight, duration of ventilation, and mental health status 5 severity levels	Payment systems focus on measuring disease complexity and impact on resource consumption, allowing for precise grouping.	(*17,28*)
England (2003)	HRG	ICD-10 OPCS-4	2900	610	Annual	Global budget	Patient classification, and increasing the transparency of hospital care	Age, sex, LoS, discharge status, and neoplasms/malignancies 6 severity levels	Tariffs include all operating expenses, staff costs and capital costs, but excludes the costs of education and research. Uses many different exclusion mechanisms and has a number of additional payments.	(*5,29, 30*)
Germany (2003)	G-DRG	ICD-10-GM OPS	1235	664	Annual	Global budget	Reducing health care costs, increasing transparency, and encouraging health system efficiency	Age, sex, birth weight, LoS, duration of ventilation, reason for discharge, and type of admission Unlimited severity levels	Payment includes all costs except costs for investing in/maintaining infrastructure and education & research. Uses many different exclusion mechanisms and has a number of additional payments.	(*22, 30-32*)
Japan (2003)	DPC	ICD-10-CM Japanese original codes	2334 (out of 4726)	1860 (out of 2552)	Biennial	FFS	Standardization, transparency, and accountability of hospital care	Age, sex, birth weight, type of care provided, and ancillary treatment 2 Severity levels	Payments are set on a per diem basis. Surgeries, endoscopies, rehabilitation therapy, devices and drugs given on the day of surgery are not included in the DPC payment, but paid on an FFS basis.	(*14,17, 33*)
Netherlands (2005)	DTC	ICD-10 Health care activity codes	5593	Over 30000	Irregular	Global budget	Facilitating negotiations between purchasers and providers	Medical specialty, type of care, demand for care, diagnosis, treatment axis	Payments could include outpatient care and post-discharge follow-up care and entail a separate fee for medical specialists. Several groups possible per hospital stay.	(*5,15, 20,22*)
China (2011)	CHS-DRG	ICD-10 ICD-9-CM	618 (2020)	108 (out of 650)	Irregular	FFS	Reducing health care costs, and improving the quality of medical care	Age, birth weight, discharge status, LoS, medical costs, and duration of ventilation 3 Severity levels	In the process of promoting DRG-based hospital payment. Payment standards are based on hospitalization fee data from the past few years and differ among provinces. There were four local DRG versions in the beginning; the national version, CHS-DRG, was launched in 2019.	(*12, 34-39*)
Thailand (1998)	Thai-DRG	ICD-10-TM ICD-9-CM	1545	511	Irregular	FFS	Addressing low hospital admission rates and increasing public expenditures on hospitals	Age, sex, discharge status, birth weight, and neoplasms/malignancies 5 Severity levels	Applied to inpatient care at public hospitals and voluntary private hospitals and global budget limits are applied.	(*17,19, 40*)
Republic of Korea (2013)	K-DRG	ICD-10-CM	- (out of 2721)	78 (out of 1880)	Irregular	FFS	Solving problems stemming from overtreatment under the FFS system	Age and sex 4 Severity levels	Facing strong opposition from private hospitals, DRG only covers seven conditions and payment at public hospitals and voluntarily participating private hospitals. Otherwise, payments are on an FFS basis.	(*17,19, 23,41*)

*Except Principal diagnosis and procedure, which are commonly used in all 10 countries. *Abbreviations*: U.S., United States; DRG, diagnosis-related groups; MS, Medicare-severity; ICD, International Classification of Diseases; CM, clinical modification; PCS, Procedure Coding System; FFS, fee-for-service; SE, Swedish modification; KVÅ, Swedish national classification system for surgery and non-surgical procedures; LoS, length of stay; AR, Australia Refined; AM, Australian modification; ACHI, Australian Classification of Health Interventions; HRG, Healthcare Resource Group; OPCS-4, the Office of Population Censuses and Surveys Classification of Interventions and Procedures, version 4; G-DRG, German-DRG; GM, German Modification; OPS, Operation and Procedure Classification; DPC, diagnosis procedure combination; DTC, diagnosis treatment combinations; CHS, China Healthcare Security; TM, Thai Modification; K-DRG, Korean DRG.

fee-for-service (FFS) basis (*19*). In order to facilitate negotiations between healthcare purchasers and providers, the Netherlands initially created more than 30,000 groups of diagnosis treatment combinations (DTCs) reflecting clinical logic more than administrative logic (*20*). However, the number of groups has been drastically reduced to around 5,000 since 2012 due to the high level of complexity and weak operability of DTC. The Japanese Diagnosis Procedure Combination (DPC) is characterized by an emphasis on classifying patients from a clinical perspective, with a total of 4,726 groups in 2023, but payment was provided for only 2,334.

4.2. Coding of diagnosis and procedures

The coding of diagnosis and procedures is important for a DRG-based payment system since this information forms the basis of the definition of patient groups. As shown in Table 1, all 10 countries use the ICD-10 for diagnosis. Significant differences exist since countries usually use the ICD-10 with country-specific modifications, such as the U.S. clinical modification, German modification, or Thailand modification. The classification system for procedures varies greatly in countries, such as the U.S. ICD-10 Procedure Coding System (ICD-10-PCS) or the Australian Classification of Health Interventions (ACHI). Almost every country has developed its own procedure coding system tailored to its needs. Consequently, these systems are very heterogeneous. Germany has converted the ACHI into the Operation and Procedure Classification (OPS) (*21*). In Sweden, the classification of surgery and non-surgical procedures is called KVÅ. The surgical procedures in KVÅ are generally the same as the procedures in the Nordic Medico-Statistical Committee Classification of Surgical Procedures (NCSP), but the medical procedures are national in scope. When introducing a DRG version from another country, attention should be paid to the bias in different coding systems.

4.3. Classification variables

The variables used to define a group and to assign a case can be complex and vary between countries. The variables required usually include clinical variables, demographic and administrative variables, and resource-use variables (*19*). Principal diagnosis and procedures are commonly used in all 10 countries because they provide the basis for the costing or pricing of treatment and resource use. Age, sex, and discharge status are commonly considered as demographic and administrative variables. Resource-use variables indicating the level of severity or complexity of the diagnosis/procedure are used in almost all 10 countries except the Netherlands. The division into severity levels within the classification is usually limited, with up to six levels in the England-Healthcare Resource Group (HRG). In the German DRG (G-DRG) system, the number of severity levels is not, in principle, limited, and up to nine levels are now used (*22*). Other variables such as LoS and the duration of ventilation are used in some countries to classify cases into economically and medically homogenous DRGs (Table 1).

4.4. Original purposes and payment characteristics

There are some differences in the purpose of adopting a DRG-based hospital payment system in various countries. European countries that introduced DRG payment are mainly oriented towards increasing the transparency and efficiency of hospitals (*5*). Japan's DPC/per diem payment system was intended to deliver quality health care and to efficiently construct a clinical database by standardizing information platforms, thereby improving the transparency of hospital care (*17*). South Korea introduced the K-DRG to solve problems stemming from overtreatment under the FFS system (*23*). Although the original purposes vary in countries, there are four main purposes of DRG-based hospital payment systems: to contain costs, increase the efficiency of inpatient care, improve the transparency of hospital care, and to improve the quality of care (Table 1). In South Korea, the K-DRG version 4.4 that was revised in 2021 has 2,721 officially identified DRGs, but only seven disease categories are covered while other care remains on an FFS basis. In the ten countries, only the Netherlands may have several DRGs per hospital stay, and payments are set on a per diem basis only in Japan. The basic characteristics of patient classification systems in the ten countries are summarized in Table 1.

5. DRG-based hospital payment for ischemic stroke

There can be large differences between countries in how they group the same disease. Taking ischemic stroke as an example, different countries have different number of groupings, classification variables, and payment methods (Figure 2). The number of DRGs is similar in China and the U.S., where ischemic stroke patients are divided into about 10 groups, while they are divided into 20 in the Netherlands. In Japan, there are 1,584 groups for ischemic stroke, accounting for 33.5% of the total DPC groups.

A principal diagnosis of stroke is considered at an early stage in the grouping algorithm in almost all 10 countries' DRG systems, while China and the U.S. have a pre-main diagnostic category (pre-MDC) only based on surgical procedures without considering the principal diagnosis. The pre-MDC is generally used for cases that consume a lot of medical resources and that are difficult to classify into other groups. In the U.S. and Netherlands, patients treated with thrombolysis are assigned into specific groups. The presence of relevant complications or comorbidities (CC) influences the

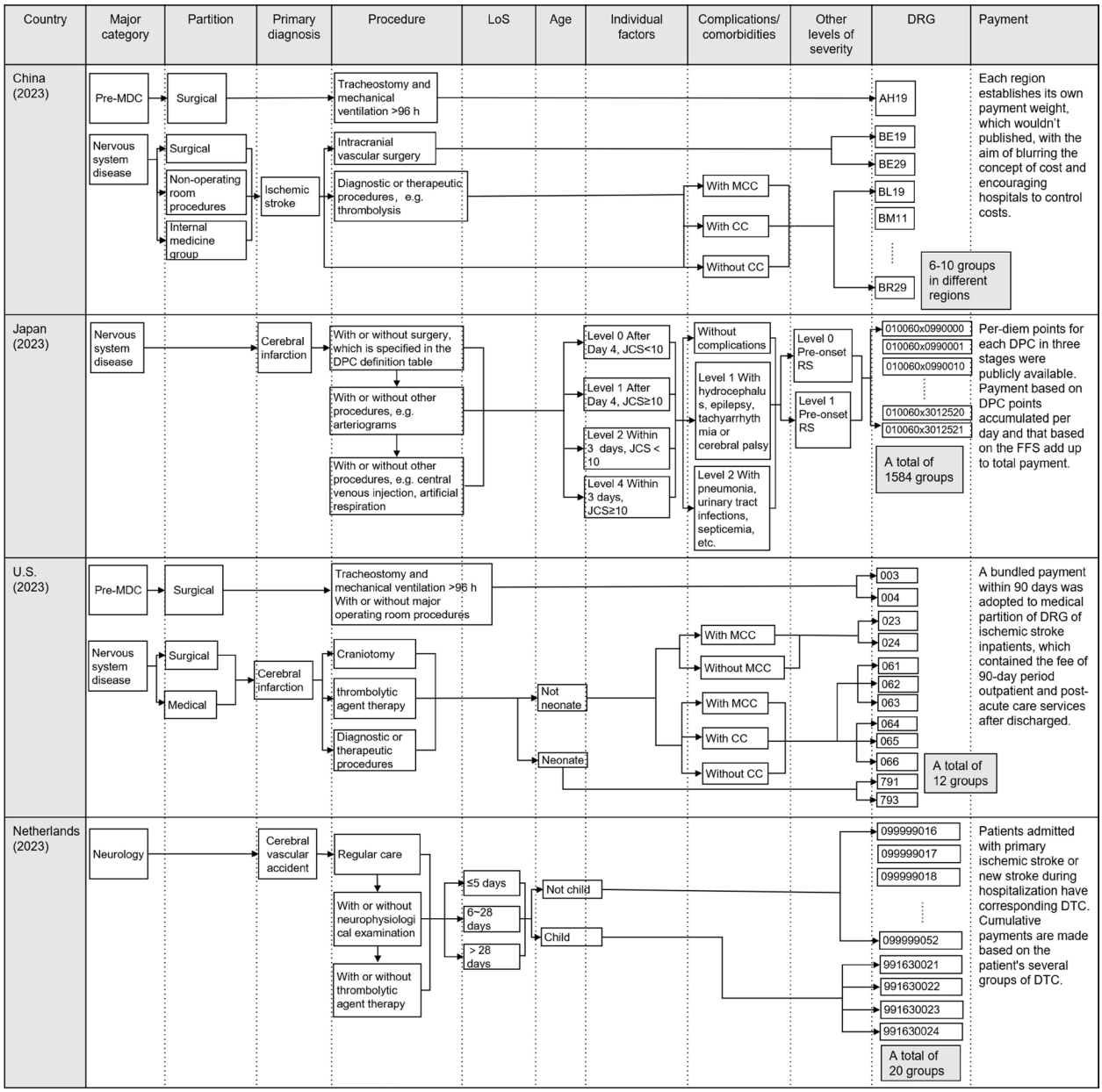

Figure 2. Graphic depiction of grouping variables and payment strategies for inpatients suffering from ischemic stroke under the DRG-based hospital payment system in China, Japan, the U.S. and the Netherlands. *Abbreviations*: LoS, length of stay; DRG, diagnosis-related groups; MDC, main diagnostic category; MCC, major complications or comorbidities; CC, complications or comorbidities; DPC, diagnosis procedure combination; JCS, Japan Coma Scale; RS, Rankin Scale; FFS, fee-for-service; U.S., United States; DTC, diagnosis treatment combinations.

classification of stroke patients in China and the U.S. In addition to diagnosis and procedure, classification variables in the Netherlands include LoS (5 days, 6–28 days, 28+ days) and patient age (child and nonchild). Japan's ischemic stroke grouping is very detailed. Various surgical procedures are grouped separately. Classification variables reflecting stroke severity are used only in Japan's DPC system such as Japan Coma Scale (JCS) scores reflecting patient consciousness and disability/dependence levels and Rankin Scale (RS) scores pre-onset. The DPC systems are finally divided into 1,584 groups after permutations and combinations, but the payment points for some groups are the same. The grouping of ischemic stroke in Japan is more based on clinical logic to facilitate disease statistics and management.

6. Implementation strategies for DRG-based hospital payment in different countries

6.1. Payment for ischemic stroke cases

Directly analyzing and comparing payments for specific diseases is complicated because different countries set DRG-based payment rates at different levels and there are different additional payments. In ischemic stroke cases (Figure 2), patients in the same DRG have the same payment standards and are only assigned to one

DRG group per hospital stay and the payment does not involve outpatient and post-acute care in China. In order to encourage coordination and cooperation between hospitals and post-acute care facilities, the U.S. adopted a bundled payment (*42*). Payment for ischemic stroke care in the medical partition of Medicare-severity DRG (MS-DRG) is bundled within 90 days of discharge, which means outpatient and post-acute care will no longer be paid for in the 90-day period after discharge. Among the 10 countries, only the Netherlands allows several DRGs per hospital stay. Patients admitted with primary ischemic stroke or new ischemic stroke during hospitalization have a corresponding DTC, and cumulative payments are made based on the several groups under the DTC under which the patient falls. In addition to very detailed grouping of ischemic stroke, Japan's DPC payment system also focuses on the quality of ischemic stroke management. Early rehabilitation for ischemic stroke patients was among the 13 quality monitoring indicators for hospitals accepting DPC payment in Japan and additional medical fee incentives are provided to hospitals that meet the quality indicators (*17*).

6.2. The impact of COVID-19 on DRG-based hospital payment systems

Before the COVID-19 pandemic, many healthcare systems around the world were already struggling to contain spending and meet the increasing demand for healthcare needs due to aging populations and a rise in chronic disease. The COVID-19 pandemic further worsened these problems and presented unique challenges to health systems. During the COVID-19 pandemic, hospitals paid by DRG systems based on activity were at financial risk because of the sudden drop in hospital admissions (*43*). Some countries, such as France, have created new DRG codes to classify payment for patients with COVID-19, and more countries have adopted a higher payment tariff or new budgets other than DRG payments to encourage hospitals to prepare for and provide care for COVID-19 patients (*43*). Japan used the existing DPC to pay for COVID-19 inpatients, while the compensation points for COVID-19 patients were updated several times every year. These points were generally 1-6 times the points for the original grouping as a result of changes in the classification of COVID-19 under the Infectious Diseases Act (*44*). COVID-19 has caused Japan to reflect on the flexibility of the DPC payment system, and the country listed "establishing a healthcare system that can flexibly respond to emergency medical needs" as a challenge in its "Basic Policy on Economic and Fiscal Management and Reform 2021" (*45*). Germany paid for COVID-19 patients according to conventional DRG payment standards but directly compensated hospitals for COVID-19-related revenue losses, such as extra financial assistance for each empty bed (*46*). England returned to global budgets in response to the COVID-19 pandemic and announced a deviation from DRG in its "National Health Service Long Term Plan" in 2019 (*47*); as of 2022, the country was moving towards a payment system that consists of three components-a fixed payment, a variable component largely based on DRGs, and a quality-related component.

7. Discussion and prospects

7.1. Differences between DRG payment versions across countries

Over the past four decades, the gradual introduction of the DRG-based hospital payment system from the U.S. to countries around the world and its continuous updating in various countries have caused differences between countries in terms of the purpose, grouping, coding, and payment mechanisms of DRG systems, even if they were based on the same classification concept. The number of groups covered by DRG-based hospital payment systems ranged from about 78 in South Korea to 5,593 in the Netherlands and has tended to increase in almost all 10 countries. In order to ensure homogeneous groups of patients, DRG systems need to consider the most important determinants of resource consumption as classification variables. The classification variables can be complex and vary between countries. Common classification variables include principal diagnosis, procedures performed, patient characteristics, and the severity of the case (*48*). The purposes of adopting a DRG-based hospital payment system differed among countries. For example, England and South Korea introduced DRG-based payment to contain costs and increase efficiency, while Australia implemented DRG-based payment to improve the transparency of resource allocation. Although the original purposes vary in countries, there are four main purposes: containing costs, increasing the efficiency of inpatient care, improving the transparency of hospital care, and improving the quality of care. Countries that need to introduce or learn from other countries' DRG-based hospital payment systems must evaluate which elements of existing DRG versions to introduce, which elements to develop on their own, and how to combine different elements consistently based on their own circumstances.

7.2. The long-term and complex process of optimizing DRG-based payment systems

Most countries that use a DRG-based payment system update their systems regularly. The process of introducing DRG-based payment systems is always carried out in stages, with gradual changes in the types of diseases covered, hospitals covered, areas covered, and payment rates. Countries that have recently introduced a DRG-based hospital payment system generally have limited DRG coverage, such as China's DRG-based payment

system that is only used in some cities and piloting hospitals and South Korea's K-DRG that only covers a limited number of disease categories. Countries that have long used DRG-based payment systems have wide coverage, such as the Thai-DRG the covers all diseases, while some European countries like the Netherlands have even extended DRG systems from inpatient to outpatient care. Some countries such as England have established additional payments that deviate from the DRG-based payment system as the goals of the healthcare system have changed. DRG payment incentivizes hospitals to control costs and improve efficiency through economic leverage, so hospitals will respond strategically to the incentives of the DRG-based payment system to explore profit maximization, which may have unintended consequences. Continuously monitoring hospital activity and dynamically updating payment rates can adjust the incentives to achieve intended goals.

7.3. Mutual promotion of DRG-based payment systems and disease management

Several studies have identified stroke severity as an important determinant of resource utilization in the treatment of stroke patients (*49-51*). When hospitals admit proportionally more patients with more severe illnesses, they are underfunded when receiving only a uniform payment per patient because of the oversimplified grouping of DRGs. Stroke severity is not included as a classification variable in most of the DRG-based hospital payment systems (*52*). There are no ICD-10 codes for ischemic stroke of differing severity at this time, which maybe a major impediment to incorporating stroke severity into DRG systems. In its DPC, Japan has incorporated the JCS score indicating patient consciousness and disability/dependence levels and RS scores pre-onset in the grouping variables for ischemic stroke since 2010. Studies have shown that adding such scores increases the usability of administrative databases and can facilitate risk-adjusted in-hospital mortality assessments, thereby promoting reform of incentive systems or payment systems (*51,53*). Learning from disease classification variables in other countries can promote the optimization of groupings in a country's DRG-based payment system. At the same time, patient prevention and treatment strategies and standardized clinical management can also benefit from classified databases that include more information about patient disease diagnosis and treatment. Groupings under DRG-based payment systems need to balance the two elements of clinical similarity and homogeneity of resource consumption.

7.4. Challenges and prospects

No hospital payment system is likely to perfectly align with the interests of payers, patients, and providers. As the spectrum of diseases changes, the population ages, and the COVID-19 public health emergency persists, the priorities and goals of countries' healthcare systems are constantly changing, and so is the proportion of DRG-based payments out of total hospital payments. Studies have shown that in some high-income countries, policymakers are searching for new ways to shift their inpatient payment systems away from a focus on volume to value-based purchasing methods (*24*). How to improve the flexibility of the DRG payment system and optimally integrate it with other payment methods to form a diversified payment system is a challenge for the sustainable development of DRG-based hospital payment systems. Many countries are exploring mechanisms for reasonable coexistence of a DRG-based payment system and various other payment methods such as global budgets, add-on payments, and episode-based payments. At present, the DRG-based payment system mainly covers inpatients in most countries. There are difficulties in controlling medical costs overall and cost transfer is a risk. A number of countries in Europe have extended the scope of DRG payments beyond 24 hours after discharge (*22*). Based on a disease spectrum featuring chronic diseases with a high incidence and a long and complex course, how to integrate inpatient care with outpatient care, day services, rehabilitation services, or nursing in the DRG-based hospital payment system may be the future direction for development of healthcare payments.

8. Conclusion

Over the past four decades, DRG-based hospital payment systems have tended to spread globally. Diversification and localization are inevitable for the sustainable development of DRG-based hospital payment systems. With the rapid changes in today's global healthcare and healthcare needs, the development of DRG-based hospital payment systems is also facing huge challenges. How to determine the level of DRG payment incentives and improve system flexibility, balance payment goals and disease management goals, and integrate development with other payment methods are areas for future research on DRG-based hospital payment systems, and they will also determine the development of DRG-based hospital payment systems over the next four decades.

Funding: None.

Conflict of Interest: The authors have no conflicts of interest to disclose.

References

1. Global Burden of Disease Health Financing Collaborator Network. Health sector spending and spending on HIV/AIDS, tuberculosis, and malaria, and development

assistance for health: Progress towards Sustainable Development Goal 3. Lancet. 2020; 396:693-724.
2. Evans DB, Elovainio R, Humphreys G. The World Health Report. In: Health Systems Financing: the Path to Universal Coverage (Etienne C, Asamoa A, eds.). World Health Organization, The World Health Report 2010, 2010; pp. 106.
3. Roger France FH. Case mix use in 25 countries: A migration success but international comparisons failure. Int J Med Inform. 2003; 70:215-219.
4. Scheller-Kreinsen D, Quentin W, Busse R. DRG-based hospital payment systems and technological innovation in 12 European countries. Value Health. 2011; 14:1166-1172.
5. Reinhard Busse AG, Wilm Quentin and Miriam Wiley. Diagnosis-related groups in Europe. In: moving towards transparency, efficiency and quality in hospitals (Josep Figueras MM, Elias Mossialos, Richard B. Saltman, Reinhard Busse, ed. World Health Organization, Open University Press, 2011; pp. 24-175.
6. Hopfe M, Stucki G, Bickenbach JE, Prodinger B. Accounting for what matters to patients in the G-DRG System: A stakeholder's perspective on integrating functioning information. Health Serv Insights. 2018; 11.
7. Peltola M. Patient classification and hospital costs of care for stroke in 10 European countries. Health Econ. 2012; 21 Suppl 2:129-140.
8. Campbell BCV, De Silva DA, Macleod MR, Coutts SB, Schwamm LH, Davis SM, Donnan GA. Ischaemic stroke. Nat Rev Dis Primers. 2019; 5:70.
9. Roth GA, Mensah GA, Johnson CO, *et al*. Global burden of cardiovascular diseases and risk factors, 1990-2019: Update from the GBD 2019 study. J Am Coll Cardiol. 2020; 76:2982-3021.
10. eClinicalMedicine. The rising global burden of stroke. EClinicalMedicine. 2023; 59:102028.
11. The World Bank Group. World Bank Group country classifications by income level. *https://blogs.worldbank.org/opendata/new-world-bank-group-country-classifications-income-level-fy24* (accessed February 5, 2024).
12. National Healthcare Security Administration. China Healthcare Security Diagnosis Related Groups (CHS-DRG) subgrouping draft version 1.0. *https://www.gov.cn/zhengce/zhengceku/2020-06/19/content_5520572.htm* (accessed January 10, 2024). (in Chinese)
13. Centers for Medicare & Medicaid Services. MS-DRG Classifications and Software. *https://www.cms.gov/medicare/payment/prospective-payment-systems/acute-inpatient-pps/ms-drg-classifications-and-software* (accessed January 25, 2023).
14. Ministry of Health Labour and Welfare of Japan. About the Diagnosis Procedure Combination (DPC) electronic score sheet. *https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/0000198757_00004.html* (accessed 19 December, 2023). (in Japanese)
15. The Dutch Healthcare Authority. Performance and rate decision for specialist medical care. *https://puc.overheid.nl/nza/doc/PUC_723219_22/* (accessed January 5, 2024). (in Dutch)
16. Mathauer I, Wittenbecher F. Hospital payment systems based on diagnosis-related groups: Experiences in low- and middle-income countries. Bull World Health Organ. 2013; 91:746-756a.
17. Asia Pacific Observatory for Health Systems and Policies & OECD. Case-based payment systems for hospital funding in Asia: An investigation of current status and future directions.Huntington PLAaD, ed. World Health Organization, WHO Press, 2015; pp. 17-68.
18. Kim S, Choi B, Lee K, Lee S, Kim S. Assessing the performance of a method for case-mix adjustment in the Korean Diagnosis-Related Groups (KDRG) system and its policy implications. Health Res Policy Syst. 2021; 19:98.
19. Annear PL, Kwon S, Lorenzoni L, Duckett S, Huntington D, Langenbrunner JC, Murakami Y, Shon C, Xu K. Pathways to DRG-based hospital payment systems in Japan, Korea, and Thailand. Health Policy. 2018; 122:707-713.
20. van Herwaarden S, Wallenburg I, Messelink J, Bal R. Opening the black box of diagnosis-related groups (DRGs): Unpacking the technical remuneration structure of the Dutch DRG system. Health Econ Policy Law. 2020; 15:196-209.
21. Bredenkamp C, Bales S, Kahur K. Transition to diagnosis-related group (DRG) payments for health: Lessons from case studies. The World Bank, International Development in Focus, 2020; pp. 1-54.
22. Quentin W, Scheller-Kreinsen D, Blümel M, Geissler A, Busse R. Hospital payment based on diagnosis-related groups differs in Europe and holds lessons for the United States. Health Aff (Millwood). 2013; 32:713-723.
23. Choi JW, Kim SJ, Park HK, Jang SI, Kim TH, Park EC. Effects of a mandatory DRG payment system in South Korea: Analysis of multi-year nationwide hospital claims data. BMC Health Serv Res. 2019; 19:776.
24. Milstein R, Schreyögg J. The end of an era? Activity-based funding based on diagnosis-related groups: A review of payment reforms in the inpatient sector in 10 high-income countries. Health Policy. 2024; 141:104990.
25. Centers for Medicare & Medicaid Services. Inpatient Prospective Payment System. *https://www.cms.gov/cms-guide-medical-technology-companies-and-other-interested-parties/payment/ipps* (accessed December 7, 2023).
26. The Swedish Board for Health and Welfare. Prospective Hospital weights, NordDRG 2024. *https://www.socialstyrelsen.se/statistik-och-data/klassifikationer-och-koder/drg/viktlistor/* (accessed December 25, 2023). (in Swedish)
27. Swedish Board for Health and Welfare. DRG Basic Concepts and Principles. *https://www.socialstyrelsen.se/statistik-och-data/klassifikationer-och-koder/drg/logiken-i-drg/* (accessed December 1, 2023). (in Swedish)
28. The Australian Independent Hospital Pricing Authority. Australian Refined Diagnosis Related Groups Version 10.0 Final Report. *https://www.ihacpa.gov.au/resources/ar-drg-version-100* (accessed February 1, 2024).
29. England NHS. 2022/23 National Tariff Payment System. *https://www.england.nhs.uk/publication/national-tariff-payment-system-documents-annexes-and-supporting-documents/* (accessed January 31, 2024).
30. Quentin W, Stephani V, Berenson RA, Bilde L, Grasic K, Sikkut R, Touré M, Geissler A. How Denmark, England, Estonia, France, Germany, and the USA pay for variable, specialized and low volume care: A cross-country comparison of in-patient payment systems. Int J Health Policy Manag. 2022; 11:2940-2950.
31. Institute for the Remuneration System in Hospitals. Care revenue catalog 2023 in accordance with the DRG Fee Catalog Ordinance 2023. *https://www.g-drg.de/ag-drg-system-2023/fallpauschalen-katalog/fallpauschalen-*

katalog-20232 (accessed November 24, 2023). (in German)
32. Geissler A, Scheller-Kreinsen D, Quentin W, Busse R. Germany: Understanding G-DRGs. Diagnosis-related groups in Europe: moving towards transparency, efficiency and quality in hospitals. 2011;243-272.
33. Hideo Y. Real World Data in Japan: Chapter II The Diagnosis Procedure Combination Database. Annals of Clinical Epidemiology. 2019.
34. Zeng JQ. The pilot results of 47 148 cases of BJ-DRGs-based payment in China. Int J Health Plann Manage. 2019; 34:1386-1398.
35. Jiao WP. Diagnosis-related groups' payment reform in Beijing. Chin Med J (Engl). 2018; 131:1763-1764.
36. Liu R, Shi J, Yang B, Jin C, Sun P, Wu L, Yu D, Xiong L, Wang Z. Charting a path forward: Policy analysis of China's evolved DRG-based hospital payment system. Int Health. 2017; 9:317-324.
37. Ji X, Fang Y, Liu J. Performance assessment of the inpatient medical services of a clinical subspecialty: A case study with risk adjustment based on diagnosis-related groups in China. Medicine (Baltimore). 2018; 97:e10855.
38. Meng Z, Ma Y, Song S, Li Y, Wang D, Si Y, Sun R, Zhang R, Xue H, Jing L, Wu H. Economic implications of Chinese diagnosis-related group-based payment systems for critically ill patients in ICUs. Crit Care Med. 2020; 48:e565-e573.
39. Zhao C, Wang C, Shen C, Wang Q. Diagnosis-related group (DRG)-based case-mix funding system, a promising alternative for fee for service payment in China. Biosci Trends. 2018; 12:109-115.
40. National Health Security Office (Thailand). Thai DRG and Relative Weight Version 6.3.3, Volume 1. *https://www.tcmc.or.th/download-tcmc* (accessed December 29, 2023). (in Thai)
41. Doroshenko O, Kuzmyn M, Kahur K, Bales S, Kasyanchuk S. Cross-country analysis of case-mix in selected hospitals. The World Bank, Health, Nutrition, and Population Discussion Paper, 2023; pp. 24-31.
42. Centers for Medicare and Medicaid Services. Bundled Payments for Care Improvement Advanced. *https://www.cms.gov/priorities/innovation/media/document/bpcia-my7-rfa* (accessed December 13, 2023).
43. Waitzberg R, Gerkens S, Dimova A, *et al*. Balancing financial incentives during COVID-19: A comparison of provider payment adjustments across 20 countries. Health Policy. 2022; 126:398-407.
44. Ministry of Health Labour and Welfare of Japan. Temporary handling of medical fees related to COVID-19. *https://www.shaho.co.jp/wp-content/uploads/2022/07/110027_040726.pdf* (accessed January 22, 2024). (in Japanese)
45. Cabinet Office. Basic Policy for Economic and Fiscal Management and Reform 2021. *https://www5.cao.go.jp/keizai-shimon/kaigi/cabinet/honebuto/2021/2021_basicpolicies_ja.pdf* (accessed January 18, 2024). (in Japanese)
46. Reif S, Schubert S. Hospital capacity reporting in Germany during COVID-19. ZEW-Centre for European Economic Research Discussion Paper. 2023.
47. England National Health Service. The NHS Long Term Plan. *https://www.longtermplan.nhs.uk/publication/nhs-long-term-plan/* (accessed January 7, 2024).
48. Klein A, Mathauer I, Stenberg K, Habich T. Diagnosis-related groups: A question and answer guide on case-based classification and payment systems. World Health Organization, Geneva, 2020; pp. 13-47.
49. Dewilde S, Annemans L, Pincé H, Thijs V. Hospital financing of ischaemic stroke: Determinants of funding and usefulness of DRG subcategories based on severity of illness. BMC Health Serv Res. 2018; 18:356.
50. Epstein D, Mason A, Manca A. The hospital costs of care for stroke in nine European countries. Health Econ. 2008; 17:S21-31.
51. Cipriano LE, Steinberg ML, Gazelle GS, González RG. Comparing and predicting the costs and outcomes of patients with major and minor stroke using the Boston Acute Stroke Imaging Scale neuroimaging classification system. AJNR Am J Neuroradiol. 2009; 30:703-709.
52. Peltola M, Quentin W. Diagnosis-related groups for stroke in Europe: Patient classification and hospital reimbursement in 11 countries. Cerebrovasc Dis. 2013; 35:113-123.
53. Lee J, Morishima T, Kunisawa S, Sasaki N, Otsubo T, Ikai H, Imanaka Y. Derivation and validation of in-hospital mortality prediction models in ischaemic stroke patients using administrative data. Cerebrovasc Dis. 2013; 35:73-80.

Received January 19, 2024; Revised February 17, 2024; Accepted February 20, 2024.

**Address correspondence to:*
Peipei Song, Center for Clinical Sciences, National Center for Global Health and Medicine, 1-21-1 Toyama Shinjuku-ku, Tokyo 162-8655, Japan.
E-mail: psong@it.ncgm.go.jp

Released online in J-STAGE as advance publication February 24, 2024.

Review

DOI: 10.5582/bst.2023.01325

Roadmap for ending TB in China by 2035: The challenges and strategies

Qishun Feng[1], Guoliang Zhang[1], Liang Chen[2], Huizhong Wu[3], Yingzhou Yang[4], Qian Gao[1, 5], Tetsuya Asakawa[1], Yanlin Zhao[6], Shuihua Lu[1], Lin Zhou[1,7,*], Hongzhou Lu[1,*]

[1] National Clinical Research Center for Infectious Diseases, Guangdong Provincial Clinical Research Center for Tuberculosis, Shenzhen Third People's Hospital, Southern University of Science and Technology, Shenzhen, Guangdong, China;
[2] Guangdong Provincial Research Center for Public Health, Guangdong Provincial Center for Diseases Control and Prevention, Guangzhou, Guangdong, China;
[3] Guangdong Provincial Center for Tuberculosis Control, Guangzhou, Guangdong, China;
[4] Shenzhen Center for Chronic Disease Control, Shenzhen, Guangdong, China;
[5] School of Basic Medical Science, Shanghai Medical College, Fudan University, Shanghai, China;
[6] National Tuberculosis Reference Laboratory, Chinese Center for Disease Control and Prevention, Beijing, China;
[7] Guangdong Provincial People's Hospital, Guangzhou, Guangdong, China.

SUMMARY Tuberculosis (TB) is one of the top ten causes of death worldwide, taking the lives of over a million people annually. In addition to being a serious health issue, TB is also closely linked to eradicating poverty according to the Sustainable Development Goals (SDGs) of the United Nations (UN). All UN members have committed to ending the TB epidemic by 2030. China has one of the highest TB loads worldwide, ranking third in the world on many TB burden indices. The national strategy for TB control is aimed at creating a collaborative network and integrating TB treatment into the medical system. According to the WHO's global TB report, China is expected to have 748,000 new cases of TB in 2022 and an incidence of 52 cases per 100,000 people. Ending TB remains a huge challenge and requires comprehensive control strategies in China. In this work, we have discussed the challenges of TB prevention and control in China and proposed specific measures to end TB.

Keywords tuberculosis, end TB, challenges, strategies

1. Introduction

More than a million people die from tuberculosis (TB) each year, making it one of the top ten global killers (*1,2*). TB has become the most common infection-related cause of death worldwide and the most common antimicrobial-resistant infection (*3,4*). In addition to being a health issue, the Sustainable Development Goal (SDG) of ending poverty is also directly related to TB. As part of the SDGs, UN member nations have pledged to eradicate the TB epidemic by 2030 (*5,6*). The WHO's "End TB Strategy" outlines benchmarks and goals to be accomplished by 2030 (*7,8*). These milestones and targets include a 90% decrease in TB deaths since 2015 and an 80% decrease in the incidence of the disease.

Recently, the 78th session of the UN General Assembly held its Second High-level Meeting on Tuberculosis earlier at UN Headquarters in New York on the theme of "Advancing science, finance, and innovation, and their benefits, to urgently end the global TB epidemic". Efforts are focused on ensuring equitable access to TB prevention, testing, treatment, and care. The meeting culminated in the adoption by world leaders, heads of government, and high-level representatives, among others, of a political declaration setting ambitious new targets for the next five years to galvanize global efforts and ultimately end TB. The newly adopted political declaration proposes to provide 90% of the population with access to TB prevention and care and to use the rapid test recommended by the WHO as the preferred method of diagnosing TB; in addition, the leaders committed to providing social welfare packages to all people with TB, licensing at least one new TB vaccine, and closing the TB response implementation and research funding gap by 2027.

In China, TB has been classified as a category II notifiable disease since 1996. A comprehensive program for TB control and prevention was started in the 1990s and has been supported by national efforts and prioritization (*9*). Because of this, fatality rates for TB in China decreased by over 80% between 1990 and 2010, while incidence decreased by 3.4% per year (*10*).

Despite these successes, China still has one of the largest global TB burdens, ranking third globally in a variety of TB burden indices (*1,11*).

The national TB control strategy sought to develop a cooperative network and incorporate TB care into the healthcare system. The need to lessen the burden of TB in China was stressed in the national plan Healthy China 2030 in 2016 (*12*). The strategy also included a goal to lower personal health spending to under 25% of overall spending on healthcare by 2030 (*13*).

In the WHO's global TB report, the estimated number of new cases of TB in China was 748,000 in 2022 (780,000 in 2021), ranking third among 30 countries with a high TB burden. The estimated number of TB deaths in China is 30,000 and the TB mortality rate is 2.0 per 100,000 people. There are estimated to be 30,000 patients with multidrug/rifampicin-resistant TB (MDR/RR-TB) (*3*). Ending TB remains a challenge and requires comprehensive control strategies in China.

2. Challenges of TB prevention and control in China

As a result of China's efforts, significant achievements have been made in TB prevention and control, but the country is still facing major challenges (*14-19*). First, socio-economic support is insufficient; TB requires adequate investment to be effectively controlled. For example, patients with TB face a heavy financial burden, and this is especially true for patients with drug-resistant TB. Although China's basic health care coverage is very high, there are still many inadequacies in the current relevant policies for the diagnosis and treatment of patients with TB, including the lack of coordination between health insurance and medical policies, and the fact that social security for TB is still unsound. A suggestion has been made that multi-channel financing should be further promoted and that the level of social security for patients should be upgraded by focusing on their actual needs. Under the Precision Poverty Alleviation Program, financial and charitable funds should be used to provide timely assistance to patients who have lost work and who are impoverished due to illness, and individualized subsidy programs should be formulated for patients with multidrug-resistant/extensively drug-resistant strains based on an understanding of the patients' actual difficulties.

Second, the current trend of TB prevalence is not well understood, there are significant regional differences in the prevalence of TB, with the western region being far more likely to have TB than the eastern and central regions (*20*). There is a high rate of TB drug resistance; prevalence in rural areas is higher than in urban areas (*10*). China's population often migrates domestically, and preventing and controlling TB is difficult because of the lack of a clear understanding of the migrant population in cities and regions. Screening should be intensified given the mobility of potential TB patient populations.

Third, there is an inaccurate understanding of key intervention points. TB prevention and control consists of three basic components: eliminating the source of infection, cutting off the means of transmission, and protecting susceptible populations. The current key measures to control the TB epidemic should include: inhibiting people with a latent infection from developing active TB, inhibiting the transmission of active TB from patients to the surrounding population, and inhibiting the infection of healthy people.

Fourth, TB prevention and control require an entire chain of services at the national level. The existing TB prevention and control service system, technology, manpower, and resource inputs do not fully meet the needs of the new circumstances, the service system's mechanism of operation is not yet perfect in some areas, and there are insufficient diagnostic and treatment facilities or equipment in designated medical facilities, as well as weak prevention and control forces at the grassroots level. The government is promoting the establishment of a new type of service system with a clear division of labor and coordination among disease prevention and control agencies, designated medical facilities, and primary medical and healthcare facilities, and it is improving the level of medical protection.

Fifth, China's TB control is not effective enough. China has a high number of patients with multidrug-resistant TB (MDR-TB) and a low cure rate of only 41%. In addition, there are difficulties in the clinical diagnosis of TB, with a 30% positive rate for patients with TB and a rate of TB diagnosis of less than 5% in children. There is also a lack of new vaccines and drugs. BCG vaccination, the only method of TB prevention, is ineffective at preventing the disease, and TB chemoprophylaxis is poorly adhered to and difficult to administer.

3. Official commitment and policy

In 2019, the National Health Commission, the National Development and Reform Commission, the Ministry of Education, the Ministry of Science and Technology, the Ministry of Civil Affairs, the Ministry of Finance, the Poverty Alleviation Office of the State Council, the National Health Insurance Bureau, and eight other departments jointly issued the "Plan of Action to Stop Tuberculosis (2019-2022) (*21*). In addition, the China Center for Disease Control and Prevention issued "The Action Plan for TB-Free Communities (2022-2027)" in September 2022, involving government advocacy and educational campaigns to mobilize society as a whole and other measures, to achieve a TB incidence of 50% by 2025 and a 90% reduction by 2027.

Ending TB requires enhancing organizational leadership, promoting the implementation of the main responsibility of local governments, and incorporating

TB prevention and treatment into local economic and social development planning and governmental target management assessment as an important component of supporting livelihoods. Measures need to be formulated and work programs need to be implemented in line with local realities, action goals and tasks need to be cascaded to specific departments, relevant institutional settings need to be created and staffing needs to be provided, and the implementation of various action measures needs to be supervised.

The State Council has also issued a series of "5-year national tuberculosis control programs" (11th, 12th, and 13th) to enhance TB control and prevention nationally and to increase funding for scientific research in terms of diagnostics, digital health, new drugs, new vaccines, and implemental and operational research (*22,23*). As a result, the Chinese Government has been paying close attention to TB control and prevention. Chinese researchers are still looking into the prevalence of TB, the traits of the most common strains (*24,25*), the effectiveness of infection control measures, ways to help vulnerable populations, novel drug therapies, and technology to support community case management using mobile phones and medical monitors. The scientific community is still aware that TB is preventing China from moving forward and achieving its Healthy China 2030 goals.

4. Comprehensive strategies and key measures for TB prevention and control

The National Tuberculosis Program (NTP), which serves as the nation's technical leading work group for TB, is dedicated to advancing research on strategies and measures for TB control and prevention, organizing and carrying out action plans, conducting scientific research on applying these techniques, and providing technical guidance, staff training, and quality control for disease control and prevention and public health services across the nation.

TB control requires comprehensive strategies (Figure 1). First, this involves prevention of TB, including BCG vaccination of newborns, infants, and young children, as well as introducing preventive treatment for people with a latent infection and implementing infection control.

Second is controlling the source of infection. Patients with TB are detected through a variety of means, including consultations regarding symptoms, proactive detection in key areas and populations, and health check-ups. Drug-resistant screening should be conducted for all patients pathologically positive for TB. In order to consolidate existing practical techniques, the use of rapid and accurate molecular diagnostic techniques is being promoted. Standardized treatment measures should be promoted based on standardized treatment protocols. Comprehensive health management and care services should be provided to patients.

Third is epidemic monitoring. All medical and healthcare facilities should report patients diagnosed with TB or patients suspected of having TB to the China Disease Prevention and Control Information System. Designated medical facilities for TB and hospitals specializing in TB should register and manage patients

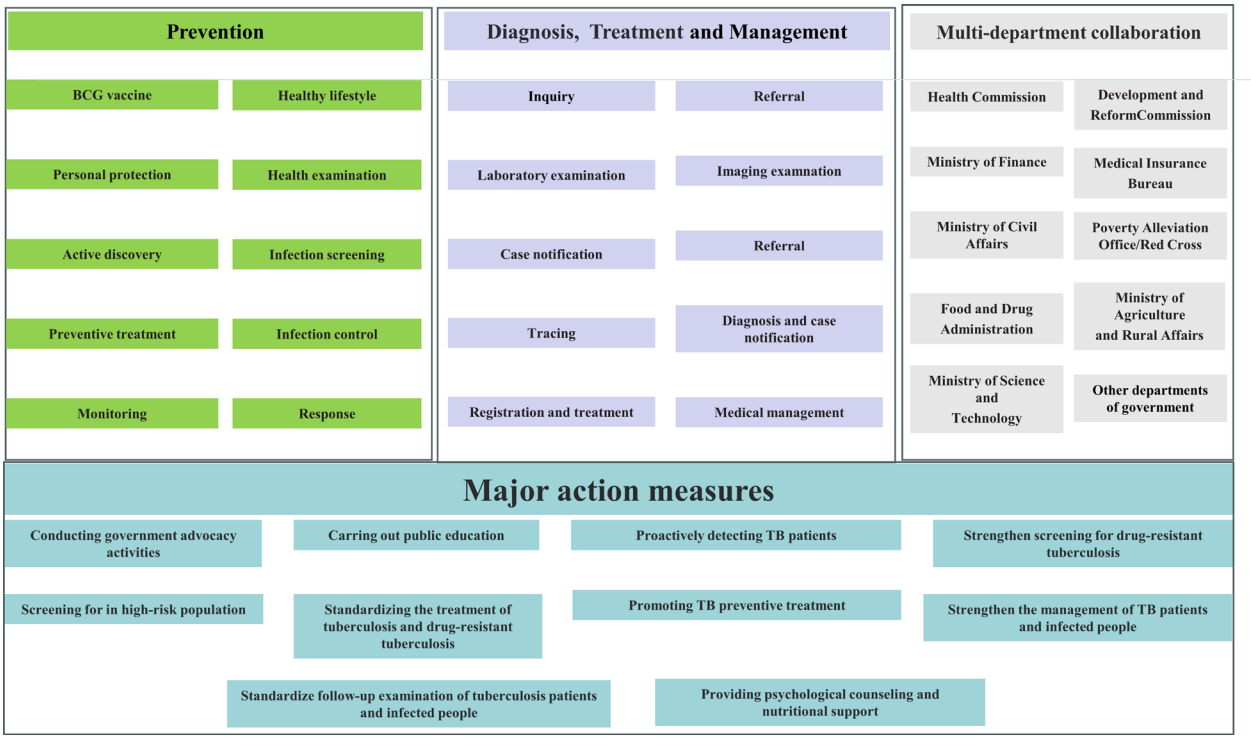

Figure 1. Ending TB strategies and key measures.

diagnosed with active TB, recording information on their diagnosis, treatment, management, and transfer. Disease prevention and control facilities should be responsible for organizing and conducting TB epidemic monitoring and dealing with epidemics.

Fourth is health promotion. Relevant social departments, enterprises and institutions, social groups, and influential people should be mobilized to participate in the prevention and treatment of TB and create good conditions for TB prevention and control. This should be led by the government, with the cooperation of multiple organizations and the participation of society as a whole. Various ways and means of communication should be used to conduct various forms of TB health education for the public and key target groups and places.

The following measures could be taken (Figure 1), including government advocacy, conducting public information and education, proactively identifying patients with TB, enhancing screening for drug-resistant TB, carrying out testing for TB infection in high-risk groups, standardizing the treatment of TB and drug-resistant TB, promoting preventive treatment of TB, enhancing the management of patients with TB and infected people, standardizing the follow-up examinations of patients with TB and infected people, and providing psychological counseling and nutritional support. In the following sections, we will discuss these measures in detail.

5. Scientific innovation is the key to ending TB

Scientific research should be enhanced to address the scientific and technological weaknesses in TB prevention and treatment (*26-28*), to promote the close integration of basic and applied research, and to accelerate the transformation of scientific and technological achievements.

5.1. TB diagnosis

Patients infected with MTB exhibit host immune responses to mycobacterial antigens in the absence of any TB symptoms. There has been much debate on the true latency and level of metabolic activity in TB-associated states. Rather than being two different disease states, latent TB infections (LTBI) and active TB are on a continuum.

Co-infection with HIV/TB continues to be a serious global public health concern. An estimated 10.4 million new cases of TB were reported in 2015, 1.2 million of which involved people who also had HIV. Patients with HIV most frequently get MTB, an opportunistic infection that puts co-infected individuals at a high risk of dying. Co-infected individuals must be diagnosed as soon as feasible due to the high mortality rate of TB-HIV co-infection; early detection of both HIV and TB is essential to saving lives (*29*).

Resistance to the antibiotics INH and RIF in patients with TB is known as MDR-TB. Resistance to at least four first- and second-line antibiotics for TB is known as extensive drug-resistant TB (XDR-TB). Data indicate that 3.3% of MDR-TB cases are among newly treated TB cases, 20% are among previously treated TB cases, and 5% of TB cases are estimated to have MDR-TB. Moreover, 9.5% of patients with MDR-TB are thought to have XDR-TB. Phenotypic drug susceptibility tests continue to be the gold standard for diagnosing MDR- and XDR-TB, despite their acknowledged limitations.

One notable feature of pediatric populations is the difficulty in diagnosing TB. Children have limited bacillary loads, so conventional bacteriological methods frequently fall short and result in misdiagnoses or confusion with other pediatric conditions. The incidence and mortality rate from TB among young people has increased due in part to this diagnostic deficit. Because of this, the World Health Organization advocated for the integration of state-of-the-art molecular diagnostic techniques beginning in 2021 (Table 1). These are intended to either replace or supplement current techniques, increasing the sensitivity and specificity of TB diagnosis in both adults and children (*30*).

5.2. TB vaccine

Currently, Bacillus Calmette-Guérin (BCG) is the only vaccine available for preventing TB. However, BCG displays only moderate effectiveness, and especially in adults. For almost a century, efforts have been made to create potent TB vaccines.

The development of new TB vaccines that are safer and more effective has been accelerated by scientific advancements in our understanding of the immune system, proteomics, and the genetics of MTB. Three elements should be included in an optimal TB vaccination strategy: preventing latent infection from reactivating, preventing primary infection and illness after exposure, and providing immunotherapeutic adjuvant therapy in addition to standard TB treatment to promote patient recovery. The inactivated, live attenuated, recombinant BCG, subunit, viral vector, and DNA vaccines are among the new TB vaccines that are presently undergoing clinical trials (*31*) (Table 2).

Research on novel TB vaccines has advanced, but several obstacles still need to be overcome. These include the low sustainability of clinical trials for the TB vaccine, challenges in selecting epitopes for antigens, the exclusion of pregnant women from current trials, disagreements over how to evaluate the endpoints of these trials, and a lack of appropriate animal models for testing TB vaccines, particularly epitope-based vaccines. That said, the use of new technologies has given TB vaccine research new directions. Examples of these include the use of deep learning and mRNA vaccines.

Table 1. Novel techniques for tuberculosis diagnosis. Data obtained from the World Health Organization's consolidated guidelines (*30*)

Diagnostic method	Detects	Use	Patient type	Technique
Xpert MTB/RIF	MTB and RIF resistance	Pulmonary TB, extra-pulmonary TB, HIV co-infection	Adults and children	PCR
Xpert MTB/RIF Ultra	MTB, minimize false RIF resistance results	TB meningitis, pulmonary TB, extra-pulmonary TB	Adults and children	PCR
Truenat MTB, MTB Plus, and MTB RIF Dx tests	Semi-quantitative detection of MTB complex, RIF resistance	Pulmonary TB, HIV co-infection	Adults	PCR
TB-mediated isothermal DNA amplification (LAMP)	MTB	Pulmonary TB	Adults	PCR
Loop-LAMP	RIF, INH, and ETO associated mutations	Pulmonary TB, extra-pulmonary TB	Adults	PCR
Lipoarabinomannan (LAM) determination by lateral flow immunochromatography	Mycobacterial LAM antigen in urine	Pulmonary TB, extra-pulmonary TB, HIV co-infection	Adults and children	Immunochromatography

Table 2. List of the candidate vaccines in clinical stages

Phase I	Phase IIa	Phase IIb	Phase III
AdHu5Ag85A	D93+GLA-SE (QTP101)	RUTI	MIP
GX-70	ACE/BC02	DAR-901	SRL172
TB/FLU-01L	ChAdOx1.85A	H56:IC31	MTBVAC
TB/FLU-04L		H4:IC31 (AERAS-404)	VPM1002
		MVA85A	M72/AS01E
		BCG (Revaccination)	GamTBvac

The creation of new TB vaccines is a public health initiative that advances human welfare, notwithstanding the many obstacles the field of TB vaccine development faces, such as financial, legal, and societal limitations. In this endeavor, governments and international organizations ought to offer strong backing and actively encourage international cooperation and exchange.

5.3. Anti-TB drugs

The discovery of drugs aimed at TB declined after the surge in the development of antibiotics. However, the WHO declared TB to be a global crisis in the 1990s due to the rise in TB cases and the development of drug resistance.

To address the urgent need to create new medications and enhance TB treatments, representatives from a variety of sectors, including academia, pharmaceutical companies, and public-private partnerships, gathered in South Africa in 2000. With seventeen drugs in various stages of clinical development and six in preclinical stages, great progress has been made after two decades of intense efforts. About 10 of these that are in the clinical stage might have novel mechanisms of action. The remaining seven contenders are enhanced iterations of previously established anti-TB (*32*). However, only 10 of the 17 candidates that are currently in the clinical stage have been tested for their capacity to eradicate intracellular MTB (Table 3), with encouraging results indicating that they may be used to treat LTBI.

Research must continue to overcome the obstacles that are in the way. The ramifications of the developments can revolutionize medicine and open the door to the possibility of worldwide TB control and even eradication.

6. Roadmap for ending the TB epidemic in China

The expected TB incidence in China is 33/100,000 in 2025, 13/100,000 in 2030, and 7/100,000 in 2035 based on the aggressive goals of the End TB strategy and the baseline incidence in 2015 (65/100,000 people) (*33*). The End TB targets are ambitious. To achieve the End TB targets at the national level, key points include the following actions (Figure 2).

Society as a whole needs to be mobilized and the public awareness of TB needs to be increased. Government advocacy is needed to actively promote the implementation of the government's main responsibility by local government departments, to clarify departmental responsibilities, enhance multisectoral cooperation, and include TB prevention and treatment as an important component of people's livelihoods in economic and social development planning and the government's target

Table 3. Drugs in phase II and phase III trials that have intracellular activity against MTB

Drug	Clinical stage	Activity	*Ref.*
OPC-167832	Phase II	Compared to RIF, its intracellular activity against the H37Rv strain was 0.0048 μg/mL, whereas it was 0.0027 μg/mL for the Kurono strain.	(*35*)
Sutezolid	Phase II	With an intracellular minimum inhibitory concentration (MIC) of 0.05 μg/mL, it has a safer profile than linezolid and also shows a strong cumulative impact.	(*36*)
BTZ-043	Phase II	At doses between 1 and 30 ng/mL, it has demonstrated effectiveness against MTB MDR and XDR strains. It specifically accumulates in foamy macrophages and penetrates necrotic nuclei with ease.	(*37*)
Delpazolid	Phase II	It shows up as intracellular activity in macrophages produced from bone marrow. Its anti-MTB activity is comparable to that of linezolid at 1 ng, but it is more effective at 1 μg.	(*38*)
Telacebec	Phase II	The antimycobacterial action of Telacebec is not affected by the stage of mycobacterial replication. With a MIC50 of 0.28 nM, it demonstrated notable intramacrophagic antibacterial action against MTB H37Rv.	(*39*)
SQ109	Phase II	MTB colony-forming units were drastically reduced to fewer than 10,000 in investigations using peritoneal macrophages infected with MTB H37Rv when SQ109 was administered at a dosage of 0.39 μg/mL for 24 hours.	(*40*)
Delamanid	Phase III	It demonstrated efficacy against intracellular mycobacteria at a dose of 0.1 μg/mL after 4 hours in experiments using THP-1 cells, making it 30 times more effective than RIF for treating LTBI.	(*41*)
Pretomanid	Phase III	After exposing MTB-infected THP-1 macrophages to pretomanid for 4 hours, LTBI was successfully eliminated at levels comparable to those of INH.	(*42*)
Bedaquiline	Phase III	It exhibits a remarkable capacity to eradicate IR, XDR, MDR, and MTB TB. It also has activity against LTBI because of its quick sterilizing action.	(*43*)

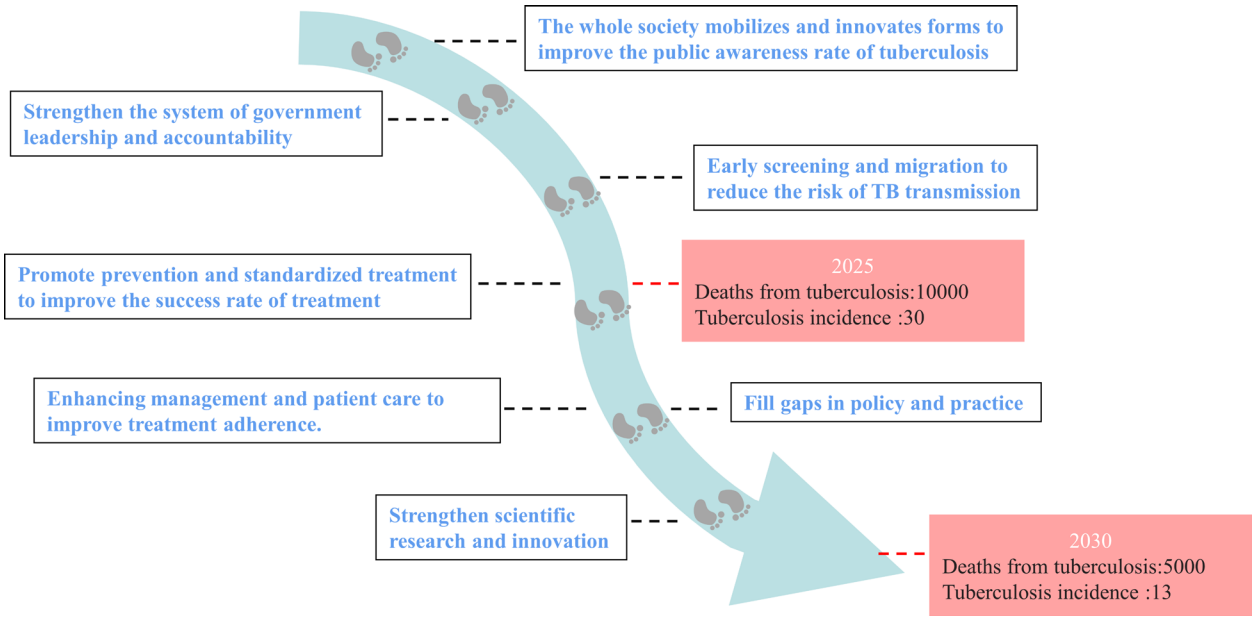

Figure 2. Roadmap for ending TB in China.

management assessment. Multifaceted financing has been provided to provide the necessary working funds, develop and implement TB medical insurance policies, reduce the economic burden on patients with TB, and ensure the smooth operation of TB-free communities. To implement educational campaigns, TB prevention and treatment professional organizations, medical facilities, educational institutions, social groups, scientific and technological associations, and volunteers need to capitalize on "3.24 World Tuberculosis Prevention and Control Day," "World Health Day," and other public awareness efforts and conduct a variety of publicity campaigns with the help of traditional media and new media. With the help of traditional media and new media, various forms of public education on TB prevention and treatment, which are popular among the people, need

to be conducted to raise the public's awareness of and concern for TB, create an awareness of being the first person responsible for public health, and create favorable conditions for society as a whole to participate in the prevention and control of TB.

The system of government leadership and accountability needs to be enhanced. To accomplish the intended objectives within set periods and designated resources, eliminating TB in children and adolescents will require high-level political will, strong leadership, and accountability. Assuring linkages and accountability for essential services (such as maternal and newborn health, HIV, and nutrition), and positioning institutions and actors within or outside the health sector to effectively use resources and sustain efforts to end TB are all important practical steps that leadership should be cognizant of. One key action utilizing the Child and Adolescent Tuberculosis Assessment and Benchmarking Tool, including operational measures, to increase institutional capacity at the national and sub-national levels to organize, direct, and implement TB control programs for children and adolescents. Another key action is enhancing focal points and broadening the Child and Adolescent Tuberculosis Working Group in national TB control programs to include important stakeholders (such as pediatric associations, neonatal and child health departments, and immunization programs). To offer a framework for coordinating the efforts of many partners, strategic planning needs to occur at the national level. Social support needs to be provided to vulnerable families, TB-related services for children and adolescents (including diagnostic services) need to be free of charge, and strategic planning needs to have clear targets, timelines, and dedicated budgets that cover the specific needs of children and adolescents for TB prevention, care, and treatment. In addition, holding health professionals, national plans, policymakers, and leaders accountable for commitments will increase participation by civil society.

Early screening needs to be conducted and target populations need to be monitored to reduce the risk of TB transmission. This includes active detection of patients with TB, active screening, early detection of patients with TB and potential sources of infection, use of innovative tools and technologies such as "Internet^{+} TB prevention and treatment," big data, artificial intelligence, and digital health according to the local situation, and active detection of TB in key populations (close contacts of patients with active TB, the elderly, diabetics, *etc*.), in key places (schools, densely populated places, *etc*.), at key times (starting a job, enrollment in schools, *etc*.). Other efforts include enhancing screening for drug-resistant TB, screening all patients pathologically positive for TB for drug resistance as early as possible, using new laboratory diagnostic techniques such as molecular biology, improving the ability to diagnose drug-resistant TB, shortening the time for drug-resistant diagnosis, maximizing the identification of patients with drug-resistant TB who can be given standardized treatment and management, and taking necessary infection control measures. Monitoring of target populations includes testing for TB infection in high-risk groups, testing for TB infection in people who have close contact with patients with active TB, people infected with HIV, patients with AIDS, people receiving tumor necrosis factor therapy, long-term dialysis therapy, people preparing for organ or bone marrow transplants, patients with silicosis, and people who have been using glucocorticosteroids or other immune-suppressing drugs for a prolonged period.

Prevention and standardized treatment need to be promoted to improve the success rate of treatment. This involves standardizing the treatment of TB and drug-resistant TB, in accordance with the requirements of the Technical Guidelines for the Prevention and Treatment of Tuberculosis in China, giving standardized treatment to patients with confirmed TB and drug-resistant TB; standardizing the isolation and treatment of patients with TB who are in the infectious stage; and actively promoting the use of new medicines and regimens to shorten the course of treatment and improve its efficacy. Other efforts are promoting preventive treatment for TB, increasing public education and mobilization efforts, raising awareness of preventive treatment among newly identified people with latent TB, actively providing preventive treatment for newly identified people with latent TB who do not have active TB and determining the contraindications for preventive treatment, and increasing acceptance of and adherence to preventive treatment for TB.

Management and patient care need to be enhanced to improve treatment adherence. To enhance the management of patients with TB and infected people, primary healthcare facilities, through models such as family doctor contracting services, need to tailor their services to local conditions and individual needs and use a combination of conventional and digital health technologies to provide patients with TB and those receiving preventive treatment with comprehensive care. Follow-up examinations for patients with TB and infected people need to be standardized. In accordance with the requirements of China's Technical Guidelines for Tuberculosis Prevention and Treatment, regular follow-up reviews are conducted for patients with TB to promptly identify adverse reactions, deal with them appropriately, and follow up on the effectiveness of treatment. For people with a latent infection receiving preventive treatment, medical examinations such as liver and kidney function tests are conducted monthly during the treatment period, and chest imaging is performed once a year after the completion of treatment. For other people with a latent infection, medical follow-ups are conducted at the end of the

third, sixth, twelfth and twenty-fourth months after screening for infection. Psychological counseling and nutritional support need to be provided. These efforts assess the nutritional needs of patients with TB, provide nutritional counseling and dietary guidance, and offer individualized nutritional support services to enhance immunity. In addition, regular activities need to be conducted through psychological support groups such as psychosocial counselors, volunteers, nurses, community doctors, social workers, and recovered patients to increase patients' confidence in treatment, channel patients' stress in their lives, and promote recovery.

Gaps in policy and practice need to be filled by policymakers, managers of pertinent local and national programs, and collaborators in implementation. To end TB and bridge the gap between policy and practice, both new and existing tools need to be capitalized upon. To that end national programs must implement the following measures. Health professionals should be trained in the diagnosis and treatment of children and adolescents with TB infections and/or suffering from TB, with an emphasis on TB prevention. Capacity-building tools regarding TB in children and adolescents should be widely disseminating and their usage increased. Supportive supervision and mentoring should be developed at all levels. Successful pilot projects should be replicated as part of routine TB and child health programs. Implementation of the TB prevention guidelines should be enhanced. The ability of child health workers to collect specimens and use available diagnostic tools (*e.g.*, digital chest radiographs and GeneXpert MTB/RIF) should be enhanced. Locally created promotional materials should be distributed to raise awareness of health workers and the public.

Research and innovation need to be enhanced. Research in TB-related science and technology needs to be promoted. Policymakers, academics, donors, implementing partners, and the corporate sector must ensure ongoing investment in research and create legal and legislative conditions that are amenable to research as well as the quick translation and implementation of research findings. To improve TB prevention, diagnosis, and treatment particularly in children and adolescents (*34*), close attention needs to be paid to the following priority areas; the creation of an effective, long-lasting TB vaccine; and the creation of precise, non-sputum-based point-of-care TB illness and infection diagnostic tests. TB prevention and treatment programs for children should be made shorter, safer, and more appropriate. Drug-resistant and susceptible TB should also be treated. Personnel should be aware of the factors that influence the development of TB as well as the main obstacles that teenagers confront while trying to access TB diagnostic and treatment services. Methods of delivering services for TB prevention, exposure, diagnosis, and treatment need to be studied.

With increased support from the government and society, the pace of eliminating TB in China will accelerate, and the expectation is that TB will be truly eliminated.

Funding: This work was supported by the National Key Research and Development Plan (No. 2021YFA1300902), the Guangdong Scientific and Technological Foundation (No. 2020B1111170014), the Guangdong Science Fund for Distinguished Young Scholars (No. 0620220214), the Shenzhen Scientific and Technological Foundation (No. JSGG20220606141001003, KCXFZ20211020163545004, RCJC20221008092726022), and the State Key Laboratory of Respiratory Diseases Open Project (No. SKLRD-OP-202324).

Conflict of Interest: The authors have no conflicts of interest to disclose.

References

1. Floyd K, Glaziou P, Zumla A, Raviglione M. The global tuberculosis epidemic and progress in care, prevention, and research: An overview in year 3 of the End TB era. Lancet Respir Med. 2018; 6:299-314.
2. Falzon D, den Boon S, Kanchar A, Zignol M, Migliori GB, Kasaeva T. Global reporting on tuberculosis preventive treatment among contacts. Eur Respir J. 2022; 59.
3. World Health Organization. Global Tuberculosis Report. 2023. *https://www.who.int/publications/i/item/9789240083851* (accessed December 27, 2023).
4. Hoagland DT, Liu J, Lee RB, Lee RE. New agents for the treatment of drug-resistant Mycobacterium tuberculosis. Adv Drug Deliv Rev. 2016; 102:55-72.
5. Carter DJ, Glaziou P, Lonnroth K, Siroka A, Floyd K, Weil D, Raviglione M, Houben R, Boccia D. The impact of social protection and poverty elimination on global tuberculosis incidence: A statistical modelling analysis of Sustainable Development Goal 1. Lancet Glob Health. 2018; 6:e514-e522.
6. Assefa Y, Gilks CF. Ending the epidemic of HIV/AIDS by 2030: Will there be an endgame to HIV, or an endemic HIV requiring an integrated health systems response in many countries? Int J Infect Dis. 2020; 100:273-277.
7. Lonnroth K, Raviglione M. The WHO's new End TB Strategy in the post-2015 era of the Sustainable Development Goals. Trans R Soc Trop Med Hyg. 2016; 110:148-150.
8. Fitchett JR, MacPherson P, Corbett EL. Implementing the End TB Strategy and the intersection with the Sustainable Development Goals, 2016-2030. Trans R Soc Trop Med Hyg. 2016; 110:145-147.
9. Long Q, Qu Y, Lucas H. Drug-resistant tuberculosis control in China: Progress and challenges. Infect Dis Poverty. 2016; 5:9.
10. Wang L, Zhang H, Ruan Y, *et al*. Tuberculosis prevalence in China, 1990-2010; A longitudinal analysis of national survey data. Lancet. 2014; 383:2057-2064.
11. Chakaya J, Khan M, Ntoumi F, *et al*. Global Tuberculosis Report 2020 - Reflections on the global TB burden, treatment and prevention efforts. Int J Infect Dis. 2021;

113 Suppl 1:S7-S12.
12. Chen P, Li F, Harmer P. Healthy China 2030: Moving from blueprint to action with a new focus on public health. Lancet Public Health. 2019; 4:e447.
13. Fu W, Zhao S, Zhang Y, Chai P, Goss J. Research in health policy making in China: Out-of-pocket payments in Healthy China 2030. BMJ. 2018; 360:k234.
14. Long Q, Guo L, Jiang W, Huan S, Tang S. Ending tuberculosis in China: Health system challenges. Lancet Public Health. 2021; 6:e948-e953.
15. Tang S, Squire SB. What lessons can be drawn from tuberculosis (TB) control in China in the 1990s? An analysis from a health system perspective. Health Policy. 2005; 72:93-104.
16. Tang S, Wang L, Wang H, Chin DP. Access to and affordability of healthcare for TB patients in China: Issues and challenges. Infect Dis Poverty. 2016; 5:10.
17. Zhao Y, Liu J. Facing the challenge of tuberculosis: Towards "End TB in China by 2035". China CDC Wkly. 2021; 3:243-246.
18. Zhao F, Cheng J, Cheng SM, Zhang H, Zhao YL, Zhang CY, Hu DM, Fan HY, Huang F, Qu Y, He GX, Wang LX. The current status and challenges regarding tuberculosis infection control in health care facilities in China. Biomed Environ Sci. 2015; 28:848-854.
19. Long Q, Jiang WX, Zhang H, Cheng J, Tang SL, Wang WB. Multi-source financing for tuberculosis treatment in China: Key issues and challenges. Infect Dis Poverty. 2021; 10:17.
20. Zhao Y, Ehiri J, Li D, Luo X, Li Y. A survey of TB knowledge among medical students in Southwest China: Is the information reaching the target? BMJ Open. 2013; 3:e003454.
21. Zhang H, Liu X, Xu C, Hu D, Li X, Li T, Zhao Y, Chen M, Liu J. Guiding tuberculosis control through the Healthy China Initiative 2019-2030. China CDC Wkly. 2020; 2:948-950.
22. Zhou Y, Anthony R, Wang S, Ou X, Liu D, Zhao Y, Soolingen DV. The epidemic of multidrug resistant tuberculosis in China in historical and phylogenetic perspectives. J Infect. 2020; 80:444-453.
23. Wang S, Zhou Y, Zhao B, Ou X, Xia H, Zheng Y, Song Y, Cheng Q, Wang X, Zhao Y. Characteristics of compensatory mutations in the rpoC gene and their association with compensated transmission of Mycobacterium tuberculosis. Front Med. 2020; 14:51-59.
24. Meehan CJ, Goig GA, Kohl TA, *et al*. Whole genome sequencing of Mycobacterium tuberculosis: Current standards and open issues. Nat Rev Microbiol. 2019; 17:533-545.
25. Hicks ND, Carey AF, Yang J, Zhao Y, Fortune SM. Bacterial genome-wide association identifies novel factors that contribute to ethionamide and prothionamide susceptibility in Mycobacterium tuberculosis. mBio. 2019; 10.
26. Liu D, He W, Jiang M, *et al*. Development of a loop-mediated isothermal amplification coupled lateral flow dipstick targeting erm(41) for detection of Mycobacterium abscessus and Mycobacterium massiliense. AMB Express. 2019; 9:11.
27. Liu L, Jiang F, Chen L, Zhao B, Dong J, Sun L, Zhu Y, Liu B, Zhou Y, Yang J, Zhao Y, Jin Q, Zhang X. The impact of combined gene mutations in inhA and ahpC genes on high levels of isoniazid resistance amongst katG non-315 in multidrug-resistant tuberculosis isolates from China. Emerg Microbes Infect. 2018; 7:183.
28. Consortium CR, the GP, Allix-Beguec C, *et al*. Prediction of susceptibility to first-line tuberculosis drugs by DNA sequencing. N Engl J Med. 2018; 379:1403-1415.
29. Lawn SD, Churchyard G. Epidemiology of HIV-associated tuberculosis. Curr Opin HIV AIDS. 2009; 4:325-333.
30. WHO Consolidated Guidelines on Tuberculosis: Module 5: Management of tuberculosis in children and adolescents (Geneva, 2022).
31. Zhuang L, Ye Z, Li L, Yang L, Gong W. Next-generation TB vaccines: Progress, challenges, and prospects. Vaccines (Basel). 2023; 11.
32. Shleider Carnero Canales C, Marquez Cazorla J, Furtado Torres AH, Monteiro Filardi ET, Di Filippo LD, Costa PI, Roque-Borda CA, Pavan FR. Advances in diagnostics and drug discovery against resistant and latent tuberculosis infection. Pharmaceutics. 2023; 15.
33. Ding C, Hu M, Shangguan Y, Guo W, Wang S, Feng X, Zhang Z, Zhang Y, Xu K. Epidemic trends in high tuberculosis burden countries during the last three decades and feasibility of achieving the global targets at the country level. Front Med (Lausanne). 2022; 9:798465.
34. Seddon JA, Johnson S, Palmer M, van der Zalm MM, Lopez-Varela E, Hughes J, Schaaf HS. Multidrug-resistant tuberculosis in children and adolescents: Current strategies for prevention and treatment. Expert Rev Respir Med. 2021; 15:221-237.
35. Sarathy JP, Zimmerman MD, Gengenbacher M, Dartois V, Dick T. Mycobacterium tuberculosis DprE1 inhibitor OPC-167832 is active against Mycobacterium abscessus *in vitro*. Antimicrob Agents Chemother. 2022; 66:e0123722.
36. Zhu T, Friedrich SO, Diacon A, Wallis RS. Population pharmacokinetic/pharmacodynamic analysis of the bactericidal activities of sutezolid (PNU-100480) and its major metabolite against intracellular Mycobacterium tuberculosis in ex vivo whole-blood cultures of patients with pulmonary tuberculosis. Antimicrob Agents Chemother. 2014; 58:3306-3311.
37. Kloss F, Krchnak V, Krchnakova A, Schieferdecker S, Dreisbach J, Krone V, Mollmann U, Hoelscher M, Miller MJ. *In Vivo* dearomatization of the potent antituberculosis agent BTZ043 *via* Meisenheimer complex formation. Angew Chem Int Ed Engl. 2017; 56:2187-2191.
38. Kim JS, Kim YH, Lee SH, *et al*. Early bactericidal activity of delpazolid (LCB01-0371) in patients with pulmonary tuberculosis. Antimicrob Agents Chemother. 2022; 66:e0168421.
39. de Jager VR, Dawson R, van Niekerk C, Hutchings J, Kim J, Vanker N, van der Merwe L, Choi J, Nam K, Diacon AH. Telacebec (Q203), a new antituberculosis agent. N Engl J Med. 2020; 382:1280-1281.
40. Singh M, Kumar S, Singh B, Jain P, Kumari A, Pahuja I, Chaturvedi S, Prasad DVR, Dwivedi VP, Das G. The 1, 2-ethylenediamine SQ109 protects against tuberculosis by promoting M1 macrophage polarization through the p38 MAPK pathway. Commun Biol. 2022; 5:759.
41. Liu Y, Matsumoto M, Ishida H, Ohguro K, Yoshitake M, Gupta R, Geiter L, Hafkin J. Delamanid: From discovery to its use for pulmonary multidrug-resistant tuberculosis (MDR-TB). Tuberculosis (Edinb). 2018; 111:20-30.
42. Stancil SL, Mirzayev F, Abdel-Rahman SM. Profiling

pretomanid as a therapeutic option for TB infection: Evidence to date. Drug Des Devel Ther. 2021; 15:2815-2830.
43. Lyu XL, Lin TT, Gao JT, Jia HY, Zhu CZ, Li ZH, Dong J, Sun Q, Shu W, Wang SS, Pan LP, Huang HR, Zhang ZD, Li Q. Effects of bedaquiline on antimicrobial activity and cytokine secretion of macrophages infected with multidrug-resistant Mycobacterium tuberculosis strains. Can J Infect Dis Med Microbiol. 2022; 2022:2703635.

Received December 28, 2023; Revised January 27, 2024; Accepted February 1, 2024.

*Address correspondence to:
Hongzhou Lu and Lin Zhou, Department of Infectious Diseases, National Clinical Research Center for Infectious Diseases, the Third People's Hospital of Shenzhen, 29 Bulan Road, Shenzhen, Guangdong 518112, China.
E-mail: luhongzhou@fudan.edu.cn (HL), ncrcid@163.com (LZ)

Released online in J-STAGE as advance publication February 8, 2024.

Review

DOI: 10.5582/bst.2023.01436

Neoadjuvant therapies in resectable hepatocellular carcinoma: Exploring strategies to improve prognosis

Ya-nan Ma[1,2], Xuemei Jiang[2], Peipei Song[1], Wei Tang[1,3,*]

[1] National Center for Global Health and Medicine, Tokyo, Japan;
[2] Department of Gastroenterology, Hainan General Hospital, Hainan Affiliated Hospital of Hainan Medical University, Haikou, China;
[3] Haikou Affiliated Hospital of Central South University Xiangya School of Medicine, Haikou, China.

SUMMARY Hepatocellular carcinoma (HCC), a challenging malignancy, often necessitates surgical intervention, notably liver resection. However, the high recurrence rate, reaching 70% within 5 years post-resection, significantly impacts patient outcomes. Neoadjuvant therapies aim to preoperatively address this challenge, reducing lesion size, improving surgical resection rates, deactivating potential micro-metastases, and ultimately lowering postoperative recurrence rates. This review concentrates on advances in research on and clinical use of neoadjuvant therapies for HCC, with particular attention to the use of immune checkpoint inhibitors (ICIs) targeting programmed cell death-1 (PD-1), programmed cell death ligand-1 (PD-L1), and cytotoxic T-lymphocyte-associated protein-4 (CTLA-4). Ongoing clinical studies exploring immunotherapy combined with a tyrosine kinase inhibitor (TKI), interventional therapy, radiotherapy, and other modalities offer promising insights into overcoming resistance to monotherapies. In summary, neoadjuvant therapies hold significant promise in terms of improving the prognosis for patients with HCC and enhancing long-term survival, particularly through innovative combination strategies.

Keywords immunotherapy, targeted therapy, recurrence, clinical trials, endpoint, response

1. Introduction

Liver cancer remains a global health challenge, and its incidence is steadily rising worldwide (*1,2*). Estimates are that by 2025, over a million individuals annually will be affected by liver cancer (*3*). Hepatocellular carcinoma (HCC) accounts for approximately 90% of liver cancer cases and is the most prevalent subtype. The primary methods for treating HCC involve surgical interventions, including liver resection (LR) and liver transplantation (LT). LT faces challenges due to organ scarcity and a prolonged waiting time, leading to patients being dropped from the waiting list due to tumor progression. Studies conducted across multiple centers in China, Italy, Japan, and the United States suggest that the likelihood of achieving a cure through resection is comparable to transplantation when the dropout rate exceeds 20% (*4*). Moreover, factors such as cancer thrombus formation, microvascular infiltration, a tumor diameter exceeding 5 cm, poor tumor differentiation, narrow surgical margins (< 1.0 cm), multifocal tumors, satellite nodules, and lymph node metastasis contribute to early recurrence following curative LR (*5-7*). Global liver cancer guidelines (as shown in Table 1), including those from the European Association for the Study of the Liver (EASL) (*8*), Barcelona Clinic Liver Cancer (BCLC) (*9*), American Association for the Study of Liver Diseases (AASLD) (*10*), National Comprehensive Cancer Network (NCCN) (*11*), China (*12*), Japan Society of Hepatology (JSH) (*13*), Korean Liver Cancer Association-National Cancer Center (KLCA-NCC) (*14*), and Indian National Association for the Study of the Liver (INASL) (*15*), indicate a recurrence rate of 10–40% after LT. Furthermore, 70% of patients with HCC experience recurrence within 5 years post-LR, with early recurrence (< 2 years) constituting 60–70% of recurrent cases. Postoperative recurrence of HCC poses a significant challenge to cure, resulting in low survival rates for patients (*16,17*). Therefore, the identification of effective approaches to reduce postoperative recurrence and enhance the curative resection rate is of paramount importance.

Traditionally, adjuvant therapy refers to postoperative treatment aimed at consolidating the surgical intervention's role in eliminating residual tumor cells. However, concerns persist regarding the effectiveness and safety of postoperative adjuvant therapy for HCC. For instance, postoperative sorafenib therapy following

Table 1. Global overview of indications, postoperative recurrence, and post-recurrence treatment strategies for hepatocellular carcinoma

Area (year) (*Ref.*)	Indications for LR	Recurrence rate after LR	Management of recurrence after LR	Indications for LT	Recurrence rate after LT	Management of recurrence after LT
Europe						
EASL (2018)(*8*)	solitary or < 3 cm × 2–3 nodules, portal hypertension (-), preserved LF, PS = 0	5-year recurrence rate of 70%. 60–70% of recurrences occur within the first 2 years.	adjuvant therapy after LR is not recommended	solitary or < 3 cm × 2–3 nodules, portal hypertension (-), preserved LF, PS = 0;	13%	preoperative NAT, including bridging and downstaging therapy, is recommended for patients eligible for LT.
BCLC (2022)(*9*)	solitary or ≤ 3 cm × 2–3 nodules, preserved LF, PS = 0;	5-year recurrence rate of 70%.	LRT; LT in patients with successful downstaging treatment; systemic therapy in patients with VI, EHM, or inappropriate TACE.	solitary or ≤ 3 cm × 2–3 nodules, preserved LF, PS = 0; multiple nodules, preserved LF, PS = 0	5-year recurrence rate 10–15%.	LRT; LT in patients with successful downstaging therapy; systemic therapy in patients with VI, EHM, or inappropriate TACE.
America						
AASLD (2023)(*10*)	solitary or ≤ 3 cm× 2–3 nodules, preserved LF, PS = 0;	5-year recurrence rate of 50–70%	remedial LT within the Milan criteria; LRT for localized recurrence beyond Milan criteria, and LT in patients with successful downstaging; systemic therapy in patients with VI, EHM, TACE is inappropriate, or advanced recurrence HCC.	solitary or ≤ 3 cm × 2–3 nodules, preserved LF, PS = 0; multiple nodules, preserved LF, PS = 0	10–15%	LRT LT in patients with successful downstaging treatment; systemic therapy in patients with VI, EHM, or inappropriate TACE.
NCCN (2023)(*11*)	Child-Pugh A/B, adequate FLR, solitary, portal hypertension (-); limited and resectable multifocal; major VI+; initially unresectable HCC that responds to therapy	5-year recurrence rate exceeds 70%	postoperative adjuvant therapy	solitary, Child-Pugh A/B, Adequate FLR, portal hypertension (-); tumor characteristics marginally outside of the UNOS criteria downstaged to within criteria	18–40%	NAT is recommended for patients eligible for LT;pre-LT NAT reduces postoperative recurrence rates
Asia						
China (2022)(*12*)	Child-Pugh A/B, PS = 0–2, solitary ≤ 5 cm (Ia); solitary > 5 cm or ≤ 3 cm × 2–3 nodules (Ib); > 3 cm × 2–3 nodules (IIa); ≥ 4 nodules (IIb); VI+ (IIIa);	5-year recurrence rate of 40–70%	Depending on the characteristics of the recurrent tumor, re-surgical resection, ablation therapy, interventional therapy, radiation therapy, or systemic anti-tumor therapy can be chosen; aggressive radical resection can be considered for patients with postoperative pure peritoneal metastases.	Child-Pugh A/B, PS = 0–2, solitary ≤ 5 cm; Child-Pugh A/B, PS = 0–2, solitary > 5 cm or ≤ 3 cm × 2–3 nodules; Child-Pugh A/B, PS = 0–2, > 3 cm × 2–3 nodules; Child-Pugh C, PS = 3–4.	5-year recurrence rate 33.7%.	early postoperative withdrawal or hormone-free regimens and reduced dosage of calcium-modulated phosphatase inhibitors in the early postoperative period after LT reduce tumor recurrence rates
JSH (2021)(*13*)	Child-Pugh A/B, ≤ 3 cm × 1–3; Child-Pugh A/B, > 3 cm × 1–3; Child-Pugh A/B, VI+;	Annual recurrence rate exceeds 10%, and 5-year recurrence rate is 70–80%. More than 90% of the initial recurrence occurs within the liver.	If recurrence occurs within the liver, treatment is considered based on the amount of liver remnant and LF. The treatment strategy is essentially the same as for the first treatment: surgical resection or, if resection is difficult, RFA, TARE, or systemic therapy	Child-Pugh C, within Milan or 5-5-500 criteria, age ≤ 65 years;	1-year recurrence rate 9.9%, 3-year recurrence rate 16.1%.	surgical resection or systemic therapy if resection is difficult.

Abbreviations: AASLD, American Association for the Study of Liver Diseases; BCLC, Barcelona Liver Cancer; CTP, Child–Turcotte–Pugh score; EASL, European Association for the Study of the Liver; EBRT, external beam radiation therapy; EHM, extrahepatic metastases; HCC, hepatocellular carcinoma; INASL, Indian National Association for Study of the Liver; JSH, Japanese Society of Hepatology; KLCA-NCC, Korean Liver Cancer Association (KLCA) and National Cancer Center (NCC) Korea; FLR, future liver remnant; LF, liver function; LR, liver resection; LRT, locoregional therapy; LT, liver transplantation; NAT, neoadjuvant therapy; NCCN, US National Comprehensive Cancer Network; PS, performance status; RFA, radiofrequency ablation; SBRT, stereotactic body radiation therapy; TACE, transarterial chemoembolization; TARE, transarterial radioembolization; TKI, tyrosine kinase inhibitor; UNOS, United Network for Organ Sharing; VI, vascular or bile duct invasion.

Table 1. Global overview of indications, postoperative recurrence, and post-recurrence treatment strategies for hepatocellular carcinoma (continued)

Area (year) (*Ref.*)	Indications for LR	Recurrence rate after LR	Management of LR recurrence	Indications for LT	Recurrence rate after LT	Management of LT recurrence
KLCA-NCC (2022)(*14*)	solitary ≤ 2 cm; solitary > 2 cm; ≤ 2 cm × 2–3 nodules; solitary ≤ 2 cm, VI+; >2 cm, ≤ 3 nodules; solitary > 2 cm, VI+;	5-year recurrence rate of 50–60%	selection of treatment regimen based on time to recurrence, residual LF, functional status, and size, location, and number of recurrent tumors	2 cm < solitary < 5 cm; multiple ≤ 2 cm; multiple > 2 cm;	8–20%	retreatment selected based on the timing of recurrence, LF, exercise status, size, location, and number of recurrent tumors.
INASL (2023)(*15*)	solitary < 5 cm, or ≤ 3 cm × 2–3 nodules (within Milan criteria), preserved LF (CTP ≤ 6), PS = 0; VI and/or EHM, moderately preserved LF (CTP ≤ 8), PS ≤ 2;	5-year recurrence rate of 70%, and intrahepatic recurrence rate of 68–98%.	TACE, TARE, or SBRT as adjuvant therapy in INASL-BCLC stage B patients TARE, SBRT or resection as adjuvant therapy in INASL-BCLC stage C1 patients	solitary or multiple nodules ≥ 5 cm (beyond Milan criteria), moderately preserved LF (CTP ≤ 8), PS ≤ 1; end stage LF (CTP ≥ 9);	15%	multimodal therapy; oligo-recurrence is resectable in one or two organs, surgical resection if R0 resection is feasible; ablation alone or combined with LRT.

Abbreviations: AASLD, American Association for the Study of Liver Diseases; BCLC, Barcelona Liver Cancer; CTP, Child–Turcotte–Pugh score; EASL, European Association for the Study of the Liver; EBRT, external beam radiation therapy; EHM, extrahepatic metastases; HCC, hepatocellular carcinoma; INASL, Indian National Association for Study of the Liver; JSH, Japanese Society of Hepatology; KLCA-NCC, Korean Liver Cancer Association (KLCA) and National Cancer Center (NCC) Korea; FLR, future liver remnant; LF, liver function; LR, liver resection; LRT, locoregional therapy; LT, liver transplantation; NAT, neoadjuvant therapy; NCCN, US National Comprehensive Cancer Network; PS, performance status; RFA, radiofrequency ablation; SBRT, stereotactic body radiation therapy; TACE, transarterial chemoembolization; TARE, transarterial radioembolization; TKI, tyrosine kinase inhibitor; UNOS, United Network for Organ Sharing; VI, vascular or bile duct invasion.

resection or ablation did not improve overall survival (OS) or disease-free survival (DFS) (*17*). In contrast to adjuvant radiotherapy, neoadjuvant radiotherapy has shown potential in improving long-term survival for patients (*18*). Considering current perspectives, early metastases often exist at the time of diagnosis, even when conventional imaging or standard diagnostic methods may not detect them. Therefore, neoadjuvant therapy (NAT), as a preoperative treatment approach, has garnered increasing attention (as shown in Figure 1). In the context of HCC, NAT presents an opportunity to reduce tumor staging and prevent early recurrence.

As many patients exhibit impaired liver function at the baseline, preserving the future liver remnant by shrinking tumors may expand the population eligible for surgery or ablation. NAT can inactivate potential micrometastases, enhance surgical resection outcomes, and reduce postoperative recurrence rates. It can also reduce lesion size, offering a chance for R0 surgical resection in potentially operable cases, thus increasing the surgical resection rate. According to the EASL (*19*), INASL (*15*), and KLCA-NCC (*14*) guidelines, the concept of NAT extends to LT, potentially lowering patients to the Milan criteria or expanding the transplant criteria.

This review specifically discusses NAT, which is intended to eliminate residual hidden cancer cells after resection and provide a means to explore the biological characteristics of tumors, for resectable HCC. For instance, the pathological response to NAT can offer prognostic information and guide the selection of adjuvant treatment regimens. Research on NAT deepens our understanding of the mechanisms of HCC pathogenesis and progression, it fosters the discovery of more effective strategies for treating HCC, and it positively influences the standardized implementation of NAT.

2. Surgical resection alone *vs.* NAT followed by surgical resection in resectable HCC

The fundamental principles for patients with HCC undergoing LR are as follows: (1) Completeness: thorough removal of the tumor with no residual tumor at the margins; (2) Safety: preservation of an adequate volume of functional liver tissue to ensure compensatory liver function postoperatively, reduce surgical complications, and lower mortality rates (*12*). However, determining resectability is a complex issue. In 2023, Japanese experts conducted relevant studies on the concept of resectability in HCC (*20*). Referring to the concept of classifying pancreatic cancer, resectability in HCC is categorized into resectable, potentially resectable, and unresectable. Unresectable HCC (uHCC) is defined as a disease with distant metastasis or the inability to achieve macroscopically radical resection (*21*). The residual liver indocyanine green clearance rate (ICG-Krem) and major vessel infiltration were

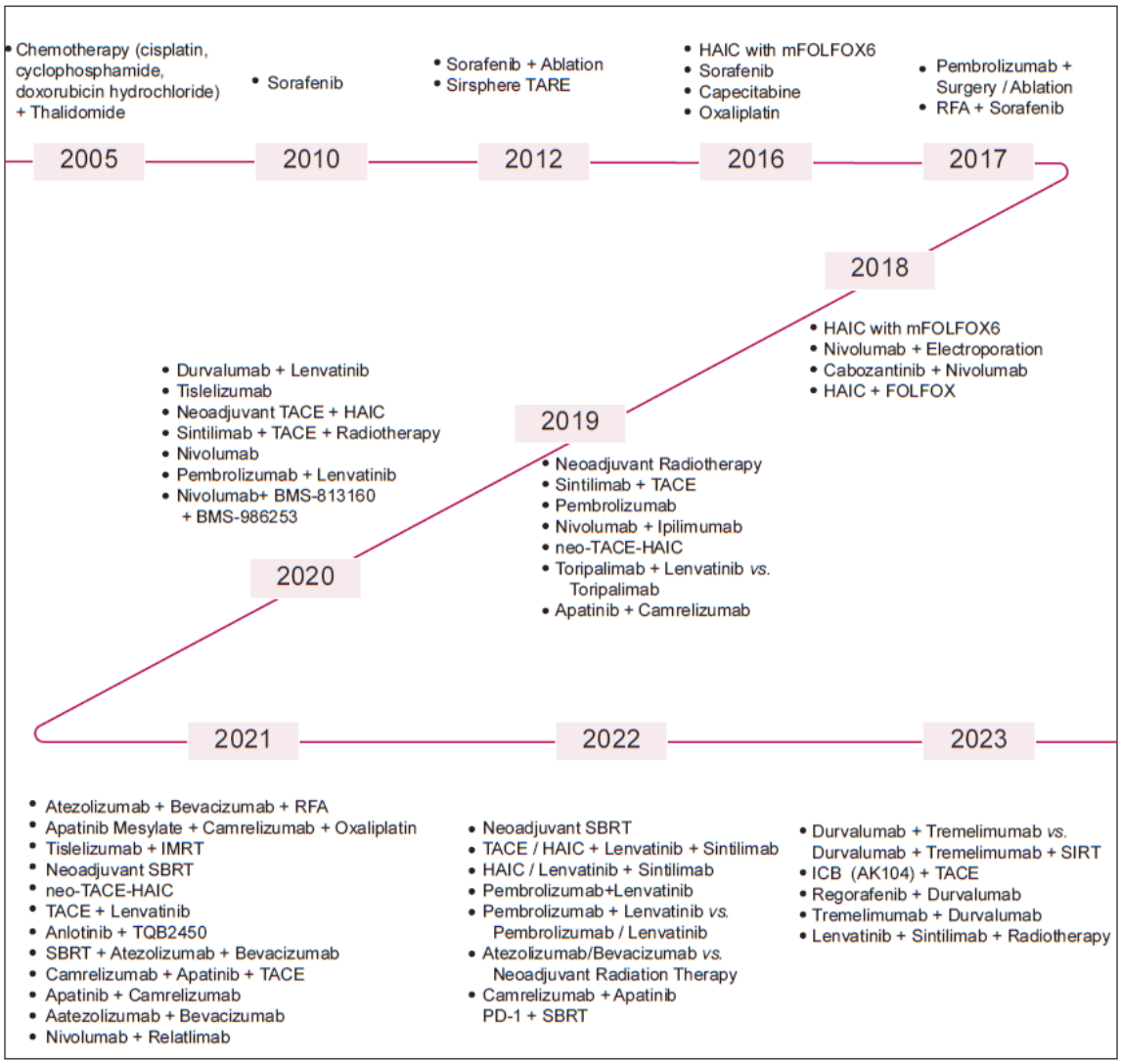

Figure 1. Evolution of clinical research on neoadjuvant therapy in hepatocellular carcinoma. The figure delineates the historical progression of neoadjuvant therapy protocols utilized in clinical studies on hepatocellular carcinoma, showcasing the emergence of novel treatment modalities over time.

selected as determinants for potentially resectable HCC, defining potentially resectable HCC as a high-risk group with clinically relevant liver failure after LR assessed with ICG-Krem and/or HCC with major vessel infiltration (*21,22*). Major vessel infiltration is defined as involvement of Vp2-Vp4 and/or Vv2-Vv3 (*23*). ICG-Krem = preoperative ICG clearance rate × future liver remnant volume (FLRV) / total liver volume (TLV). According to studies and relevant guidelines (*21,22,24*), ICG-Krem < 0.03 is defined as uHCC, ICG-Krem < 0.05–≥ 0.03 is defined as potentially resectable HCC, and the rest are classified as resectable HCC. Chinese experts have implemented a more detailed classification of uHCC, identifying two primary types (*25*). The first type is characterized by surgical unresectability, encompassing patients who are unable to endure surgical trauma due to factors such as their general condition, liver function, and insufficient FLRV. The second type of uHCC is technically resectable but effectiveness cannot be achieved compared to non-surgical treatment after resection, rendering it oncologically/biologically unresectable.

In patients with resectable HCC, those who underwent NAT before surgery have significantly improved survival rates and outcomes compared to those underwent surgery alone. Findings from a multicenter randomized controlled clinical trial involving 208 patients with resectable HCC in stage III revealed that patients in the neoadjuvant hepatic arterial infusion chemotherapy (HAIC) group had markedly higher 1-year, 2-year, and 3-year OS rates (92.9%, 78.6%, and 63.5%, respectively) in contrast to the surgery-alone group (79.5%, 62.0%, and 46.3%, respectively) (p = 0.016) (*26*). In another multicenter phase III clinical trial involving 487 patients with resectable HCC, those who underwent neoadjuvant HAIC had significantly higher 1-year, 2-year, and 3-year OS rates (97.7%, 86.3%, and 77.1%, respectively) compared to the surgery-alone group (90.0%, 80.9%, and 70.6%, respectively) (p = 0.032) (*27*). A retrospective analysis of 100 high-risk

patients with resectable HCC at various centers indicated that patients who received triple NAT consisting of lenvatinib, anti-programmed cell death-1 (PD-1) antibody, and transcatheter arterial chemoembolization (TACE) had a significantly improved DFS and OS compared to the surgery-alone group. The OS rates at 6, 12, 18, and 24 months in the NAT group were 100.0%, 100.0%, 100.0%, and 85.7%, respectively, whereas the surgery group's OS rates were 92.1%, 73.7%, 53.9%, and 48.7%, respectively ($p < 0.001$) (*28*). Moreover, the NAT group had markedly superior DFS rates at 6, 12, 18, and 24 months (82.2%, 66.95%, 48.8%, and 48.8%, respectively) compared to the surgery-alone group (41.92%, 28.34%, 27.05%, and 22.99%, respectively) ($p = 0.003$) (*28*). In patients with Chinese Liver Cancer (CNLC) stage IIb-IIIa resectable HCC, those who underwent NAT with camrelizumab plus apatinib for 1 year had significantly higher OS rates than the surgery-alone group (100% *vs.* 74.2%) ($p = 0.023$). In addition, the NAT group had a substantially lower 1-year recurrence rate than the surgery-alone group (42.9% *vs.* 64.0%, $p = 0.050$) (*29*). In the subset of HCC patients with a single tumor, the 1-year recurrence rate in the surgery-alone group was notably higher compared to the NAT group (71.0% *vs.* 25.0%, $p = 0.022$) (*29*).

NAT results in enhanced OS and DFS outcomes, particularly in patients with massive resectable HCC (≥ 10 cm). A 10-year retrospective analysis over the period from 2004 to 2014 revealed that patients with massive resectable HCC (≥ 10 cm) who underwent neoadjuvant TACE had a significantly improved median OS compared to the surgery-alone group (32.8 months *vs.* 22.3 months , $p = 0.035$) and a better DFS (12.9 months *vs.* 6.4 months, $p = 0.016$) (*30*). In patients with resectable HCC and portal vein tumor thrombus (PVTT), the neoadjuvant radiation therapy group had significantly improved 6-month, 12-month, 18-month, and 24-month OS rates (89.0%, 75.2%, 43.9%, and 27.4%, respectively) compared to the surgery-alone group (81.7%, 43.1%, 16.7%, and 9.4%, respectively) ($p < 0.001$) (*31*). Moreover, the corresponding DFS rates for the neoadjuvant radiation therapy group at 6, 12, 18, and 24 months (56.9%, 33.0%, 20.3%, and 13.3%, respectively) were superior to those of the surgery-alone group (42.1%, 14.9%, 5.0%, and 3.3%, respectively) ($p < 0.001$) (*31*). Among patients with HCC and PVTT, the neoadjuvant folinic acid, fluorouracil, and oxaliplatin (FOLFOX)-HAIC treatment group had significantly higher 1-year, 3-year, and 5-year OS rates (94.9%, 78%, and 66.4%, respectively) compared to the surgery-alone group (84.6%, 47.6%, and 37.2%, respectively) ($p < 0.001$) (*32*). In addition, the neoadjuvant FOLFOX-HAIC group had superior 1-year, 3-year, and 5-year recurrence-free survival (RFS) rates (88.7%, 56.2%, and 38.6%, respectively) compared to the surgery-alone group (84.9%, 38.3%, and 22.6%, respectively) ($p = 0.002$) (*32*).

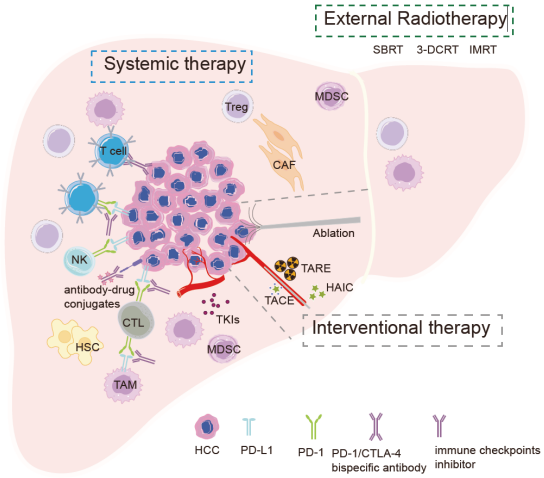

Figure 2. Schematic representation of neoadjuvant therapeutic mechanisms in hepatocellular carcinoma (HCC). Current neoadjuvant therapeutic strategies for HCC are characterized by the predominant utilization of interventional, radiation, and systemic modalities, with a discernible escalation in the prevalence of combination therapies within these treatment paradigms. Interventional approaches are exemplified by transarterial chemoembolization (TACE) and hepatic arterial infusion chemotherapy (HAIC). Radiation therapies include transarterial radioembolization (TARE), stereotactic body radiation therapy (SBRT), 3-dimensional conformal radiotherapy (3-DCRT), and intensity-modulated radiation therapy (IMRT). Systemic treatments predominantly involve tyrosine kinase inhibitors (TKI) and immune checkpoint inhibitors (ICI). An observable trend is the increasing adoption of diverse combination strategies among these therapeutic modalities.

3. NAT strategies prior to LR in HCC

One of the key objectives of NAT in resectable HCC is to enhance overall efficacy and prevent early postoperative metastasis. In addition, NAT functions as a biological assay to evaluate the feasibility of surgery and the tumor's responsiveness to treatment. The choice of an optimal NAT regimen is pivotal, given its substantial impact on patient prognosis. Through a thorough examination of current studies, as depicted in Figure 2, primary neoadjuvant strategies for HCC before LR encompass interventional therapy, radiation therapy, systemic treatment, and combination therapy.

3.1. Interventional therapy

The utilization of neoadjuvant TACE was initially reported by Monden *et al.* in 1989 (*33*). A retrospective analysis from 1990 to 1995 subsequently revealed that the 5-year DFS rate was 51.0% for the group that underwent TACE treatment two or more times preoperatively, 35.5% for the group that underwent TACE treatment once preoperatively, and 21.4% for the group that underwent no preoperative TACE treatment. The average DFS for these groups were 66.4 months, 22.5 months, and 12.5 months, respectively, suggesting a significant improvement in patient prognosis with preoperative TACE neoadjuvant therapy (*34*). In

2009, a study investigated the impact of preoperative TACE on the surgical outcomes of patients with resectable large HCC (diameter ≥ 5 cm). Although not statistically significant, the preoperative TACE group had a seemingly better DFS and OS than the control group (*35*). In 2010, a retrospective analysis of Korean patients with resectable HCC compared the survival outcomes of patients who received preoperative TACE treatment with those who underwent LR alone. The study, involving 1,530 patients with HCC, indicated that patients who underwent TACE before resection had similar 1-year, 2-year, and 5-year OS rates compared to those who did not receive preoperative treatment (p = 0.11) (*36*). However, patients in the preoperative TACE group had lower rates of DFS (*36*). In 2014, findings from a single-center study in China, encompassing 183 patients who received neoadjuvant TACE and 405 patients who underwent LR alone, had similar 1-year, 3-year, and 5-year OS rates (p = 0.739) (*37*). A phase III clinical study involving seven centers in China revealed that neoadjuvant FOLFOX-HAIC could improve the prognosis for patients with resectable BCLC A/B stage HCC beyond the Milan criteria. The disease control rate (DCR) in the NAT group reached 97.4% (*27*). A safety assessment indicated that neoadjuvant HAIC treatment was relatively safe, with rates of surgery-related adverse events (AEs) being similar between the NAT and control groups (p = 0.265) (*27*). Another phase III clinical trial, conducted between 2016 and 2020 at five hospitals in China, yielded comparable results. Patients in the NAT group had significantly better 6-month, 12-month, and 18-month PFS rates (77.6%, 50.4%, and 47.4%, respectively) than patients in the control group (52.7%, 42.8%, and 34.8%, respectively) (p = 0.017) (*26*). Preoperative ^{90}Y transarterial radioembolization (TARE) has demonstrated benefits in increasing the functional residual liver volume (*38*). Findings from a clinical study in 2023 revealed that patients with locally advanced HCC treated with ^{90}Y-selective internal radiation therapy (SIRT) before LR had a significantly improved 5-year OS and RFS compared to those underwent early LR (5-year OS 69.0% *vs.* 47.5%, p = 0.048; 5-year RFS 53.5% *vs.* 27.0%, p = 0.047) (*39*). Moreover, the 5-year OS and RFS in the NAT group were comparable to those of patients who underwent early LR (5-year OS 69.0% *vs.* 62.6%, p = 0.475; 5-year RFS 53.5% *vs.* 39.0%, p = 0.736) (*39*).

3.2. Radiation therapy

Preoperative treatment with ^{131}I-lipiodol has been found to lead to a reduction in serum alpha-fetoprotein (AFP) levels by more than 50% in 70% of patients (*40*). Out of 34 patients from whom postoperative tumor tissue samples were obtained, 25 displayed an objective response or tumor necrosis exceeding 90% (*40*). The RFS rates of the patients at 1, 2, and 3 years after surgery were 94%, 48%, and 48%, respectively (*40*).

Compared to the surgery-alone group, preoperative neoadjuvant three-dimensional conformal radiotherapy (3-DCRT) significantly reduced the recurrence rate and HCC-related mortality in patients with HCC and main portal vein thrombus (*41*). A clinical trial in 2018 indicated that preoperative SIRT can improve outcomes in patients with cirrhotic HCC; a major pathological response (MPR) was achieved postoperatively in 80% of patients treated with neoadjuvant radiotherapy and a pathological complete response (pCR) was achieved in 40% (*42*). In 2019, a study indicated that neoadjuvant 3-DCRT significantly reduced HCC-associated mortality and recurrence rates compared to surgery alone in patients with resectable HCC and PVTT (hazard ratio (HR): 0.35 *vs.* 0.45, $p < 0.001$) (*31*).

In 2020, a study indicated that preoperative treatment with SIRT facilitated the recruitment/activation of effector immune cells within the tumor. This resulted in a significant increase in tumor-infiltrating lymphocytes (TILs), CD4(+) T cells, CD8(+) T cells, and granzyme B (GZB) compared to patients in either the surgery-alone group or the group undergoing TACE preoperatively (*43*). A 2021 clinical study indicated a 65.3% 5-year OS rate for patients receiving neoadjuvant radiation therapy, compared to 46.6% in the surgery-alone group. In addition, the study found that neoadjuvant radiation therapy was significantly associated with improved OS (HR 0.549; p = 0.023) (*44*). In 2022, a phase II clinical trial investigated the use of neoadjuvant intensity-modulated radiotherapy (IMRT) for centrally located HCC. Results revealed notable outcomes, with 1-year, 3-year, and 5-year OS rates of 94.6%, 75.4%, and 69.1%, respectively. The DFS rates were 70.3%, 54.1%, and 41.0%, with a median DFS of 45.8 months. Moreover, an MPR was achieved in 34.2% of patients, and a pCR was achieved in 13.2% (*45*).

3.3. Systemic therapy

The advent of tyrosine kinase inhibitors (TKIs) marks a new era in systemic therapy for HCC, and sorafenib, which is a NAT, has exhibited a favorable safety profile in patients with resectable HCC (*46*). In recent years, immunotherapy has emerged as a prominent area of research for the treatment of HCC, and its main mechanisms include induction of immune responses, promotion of immunogenicity, regulation of immune responses, recruitment of cytotoxic immune cells, stimulation of cytotoxic T cell proliferation, reduction of immune tolerance, and other related factors. Extensive research is currently being conducted on monotherapy immunotherapy, and specifically immune checkpoint inhibitors (ICIs) that target PD-1, programmed cell death ligand-1 (PD-L1), and cytotoxic T-lymphocyte-associated protein-4 (CTLA-4).

As an example, the PD-1-targeting antibody

cemiplimab is used in NAT for patients with resectable HCC, yielding an R0 resection rate of 95.2%. Notably, a pCR with over 70% necrosis was observed in 20% of patients, and an MPR with 50–70% necrosis was observed in 15% (*47*). Following NAT with toripalimab, 80% of patients (8/10) underwent LR, with an incidence of MPR of 20% (*48*). In patients with resectable HCC receiving monotherapy with nivolumab, an MPR was achieved in approximately 33% (*49*).

3.4. Combination therapy

Combination therapy has shown promise in enhancing the efficacy of HCC treatment compared to monotherapy, making it a prospective approach to address the challenge of resistance to monotherapy as more clinical trials are conducted.

(1) Anti-PD-1 antibody combined with TKIs

In a study of 24 patients with resectable HCC receiving NAT with tislelizumab combined with lenvatinib, 17 patients (70.8%) underwent R0 resection, a pCR was achieved in 17.6%, and an MPR was achieved in 35.3% (necrosis >70%) (*50*). After undergoing NAT with nivolumab combined with cabozantinib, R0 resection was successfully performed in approximately 85.7% of patients. In addition, an MPR or a cPR was observed in 41.7% of tumor specimens (*51*). In the NAT group receiving toripalimab combined with lenvatinib, all 8 patients underwent surgical resection, and immunohistochemical analysis of tumor tissue infiltration revealed increased T-cell infiltration in responsive tumor tissue compared to non-responsive tumor tissue (*48*).

(2) Anti-PD-1 antibody combined with a vascular endothelial growth factor receptor (VEGFR) antagonist

In HCC patients with a high risk of recurrence, NAT combining camrelizumab and apatinib resulted in a favorable pathological response. In a study focusing on patients with resectable HCC with a high risk of recurrence, a 100% R0 surgical resection rate was achieved in those who underwent NAT with camrelizumab combined with apatinib (*52*). A MPR was observed in 38.5% of those patients, and a pCR was noted in 7.7%. Another clinical trial involving HCC patients with an intermediate to high risk of recurrence reported that approximately 89% of patients successfully underwent LR after receiving camrelizumab combined with apatinib, with a corresponding MPR rate of 46.2% in patients who underwent LR (*53*). In patients who underwent NAT with camrelizumab combined with apatinib, the LR rate was 94.4%, the MPR rate was 29.4% (5/17), and the pCR rate was 5.9% (1/17) (*54*).

(3) Anti-PD-1 antibody combined with anti-CTLA-4 antibody

In patients with resectable HCC, the combination of nivolumab and ipilimumab in NAT resulted in a significantly improved median PFS compared to nivolumab monotherapy (19.53 months *vs.* 9.4 months) (*49*). In patients who received combination NAT, the MPR rate was 27% (*49*). However, the incidence of grade 3–4 AEs with combination therapy was higher than that observed with nivolumab alone (43% *vs.* 23%) (*49*). Following NAT with ipilimumab combined with nivolumab, the DCR was 95%, and the MPR rate was 56% (*55*). In an ongoing phase II randomized controlled clinical trial, a pCR was achieved in approximately 25% of patients receiving neoadjuvant immunotherapy with nivolumab combined with ipilimumab (*56*).

4. Predictive biomarkers for NAT response and prognosis in HCC

Identifying robust biomarkers to predict NAT response and prognosis is pivotal to guiding treatment selection, optimizing intervention timing, and assessing surgical outcomes in HCC. Despite the evolving landscape of NAT for HCC, the scarcity of extensively validated biomarkers capable of reliably predicting efficacy and surgical success remains a challenge. Figure 3 provides an overview of biomarkers associated with treatment and prognosis in advanced HCC, serving as a foundation for an expanded exploration of NAT-related biomarkers in HCC.

4.1. Circulating biomarkers

Several studies have underscored the utility of changes in AFP levels as surrogate biomarkers, reflecting both systemic and local treatment responses throughout

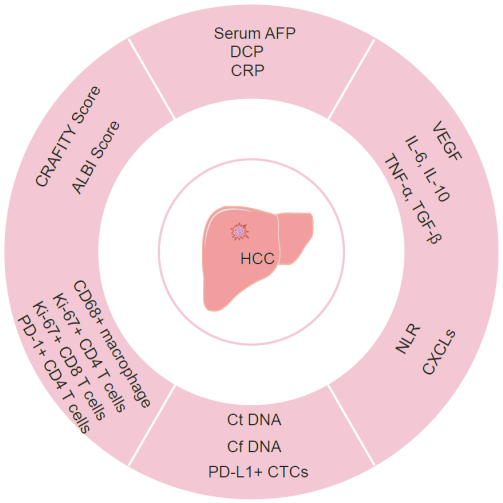

Figure 3. Overview of biomarkers associated with treatment and prognosis in hepatocellular carcinoma (HCC). Referencing studies on biomarkers associated with treatment and prognosis in advanced HCC provides valuable insights to further expand clinical research on neoadjuvant therapy (NAT) biomarkers in HCC. This includes circulating biomarkers such as serum AFP, DCP, CRP, the ALBI score, CRAFITY score, VEGF, IL-6, IL-8, IL-10, TGF-β, TNF-α, NLR, and CXCL9 and tumor microenvironment-related biomarkers such as Ki-67+ CD4 T cells and CD8 T cells, PD-1+ CD4 T cells, and CD68 macrophages. Liquid biopsy components consist of ctDNA, cfDNA, and circulating tumor cells (CTCs).

various stages of HCC treatment. Specifically, an early AFP response, defined by a >20% decline in serum AFP levels within the initial 4 weeks of treatment compared to baseline, has emerged as an independent predictor associated with a prolonged OS and PFS in advanced HCC treated with ICI (*57*).

Monitoring AFP levels during atezolizumab plus bevacizumab (Atez/Bev) treatment is essential for real-time assessment and treatment optimization. In a prospective multicenter study, researchers defined optimal thresholds for AFP response in patients with uHCC who received Atez/Bev treatment. An AFP response of 50% or more and 20% or more was associated with the objective response rate (ORR) and the DCR, respectively. Both responses were also associated with PFS (*58*). The phase Ib GO30140 study proposed using AFP criteria at 6 weeks to identify responders and disease controllers for Atez/Bev treatment (*59*). AFP thresholds delineated in the study, involving a decline of at least 75% and a rise of no more than 10% from the baseline at 6 weeks, were used to discern responders to Atez/Bev and disease controllers, respectively. In HCC patients with AFP levels exceeding 20 ng/mL, a decrease of ≥20% in AFP at 3 weeks can predict the tumor prognosis in patients undergoing Atez/Bev treatment. Combining this with the albumin-bilirubin (ALBI) score enhances the accuracy of prognostic discrimination (*59*).

The CELESTIAL phase III study established that maintaining AFP levels without an increase from the baseline at 8 weeks serves as the most reliable predictor for prolonged OS and PFS in patients with advanced HCC treated with cabozantinib (*60*). Outcomes from the REACH and REACH-2 phase III trials revealed that patients treated with ramucirumab had a prolonged OS when manifesting an AFP response, defined as a reduction of at least 20% from the baseline (*61*). Ramucirumab treatment was considered suitable for patients with AFP levels of at least 400 ng/ml (*61*). In patients with a baseline AFP level of ≥10 ng/mL, an AFP response (defined as a reduction of ≥10% from the baseline) may have a significant effect on the treatment outcomes of patients with HCC who underwent lenvatinib therapy. For patients with an AFP level <10 ng/mL, the baseline ALBI score and the change in ALBI score from the baseline to the one-month post-treatment estimate could play a crucial role in determining treatment response (*62*).

The C-reactive protein and alpha-fetoprotein in immunotherapy (CRAFITY) score, derived from a multicenter retrospective study in Japan, is designed to predict treatment outcomes and treatment-associated AEs among patients with diverse stages of HCC undergoing Atez/Bev therapy. Patients with an AFP level ≥ 100 ng/mL and C-reactive protein (CRP) ≥ 1 mg/dL received a CRAFITY score of 1 (*63*). Concurrently, a multicenter retrospective study in Europe, in line with the Japanese study, found that the CRAFITY score correlated with patient survival and radiographic response during PD-(L)1 immunotherapy (*64*). In Japan, another retrospective multicenter study, encompassing 426 patients with HCC treated with Atez/Bev, established the mALF score based on a baseline mALBI grade of 2b or 3 (HR 2.36, $p = 0.002$) and AFP ≥ 100 ng/ml (HR 2.61, $p < 0.001$). This study validated the mALF score's robust predictive capability for survival in patients undergoing Atez/Bev treatment for HCC (*65*). A retrospective analysis evaluating disease response rate and changes in AFP and des-gamma-carboxy prothrombin (DCP) levels at 1, 2, 3, and 6 weeks, respectively, suggested that an AFP/DCP ratio of 1.4 or higher at 3 weeks may serve as an early predictor for advanced HCC treated with Atez/Bev (*66*). Studies indicated that changes in the Response Evaluation Criteria in Solid Tumors (RECIST), AFP, and DCP can be scrutinized for early response assessment in HAIC (*67*). Within clinical trials of neoadjuvant FOLFOX-HAIC therapy, a logistic regression model integrating AFP and CRP resulted in enhanced precision in predicting neoadjuvant FOLFOX-HAIC response, boasting a sensitivity of 72.2% and specificity of 72.4% (*32*).

In a phase II clinical study evaluating pembrolizumab for uHCC, Lynn *et al.* identified a correlation between reduced efficacy of pembrolizumab treatment and higher plasma levels of transforming growth factor-beta (TGF-β) (≥ 200 pg/mL) in patients (*68*). Moreover, elevated serum interleukin 6 (IL-6) (> 18.49 pg/mL) was linked to diminished clinical benefits, defined as achieving complete or partial remission or disease stabilization for ≥6 months, in patients receiving Atez/Bev for uHCC (*69*). Patients with a lower baseline IL-6 levels had increased response rates and prolonged PFS and OS following Ate/Bev treatment compared to those with elevated baseline IL-6 levels (*70*).

HCC patients with elevated serum IL-10 levels exhibit a substantial suppression of peripheral blood lymphokine-activated killer (LAK) and natural killer (NK) cell activity (*71*). In a prospective study, patients with serum IL-10 levels exceeding 1 pg/mL had a shorter OS (5.0 months *vs.* 14.9 months; $p < 0.0001$), and the IL-10 level emerged as an independent prognostic factor (HR 1.824; $p = 0.0005$) (*72*). In a multicenter phase II pilot study, researchers identified baseline levels of IL-6 at 8.58 pg/mL and IL-8 at 57.9 pg/mL as effective thresholds for predicting OS in uHCC patients treated with sorafenib (*73*). Baseline IL-6 and IL-8, with their respective cut-off values, can serve as predictors for ORR based on modified RECIST (mRECIST) in a subset of 42 patients with available follow-up imaging (IL-6, 46.6% *vs.* 19.2%, $p = 0.007$; IL-8, 50.0% *vs.* 17.4%, $p = 0.011$) (*73*). Moreover, plasma IL-8 and tumor necrosis factor-alpha (TNF-α) levels may serve as predictors of response to sorafenib in uHCC patients during early treatment (5–10 days) (*74*).

The pre-treatment assessment of serum vascular

endothelial growth factor (VEGF) levels has emerged as a promising prognostic biomarker for ablative interventions in HCC. Patients with serum VEGF levels surpassing 240 pg/mL had diminished OS and RFS rates (*75*).

The neutrophil-to-lymphocyte ratio (NLR) is instrumental in evaluating the effectiveness of neoadjuvant TACE therapy (*76*). Notably, patients with a high NLR (≥ 1.6) within the TACE plus sequential resection cohort had a markedly lower 5-year OS rate compared to those with a low NLR (78.4% *vs.* 100%, $p = 0.027$) (*76*). Robust evidence supports the pivotal role of NLR in predicting outcomes of Atez/Bev therapy in HCC patients. As a predictive marker for Atez/Bev response in HCC, pre-treatment NLR was significantly lower in patients in whom disease control was achieved compared to that in patients experiencing disease progression (2.47 *vs.* 4.48, $p = 0.013$). Moreover, patients with NLR ≤ 3.21 had a significantly superior PFS compared to those with NLR > 3.21 ($p < 0.0001$) (*77*). In a separate study, the observed difference in cumulative OS at 2, 4, 6, and 8 months between patients with low (< 3.0) and high NLR (≥ 3.0) in HCC patients treated with Atez/Bev was statistically significant ($p = 0.001$) (*78*). Nonetheless, there were no statistically significant differences in the response to combination therapy between patients with a low and high NLR (*78*). In terms of AEs, notable differences were noted in immune-related liver injury, decreased appetite of any grade, proteinuria of at least grade 3, and AEs of any other grade between patients with a low and high NLR (*78*).

Recent findings have elucidated the optimal threshold for NLR-2c initiation at the outset of the second therapeutic course in patients with uHCC who underwent Atez/Bev treatment, identifying it as 1.97 (*79*). Notably, patients with an NLR-2c < 1.97 had a superior OS and PFS compared to those with NLR-2c ≥ 1.97 (*79*). In a cohort of Japanese patients with HCC treated with Atez/Bev, a baseline NLR ≥ 3 emerged as the exclusive independent factor associated with highly progressive disease (*80*). A German study corroborated NLR > 3.2 as the most critical predictor of poorer ORR and PFS (*81*). Moreover, a multicenter international retrospective cohort study independently established NLR ≥ 5 as a predictor of inferior survival outcomes (*82*).

A study conducted in Japan has validated the potential of serum chemokine C-X-C motif ligand 9 (CXCL9) as a predictive indicator for early disease progression after Atez/Bev treatment (*83*). The research established that the optimal serum CXCL9 threshold for predicting early disease progression in uHCC treated with Atez/Bev is 333 pg/mL, with a sensitivity of 60.0% and specificity of 92.3%. Patients with lower serum CXCL9 levels (< 333 pg/mL) had a higher likelihood of early disease progression, accompanied by a significantly shorter median PFS compared to those with higher levels (126 days *vs.* 227 days, $p = 0.0084$). Notably, patients exhibiting an objective response to lenvatinib displayed notably lower baseline serum CXCL9 levels than those without an objective response (*83*).

4.2. Tumor microenvironment (TME)-related biomarkers

As an immune organ housing a diverse array of immune cells, the liver is particularly prone to developing immunotherapy tolerance. Early recurrent HCC displays reduced levels of T regulatory cells (Tregs) and elevated levels of dendritic cells (DCs) and CD8(+) T cells in comparison to primary HCC (*84*) . An immunohistochemical examination of human HCC tissues has revealed that PD-L1 is preferentially expressed in CD68 macrophages within the TME. Among patients undergoing nivolumab treatment, 3 out of 8 had a positive response to anti-PD-1 therapy. Responders had a higher proportion of Ki-67+ CD4 and CD8 T cells in their blood compared to non-responders (*85*). The greater the number of cells expressing CD68 and PD-L1 in the tumor, the more favorable the response to multikinase inhibitors in patients with HCC (*86*). In patients receiving neoadjuvant treatment with camrelizumab in combination with apatinib, tumor immune microenvironment (TIME) cell infiltration, particularly of DCs, was observed to be more favorable in responding tumors than in non-responding tumors (*54*). A recent study has indicated that patients with a higher baseline frequency of PD-1+ CD4 cells are more likely to exhibit positive responses to anti-CTLA4 therapy, including trastuzumab (*87*). Moreover, studies integrating single-cell and spatial transcriptomics data have found that the structural composition of the tumor immune barrier within the TME may also influence the efficacy of immunotherapy (*88*).

4.3. Liquid biopsy

In patients with HCC, the levels of circulating tumor DNA (ctDNA) are correlated with tumor size, extrahepatic spread, and vascular infiltration. Liquid biopsy, utilizing ctDNA and circulating tumor cells (CTCs), has emerged as a promising method for predicting treatment response and prognosis. In a phase II study involving camrelizumab plus apatinib for HCC treatment, ctDNA played a crucial role in predicting pathological response and RFS (*54*). A Japanese study explored the potential for cell-free DNA (cfDNA)/ctDNA in peripheral blood to serve as a biomarker with which to predict treatment response in patients with uHCC treated with Atez/Bev (*89*). The study revealed that elevated cfDNA levels pretreatment were linked to lower response rates and a shorter PFS and OS. Telomerase reverse transcriptase (TERT) mutations in peripheral blood cfDNA and serum AFP levels ≥400 ng/mL were identified as independent predictors of poor OS following Atez/Bev treatment (*89*). These factors provide

a basis for stratifying patients undergoing Atez/Bev therapy based on prognosis (*89*).

A phase II study indicated that ctDNA can serve as a predictor of pathological response and relapse following treatment with camrelizumab and apatinib (*54*). Patients in whom a pCR/MPR was achieved at the baseline had a higher mutation burden compared to patients in whom a pCR/MPR was not achieved (6 mutations *vs.* 2.5 mutations, $p = 0.025$). There was a noticeable trend towards a shorter RFS in ctDNA-positive patients after adjuvant therapy compared to ctDNA-negative patients (*54*). The clinical predictive significance of mutations in the human TERT (hTERT) promoter in free DNA for the treatment of advanced HCC has been established. Responders who had peak DNA levels within one week of TKI initiation had a significantly improved PFS compared to non-responders ($p = 0.004$). The extent of mutant DNA changes after TACE was significantly correlated with tumor volume ($p < 0.001$) (*90*).

CTCs are regarded as ideal biomarkers due to their cancer-specific characteristics. PD-L1+ CTCs can serve as an independent predictor of OS ($p = 0.010$). Patients with PD-L1+ CTCs have a worse OS compared to those lacking PD-L1+ CTCs (14.0 months *vs.* not achieved, $p = 0.001$) (*91*). In patients with HCC treated with anti-PD-1/PD-L1, the presence of PD-L1+ CTCs was strongly associated with a favorable treatment response (*91,92*). Specifically, in patients with uHCC receiving a combined regimen of IMRT, anti-PD-1 antibodies, and antiangiogenic drugs, those with PD-L1+ CTC counts below 2 have a prolonged ORR and OS in comparison to patients with counts above 2 (ORR: 56. 5% *vs.* 16.7%, $p = 0.007$; OS: not reached *vs.* 10.8 months, $p = 0.001$) (*93*).

5. Ongoing clinical studies on preoperative NAT for HCC

In the current landscape of global clinical trials exploring neoadjuvant locoregional therapy for HCC, various treatment modalities take precedence, including TACE-HAIC (FOLFOX), (oxaliplatin, leucovorin, and 5-fluorouracil) mFOLFOX6-TAI, FOLFOX-HAIC, and (cisplatin, doxorubicin hydrochloride, and thalidomide) PLADOTH-TACE (as shown in Table 2A). Ongoing clinical trials of NAT for HCC also involve sorafenib monotherapy, sorafenib combined with capecitabine and oxaliplatin, and lenvatinib in conjunction with TACE (as shown in Table 2B). In addition, ongoing studies into neoadjuvant radiation therapy for HCC are outlined in Table 2C. Neoadjuvant immunotherapy studies, both as monotherapy and as combination therapy, constitute a significant focus (as shown in Table 3). Combination therapies are broadly categorized into two- and three-agent combinations. Noteworthy combinations involve anti-PD-1 antibodies paired with interventional therapy, radiation therapy, VEGFR antagonists, VEGF/VEGFR monoclonal antibodies, anti-CTLA-4 antibodies, C-C chemokine receptor 2/5 (CCR2/5) inhibitors, or anti-IL-8 antibodies. Triple NAT options include combinations of anti-PD-L1 antibodies with VEGF/VEGFR monoclonal antibodies and interventional therapy, anti-PD-1 antibodies with VEGFR antagonists and chemotherapeutic agents, anti-PD-L1 antibodies with radiation therapy and interventional therapy, anti-PD-L1 antibodies with anti-CTLA-4 antibodies and radiation therapy, anti-PD-1 antibodies with VEGFR antagonists and radiation therapy, and anti-PD-L1 antibodies with TKIs and radiation therapy.

In a phase 1b study evaluating neoadjuvant cabozantinib and nivolumab in patients with locally advanced or borderline resectable HCC, an R0 resection was achieved in 85.7% of patients (12/14) who completed NAT (*51*). A pathological examination revealed a MPR with over 90% tumor necrosis in 42% of patients (5/12). Immunoassays revealed a significant enrichment in the spatial arrangement of T effector cells, tertiary lymphoid structures, and CD138+ plasma cells and B cells in responders compared to non-responders (*51*).

6. Appropriate research endpoints

In the context of current clinical research paradigms, as delineated in Tables 2 and 3, the primary research endpoints in neoadjuvant local chemotherapy clinical trials encompass OS, PFS, and event-free survival (EFS), with a study duration spanning 3-5 years. For neoadjuvant TKI monotherapy or combination therapy, key endpoints include significant pathologic response, the surgical resection rate, and DFS in a study, conducted over a period of 56 days to 1 year. In neoadjuvant radiotherapy trials, primary endpoints consist of OS, the dropout rate, and treatment-related adverse events (trAEs) according to the Common Terminology Criteria for Adverse Events (CTCAE v5.0), with a study period ranging from 3 months to 1 year. Neoadjuvant immunotherapy with ICIs is characterized by a comprehensive set of primary research endpoints, including OS, RFS, EFS, pCR, MPR, DFS, significant tumor necrosis (STN), ORR, the resection rate, delayed surgery rate, immune-related AEs (irAEs, CTCAE v5.0), and lesion reduction >10% (RECIST v1.1), in a study over a period of 6 weeks to 4 years.

In phase III clinical studies, the primary endpoint emphasis is on OS and RFS. Phase II studies commonly use RFS, STN, ORR, MPR, the tumor response rate (mRECIST), time to recurrence, time to progression, resection rate, and DFS as primary endpoints. Phase I studies prioritize the assessment of AEs, irAEs, the number of patients completing preoperative treatment and undergoing surgical intervention, recurrence rate, ORR, and pathological response.

Treatment with TKIs has been found to be associated with a reduced probability of tumor shrinkage, whereas

Table 2A. Clinical trials investigating neoadjuvant locoregional chemotherapy for hepatocellular carcinoma

Research phase	Treatment	Disease	Number of cases	Treatment modality	Primary endpoint	Time window	NAT duration	Interval from NAT completion to surgery	NCT number	Region
III	mFOLFOX6-TAI	HCC with PVTT	230	NAT	OS	5 years	N/A	N/A	NCT03368651	China
III	mFOLFOX6-TAI	resectable HCC beyond Milan criteria	344	NAT	OS	5 years	N/A	N/A	NCT03851913	China
III	FOLFOX-HAIC	resectable HCC beyond Milan criteria	252	NAT	OS	5 years	4 cycles	N/A	NCT03469479	China
II	PLADOTH-TACE	resectable HCC	47	NAT/AT	1.EFR 2.OS	N/A	N/A	N/A	NCT00276705	UK
N/A	TACE-HAIC(FOLFOX)	resectable HCC	320	NAT	PFS	3 years	N/A	N/A	NCT04777942	China
N/A	TACE-HAIC(FOLFOX)	resectable HCC	280	NAT	PFS	3 years	N/A	N/A	NCT04424043	China
N/A	TACE-HAIC(FOLFOX)	HCC with PVTT	320	NAT	PFS	3 years	N/A	N/A	NCT04181931	China

Table 2B. Ongoing clinical trials on neoadjuvant treatment of hepatocellular carcinoma with tyrosine kinase inhibitors

Research phase	Treatment	Disease	HCC stage/liver function	Number of cases	Treatment modality	Primary endpoint	Time window	NAT duration	Interval from NAT completion to surgery	NCT number	Region
II	Sorafenib	resectable HCC	Child-Pugh B/C	36	NAT	1.Antiangiogenic effects 2.Significant pathological changes	on day 50 and at 3 months after surgery	4 weeks	7 days	NCT01182272	France
II	Sorafenib, capecitabine, oxaliplatin	resectable HCC	Child-Pugh A	15	NAT	Resectability	at the end of cycle 4 (each cycle is 14 days)	56 days		NCT03578874	Hong Kong
N/A	TACE+lenvatinib	resectable HCC	CNLC III, Child-Pugh A/B	164	NAT	DFS	1 year			NCT04961138	China

Table 2C. Ongoing clinical trials on neoadjuvant radiotherapy for hepatocellular carcinoma

Research phase	Treatment	Disease	HCC stage/liver function	Number of cases	Treatment modality	Primary endpoint	Time window	NAT duration	Interval from NAT completion to surgery	NCT number	Region
III	Radiotherapy	HCC with PVTT	Child-Pugh A/B	214	NAT	OS	1 year	N/A	4 weeks	NCT04025437	China
I	SBRT	resectable HCC	Child-Pugh 0/A	30	NAT	Drop-out rate	5 months	N/A	4-6 weeks	NCT04587739	France
I	SBRT	resectable HCC	BCLC A, Child-Pugh A/B	15	NAT	trAE (CTCAE v5.0)	3 months after surgery	N/A	N/A	NCT05598060	China

Abbreviations: AE, adverse event; AT, adjuvant therapy; BCLC, Barcelona Clinic Liver Cancer; CR, complete response; CTCAE, Common Terminology Criteria for Adverse Events; DFS, disease-free survival; EFR, event-free survival; FOLFOX, infusional fluorouracil, leucovorin, and oxaliplatin; HAIC, hepatic arterial infusion chemotherapy; HCC, hepatocellular carcinoma; mFOLFOX6, infusional oxaliplatin, calcium folinate, and 5-FU; N/A, not applicable; NAT, neoadjuvant therapy; OS, overall survival; PFS, progression-free survival; PLADOTH, cisplatin, doxorubicin hydrochloride, and thalidomide; PR, partial response; PVTT, portal vein tumor thrombus; RECIST, Response Evaluation Criteria in Solid Tumors; SBRT, stereotactic body radiotherapy; TACE, transcatheter arterial chemoembolization; TAI, transarterial chemoinfusion; TARE, transarterial radioembolization; 3D, three-dimensional; trAE, treatment-related adverse event.

Table 3. Ongoing clinical trials on neoadjuvant immune checkpoint inhibitors for hepatocellular carcinoma

Research phase	Treatment	Disease	HCC stage/liver function	Number of cases	Treatment modality	Primary endpoint	Time window	NAT duration	Interval from NAT completion to surgery	NCT number	Region
III	TACE/HAIC combined with lenvatinib+sintilimab	resectable HCC	BCLC B/C	90	NAT	RFS	1 year	N/A	3 months from start	NCT05250843	China
III	Camrelizumab combined with apatinib	resectable HCC	CNLC Ib/IIa/IIb/IIIa	130	NAT/AT	RFS	3 years	4 weeks	≥ 1 week	NCT05613478	China
II	Pembrolizumab+ surgery/ ablation	resectable HCC	BCLC 0/A	50	NAT/AT	RFS	1 year	once	N/A	NCT03337841	Japan
II	Tislelizumab+IMRT	resectable HCC	BCLC 0/A	30	NAT	RFS	2 years	N/A	N/A	NCT04850157	China
II	Apatinib+camrelizumab+oxaliplatin	resectable HCC	BCLC 0/A	15	NAT	MPR#	2 years	N/A	N/A	NCT04850040	China
II	Tislelizumab	resectable HCC	BCLC A/B	80	NAT	DFS	1 year	4 weeks	2 weeks	NCT04615143	China
II	Immune-checkpoint blockade therapy (AK104) combined with TACE	resectable HCC	BCLC A/B	54	NAT	MPR#	2 years	4 weeks	N/A	NCT05578430	China
II	Cemiplimab+SBRT	resectable HCC	BCLC A/B/C	73	NAT/AT	STN	Upon surgery	N/A	N/A	NCT03916627	USA
II	Cemiplimab	resectable HCC	BCLC A/B/C	73	NAT/AT	STN	Upon surgery	N/A	N/A	NCT03916627	USA
II	Cemiplimab+fianlimab	resectable HCC	BCLC A/B/C	73	NAT/AT	STN	Upon surgery	N/A	N/A	NCT03916627	USA
II	Sintilimab+TACE+radiotherapy	resectable HCC	BCLC B/C	10	NAT	EFS	4 years	N/A	N/A	NCT04653389	China
II	Apatinib+camrelizumab	resectable HCC	BCLC B/C	78	NAT/AT	RFS	1 year	7 weeks	N/A	NCT04930315	China
II	Nivolumab+ipilimumab	resectable HCC	BCLC B/C	40	NAT	Tumor shrinkage**	4 years	6/12 weeks	N/A	NCT03510871	China
II	Aatezolizumab+bevacizumab	resectable HCC	BCLC B/C	45	NAT	pCR	6 months	4 weeks	N/A	NCT04954339	South Korea
II	Regorafenib+durvalumab	resectable HCC	Child-Pugh A Child-Pugh A	27	NAT	ORR	16 weeks	3 weeks	1 week	NCT05194293	USA
II	Nivolumab	Potentially resectable HCC	Child-Pugh A	20	NAT	pRR	After surgery (normally 6 weeks after the start of nivolumab)	5 weeks	2 weeks	NCT05471674	Hong Kong
II	Pembrolizumab+lenvatinib	resectable HCC	Child-Pugh A	43	NAT/AT	MPR*	24 weeks	9 weeks	1week	NCT05389527	China
II	Pembrolizumab+lenvatinib *vs.* pembrolizumab/lenvatinib alone	resectable HCC	Child-Pugh A	60	NAT	MPR#	4 months	6/4/6 weeks	N/A	NCT05185739	UK
II	Atezolizumab/bevacizumab *vs.* neoadjuvant SBRT	HCC with PVTT	Child-Pugh A	70	NAT	proportion of LR	17 weeks	10 weeks/10 days	N/A	NCT05137899	Canada

Abbreviations: AE, adverse event; AT, adjuvant therapy; BCLC, Barcelona Clinic Liver Cancer; CNLC, China Liver Cancer Staging System; CR, complete response; CTCAE, Common Terminology Criteria for Adverse Events; DFS, disease-free survival; EFS, event-free survival; HAIC, hepatic arterial infusion chemotherapy; HCC, hepatocellular carcinoma; HKLC, Hong Kong Liver Cancer staging system; IMRT, intensity modulated photon therapy; irAE, immune related adverse event; MPR, major pathological response; MPR$^{\#}$, defined as <10% viable tumor within resection; MPR*, defined as ≥ 50% necrosis pathologically in the resected specimen; mRECIST, modified Response Evaluation Criteria in Solid Tumors; N/A, not applicable; NAT, neoadjuvant therapy; ORR, objective response rate; OS, overall survival; pCR, pathologically complete response; PFS, progression-free survival; pPR, pathological partial response; PR, partial response; pRR, pathological response rate; PVTT, portal vein tumor thrombus; RECIST, Response Evaluation Criteria in Solid Tumors; RFS, relapse-free survival; SBRT, stereotactic body radiotherapy; STN, significant tumor necrosis; SIRT, selective internal radiation treatment; TACE, transcatheter arterial chemoembolization; TAI, transarterial chemoinfusion; trAE, treatment-related adverse event; UICC, Union for International Cancer Control classification system.

Table 3. Ongoing clinical trials on neoadjuvant immune checkpoint inhibitors for hepatocellular carcinoma (continued)

Research phase	Treatment	Disease	HCC stage/liver function	Number of cases	Treatment modality	Primary endpoint	Time window	NAT duration	Interval from NAT completion to surgery	NCT number	Region
II	Tremelimumab+durvalumab	resectable HCC	Child-Pugh A	28	NAT/AT	AE	4 years	5 weeks	N/A	NCT05440864	Spain
II	Nivolumab	resectable HCC	N/A	36	NAT	(1) MPR# (2) STN	2 years	5 weeks	N/A	NCT04123379	USA
II	Nivolumab+BMS-813160 (CCR2/5 inhibitor)	resectable HCC	N/A	36	NAT	(1) MPR# (2) STN	2 years	5 weeks	N/A	NCT04123379	USA
II	Nivolumab+BMS-986253 (anti IL-8)	resectable HCC	N/A	36	NAT	(1) MPR# (2) STN	2 years	5 weeks	N/A	NCT04123379	USA
Ib/II	Anlotinib hydrochloride +TQB2450	resectable HCC	BCLC A/B	20	NAT	(1) pCR (2) ORR	6 months	N/A	N/A	NCT04888546	China
I/II	Ipilimumab/ipilimumab+nivolumab	resectable HCC	Child-Pugh A	32	NAT	(1) Delay to surgery (2) trAE	(1) 89 days (2) 127 days	/6 weeks	N/A	NCT03682276	UK
I	Durvalumab+tremelimumab *vs.* durvalumab+tremelimumab+SIRT	resectable HCC	BCLC 0/A	20	NAT /AT	AE (at least one grade 3-5 trAE according to CTCAE v5.0)	18 months	4 weeks	3 weeks/24 days	NCT05701488	USA
I	SBRT+atezolizumab+bevacizumab	resectable HCC	BCLC A/B	20	NAT	AE (grade 3-4 trAE according to CTCAE v5.0)	6 months	6 weeks	6-8 weeks	NCT04857684	USA
I	Pembrolizumab	resectable HCC	BCLC B	45	NAT	(1) Recurrence rate (2) Number of CD8+ Ki67+ T cells in tumor	2 years after operation	one dose	4 weeks	NCT04224480	Singapore
I	Lenvatinib+sintilimab+radiotherapy	HCC with PVTT	BCLC C	20	NAT	(1) AE (2) Number of patients who complete surgery	5 years	N/A	N/A	NCT05225116	China
I	Tislelizumab+SBRT	resectable HCC	N/A	20	NAT	(1) Delay to surgery (number of patients experiencing a surgical delay of 6 weeks or longer) (2) ORR (3) pCR, pPR, MPR (4) irAE (CTCAE v5.0)	(1) 92 days (2) 1 day before LR (3) 1 month after LR (4) 2 months after LR	N/A	N/A	NCT05185531	China

Abbreviations: AE, adverse event; AT, adjuvant therapy; BCLC, Barcelona Clinic Liver Cancer; CNLC, China Liver Cancer Staging System; CR, complete response; CTCAE, Common Terminology Criteria for Adverse Events; DFS, disease-free survival; EFS, event-free survival; HAIC, hepatic arterial infusion chemotherapy; HCC, hepatocellular carcinoma; HKLC, Hong Kong Liver Cancer staging system; IMRT, intensity modulated photon therapy; irAE, immune related adverse event; MPR, major pathological response; MPR$^{\#}$, defined as <10% viable tumor within resection; MPR*, defined as ≥ 50% necrosis pathologically in the resected specimen; mRECIST, modified Response Evaluation Criteria in Solid Tumors; N/A, not applicable; NAT, neoadjuvant therapy; ORR, objective response rate; OS, overall survival; pCR, pathologically complete response; PFS, progression-free survival; pPR, pathological partial response; PR, partial response; pRR, pathological response rate; PVTT, portal vein tumor thrombus; RECIST, Response Evaluation Criteria in Solid Tumors; RFS, relapse-free survival; SBRT, stereotactic body radiotherapy; STN, significant tumor necrosis; SIRT, selective internal radiation treatment; TACE, transcatheter arterial chemoembolization; TAI, transarterial chemoinfusion; trAE, treatment-related adverse event; UICC, Union for International Cancer Control classification system.

Table 3. Ongoing clinical trials on neoadjuvant immune checkpoint inhibitors for hepatocellular carcinoma (continued)

Research phase	Treatment	Disease	HCC stage/liver function	Number of cases	Treatment modality	Primary endpoint	Time window	NAT duration	Interval from NAT completion to surgery	NCT number	Region
I	Nivolumab *vs.* nivolumab +relatlimab	resectable HCC	N/A	20	NAT	Number of patients who complete surgery	4 years	N/A	N/A	NCT04658147	USA
N/A	HAIC/lenvatinib+sintilimab	resectable HCC	Child-Pugh A	60	NAT	DFS	1 year	7 weeks	N/A	NCT05621499	China
N/A	Camrelizumab+apatinib+TACE	resectable HCC	BCLC B/C	290	NAT	EFS	3 years	3 weeks	2-4 weeks	NCT04521153	China

Abbreviations: AE, adverse event; AT, adjuvant therapy; BCLC, Barcelona Clinic Liver Cancer; CNLC, China Liver Cancer Staging System; CR, complete response; CTCAE, Common Terminology Criteria for Adverse Events; DFS, disease-free survival; EFS, event-free survival; HAIC, hepatic arterial infusion chemotherapy; HCC, hepatocellular carcinoma; HKLC, Hong Kong Liver Cancer staging system; IMRT, intensity modulated photon therapy; irAE, immune related adverse event; MPR, major pathological response; MPR#, defined as <10% viable tumor within resection; MPR*, defined as ≥ 50% necrosis pathologically in the resected specimen; mRECIST, modified Response Evaluation Criteria in Solid Tumors; N/A, not applicable; NAT, neoadjuvant therapy; ORR, objective response rate; OS, overall survival; pCR, pathologically complete response; PFS, progression-free survival; pPR, pathological partial response; PR, partial response; pRR, pathological response rate; PVTT, portal vein tumor thrombus; RECIST, Response Evaluation Criteria in Solid Tumors; RFS, relapse-free survival; SBRT, stereotactic body radiotherapy; STN, significant tumor necrosis; SIRT, selective internal radiation treatment; TACE, transcatheter arterial chemoembolization; TAI, transarterial chemoinfusion; trAE, treatment-related adverse event; UICC, Union for International Cancer Control classification system.

ICIs may result in unconventional radiologic response patterns, such as delayed responses or pseudoprogression, initially appearing as an increased tumor burden and later transforming into radiologic shrinkage (*94*). This poses a challenge to the use of conventional response criteria such as RECIST v1.1 (*95*) and mRECIST (*96*). RECIST v1.1, for instance, fails to account for complete pathologic necrosis of HCC with lipiodol deposition as a result of conventional TACE (*97*). In addition, the mRECIST criteria necessitate subtraction imaging for an accurate assessment of complete pathologic necrosis (*98*).

Immunotherapy-related imaging tumor response assessment criteria, such as immune-related response criteria (irRC) (*99*), immune-related RECIST (irRECIST) (*100*), immune RECIST (iRECIST) (*101*), immune-modified RECIST (imRECIST) (*102*), and intra-tumoral RECIST (itRECIST) (*103*), are designed to measure treatment response or disease progression in patients who underwent immunotherapy, and the use of those criteria shows promise (as shown in Table 4). A recent proposal by Japanese researchers regarding combination therapy involving systemic and local therapies outlined clinical complete response (cCR) criteria (*104*): (1) Attainment of a complete response (CR) according to the mRECIST/RECIST v1.1 criteria assessed with CT/MRI, and (2) Attainment of a CR indicated according to three tumor markers (AFP, vitamin K absence II (PIVKA-II), and AFP bound to Lens culinaris agglutinin (AFP-L3)) that have remained continuously normalized for more than 6 weeks.

Determining the optimal duration of therapeutic intervention is a crucial consideration within the clinical landscape. Ordinarily, cemiplimab is used as a neoadjuvant within a concise 22-day protocol (*47*), while the administration of nivolumab as a neoadjuvant, whether as a monotherapy or in conjunction with ipilimumab, entails a more protracted 6-week regimen (*49*). A point worth noting is that there is a discernible inverse correlation between the duration of treatment administered to patients before surgery and the subsequent pathological response rate. The main goals of NAT are reducing the risk of recurrence by eliminating micro-metastatic disease that cannot be detected by imaging and facilitating treatment of the primary tumor through cytoreductive surgery. Given these goals, the primary reason for using NAT is to stimulate an immune response against micro-metastatic disease rather than directly killing the tumor. Consequently, interventions of a shorter duration may offer comparable benefits while potentially mitigating the risk of preoperative irAEs that could compromise planned surgical procedures. According to the Chinese expert consensus and related studies, NAT should typically last 1.5–3 months, with a maximum duration of 4 months. The goal of this therapy is to achieve the surgical objective as soon as possible, regardless of whether the lesion has shrunk or not (*25*).

Table 4. Radiological assessment criteria for tumor response

Criteria (*Ref.*)	RECIST v1.1 (*95*)	mRECIST (*96*)	irRC (*99*)	irRECIST (*100*)	iRECIST (*101*)	imRECIST (*102*)	itRECIST (*103*)
Lesion definition	Uni-dimensional, largest diameter	Uni-dimensional, enhancing tumor	Bi-dimensional	Uni-dimensional	Uni-dimensional, largest diameter	Uni-dimensional	Uni-dimensional
Target lesions	Measurable (> 10 mm), largest lesions; Lymph node ≥ 15 mm in short axis	Measurable (> 10 mm), largest lesions with arterial-phase enhancement; Lymph node ≥ 20mm in short axis at the porta hepatis	Measurable (≥ 5 × 5 mm), 15 lesions	Measurable (> 10 mm), largest lesions	Measurable (> 10 mm), largest lesions; Lymph node ≥ 15mm in short axis	Measurable (> 10 mm), largest lesions; Lymph node ≥ 15 mm in short axis	Measurable (> 10 mm), largest lesions
Number of targets	Five (two per organ)	Five (two per organ)	Five (per organ)	Five (two per organ)	Five (two per organ)	Five (two per organ)	Ten (five injected, five not injected)
Complete Response	Disappearance of all target lesions; Lymph node with short axis < 10 mm	Disappearance of any intra-tumoral arterial enhancement in all target lesions	Disappearance of all target lesions	Disappearance of all target lesions	Disappearance of all target lesions; Lymph node with short axis < 10mm	Disappearance of all target lesions	Disappearance of all target lesions
Partial Response	≥ 30% decrease in sum of maximum diameter of target lesions	≥ 30%decrease in sum of maximum diameter of enhancing target lesions	≥ 50% decrease from the baseline	≥ 30% decrease from the baseline	≥ 30% decrease in sum of maximum diameter of target lesions	≥ 30% decrease from the baseline	≥ 30% decrease for injected lesions, ≥ 30% decrease for not injected lesions
Progressive Disease	≥ 20% increase in sum of diameters and at least 5 mm absolute increase in sum and/or new lesions	≥ 20%increase in sum of diameters of enhancing target lesions and/or new lesions	≥ 25% increase from the nadir	≥ 20% increase from the nadir (≥ 5 mm)	iUPD; iCPD	≥ 20% increase from the nadir (≥ 5 mm)	≥ 20% increase from the nadir (≥ 5 mm)
Confirmation of Progressive Disease	Not applicable	Not applicable	At least 4 weeks	4-12 weeks	4-8 weeks	At least 4 weeks	4-12 weeks

Abbreviations: iCPD, immune-confirmed progressive disease, Confirmed progression with increase in sum of measures ≥ 5 mm; imRECIST, immune-modified RECIST; iRECIST, immune RECIST; irRC, immune-related response criteria; irRECIST, immune-related RECIST; itRECIST, intra-tumoral RECIST; iUPD, immune-unconfirmed progressive disease, ≥ 20% increase in sum of diameters and at least 5 mm absolute increase in sum; mRECIST, modified RECIST; PD, Progressive Disease; RECIST, Response Evaluation Criteria in Solid Tumors.

7. Challenges with NAT for LR in HCC

The potential drawbacks associated with NAT involve significant challenges, as evinced by a phase II clinical trial that evaluated the perioperative efficacy and safety of camrelizumab in combination with apatinib for resectable HCC (*54*). Despite the notable pathological response observed in resected specimens, a substantial proportion of patients completing NAT - 89% (16/18) - experienced AEs. Of particular concern, 16.7% (3/18) of patients experienced grade 3 or higher AEs, necessitating a dose adjustment of apatinib in 5.6% (1/18) of patients due to high blood pressure. Moreover, 11.1% (2/18) of patients required preoperative steroid therapy to deal with severe liver dysfunction or a severe rash. There were additional challenges preoperatively, with 38.8% (7/18) of patients experiencing post-hepatectomy grade A liver failure, 16.7% (3/18) developing postoperative bile leakage, 11.1% (2/18) requiring blood transfusions, and 5.6% (1/18) reporting chest tightness. These findings underscore the intricate balance between efficacy and potential complications associated with NAT, emphasizing the need for comprehensive efforts through expanded clinical research (as shown in Figure 4).

For early-stage HCC, the efficacy of NAT in improving patient survival and reducing cancer recurrence remains uncertain. There are concerns regarding the risks of tumor progression during NAT and the potential for delayed curative surgery due to AEs during treatment. Neoadjuvant immunotherapy in particular carries the possibility of reactivating the hepatitis B virus in patients (*105*). Addressing the challenges associated with potential delayed tumor responses and the chemotherapy-free interval (CFI) during NAT is paramount. Moreover, clinical predictors to distinguish patients who may derive optimal benefits from NAT need to be promptly identified. The absence of a standardized definition for MPR in HCC adds complexity, and its prognostic significance remains unclear. The lack of validated biomarkers predicting surgical success further contributes to the existing challenges. There is still considerable heterogeneity in the selection of a treatment plan among different cancer centers. Additional clinical evidence is needed to guide decisions on whether patients who underwent NAT should proceed to immediate surgery upon disease progression or opt for delayed surgery.

8. Prospects for NAT

In the evolving landscape of clinical trials, advances in research are gradually revealing more efficacious NAT options for HCC. Treatment decisions for patients with HCC should be collaboratively determined through a multidisciplinary team (MDT) approach involving surgery, oncology, radiation therapy, pathology, interventional radiology, and other specialties. This ensures the formulation of optimal treatment strategies and enhances overall patient survival rates. NAT plays a pivotal role in bolstering local control, targeting latent micrometastases in the early stages of the disease, facilitating preoperative recovery, and enhancing the probability of completing multimodal treatment. The assessment of response post-NAT furnishes valuable

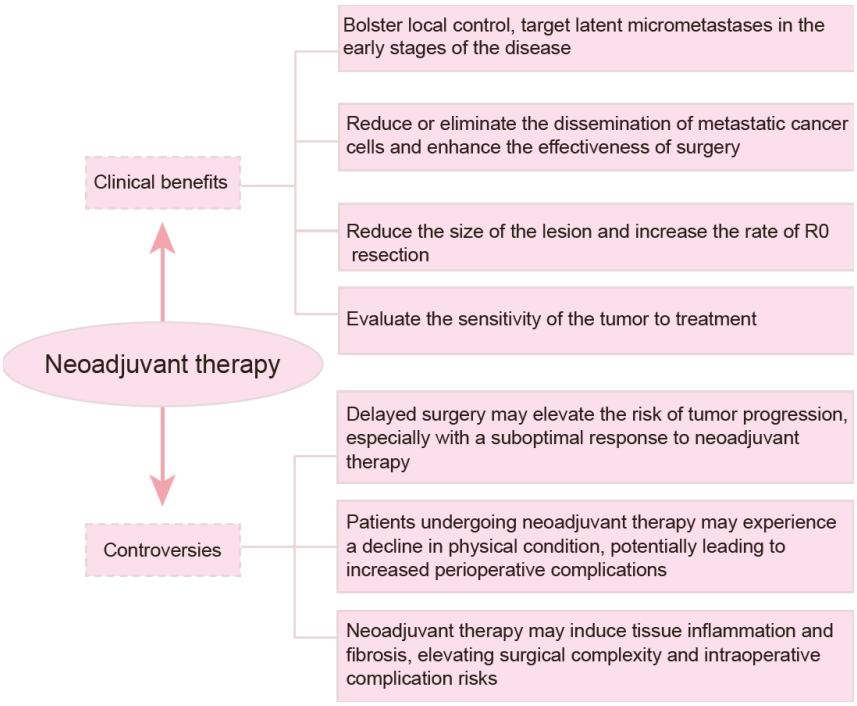

Figure 4. Clinical benefits and controversies of neoadjuvant therapy in hepatocellular carcinoma (HCC).

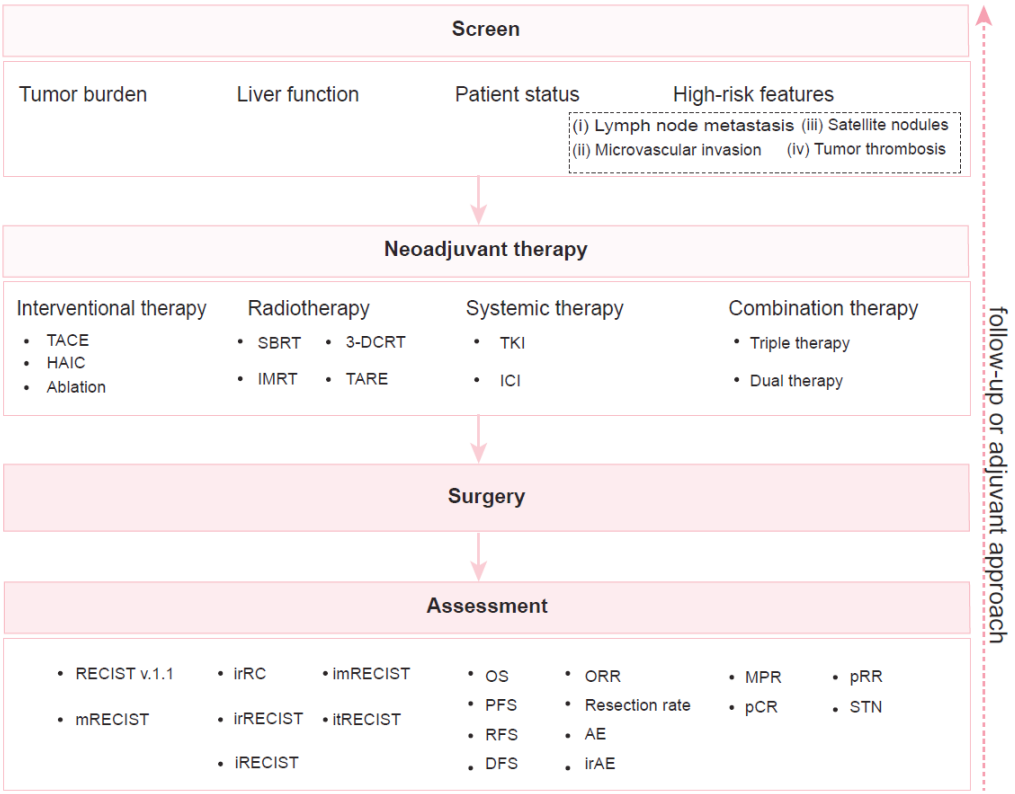

Figure 5: Recommended paradigm for neoadjuvant therapy in hepatocellular carcinoma (HCC). A comprehensive assessment, incorporating patient tumor burden, hepatic function, patient status, and high-risk features, is advocated for the identification of resectable hepatocellular carcinoma patients who may benefit from neoadjuvant therapy (NAT). NAT encompasses interventional, radiation, systemic, and combination modalities. Subsequent to the completion of neoadjuvant treatment, the next step involves surgical intervention. Postoperative evaluation should include a holistic approach, integrating imaging studies, biomarkers, pathological response, disease status, and adverse events. A thorough assessment should be performed and a subsequent adjuvant therapeutic strategy should be formulated in a multidisciplinary collaborative framework, followed by diligent follow-up.

insights into planning subsequent treatments. Preliminary outcomes show promise, but further research is needed to delineate the optimal duration of treatment, to validate pertinent endpoints, and to identify biomarkers that can adeptly help to decide treatments.

In summary, NAT for HCC has significant advantages in improving pCR, MPR, ORR, DFS, and OS. NAT for HCC represents a paradigm shift in the treatment of HCC (as shown in Figure 5), requiring multidisciplinary collaboration for assessing disease and deciding treatment. In addition, the incorporation of immunotherapy in NAT poses new challenges regarding endpoints of radiological, pathological, and clinical research. Therefore, further research is essential to enhancing treatment options guided by biomarkers, to determining the optimal duration of treatment, and to ultimately improving survival time and the quality of life for patients with HCC.

Funding: This work was supported by a grant from the National Natural Science Foundation of China (no. 81960446), Innovative Research Projects for Graduate Students of Hainan Medical College, and the 2023 Foreign Experts Program in Hainan Province (SQ2023WGZJ0002).

Conflict of Interest: The authors have no conflicts of interest to disclose.

References

1. Villanueva A. Hepatocellular Carcinoma. N Engl J Med. 2019; 380:1450-1462.
2. Llovet JM, Zucman-Rossi J, Pikarsky E, Sangro B, Schwartz M, Sherman M, Gores G. Hepatocellular carcinoma. Nat Rev Dis Primers. 2016; 2:16018.
3. Organization WH. Estimated age-standardized incidence rates (World) in 2020, liver, both sexes, all ages. 2020.
4. Chan AWH, Zhong J, Berhane S, *et al.* Development of pre and post-operative models to predict early recurrence of hepatocellular carcinoma after surgical resection. J Hepatol. 2018; 69:1284-1293.
5. Wakayama K, Kamiyama T, Yokoo H, Orimo T, Shimada S, Einama T, Kamachi H, Taketomi A. Huge hepatocellular carcinoma greater than 10 cm in diameter worsens prognosis by causing distant recurrence after curative resection. J Surg Oncol. 2017; 115:324-329.
6. Liu L, Shui Y, Yu Q, Guo Y, Zhang L, Zhou X, Yu R, Lou J, Wei S, Wei Q. Narrow-margin hepatectomy resulted in higher recurrence and lower overall survival for R0 resection hepatocellular carcinoma. Front Oncol. 2020; 10:610636.
7. Erstad DJ, Tanabe KK. Prognostic and therapeutic

implications of microvascular invasion in hepatocellular carcinoma. Ann Surg Oncol. 2019; 26:1474-1493.
8. European Association for the Study of the Liver. Electronic address eee, European Association for the Study of the L. EASL clinical practice guidelines: Management of hepatocellular carcinoma. J Hepatol. 2018; 69:182-236.
9. Reig M, Forner A, Rimola J, *et al.* BCLC strategy for prognosis prediction and treatment recommendation: The 2022 update. J Hepatol. 2022; 76:681-693.
10. Singal AG, Llovet JM, Yarchoan M, *et al*. AASLD practice guidance on prevention, diagnosis, and treatment of hepatocellular carcinoma. Hepatology. 2023; 78:1922-1965.
11. NCCN clinical practice guidelines in oncology (NCCN Guidelines®): Hepatocellular carcinoma (Version 1.2023). *https://medfind.link/wp-content/uploads/2023/07/%E8%82%9D%E7%BB%86%E8%83%9E%E7%99%8C_2023.V1_EN.pdf* (accessd August 19, 2023)
12. Guidelines for the diagnosis and treatment of primary HCC in China. *http://www.nhc.gov.cn/yzygj/s7659/202201/a01ceb75c62b486fa459e36ba0fdfdbc/files/e8b02c99ab2d4ebea07a4c636eace9c9.pdf (in Chinese)* (accessd August 19, 2023).
13. Kudo M, Kawamura Y, Hasegawa K, Tateishi R, Kariyama K, Shiina S, Toyoda H, Imai Y, Hiraoka A, Ikeda M. Management of hepatocellular carcinoma in Japan: JSH consensus statements and recommendations 2021 update. Liver cancer. 2021; 10:181-223.
14. Association KLC. 2022 KLCA-NCC Korea practice guidelines for the management of hepatocellular carcinoma. Korean Journal of Radiology. 2022; 23:1126.
15. Kumar A, Acharya SK, Singh SP, *et al*. 2023 update of Indian national association for study of the liver consensus on management of intermediate and advanced hepatocellular carcinoma: The Puri III recommendations. J Clin Exp Hepatol. 2024; 14:101269.
16. Singal AG, Kudo M, Bruix J. Breakthroughs in hepatocellular carcinoma therapies. Clin Gastroenterol Hepatol. 2023; 21:2135-2149.
17. Bruix J, Takayama T, Mazzaferro V, *et al*. Adjuvant sorafenib for hepatocellular carcinoma after resection or ablation (STORM): A phase 3, randomised, double-blind, placebo-controlled trial. Lancet Oncol. 2015; 16:1344-1354.
18. Lin H, Li X, Liu Y, Hu Y. Neoadjuvant radiotherapy provided survival benefit compared to adjuvant radiotherapy for hepatocellular carcinoma. ANZ J Surg. 2018; 88:E718-E724.
19. Galle PR, Forner A, Llovet JM, Mazzaferro V, Piscaglia F, Raoul J-L, Schirmacher P, Vilgrain V. EASL clinical practice guidelines: Management of hepatocellular carcinoma. J Hepatol. 2018; 69:182-236.
20. Yoh T, Ishii T, Nishio T, Koyama Y, Ogiso S, Fukumitsu K, Uchida Y, Ito T, Seo S, Hata K, Hatano E. A conceptual classification of resectability for hepatocellular carcinoma. World J Surg. 2023; 47:740-748.
21. Iguchi K, Hatano E, Yamanaka K, Tanaka S, Taura K, Uemoto S. Validation of the conventional resection criteria in patients with hepatocellular carcinoma in terms of the incidence of posthepatectomy liver failure and long-term prognosis. Dig Surg. 2015; 32:344-351.
22. Kobayashi Y, Kiya Y, Sugawara T, Nishioka Y, Hashimoto M, Shindoh J. Expanded Makuuchi's criteria using estimated indocyanine green clearance rate of future liver remnant as a safety limit for maximum extent of liver resection. HPB (Oxford). 2019; 21:990-997.
23. Costentin CE, Ferrone CR, Arellano RS, Ganguli S, Hong TS, Zhu AX. Hepatocellular carcinoma with macrovascular invasion: Defining the optimal treatment strategy. Liver Cancer. 2017; 6:360-374.
24. European Association For The Study Of The Liver; European Organisation For Research And Treatment Of Cancer. EASL-EORTC clinical practice guidelines: Management of hepatocellular carcinoma. J Hepatol. 2012; 56:908-943.
25. Sun HC, Zhou J, Wang Z, *et al.* Chinese expert consensus on conversion therapy for hepatocellular carcinoma (2021 edition). Hepatobiliary Surg Nutr. 2022; 11:227-252.
26. Li S, Zhong C, Li Q, *et al.* Neoadjuvant transarterial infusion chemotherapy with FOLFOX could improve outcomes of resectable BCLC stage A/B hepatocellular carcinoma patients beyond Milan criteria: An interim analysis of a multi-center, phase 3, randomized, controlled clinical trial. Journal of Clinical Oncology. 2021; 39:4008-4008.
27. Wei W, Li S, Zhao R, *et al.* Neoadjuvant hepatic arterial infusion chemotherapy with FOLFOX could improve outcomes of resectable BCLC stage A/B hepatocellular carcinoma patients beyond Milan criteria: A multi-center, phase 3, randomized, controlled clinical trial. Journal of Clinical Oncology. 2023; 41:4023-4023.
28. Wu JY, Yin ZY, Bai YN, Chen YF, Zhou SQ, Wang SJ, Zhou JY, Li YN, Qiu FN, Li B, Yan ML. Lenvatinib combined with anti-PD-1 antibodies plus transcatheter arterial chemoembolization for unresectable hepatocellular carcinoma: A multicenter retrospective study. J Hepatocell Carcinoma. 2021; 8:1233-1240.
29. Xia YX, Zhang H, Zhang F, *et al.* Efficacy and safety of neoadjuvant immunotherapy for hepatocellular carcinoma. Chinese Journal of Surgery (in Chinese). 2022; 60:688-694.
30. Li C, Wang MD, Lu L, *et al.* Preoperative transcatheter arterial chemoembolization for surgical resection of huge hepatocellular carcinoma (>/= 10 cm): A multicenter propensity matching analysis. Hepatol Int. 2019; 13:736-747.
31. Wei X, Jiang Y, Zhang X, *et al*. Neoadjuvant three-dimensional conformal radiotherapy for resectable hepatocellular carcinoma with portal vein tumor thrombus: A randomized, open-label, multicenter controlled study. J Clin Oncol. 2019; 37:2141-2151.
32. Hu Z, Yang Z, Wang J, Fu Y, Hu Z, Zhou Z, Chen M, Zhang Y. Survival benefit of neoadjuvant hepatic arterial infusion chemotherapy followed by hepatectomy for hepatocellular carcinoma with portal vein tumor thrombus. Front Pharmacol. 2023; 14:1223632.
33. Monden M, Okamura J, Sakon M, Gotoh M, Kobayashi K, Umeshita K, Yamada T, Kuroda C, Sakurai M, Mori T. Significance of transcatheter chemoembolization combined with surgical resection for hepatocellular carcinomas. Cancer Chemother Pharmacol. 1989; 23 Suppl:S90-S95.
34. Zhang Z, Liu Q, He J, Yang J, Yang G, Wu M. The effect of preoperative transcatheter hepatic arterial chemoembolization on disease-free survival after hepatectomy for hepatocellular carcinoma. Cancer. 2000; 89:2606-2612.
35. Zhou WP, Lai EC, Li AJ, Fu SY, Zhou JP, Pan ZY, Lau WY, Wu MC. A prospective, randomized, controlled trial of preoperative transarterial chemoembolization for

resectable large hepatocellular carcinoma. Ann Surg. 2009; 249:195-202.
36. Kang JY, Choi MS, Kim SJ, Kil JS, Lee JH, Koh KC, Paik SW, Yoo BC. Long-term outcome of preoperative transarterial chemoembolization and hepatic resection in patients with hepatocellular carcinoma. Korean J Hepatol. 2010; 16:383-388.
37. Jianyong L, Jinjing Z, Wentao W, Lunan Y, Qiao Z, Bo L, Tianfu W, Minqing X, Jiaying Y, Yongang W. Preoperative transcatheter arterial chemoembolization for resectable hepatocellular carcinoma: a single center analysis. Ann Hepatol. 2014; 13:394-402.
38. Edeline J, Lenoir L, Boudjema K, Rolland Y, Boulic A, Le Du F, Pracht M, Raoul JL, Clement B, Garin E, Boucher E. Volumetric changes after (90)y radioembolization for hepatocellular carcinoma in cirrhosis: an option to portal vein embolization in a preoperative setting? Ann Surg Oncol. 2013; 20:2518-2525.
39. Hoang M, Chow PK. Downstaging locally advanced hepatocellular carcinoma with selective internal radiation therapy. American Society of Clinical Oncology, 2023: 4_ suppl:536.
40. Raoul JL, Messner M, Boucher E, Bretagne JF, Campion JP, Boudjema K. Preoperative treatment of hepatocellular carcinoma with intra-arterial injection of 131I-labelled lipiodol. Br J Surg. 2003; 90:1379-1383.
41. Li N, Feng S, Xue J, Wei XB, Shi J, Guo WX, Lau WY, Wu MC, Cheng SQ, Meng Y. Hepatocellular carcinoma with main portal vein tumor thrombus: A comparative study comparing hepatectomy with or without neoadjuvant radiotherapy. HPB (Oxford). 2016; 18:549-556.
42. Lemaire M, Lucidi V, Bouazza F, *et al.* Selective internal radiation therapy (SIRT) before partial hepatectomy or radiofrequency destruction for treatment of hepatocellular carcinoma in cirrhotic patients: A feasibility and safety pilot study. HPB (Oxford). 2018; 20:641-648.
43. Craciun L, de Wind R, Demetter P, *et al.* Retrospective analysis of the immunogenic effects of intra-arterial locoregional therapies in hepatocellular carcinoma: a rationale for combining selective internal radiation therapy (SIRT) and immunotherapy. BMC Cancer. 2020; 20:135.
44. Luo Z, Che X, Cai J, Zhao H, Jin J, Tang Y, Chen B, Bi X. Neoadjuvant radiotherapy to improve overall survival in resectable hepatocellular carcinoma. Journal of Clinical Oncology. 2021; 39:e16178-e16178.
45. Wu F, Chen B, Dong D, *et al.* Phase 2 evaluation of neoadjuvant intensity-modulated radiotherapy in centrally located hepatocellular carcinoma: A nonrandomized controlled trial. JAMA Surg. 2022; 157:1089-1096.
46. Bouattour M, Fartoux L, Rosmorduc O, Scatton O, Vibert E, Costentin C, Soubrane O, Ronot M, Granier MM, De Gramont A. BIOSHARE multicenter neoadjuvant phase 2 study: Results of pre-operative sorafenib in patients with resectable hepatocellular carcinoma (HCC)—From GERCOR IRC. American Society of Clinical Oncology, 2016; 34:252.
47. Marron TU, Fiel MI, Hamon P, *et al.* Neoadjuvant cemiplimab for resectable hepatocellular carcinoma: a single-arm, open-label, phase 2 trial. Lancet Gastroenterol Hepatol. 2022; 7:219-229.
48. Shi Y-H, Ji Y, Liu W-R, Pang Y-R, Ding Z-B, Fu X-T, Zhang X, Huang C, Sun Y-F, Zhu X-D, Sun H-C, Zhou J, Fan J. Abstract 486: A phase Ib/II, open-label study evaluating the efficacy and safety of toripalimab injection (JS001) or combination with lenvatinib as a neoadjuvant therapy for patients with resectable hepatocellular carcinoma (HCC). Cancer Research. 2021; 81:486-486.
49. Kaseb AO, Hasanov E, Cao HST, *et al.* Perioperative nivolumab monotherapy versus nivolumab plus ipilimumab in resectable hepatocellular carcinoma: A randomised, open-label, phase 2 trial. Lancet Gastroenterol Hepatol. 2022; 7:208-218.
50. Song T. A prospective, single-arm, phase II clinical study of tislelizumab in combination with lenvatinib for perioperative treatment of resectable primary hepatocellular carcinoma with high risk of recurrence. Journal of Clinical Oncology. 2023; 41:e16218-e16218.
51. Ho WJ, Zhu Q, Durham J, *et al.* Neoadjuvant cabozantinib and nivolumab converts locally advanced HCC into resectable disease with enhanced antitumor immunity. Nat Cancer. 2021; 2:891-903.
52. Cui Y, Bao X, Yu G, Li H, Fang F, Li Q, Zhang W, Wu Q, Chen L, Liu C, Song T. Camrelizumab in combination with apatinib as a perioperative treatment for patients with hepatocellular carcinoma at high risk of recurrence: A prospective, single-arm, phase 2 study. Journal of Clinical Oncology. 2023; 41:4120-4120.
53. Zhou J, Fan J, Gu F-M, *et al.* A phase II/III study of camrelizumab plus apatinib as perioperative treatment of resectable hepatocellular carcinoma at intermediate-high risk of recurrence: Primary results of major pathologic response from phase II stage. Journal of Clinical Oncology. 2023; 41:4126-4126.
54. Xia Y, Tang W, Qian X, *et al.* Efficacy and safety of camrelizumab plus apatinib during the perioperative period in resectable hepatocellular carcinoma: A single-arm, open label, phase II clinical trial. J Immunother Cancer. 2022; 10:e004656.
55. Pinato DJ, Cortellini A, Sukumaran A, *et al.* PRIME-HCC: Phase Ib study of neoadjuvant ipilimumab and nivolumab prior to liver resection for hepatocellular carcinoma. BMC Cancer. 2021; 21:301.
56. Kaseb A, Duda DG, Tran Cao HS, *et al.* LBA47 - Randomized, open-label, perioperative phase II study evaluating nivolumab alone versus nivolumab plus ipilimumab in patients with resectable HCC. Annals of Oncology. 2019; 30:v880.
57. Tamaki N, Tada T, Kurosaki M, *et al.* Optimal threshold of alpha-fetoprotein response in patients with unresectable hepatocellular carcinoma treated with atezolizumab and bevacizumab. Invest New Drugs. 2022; 40:1290-1297.
58. Zhu AX, Dayyani F, Yen CJ, Ren Z, Bai Y, Meng Z, Pan H, Dillon P, Mhatre SK, Gaillard VE, Hernandez S, Kelley RK, Sangro B. Alpha-fetoprotein as a potential surrogate biomarker for atezolizumab + bevacizumab treatment of hepatocellular carcinoma. Clin Cancer Res. 2022; 28:3537-3545.
59. Campani C, Bamba-Funck J, Campion B, *et al.* Baseline ALBI score and early variation of serum AFP predicts outcomes in patients with HCC treated by atezolizumab-bevacizumab. Liver Int. 2023; 43:708-717.
60. Kelley RK, Meyer T, Rimassa L, *et al.* Serum alpha-fetoprotein levels and clinical outcomes in the phase III CELESTIAL study of cabozantinib versus placebo in patients with advanced hepatocellular carcinoma. Clin Cancer Res. 2020; 26:4795-4804.
61. Zhu AX, Finn RS, Kang YK, *et al.* Serum alpha-fetoprotein and clinical outcomes in patients with advanced hepatocellular carcinoma treated with ramucirumab. Br J Cancer. 2021; 124:1388-1397.

62. Saeki I, Yamasaki T, Yamashita S, *et al.* Early predictors of objective response in patients with hepatocellular carcinoma undergoing lenvatinib treatment. Cancers (Basel). 2020; 12:779.
63. Hatanaka T, Kakizaki S, Hiraoka A, *et al.* Prognostic impact of c-reactive protein and alpha-fetoprotein in immunotherapy score in hepatocellular carcinoma patients treated with atezolizumab plus bevacizumab: A multicenter retrospective study. Hepatol Int. 2022; 16:1150-1160.
64. Scheiner B, Pomej K, Kirstein MM, *et al.* Prognosis of patients with hepatocellular carcinoma treated with immunotherapy - development and validation of the CRAFITY score. J Hepatol. 2022; 76:353-363.
65. Hatanaka T, Kakizaki S, Hiraoka A, *et al.* Development and validation of a modified albumin-bilirubin grade and alpha-fetoprotein score (mALF score) for hepatocellular carcinoma patients receiving atezolizumab and bevacizumab. Hepatol Int. 2023; 17:86-96.
66. Kuzuya T, Kawabe N, Hashimoto S, *et al.* Early changes in alpha-fetoprotein are a useful predictor of efficacy of atezolizumab plus bevacizumab treatment in patients with advanced hepatocellular carcinoma. Oncology. 2022; 100:12-21.
67. Miyaki D, Kawaoka T, Aikata H, *et al.* Evaluation of early response to hepatic arterial infusion chemotherapy in patients with advanced hepatocellular carcinoma using the combination of response evaluation criteria in solid tumors and tumor markers. J Gastroenterol Hepatol. 2015; 30:726-732.
68. Feun LG, Li YY, Wu C, Wangpaichitr M, Jones PD, Richman SP, Madrazo B, Kwon D, Garcia-Buitrago M, Martin P, Hosein PJ, Savaraj N. Phase 2 study of pembrolizumab and circulating biomarkers to predict anticancer response in advanced, unresectable hepatocellular carcinoma. Cancer. 2019; 125:3603-3614.
69. Yang H, Kang B, Ha Y, Lee SH, Kim I, Kim H, Lee WS, Kim G, Jung S, Rha SY, Gaillard VE, Cheon J, Kim C, Chon HJ. High serum IL-6 correlates with reduced clinical benefit of atezolizumab and bevacizumab in unresectable hepatocellular carcinoma. JHEP Rep. 2023; 5:100672.
70. Myojin Y, Kodama T, Sakamori R, *et al.* Interleukin-6 is a circulating prognostic biomarker for hepatocellular carcinoma patients treated with combined immunotherapy. Cancers (Basel). 2022; 14:883.
71. Hattori E, Okumoto K, Adachi T, Takeda T, Ito J, Sugahara K, Watanabe H, Saito K, Saito T, Togashi H, Kawata S. Possible contribution of circulating interleukin-10 (IL-10) to anti-tumor immunity and prognosis in patients with unresectable hepatocellular carcinoma. Hepatol Res. 2003; 27:309-314.
72. Chan SL, Mo FK, Wong CS, Chan CM, Leung LK, Hui EP, Ma BB, Chan AT, Mok TS, Yeo W. A study of circulating interleukin 10 in prognostication of unresectable hepatocellular carcinoma. Cancer. 2012; 118:3984-3992.
73. Ocal O, Schutte K, Kupcinskas J, *et al.* Baseline interleukin-6 and -8 predict response and survival in patients with advanced hepatocellular carcinoma treated with sorafenib monotherapy: An exploratory post hoc analysis of the SORAMIC trial. J Cancer Res Clin Oncol. 2022; 148:475-485.
74. Iida-Ueno A, Enomoto M, Uchida-Kobayashi S, Hagihara A, Teranishi Y, Fujii H, Morikawa H, Murakami Y, Tamori A, Thuy LTT, Kawada N. Changes in plasma interleukin-8 and tumor necrosis factor-alpha levels during the early treatment period as a predictor of the response to sorafenib in patients with unresectable hepatocellular carcinoma. Cancer Chemother Pharmacol. 2018; 82:857-864.
75. Poon RT, Lau C, Pang R, Ng KK, Yuen J, Fan ST. High serum vascular endothelial growth factor levels predict poor prognosis after radiofrequency ablation of hepatocellular carcinoma: Importance of tumor biomarker in ablative therapies. Ann Surg Oncol. 2007; 14:1835-1845.
76. Hong YM, Cho M, Yoon KT, Ryu JH, Yang KH, Jeon UB, Hwang TH. Neutrophil-lymphocyte ratio predicts the therapeutic benefit of neoadjuvant transarterial chemoembolization in patients with resectable hepatocellular carcinoma. Eur J Gastroenterol Hepatol. 2020; 32:1186-1191.
77. Eso Y, Takeda H, Taura K, Takai A, Takahashi K, Seno H. Pretreatment neutrophil-to-lymphocyte ratio as a predictive marker of response to atezolizumab plus bevacizumab for hepatocellular carcinoma. Curr Oncol. 2021; 28:4157-4166.
78. Tada T, Kumada T, Hiraoka A, *et al.* Neutrophil-lymphocyte ratio predicts early outcomes in patients with unresectable hepatocellular carcinoma treated with atezolizumab plus bevacizumab: A multicenter analysis. Eur J Gastroenterol Hepatol. 2022; 34:698-706.
79. Matoya S, Suzuki T, Matsuura K, *et al.* The neutrophil-to-lymphocyte ratio at the start of the second course during atezolizumab plus bevacizumab therapy predicts therapeutic efficacy in patients with advanced hepatocellular carcinoma: A multicenter analysis. Hepatol Res. 2023; 53:511-521.
80. Maesaka K, Sakamori R, Yamada R, *et al.* Hyperprogressive disease in patients with unresectable hepatocellular carcinoma receiving atezolizumab plus bevacizumab therapy. Hepatol Res. 2022; 52:298-307.
81. Jost-Brinkmann F, Demir M, Wree A, Luedde T, Loosen SH, Muller T, Tacke F, Roderburg C, Mohr R. Atezolizumab plus bevacizumab in unresectable hepatocellular carcinoma: Results from a German real-world cohort. Aliment Pharmacol Ther. 2023; 57:1313-1325.
82. Wu YL, Fulgenzi CAM, D'Alessio A, *et al.* Neutrophil-to-lymphocyte and platelet-to-lymphocyte ratios as prognostic biomarkers in unresectable hepatocellular carcinoma treated with atezolizumab plus bevacizumab. Cancers (Basel). 2022; 14:5834.
83. Hosoda S, Suda G, Sho T, *et al.* Low baseline CXCL9 predicts early progressive disease in unresectable HCC with atezolizumab plus bevacizumab treatment. Liver Cancer. 2023; 12:156-170.
84. Sun Y, Wu L, Zhong Y, *et al.* Single-cell landscape of the ecosystem in early-relapse hepatocellular carcinoma. Cell. 2021; 184:404-421.e16.
85. Park DJ, Sung PS, Lee GW, *et al.* Preferential expression of programmed death ligand 1 protein in tumor-associated macrophages and its potential role in immunotherapy for hepatocellular carcinoma. Int J Mol Sci. 2021; 22:4710.
86. Han JW, Kim JH, Kim DH, Jang JW, Bae SH, Choi JY, Yoon SK, Ahn J, Yang H, Sung PS. Higher number of tumor-infiltrating PD-L1+ cells is related to better response to multikinase inhibitors in hepatocellular carcinoma. Diagnostics (Basel). 2023; 13:1453.
87. Agdashian D, ElGindi M, Xie C, *et al.* The effect of anti-CTLA4 treatment on peripheral and intra-tumoral T

cells in patients with hepatocellular carcinoma. Cancer Immunol Immunother. 2019; 68:599-608.
88. Liu Y, Xun Z, Ma K, *et al.* Identification of a tumour immune barrier in the HCC microenvironment that determines the efficacy of immunotherapy. J Hepatol. 2023; 78:770-782.
89. Matsumae T, Kodama T, Myojin Y, *et al.* Circulating cell-free DNA profiling predicts the therapeutic outcome in advanced hepatocellular carcinoma patients treated with combination immunotherapy. Cancers (Basel). 2022; 14:3367.
90. Muraoka M, Maekawa S, Katoh R, *et al.* Usefulness of cell-free human telomerase reverse transcriptase mutant DNA quantification in blood for predicting hepatocellular carcinoma treatment efficacy. Hepatol Commun. 2021; 5:1927-1938.
91. Winograd P, Hou S, Court CM, *et al.* Hepatocellular carcinoma-circulating tumor cells expressing PD-L1 are prognostic and potentially associated with response to checkpoint inhibitors. Hepatol Commun. 2020; 4:1527-1540.
92. Tan Z, Yue C, Ji S, Zhao C, Jia R, Zhang Y, Liu R, Li D, Yu Q, Li P, Hu Z, Yang Y, Xu J. Assessment of PD-L1 expression on circulating tumor cells for predicting clinical outcomes in patients with cancer receiving PD-1/PD-L1 blockade therapies. Oncologist. 2021; 26:e2227-e2238.
93. Su K, Guo L, He K, *et al.* PD-L1 expression on circulating tumor cells can be a predictive biomarker to PD-1 inhibitors combined with radiotherapy and antiangiogenic therapy in advanced hepatocellular carcinoma. Front Oncol. 2022; 12:873830.
94. Cannella R, Lewis S, da Fonseca L, Ronot M, Rimola J. Immunotherapy-based treatments of hepatocellular carcinoma: AJR expert panel narrative review. AJR Am J Roentgenol. 2022; 219:533-546.
95. Eisenhauer EA, Therasse P, Bogaerts J, *et al.* New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). Eur J Cancer. 2009; 45:228-247.
96. Lencioni R, Llovet JM. Modified RECIST (mRECIST) assessment for hepatocellular carcinoma. Semin Liver Dis. 2010; 30:52-60.
97. Gregory J, Dioguardi Burgio M, Corrias G, Vilgrain V, Ronot M. Evaluation of liver tumour response by imaging. JHEP Rep. 2020; 2:100100.
98. Gordic S, Corcuera-Solano I, Stueck A, Besa C, Argiriadi P, Guniganti P, King M, Kihira S, Babb J, Thung S, Taouli B. Evaluation of HCC response to locoregional therapy: Validation of MRI-based response criteria versus explant pathology. J Hepatol. 2017; 67:1213-1221.
99. Wolchok JD, Hoos A, O'Day S, Weber JS, Hamid O, Lebbe C, Maio M, Binder M, Bohnsack O, Nichol G, Humphrey R, Hodi FS. Guidelines for the evaluation of immune therapy activity in solid tumors: Immune-related response criteria. Clin Cancer Res. 2009; 15:7412-7420.
100. Nishino M, Giobbie-Hurder A, Gargano M, Suda M, Ramaiya NH, Hodi FS. Developing a common language for tumor response to immunotherapy: Immune-related response criteria using unidimensional measurements. Clin Cancer Res. 2013; 19:3936-3943.
101. Seymour L, Bogaerts J, Perrone A, *et al.* iRECIST: Guidelines for response criteria for use in trials testing immunotherapeutics. Lancet Oncol. 2017; 18:e143-e152.
102. Hodi FS, Ballinger M, Lyons B, *et al.* Immune-modified response evaluation criteria in solid tumors (imRECIST): Refining guidelines to assess the clinical benefit of cancer immunotherapy. J Clin Oncol. 2018; 36:850-858.
103. Goldmacher GV, Khilnani AD, Andtbacka RHI, Luke JJ, Hodi FS, Marabelle A, Harrington K, Perrone A, Tse A, Madoff DC, Schwartz LH. Response criteria for intratumoral immunotherapy in solid tumors: itRECIST. J Clin Oncol. 2020; 38:2667-2676.
104. Kudo M, Aoki T, Ueshima K, *et al.* Achievement of complete response and drug-free status by atezolizumab plus bevacizumab combined with or without curative conversion in patients with transarterial chemoembolization-unsuitable, intermediate-stage hepatocellular carcinoma: A multicenter proof-of-concept study. Liver Cancer. 2023; 12:321-338.
105. Burns EA, Muhsen IN, Anand K, Xu J, Umoru G, Arain AN, Abdelrahim M. Hepatitis B virus reactivation in cancer patients treated with immune checkpoint inhibitors. J Immunother. 2021; 44:132-139.

Received December 12, 2023; Revised February 8, 2024; Accepted February 14, 2024.

**Address correspondence to:*
Wei Tang, International Health Care Center, National Center for Global Health and Medicine, Tokyo, Japan.
E-mail: politang-tky@umin.ac.jp

Released online in J-STAGE as advance publication February 20, 2024.

Review

DOI: 10.5582/bst.2023.01326

Effect of transcatheter arterial chemoembolization combined with lenvatinib plus anti–PD-1 antibodies in patients with unresectable hepatocellular carcinoma: A treatment with Chinese characteristics

Jiayi Wu[1,2], Junyi Wu[1,2], Shuqun Li[1,3], Mengchao Luo[1,4], Zhenxin Zeng[1], Yinan Li[1], Yangkai Fu[1], Han Li[1], Deyi Liu[1], Xiangye Ou[1], Zhongtai Lin[1,4], Shaoming Wei[1,4], Maolin Yan[1,2,*]

[1] Shengli Clinical Medical College of Fujian Medical University, Fuzhou, Fujian, China;
[2] Department of Hepatobiliary Pancreatic Surgery, Fujian Provincial Hospital, Fuzhou, Fujian, China;
[3] Department of Hepatobiliary Pancreatic Surgery, Affiliated Hospital of Guilin Medical University, Guilin, Guangxi, China;
[4] Department of General Surgery, Fujian Provincial Hospital, Fuzhou, Fujian, China.

SUMMARY Therapies for patients with unresectable hepatocellular carcinoma (uHCC) are currently popular. Current first-line standard-of-care treatments for uHCC are systematic therapies. However, treatments that combine locoregional therapy with systemic therapy are widely accepted in China and have demonstrated high rates of tumor response and conversion to resection with manageable toxicity. A literature review was performed by searching published literature in PubMed and Web of Science up to December 2023 for relevant articles on the use of triple therapy (transarterial chemoembolization combined with lenvatinib and anti–PD-1 antibodies) in uHCC. This review concentrates on the efficacy and safety of triple therapy with Chinese characteristics in patients with uHCC and describes the outcome of conversion surgery, degree of pathological necrosis, and effect prediction. This article will contribute to a comprehensive understanding of the role of triple therapy with Chinese characteristics in patients with uHCC.

Keywords hepatocellular carcinoma (HCC), conversion therapy, transcatheter arterial chemoembolization, lenvatinib, programmed death-1

1. Introduction

Because of the insidious onset of hepatocellular carcinoma (HCC), unresectable HCC (uHCC) accounts for a large proportion of cases (*1,2*). In general, there are two main types of uHCC: surgically and oncologically unresectable (*3*). The definition of surgically uHCC is widely accepted and includes cases where R0 resection cannot be achieved due to extrahepatic metastasis, bilobar tumor locations, main vascular invasion, insufficient residual liver volume, and poor general condition or liver function. However, the definition of oncologically uHCC varies and is controversial; it includes cases that may be technically resectable but have a high risk of recurrence, precluding them from benefitting from surgery. Most references to uHCC usually refer to surgically uHCC.

Recent progress in systematic therapy, the primary treatment for uHCC, and especially the success of the REFLECT and IMBRAVE 150 trials *(4,5*), has greatly improved the treatment of uHCC. The Barcelona Clinic Liver Cancer (BCLC) staging system recommends atezolizumab-bevacizumab/durvalumab-tremelimumab as the first-line standard-of-care treatments for uHCC; if this treatment is not feasible, sorafenib, lenvatinib, or durvalumab is considered (*1*). As shown in Table 1, first-line systemic treatment for uHCC improves prognosis, with a median overall survival (OS) of 6.4–22.1 months and progression-free survival (PFS) of 2.1–7.3 months (*4-12*). However, the outcomes have not been satisfactory.

In China, locoregional therapy (LRT), and especially transcatheter arterial chemoembolization (TACE), plays a critical role in managing patients with uHCC and is widely used for intermediate- and advanced-stage HCC (*13*). LRT combined with systemic therapy has yielded impressive outcomes. The CHANCE 001 study (*14*), a multicenter retrospective matched-cohort study of patients with uHCC from 59 academic hospitals across 22 provinces in China, found that combining TACE with anti–programmed death-(ligand) 1 (anti–PD-[L]1) antibodies and molecular targeted treatments (MTT) significantly improved the objective response rate (ORR),

Table 1. First-line systemic treatment for uHCC

Study (*Ref.*)	Regimen	Number of patients	Main characteristics of patients	mOS (months)	mPFS (months)	ORR (RECIST 1.1)	DCR (RECIST 1.1)	Treatment-related adverse effects grade 3/4
SHARP (*6*)	Sorafenib *vs.* placebo	299 *vs.* 303	BCLC-B/C	10.7 *vs.* 7.9	5.5 *vs.* 2.8	2% *vs.* 1%	43% *vs.* 32%	52% *vs.* 54%
REFLECT (*4*)	Lenvatinib *vs.* Sorafenib	478 *vs.* 476	BCLC-B/C	13.6 *vs.* 12.3	7.3 *vs.* 3.6	18.8% *vs.* 6.5%	72.8% *vs.* 59.0%	57% *vs.* 49%
IMbrave150 (*5*)	Atezolizumab-Bevacizumab *vs.* Sorafenib	336 *vs.* 165	BCLC-A/B/C	19.2 *vs.* 13.4	6.8 *vs.* 4.3	27.3% *vs.* 11.9%	73.6% *vs.* 55.3%	56.5% *vs.* 55.1%
HIMALAYA (*7*)	Tremelimumab-Durvalumab *vs.* Sorafenib	393 *vs.* 389	BCLC-B/C	16.4 *vs.* 13.8	3.8 *vs.* 4.1	20.1% *vs.* 5.1%	60.1% *vs.* 60.7%	50.5% *vs.* 52.4%
HIMALAYA (*7*)	Durvalumab *vs.* Sorafenib	389 *vs.* 389	BCLC-B/C	16.6 *vs.* 13.8	3.7 *vs.* 4.1	17.0% *vs.* 5.1%	54.8% *vs.* 60.7%	37.1% *vs.* 52.4%
EACH (*8*)	FOLFOX4 *vs.* Doxorubicin	184 *vs.* 187	BCLC-B/C	6.40 *vs.* 4.97	2.93 *vs.* 1.77	8.15% *vs.* 2.67%	52.17% *vs.* 31.55%	55.74% *vs.* 45.40%
ZGDH3 (*9*)	Donafenib *vs.* Sorafenib	328 *vs.* 331	BCLC-B/C	12.1 *vs.* 10.3	3.7 *vs.* 3.6	4.6% *vs.* 2.7%	30.8% *vs.* 28.7%	57% *vs.* 67%
RATIONALE-301 (*10*)	Tislelizumab *vs.* Sorafenib	342 *vs.* 332	BCLC-B/C	15.9 *vs.* 14.1	2.1 *vs.* 3.4	14.3% *vs.* 5.4%	44.2% *vs.* 50.3%	22.2% *vs.* 53.4%
CARES-310 (*11*)	Camrelizumab-Rivoceranib *vs.* Sorafenib	272 *vs.* 271	BCLC-B/C	22.1 *vs.* 15.2	5.6 *vs.* 3.7	25% *vs.* 6%	78% *vs.* 54%	81% *vs.* 52%
ORIENT-32 (*12*)	Sintilimab plus IBI305 *vs.* Sorafenib	380 *vs.* 191	BCLC-B/C	NR *vs.* 10.4	4.6 *vs.* 2.8	21% *vs.* 4%	72% *vs.* 64%	53% *vs.* 45%

Abbreviations: OS, overall survival; PFS, progression-free survival; ORR, objective response rate; RECIST, response evaluation criteria in solid tumors; DCR, disease control rate; BCLC, Barcelona clinic liver cancer; NR, not reached.

PFS, and OS compared to TACE alone in patients with predominantly advanced HCC. A systematic review and meta-analysis (*15*) also confirmed that combining MTT with anti–PD-1 antibodies and LRT is an effective conversion therapy regimen with a significant ORR, conversion potential, and satisfactory safety profile.

Because of the heterogeneity of MTT, the current study focused on the triple therapy of TACE combined with lenvatinib (an MTT) plus anti–PD-1 antibodies. Searches on PubMed and Web of Science conducted on December 1, 2023 revealed that all articles on triple therapy were written by Chinese researchers (*16-39*). Therefore, this review aims to explain triple therapy with Chinese characteristics and to examine its role in managing uHCC.

2. Triple therapy in unresectable HCC

In 2021, the first-line efficacy of triple therapy for uHCC was analyzed based on triple therapy's clinical presentation (*16*). The study enrolled 62 patients with initial uHCC from four centers in China: 35, 21, and 6 patients with BCLC stages C, B, and A, respectively. Based on the modified Response Evaluation Criteria in Solid Tumors (mRECIST), the ORR was 80.6% per investigator and 77.4% per blinded independent central review. Twenty-nine patients underwent conversion surgery with a median follow-up time of 12.2 months. A pathological complete response (PCR) and major pathological response (MPR) were achieved in 16 and 24 patients, respectively. Because of the relatively short follow-up time, the median PFS and OS times were not reached.

As shown in Table 2 (Online Data: *http://www.biosciencetrends.com/action/getSupplementalData.php?ID=185*), triple therapy resulted in an ORR of 26.1–87.2%, disease control rate (DCR) of 70–100%, median PFS of 6.3–22.5 months, and median OS of 15.7–29 months. Despite a lack of final results from randomized controlled phase III trials, triple therapy was found to be effective, with a median OS comparable to that of current first-line treatment regimens.

Triple therapy in uHCC with portal vein tumor thrombosis. Portal vein tumor thrombosis (PVTT) is a dismal prognostic factor for HCC, with a median survival period of 2.7–4.0 months without treatment (*40*). Despite the short survival, triple therapy's effectiveness was able to be determined in patients with HCC and main trunk PVTT. Our retrospective study (*37*) enrolled 41 patients with main trunk PVTT who received triple therapy as the first-line therapy. The intrahepatic tumor ORR was 68.3% (5 complete responses [CR] and 23 partial responses [PR]) per mRECIST. PVTT was considered to have regressed in 8 patients, and 4 patients had complete necrosis. After a median follow-up of 18 months, the median PFS was 14.5 (range 1.3–27.6) months, and the median OS was 21.7 (range 2.8–30.5) months; 12

patients (29.3%) underwent conversion surgery. Of the 12 patients, three had an intrahepatic tumor PCR and seven had a PVTT PCR as determined by a pathological examination of the resected specimen.

Two studies (Zou *et al*. and Li *et al*.) analyzed triple therapy's safety and clinical efficacy in patients with uHCC and PVTT. In the study by Zou *et al*. (*36*), patients with uHCC and PVTT (53.75% PVTT type I, 46.25% type II/III/IV, per Cheng's classification) after triple therapy had a median OS of 21.7 months and a PFS of 6.3 months. The multicenter prospective study by Li *et al* (39). enrolled 69 patients with uHCC and PVTT (13% PVTT type I, 87% PVTT type II/III/IV, per Cheng's classification). After a median follow-up of 17.3 months, the ORR was 26.1%, and the DCR was 78.3% per mRECIST. The median PFS and OS were 9.3 and 18.2 months, respectively.

Although patients with HCC and PVTT have poor prognoses, promising results are obtained after triple therapy.

3. Triple therapy in conversion surgery

Although participants' baseline characteristics and the definition of conversion to resectable HCC varied among studies, conversion rates were 25–50%, based on the good ORR performance of triple therapy (*16-19*).

A meta-analysis (*15*) evaluating the efficacy and safety of different conversion regimens found that combining LRT and MTT plus anti–PD-1 antibodies resulted in a significantly greater conversion rate (33%, 95% confidence interval [CI] 17–52%) than combinations of LRT and MTT without anti–PD-1 antibodies (12%, 95% CI: 8–17%; $P = 0.01$).

The prognosis after conversion surgery was also a topic worthy of attention. Therefore, we conducted a study that enrolled patients with uHCC who received first-line triple therapy and underwent conversion surgery at five major cancer centers in China (*41*). Ultimately, the study included 70 patients. After a median follow-up of 12.9 months, the 1-year recurrence-free survival (RFS) and OS rates were 68.9% and 97.1%, respectively; the 2-year RFS and OS rates were 54.4% and 94.4%, respectively. The prognosis for patients undergoing conversion surgery was similar to that of patients with initially resectable intermediate-stage HCC (*1,13*).

4. Pathological results of triple therapy

The pathological results of conversion surgery after triple therapy were notable. In our study of conversion surgery (*41*), a PCR after triple therapy was observed in 29 (41.4%) patients and an MPR in 59 patients (84.3%). Achieving a PCR was associated with a favorable RFS (hazard ratio [HR] = 0.113, 95% CI: 0.031–0.409, $P =$ 0001). In other studies (*42,43*) on the degree of tumor necrosis after conversion therapy, an MPR or PCR was suggested to improve the prognosis for conversion surgery. Deep tumor cell necrosis after triple therapy may reduce the risk of recurrence.

Based on the triple therapy responses, many patients had a PCR. Since patients had a CR, whether conversion surgery remains necessary was questionable. Therefore, a clinical study (*44*) was conducted to determine whether conversion surgery offers prognostic advantages for patients with uHCC with a clinical complete response (cCR) after conversion therapy. A cCR was defined as 1) serum tumor marker normalization (α-fetoprotein [AFP] < 7 ng/mL and des-gamma-carboxyprothrombin [DCP] < 40 mAU/mL) for ≥ 4 weeks and 2) radiographic CR per mRECIST for ≥ 4 weeks. Ultimately, the study included 74 patients who had cCR; 52 (70.3%) received triple therapy as described in this review. Propensity score matching (PSM) was performed to minimize the influence of potential confounders. Before PSM, 45 patients (60.8%) underwent conversion surgery; 29 (39.2%) received nonsurgical treatment. No statistically significant differences in disease-free survival (DFS) or OS were noted between the two cohorts (HR = 0.715, 95% CI: 0.250–2.043, $P = 0.531$; HR = 0.980, 95% CI: 0.177–5.418, $P = 0.982$, respectively). After PSM, 26 pairs of patients were matched; no significant differences in DFS and OS were noted between the two cohorts (HR = 1.547, 95% CI: 0.51–4.669, $P = 0.439$; HR = 1.024, 95% CI: 0.168–6.242, $P = 0.979$, respectively). This finding suggests that conversion surgery may not be essential for patients with uHCC with cCR.

5. Prognostic prediction of triple therapy

Despite a high ORR, some patients experience disease progression. Therefore, the early prediction of the prognosis for triple therapy is important.

ORR and OS are closely related; therefore, a nomogram model was developed to predict early ORR in patients with uHCC receiving triple therapy after 3 months (*45*). The ORR was 60.9%, and early ORR was predicted independently by AFP, PVTT, tumor number, and tumor size. The nomogram model was highly consistent and clinically useful in the training cohort (C-index = 0.853, 95% CI: 77.50–93.07%). These findings were confirmed in an external validation cohort from three cancer centers in China (C-index = 0.800, 95% CI: 63.52−87.83%).

Moreover, we found that AFP and DCP responses at 6 weeks were predictors for patients with uHCC receiving triple therapy (*46*). After 6 weeks of triple therapy, a > 50% reduction in AFP or DCP levels predicted better treatment outcomes. However, predicting outcomes by the responses of tumor markers remains problematic. Therefore, a prognostic scoring model based on pretreatment baseline levels was developed to predict

outcomes and facilitate earlier treatment decisions (*47*). Patients who received triple therapy at eight centers in China were assigned to training (n = 126) and validation cohorts (n = 84). Baseline patient demographics were collected. In a multivariate analysis, TAE scores (total bilirubin ≥ 17 μmol/L, AFP ≥ 400 ng/mL, and extrahepatic metastasis) were independent predictors of survival in the training cohort. The TAE scoring model was calculated by summing the scores of each of these 1-point risk factors and categorizing the results into three groups: favorable (0 points), intermediate (1 point), and dismal (2–3 points). The TAE score predicted the OS of patients who received triple therapy in both the training (C-index = 0.738, 95% CI: 0.640–0.836) and validation cohorts (C-index = 0.771, 95% CI: 0.689–0.853). The TAE score also stratified PFS well in the training and validation cohorts.

6. The mechanism of triple therapy

Many researchers have sought to explain the potential mechanism of triple therapy (*48-50*). Anti–PD-1 antibodies inhibit the binding of PD-1 and PD-L1, leading to antitumor action by restoring the activity of T cells (*51*). TACE leads to ischemia and tumor tissue necrosis *via* transarterial embolization, and converting "cold" tumors to "hot" tumors by releasing tumor-specific antigens that further enhance the anti-tumor efficacy of anti–PD-1 antibodies. However, a hypoxic microenvironment caused by TACE leads to upward regulation of hypoxia-inducible factor-1, vascular endothelial growth factor, and platelet-derived growth factor receptor, resulting in tumor angiogenesis and progression (*48,52,53*). Lenvatinib is a multi-kinase inhibitor of vascular endothelial growth factors 1-3, fibroblast growth factor receptors 1-4, platelet-derived growth factor receptor alpha, RET, and KIT, which modifies hypoxia and immunosuppression tumor microenvironments by normalizing tumor blood vessels while also enhancing the efficacy of TACE and PD-1 antibodies (*4,54*). The mechanism of triple therapy is complex and synergistic and requires further study to advance our understanding.

7. Treatment-related adverse effects

In addition to the effectiveness of the triple therapy, treatment-related adverse effects (TRAEs) should be considered. Currently, the incidence of grade 3/4 TRAEs in first-line therapy is as high as 37.1–57% (Table 1); the incidence of TRAEs in triple therapy is similar (Table 2). In retrospective cohort studies, no statistically significant differences were noted in the incidence of TRAEs between the triple therapy group and the dual therapy or monotherapy group (*18,19,25-36*).

Identifying the cause of a patient's TRAEs is important since it affects the patient's treatment plan. The most common TRAEs of lenvatinib were hypertension, diarrhea, decreased appetite, and weight loss (*4*). The common TRAEs of TACE included post-thrombotic syndrome (fever, nausea, vomiting, and abdominal pain), liver function damage, allergic reactions, and ectopic embolism (*55*). However, immune toxicities related to the anti–PD-1 antibodies were more extensive (*56*), including almost every organ or system: the skin, endocrine glands (abnormal thyroid function, hypophysitis, primary adrenal insufficiency, type 1 diabetes), lungs (pneumonitis), the gastrointestinal tract (enterocolitis), liver, nervous system, heart, and kidneys. Therefore, scientific monitoring, early detection, correct identification, and effective treatment of TRAEs are very important and could maximize the survival and quality of life for patients with uHCC.

8. Potential problems in triple therapy

Triple therapy has shown promising antitumor activity as a first-line treatment for patients with uHCC; however, several problems remain unsolved. First, triple therapy is not a first-line treatment option because of the lack of randomized phase-III case-controlled trials. Second, combination therapy is not always better than monotherapy. The Leap 002 study (*57*) found that although combining lenvatinib and pembrolizumab showed promising clinical outcomes for uHCC, the OS and PFS for the combination did not meet the prespecified statistical significance compared to lenvatinib monotherapy. Third, triple therapy treatment has Chinese characteristics; the patients in these studies were diagnosed predominantly with hepatitis B virus–related HCC. Therefore, whether triple therapy is as effective for other HCC etiologies requires further investigation. Fourth, the triple therapy in this study combined LRT and MTT plus anti–PD-1 antibodies; the effect of combining other types of LRT or MTT requires further research.

9. Conclusion

Triple therapy shows good clinical outcomes and improves outcomes in patients with uHCC because of its strong antitumor action. However, prospective clinical studies are required to validate triple therapy's effects and provide promising guidance for clinical treatment.

Funding: This study was supported by the Medical Innovation Project of the Health and Family Planning Commission of Fujian Province (Grant number: 2022CXA002) and the Fujian Provincial Health Technology Project (Grant number: 2023CXA005).

Conflict of Interest: The authors have no conflicts of interest to disclose.

References

1. Reig M, Forner A, Rimola J, *et al*. BCLC strategy for prognosis prediction and treatment recommendation: The 2022 update. J Hepatol. 2022; 76:681-693.
2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021; 71:209- 249.
3. Sun HC, Zhou J, Wang Z, *et al*. Chinese expert consensus on conversion therapy for hepatocellular carcinoma (2021 edition). Hepatobiliary Surg Nutr. 2022; 11:227-252.
4. Kudo M, Finn RS, Qin S, *et al*. Lenvatinib versus sorafenib in first-line treatment of patients with unresectable hepatocellular carcinoma: A randomised phase 3 non-inferiority trial. Lancet. 2018; 391:1163-1173.
5. Finn RS, Qin S, Ikeda M, *et al*. Atezolizumab plus bevacizumab in unresectable hepatocellular carcinoma. N Engl J Med. 2020; 382:1894-1905.
6. Llovet JM, Ricci S, Mazzaferro V, *et al*. Sorafenib in advanced hepatocellular carcinoma. N Engl J Med. 2008; 359:378-390.
7. Abou-Alfa GK, Chan SL, Kudo M, *et al*. Phase 3 randomized, open-label, multicenter study of tremelimumab and durvalumab as first-line therapy in patients with unresectable hepatocellular carcinoma: HIMALAYA. J Clin Oncol. 2022; 40(4 Suppl.):379.
8. Qin S, Bai Y, Lim HY, Thongprasert S, Chao Y, Fan J, Yang TS, Bhudhisawasdi V, Kang WK, Zhou Y, Lee JH, Sun Y. Randomized, multicenter, open-label study of oxaliplatin plus fluorouracil/leucovorin versus doxorubicin as palliative chemotherapy in patients with advanced hepatocellular carcinoma from Asia. J Clin Oncol. 2013; 31:3501-3508.
9. Qin S, Bi F, Gu S, *et al*. Donafenib versus sorafenib in first-line treatment of unresectable or metastatic hepatocellular carcinoma: A randomized, open-label, parallel-controlled phase II-III trial. J Clin Oncol. 2021; 39:3002-3011.
10. Qin S, Kudo M, Meyer T, *et al*. Tislelizumab *vs* sorafenib as first-line treatment for unresectable hepatocellular carcinoma: a phase 3 randomized clinical trial. JAMA Oncol. 2023; 9:1651-1659.
11. Qin S, Chan SL, Gu S, *et al*. Camrelizumab plus rivoceranib versus sorafenib as first-line therapy for unresectable hepatocellular carcinoma (CARES-310): A randomised, open-label, international phase 3 study. Lancet. 2023; 402:1133-1146.
12. Ren Z, Xu J, Bai Y, *et al*. Sintilimab plus a bevacizumab biosimilar (IBI305) versus sorafenib in unresectable hepatocellular carcinoma (ORIENT-32): A randomised, open-label, phase 2-3 study. Lancet Oncol. 2021; 22:977-990.
13. Zhou J, Sun H, Wang Z, *et al*. Guidelines for the diagnosis and treatment of hepatocellular carcinoma (2019 edition). Liver Cancer. 2020; 9:682-720.
14. Zhu HD, Li HL, Huang MS, *et al*. Transarterial chemoembolization with PD-(L)1 inhibitors plus molecular targeted therapies for hepatocellular carcinoma (CHANCE001). Signal Transduct Target Ther. 2023; 8:58.
15. Pei Y, Li W, Wang Z, Liu J. Successful conversion therapy for unresectable hepatocellular carcinoma is getting closer: A systematic review and meta-analysis. Front Oncol. 2022; 12:978823.
16. Wu JY, Yin ZY, Bai YN, Chen YF, Zhou SQ, Wang SJ, Zhou JY, Li YN, Qiu FN, Li B, Yan ML. Lenvatinib combined with anti-PD-1 antibodies plus transcatheter arterial chemoembolization for unresectable hepatocellular carcinoma: A multicenter retrospective study. J Hepatocellular Carcinoma. 2021; 8:1233-1240.
17. Li X, Chen J, Wang X, Bai T, Lu S, Wei T, Tang Z, Huang C, Zhang B, Liu B, Li L, Wu F. Outcomes and prognostic factors in initially unresectable hepatocellular carcinoma treated using conversion therapy with lenvatinib and TACE plus PD-1 inhibitors. Front Oncol. 2023; 13:1110689.
18. Qu WF, Ding ZB, Qu XD, *et al*. Conversion therapy for initially unresectable hepatocellular carcinoma using a combination of toripalimab, lenvatinib plus TACE: real-world study. BJS Open. 2022; 6:zrac114.
19. Chen S, Wu Z, Shi F, Mai Q, Wang L, Wang F, Zhuang W, Chen X, Chen H, Xu B, Lai J, Guo W. Lenvatinib plus TACE with or without pembrolizumab for the treatment of initially unresectable hepatocellular carcinoma harbouring PD-L1 expression: A retrospective study. J Cancer Res Clin Oncol. 2022; 148: 2115-2125.
20. Liu J, Li Z, Zhang W, Lu H, Sun Z, Wang G, Han X. Comprehensive treatment of trans-arterial chemoembolization plus lenvatinib followed by camrelizumab for advanced hepatocellular carcinoma patients. Front Pharmacol. 2021; 12: 709060.
21. Cao F, Yang Y, Si T, Luo J, Zeng H, Zhang Z, Feng D, Chen Y, Zheng J. The efficacy of TACE combined with lenvatinib plus sintilimab in unresectable hepatocellular carcinoma: A multicenter retrospective study. Front Oncol. 2021;11: 783480.
22. Teng Y, Ding X, Li W, Sun W, Chen J. A retrospective study on therapeutic efficacy of transarterial chemoembolization combined with immune checkpoint inhibitors plus lenvatinib in patients with unresectable hepatocellular carcinoma. Technol Cancer Res Treat. 2022; 21:15330338221075174.
23. Qu S, Wu D, Hu Z. Neutrophil-to-lymphocyte ratio and early tumor shrinkage as predictive biomarkers in unresectable hepatocellular carcinoma patients treated with lenvatinib, PD-1 inhibitors, in combination with TACE. Technol Cancer Res Treat. 2023; 22:15330338231206704.
24. Li X, Fu Z, Chen X, Cao K, Zhong J, Liu L, Ding N, Zhang X, Zhai J, Qu Z. Efficacy and safety of lenvatinib combined with PD-1 inhibitors plus TACE for unresectable hepatocellular carcinoma patients in China real-world. Front Oncol. 2022; 12:950266.
25. Zhao S, Zhou M, Wang P, Yang J, Zhang D, Yin F, Song P. Sorafenib, lenvatinib, or lenvatinib combining PD-1 inhibitors plus TACE in unresectable hepatocellular carcinoma: A retrospective analysis. Technol Cancer Res Treat. 2022; 21:15330338221133640.
26. Sun B, Zhang L, Sun T, Ren Y, Cao Y, Zhang W, Zhu L, Guo Y, Gui Y, Liu F, Chen L, Xiong F, Zheng C. Safety and efficacy of lenvatinib combined with camrelizumab plus transcatheter arterial chemoembolization for unresectable hepatocellular carcinoma: A two-center retrospective study. Front Oncol. 2022; 12:982948.
27. Qu S, Zhang X, Wu Y, Meng Y, Pan H, Fang Q, Hu L, Zhang J, Wang R, Wei L, Wu D. Efficacy and safety of TACE combined with lenvatinib plus PD-1 inhibitors compared with TACE alone for unresectable hepatocellular carcinoma patients: A prospective cohort

study. Front Oncol. 2022; 12:874473.
28. Guo P, Pi X, Gao F, Li Q, Li D, Feng W, Cao W. Transarterial chemoembolization plus lenvatinib with or without programmed death-1 inhibitors for patients with unresectable hepatocellular carcinoma: A propensity score matching study. Front Oncol. 2022; 12:945915.
29. Yang H, Yang T, Qiu G, Liu J. Efficacy and safety of TACE combined with lenvatinib and PD-(L)1 inhibitor in the treatment of unresectable hepatocellular carcinoma: A retrospective study. J Hepatocell Carcinoma. 2023; 10:1435-1443.
30. Wang YY, Yang X, Wang YC, Long JY, Sun HS, Li YR, Xun ZY, Zhang N, Xue JN, Ning C, Zhang JW, Zhu CP, Zhang LH, Yang XB, Zhao HT. Clinical outcomes of lenvatinib plus transarterial chemoembolization with or without programmed death receptor-1 inhibitors in unresectable hepatocellular carcinoma. World J Gastroenterol. 2023; 29:1614-1626.
31. Xin Y, Zhang X, Liu N, Peng G, Huang X, Cao X, Zhou X, Li X. Efficacy and safety of lenvatinib plus PD-1 inhibitor with or without transarterial chemoembolization in unresectable hepatocellular carcinoma. Hepatol Int. 2023; 17:753-764.
32. Xiang Z, Li G, Mu L, Wang H, Zhou C, Yan H, Huang M. TACE combined with lenvatinib and camrelizumab for unresectable multiple nodular and large hepatocellular carcinoma (> 5 cm). Technol Cancer Res Treat. 2023; 22:15330338231200320.
33. Xiang YJ, Wang K, Yu HM, Li XW, Cheng YQ, Wang WJ, Feng JK, Bo MH, Qin YY, Zheng YT, Shan YF, Zhou LP, Zhai J, Cheng SQ. Transarterial chemoembolization plus a PD-1 inhibitor with or without lenvatinib for intermediate-stage hepatocellular carcinoma. Hepatol Res. 2022; 52:721-729.
34. Cai M, Huang W, Huang J, Shi W, Guo Y, Liang L, Zhou J, Lin L, Cao B, Chen Y, Zhou J, Zhu K. Transarterial chemoembolization combined with lenvatinib plus PD-1 inhibitor for advanced hepatocellular carcinoma: A retrospective cohort study. Front Immunol. 2022; 13:848387.
35. Wu J, Zeng J, Wang H, Huo Z, Hou X, He D. Efficacy and safety of transarterial chemoembolization combined with lenvatinib and camrelizumab in patients with BCLC-defined stage C hepatocellular carcinoma. Front Oncol. 2023; 13:1244341.
36. Zou X, Xu Q, You R, Yin G. Evaluating the benefits of TACE combined with lenvatinib plus PD-1 inhibitor for hepatocellular carcinoma with portal vein tumor thrombus. Adv Ther. 2023; 40:1686-1704.
37. Li SQ, Wu JY, Wu JY, Xie H, Li JH, Zeng ZX, Fu YK, Liu DY, Li H, Chen WZ, Huang JY, Yan ML. Transarterial chemoembolization plus lenvatinib and PD-1 inhibitors for hepatocellular carcinoma with main trunk portal vein tumor thrombus: A multicenter retrospective study. J Hepatocell Carcinoma. 2023; 10:1799-1811.
38. Ning S, Li X, Ma X, Liu J, Chang X. Efficacy of TACE combined with lenvatinib plus sintilimab for hepatocellular carcinoma with tumor thrombus in the inferior vena cava and/or right atrium. J Hepatocell Carcinoma. 2023; 10:1511-1525.
39. Li X, Ding X, Liu M, Wang J, Sun W, Teng Y, Xu Y, Wu H, Li W, Zhou L, Chen J. A multicenter prospective study of TACE combined with lenvatinib and camrelizumab for hepatocellular carcinoma with portal vein tumor thrombus. Cancer Med. 2023; 12:16805-16814.
40. Qiu G, Xie K, Jin Z, Jiang C, Liu H, Wan H, Huang J. The multidisciplinary management of hepatocellular carcinoma with portal vein tumor thrombus. Biosci Trends. 2021; 15:148-154.
41. Wu JY, Zhang ZB, Zhou JY, Ke JP, Bai YN, Chen YF, Wu JY, Zhou SQ, Wang SJ, Zeng ZX, Li YN, Qiu FN, Li B, Yan ML. Outcomes of salvage surgery for initially unresectable hepatocellular carcinoma converted by transcatheter arterial chemoembolization combined with lenvatinib plus anti-PD-1 antibodies: A multicenter retrospective study. Liver cancer. 2023; 12:229-237
42. Yang K, Sung PS, You YK, Kim DG, Oh JS, Chun HJ, Jang JW, Bae SH, Choi JY, Yoon SK. Pathologic complete response to chemoembolization improves survival outcomes after curative surgery for hepatocellular carcinoma: Predictive factors of response. HPB. 2019; 21:1718-1726.
43. Allard MA, Sebagh M, Ruiz A, Guettier C, Paule B, Vibert E, Cunha AS, Cherqui D, Samuel D, Bismuth H, Castaing D, Adam R. Does pathological response after transarterial chemoembolization for hepatocellular carcinoma in cirrhotic patients with cirrhosis predict outcome after liver resection or transplantation? J Hepatology. 2015; 63:83-92.
44. Wu JY, Wu JY, Liu DY, Li H, Zhuang SW, Li B, Zhou JY, Huang JY, Zhang ZB, Li SQ, Yan ML, Wang YD. Clinical complete response after conversion therapy for unresectable hepatocellular carcinoma: Is salvage hepatectomy necessary? J Hepatocell Carcinoma. 2023; 10:2161-2171.
45. Li S, Wu J, Wu J, Fu Y, Zeng Z, Li Y, Li H, Liao W, Yan M. Prediction of early treatment response to the combination therapy of TACE plus lenvatinib and anti-PD-1 antibody immunotherapy for unresectable hepatocellular carcinoma: Multicenter retrospective study. Front Immunol. 2023; 14:1109771.
46. Luo MC, Wu JY, Wu JY, Lin ZT, Li YN, Zeng ZX, Wei SM, Yan ML. Early tumor marker response predicts treatment outcomes in patients with unresectable hepatocellular carcinoma receiving combined lenvatinib, immune checkpoint inhibitors, and transcatheter arterial chemoembolization therapy. J Hepatocell Carcinoma. 2023; 10:1827-1837.
47. Zeng ZX, Wu JY, Wu JY, Li YN, Fu YK, Zhang ZB, Liu DY, Li H, Ou XY, Zhuang SW, Yan ML. The TAE score predicts prognosis of unresectable HCC patients treated with TACE plus lenvatinib with PD-1 inhibitors. Hepatol Int. 2023; doi: 10.1007/s12072-023-10613-x.
48. Llovet JM, Castet F, Heikenwalder M, Maini MK, Mazzaferro V, Pinato DJ, Pikarsky E, Zhu AX, Finn RS. Immunotherapies for hepatocellular carcinoma. Nat Rev Clin Oncol. 2022; 19:151-172.
49. Lu Y, Jin J, Du Q, Hu M, Wei Y, Wang M, Li H, Li Q. Multi-omics analysis of the anti-tumor synergistic mechanism and potential application of immune checkpoint blockade combined with lenvatinib. Front Cell Dev Biol. 2021; 9: 730240.
50. Xing R, Gao J, Cui Q, Wang Q. Strategies to improve the antitumor effect of immunotherapy for hepatocellular carcinoma. Front Immunol. 2021; 26:783236.
51. El-Khoueiry AB, Sangro B, Yau T, *et al*. Nivolumab in patients with advanced hepatocellular carcinoma (CheckMate 040): An open-label, non-comparative, phase 1/2 dose escalation and expansion trial. Lancet. 2017; 389:2492-2502.

52. Hack SP, Zhu AX, Wang Y. Augmenting anticancer immunity through combined targeting of angiogenic and PD-1/PD-L1 pathways: Challenges and opportunities. Front Immunol. 2020; 5:598877.
53. Kudo M. A new treatment option for intermediate-stage hepatocellular carcinoma with high tumor burden: Initial lenvatinib therapy with subsequent selective TACE. Liver Cancer. 2019; 8:299-311.
54. Yamamoto Y, Matsui J, Matsushima T, *et al*. Lenvatinib, an angiogenesis inhibitor targeting VEGFR/FGFR, shows broad antitumor activity in human tumor xenograft models associated with microvessel density and pericyte coverage. Vasc Cell. 2014; 6:18.
55. Jiang JQ, Huang JT, Zhong BY, Wang WD, Sun JH, Wang Q, Ding WB, Ni CF, Zhu XL. Transarterial chemoembolization for patients with unresectable hepatocellular carcinoma with Child-Pugh B7. J Hepatocell Carcinoma. 2023; 10:1629-1638.
56. Thuny F, Alexandre J, Salem JE, Mirabel M, Dolladille C, Cohen-Solal A, Cohen A, Ederhy S, Cautela J; French Working Group of Cardio-Oncology. Management of immune checkpoint inhibitor-induced myocarditis: The French Working Group's plea for a pragmatic approach. JACC CardioOncol. 2021; 3:157-161.
57. Llovet JM, Kudo M, Merle P, *et al*. Lenvatinib plus pembrolizumab versus lenvatinib plus placebo for advanced hepatocellular carcinoma (LEAP-002): A randomised, double-blind, phase 3 trial. Lancet Oncol. 2023; 24:1399-1410.

Received December 28, 2023; Revised January 27, 2024; Accepted February 2, 2024.

**Address correspondence to:*
Maolin Yan, The Shengli Clinical Medical College of Fujian Medical University, Department of Hepatobiliary Pancreatic Surgery, Fujian Provincial Hospital, No. 134 Dongjie Road, Fuzhou, Fujian 350001, China.
E-mail: yanmaolin74@163.com

Released online in J-STAGE as advance publication February 8, 2024.

Review

DOI: 10.5582/bst.2023.01288

Monoclonal antibody therapy for Alzheimer's disease focusing on intracerebral targets

Xiaolei Gu[1,§], Long Qi[2,§], Qing Qi[3,4,5], Jing Zhou[3,4,5], Song Chen[6,*], Ling Wang[3,4,5,*]

[1] College of Acupuncture and Orthopedics, Hubei University of Chinese Medicine, Wuhan, Hubei, China;
[2] New Drug Screening Center, Jiangsu Center for Pharmacodynamics Research and Evaluation, China Pharmaceutical University, Nanjing, China;
[3] Laboratory for Reproductive Immunology, Obstetrics and Gynecology Hospital of Fudan University, Shanghai, China;
[4] The Academy of Integrative Medicine of Fudan University, Shanghai, China;
[5] Shanghai Key Laboratory of Female Reproductive Endocrine-related Diseases, Shanghai, China;
[6] Postdoctoral Station of Xiamen University, Fujian, China.

SUMMARY Alzheimer's disease (AD) is one of the most common neurodegenerative diseases. Due to the complexity of the disorder and the presence of the blood-brain barrier (BBB), its drug discovery and development are facing enormous challenges, especially after several failures of monoclonal antibody (mAb) trials. Nevertheless, the Food and Drug Administration's approval of the mAb aducanumab has ushered in a new day. As we better understand the disease's pathogenesis and identify novel intracerebral therapeutic targets, antibody-based therapies have advanced over the past few years. The mAb drugs targeting β-amyloid or hyperphosphorylated tau protein are the focus of the current research. Massive neuronal loss and glial cell-mediated inflammation are also the vital pathological hallmarks of AD, signaling a new direction for research on mAb drugs. We have elucidated the mechanisms by which AD-specific mAbs cross the BBB to bind to targets. In order to investigate therapeutic approaches to treat AD, this review focuses on the promising mAbs targeting intracerebral dysfunction and related strategies to cross the BBB.

Keywords Alzheimer's disease, blood-brain barrier, pathological mechanism, monoclonal antibodies, target

1. Introduction

Alzheimer's disease (AD), as the leading neurodegenerative disease, is a major cause of dementia that occurs in the middle-aged and elderly population (*1*). As aging of the population intensifies, the disease's incidence increases yearly, seriously affecting the life quality of patients and their families (*2*). The pathogenesis of AD is complex and has yet to be fully elucidated. Hyperphosphorylated tau, β-amyloid (Aβ) plaques, and neuroinflammation are considered core pathological factors (*3,4*).

Over the past few years, research has focused on early detection, diagnosis, and treatment of AD (*5*). Immune therapy, and especially disease-modifying therapy, has played an essential role in disease prevention and treatment due to its unique target specificity. Along with the discovery of the intracerebral targets and biomarkers of AD, antibody-based drugs have shed new light on AD, and this is especially true since the approval of the monoclonal antibody (mAB) aducanumab for AD. As a highly homogeneous antibody produced from a single B cell clone, mAbs work at a specific epitope (*6*), which means they have a high level of target specificity. However, the blood-brain barrier (BBB), a highly selective membrane barrier that prevents 98% of small-molecule drugs and almost 100% of large-molecule drugs from crossing (*7*), poses a considerable challenge to drug development.

Several reviews have discussed advances in research on mAbs against Aβ (*8-11*). However, new etiologic and pathological factors have been uncovered based on the Aβ hypothesis over the last two decades. At the same time, new strategies have emerged to bypass the BBB for better efficacy. Based on the pathogenesis and targets of AD, this review has summarized the intracerebral dysfunction of the disease, outlined the use of AD-specific mAbs, and discussed the strategies by which antibody drugs cross the BBB to treat AD.

2. Pathogenesis of AD

Although the pathogenesis of AD remains unclear, intracerebral senile plaques, neurofibrillary tangles (NFTs), and concomitant neuroinflammation are believed

to play significant roles (*12,13*). Later, morphological changes happen, including atrophied brain tissue, reduced weight, and even numerous neuronal losses in the brain (*14*). Although the factors and mechanisms that cause these changes remain unclear, age is undoubtedly the most important one. Several factors, including sex, genetic mutations, and lifestyle habits, affect neuronal regeneration through changes in hormone levels, systemic validation, and so on, contributing to the formation of senile plaques and NFS and activation of neuroinflammation in the brain (*15-17*).

2.1. Aβ hypothesis

Evidence has indicated that Aβ peptide plays a vital role in the pathogenesis of AD, so Aβ is often used as a neuropathological diagnostic criterion for AD (*18,19*). Physiologically, Aβ is widely present in the body and brain and is involved in neuronal growth, regulation of synaptic function, protection against oxidative stress, and even the innate immune system (*20,21*). Aβ production and elimination are altered with aging, leading to downstream activation. A study has indicated that Aβ promotes astrocyte senescence *via* NLRP3 pathway activation (*22*). Moreover, Aβ1-42 oligomers induce secretion of senescence factors such as p16 and SA-β-gal in adult mouse hippocampal neural stem cells (*23*). Presenilin (PSEN) 1 and PSEN2, two presenilin genes, can affect amyloid precursor protein (APP) cleavage by influencing the expression of β-secretase and γ-secretase. In AD encephalopathy, APP is successively cleaved by β-secretase and γ-secretase to produce neurotoxic monomer Aβ (*24*). These products from anomalous enzyme shearing lead to monomeric Aβ misfolding, and the consequent oligomers further fold into protofibrils and fibrils, eventually forming amyloid plaques to induce a cytotoxic effect (Figure 1).

Aβ accumulation is associated with AD processes, both in familial and sporadic AD (*25-27*). The gene mutations, such as PSEN 1, PSEN2, and APP, are also associated with Aβ aggregation in familial AD, causing cognitive impairment in patients with AD carrying these mutations (*28*). Based on the amyloid hypothesis, excessive Aβ plaque deposition in the brain can induce a series of downstream pathological changes, such as tau-associated network disruption, chronic inflammation, failure of metabolic activity, abnormal microglial activation, oxidative stress, cholesterol-associated neuronal distress, and autophagy deficit (*8,29,30*). Additional studies have indicated that the

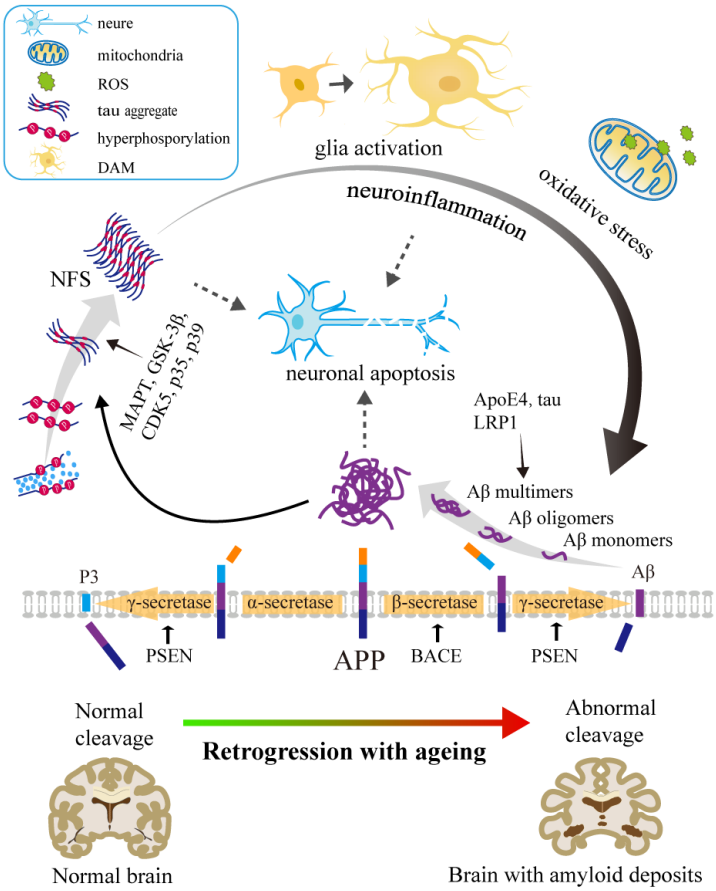

Figure 1. AD pathological processes. Accompanying aging, abnormal APP clearance contributes to Aβ accumulation, which induces AD pathological processes such as tau accumulation, abnormal glial activation, release of inflammatory factors, and neuronal damage. PSEN and BACE genes affect the formation of Aβ monomers by regulating γ-secretase and β-secretase, while ApoE, tau, and neuroinflammation promote the accumulation of Aβ. P3, a non-amyloidogenic peptide, is a cleavage product of APP and lacks pathological effects due to its solubility.

damage is caused by soluble Aβ, including oligomers consisting of a small number of Aβ peptides, rather than an accumulation of aggregated Aβ (*31,32*), and especially in early neuronal toxicity (*33*). Injecting soluble Aβ42 oligomers isolated from the brains of patients with AD into healthy rats impairs memory, reduces the number of synapses, and enhances long-term synaptic inhibition in the rodent hippocampus (*34*). This may be because Aβ localization induces mitochondrial dysfunction by impeding the electron transport chain of the mitochondrial membrane and by increasing reactive oxygen species (ROS) production, finally leading to neuronal death (*35,36*).

Large-scale genome-wide studies have identified sporadic AD as the most commonly occurring type of dementia worldwide and a multifactorial disease caused by diverse genetic factors (*37,38*). ApoE, a lipid and cholesterol carrier that responds to transporting nutrients from astrocytes to neurons *via* transmembrane transport and regulating neurons in the central nervous system, is highly expressed by glial cells around Aβ plaque (*39*). *In vitro* and *in vivo* studies have indicated that ApoE isoform, especially ApoE4, promotes the progression of Aβ peptides to Aβ oligomers, protofibrils, and fibrils and inhibits the clearance and enzymatic degradation of intracerebral Aβ (*40,41*). Neurons carrying ApoE4 grow at a lower rate and density than those carrying ApoE3, making individuals with the ApoE4 more vulnerable to attacks of AD (*42*). Moreover, neprilysin and insulin-degrading enzymes are required for Aβ clearance, and the expression of ApoE4 appears to decrease the activity of both enzymes, unlike in cadavers carrying the non-ApoE gene (*43,44*).

In summary, Aβ has a wide range of pathologic roles in AD progression. Although the exact mechanism remains unclear, clinical trial results have indicated that reducing either soluble Aβ, or amyloid plaques, or both, in the brain to non-pathological levels — that is, below the level that provokes tau pathology spread — may be of therapeutic benefit to patients with AD (*45*).

2.2. Tau protein-related mechanisms

Tau protein, enriched in axons, regulates intraneuronal transport, microtubule dynamics, and synaptic transmission. However, pathological tau protein is the basis of intracellular NFTs, which are another contributor to AD pathogenesis. Tau protein has a microtubule-binding domain expressed by a continuous sequence of repetitive conserved sequences at the carboxyl terminus (*46,47*). These conserved sequences, as the sites for microtubule binding, constitute the structure of mature and stable microtubules (*48*). Tau protein usually occurs as monomers, small oligomers, and pairs of helical and straight filaments. They tend to dissociate from microtubules and tangle in the neuro when their excessive or abnormal phosphorylation forms pathological structures (*49,50*). Once the tau protein loses its ability to bind to microtubules, the neural cell architecture is destroyed, leading to disruption of signal processing and transport of the substance between neuronal synapses, eventually inducing neuronal apoptosis (*51*). Therefore, hyperphosphorylated pathological tau protein can degenerate neurons due to its cytotoxicity and disturbance of microtubules (*52*). Hyperphosphorylated tau proteins are also thought to spread because they are taken up by surrounding normal neurons *via* synaptic transmission, exosomal release, or direct extracellular secretion, leading to abnormal aggregation of tau proteins in healthy neurons and the continuous production of pathological tau (*46*). Thus, hyperphosphorylated tau protein has become a biomarker for disease diagnosis (*53*). Although the exact mechanisms of tau hyperphosphorylation are still unclear, Aβ appears to be involved, according to the amyloid hypothesis during cascade (*54*).

However, a differing view is that tau pathology may be a prerequisite for causing Aβ in AD. Autopsies have revealed that tau pathology appears to precede Aβ accumulation and that it is closely related to the patient's cognitive impairment. The reason why is related to the regulation of kinases and genetic variation (*55*). Several kinases are involved in accomplishing the phosphorylation of tau protein. Glycogen synthase kinase-3β (GSK-3β) and cyclin dependent kinase 5 (CDK5) are two of the most critical kinases. GSK-3β regulates tau phosphorylation mainly *via* PI3K/AKT/GSK-3β pathway, while CDK5 is regulated by p35 and p39. Once overactivated, the kinases promote tau hyperphosphorylation and cause neuronal injury, Aβ aggregation, inflammation, and mitochondrial dysfunction (*56*). MAPT, another gene associated with tau pathology, prompts tau expression as 3R (exon ten exclusion) or 4R (exon ten inclusion), two isoforms associated with tau aggregation (*57*). The pathologic changes of AD contain approximately equal amounts of 3R-tau and 4R-tau, which are thought to play an important role in AD pathology (*58*). Misrepresentation and mistranslation of MAPT lead to 3R and 4R expression (*59*). Moreover, ApoE4 promotes tau-induced neurodegeneration and atrophy in that ApoE affects the ability of tau to bind to LRP1, accelerating tau diffusion (*60,61*). In addition, inflammation is a factor that promotes tau hyperphosphorylation (*62*). Tau protein aggregation is associated with cellular senescence in the brain. Aging-associated secretory phenotypes and NF-κB activation upregulate tau by impairing mitochondrial function (*63*). In addition, upregulation of the Cdkn2a gene is related to NFT formation (*63*).

2.3. The roles of inflammatory factors

Complex and prolonged neuroinflammation also contributes to AD pathogenesis. Neuroinflammation

in the early development of AD clears neurotoxins such as Aβ and tau by phagocytosis; in the later stages, however, persistent neuroinflammation is a facilitator of neurological damage (*64*). As the resident immune cells, microglia and astrocytes are housekeeping phagocytes, playing a crucial role in neuroinflammation. Typically, microglia can remove diseased neurons by phagocytosis and endocytosis (*65*). Astrocytes play an important role in the maintenance of cellular metabolism homeostasis, providing energy substrates and nutrition to neurons. In addition, astrocytes, as the composition of the BBB, are involved in regulating the pH level, energy balance, neurotransmitter removal, and metal ion balance of the brain (*66*). Microglia and astrocytes can respond accordingly to poisons such as Aβ or tau. They release pro-inflammatory factors to accelerate the metabolism of these toxins, thereby maintaining homeostasis in the brain (*13*). Moreover, microglia secrete glial-derived and brain-derived neurotrophic factors to repair neurons (*67*). Once neuroprotective feedback is disrupted, however, microglia and astrocytes are overactivated in response to oligomer Aβ and tau, leading to chronic neuroinflammation that can damage neurons and induce synaptic death and nerve cell senescence (*68,69*). The complement system and multiple gene mutations are also involved (*4*). Furthermore, lifestyle habits and diseases may be other promoters (*70*).

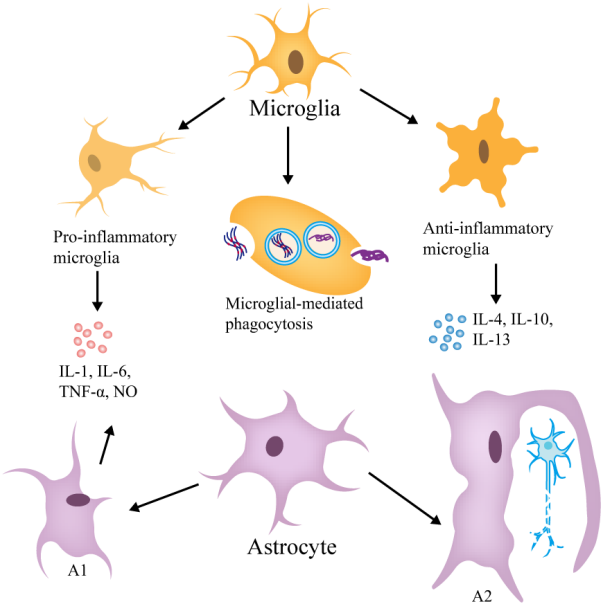

Figure 2. The roles of microglia and astrocytes in different phenotypes after activation. Pro-inflammatory microglia and A1 astrocytes secrete pro-inflammatory factors such as IL-1, IL-6, TNF-α, and NO. Correspondingly, anti-inflammatory microglia release IL-4, IL-10, and IL-13 to counteract inflammation. Astrocyte cells are activated in the A2 phenotype and then grow to form a physical barrier around the lesion, blocking the propagation of pathologic products.

2.3.1. Abnormal activation of glial cells

As the resident immune cells, microglia and astrocytes are housekeeping phagocytes, playing a crucial role in neuroinflammation (Figure 2). Disease-associated microglia (DAM) are aberrantly activated and associated with neuroinflammation and Aβ aggregates (*71*). They interact with Aβ to release pro-inflammatory factors, including interleukin (IL)-1β, IL-6, tumor necrosis factor (TNF)-α, and reactive nitrogen oxides, mediating neuroinflammation and interfering with synaptic sprouting and axonal growth (*72*). Moreover, abnormal microglia activation may be associated with tau hyperphosphorylation (*73*), which may be related to the fact that microglia trigger the NLRP3 pathway during Aβ clearance (*74*). Astrocytes play an equally important role in AD pathology and chronic inflammation. Similar to microglia, astrocytes have different phenotypes after pathological stimulation in AD. The A1 astrocyte phenotype is dominated by NF-κB pathway-mediated inflammation, and the A2 astrocyte is dominated by gliosis dependent on the signal transducer and activator of transcription 3 pathway (*69*). Typically, A1 astrocytes are associated with neurological injury, while A2 is associated with neuroprotection in AD. A1 astrocytes induce neuronal death by releasing inflammatory factors and activating the complement system (*75*). Stimulated by IL-1α and TNFα released by microglia, A1 astrocytes are activated, which further induces neuronal death and phagocytosis (*75*). In addition, they activates the complement system and induce long-term neuroinflammation (*75*). A2 astrocytes enlarge upon activation and exhibit overexpression of glial fibrillary acidic protein. These proteins form a physical barrier around the lesion, blocking the spread of pathologic products (*76*). In turn, neuroinflammation exacerbates the accumulation of Aβ and tau. Inflammatory factors such as IL-1β, lipopolysaccharide, and prostaglandin E2 reduce the oligomeric Aβ and tau uptake by microglia, promoting the accumulation of pathological products (*76,77*).

2.3.2. Involvement of the complement system

The complement system plays a regulatory role in glial cell-mediated neuroinflammation. Aβ and tau can activate the complement system. When Aβ activates the NF-κB pathway, A1 astrocytes release C3, which can aggravate tau accumulation and microglia activation (*78,79*). In contrast, complement C3aR inactivation on the surface of microglia and astrocytes attenuates tau pathology and reverses dysregulated immune networks in a model of tau disease (*80*). DAM in turn mediate neuronal clearance by astrocytes through the release of C1q (*69,81*). Inhibition of C1q, C3, or the microglia complement receptor CR3 reduces the number of phagocytic microglia (*82*). C5a induces a chronic inflammatory state in microglia by binding to the receptor C5aR1, which causes cell lysis by forming the membrane attack complex MAC. Inhibition

of the C5a cascade response blocks neuronal damage by the complement system (*83*). Therefore, the complement system may be a potential therapeutic target for AD.

2.3.3. Receptors and genes regulate neuroinflammation

Specific receptor expression and genetic variants are associated with aberrantly activated glial cells. Toll-like receptor (TLR) 4 on the surface of microglia is recognized by Aβ, which can prolong microglia activation, increase phagocytosis and cytokine production, and stimulate Aβ accumulation (*76,84*). Inflammasomes are involved in initiating and maintaining the innate immune response and activating IL-1β and IL-18 (*85*). The NLRP3 inflammasome consists of a sensor (NLRP3), an adaptor (apoptosis-associated speck-like protein (ASC) or PYARD), and an effector (caspase 1) (*86*). Aβ binds to ASC to form the ACS-Aβ complex and activates NF-κB (*87*), which promotes NLRP3 activation and the release of inflammatory mediators such as IL-1β and IL-18 from microglia to induce apoptosis (*88*). Studies have indicated that ACS inhibition can mitigate Aβ aggregation (*89,90*). Crosstalk among TLR4, the NLRP3 inflammasome, and complements promotes neuroinflammation in AD. In addition, aggregated Aβ phagocytosed by microglia damages lysosomes and leaks into the cytoplasm, also contributing to inflammasome activation. Inflammasome-induced cellular pyroptosis leads to the secretion of IL-1β and ASC speckles, which bind to Aβ, leading to further Aβ aggregation. This vicious cycle exacerbates the pathology of Aβ (*87*).

Triggering receptor expressed on myeloid cells 2 (TREM2) may be an AD protective receptor because of its anti-inflammatory functions. TREM2 interacts with the junction proteins DAP12 and DAP10 to affect tyrosine phosphorylation and promote microglial chemotaxis and phagocytosis of toxicants and damaged synapses (*91*). TREM2-associated pathway activation modulates microglia phagocytosis, accelerates Aβ plaque clearance, and ameliorates AD pathology (*92-94*). In addition, TREM2 inhibits tau protein phosphorylation by inhibiting GSK-3β, a tau phosphorylation-related kinase (*95*). In contrast, reduced expression or mutation of TREM2 is associated with high levels of pro-inflammatory mediators such as TNF-α, IL-6, and IL-1 (*96*). Mutations and polymorphisms of the TREM2 gene have been associated with a significantly increased risk of AD; among the various variants, the R47H variant is associated with decreased Aβ, tau clearance, and severe neuroinflammation (*97*). Thus, activation of TREM2 may be a potential therapeutic target of AD (*98*).

ApoE also plays a crucial role in neuroinflammation. Microglia and astrocytes encoded with the ApoE4 variant display an immunomodulatory effect, actively participating in neuroinflammation. Transcriptomics studies have suggested that gene expression of ApoE is altered during DAM activation (*99,100*). In addition, ApoE4 mediates higher levels of inflammatory factors (TNF-α, IL-6, and IL-1β) than other phenotypes (*101*). In addition, microglia expressing ApoE4 but not ApoE3 inhibit CNS-associated macrophages from responding to pathological changes in AD (*102*). Animal models have indicated that mice expressing ApoE4 exhibit increased lipid droplet formation and synaptic dysfunction in the brain, while mice expressing ApoE3 have improved memory and synaptic plasticity (*102,103*). ApoE4 is also associated with astrocytes with a lower autophagic flux, reducing the clearance of neurotoxic substances such as Aβ (*104*). In addition, the interplay between ApoE4 and Aβ oligomers may result in synaptic loss. When ApoE expression is eliminated in astrocytes during the phase of Aβ accumulation, the burden of Aβ plaques can be reduced (*105*), indicating its vital role in inflammatory response (*106*). This may explain the poor efficacy and high rate of edema as a complication of existing monotherapy antibody drugs for patients with the ApoE4 mutation (*107,108*). In conclusion, ApoE is an essential modulator in the AD process and is mainly considered to be a risk factor, while ApoE2 has more protective properties.

2.3.4. Impact of lifestyle and diseases

Neuroinflammation in AD is also associated with other factors such as aging, alcoholism, and chronic life stress. Animal experiments have indicated that sustained proliferation of DAMs increases βgal activity, a senescence-associated transcriptional signature, and telomere shortening, leading to cellular senescence (*109*). Kanchan discovered dark microglia, a new phenotype predominantly associated with pathological states, that are expressed primarily in chronic stress, aging, and AD (*110*). In addition, alcohol abuse affects microglia-mediated neuroinflammation, leading to synapse elimination and exacerbating cognitive impairment (*111*). Chronic life stress induces stressful heterogeneity in astrocytes, leading to the pyrolytic death of astrocytes, which may be related to stress-induced changes in glucocorticoids that affect astrocyte cell-associated targets (*112*). Interestingly, aside from brain inflammation, systemic inflammation affects AD as well. Gut flora interact with AD *via* multiple pathways. Alterations in gut flora activate pro-inflammatory cytokines, alter the systemic inflammatory milieu, and trigger systemic inflammation-derived pro-inflammatory factors that enhance neuroinflammation (*113,114*). Intestinal permeability is also increased when gut flora are altered, resulting in the transfer of Aβ oligomers from the gut to the brain (*115*). In addition, AD may be complicated by obesity and type 2 diabetes. Those conditions cause aberrant activation of the NLRP3 signaling pathway, which contributes to the activation of the inflammatory vesicle complex and the release of IL-1β and IL-18; NLRP3 pathway activation may be key

to the link between AD and obesity and type 2 diabetes (*116*). In short, inflammation is part of AD development that cannot be ignored.

3. Use of mAbs in AD

A breakthrough in immunology was the hybridoma technology developed by Milstein and Kohler in 1975, allowing the production of unlimited quantities of mAbs (*117*). Due to their target specificity, mAbs have become powerful tools in biochemistry, molecular biology, and medicine today, and especially for preventing and treating AD. Passive immunization (antibodies) and active immunization (vaccines) are currently used clinically to inhibit or clear Aβ accumulation and hyperphosphorylated tau protein. In addition, antioxidants or anti-free radical drugs are also considered potential drugs for the induction of oxidative damage in current targeted therapy (*118*) (Table 1).

3.1. mAbs against the Aβ proteins

3.1.1. mAbs reduce Aβ aggregation

Therapies targeting Aβ have been the focus of AD treatment for the past 30 years, along with passive immunization using exogenous mAbs (*119*). Over the past few years, Aβ immunotherapy has mainly included γ-synthase inhibitors, γ-secretase and β-secretase inhibitors, and Aβ aggregation inhibitors. These inhibitors are designed to reduce Aβ formation and promote Aβ degradation in the brain, that is, to regulate the cell signaling mechanism, promote the α-secretase pathway, and inhibit the β-secretase pathway. mAb drugs targeting Aβ proteins have been extensively developed and have undergone human trials over the past 15 years.

Drugs for Aβ clearance include aducanumab, lecanemab, donanemab, crenezumab, solanezumab, gantenerumab, bapineuzumab, and GSK933776. Their targets are different stages of Aβ formation: monomers, oligomers, and plaques. Due to differences in targeted epitopes, drugs have differing ability to bind to and clear different forms of Aβ (*3*). Aβ is highly heterogeneous. Broadly speaking, the N-terminus of Aβ is exposed during aggregation and folding, while the C-terminus is hidden inside (*120*). Therefore, most of the drugs target epitopes at the N-terminus of Aβ, *e.g.*, aducanumab (3-7), lecanemab (1-16), and donanemab (p3-7). Aducanumab and lecanemab can bind equally to monomers, oligomers, and plaques of Aβ, while donanemab is designed to target pyroglutamate-modification of Aβ(AβN3pE) found almost exclusively in Aβ plaque. Thus, donanemab targets existing amyloid plaques. In addition, gantenerumab can bind two discontinuous regions of Aβ located at the N- and C-termini, respectively.

The strategy for selecting the N-terminal as an epitope was found to be feasible in advanced clinical trials. In two large randomized double-blind controlled phase 3 trials with more than 3,200 patients with early AD, results indicated a significant dose- and time-dependent reduction in pathophysiological markers but a low clinical response rate after aducanumab treatment (*121*). In contrast, lecanemab is designed to respond to the E22G (APP E693G) mutation, a pathogenic missense mutation that affects the twenty-second amino acid of Aβ peptides. Patients carrying the E22G mutation have a high level of Aβ protofibrils and lecanemab has a higher selectivity for it. An 18-month multicenter double-blind phase 3 trial including 1,795 subjects with early AD found that intravenous lecanemab every two weeks (10 mg/kg) significantly reduced the brain amyloid burden and score on the 14-item cognitive subscale of the AD Assessment Scale (ADAS-cog14). However, the sum of boxes on the Clinical Dementia Rating (CDR-SB) scale, which was the primary outcome, did not change markedly (*122*). A study reported infusion-related reactions in 26.4% of subjects and amyloid-related imaging abnormalities (ARIA) with edema in 12.6% (*122*). In an open-label trial, disease progression in patients with early AD can be mitigated by lecanemab over 24 months after the drug was discontinued (*123*). The study also suggested that lecanemab is better at maximum plaque removal, regardless of the ApoE4 status (*123*). Even APOEε4 carriers might respond better to lecanemab (*120*). In addition, TRAILBLAZER-ALZ2, a recent 76-week phase 3 randomized double-blind parallel multicenter placebo-controlled trial, suggested that donanemab significantly slowed AD progression in both low/medium tau and high tau populations (*124*). In the low/medium tau patients in particular, the clinical outcome was achieved in 52% of subjects based on amyloid clearance criteria; what is exciting is that a pronounced improvement in cognition was observed (*124*). This may be related to the high selectivity of donanemab for plaque.

Unlike the three aforementioned drugs that target the N-terminus, solanezumab and crenezumab target mid-sequence epitopes of amyloid fibrils. Solanezumab (LY2062430), which has a Fab fragment that can specifically bind to amino acid residues 16-26 of the Aβ protein, recognizes monomeric Aβ. It is supposed to capture and eliminate peripheral and central Aβ proteins, leading to the degradation of Aβ protein plaques (*125*). A phase 2 trial suggested that solanezumab seemed to shift Aβ equilibria to mobilize Aβ(1-42) from amyloid plaques, accompanied by increased unbound Aβ(1-42) in cerebrospinal fluid (CSF) in a dose-dependent manner (*126*). However, the cognitive scores of the trial's subjects did not change significantly, which was consistent with the results of a phase 1 trial (*127*). However, several subsequent phase 3 clinical trials were terminated after failing to demonstrate clinical efficacy (*128-130*). Crenezumab has a high affinity for higher molecular weight species such as fibrils, plaques, and oligomers but

Table 1. Monoclonal antibodies against Aβ and tau proteins in clinical trials (current as of February 2, 2024)

Drugs	Category	Major targets	Study phases	Subjects	Effectiveness	Main adverse events	*Ref.*
Aducanumab	Removes Aβ	Aβ multimers	3	Early AD	Slowing cognitive decline; improved markers of amyloid	ARIA	Budd *et al*., 2022
Lecanemab	Removes Aβ	Aβ oligomers	3	Early AD	Improved markers of amyloid	ARIA	van Dyck *et al*., 2023
Gantenerumab	Removes Aβ	Aβ multimers or monomers	3	Early AD	Reduced Aβ plaque	ARIA	Bateman *et al*., 2023
Crenezumab	Removes Aβ	Aβ multimers or monomers	3	Early AD	Neither cognitive decline nor markers of amyloid	Rare, mild ARIA	Ostrowitzki *et al*., 2022
Solanezumab	Removes Aβ	Aβ monomers	3	Preclinical AD	Neither cognitive decline nor markers of amyloid	ARIA with microhemorrhage or hemosiderosis	Sperling *et al*., 2023
Bapineuzumab	Removes Aβ	soluble Aβ; Aβ protein	3	Mild-to-moderate AD	Neither cognitive decline	Rare ARIA; falls, agitation, and urinary tract infections	Salloway *et al*., 2018
GSK933776	Removes Aβ	soluble Aβ	1	Mild AD	Improved markers of amyloid	Increased levels of blood creatine phosphokinase	Andreasen *et al*., 2015
Donanemab	Removes Aβ	Aβ plaque	3	Early AD	Slowing cognitive decline; Reduced Aβ plaque	ARIA with microhemorrhages and hemosiderin	Sims *et al*., 2023
Semagacestat	Reduces Aβ production	γ-secretase	3	Mild-to-moderate AD	Exacerbated cognitive impairment	Skin cancer and infections	Doody *et al*., 2013
Verubecestat	Reduces Aβ production	BACE1	3	Mild-to-moderate AD	Improved markers of amyloid	Rashes and hair color changes	Egan *et al*., 2018
Atabecestat	Reduces Aβ production	BACE1/2	2b/3	Preclinical AD	Exacerbated cognitive impairment	Liver toxicity	Sperling *et al*., 2021
Lanabecestat	Reduces Aβ production	BACE1/2	3	Early or mild AD	Improved markers of amyloid	Psychiatric adverse events, weight loss, and hair color changes	Wessels *et al*., 2020
JNJ-63733657	Inhibits aggregation	Thr217	2	Early AD	Not yet determined	Not yet determined	Janssen Research & Development, LLC, 2020
Zagotenemab	Inhibits aggregation	aggregated misfolded tau	2	Early AD	Not yet determined	Sinus bradycardia, headaches, falls, and bronchitis	Willis *et al*., 2023
E2814	Inhibits aggregation	microtubule-binding domain	2	Early AD	Not yet determined	Not yet determined	Eisai Inc., 2021
Gosuranemab	Eliminates tau aggregates	tau monomers and fibrils	2	Early AD	Improved markers of amyloid	Falls, nasopharyngitis, arthralgia, headaches, diarrhea, and constipation	Shulman *et al*., 2023
Semorinemab	Eliminates tau aggregates	full-length tau	2	Mild-to-moderate AD	Neither cognitive decline nor markers of amyloid	Falls, nasopharyngitis, and injection-related reactions	Teng *et al*., 2022

Not yet determined: Clinical results are not yet determined because the trial has not concluded.

a low affinity for monomers (*120*). Notably, crenezumab causes a small incidence of ARIA, which may be related to its selection of the IgG4 backbone isotype. In a multicenter, double-blind 1b study, most adverse events were mild or moderate, and ARIA with edema was not reported, indicating that it is well-tolerated and safe (*131*). The same result was observed in two phase 3 multicenter randomized double-blind placebo-controlled parallel-group trials, which found that ARIA with edema is rare, mild, and transient (*132*). Unfortunately, however, there were no significant changes in clinical outcomes or biomarkers (*132*). However, a phase 2 trial of high-dose crenezumab in patients with mild AD yielded better results (*133*).

Gantenerumab mainly targets Aβ plaques and removes them *via* Fcγ receptor-mediated microglial phagocytosis (*134,135*). It causes little significant discomfort when injected and appears to help with the clinical progression of a high-dose and low-frequency strategy (*136,137*). A monthly high subcutaneous dose of 1,200 mg gantenerumab demonstrated acceptable long-term safety and robust plaque clearance in patients with prodromal to moderate AD; after two years of treatment, 51% of patients had sparse-to-no neuritic Aβ plaques with clinical decline trending in the same direction, suggesting its potential clinical benefit (*138*). However, a study with a large sample is necessary for validation.

ARIA is the most common adverse reaction to Aβ-clearing antibodies, which may be related to the stress-induced immune clearance of antibodies. Insoluble Aβ is present in the brains of both healthy older adults and patients with AD. mAbs also recognize it in addition to the toxic soluble Aβ and activate monocyte and lymphocyte recruitment and clearance (*139*). In this process, the antibody conjugate stimulates the expression of macrophage proteases through the Fc receptor, which degrades the extracellular matrix, disrupting the BBB and entry of tissue fluid into the brain. Ultimately, it manifests as ARIA with edema or hemorrhage (*140*). Bapineuzumab (AAB-001), a long-acting antibody administered in 13-week cycles, can clear Aβ by binding to 5 N-terminal residues of the Aβ protein (*141*). Still, it has an unsatisfactory incidence of AIRA and associated symptoms such as headaches (*142*). AAB-003, a derivative of bapineuzumab, has three amino acid residues in the lower hinge region of bapineuzumab to reduce inflammatory activation and cellular damage by the Fc-receptor, reducing the risk of ARIA and increasing the safe dosage. A first-in-human study of AAB-003 evaluated its safety, tolerability, and pharmacokinetic data at five dose levels (0.5, 1, 2, 4, and 8 mg/kg) (*142*). Results indicated that AAB-003 was safe and well-tolerated in patients with mild to moderate AD at up to 8 mg/kg for up to 91 weeks (*143*), with similar adverse events to bapineuzumab after the first or second injection. GSK933776, another drug that targets the N-terminus of the Aβ protein (amino acid disabled sequences 1-5) and that provides passive immunity, is thought to reduce the incidence of ARIA since it contains a variant amino acid sequence that substantially reduces the antibody's effect on the Fc region. In a two-part placebo-controlled first-in-human study in patients with mild AD, total plasma Aβ levels increased after single-dose and repeated-dose intravenous administration decreased free Aβ levels in a dose-dependent manner (*144*). No subjects in any of the dose groups had drug-related ARIA with edema or hemorrhage (*144*). Further studies are needed to prove the safety and clinical efficacy of GSK933776 in patients with AD. Put simply, reducing ARIA during antibody treatment by inhibiting the Fc receptor seems desirable.

3.1.2. mAbs reduce Aβ production

In addition to removing Aβ, counteracting amyloid deposition also reduces its production. As mentioned earlier, Aβ production requires the involvement of γ-secretase and β-secretase. Therefore, mAbs have been designed against γ-secretase inhibitors, like semagacestat, and β-secretase inhibitors, like verubecestat, atabecestat, and lanabecestat. Semagacestat reduces Aβ40 and Aβ42 production and secretion from its substrate APP, which seems to get to the root of the problem. Unfortunately, a phase 3 trial found that semagacestat led to a worse outcome and was associated with more adverse events, including skin cancer and infections, leading to the termination of the drug's development (*145*). Semagacestat may not be an γ-secretase inhibitor because it does not inhibit intracellular levels of γ-byproducts (product peptides of γ-secretase) (*146*). Subsequent validation found that semagacestat inhibited the transport of γ-byproducts and Aβ to the extracellular compartment, leading to intracellular accumulation and cytotoxicity (*146*). Given its high toxicity and low efficacy, clinical studies of semagacestat have been discontinued.

Another target for Aβ production is the beta-site amyloid precursor protein cleaving enzyme (BACE), also known as β-secretase. Verubecestat is an inhibitor of the β-secretase enzyme, which cleaves APP proteins into various Aβ peptides. As an oral medication, it is envisioned as long-term maintenance therapy to limit Aβ production (*147*). A preliminary clinical trial demonstrated its safety - adverse reactions were primarily rashes and hair color changes rather than ARIA (*148*). Similarly, no adverse effects, such as neurodegeneration or altered glucose homeostasis, were observed in an animal study (*147*). A randomized placebo-controlled phase 3 study conducted at 238 centers in 21 countries and involving 1,958 patients with mild-to-moderate AD indicated that verubecestat reduced biomarkers (Aβ:40: 71.1-80.6%, Aβ:42: 62.7-76.4%, sAPPβ: 76.6-86.1%); however, it did not alleviate cognitive decline in patients (*148*). Lanabecestat is another orally administered mAb that targets both BACE1 and BACE2. Biomarker data indicated that lanabecestat reduced blood Aβ40 and Aβ42

levels by 70 to 80% in both trials. Aβ levels measured in CSF dropped by 50 and 73% at the low and high doses, respectively (*149*). In addition, lanabecestat reduced brain amyloid on PET imaging in a dose-dependent manner (*149*). However, there was no significant alleviation of clinical symptoms. The high dose caused more dropouts because of psychiatric adverse reactions, weight loss, and hair discoloration (*149*). In addition, atabecestat, another β-secretase inhibitor that has similar pharmacologic effects to verubecestat, has worse efficacy and causes more serious adverse events (*150,151*). The drug induces liver toxicity. In one subject who discontinued treatment because of elevated liver enzymes, a liver biopsy revealed inflammation, infiltration of immune T and B cells, and hepatocyte death (*152*). Subsequently, drug-responsive T cells were detected in the subject with liver injury; these cells were generated by binding to atabecestat or its metabolites to antigen-presenting cells (*153*). Eventually, the development of the drug was terminated. In conclusion, γ-secretase and β-secretase inhibitors are promising as oral antibodies for ultra-early therapeutic use in AD. However, efficacy and dose-dependent adverse events need to be urgently addressed.

3.2. mAbs against the tau protein

Tau protein pathology has recently received more attention in AD treatment following the issues with Aβ immunotherapy (*154*). Tau pathology can be blocked in four ways: inhibition of phosphorylation/acetylation, inhibition of aggregation, elimination of tau aggregates, and promotion of microtubule stabilization (*120*). Development of mAbs focuses on inhibition of tau aggregation and removal of tau aggregates. The first study on passive immunization with tau used PHF1 antibodies against the pSer396/404 epitope in a mouse model of tauopathy (*155*). mAbs targeting tau may inhibit AD progression by retarding the accumulation of pathological tau. Table 1 summarizes the mAb drugs targeting tau protein in clinical trials over the past few years.

3.2.1. mAbs inhibit tau aggregation

Zagotenemab (LY3303560), derived from mouse mAb MCI-1, binds to and neutralizes soluble tau aggregates. MCI-1 is a conformationally selective anti-tau antibody that binds to an early pathological form of soluble tau conformation to avoid its accumulation (*156*). Thus, it might be useful for early prevention. A 16-week study evaluated the safety and tolerability of the doses of zagotenemab (70 mg or 210 mg, q4w with 49 weeks) in patients with AD and early cognitive impairment and healthy volunteers. Results revealed a dose-dependent increase in plasma tau concentration in the SUBJECTS study after zagotenemab was administered, but there was no significant alleviation of clinical manifestations. A phase 2 clinical study to evaluate its safety and efficacy in patients with early AD symptoms was completed in 2021 and yielded similar results (*157*).

JNJ-63733657 targets tau phosphorylated at Thr217, an epitope in the middle region of the tau protein. JNJ-63733657 recognizes the microtubule-binding region of tau and therefore interferes with the intercellular proliferation of pathogenic aggregated tau proteins more effectively than other antibodies. The latest phase 2 clinical trial to determine its safety and tolerability in patients with early AD is being conducted and is scheduled to conclude in 2035. The trial is divided into two parts: in the first part, healthy subjects received a single ascending dose of JN-63733657 or a placebo, and in the second part, patients with early AD received three doses of escalating intravenous injections for eight weeks (*158*). In addition, a study of JNJ-63733657 in healthy Chinese subjects and participants with early AD is currently underway (*159*).

E2814 targets a mid-range epitope in the microtubule-binding domain named HVPGG near the mid-structural domain of tau. This region is a major component of tau tangles and is involved in seeding and spreading pathogenic tau aggregates. E2814 is designed to bind extracellular tau, inhibit tau aggregation and seeding, prevent further accumulation of NFTs, and mediate microglia clearance (*58*). A phase 2 clinical trial, scheduled for 2024, is underway to evaluate the safety and tolerability of the drug administered intravenously to patients with dominant AD (*160*).

3.2.2. mAbs remove tau aggregates

Gosuranemab (BIIB092), a humanized IgG4 antibody that selectively binds to extracellular N-terminal tau fragments (residues 15-22), targets extracellular tau fragments, which may affect neurons and glial cells and seed neuropathology (*161,162*). A double-blind placebo-controlled parallel-group phase 2 trial involving subjects with mild cognitive impairment found it to be well-tolerated and safe overall; however, the trial was terminated based on a lack of effectiveness despite a significant reduction in the CSF levels of unbound N-terminal tau at 76 weeks (*163*). Just like gosuranemab targets the N-terminus of extracellular tau, semorinemab, a humanized IgG4 antibody, can bind to all forms of hyperphosphorylated and oligomeric tau with a high affinity and specificity (*164*). In preclinical studies, semorinemab reduced tau pathological changes in a transgenic mouse model. However, in a 73-week phase 2 randomized clinical trial, semorinemab did not prevent disease progression in patients with mild AD compared to a placebo. Nonetheless, it did have acceptable and well-tolerated safety (*164*). Data revealed the drug's favorable safety profile, and the most common adverse events included falls, nasopharyngitis, and injection-related reactions (*164,165*). Another phase 2 clinical trial

noted a 42.2% reduction in the rate of decline according to the 11-item Alzheimer's Disease Assessment Scale–Cognitive subscale (ADAS-Cog11) in the semorinemab treatment group compared to the placebo group.

4. Influence of the BBB

BBB, a highly selective semipermeable membrane structural and chemical barrier between the peripheral circulation and the central nervous system, mainly consists of capillary endothelial cells, pericytes, astrocytes, neurons, and tight junctions (*166,167*). The barrier prevents substances from reaching the brain and it stops specific macromolecules from entering the blood (*168*). Issues with AD treatment are the lack of effective therapeutic molecules as well as the difficulty of penetrating the BBB and reaching specific targets for disease treatment (*169*). More than 98% of small-molecule drugs and almost 100% of large-molecule medicines were precluded from reaching the brain during treatment (*170,171*). Conventional mAbs in particular cross less than 0.5% of the BBB (*69*).

Molecules are transported across the BBB in multiple ways. Some small hydrophilic and lipophilic molecules can enter the brain tissue by paracellular and extracellular diffusion (*172*). Other molecular substances that cannot diffuse through the cell membrane, such as glucose, amino acids, and nucleosides, can enter the brain *via* carrier-mediated transport systems, receptor-mediated endocytosis, and adsorption-mediated exocytosis (*173*). Like ordinary large-molecule substances, the passage of mAbs across the BBB is mainly accomplished by endogenous transport on endothelial cells, including adsorption-mediated cytosis, carrier-mediated cytosis (CMT), and receptor-mediated cytosis (RMT). Cytosis is a standard mode of drug uptake. Upon a receptor or carrier's recognition of a ligand signal, the cell membrane invaginates and encapsulates the drug and detaches from the plasma membrane to form a vesicle. Then, the drug is digested intracellularly by lysosomes and released into the brain; at the same time, the receptor and ligand can be transported back to the original plasma membrane by carrier vesicles for reuse (*174*). Several carrier proteins hold promise for research, including transferrin receptor (TfR), insulin receptor, and melanin transferrin. Trontinemab, a new drug for AD, consists of gantenerumab and a human TfR1-directed Brainshuttle™ module. *In vitro* experiments have indicated that it has a similar capacity to bind to Aβ fibers and plaques as gantenerumab. Animal experiments have indicated better brain and plasma pharmacokinetic parameters, approximately 4-18 times better than gantenerumab, according to nonlinear mixed-effects modeling with correction for tissue residual blood (*175*). In a 44-person study, trontinemab markedly cleared plaque in three-quarters of participants within six months with negligible to no AIRA (*176*). Pharmacokinetics indicated the concentration of mAbs in the brain is expected to be much lower than that in plasma because peripherally administered mAbs have difficulty crossing the BBB, leading to low safety and efficacy in AD treatment.

Nanoparticles have been widely used as auxiliary tools to treat neurological diseases because their physicochemical properties and multifunctionalities enable them to cross the BBB (*177*). They can work as a carrier to treat several contributors to AD, including tau pathology, Aβ accumulation, and AD-associated neuroinflammation (*178*). Several nanoparticles (solid lipid particles, dendrimers, nanofibers, nanotubes, PLA/PLGA NPs, *etc.*) are currently being developed for preclinical or biomedical use (*179*). Advantageously, nanoparticles are more likely to target specific tissues by covalently binding to various ligands (*180*); a liposome, a highly flexible and biocompatible drug delivery system, has the potential to carry biologically active molecules, effectively improving the bioavailability of drugs. A phase 1b randomized clinical trial indicated that liposome-based anti-amyloid ACI-24 has good safety and specific efficiency in treating Down syndrome. There was no ARIA with edema or cerebral microhemorrhage, and increases in anti-Aβ immunoglobulin G titers were observed in 4 of 12 participants (33.3%) receiving ACI-24 (*181*). Intranasal drug delivery has been developed and utilized since first-pass metabolism, systemic clearance, and enzymatic degradation of drugs are other essential factors affecting efficacy and safety. Intranasal administration brings the drug to the brain in two ways. One is absorption through the venous vessels of the nose into the cavernous sinus and subsequently directly into the brain's arteries. This approach, although inevitably passing through the BBB, avoids the first-pass metabolism and degradation (*182*). The other way is diffusion through the nervous system, including intracellular and extracellular pathways. The drug passes through the olfactory epithelium trigeminal nerve and olfactory nerve receptors to reach the nerve endings and enters the neurons by pinocytosis. It is encapsulated in vesicles as exosomes, translocated to axon terminals, and discharged into the postsynaptic cells of the olfactory bulb. There, the lipophilic molecules can exert their pharmacological action by reaching the lamina propria *via* transcellular transport. Transportation of the drug to the brain in this way takes about 24 hours, while other peptides and hydrophilic drugs can reach the central system *via* the extracellular pathway in less than 30 minutes (*183*). In this way, drugs diffuse through the perineural space between the olfactory neuron-ensheathing cells and the olfactory neural fibroblasts. This space continues from the olfactory epithelium to the olfactory bulb and, more importantly, connects to the subarachnoid space, meaning that the drug directly enters the CSF (*183*). Animal studies have indicated that nanoparticles for nasal administration require a lower dosage and have a higher bioavailability than those for

conventional oral administration, which means they are safer and more efficient (*184,185*). Regrettably, however, there appear to be no clinical trials on nanoparticle drugs yet.

5. Conclusions and prospects

Over the past few decades, our understanding of the pathogenesis, diagnosis, and treatment of AD and other related neurodegenerative diseases has improved significantly. Nowadays, the higher incidence of AD and its impact on patients and their families has undoubtedly attracted more attention from researchers. Brain targets and biomarkers have been gradually discovered, significantly promoting the diagnosis and treatment of AD. mAbs have emerged as a promising tool for precisely binding drugs to their targets, blocking or modulating their effects, reducing the intensity of the immune response, and potentially eliminating the pathological process of AD. However, clinical trial data suggests that while these therapies can reduce AD-related markers such as Aβ, they do not consistently produce satisfactory clinical outcomes. This may be linked to the complex and incompletely understood pathology of AD. Recently, the question has been raised as to whether the priority of the Aβ hypothesis in the progression of AD needs to be reconsidered (*45*).

Immunoprophylaxis is receiving more attention based on setbacks in immunotherapy. AD presents with intracerebral pathology, such as accumulation of Aβ, 20-30 years before the onset of cognitive impairment (*186,187*). Given the growing pathologic complexity of the disease, the clinical benefit that can be expected from Aβ-removing therapies alone is unclear at this point in time. Thus, targeting aberrant Aβ as immunoprevention is likely to be most successful when initiated during or prior to the stage characterized by the emergence and seeded propagation of aberrant Aβ and Aβ-associated pathologies (*3*). As anti-Aβ therapy, KHK6640, a novel anti-amyloid beta oligomer-specific antibody, tended to perform well in the treatment of patients included prodromal AD, with nonsignificant AIRA (*188*). In contrast, anti-tau mAbs have made better progress. A phase 1 study demonstrated that atabecestat reduces Aβ isoforms and precursor protein (sAPPβ) in the CSF of patients with preclinical AD by an average of 67% (*189*). Semorinemab is safe but not effective (*164*). In addition, several trials on antibodies for AD are currently in the first phase to demonstrate the effectiveness of immunoprophylaxis (*190,191*).

In addition, the existence of the BBB poses enormous challenges to AD treatment. Antibodies are highly attractive due to their target specificity, long serum half-life, precise mechanism of action, and limited target deficiencies, compared to small molecule drugs for treating neurological diseases. However, crossing the BBB and penetration efficiency is a major problem in developing monoclonal drugs for AD treatment. In addition, the efficiency of their penetration should not be overlooked. If the concentration of mAbs is too high in the periphery, this will cause adverse events such as AIRA. Therefore, in-depth pharmacokinetic studies are necessary. Researchers have proposed using bifunctional IgG fusion proteins, nanoparticles, and other methods to address these issues. Nasal administration has also been used to bypass hepatic first-pass metabolism and degradation by gastrointestinal enzymes. However, further clinical data on these approaches are needed. Although most therapeutic agents can prevent disease progression in preclinical studies, expected results were not yielded in clinical trials. Most drugs have only reached clinical phase 1 or 2 trials, and many of those that undergo phase 3 clinical trials are terminated due to poor efficacy. Although the FDA hastily approved aducanumab without conclusive clinical evidence, this may encourage patients with AD, indicating that drug development is advancing (*192*). However, we must remember that there may be a long way to go.

Funding: None.

Conflict of Interest: The authors have no conflicts of interest to disclose.

References

1. Nasb M, Tao W, Chen N. Alzheimer's disease puzzle: Delving into pathogenesis hypotheses. Aging Dis. 2024; 15:43-73.
2. Shah A, Kishore U, Shastri A. Complement system in Alzheimer's disease. Int J Mol Sci. 2021; 22:13647.
3. Jucker M, Walker LC. Alzheimer's disease: From immunotherapy to immunoprevention. Cell. 2023; 186:4260-4270.
4. Scheltens P, De Strooper B, Kivipelto M, Holstege H, Chételat G, Teunissen CE, Cummings J, van der Flier WM. Alzheimer's disease. Lancet. 2021; 397:1577-1590.
5. Khan S, Barve KH, Kumar MS. Recent advancements in pathogenesis, diagnostics and treatment of Alzheimer's disease. Curr Neuropharmacol. 2020; 18:1106-1125.
6. Pedrioli A, Oxenius A. Single B cell technologies for monoclonal antibody discovery. Trends Immunol 2021; 42:1143-1158.
7. Pardridge WM. Treatment of Alzheimer's disease and blood-brain barrier drug delivery. Pharmaceuticals (Basel). 2020; 13:394.
8. Imbimbo BP, Ippati S, Watling M, Imbimbo C. Role of monomeric amyloid-β in cognitive performance in Alzheimer's disease: Insights from clinical trials with secretase inhibitors and monoclonal antibodies. Pharmacol Res. 2023; 187:106631.
9. Jeremic D, Navarro-López JD, Jiménez-Díaz L. Efficacy and safety of anti-amyloid-β monoclonal antibodies in current Alzheimer's disease phase III clinical trials: A systematic review and interactive web app-based meta-analysis. Ageing Res Rev. 2023; 90:102012.
10. Ramanan VK, Armstrong MJ, Choudhury P, Coerver KA, Hamilton RH, Klein BC, Wolk DA, Wessels SR,

Jones LK, Jr. Antiamyloid monoclonal antibody therapy for Alzheimer disease: Emerging issues in neurology. Neurology. 2023; 101:842-852.
11. Vitek GE, Decourt B, Sabbagh MN. Lecanemab (BAN2401): An anti-beta-amyloid monoclonal antibody for the treatment of Alzheimer disease. Expert Opin Investig Drugs. 2023; 32:89-94.
12. Huang J, Beach P, Bozoki A, Zhu DC. Alzheimer's disease progressively reduces visual functional network connectivity. J Alzheimers Dis Rep. 2021; 5:549-562.
13. Fakhoury M. Microglia and astrocytes in Alzheimer's disease: Implications for therapy. Curr Neuropharmacol. 2018; 16:508-518.
14. Ju Y, Tam KY. Pathological mechanisms and therapeutic strategies for Alzheimer's disease. Neural Regen Res. 2022; 17:543-549.
15. Serrano-Pozo A, Growdon JH. Is Alzheimer's disease risk modifiable? J Alzheimers Dis. 2019; 67:795-819.
16. Scheyer O, Rahman A, Hristov H, Berkowitz C, Isaacson RS, Diaz Brinton R, Mosconi L. Female sex and Alzheimer's risk: The menopause connection. J Prev Alzheimers Dis. 2018; 5:225-230.
17. Irwin MR, Vitiello MV. Implications of sleep disturbance and inflammation for Alzheimer's disease dementia. Lancet Neurol. 2019; 18:296-306.
18. Gouras GK, Olsson TT, Hansson O. β-Amyloid peptides and amyloid plaques in Alzheimer's disease. Neurotherapeutics. 2015; 12:3-11.
19. Cisternas P, Zolezzi JM, Lindsay C, Rivera DS, Martinez A, Bozinovic F, Inestrosa NC. New insights into the spontaneous human Alzheimer's disease-like model Octodon degus: Unraveling amyloid-β peptide aggregation and age-related amyloid pathology. J Alzheimers Dis. 2018; 66:1145-1163.
20. Reinsfelt B, Westerlind A, Blennow K, Zetterberg H, Ricksten SE. Open-heart surgery increases cerebrospinal fluid levels of Alzheimer-associated amyloid β. Acta Anaesthesiol Scand. 2013; 57:82-88.
21. Chételat G, La Joie R, Villain N, Perrotin A, de La Sayette V, Eustache F, Vandenberghe R. Amyloid imaging in cognitively normal individuals, at-risk populations and preclinical Alzheimer's disease. Neuroimage Clin. 2013; 2:356-365.
22. Shang D, Hong Y, Xie W, Tu Z, Xu J. Interleukin-1β drives cellular senescence of rat astrocytes induced by oligomerized amyloid β peptide and oxidative stress. Front Neurol. 2020; 11:929.
23. He N, Jin WL, Lok KH, Wang Y, Yin M, Wang ZJ. Amyloid-β(1-42) oligomer accelerates senescence in adult hippocampal neural stem/progenitor cells *via* formylpeptide receptor 2. Cell Death Dis. 2013; 4:e924.
24. Chen JX, Yan SS. Role of mitochondrial amyloid-beta in Alzheimer's disease. J Alzheimers Dis. 2010; 20 Suppl 2:S569-578.
25. O'Brien RJ, Wong PC. Amyloid precursor protein processing and Alzheimer's disease. Annu Rev Neurosci. 2011; 34:185-204.
26. Tiwari S, Atluri V, Kaushik A, Yndart A, Nair M. Alzheimer's disease: Pathogenesis, diagnostics, and therapeutics. Int J Nanomedicine. 2019; 14:5541-5554.
27. Matuszyk MM, Garwood CJ, Ferraiuolo L, Simpson JE, Staniforth RA, Wharton SB. Biological and methodological complexities of beta-amyloid peptide: Implications for Alzheimer's disease research. J Neurochem. 2022; 160:434-453.
28. Takuma K, Yao J, Huang J, Xu H, Chen X, Luddy J, Trillat AC, Stern DM, Arancio O, Yan SS. ABAD enhances Abeta-induced cell stress *via* mitochondrial dysfunction. Faseb J. 2005; 19:597-598.
29. Chen Z, Zhong C. Oxidative stress in Alzheimer's disease. Neurosci Bull. 2014; 30:271-281.
30. Latulippe J, Lotito D, Murby D. A mathematical model for the effects of amyloid beta on intracellular calcium. PLoS One. 2018; 13:e0202503.
31. Fagan AM, Head D, Shah AR, Marcus D, Mintun M, Morris JC, Holtzman DM. Decreased cerebrospinal fluid Abeta(42) correlates with brain atrophy in cognitively normal elderly. Ann Neurol. 2009; 65:176-183.
32. Hong W, Wang Z, Liu W, O'Malley TT, Jin M, Willem M, Haass C, Frosch MP, Walsh DM. Diffusible, highly bioactive oligomers represent a critical minority of soluble Aβ in Alzheimer's disease brain. Acta Neuropathol. 2018; 136:19-40.
33. Espay AJ, Sturchio A, Schneider LS, Ezzat K. Soluble amyloid-β consumption in Alzheimer's disease. J Alzheimers Dis. 2021; 82:1403-1415.
34. Selkoe DJ, Hardy J. The amyloid hypothesis of Alzheimer's disease at 25 years. EMBO Mol Med. 2016; 8:595-608.
35. Spuch C, Ortolano S, Navarro C. New insights in the amyloid-beta interaction with mitochondria. J Aging Res. 2012; 2012:324968.
36. Chen JX, Yan SD. Amyloid-beta-induced mitochondrial dysfunction. J Alzheimers Dis. 2007; 12:177-184.
37. Streit WJ, Khoshbouei H, Bechmann I. The role of microglia in sporadic Alzheimer's disease. J Alzheimers Dis. 2021; 79:961-968.
38. Arnsten AFT, Datta D, Del Tredici K, Braak H. Hypothesis: Tau pathology is an initiating factor in sporadic Alzheimer's disease. Alzheimers Dement. 2021; 17:115-124.
39. Raulin AC, Doss SV, Trottier ZA, Ikezu TC, Bu G, Liu CC. ApoE in Alzheimer's disease: Pathophysiology and therapeutic strategies. Mol Neurodegener. 2022; 17:72.
40. Deane R, Sagare A, Hamm K, Parisi M, Lane S, Finn MB, Holtzman DM, Zlokovic BV. ApoE isoform-specific disruption of amyloid beta peptide clearance from mouse brain. J Clin Invest. 2008; 118:4002-4013.
41. Liu CC, Zhao N, Fu Y, Wang N, Linares C, Tsai CW, Bu G. ApoE4 accelerates early seeding of amyloid pathology. Neuron. 2017; 96:1024-1032.e1023.
42. Marais AD. Apolipoprotein E in lipoprotein metabolism, health and cardiovascular disease. Pathology. 2019; 51:165-176.
43. Miners JS, Van Helmond Z, Chalmers K, Wilcock G, Love S, Kehoe PG. Decreased expression and activity of neprilysin in Alzheimer disease are associated with cerebral amyloid angiopathy. J Neuropathol Exp Neurol. 2006; 65:1012-1021.
44. Cook DG, Leverenz JB, McMillan PJ, Kulstad JJ, Ericksen S, Roth RA, Schellenberg GD, Jin LW, Kovacina KS, Craft S. Reduced hippocampal insulin-degrading enzyme in late-onset Alzheimer's disease is associated with the apolipoprotein E-epsilon4 allele. Am J Pathol. 2003; 162:313-319.
45. Karran E, De Strooper B. The amyloid hypothesis in Alzheimer disease: New insights from new therapeutics. Nat Rev Drug Discov. 2022; 21:306-318.
46. Brunello CA, Merezhko M, Uronen RL, Huttunen HJ. Mechanisms of secretion and spreading of pathological

tau protein. Cell Mol Life Sci. 2020; 77:1721-1744.
47. Marcus JN, Schachter J. Targeting post-translational modifications on tau as a therapeutic strategy for Alzheimer's disease. J Neurogenet. 2011; 25:127-133.
48. Claeysen S, Cochet M, Donneger R, Dumuis A, Bockaert J, Giannoni P. Alzheimer culprits: Cellular crossroads and interplay. Cell Signal. 2012; 24:1831-1840.
49. Meraz-Ríos MA, Lira-De León KI, Campos-Peña V, De Anda-Hernández MA, Mena-López R. Tau oligomers and aggregation in Alzheimer's disease. J Neurochem. 2010; 112:1353-1367.
50. Arima K. Ultrastructural characteristics of tau filaments in tauopathies: Immuno-electron microscopic demonstration of tau filaments in tauopathies. Neuropathology. 2006; 26:475-483.
51. Aiken J, Holzbaur ELF. Cytoskeletal regulation guides neuronal trafficking to effectively supply the synapse. Curr Biol. 2021; 31:R633-r650.
52. Hung SY, Fu WM. Drug candidates in clinical trials for Alzheimer's disease. J Biomed Sci. 2017; 24:47.
53. Sinsky J, Pichlerova K, Hanes J. Tau protein interaction partners and their roles in Alzheimer's disease and other tauopathies. Int J Mol Sci. 2021; 22:9207.
54. Sehar U, Rawat P, Reddy AP, Kopel J, Reddy PH. Amyloid beta in aging and Alzheimer's disease. Int J Mol Sci. 2022; 23:12924.
55. Karimi N, Bayram Çatak F, Arslan E, Saghazadeh A, Rezaei N. Tau immunotherapy in Alzheimer's disease and progressive supranuclear palsy. Int Immunopharmacol. 2022; 113:109445.
56. Gao Y, Tan L, Yu JT, Tan L. Tau in Alzheimer's disease: Mechanisms and therapeutic strategies. Curr Alzheimer Res. 2018; 15:283-300.
57. Sexton C, Snyder H, Beher D, *et al*. Current directions in tau research: Highlights from tau 2020. Alzheimers Dement. 2022; 18:988-1007.
58. Roberts M, Sevastou I, Imaizumi Y, *et al*. Pre-clinical characterisation of E2814, a high-affinity antibody targeting the microtubule-binding repeat domain of tau for passive immunotherapy in Alzheimer's disease. Acta Neuropathol Commun. 2020; 8:13.
59. Strang KH, Golde TE, Giasson BI. MAPT mutations, tauopathy, and mechanisms of neurodegeneration. Lab Invest. 2019; 99:912-928.
60. Rauch JN, Luna G, Guzman E, Audouard M, Challis C, Sibih YE, Leshuk C, Hernandez I, Wegmann S, Hyman BT, Gradinaru V, Kampmann M, Kosik KS. LRP1 is a master regulator of tau uptake and spread. Nature. 2020; 580:381-385.
61. Xian X, Pohlkamp T, Durakoglugil MS, Wong CH, Beck JK, Lane-Donovan C, Plattner F, Herz J. Reversal of ApoE4-induced recycling block as a novel prevention approach for Alzheimer's disease. Elife. 2018; 7:e40048.
62. Xie J, Van Hoecke L, Vandenbroucke RE. The impact of systemic inflammation on Alzheimer's disease pathology. Front Immunol. 2021; 12:796867.
63. Musi N, Valentine JM, Sickora KR, Baeuerle E, Thompson CS, Shen Q, Orr ME. Tau protein aggregation is associated with cellular senescence in the brain. Aging Cell. 2018; 17:e12840.
64. Fan Z, Brooks DJ, Okello A, Edison P. An early and late peak in microglial activation in Alzheimer's disease trajectory. Brain. 2017; 140:792-803.
65. Nimmerjahn A, Kirchhoff F, Helmchen F. Resting microglial cells are highly dynamic surveillants of brain parenchyma *in vivo*. Science. 2005; 308:1314-1318.
66. Verkhratsky A, Nedergaard M. Physiology of astroglia. Physiol Rev. 2018; 98:239-389.
67. Imai F, Suzuki H, Oda J, Ninomiya T, Ono K, Sano H, Sawada M. Neuroprotective effect of exogenous microglia in global brain ischemia. J Cereb Blood Flow Metab. 2007; 27:488-500.
68. Dani M, Wood M, Mizoguchi R, Fan Z, Walker Z, Morgan R, Hinz R, Biju M, Kuruvilla T, Brooks DJ, Edison P. Microglial activation correlates *in vivo* with both tau and amyloid in Alzheimer's disease. Brain. 2018; 141:2740-2754.
69. Leng F, Edison P. Neuroinflammation and microglial activation in Alzheimer disease: Where do we go from here? Nat Rev Neurol. 2021; 17:157-172.
70. Dhana K, Franco OH, Ritz EM, Ford CN, Desai P, Krueger KR, Holland TM, Dhana A, Liu X, Aggarwal NT, Evans DA, Rajan KB. Healthy lifestyle and life expectancy with and without Alzheimer's dementia: Population based cohort study. BMJ. 2022; 377:e068390.
71. Bivona G, Iemmolo M, Agnello L, Lo Sasso B, Gambino CM, Giglio RV, Scazzone C, Ghersi G, Ciaccio M. Microglial activation and priming in Alzheimer's disease: State of the art and future perspectives. Int J Mol Sci. 2023; 24:884.
72. Guo T, Zhang D, Zeng Y, Huang TY, Xu H, Zhao Y. Molecular and cellular mechanisms underlying the pathogenesis of Alzheimer's disease. Mol Neurodegener. 2020; 15:40.
73. Sekiya M, Wang M, Fujisaki N, Sakakibara Y, Quan X, Ehrlich ME, De Jager PL, Bennett DA, Schadt EE, Gandy S, Ando K, Zhang B, Iijima KM. Integrated biology approach reveals molecular and pathological interactions among Alzheimer's Aβ42, Tau, TREM2, and TYROBP in Drosophila models. Genome Med. 2018; 10:26.
74. Ising C, Venegas C, Zhang S, *et al*. NLRP3 inflammasome activation drives tau pathology. Nature. 2019; 575:669-673.
75. Arranz AM, De Strooper B. The role of astroglia in Alzheimer's disease: Pathophysiology and clinical implications. Lancet Neurol. 2019; 18:406-414.
76. Singh D. Astrocytic and microglial cells as the modulators of neuroinflammation in Alzheimer's disease. J Neuroinflammation. 2022; 19:206.
77. Hansen DV, Hanson JE, Sheng M. Microglia in Alzheimer's disease. J Cell Biol. 2018; 217:459-472.
78. Nanclares C, Baraibar AM, Araque A, Kofuji P. Dysregulation of astrocyte-neuronal communication in Alzheimer's disease. Int J Mol Sci. 2021; 22:7887.
79. Wu T, Dejanovic B, Gandham VD, *et al*. Complement C3 is activated in human AD brain and is required for neurodegeneration in mouse models of amyloidosis and tauopathy. Cell Rep. 2019; 28:2111-2123.e2116.
80. Litvinchuk A, Wan YW, Swartzlander DB, Chen F, Cole A, Propson NE, Wang Q, Zhang B, Liu Z, Zheng H. Complement C3aR inactivation attenuates tau pathology and reverses an immune network deregulated in tauopathy models and Alzheimer's disease. Neuron. 2018; 100:1337-1353.e1335.
81. Dejanovic B, Huntley MA, De Mazière A, *et al*. Changes in the synaptic proteome in tauopathy and rescue of tau-Induced synapse loss by C1q antibodies. Neuron. 2018; 100:1322-1336.e1327.
82. Hong S, Beja-Glasser VF, Nfonoyim BM, Frouin A, Li S, Ramakrishnan S, Merry KM, Shi Q, Rosenthal A, Barres

BA, Lemere CA, Selkoe DJ, Stevens B. Complement and microglia mediate early synapse loss in Alzheimer mouse models. Science. 2016; 352:712-716.
83. Tenner AJ. Complement-mediated events in Alzheimer's disease: Mechanisms and potential therapeutic targets. J Immunol. 2020; 204:306-315.
84. Calvo-Rodriguez M, García-Rodríguez C, Villalobos C, Núñez L. Role of toll like receptor 4 in Alzheimer's disease. Front Immunol. 2020; 11:1588.
85. Guo H, Callaway JB, Ting JP. Inflammasomes: Mechanism of action, role in disease, and therapeutics. Nat Med. 2015; 21:677-687.
86. Swanson KV, Deng M, Ting JP. The NLRP3 inflammasome: Molecular activation and regulation to therapeutics. Nat Rev Immunol. 2019; 19:477-489.
87. Yang J, Wise L, Fukuchi KI. TLR4 cross-talk with NLRP3 inflammasome and complement signaling pathways in Alzheimer's disease. Front Immunol. 2020; 11:724.
88. Van Zeller M, Dias D, Sebastião AM, Valente CA. NLRP3 inflammasome: A starring role in amyloid-β- and tau-driven pathological events in Alzheimer's disease. J Alzheimers Dis. 2021; 83:939-961.
89. Venegas C, Kumar S, Franklin BS, *et al*. Microglia-derived ASC specks cross-seed amyloid-β in Alzheimer's disease. Nature. 2017; 552:355-361.
90. Zheng R, Yan Y, Dai S, Ruan Y, Chen Y, Hu C, Lin Z, Xue N, Song Z, Liu Y, Zhang B, Pu J. ASC specks exacerbate α-synuclein pathology *via* amplifying NLRP3 inflammasome activities. J Neuroinflammation. 2023; 20:26.
91. Zheng H, Jia L, Liu CC, Rong Z, Zhong L, Yang L, Chen XF, Fryer JD, Wang X, Zhang YW, Xu H, Bu G. TREM2 promotes microglial survival by activating wnt/β-catenin pathway. J Neurosci. 2017; 37:1772-1784.
92. Rueda-Carrasco J, Sokolova D, Lee SE, *et al*. Microglia-synapse engulfment *via* PtdSer-TREM2 ameliorates neuronal hyperactivity in Alzheimer's disease models. Embo j. 2023; 42:e113246.
93. Shi Q, Chang C, Saliba A, Bhat MA. Microglial mTOR activation upregulates Trem2 and enhances β-amyloid plaque clearance in the 5XFAD Alzheimer's disease model. J Neurosci. 2022; 42:5294-5313.
94. Sanjay, Shin JH, Park M, Lee HJ. Cyanidin-3-o-glucoside regulates the M1/M2 polarization of microglia *via* PPARγ and Aβ42 phagocytosis through TREM2 in an Alzheimer's disease model. Mol Neurobiol. 2022; 59:5135-5148.
95. Peng X, Guo H, Zhang X, Yang Z, Ruganzu JB, Yang Z, Wu X, Bi W, Ji S, Yang W. TREM2 inhibits tau hyperphosphorylation and neuronal apoptosis *via* the PI3K/Akt/GSK-3β signaling pathway *in vivo* and *in vitro*. Mol Neurobiol. 2023; 60:2470-2485.
96. Yin J, Liu X, He Q, Zhou L, Yuan Z, Zhao S. Vps35-dependent recycling of Trem2 regulates microglial function. Traffic. 2016; 17:1286-1296.
97. Yuan P, Condello C, Keene CD, Wang Y, Bird TD, Paul SM, Luo W, Colonna M, Baddeley D, Grutzendler J. TREM2 haplodeficiency in mice and humans impairs the microglia barrier function leading to decreased amyloid compaction and severe axonal dystrophy. Neuron. 2016; 92:252-264.
98. Zhao P, Xu Y, Jiang L, Fan X, Li L, Li X, Arase H, Zhao Y, Cao W, Zheng H, Xu H, Tong Q, Zhang N, An Z. A tetravalent TREM2 agonistic antibody reduced amyloid pathology in a mouse model of Alzheimer's disease. Sci Transl Med. 2022; 14:eabq0095.
99. Keren-Shaul H, Spinrad A, Weiner A, Matcovitch-Natan O, Dvir-Szternfeld R, Ulland TK, David E, Baruch K, Lara-Astaiso D, Toth B, Itzkovitz S, Colonna M, Schwartz M, Amit I. A unique microglia type associated with restricting development of Alzheimer's disease. Cell. 2017; 169:1276-1290.e1217.
100. Lin YT, Seo J, Gao F, *et al*. APOE4 causes widespread molecular and cellular alterations associated with Alzheimer's disease phenotypes in human iPSC-derived brain cell types. Neuron. 2018; 98:1141-1154.e1147.
101. Fan YY, Cai QL, Gao ZY, Lin X, Huang Q, Tang W, Liu JH. APOE ε4 allele elevates the expressions of inflammatory factors and promotes Alzheimer's disease progression: A comparative study based on Han and She populations in the Wenzhou area. Brain Res Bull. 2017; 132:39-43.
102. Wogram E, Prinz M. APOE set the microglia free. Nat Immunol. 2023; 24:1790-1791.
103. Liu C-C, Wang N, Chen Y, *et al*. Cell-autonomous effects of APOE4 in restricting microglial response in brain homeostasis and Alzheimer's disease. Nature Immunology. 2023; 24:1854-1866.
104. Simonovitch S, Schmukler E, Bespalko A, Iram T, Frenkel D, Holtzman DM, Masliah E, Michaelson DM, Pinkas-Kramarski R. Impaired autophagy in ApoE4 astrocytes. J Alzheimers Dis. 2016; 51:915-927.
105. Serrano-Pozo A, Das S, Hyman BT. APOE and Alzheimer's disease: Advances in genetics, pathophysiology, and therapeutic approaches. Lancet Neurol. 2021; 20:68-80.
106. Zhao N, Ren Y, Yamazaki Y, *et al*. Alzheimer's risk factors age, ApoE genotype, and sex drive distinct molecular pathways. Neuron. 2020; 106:727-742.e726.
107. Shcherbinin S, Evans CD, Lu M, Andersen SW, Pontecorvo MJ, Willis BA, Gueorguieva I, Hauck PM, Brooks DA, Mintun MA, Sims JR. Association of amyloid reduction after donanemab treatment with tau pathology and clinical outcomes: The TRAILBLAZER-ALZ randomized clinical trial. JAMA Neurol. 2022; 79:1015-1024.
108. Waring JF, Tang Q, Robieson WZ, King DP, Das U, Dubow J, Dutta S, Marek GJ, Gault LM. ApoE-ε4 carrier status and donepezil response in patients with Alzheimer's disease. J Alzheimers Dis. 2015; 47:137-148.
109. Hu Y, Fryatt GL, Ghorbani M, Obst J, Menassa DA, Martin-Estebane M, Muntslag TAO, Olmos-Alonso A, Guerrero-Carrasco M, Thomas D, Cragg MS, Gomez-Nicola D. Replicative senescence dictates the emergence of disease-associated microglia and contributes to Aβ pathology. Cell Rep. 2021; 35:109228.
110. Bisht K, Sharma KP, Lecours C, *et al*. Dark microglia: A new phenotype predominantly associated with pathological states. Glia. 2016; 64:826-839.
111. Lan L, Wang H, Zhang X, Shen Q, Li X, He L, Rong X, Peng J, Mo J, Peng Y. Chronic exposure of alcohol triggers microglia-mediated synaptic elimination inducing cognitive impairment. Exp Neurol. 2022; 353:114061.
112. Dolotov OV, Inozemtseva LS, Myasoedov NF, Grivennikov IA. Stress-induced depression and Alzheimer's disease: Focus on astrocytes. Int J Mol Sci. 2022; 23:4999.
113. Saha P, Guha S, Biswas SC. P38K and JNK pathways are induced by amyloid-β in astrocyte: Implication of MAPK pathways in astrogliosis in Alzheimer's disease. Mol Cell Neurosci. 2020; 108:103551.

114. Ghosh SS, Wang J, Yannie PJ, Ghosh S. Intestinal barrier dysfunction, LPS translocation, and disease development. J Endocr Soc. 2020; 4:bvz039.
115. Naomi R, Embong H, Othman F, Ghazi HF, Maruthey N, Bahari H. Probiotics for Alzheimer's disease: A systematic review. Nutrients. 2021; 14:20.
116. Litwiniuk A, Bik W, Kalisz M, Baranowska-Bik A. Inflammasome NLRP3 potentially links obesity-associated low-grade systemic inflammation and insulin resistance with Alzheimer's disease. Int J Mol Sci. 2021; 22:5603.
117. Van Hoecke L, Roose K. How mRNA therapeutics are entering the monoclonal antibody field. J Transl Med. 2019; 17:54.
118. Cummings J, Lee G, Nahed P, Kambar M, Zhong K, Fonseca J, Taghva K. Alzheimer's disease drug development pipeline: 2022. Alzheimers Dement (N Y). 2022; 8:e12295.
119. Van Dyck CH. Anti-amyloid-β monoclonal antibodies for Alzheimer's disease: Pitfalls and promise. Biol Psychiatry. 2018; 83:311-319.
120. Panza F, Solfrizzi V, Daniele A, Lozupone M. Passive tau-based immunotherapy for tauopathies. Handb Clin Neurol. 2023; 196:611-619.
121. Budd Haeberlein S, Aisen PS, Barkhof F, *et al*. Two randomized phase 3 studies of aducanumab in early Alzheimer's disease. J Prev Alzheimers Dis. 2022; 9:197-210.
122. Van Dyck CH, Swanson CJ, Aisen P, *et al*. Lecanemab in early Alzheimer's disease. N Engl J Med. 2023; 388:9-21.
123. McDade E, Cummings JL, Dhadda S, Swanson CJ, Reyderman L, Kanekiyo M, Koyama A, Irizarry M, Kramer LD, Bateman RJ. Lecanemab in patients with early Alzheimer's disease: Detailed results on biomarker, cognitive, and clinical effects from the randomized and open-label extension of the phase 2 proof-of-concept study. Alzheimers Res Ther. 2022; 14:191.
124. Sims JR, Zimmer JA, Evans CD, *et al*. Donanemab in early symptomatic Alzheimer disease: The TRAILBLAZER-ALZ 2 randomized clinical trial. JAMA. 2023; 330:512-527.
125. Le Couteur DG, Hunter S, Brayne C. Solanezumab and the amyloid hypothesis for Alzheimer's disease. BMJ. 2016; 355:i6771.
126. Gold M. Phase II clinical trials of anti-amyloid β antibodies: When is enough, enough? Alzheimers Dement (N Y). 2017; 3:402-409.
127. Siemers ER, Friedrich S, Dean RA, Gonzales CR, Farlow MR, Paul SM, Demattos RB. Safety and changes in plasma and cerebrospinal fluid amyloid beta after a single administration of an amyloid beta monoclonal antibody in subjects with Alzheimer disease. Clin Neuropharmacol. 2010; 33:67-73.
128. Honig LS, Vellas B, Woodward M, *et al*. Trial of solanezumab for mild dementia due to Alzheimer's disease. N Engl J Med. 2018; 378:321-330.
129. Sperling RA, Donohue MC, Raman R, *et al*. Trial of solanezumab in preclinical Alzheimer's disease. N Engl J Med. 2023; 389:1096-1107.
130. Doody RS, Thomas RG, Farlow M, Iwatsubo T, Vellas B, Joffe S, Kieburtz K, Raman R, Sun X, Aisen PS, Siemers E, Liu-Seifert H, Mohs R. Phase 3 trials of solanezumab for mild-to-moderate Alzheimer's disease. N Engl J Med. 2014; 370:311-321.
131. Guthrie H, Honig LS, Lin H, Sink KM, Blondeau K, Quartino A, Dolton M, Carrasco-Triguero M, Lian Q, Bittner T, Clayton D, Smith J, Ostrowitzki S. Safety, tolerability, and pharmacokinetics of crenezumab in patients with mild-to-moderate Alzheimer's disease treated with escalating doses for up to 133 weeks. J Alzheimers Dis. 2020; 76:967-979.
132. Ostrowitzki S, Bittner T, Sink KM, *et al*. Evaluating the safety and efficacy of crenezumab vs placebo in adults with early Alzheimer disease: Two phase 3 randomized placebo-controlled trials. JAMA Neurol. 2022; 79:1113-1121.
133. Cummings JL, Cohen S, van Dyck CH, *et al*. ABBY: A phase 2 randomized trial of crenezumab in mild to moderate Alzheimer disease. Neurology. 2018; 90:e1889-e1897.
134. Söderberg L, Johannesson M, Nygren P, Laudon H, Eriksson F, Osswald G, Möller C, Lannfelt L. Lecanemab, aducanumab, and gantenerumab - Binding profiles to different forms of amyloid—beta might explain efficacy and side effects in clinical trials for Alzheimer's disease. Neurotherapeutics. 2023; 20:195-206.
135. Bohrmann B, Baumann K, Benz J, *et al*. Gantenerumab: A novel human anti-Aβ antibody demonstrates sustained cerebral amyloid-β binding and elicits cell-mediated removal of human amyloid-β. J Alzheimers Dis. 2012; 28:49-69.
136. Portron A, Jordan P, Draper K, Muenzer C, Dickerson D, van Iersel T, Hofmann C. A phase I study to assess the effect of speed of injection on pain, tolerability, and pharmacokinetics after high-volume subcutaneous administration of gantenerumab in healthy volunteers. Clin Ther. 2020; 42:108-120.e101.
137. Bateman RJ, Smith J, Donohue MC, *et al*. Two phase 3 trials of gantenerumab in early Alzheimer's disease. N Engl J Med. 2023; 389:1862-1876.
138. Klein G, Delmar P, Voyle N, Rehal S, Hofmann C, Abi-Saab D, Andjelkovic M, Ristic S, Wang G, Bateman R, Kerchner GA, Baudler M, Fontoura P, Doody R. Gantenerumab reduces amyloid-β plaques in patients with prodromal to moderate Alzheimer's disease: A PET substudy interim analysis. Alzheimers Res Ther. 2019; 11:101.
139. Tagliavini F, Giaccone G, Frangione B, Bugiani O. Preamyloid deposits in the cerebral cortex of patients with Alzheimer's disease and nondemented individuals. Neurosci Lett. 1988; 93:191-196.
140. Taylor X, Clark IM, Fitzgerald GJ, Oluoch H, Hole JT, DeMattos RB, Wang Y, Pan F. Amyloid-β (Aβ) immunotherapy induced microhemorrhages are associated with activated perivascular macrophages and peripheral monocyte recruitment in Alzheimer's disease mice. Mol Neurodegener. 2023; 18:59.
141. Miles LA, Crespi GA, Doughty L, Parker MW. Bapineuzumab captures the N-terminus of the Alzheimer's disease amyloid-beta peptide in a helical conformation. Sci Rep. 2013; 3:1302.
142. Salloway S, Marshall GA, Lu M, Brashear HR. Long-term safety and efficacy of bapineuzumab in patients with mild-to-moderate Alzheimer's disease: A phase 2, open-label extension study. Curr Alzheimer Res. 2018; 15:1231-1243.
143. Delnomdedieu M, Duvvuri S, Li DJ, Atassi N, Lu M, Brashear HR, Liu E, Ness S, Kupiec JW. First-in-human safety and long-term exposure data for AAB-003 (PF-05236812) and biomarkers after intravenous infusions of escalating doses in patients with mild to moderate

Alzheimer's disease. Alzheimers Res Ther. 2016; 8:12.
144. Andreasen N, Simeoni M, Ostlund H, *et al*. First administration of the Fc-attenuated anti-β amyloid antibody GSK933776 to patients with mild Alzheimer's disease: A randomized, placebo-controlled study. PLoS One. 2015; 10:e0098153.
145. Doody RS, Raman R, Farlow M, Iwatsubo T, Vellas B, Joffe S, Kieburtz K, He F, Sun X, Thomas RG, Aisen PS, Siemers E, Sethuraman G, Mohs R. A phase 3 trial of semagacestat for treatment of Alzheimer's disease. N Engl J Med. 2013; 369:341-350.
146. Tagami S, Yanagida K, Kodama TS, *et al*. Semagacestat is a pseudo-inhibitor of γ-secretase. Cell Rep. 2017; 21:259-273.
147. Kennedy ME, Stamford AW, Chen X, *et al*. The BACE1 inhibitor verubecestat (MK-8931) reduces CNS β-amyloid in animal models and in Alzheimer's disease patients. Sci Transl Med. 2016; 8:363ra150.
148. Egan MF, Kost J, Tariot PN, *et al*. Randomized trial of verubecestat for mild-to-moderate Alzheimer's disease. N Engl J Med. 2018; 378:1691-1703.
149. Wessels AM, Tariot PN, Zimmer JA, *et al*. Efficacy and safety of lanabecestat for treatment of early and mild Alzheimer disease: The AMARANTH and DAYBREAK-ALZ randomized clinical trials. JAMA Neurol. 2020; 77:199-209.
150. Sperling R, Henley D, Aisen PS, *et al*. Findings of efficacy, safety, and biomarker outcomes of atabecestat in preclinical Alzheimer disease: A truncated randomized phase 2b/3 clinical trial. JAMA Neurol. 2021; 78:293-301.
151. Papp KV, Rofael H, Veroff AE, *et al*. Sensitivity of the Preclinical Alzheimer's Cognitive Composite (PACC), PACC5, and Repeatable Battery for Neuropsychological Status (RBANS) to amyloid status in preclinical Alzheimer's disease - Atabecestat phase 2b/3 early clinical trial. J Prev Alzheimers Dis. 2022; 9:255-261.
152. De Jonghe S, Weinstock D, Aligo J, Washington K, Naisbitt D. Biopsy pathology and immunohistochemistry of a case of immune-mediated drug-induced liver injury with atabecestat. Hepatology. 2021; 73:452-455.
153. Thomson PJ, Kafu L, Meng X, Snoeys J, De Bondt A, De Maeyer D, Wils H, Leclercq L, Vinken P, Naisbitt DJ. Drug-specific T-cell responses in patients with liver injury following treatment with the BACE inhibitor atabecestat. Allergy. 2021; 76:1825-1835.
154. Valera E, Spencer B, Masliah E. Immunotherapeutic approaches targeting amyloid-β, α-synuclein, and tau for the treatment of neurodegenerative disorders. Neurotherapeutics. 2016; 13:179-189.
155. Boutajangout A, Ingadóttir J, Davies P, Sigurdsson EMJAs, Dementia. Passive tau immunotherapy diminishes functional decline and clears tau aggregates in a mouse model of tauopathy. In: Alzheimer's Dementia 2010; 4:S578-S578.
156. Weaver CL, Espinoza M, Kress Y, Davies P. Conformational change as one of the earliest alterations of tau in Alzheimer's disease. Neurobiol Aging. 2000; 21:719-727.
157. Eli Lilly and Company. A study of LY3303560 in participants with early symptomatic Alzheimer's disease. *https://clinicaltrials.gov/ct2/show/NCT03518073?term=LY3303560&draw=2&rank=3* (accessed December 16, 2023).
158. Janssen Research & Development, LLC. A study of JNJ-63733657 in participants with early Alzheimer's disease (Autonomy). *https://clinicaltrials.gov/ct2/show/NCT04619420?term=JNJ-63733657&draw=2&rank=2* (accessed October 18, 2023).
159. Janssen Research & Development, LLC. A study of JNJ-63733657 in healthy Chinese participants. *https://clinicaltrials.gov/ct2/show/NCT05407818?term=JNJ-63733657&draw=2&rank=1* (accessed February 2, 2024).
160. Eisai Inc. A study to assess safety and target engagement of E2814 in participants with mild to moderate cognitive impairment due to dominantly inherited Alzheimer's disease. *https://clinicaltrials.gov/ct2/show/NCT04971733?term=E2814&draw=2&rank=1* (accessed October 18, 2023).
161. Sopko R, Golonzhka O, Arndt J, *et al*. Characterization of tau binding by gosuranemab. Neurobiol Dis. 2020; 146:105120.
162. Yanamandra K, Kfoury N, Jiang H, Mahan TE, Ma S, Maloney SE, Wozniak DF, Diamond MI, Holtzman DM. Anti-tau antibodies that block tau aggregate seeding *in vitro* markedly decrease pathology and improve cognition *in vivo*. Neuron. 2013; 80:402-414.
163. Shulman M, Kong J, O'Gorman J, Ratti E, Rajagovindan R, Viollet L, Huang E, Sharma S, Racine AM, Czerkowicz J, Graham D, Li Y, Hering H, Haeberlein SB. TANGO: A placebo-controlled randomized phase 2 study of efficacy and safety of the anti-tau monoclonal antibody gosuranemab in early Alzheimer's disease. Nature Aging. 2023; 3:1591-1601.
164. Teng E, Manser PT, Pickthorn K, *et al*. Safety and efficacy of semorinemab in individuals with prodromal to mild alzheimer disease: A randomized clinical trial. JAMA Neurol. 2022; 79:758-767.
165. Ayalon G, Lee SH, Adolfsson O, *et al*. Antibody semorinemab reduces tau pathology in a transgenic mouse model and engages tau in patients with Alzheimer's disease. Sci Transl Med. 2021; 13:eabb2639.
166. Najjar S, Pahlajani S, De Sanctis V, Stern JNH, Najjar A, Chong D. Neurovascular unit dysfunction and blood-brain barrier hyperpermeability contribute to schizophrenia neurobiology: A theoretical integration of clinical and experimental evidence. Front Psychiatry. 2017; 8:83.
167. Xu L, Nirwane A, Yao Y. Basement membrane and blood-brain barrier. Stroke Vasc Neurol. 2019; 4:78-82.
168. Sweeney MD, Sagare AP, Zlokovic BV. Blood-brain barrier breakdown in Alzheimer disease and other neurodegenerative disorders. Nat Rev Neurol. 2018; 14:133-150.
169. Chen Y, Liu L. Modern methods for delivery of drugs across the blood-brain barrier. Adv Drug Deliv Rev. 2012; 64:640-665.
170. Pardridge WM. The blood-brain barrier: Bottleneck in brain drug development. NeuroRX. 2005; 2:3-14.
171. Breijyeh Z, Karaman R. Comprehensive review on Alzheimer's disease: Causes and treatment. Molecules. 2020; 25:5789.
172. Fong CW. Permeability of the blood-brain barrier: Molecular mechanism of transport of drugs and physiologically important compounds. J Membr Biol. 2015; 248:651-669.
173. Banks WA. Characteristics of compounds that cross the blood-brain barrier. BMC Neurol. 2009; 9:S3.
174. Arpino G, Somasundaram A, Shin W, Ge L, Villareal S, Chan CY, Ashery U, Shupliakov O, Taraska JW, Wu LG. Clathrin-mediated endocytosis cooperates with bulk endocytosis to generate vesicles. iScience. 2022;

25:103809.
175. Grimm HP, Schumacher V, Schäfer M, *et al*. Delivery of the Brainshuttle™ amyloid-beta antibody fusion trontinemab to non-human primate brain and projected efficacious dose regimens in humans. MAbs. 2023; 15:2261509.
176. Götz J. Unlocking blood-brain barrier boosts immunotherapy efficacy, lowers ARIA. *https://www.alzforum.org/news/conference-coverage/unlocking-blood-brain-barrier-boosts-immunotherapy-efficacy-lowers-aria* (accessed December 15, 2023).
177. Jena L, McErlean E, McCarthy H. Delivery across the blood-brain barrier: Nanomedicine for glioblastoma multiforme. Drug Deliv Transl Res. 2020; 10:304-318.
178. Yang L, Zhou Z, Song J, Chen X. Anisotropic nanomaterials for shape-dependent physicochemical and biomedical applications. Chem Soc Rev. 2019; 48:5140-5176.
179. Poudel P, Park S. Recent advances in the treatment of Alzheimer's disease using nanoparticle-based drug delivery systems. Pharmaceutics. 2022; 14:835.
180. González LF, Bevilacqua LE, Naves R. Nanotechnology-based drug delivery strategies to repair the mitochondrial function in neuroinflammatory and neurodegenerative diseases. Pharmaceutics. 2021; 13:2055.
181. Rafii MS, Sol O, Mobley WC, *et al*. Safety, tolerability, and immunogenicity of the ACI-24 vaccine in adults with down syndrome: A Phase 1b randomized clinical trial. JAMA Neurol. 2022; 79:565-574.
182. Lombardo R, Musumeci T, Carbone C, Pignatello R. Nanotechnologies for intranasal drug delivery: An update of literature. Pharm Dev Technol. 2021; 26:824-845.
183. Crowe TP, Greenlee MHW, Kanthasamy AG, Hsu WH. Mechanism of intranasal drug delivery directly to the brain. Life Sci. 2018; 195:44-52.
184. Ravi PR, Aditya N, Patil S, Cherian L. Nasal in-situ gels for delivery of rasagiline mesylate: Improvement in bioavailability and brain localization. Drug Deliv. 2015; 22:903-910.
185. Sunena, Singh SK, Mishra DN. Nose to brain delivery of galantamine loaded nanoparticles: *In-vivo* pharmacodynamic and biochemical study in mice. Curr Drug Deliv. 2019; 16:51-58.
186. Jansen WJ, Ossenkoppele R, Knol DL, *et al*. Prevalence of cerebral amyloid pathology in persons without dementia: A meta-analysis. JAMA. 2015; 313:1924-1938.
187. McDade E, Wang G, Gordon BA, *et al*. Longitudinal cognitive and biomarker changes in dominantly inherited Alzheimer disease. Neurology. 2018; 91:e1295-e1306.
188. Cantillon M, Andreasen N, Prins N. Phase 1/2a intravenous and subcutaneous oligomer-specific antibody KHK6640 in mild to moderate Alzheimer's disease. J Prev Alzheimers Dis. 2024; 11:65-70.
189. Timmers M, Streffer JR, Russu A, *et al*. Pharmacodynamics of atabecestat (JNJ-54861911), an oral BACE1 inhibitor in patients with early Alzheimer's disease: Randomized, double-blind, placebo-controlled study. Alzheimers Res Ther. 2018; 10:85.
190. Eisai Inc. AHEAD 3-45 study: A study to evaluate efficacy and safety of treatment with lecanemab in participants with preclinical Alzheimer's disease and elevated amyloid and also in participants with early preclinical Alzheimer's disease and intermediate amyloid. *https://clinicaltrials.gov/study/NCT04468659?term=Lecanemab%20in%20Early%20Alzheimer%27s%20Disease&rank=2* (accessed December 1, 2023).
191. Eli Lilly and Company. A donanemab (LY3002813) prevention study in participants with Alzheimer's disease (TRAILBLAZER-ALZ 3). *https://clinicaltrials.gov/study/NCT05026866?cond=Alzheimer%27s%20Disease&term=preclinical&aggFilters=status:not%20rec&rank=10* (accessed February 2, 2024).
192. Lalli G, Schott JM, Hardy J, De Strooper B. Aducanumab: A new phase in therapeutic development for Alzheimer's disease? EMBO Mol Med. 2021; 13:e14781.

Received November 19, 2023; Revised February 9, 2024; Accepted February 19, 2024.

[§]These authors contributed equally to this work.

**Address correspondence to:*
Song Chen, Postdoctoral Station of Xiamen University, No. 422, Siming South Road, Xiamen 361005, China.
E-mail: 2775@hbtcm.edu.cn

Ling Wang, Laboratory for Reproductive Immunology, Obstetrics and Gynecology Hospital of Fudan University, 419 Fangxie Road, Shanghai 200011, China.
E-mail: Dr.wangling@fudan.edu.cn

Released online in J-STAGE as advance publication February 21, 2024.

Review

DOI: 10.5582/bst.2024.01026

Predictive deep learning models for cognitive risk using accessible data

Kenji Karako*

Department of Human and Engineered Environmental Studies, Graduate School of Frontier Sciences, The University of Tokyo, Chiba, Japan.

SUMMARY The early detection of mild cognitive impairment (MCI) is crucial to preventing the progression of dementia. However, it necessitates that patients voluntarily undergo cognitive function tests, which may be too late if symptoms are only recognized once they become apparent. Recent advances in deep learning have improved model performance, leading to applied research in various predictive problems. Studies attempting to estimate dementia and the risk of MCI based on readily available data are being conducted, with the hope of facilitating the early detection of MCI. The data used for these predictions vary widely, including facial imagery, voice recordings, blood tests, and inertial information during walking. Deep learning models that make predictions based on these data sources have been proposed. This article summarizes recent research efforts to predict the risk of dementia using easily accessible data. As research progresses and more accurate predictions become feasible, simple tests could be incorporated into daily life to monitor one's personal health status and to facilitate an early intervention.

Keywords dementia, deep learning, mild cognitive impairment, predictive model

1. Introduction

Globally, the population is aging, with the number of people age 65 and above reaching 727 million, representing 9.3% of the total population of 7.7 billion in 2020 (*1*). Japan has the world's highest rate of aging, with its elderly population accounting for 28.6% of its total population in 2020. Dementia, and especially Alzheimer's disease, is a significant challenge in such aging societies. The Cabinet Office predicts that by 2025, around 7 million elderly Japanese will have dementia, accounting for 20% of those age 65 and over (*2*). Globally, dementia cases are expected to rise to 152 million by 2050 (*3*). Early detection is crucial as many cases progress significantly before becoming apparent, particularly in the early stages of mild cognitive impairment (MCI), which often goes unnoticed due to its minimal impact on daily or social activities. Identifying MCI early is essential to preventing and halting the progression of dementia.

Over the past few years, various studies haves been conducted to detect MCI early. A technology called deep learning has been particularly highlighted and utilized. Deep learning is one of the methods in the field known as machine learning. Essentially, machine learning techniques involve using an algorithm to discover features, rules, or patterns existing in the background of the data collected with regard to a certain event or task, and then using those features or rules to make inferences. Deep learning is an improved method of machine learning based on a technique called neural networks. A characteristic of deep learning is its ability to learn features, rules, or patterns from a large amount of data collected on complex problems, enabling high-performance inference. Conventional machine learning algorithms have difficulty dealing with such a large amount of input information, but one of the deep learning technologies, convolutional neural networks (CNNs) (*4*), can locally extract image information and convert it into data of a smaller size. For instance, in tasks where the goal is to discern whether an image contains a cat or a dog, a CNN learn to recognize essential patterns such as eyes, ears, and the mouth. This learning process involves repeatedly extracting relevant information, allowing the network to focus only on the data necessary for image recognition tasks. The CNN learns from a dataset designed for the specific task, including images of cats and dogs alongside the correct identification of each. Deep learning utilizes vast amounts of task-related data and correct answers to develop an algorithm capable of high-performance predictions and feature extraction. Over the past decade,

deep learning has advanced significantly, demonstrating human-like or superior performance in areas such as image recognition, text generation, autonomous driving, facial recognition, and AI systems like ChatGPT.

Deep learning is increasingly used in medical research, including predicting dementia. Here, studies using deep learning from various perspectives to detect dementia early are described. Conventionally, dementia is assessed using the Mini-Mental State Examination (MMSE) to evaluate cognitive function (*5*). In addition, brain MRI scans and biomarker tests are used. However, markers like amyloid-beta require invasive procedures, making them impractical for widespread screening and early detection of dementia (*6*). This highlights the significant challenge of early detection, as opportunities for testing are limited unless patients proactively seek medical help. Moreover, administering the MMSE and performing an MRI scan are costly and time-consuming, making their use as screening tests impractical. Therefore, recent research has focused on developing more affordable and convenient methods of detecting dementia using deep learning. This approach differs from conventional testing methods by focusing on easily obtainable information, such as facial expressions, voice, basic blood tests, and gait data. The potential of these data types to detect dementia early will be detailed further. The key advantage of these sources is their ease of acquisition. If these prediction models evolve to offer a high level of accuracy, they could enable immediate on-site testing, known as point-of-care testing (PoCT), and these tests could be incorporated into daily life. Here, the potential to use deep learning-based methods of estimation for PoCT to detect dementia is summarized.

2. Estimation of MCI using facial images

Research has attempted to estimate dementia based on facial video. The field of image recognition, which has particularly advanced as a result of deep learning, encompasses object estimation, facial recognition, facial expression recognition, and detecting human figures in video. Predominantly developed through CNNs, models like AlexNet (*7*), ResNet (*8*), and VGG (*9*) have emerged to extract features from images, alongside object detection models such as Faster R-CNN (*10*) and YOLO (*11*) for real-time detection. These technologies are used in studies to estimate cognitive function based on facial videos (*12*). Prior studies reported younger-looking facial impressions in individuals without dementia (*13*), suggesting potential facial indicators of cognitive decline. This research focuses on estimating cognitive functions based on facial videos. For the study, videos ranging from 3 to 30 minutes in length were recorded of 34 elderly individuals age 65 and above, including 10 with MCI. Images were extracted from these videos at a rate of 5 frames per second, with 10 frames over 2 seconds forming one set for the model's input. A total of 3,822 data sets were created, with 3,058 sets used for training and 764 sets for evaluation, to solve a binary classification problem of distinguishing between MCI and health using deep learning. The study used ResNet, which is based on a CNN, to extract facial structure and motion information from facial videos (Figure 1). ResNet, a deep learning model linking over 50 layers of CNNs, was developed for image recognition tasks and is highly effective at extracting features from two-dimensional spatial information. The model to estimate MCI was created using two instances of ResNet: one as a model to extract spatial features from the face to estimate MCI and the other as a model to extract dynamic features based on facial dynamics to predict MCI. The spatial model randomly selects one image from a set of 10 frames over 2 seconds for input, focusing on static facial features. The dynamic model generates an optical flow from the same frame set, reflecting facial movements over 2 seconds, which ResNet then uses to extract features. Optical flow (*14*), represented by a three-dimensional vector for each

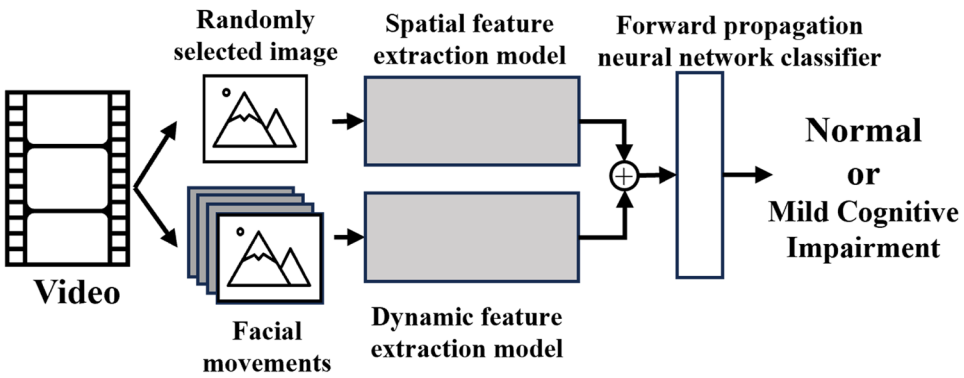

Figure 1. Deep learning structure to estimate MCI based on image and motion information. Facial videos are divided into still images and motion information using an optical flow. Deep learning models are created for each to extract spatial and dynamic features, which are then used to estimate MCI.

pixel, is analogous to the RGB structure of images, making ResNet suitable for extracting features from these data. The model ultimately estimates whether an individual has MCI based on two dynamically and spatially obtained features. The final model had a precision of 0.94, recall of 0.78, accuracy of 0.91, and an F1 score of 0.85. Despite the low recall and concerns over data on a small number of individuals and the balance between MCI and normal data, the ability to determine MCI at a certain level using deep learning represents a significant advance in early detection. Estimation of cognitive functions based on video data, such as this, is also being performed in another study (*15*) and is an area of growing interest. If MCI can be estimated based on approximately two seconds of video data, this could allow for testing without a significant burden in everyday life or visits to hospitals and care facilities, enabling immediate examinations on-site.

3. Estimating Alzheimer's disease using speech information

Estimating dementia based on speech information is one of the most extensively studied tasks in the field of deep learning-based estimation of dementia (*16*). Alzheimer's disease, a type of dementia, initially manifests as language impairments. Focusing on this characteristic, the goal is to estimate the presence of Alzheimer's disease using speech data. Previous studies have reported that Alzheimer's patients tend to pause more frequently between words and speak more slowly than healthy individuals (*17*). Moreover, Alzheimer's patients are reported to have difficulties in finding appropriate words or expressions to match a sentence (*18,19*). Deep learning models are used to extract various vocal features from speech data. In order to estimate Alzheimer's disease, two primary features are extracted: features from continuous speech signals and features from speech converted to text to analyze the context and content of conversations. These features are then used for the final estimation task.

In order to extract features from speech signals, studies have used deep learning algorithms that are effective at continuous signal processing (*20*), such as long short-term memory (*21*) and recurrent neural networks (RNN) (*22*). These algorithms have the capability to internally retain a memory of past inputs, allowing the neural network to maintain information over a certain duration. This capacity enables the extraction of features needed to estimate Alzheimer's disease not just based on a single speech sample but also based on historical data. However, they have limitations in terms of storing information over extended periods, such as tens of minutes.

The second method involves converting speech into textual data and then extracting Alzheimer's disease characteristics from the context and content of the text. This approach estimates Alzheimer's based on the coherence and expressiveness of the text. A drawback is that features unique to speech might be missed. However, unlike with direct extraction of speech features, this approach allows for estimation based on lengthy dialogues that have been converted to text. Recent advances in deep learning for natural language processing, such as the use of the high-performance natural language model BERT (*23*), have led to proposed methods of estimating Alzheimer's using those technologies (*24*).

Data used to train and evaluate models come from tasks performed during studies. Primarily, tasks include semantic verbal fluency (*25*), where subjects list as many items as possible from a category like animals or vegetables within one minute (*26,27*), a natural speech task involving conversation without direct questions (*28*), and a picture description task where participants orally describe the content of a picture within a set time (*29*). Notably, the ADReSS database (*30*) offers open access to data from these tasks, including voice recordings, transcribed texts, and MMSE scores. Such databases are valuable for developing deep learning models to estimate Alzheimer's based on speech data.

4. Estimation of MCI using blood test information

One unique area of research to detect dementia early using deep learning involves blood test information (*31*). This research focuses on the relationship between systemic disorders like arteriosclerosis, which is the result of lifestyle diseases, and cognitive impairments, which include both MCI and severe dementia (*32-34*). It also considers other systemic disorders that might affect cognitive function, such as malnutrition (*35*), anemia (*36*), lipid metabolism (*37*), purine metabolism (38), and renal dysfunction (*39*). These can be detected *via* basic blood tests obtained during health check-ups. The research attempts to estimate MCI using blood test data, including 23 items like red and white blood cell counts, hemoglobin levels, hematocrit, albumin levels, and age, using a feedforward neural network, a basic form of deep learning, to predict MMSE scores. The input items obtained from the blood tests used are shown in Table 1. This neural network consists of a four-layer structure with intermediate layers as shown in Figure 2. Each intermediate layer has a neural network with 400 nodes, solving a regression problem that estimates the MMSE in the range of 0 to 30 based on 24 numerical items. Data used to train and evaluate the data were collected from 202 patients (average age: 73.48 ± 13.1 years). All patients received inpatient treatment including rehabilitation and pharmacotherapy for lifestyle-related diseases, with 142 patients having cerebrovascular diseases and 174 patients having at least one lifestyle-related disease. The feedforward neural network was trained and evaluated using the leave-one-out method,

which was applied to the blood test results and MMSE scores from the 202 patients. Actual MMSE scores and predicted MMSE scores were correlated ($r = 0.85$, $p < 0.001$). The mean absolute error was 2.02. Blood tests, primarily obtained during medical examinations and health check-ups, serve as the main data for this research. A cognitive function estimation model based on blood tests could effectively be utilized as a test to screen for dementia in medical facilities and during regular health check-ups. For instance, when elderly individuals undergo blood tests during a health examination or medical visit, their cognitive function can be estimated using deep learning in no time at all. If MCI or dementia is suspected, a medical facility could then encourage the individual to undergo a more detailed examination or visit an outpatient clinic. This estimation model could be an effective means for early detection of dementia, simply by undergoing a regular medical consultation or health check-up.

Table 1. Test items used to estimate the MMSE score based on blood test data

Blood test items
White blood cell count
Red blood cell count
Mean corpuscular volume
Mean corpuscular hemoglobin
Mean corpuscular hemoglobin concentration
Platelet count
Hematocrit
Hemoglobin
Total protein
Albumin
A/G ratio
AST (GOT)
ALT (GPT)
γ-GTP
Total cholesterol
Triglyceride
Blood urea nitrogen
Creatinine
Uric Acid
Glucose
Na
K
CI

5. Estimation of MCI using inertial information during walking

Compared to the previously described models to estimate MCI, there is another approach that is more similar to everyday life, and it has the potential to be used for the early detection of MCI by estimating cognitive decline on the spot in everyday situations. Studies have estimated MCI using inertial sensor data collected by a wearable device when the wearer walks (*40,41*). In those studies, a small inertial sensor was affixed to the shin of 30 cognitively normal individuals and 30 individuals with MCI, and they were asked to perform a simple task of walking 20 meters, as well as a complex task of walking 20 meters while simultaneously performing cognitive tasks such as subtracting numbers or naming animals. Moreover, subjects were asked to always keep walking while performing the task. The device used for measurement was the Shimmer3 GSR+ Unit (*42*), which is equipped with a 3-axis accelerometer and a 3-axis gyroscope. Eight pieces of information, including three forms of acceleration, three angular velocities, and the total magnitude of the signal vectors of both the accelerometer and gyroscope sensors, were used to estimate MCI. A six-layer CNN and three types of RNN were used to estimate MCI, as shown in Figure 3. The eight signals input are time-series data, and the 8×T input information, segmented

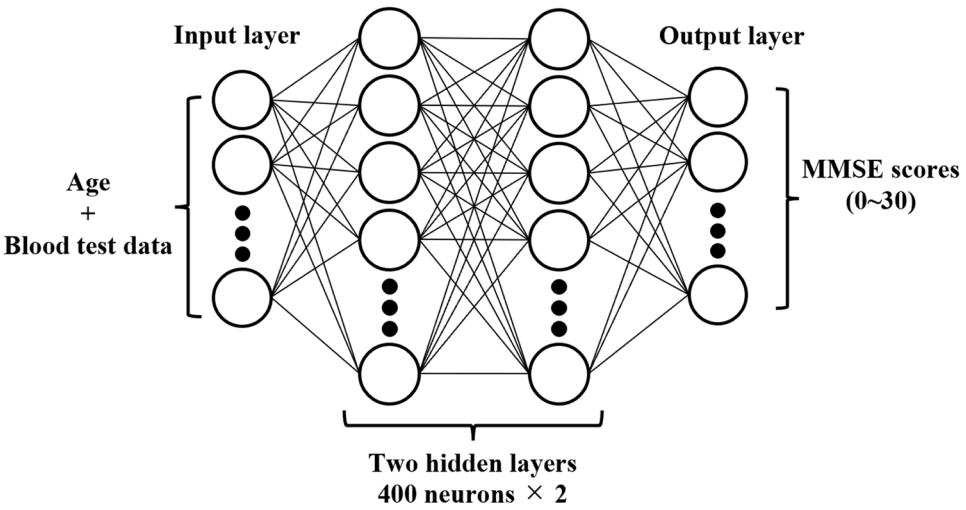

Figure 2. Structure of deep learning used to estimate the MMSE score based on blood test data. It consists of a forward propagation neural network with a four-layer structure.

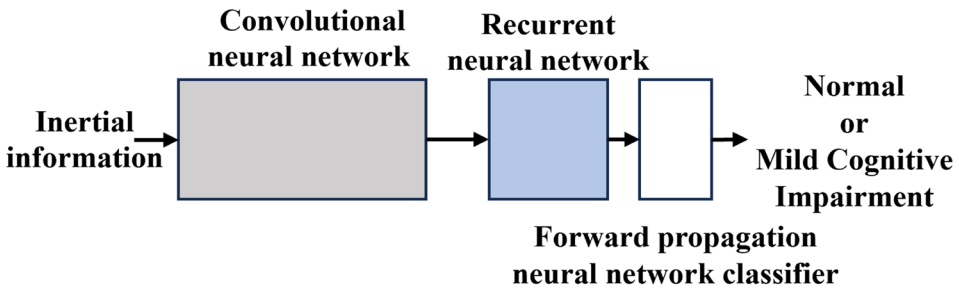

Figure 3. The structure of deep learning to estimate MCI based on inertial information during walking. The CNN handles inertial information in image format and extracts features. The recurrent neural network subsequently extracts features based on those from the past and present, and these are used these to perform the final estimation of MCI.

by a certain time T, is input into the CNN as an image. The features extracted by this process are then input into the RNN. Since the RNN has the characteristic of retaining past input information, the features extracted by the CNN from information before the most recent time T are retained, and features that incorporate time-series information are ultimately extracted. Finally, a binary classification of MCI is performed by a feedforward neural network. The leave-one-out method was used for model training and evaluation, achieving an accuracy of 73.33%, a sensitivity of 83.33%, and a specificity of 63.3%. The walking information used in this study was obtained by attaching a measuring device to the shin, which differs from walking data that can be easily obtained with commonly carried devices such as smartphones or smartwatches. Therefore, this method has not yet reached the point where it can be used for early detection in everyday life as it is. However, as research and data collection progress, this method could be effectively utilized as a method of detecting MCI early since it the sensor is easy to attach and measurement is performed simply by walking, potentially serving as a prompt before visiting a medical facility.

6. Conclusion

PoCT refers to methods that allow for immediate testing on the spot at the appropriate time. Conventional tests for dementia primarily involve brain imaging with MRI, peripheral biomarkers like amyloid-beta, and the MMSE, which are used for final diagnosis. These tests require a certain amount of time to conduct, and moreover, they are opportunities that will not arise unless individuals are aware of their symptoms and go to a hospital voluntarily. Due to the inconvenience of such tests, research has been conducted on methods that can estimate MCI using deep learning based on information that can be acquired more easily, without hassle, and without posing a burden. Deep learning has a high level of inferential performance and can learn from complex data, so data measured during events that indirectly reflect the impact of dementia could be used effectively, something that was difficult to achieve in the past. If high-precision estimation of MCI becomes possible based on information that can be obtained relatively easily, such as speech, facial expressions, blood, and gait, deep learning models will likely be incorporated into testing systems for PoCT. For example, estimation of MCI using facial recognition or blood tests could be combined with regular health checkups for the elderly, allowing for effortless estimation of MCI. Moreover, if estimation of MCI can be achieved using diverse data sources, this could lead to more accurate estimates. Furthermore, if estimation of MCI is widely adopted for PoCT, it could easily be configured into a smartphone or web app. Estimation could be performed at medical facilities and also in nursing homes and at home, so this test could be integrated into one's daily life. Including simple tests using deep learning in daily life could allow for immediate detection of abnormalities, leading to the discovery of cognitive decline at an earlier stage compared to conventional methods. If tests are conducted daily and the data collected and used for research, this could lead to estimates of future changes in cognitive functions, such as one year or five years later, based on an analysis of daily data collected over time.

Funding: None.

Conflict of Interest: The author has no conflicts of interest to disclose.

References

1. Cabinet Office, 2022 Edition of the White Paper on an Aging Society: Chapter 1 The State of Aging (Section 1, 2) *https://www8.cao.go.jp/kourei/whitepaper/w-2022/html/zenbun/s1_1_2.html* (accessed February 10, 2024) (in Japanese)
2. Cabinet Office, 2022 Edition of the White Paper on an Aging Society: Chapter 1 The State of Aging (Section 1, 1) *https://www8.cao.go.jp/kourei/whitepaper/w-2022/html/zenbun/s1_1_1.html* (accessed February 10, 2024) (in Japanese)

3. World Health Organization, Dementia: A public health priority. World Health Organization *https://www.who.int/publications/i/item/dementia-a-public-health-priority* (accessed February 10, 2024)
4. LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W, Jackel LD. Backpropagation applied to handwritten zip code recognition. Neural Comput. 1989; 1: 541-551.
5. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state" A practical method for grading the cognitive state of patients for clinicians. J Psychiatr Res. 1975; 12: 189-98.
6. Dumurgier J, Hanseeuw BJ, Hatling FB, Judge KA, Schultz AP, Chhatwal JP, Blacker D, Sperling RA, Johnson KA, Hyman BT, Gómez-Isla T. Alzheimer's disease biomarkers and future decline in cognitive normal older adults. J. Alzheimer's Dis. 2017; 60:1451-1459.
7. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Commun. ACM. 2017; 60:84-90.
8. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. 2016:770-778.
9. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. 2014; https://doi.org/10.48550/arXiv.1409.1556
10. Ren S, He K, Girshick R, Sun J. Faster R-CNN: Towards real-time object detection with region proposal networks. 2015; https://doi.org/10.48550/arXiv.1506.01497
11. Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: Unified, real-time object detection. Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. 2016:779-788.
12. Lee C, Chau H, Wang H, Chuang Y, Chau Y. Detection of mild cognitive impairment by facial videos. IEEE Int. Conf. Consum. Electron.-Taiwan. 2022:197-198.
13. Tufekcioglu Z, Bilgic B, Zeylan AE, Salah AA, Dibeklioglu H, Emre M. Do Alzheimer's disease patients appear younger than their real age? Dement Geriatr Cogn Disord. 2020; 49:483-488.
14. Lai WS, Huang JB, Yang MH. Semi-supervised learning for optical flow with generative adversarial networks. Adv. Neural Inf. Process. Syst. 2017; 30.
15. Sun J, Dodge HH, Mahoor MH. MC-ViViT: Multi-branch classifier-ViViT to detect mild cognitive impairment in older adults using facial videos. 2023; https://doi.org/10.48550/arXiv.2304.05292
16. Yang Q, Li X, Ding X, Xu F, Ling Z. Deep learning-based speech analysis for Alzheimer's disease detection: A literature review. Alzheimers Res Ther. 2022; 14:186-186.
17. Croisile B, Brabant M, Carmoi T, Lepage Y, Aimard G, Trillet M. Comparison between oral and written spelling in Alzheimer's disease. Brain Lang. 1996; 54:361-387.
18. Croisile B, Ska B, Brabant M, Duchene A, Lepage Y, Aimard G, Trillet M. Comparative study of oral and written picture description in patients with Alzheimer's disease. Brain Lang. 1996; 53:1-19.
19. Cuetos F, Arango-Lasprilla JC, Uribe C, Valencia C, Lopera F. Linguistic changes in verbal expression: A preclinical marker of Alzheimer's disease. J Int Neuropsychol Soc. 2007; 13:433-439.
20. Rohanian M, Hough J, Purver M. Alzheimer's dementia recognition using acoustic, lexical, disfluency and speech pause features robust to noisy inputs. 2021; https://doi.org/10.48550/arXiv.2106.15684
21. Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput. 1997; 9:1735-1780.
22. Medsker L, Jain LC. Recurrent neural networks: Design and applications. CRC press. 1999.
23. Devlin J, Chang MW, Lee K, Toutanova K. Bert: Pre-training of deep bidirectional transformers for language understanding. 2018; https://doi.org/10.48550/arXiv.1810.04805
24. Liu Z, Proctor L, Collier PN, Zhao X. Automatic diagnosis and prediction of cognitive decline associated with Alzheimer's dementia through spontaneous speech. IEEE Int. Conf. Signal Image Processing Appl. 2021:39-43.
25. Lopes M, Brucki SMD, Giampaoli V, Mansur LL. Semantic verbal fluency test in dementia: Preliminary retrospective analysis. Dement Neuropsychol. 2009; 3:315-20.
26. Campagna F, Montagnese S, Ridola L, Senzolo M, Schiff S, De Rui M, et al. The animal naming test: An easy tool for the assessment of hepatic encephalopathy. Hepatology. 2017; 66:198-208.
27. Chien YW, Hong SY, Cheah WT, Fu LC, Chang YL. An assessment system for Alzheimer's disease based on speech using a novel feature sequence design and recurrent neural network. IEEE Int. Conf. Syst. Man Cybern. 2018; 3289-3294.
28. Illes J. Neurolinguistic features of spontaneous language production dissociate three forms of neurodegenerative disease: Alzheimer's, Huntington's, and Parkinson's. Brain Lang. 1989; 37:628-642.
29. Becker JT, Boller F, Lopez OL, Saxton J, McGonigle KL. The natural history of Alzheimer's disease: Description of study cohort and accuracy of diagnosis. Arch. Neurol. 1994; 51:585-594.
30. Luz S, Haider F, de la Fuente S, Fromm D, MacWhinney B. Alzheimer's dementia recognition through spontaneous speech: The ADReSS challenge. 2020; https://doi.org/10.48550/arXiv.2004.06833
31. Sakatani K, Oyama K, Hu L. Deep learning-based screening test for cognitive impairment using basic blood test data for health examination. Front. Neurol. 2020; 11:588140.
32. Van Der Flier WM, Skoog I, Schneider JA, Pantoni L, Mok V, Chen CL, Scheltens P. Vascular cognitive impairment. Nat. Rev. Dis. Primers. 2018; 4:1-16.
33. Al-Qazzaz NK, Ali SH, Ahmad SA, Islam S, Mohamad K. Cognitive impairment and memory dysfunction after a stroke diagnosis: A post-stroke memory assessment. Neuropsychiatr Dis Treat. 2014; 1677-1691.
34. Gorelick PB, Scuteri A, Black SE, *et al*. Vascular contributions to cognitive impairment and dementia: A statement for healthcare professionals from the American Heart Association/American Stroke Association. Stroke. 2011; 42:2672-2713.
35. Brooke J, Ojo O Enteral nutrition in dementia: A systematic review. Nutrients. 2015; 7:2456-2468.
36. Hong CH, Falvey C, Harris TB, Simonsick EM, Satterfield S, Ferrucci L, Metti AL, Patel KV, Yaffe K. Anemia and risk of dementia in older adults: Findings from the Health ABC study. Neurology. 2013; 81:528-533.
37. Qizilbash N, Gregson J, Johnson ME, Pearce N, Douglas I, Wing K, Evans SJW, Pocock SJ. BMI and risk of dementia in two million people over two decades: A retrospective cohort study. Lancet Diabetes Endocrinol. 2015; 3:431-436.
38. Delgado-Alvarado M, Gago B, Navalpotro-Gomez I,

Jiménez-Urbieta H, Rodriguez-Oroz MC. Biomarkers for dementia and mild cognitive impairment in Parkinson's disease. Mov. Disord. 2016; 31:861-881.
39. Miranda AS, Cordeiro TM, dos Santos Lacerda Soares TM, Ferreira RN, Simoes e Silva AC. Kidney–brain axis inflammatory cross-talk: from bench to bedside. Clin. Sci. 2017; 131:1093-1105.
40. Lee H, Shahzad A, Kim K. Automated prescreening of MCI through deep learning models based on wearable inertial sensors data. Alzheimers. Dement. 2021; 17: e052744.
41. Shahzad A, Dadlani A, Lee H, Kim K. Automated prescreening of mild cognitive impairment using shank-mounted inertial sensors based gait biomarkers. IEEE Access. 2022; 10:15835-15844.
42. Burns A, Greene BR, McGrath MJ, O'Shea TJ, Kuris B, Ayer SM, Stroiescu F, Cionca V. SHIMMER™ A wireless sensor platform for noninvasive biomedical research. IEEE Sens. J. 2010; 10:1527-1534.

Received January 12, 2024; Revised February 10, 2024; Accepted February 16, 2024.

*Address correspondence to:
Kenji Karako, Department of Human and Engineered Environmental Studies, Graduate School of Frontier Sciences, The University of Tokyo, 5-1-5 Kashiwanoha, City of Kashiwa, Chiba 227-8568, Japan.
E-mail: tri.leafs@gmail.com

Released online in J-STAGE as advance publication February 20, 2024.

Original Article

DOI: 10.5582/bst.2023.01319

Socioeconomic disparities in education placement for children of primary school age with autism spectrum disorder in China

Yanan Zhao[1,2], Rong Zhang[3,4,5,6,7], Xiaoying Zheng[8,*]

[1] School of Public Health, Peking University, Beijing, China;
[2] China Center for Health Development Studies, Peking University, Beijing, China;
[3] Department of Neurobiology, School of Basic Medical Sciences, Peking University, Beijing, China;
[4] Neuroscience Research Institute, Peking University, Beijing, China;
[5] Key laboratory for Neuroscience, Ministry of Education/National Health and Family Planning Commission, Peking University, Beijing, China;
[6] Autism Research Center, Peking University Health Science Center, Beijing, China;
[7] Department of Integration of Chinese and Western Medicine, School of Basic Medical Sciences, Peking University, Beijing, China;
[8] School of Population Medicine and Public Health, Chinese Academy of Medical Sciences/Peking Union Medical College, Beijing, China

SUMMARY Relatively little is known about education placements for children with autism spectrum disorder (ASD) in China. While disparities in ASD diagnoses and services for the population broadly are often documented, the presence and determinants of differences in the educational placement of ASD children are less studied and understood. By identifying who is likely to be in segregated settings, we can discern how to best support them and facilitate a possible transition to a less restrictive setting. This study describes four placements (regular schools, special schools, institutions, homes) and their influencing factors retrospectively in a large sample (n = 2,190) of Chinese primary school-aged children (6-12 years old). We divided ASD into severe and mild to moderate categories for analysis. Children with ASD were more likely to study in a regular school (48.60%), while 13.88% were in a special school. Children with severe ASD were placed in less regular settings than children with mild to moderate ASD. However, families with higher socioeconomic status (SES) were more likely to place their children in regular schools than lower SES families if their children experienced mild to moderate symptoms. Children with severe ASD were more likely to be placed in expensive institutions for families with higher SES than those with lower SES. SES disparities in educational placement existed and had two manifestations. It is important to characterize educational placements of students with ASD to determine the extent to which they are placed in general education settings, which are often the preferred placement.

Keywords autism spectrum disorder (ASD), primary school, inclusive education, socioeconomic disparities, China

1. Introduction

Autism spectrum disorders (ASDs) are a range of neurodevelopmental disorders that are characterized by the following core deficits: impairments in social interaction and communication and restricted, repetitive behaviors (*1*). According to the recent evidence from China, the prevalence of ASD among children aged 6 to 12 years was 0.7% (*2*), which was much higher than most previous research findings in China. As a developing country with a large population, China faces the challenge of providing sufficient educational supports for individuals with ASD. The required supports include direct cognitive instruction, behavior cultivation, as well as necessary social-emotional and mental health services (*3*). In recent years, countries worldwide have explored and supported inclusive education, which is viewed as a moral and judicial imperative (*4*) and a reflection of a fair society (*5*). In the context of a growing emphasis on inclusive education, the increasing number of children with ASD diagnosed has led to a competition for inclusive education resources.

Educational placement is not a simple choice or allocation. Numerous factors can influence the placement of education among children with special need (*6-8*). ASD students with greater support needs face a series of barriers that may prevent them from making the most of their inclusive education (*9*). In addition to children's characteristics, the socio-economic background (SES) of families can also impact the decisions. A higher family

income has been found to be associated with a greater probability of attending non-public schools rather than public schools for ASD children (*10*). When mothers have a higher level of education, the children with Down syndrome are more likely to choose mainstream schools (*11*). Children from higher SES families are more likely to receive education from less restrictive placements than children from lower SES backgrounds (*12*). But there have also been contrasting findings. For example, students with vision impairment from families with higher SES are more frequently placed in special than regular schools (*13*). Maternal educational attainment is not significantly related to attending non-public or public school (*10*). Nevertheless, the growing literature have noted that improved service access among families with more resources (*14*). The SES of parents differentiate the forms of educating students with disabilities.

While these studies have contributed significantly to our understanding of this topic, they are limited in the following ways. First, although SES disparities in the diagnosis and utilization of healthcare services for ASD are studied, there is limited research regarding the SES differences in educational placement for students with ASD, and no consistent conclusions have been reached (*15,16*). The placement in less-restrictive settings varied along a number of factors, such as parents' level of education, suggesting an inequitable access to the inclusive educational resources for children with ASD. Second, symptoms of diseases exhibit heterogeneity, while abilities may serve as a starting point for research. Higher functional skills were associated with greater likelihood of attending postsecondary education or earning above minimum wage (*17*). However, existing analyses of educational placement lack exploration into the influence of ASD symptoms. Third, previous studies usually focused on one or two types of educational placements, but it was far more than just schools (regular or special). Home is also an educational arrangement that can be handled by parents for training or given up training (*18*) and it requires analysis of more categories of educational placement. Fourth, most studies have been conducted in other countries with different health systems, which were likely to differ from those in China. China's Sui Ban Jiu Du or learning in regular classrooms (LRC) program, implemented since the 1980s, aims to integrate children with special needs into regular classrooms (*19*). But many children with ASD were still excluded due to factors like not meeting criteria or limited resources of this policy (*20*). The current understanding of how families in China place children with ASD in educational settings is not fully clear.

The SES difference may affect resource allocation and children's health outcomes, which necessitate more reasonable public health and education initiatives, as their goal is to decrease the disadvantages of lower SES households. Therefore, we examine the educational placements for children with ASD in China using a nationwide survey data. It is important to characterize educational placements to determine the extent to which they are placed in general education settings which are often the preferred placement. It is also important to identify correlates of placement in general education settings; by identifying who is likely to be in segregated settings, we can discern how to support them and make a potential transition to a less restrictive setting. It therefore remains unclear what factors are truly influential and effective in developing inclusive education programs for children with ASD on a policy level. Research on China can not only shed light on the current status of inclusive education in developing countries, thereby facilitating the development of more appropriate policies, but also enhance understanding of inequality.

2. Methods

2.1. Participants

This study used data from the Survey on Family Circumstances and Demand for Support and Resources among Autistic Children in China (FCDSR). It was a survey that was distributed to members of the AlsoLife online patient community. More than 200,000 parents of ASD children can share information about their conditions, treatments, symptoms, and comorbidities on that platform, which is the largest online gathering place for parents with children and adolescents with ASD in China. The Quality Assurance staff at China Association of Rehabilitation of Disabled Persons (CARDP) reviewed the survey for editorial and technical suggestions, which aimed to describe the family information, treatment, education and health expenditure of ASD children. The survey was available online from 15th September to 30th September 2020. The other details of survey have been described elsewhere (*21*).

2.2. Data collection

Families having children diagnosed with ASD were recruited if they met the following criteria: (1) the children were between the ages of 6 and 12 that the age of primary school; (2) the hospital had diagnostic qualifications and followed a Diagnostic and Statistical Manual of Mental Disorders 5th ed (DSM-5) standard, not only through scale measurement but also *via* medical professional diagnosis. Exclusion criteria were individuals with physical support needs such as those who have a diagnosis of cerebral palsy. Children with intellectual disabilities were not excluded in this analysis. There were 8014 households investigated, with 2190 households included in this study. The selection procedure was depicted in Figure 1. The family location distribution was consistent with China's overall population distribution. 31 provinces in China and a total of 216 cities or districts were included (see Supplemental

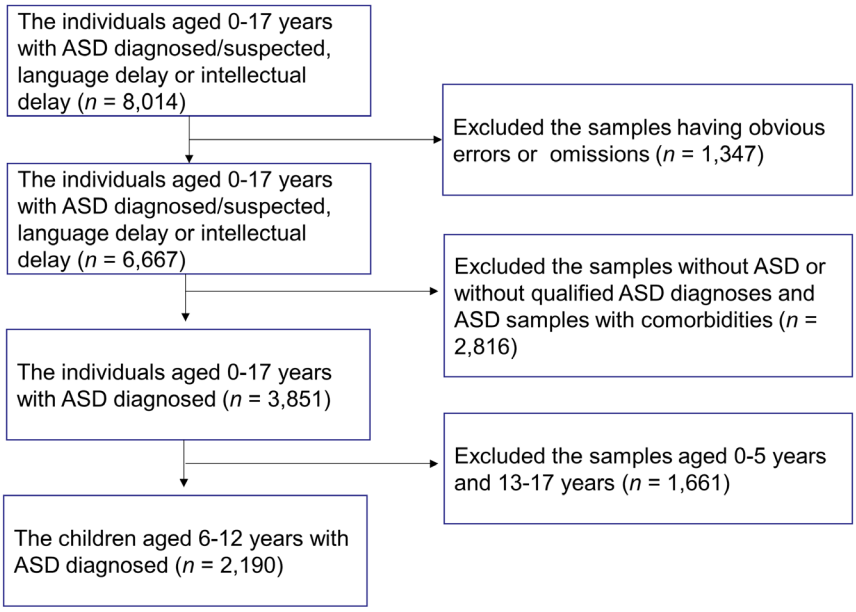

Figure 1. Flow chart.

Table S1 for details, *http://www.biosciencetrends.com/action/getSupplementalData.php?ID=186*). The sample distribution is relatively consistent with the national population distribution, and the sample is representative of the country.

2.3. Measures

2.3.1. Educational placements

It includes four educational placements, namely: "regular school", which means the most of time the students' study were in ordinary or regular schools that serviced general education students; "Special schools", which means that the most of time the students' study were in public special schools that serviced primarily students with special needs; "Institution", which means that the most of time the child were in private institutions that serviced students with special needs, especially for those developmental disorders. The institutions were those more restrictive training agencies, with the majority being private due to inadequate or unsatisfactory services provided by public special and regular schools (*22*). "Home", which means that the children had no other placements but home. Compared to regular schools, the other three placements were more restrictive environment for children with ASD. Because of variances between institutions, we have designated the top 25th percentile of monthly fees as the expensive ones.

2.3.2. Socioeconomic and Demographic Variables

The age of the children was their age at the survey point. The age of children was divided into **two** age groups: 6-8 years old (the primary grades) and 9-12 years old (high-grades in primary school).

The severity of ASD was judged according to professional evaluation or parents' subjective judgment. Due to the fact that severe symptoms were often easier to distinguish, while moderate and high function were more difficult to accurately distinguish, we divided the severity into two levels: (1) severe or need lots of supports (the children with low function ASD (LFA)), (2) mild/moderate or need some supports (the children with middle function autism (MFA) and high function autism (HFA). The regional variables were "eastern", "central" and "western". The provinces in the eastern region were among the first to implement the coastal opening-up policy and have a high level of economic development. The provinces of the central region are economically underdeveloped, while those of the western region are even less so. We classified family income into three categories. According to the data distribution, the below-average group had an annual income of less than $12,327 (80,000 yuan), the around-average group had an annual income of between $12,327 (80,001 Yuan) and $23,112 (150,000 Yuan), and the above-average group had an annual income of more than $23,112 (150,000 Yuan), (*21*). Other background information was collected on children's sex, children's number in the family, parents' education levels.

2.4. Statistical analysis

We use frequencies and percentages to reported for categorical variables, and means/SDs (standard deviations) for continuous variables. Logistic regression models were used to identify the factors influencing educational placements. Associations between predictors and independent variables were reported by odds ratios (*OR*s) and their 95% confidence intervals (*CI*s). All statistical analyses were conducted using SPSS 22.0 for

Windows (SPSS Inc, Chicago, IL, USA).

2.5. Consent and ethics approval

All families provided electronic informed consent before enrollment. All procedures involving human subjects/patients were approved by the ethics committee of Peking University Institutional Review Board and approval number is IRB00001052-20016.

3. Results

3.1. Sample descriptive statistics

A total of 2190 households were included in this survey. Most of the children (86.12%) were boys, and the mean age was 7.44 (SD: 1.45) years old, with the leading severity being mild/moderate (73.84%). 26.16% of the children had severe ASD symptoms. Most of the parents had a college degree (65.34%). Most families lived in the eastern region (62.92%), which was in line with China's population distribution. A total of 48.26% of children were in regular schools, while a total of 13.88% were in special schools, a total of 29.86% were in institutions, and a total of 7.99% were at home. The study population was further described in Table 1.

Figure 2 depicts the proportions of children with ASD who had different accommodations stratified by gender, maternal education level, household income and resident districts. There was no significant difference in the proportion of boys and girls entering the four placement categories. When a child had mild or moderate symptoms, had higher maternal education level he or she was more likely to enter regular schools and les s less likely to be institutions and home. For children with milder symptoms, the proportion of entering regular schools was higher when they came from higher-income

Table 1. Characteristics of study sample (n = 2,190)

Characteristic	N/M	%/SD
Age	7.44	1.45
Sex		
Boy	1,886	86.12
Girl	304	13.88
Only child		
No	1,056	48.22
Yes	1,134	51.78
Severity		
Severe	573	26.16
Mild/moderate	1,617	73.84
Maternal Education level		
High school or below	759	34.66
College degree or higher	1,431	65.34
Household income		
Low	728	33.24
Middle	761	34.75
High	701	32.01
Resident districts		
Eastern	1,378	62.92
Central	594	27.12
Western	218	9.95
Placements		
Regular school	1,057	48.26
Special school	304	13.88
Institution	654	29.86
Home	175	7.99

N: number; M: mean; SD: standard deviation.

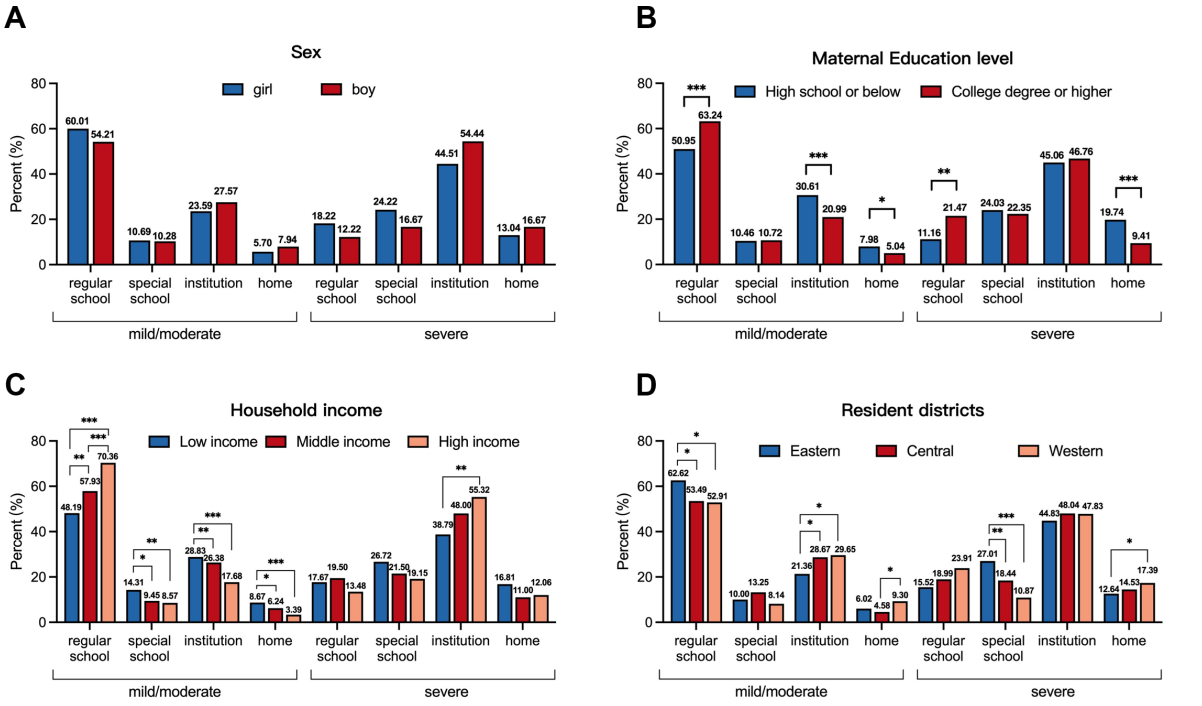

Figure 2. Percentage of children with ASD who had different educational placement stratified by (A) sex, (B) maternal education level, (C) household income, (D) resident districts. *$p \leq 0.05$, **$p < 0.01$, ***$p < 0.001$.

families (70.36% *v*s. 57.93% *v*s. 48.19%), while the proportions of special schools, institutions, and home-based care were lower for children with high family income. For children with more severe symptoms, the proportion of entering institutions was higher for children from higher-income families (55.32% *vs*. 38.79%). Children with milder symptoms from families in the eastern region were more likely to enter regular schools (62.62% *vs*. 53.49% *vs*. 52.91%) and less likely to enter institutions (21.36% *vs*. 28.57% *vs*. 29.65%). For children with more severe symptoms, when children from families in the eastern region, there is a higher proportion of children attending special schools (27.01% *vs*. 18.44% *vs*. 10.87%).

3.2. Predictors of educational placements

The probability of older age group (9-12 years) entering regular schools was higher than that of younger age group (6-8 years) (OR 1.51, 95% CI 1.20–1.90). The only child was more likely to be in regular school than child from multi-children family (OR 1.21, 95% CI 1.00–1.46). The children with mothers who had a college degree or above were 1.42 times more likely to be in regular school than the mothers who had no college degrees (OR 1.42, 95% CI 1.15–1.76; Table 2). A child with higher family income was more likely to be in regular school than child from low-income family (OR 1.30, 95% CI 1.04–1.64 for middle income family; OR 1.74, 95% CI 1.36–2.23 for high income family). Children with higher severity of ASD were less likely to enter regular schools when compared to mild/moderate severity (OR 0.15, 95% CI 0.11–0.19). Girls had a lower probability of entering regular schools compared to boys (OR 0.75, 95% CI 0.57–0.98).

For special school placement, the severity of ASD, age, family income, and the location of the family had an impact. Children in older age group were 2.31 times more likely to be in special schools than younger children (OR 2.31, 95% CI 1.77–3.02). Children with higher severity of ASD were 2.22 times more likely to be in special schools when compared to mild/moderate severity (OR 2.22, 95% CI 1.71–2.87). Children whose families reside in the western region were less likely to be enrolled in special schools compared to those from eastern region (OR 0.53, 95% CI 0.32–0.88). Children from middle or high income families were less likely to be special schools compared to those from low-income families (OR 0.65, 95% CI 0.48–0.88 for middle income family; OR 0.55, 95% CI 0.39–0.77 for high income family).

For institution placement, child sex, age, the severity of ASD, maternal education, and the location of the family had an impact. Girls had a higher probability of

Table 2. Multivariable logistic regression models for four placements

Characteristics	Model 1 regular school			Model 2 special school			Model 3 institution			Model 4 home		
	OR	95%CI		OR	95%CI		OR	95%CI		OR	95%CI	
		Low	High		Low	High		Low	High		Low	High
Sex												
Boy	1.00			1.00			1.00			1.00		
Girl	0.75	0.57	0.98	0.83	0.57	1.21	1.31	1.00	1.71	1.44	0.96	2.18
Age												
6-8 years	1.00			1.00			1.00			1.00		
9-12 years	1.51	1.20	1.90	2.31	1.77	3.02	0.26	0.20	0.35	1.34	0.94	1.91
Only child												
No	1.00			1.00			1.00			1.00		
Yes	1.21	1.00	1.46	0.80	0.62	1.04	1.00	0.82	1.23	0.77	0.57	1.07
Severity												
Mild/moderate	1.00			1.00			1.00			1.00		
severe	0.15	0.11	0.19	2.22	1.71	2.87	3.08	2.48	3.81	2.19	1.59	3.03
Maternal Education level												
High school or below	1.00			1.00			1.00			1.00		
College degree or higher	1.42	1.15	1.76	1.28	0.96	1.71	0.71	0.57	0.88	0.63	0.45	0.89
Household income												
Low	1.00			1.00			1.00			1.00		
Middle	1.30	1.04	1.64	0.65	0.48	0.88	1.13	0.89	1.43	0.73	0.51	1.06
High	1.74	1.36	2.23	0.55	0.39	0.77	0.94	0.72	1.22	0.58	0.37	0.92
Resident district												
Eastern	1.00			1.00			1.00			1.00		
Central	0.83	0.67	1.03	0.94	0.71	1.26	1.36	1.09	1.70	0.87	0.59	1.26
Western	0.74	0.55	1.01	0.53	0.32	0.88	1.62	1.17	2.23	1.60	1.01	2.57

OR: odds ratio; CI: confidence interval. Independent variables were entered using the stepwise forward method. Model 1: regular school in comparison with all other placements; Model 2: special school in comparison with all other placements; Model 3: institution in comparison with all other placements; Model 4: home in comparison with all other placements.

entering institutions compared to boys (OR 1.31, 95% CI 1.00–1.71). Children with higher severity of ASD were 3.08 times more likely to be institutions when compared to mild/moderate severity (OR 3.08, 95% CI 2.48–3.81). Children with mothers who had a college degree or above were less likely to be in institutions when compared with the mothers without college degrees (OR 0.71, 95% CI 0.57–0.88). Children with families reside in the central or western region were more likely to be institutions compared to those reside in the eastern region (OR 1.36, 95% CI 1.09–1.70 for central; OR 1.62, 95% CI 1.17–2.23 for western).

For home placement, severity of ASD, maternal education, family income and the location of the family had an impact. The children with mothers who had a college degree or above were less likely to be in homes than the mothers who had no college degrees (OR 0.63, 95% CI 0.45–0.89). The children with high family income were less likely to be in homes than children from low-income family (OR 0.58, 95% CI 0.37–0.92). Children with higher severity of ASD were 2.19 times more likely to be at their homes when compared to mild/moderate severity (OR 2.19, 95% CI 1.59–3.03). Children whose families reside in the western region were more likely to be homes compared to the eastern region (OR 1.60, 95% CI 1.01–2.57).

3.3. The association between severity of ASD and family income in the educational placements

Model 5 and model 6 in Table 3 added the interaction between severity of ASD and family income. Compared with children with mild/moderate severity in low-income family, the odds ratio to be in regular school of severe children in high income families were lower, with OR of 0.30 (95%CI 0.16–0.56). CCompared with children with mild/moderate severity in low-income family, the odds ratio to be in expensive institutions of severe children in high income families were higher, with OR of 3.43 (95%CI 1.72-6.84) (Table 3). Figure 3 further illustrated the interaction between severity of ASD and family income, which presented that regular school's negative slope with respect to severity was steeper for high income family than for low-income family, and expensive institution's slope was in different directions. It indicated that as the level of severity of ASD increased, the possibility for regular schools' placement for children from high-income families decreased faster than whom from low-income families. What's more, as the level of severity of ASD increased, the possibility for expensive institution placement for children from high-income family increased faster.

Table 3. household income difference of the association between comorbidities and accommodations

Characteristics	Model 5 regular school			Model 6 institute (expensive ones)		
	OR	95%CI		OR	95%CI	
		Low	High		Low	High
Sex						
Boy	1.00			1.00		
Girl	0.74	0.57	0.97	1.14	0.80	1.61
Age group						
6-8 years	1.00			1.00		
9-12 years	1.50	1.19	1.89	0.23	0.15	0.37
Only child						
No	1.00			1.00		
Yes	1.19	0.99	1.44	1.17	0.90	1.52
Severity						
Mild/moderate	1.00			1.00		
severe	0.22	0.15	0.32	1.37	0.80	2.53
Maternal Education level						
High school and below	1.00			1.00		
College degree or higher	1.42	1.15	1.76	1.05	0.77	1.41
Household Income						
Low	1.00			1.00		
Middle	1.37	1.07	1.76	1.44	0.96	2.17
High	2.13	1.62	2.80	1.34	0.87	2.06
Severity*Household Income						
Low* Mild/moderate	1.00			1.00		
Middle *Severe	0.78	0.45	1.35	1.74	0.88	3.43
High * Severe	0.30	0.16	0.56	3.43	1.72	6.84
Resident district						
Eastern	1.00			1.00		
Central	0.84	0.67	1.04	0.86	0.64	1.17
Western	0.76	0.55	1.03	1.39	0.93	2.09

OR: odds ratio; CI: confidence interval. Independent variables were entered using the stepwise forward method. Model 5: regular school in comparison with all other placements; Model 6: expensive institute in comparison with all other placements.

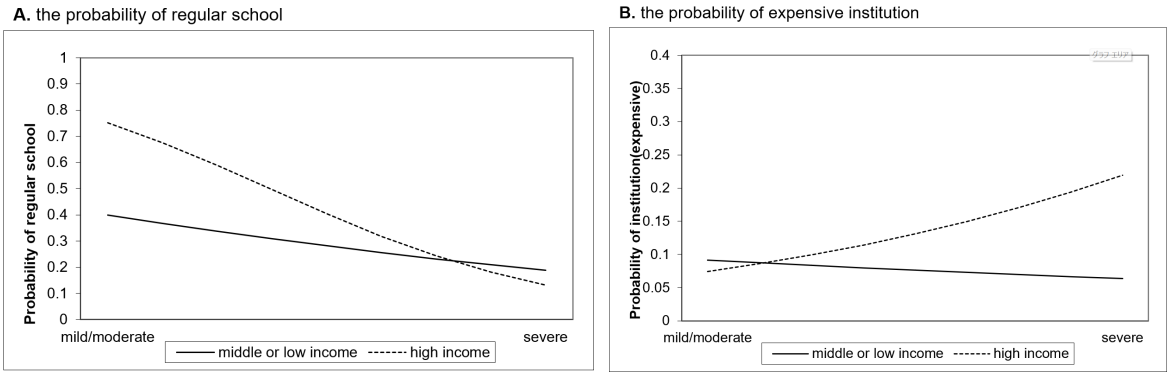

Figure 3. Predicted probability of regular school (A) and expensive institutions (B) by severity of ASD and household income.

4. Discussion

This study revealed the socioeconomic differences in the educational placement of ASD children in China. This was the first study to investigate potential socioeconomic disparities within Chinese families and explore variations in educational placement among children with different severity levels of ASD, providing new insights to the field. A total of 48.26% of children have entered regular school, while 29.86% have entered institutions, with a relatively low proportion entering special schools (13.88%) and being their homes (7.99%). Similar to previous research (*23,24*), this study found that the majority of the mild ASD group were placed in regular school. Including children with special education needs in mainstream classrooms was found to benefit their academic and social skills, as well as their well-being (*25,26*). However, inclusive education was seen as challenging since educational systems usually emerged from highly particular circumstances, both in terms of practice and policy, making each one distinct in its operation (*27*). Systems of inclusive education were often integrated into frameworks for both special education and mainstream education in a country.

This research highlights the presence of SES disparities in educational placements, which manifest in two ways. Prior research has mostly concentrated on one type of disparities, which is that attending regular schools is positively related with family SES (*12,28-31*). In this study, however, we differentiated the influences in two directions. For children with mild or moderate symptoms, regular schools were more accessible for those with high family SES. Conversely, for children with severe symptoms, expensive institutions were favored over regular schools for those with high family SES. No significant relationships were observed between SES and parental placement preferences (see Supplemental Table S2, *http://www.biosciencetrends.com/action/getSupplementalData.php?ID=186*). Most Chinese parents would like their child to attend a regular school, but we did find clear differences in parental SES and their educational placement. Previous studies found that parents of children with disabilities valued inclusive forms more than special schools (*11,32*). But individual outcomes may vary as the population is notably heterogeneous. From the perspective of parents' choices, it may be that regular school is a better arrangement for children with mild symptoms.

For children with severe symptoms, however, inclusive education is not a priority for families with high SES in China. The child's developmental level was considered to be a critical factor for successful engagement in inclusive settings by parents, teachers, and clinical practitioners (*9,33*). The demands of students perform well on academic tests may affect the school's quality for students with ASD (*34,35*). As children with mild or moderate severity appeared to do equally well across settings, whilst those with severe ASD made smaller gains in inclusive settings (*36*). Our study found that among those enrolled in regular schools, the majority of children with severe symptoms had a higher proportion of poor academic performance (see Supplemental Table S3, *http://www.biosciencetrends.com/action/getSupplementalData.php?ID=186*). This indicated that the regular schools are unable to meet these students' educational needs. Previous studies showed that the positive relationship between regular schools and higher parental higher education levels only occurred in mildly disabled children (*12*). Actually, parents' high SES might influence not the inclusive education decision, instead they would choose a more suitable institution for their children. High quality restrictive placements had many advantages, including access to distraction-free environments, specialized curriculum, behavioral supports, which were rarely realized in regular settings (*6,37*) and drove privileged families to pursue these placements. Similar with previous research (*6,38*), children with severe ASD symptom were more likely to be in a less-inclusive placement in China. This mainly due to the relatively average quality and limited quantity of special education in China (*39-41*).

From a supply perspective, regional resources affect the placement of children. In this study, children in the central or western were more likely to stay at home and

had less access to special schools, which might be related to insufficient local educational resources. Most studies have come to a conclusion that inclusive education was more frequently created in areas inhabited by more affluent people who have achieved higher levels of education (*28*), similar to our study. It was worth noting that rehabilitation resources were unevenly distributed in China (*21*). The resources distribution within a country affects health output and China's insufficient allocation of resources to the central and western regions may result in unfairness. These findings underscored the fact that older children with ASD in China were more likely to attend special schools and regular school than the younger groups, similar findings from previous studies for special school (*24*) and regular school (*42*). Although research have shown that children in their homes perform equally to or better than their conventionally educated peers (*18*), more than 70% of children in our sample with home placement received less than 2 hours a day of training at home.

What the government provides is not always the best, but the government's supply model often determines many things, especially for the poor. Parents with higher SES have more resources with which to implement their preferences and make it easier for them to meet expensive rehabilitation needs (*43,44*). Families with lower SES lacked the resources for sustained advocacy for less restrictive placements and expensive institutions (*28,45*). Improving the accessibility and quality of inclusive education, providing more high-quality special education institutions, may be the solution to the problem. What's more, simply discussing placement is not enough. Current inclusion practices might not benefit all children equally (*9,46-47*). The mere physical integration of autistic children in mainstream classrooms is widely considered insufficient for a successful educational experience (*25,48*), but that does not mean that the solution is to place them in a segregated placement. It must move toward ensuring students with ASD are served in inclusive, general education classrooms, where they can access academic instruction, meaningful interactions and relationships with peers, and supplementary aids and services (*49-51*).

It is critical to provide inclusive education in mainstream schools. Especially given that ASD symptoms are not binary, but rather a continuous continuum, there are still a large number of youngsters who have not been identified with autistic symptoms. However, the creation of inclusive education is a complicated process that may necessitate incremental progress. For example, many students with ASD and a normal intelligence quotient (IQ) but impaired social skills are not eligible for LRC plan in China (*52*). ASD should be considered as an independent special needs education category in order to address these practical issues. The research findings are important for the development of the concept of equity in inclusive education, as well as for helping policymakers focus on more vulnerable people.

There are limitations of this study. First, the data were only relevant to China. In countries with longer and more deeply developed inclusive education traditions, school systems may differ. Second, data on placement were based on parent-reported historical information. There is always the possibility that parents do not remember information accurately, or that they have misinformation about placement and services. Third, there is a need for a more comprehensive analysis of inclusivity. It is important to explore the extent to which children and adolescents with ASD are included in regular schools, whether on a full-time or part-time basis. Further research should aim to deepen our understanding of inclusivity in this context. Fourth, this study focuses on primary school samples, but it is crucial to acknowledge that there may be significant differences in educational facilities for older children. Therefore, future research should include a detailed analysis of educational placement in secondary schools.

In conclusion, this research revealed two socioeconomic disparities in the placement of children with ASD. For severe cases, high SES families tend to choose expensive institutions, while low SES families may opt for special schools or home-based education. For mild to moderate cases, low SES families have less access to regular schools compared to high SES families. To promote equal access to educational services for all families of children with ASD, it is crucial to enhance the availability of inclusive schools or classes, increase the number of high-quality special schools or institutions. Furthermore, future research should focus on strengthening the education of children with ASD, seeking placement facilities and educational intervention methods that are more suitable for children with different symptoms.

Acknowledgements

The authors are grateful to Liu Daiyue and Xu Disha at ALSOLIFE for their contributions.

Funding: All phases of this study were supported by the Key Realm R&D Program of GuangDong Province (2019B030335001) & China University IUR Innovation Foundation (Dezhou) (2021DZ021).

Conflict of Interest: The authors have no conflicts of interest to disclose.

References

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders: DSM 5. American Psychiatric Publishing, New York, US. 2013; pp. 50-58.
2. Zhou H, Xu X, Yan W, *et al*. Prevalence of Autism

Spectrum Disorder in China: A Nationwide Multi-Center Population-based Study Among Children Aged 6 to 12 Years. Neurosci. Bull. 2020; 36:961-971.
3. Bitterman A, Daley TC, Misra S, Carlson E, Markowitz J. A national sample of preschoolers with autism spectrum disorders: special education services and parent satisfaction. J Autism Dev Disord. 2008; 38:1509-1517.
4. Kefallinou A, Symeonidou S, Meijer CJW. Understanding the value of inclusive education and its implementation: A review of the literature. Prospects. 2020; 49:135-152.
5. Simón C, Palomo R, Echeita G. The duty to promote an inclusive educational system: A phenomenological study on the assessment procedures of pupils with special educational needs in Madrid (Spain). International Journal of Inclusive Education. 2021; 1:1-17.
6. White SW, Scahill L, Klin A, Koenig K, Volkmar FR. Educational placements and service use patterns of individuals with autism spectrum disorders. J Autism Dev Disord. 2007; 37:1403-1412.
7. Segall MJ and Campbell JM. Factors influencing the educational placement of students with autism spectrum disorders. Research in Autism Spectrum Disorders. 2014; 8:31-43.
8. Spaulding CJ, Lerner MD, Gadow KD. Trajectories and correlates of special education supports for youth with autism spectrum disorder and psychiatric comparisons. Autism. 2017; 21:423-435.
9. Humphrey N, Symes W. Inclusive education for pupils with autistic spectrum disorders in secondary mainstream schools: Teacher attitudes, experience and knowledge. International Journal of Inclusive Education. 2013; 17:32-46.
10. Lauderdale-Littin S, Howell E, Blacher J. Educational placement for children with autism spectrum disorders in public and non-public school settings: The impact of social skills and behavior problems. Education and Training in Autism and Developmental Disabilities. 2013; 48:469-478.
11. Turner S, Alborz A, Gayle V. Predictors of academic attainments of young people with Down's syndrome. J Intellect Disabil Res. 2008; 52:380-392.
12. Szumski G, & Karwowski M. School achievement of children with intellectual disability: The role of socioeconomic status, placement, and parents' engagement. Res Dev Disabil. 2012; 33:1615-1625.
13. Ajuwon PM, & Olu Oyinlade A. Educational placement of children who are blind or have low vision in residential and public schools: A national study of parents' perspectives. Journal of Visual Impairment & Blindness.2008; 6:325-339.
14. Irvin DW, McBee M, Boyd BA, Hume K, Odom SL. Child and family factors associated with the use of services for preschoolers with autism spectrum disorder. Research in Autism Spectrum Disorders. 2012; 6:565-572.
15. Suhrheinrich J, Brittney VR, Melgarejo M, Kelsey D, Vejnoska S, Reith SR. Exploring differences and disparities in school-based services received by students with autism: A systematic review. Research in Autism Spectrum Disorders. 2021; 85:101791.
16. Szumski G, Smogorzewska J, Grygiel P. Academic achievement of students without special educational needs and disabilities in inclusive education-Does the type of inclusion matter? PloS One. 2022; 6:17: e0270124.
17. Bouck EC, & Joshi GS. Does curriculum matter for secondary students with autism spectrum disorders: Analyzing the NLTS2. J Autism Dev Disord. 2015; 45:1204-1212.
18. Valiente C, Spinrad TL, Ray BD, Eisenberg N, Ruof A. Homeschooling: What do we know and what do we need to learn? Child Development Perspectives. 2022; 16:48-53.
19. Deng M. Inclusive education in the eyes of special needs education administrators: A study on the implementation of China's learning in regular classrooms policy. Ed Res and Experiment. 2004; 4:41-47. (in Chinese)
20. Xu Y, Zhu M. The "pain" and "difficulty" of inclusive education for children with autism in China. A J Modern Special Ed. 2016; 10:24-27. (in Chinese)
21. Zhao Y, Luo Y, Zhang R, Zheng X. Direct and indirect costs for families of children with autism spectrum disorder in China. Autism. 2023; 7:13623613231158862.
22. Su XY, Guo JJ, Wang XH. Different stakeholders' perspectives on inclusive education in China: parents of children with ASD, parents of typically developing children, and classroom teachers. International Journal of Inclusive Education. 2020; 24:968-963.
23. Towle PO, Vacanti-Shova K, Shah S, Higgins-D'alessandro A. School-aged functioning of children diagnosed with autism spectrum disorder before age three: Parent-reported diagnostic, adaptive, medication, and school placement outcomes. J Autism Dev Disord. 2014; 44:1357-1372.
24. Rattaz C, Munir K, Michelon C, Picot MC, Baghdadli A; ELENA study group. School Inclusion in Children and Adolescents with Autism Spectrum Disorders in France: Report from the ELENA French Cohort Study. J Autism Dev Disord. 2020; 50:455-466.
25. Lai MC, Anagnostou E, Wiznitzer M, Allison C, Baron Cohen S. Evidence-based support for autistic people across the lifespan: Maximising potential, minimising barriers, and optimising the person–environment fit. Lancet Neurol. 2020; 19:434-451.
26. Van Mieghem A, Verschueren K, Petry K, Struyf E. An analysis of research on inclusive education: A systematic search and meta review. International Journal of Inclusive Education. 2018; 1-15.
27. D'alessio S, Watkins A. International comparisons of inclusive policy and practice: Are we talking about the same thing? Research in Comparative & International Education. 2009; 4:233.
28. Kurth J, Mastergeorge A, Paschall K. Economic and demographic factors impacting placement of students with autism. Education and Training in Autism and Developmental Disabilities. 2016; 51:3-12.
29. Skiba RJ, Simmons AB, Ritter S, Gibb AC, Chung CG. Achieving equity in special education: History, status and current challenges. Exceptional Children. 2008; 74:264-288.
30. Cosier M & Causton-Theoharis J. Economic and demographic predictors of inclusive education. Remedial & Special Education. 2011; 32:496-505.
31. Phillips DA, Johnson AD, Iruka IU. Early care and education settings as contexts for socialization: New directions for quality assessment. Child Development Perspectives. 2022; 16:127-133.
32. Bulletin N, Fall VN. National Study on Inclusion: Overview and Summary Report. 1995. pp:35-36.
33. Larcombe TJ, Joosten AV, Cordier R, Vaz S. Preparing children with autism for transition to mainstream school and perspectives on supporting positive school

experiences. J Autism Dev Disord. 2019; 49:3073-3088.
34. Kraemer BR, Odom SL, Tomaszewski B, Hall LJ, Dawalt L, Hume KA, Steinbrenner J, Szidon K, Brum C. Quality of high school programs for students with autism spectrum disorder. Autism. 2020; 24:707-717.
35. Snell-Rood C, Ruble L, Kleinert H, McGrew JH, Adams M, Rodgers A, Odom J, Wong WH, Yu Y. Stakeholder perspectives on transition planning, implementation, and outcomes for students with autism spectrum disorder. Autism. 2020; 24:1164-1176.
36. Vivanti G, Bent C, Capes K, Upson S, Hudry K, Dissanayake C; Victorian ASELCC Team. Characteristics of children on the autism spectrum who benefit the most from receiving intervention in inclusive versus specialized early childhood education settings. Autism Res. 2022; 15:2200-2209.
37. Causton-Theoharis JN, Theoharis GT, Orsait F, Cosier M. Does self-contained special education deliver on its promises? A critical inquiry into research and practice. Journal of Special Education Leadership. 2011; 24:61-78.
38. Lavelle TA, Weinstein MC, Newhouse JP, Munir K, Kuhlthau KA, Prosser LA. Economic burden of childhood autism spectrum disorders. Pediatrics. 2014; 133:e520-e529.
39. Sun X, Allison C, Auyeung B, Baron-Cohen S, Brayne C. A review of healthcare service and education provision of Autism Spectrum Condition in mainland China. Res Dev Disabil. 2013; 34:469-79.
40. Yang GX, Yang FY, HP Tan. National survey and policy analysis for teacher professional development in special education school. Shanghai: East China Normal University Press. 2014. China. pp:14-19.
41. Zhang D, Spencer VG. Addressing the Needs of Students with Autism and Other Disabilities in China: Perspectives from the Field. International Journal of Disability Development & Education.2015; 62:168-181.
42. Eaves LC, Ho HH. Young adult outcome of autism spectrum disorders. J Autism Dev Disord. 2008; 38:739-747.
43. Kalyanpur M, Harry B, Skrtic T. Equity and advocacy expectations of culturally diverse families' participation in special education. International Journal of Disability, Development and Education. 2000; 47:119-136.
44. Heward WL. Exceptional children. An instruction to special education (ninth edition). New Jersey: Pearson Education, US. 2009; pp. 26-30.
45. Wakelin MM. Challenging disparities in special education: Moving parents from disempowered team members to ardent advocates. Nw.j.l. & Soc.poly. 2008; 3:263-288.
46. Waddington EM, Reed P. Comparison of the effects of mainstream and special school on National Curriculum outcomes in children with autism spectrum disorder: An archive-based analysis. Journal of Research in Special Educational Needs. 2017; 17:132-142.
47. Agran M, Jackson L, Kurth JA, Ryndak D, Wehmeyer M. Why Aren't Students with Severe Disabilities Being Placed in General Education Classrooms: Examining the Relations Among Classroom Placement, Learner Outcomes, and Other Factors. Research and Practice for Persons with Severe Disabilities. 2020; 45:4-13.
48. Pellicano L, Bölte S, Stahmer A. The current illusion of educational inclusion. Autism. 2018; 22:386-387.
49. Giangreco MF. How Can a Student with Severe Disabilities Be in a Fifth-Grade Class When He Can't Do Fifth-Grade Level Work? Misapplying the Least Restrictive Environment. Research and Practice for Persons with Severe Disabilities. 2020; 45:23-27.
50. Jameson M, Hicks T, Lansey K, Kurth JA, Jackson L, Zagona AL, Burnette K, Agran M, Shogren K, Pace J, Gerasimova D. Predictions on the frequency and importance of social contacts across placements: A bayesian multilevel model analysis. Research and Practice for Persons with Severe Disabilities. 2022; 47:229-243.
51. Zagona AL, Kurth JA, Lockman Turner E, Pace J, Shogren K, Lansey K, Jameson M, Burnette K, Mansouri M, Hicks T, Gerasimova D. Ecobehavioral analysis of the experiences of students with complex support needs in different classroom types. Research and Practice for Persons with Severe Disabilities. 2022; 47:209-228.
52. Li M, Lin Y, Bao T, Zhao Q, Wang Y, Li M, Chen Y, Qian Y, Chen L, Zhu D. Inclusive education of elementary students with autism spectrum disorders in Shanghai, China: From the teachers' perspective. Biosci Trends. 2022; 16:142-150.

Received December 19, 2023; Revised January 25, 2024; Accepted February 4, 2024.

Address correspondence to:
Xiaoying Zheng, School of Population Medicine and Public Health, Chinese Academy of Medical Sciences/Peking Union Medical College; 100730, No.31, Road 3rd, Bei-Ji-Ge, Dongcheng District, Beijing, China.
E-mail: zhengxiaoying@sph.pumc.edu.cn

Released online in J-STAGE as advance publication February 8, 2024.

Original Article

DOI: 10.5582/bst.2023.01327

Automated machine learning-based model for the prediction of pedicle screw loosening after degenerative lumbar fusion surgery

Feng Jiang[1], Xinxin Li[1], Lei Liu[2], Zhiyang Xie[2], Xiaotao Wu[1,2], Yuntao Wang[1,2,*]

[1] Southeast University Medical College, Nanjing, Jiangsu, China;
[2] Department of Spine Surgery, Southeast University ZhongDa Hospital, Nanjing, Jiangsu, China.

SUMMARY The adequacy of screw anchorage is a critical factor in achieving successful spinal fusion. This study aimed to use machine learning algorithms to identify critical variables and predict pedicle screw loosening after degenerative lumbar fusion surgery. A total of 552 patients who underwent primary transpedicular lumbar fixation for lumbar degenerative disease were included. The LASSO method identified key features associated with pedicle screw loosening. Patient clinical characteristics, intraoperative variables, and radiographic parameters were collected and used to construct eight machine learning models, including a training set (80% of participants) and a test set (20% of participants). The XGBoost model exhibited the best performance, with an AUC of 0.884 (95% CI: 0.825–0.944) in the test set, along with the lowest Brier score. Ten crucial variables, including age, disease diagnosis: degenerative scoliosis, number of fused levels, fixation to S1, HU value, preoperative PT, preoperative PI-LL, postoperative LL, postoperative PT, and postoperative PI-LL were selected. In the prospective cohort, the XGBoost model demonstrated substantial performance with an accuracy of 83.32%. This study identified crucial variables associated with pedicle screw loosening after degenerative lumbar fusion surgery and successfully developed a machine learning model to predict pedicle screw loosening. The findings of this study may provide valuable information for clinical decision-making.

Keywords CT Hounsfield units, osteoporosis, lumbar degenerative disease, screw loosening, explainable machine learning

1. Introduction

Pedicle screw fixation is a commonly utilized surgical technique for thoracolumbar disease treatment, which can stabilize the spine before solid fusion and restore spinal balance (*1*). However, screw loosening is one of the common complications associated with this treatment (*2,3*) and may lead to fixation failure, chronic low back pain, non-union and pseudarthrosis, and in severe cases may even require revision surgery (*4-7*), affecting the patient's quality of life. Thus, it is crucial to prevent screw loosening.

Osteoporosis has been identified as the predominant risk factor for screw loosening. In the osteoporotic spine, the bone-screw interface tends to be unstable, resulting in diminished pullout force and cutout force. Clinical studies indicated a pedicle screw loosening rate of less than 15% in non-osteoporotic patients, whereas it could escalate to as much as 60% in osteoporotic patients (*6,8,9*). Dual-energy X-ray absorptiometry (DXA) is currently considered the gold standard for assessing bone mineral density (BMD), with osteoporosis defined by the lowest T-score≤-2.5 (*10,11*). To prevent screw loosening, most spine surgeons have opted for target patients with a T-score of ≤ -2.5 for the application of pedicle screw augmentation techniques (*2, 12-14*). However, the lumbar degenerative changes in patients with lumbar degenerative disease (LDD) can result in an overestimation of T-scores, leading to potential false negative results (*15,16*). As a consequence, DXA outcomes may misguide spine surgeons in their preoperative surgical planning. In recent years, preoperative computed tomography (CT) measurements of vertebral body Hounsfield unit (HU) values have been widely used for the prediction of screw loosening. HU values are measured in the vertebral body, at the midsagittal plane, central transverse plane, and transverse planes close to the superior and inferior endplates separately (*17,18*). In this process, the region of interest (ROI) is expanded as much as possible within the cancellous bone but excluding other bony structures, such as cortical, bony endplates, and osteophytes. The

confusion caused by pathological bone formations can be eliminated (*18*), and the specific BMD of cancellous bone can be measured (*17*). Clinical research showed that it was a better predictor of postoperative complications than the T-score (*10,11,15*), and its predictive performance was superior to that of DXA (*10,15,16*). In addition, it was reported that gender, age, number of fused segments, and imaging parameters (fixation to S1, sagittal imbalance) were associated with pedicle screw loosening (*19-21*). Nevertheless, previous studies on predicting screw loosening have predominantly relied on a single statistical approach, potentially limiting their predictive performance (*10,19*). A study by Da *et al*. reported an AUC of only 0.666, below 0.75, for predicting pedicle screw loosening using Hounsfield units in patients with LDD (*10*). Thus the predictive performance is insufficient to meet the clinical needs in the existing models.

Recently, there have been increasing reports that applying machine learning techniques to develop various disease prediction models could improve their predictive performance (*22,23*). Machine learning is an advanced predictive modeling technique founded in computer science, utilizing artificial intelligence to create algorithms trained on data to perform diverse tasks. By employing validation measures, machine learning enhances model robustness, enabling predictions beyond the scope of traditional inferential statistics (*24,25*). Machine learning has been applied to spinal deformity and tumor patients for enhancing the clinical decision-making process (*26,27*). Currently, there are no studies on machine learning models and postoperative pedicle screw loosening in degenerative lumbar fusion surgery. Therefore, in this study, we used multiple artificial intelligence algorithms to construct predictive models for screw loosening and compared these models to finalize the model with the best predictive performance to support clinical decision making.

2. Materials and Methods

2.1. Patient data collection

This study was approved by the Ethics Committee of Southeast University ZhongDa Hospital and conforms to the provisions of the Declaration of Helsinki (as

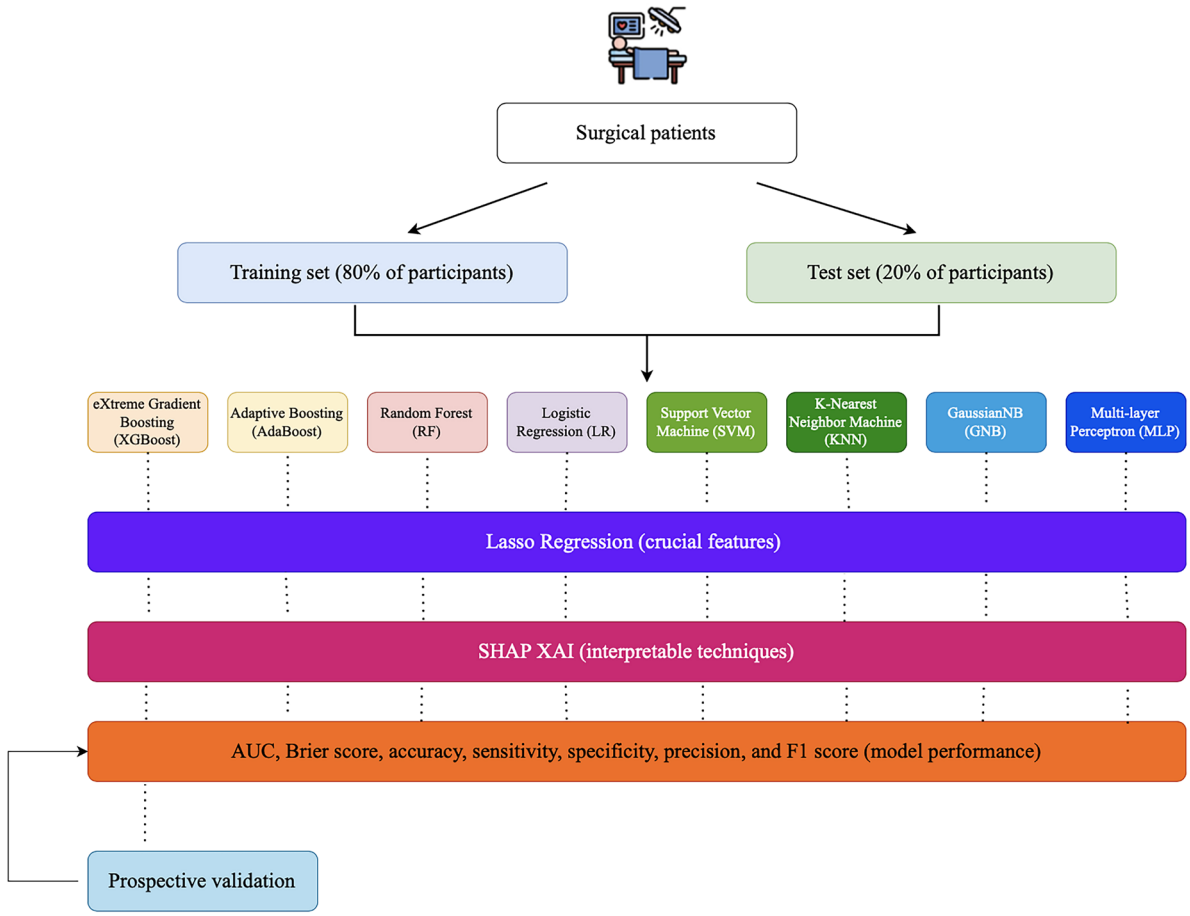

Figure 1. Flow chart of the study design. The figure shows the relevant data collected from patients undergoing surgery for degenerative lumbar fusion in hospitals, including demographic characteristics, radiological measurement parameters, and surgical information. A total of 32 variables were collected, out of which 10 non-zero features were selected through LASSO regression for building machine learning models. Subsequently, the model's performance was evaluated to determine the optimal predictive model. The data of 45 patients were prospectively collected for further validation. Finally, SHAP interpretability analysis was conducted based on the best predictive model.

revised in 2013). Informed consent was waived for this retrospective study. The workflow of our study design and its corresponding analyses are depicted in Figure 1. We retrospectively analyzed the clinical data and radiographic data of 552 patients who underwent primary transpedicular lumbar fixation for LDD at the Spine Surgery Center of Southeast University ZhongDa Hospital from January 2018 to December 2021. The inclusion criteria were as follows: *i*) patient's age over 50 years; *ii*) patients who underwent primary pedicle screw fixation for LDD, including lumbar disc herniation, degenerative lumbar spondylolisthesis, degenerative lumbar spinal stenosis, and degenerative lumbar scoliosis; *iii*) the number of fused levels ≤4 segments; *iv*) patients who underwent lumbar X-ray, CT, and DXA within 1 month prior to surgery at our institution; and *v*) patients were followed up for 3-12 months after surgery, and the follow-up data were complete. The exclusion criteria were as follows: *i*) patients with congenital spinal deformities, spinal trauma, spinal tumors, spinal tuberculosis, spine infection, ankylosing spondylitis, or a history of previous spinal surgery; *ii*) the presence of metabolic bone disease or long-term use of drugs such as corticosteroids that affect bone density; and *iii*) patients with screw loosening due to surgical site infection.

To further validate the model's accuracy, we prospectively collected data from patients who underwent primary transpedicular lumbar fixation for LDD at the Spine Surgery Center of Southeast University ZhongDa Hospital from January 2022 to April 2022.

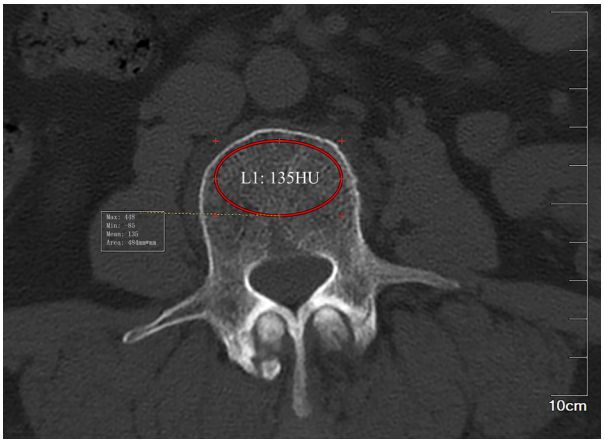

Figure 2. The measurement of HU value: the HU value of L1 was 135.

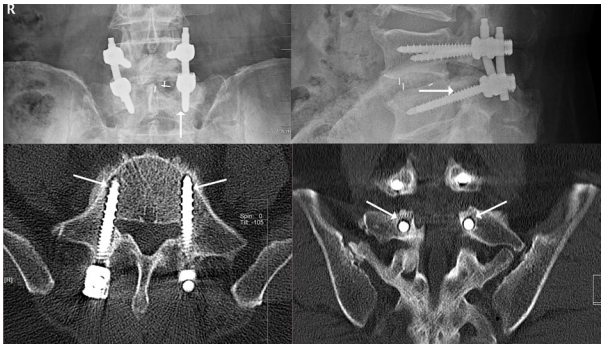

Figure 3. Postoperative follow-up radiographs and CT scans shows screw loosening.

2.2. Evaluation of BMD and screw loosening

All patients received DXA scanning and lumbar CT with three-dimensional reconstruction examination at our radiology center one month prior to the surgery. The tube voltage for the CT scan was 120 kV. The HU values for L1 – L4 were independently measured for each patient by two authors (FJ and XXL), adhering to the methodology outlined in prior studies (*16,28*). This involved placing an elliptical region of interest (ROI) on the central cross-sectional CT image of the vertebral body, with the inclusion of trabecular bone within the ROI and the exclusion of cortical bone, osteophytes, bone endplates, and the posterior venous plexus (Figure 2). Subsequently, the HU value was automatically calculated by the picture archiving and communication system (PACS). The mean HU values of L1 to L4 represented the lumbar BMD. In addition, DXA scans were conducted at the lumbar vertebrae (L1 – L4), as well as the total hips and femoral necks, and the lowest lumbar BMD and the lowest T-score were documented for subsequent analysis.

Patients were followed up with a lumbar X-ray at 3–12 months postoperatively. Lumbar CT was not routinely conducted throughout the follow-up duration; consequently, if abnormalities were detected on the lumbar X-ray, supplementary lumbar CT scans were conducted to confirm the presence of screw loosening (Figure 3). In the current study, screw loosening was defined as the presence of a radiolucent zone with a minimum width of 1 mm around the pedicle screw on radiographs taken during the 3–12 month follow-up period (*29,30*). Patients were categorized into two groups based on the presence or absence of screw loosening at the 12-month follow-up examination: the loosening group and the non-loosening group.

In order to assess reliability, a random selection of 30 patients was made to evaluate the measurement of HU values and the judgment of screw loosening. Two authors independently measured the HU values of LI – L4 for all patients and judged screw loosening for each patient. Two weeks later, the HU values of these 30 patients were measured again and screw loosening was reevaluated. Throughout the process of HU value measurement and screw loosening assessment, two authors were kept blinded to both the DXA results of the patients and the measurements recorded by the other author.

2.3. Lumbar X-ray assessment

Patients underwent lumbar X-ray examination one month before surgery and prior to discharge. The lumbar lordosis (LL), pelvic incidence (PI), pelvic tilt (PT),

sacral slope (SS), and the difference between pelvic incidence and lumbar lordosis (PI-LL) were measured and recorded.

2.4. Model input features and model development

We collected 32 potential characteristics, including basic patient characteristics: age, gender, height, weight, BMI, hypertension, diabetes, history of smoking, and history of alcoholism; surgery-related information: duration of surgery, intraoperative blood loss, number of fused levels, fixation to S1, hospitalization time; and preoperative and postoperative radiographic parameters. To identify the crucial factors attributed to screw loosening, the least absolute shrinkage and selection operator (LASSO) technique was employed for feature selection (*31,32*).

In order to maximize predictive performance, we developed eight machine learning models: the eXtreme Gradient Boosting (XGBoost) algorithm, Adaptive Boosting (AdaBoost), Random Forest (RF), Logistic Regression (LR), Support Vector Machine (SVM), K-Nearest Neighbor Machine (KNN), GaussianNB (GNB), and Multi-layer Perceptron (MLP).

2.5. Sample size and statistical analysis

For the binary prediction model, the sample size calculation formula is obtained according to the previous study (*33*), which is:

$$N = \exp\left(\frac{-0.508+0.259\ln(\varphi)+0.504\ln(P)-\ln(MAPE)}{0.544}\right)$$

Here, φ denotes the ratio of positive events, P denotes the number of model input features, and $MAPE$ denotes the mean absolute percentage error between the observed and actual outcome probability. Based on the above formula, the minimum sample size was estimated to be 406. Thus, we performed a random partition of the complete dataset (n = 552) into a training set (n = 442) and a test set (n = 110) using an 8:2 ratio.

In this study, all analyses were performed using Python version 3.9.0 (*34*). Interobserver and intraobserver reliability of the HU values were assessed using the Intraclass correlation coefficient (ICC). Excellent reliability was defined as ICC ≥ 0.8. The agreement in determining screw loosening was evaluated using a kappa statistics test. The Shapiro-Wilk test was used to test the normality of the distribution of continuous variables. Continuous variables that conformed to a normal distribution were reported as mean ± standard deviation (SD) and compared using independent-samples *t*-test. Continuous variables that were not normally distributed were expressed as the median and interquartile range (IQR) and analyzed using the Mann-Whitney *U* test. Categorical variables were described as frequencies and percentages, and compared using chi-square tests or Fisher's exact probability tests. Lastly, crucial features were selected through LASSO regression analysis, and based on these features, eight models were developed.

For the selection of model hyperparameters, ten-fold cross-validation was performed on the training datasets. The approach for handling missing data was as follows: missing values were imputed using the random forest regression method if the percentage of missing values was less than 20%; otherwise, the missing cases were excluded from the analysis. The predictive performance of the model was assessed through discrimination and calibration. Discrimination was quantified using the AUROC and Brier score, and model performance was assessed by accuracy, sensitivity, specificity, precision, and F1 score. The Brier score, representing the average squared difference between predicted probabilities and true labels, served as an indicator of model performance, with lower scores indicating higher accuracy. Following the identification of the optimal model, the Python-based SHAP package was utilized to illustrate the significance of individual features (*35*). At last, the selected model was employed to visualize prospective validations.

3. Results

3.1. Patient characteristics and pedicle screw loosening rates

A total of 552 patients were included in this study. Patients were divided into the loosening group (n = 128) and the non-loosening group (n = 424) based on the presence or absence of screw loosening within 12 months of postoperative follow-up. Table 1 shows the demographic characteristics and surgical information of the study participants who underwent transpedicular lumbar fixation surgery for LDD. The radiographic data for the loosening group and the non-loosening group are shown in Table 2. The incidence rate of pedicle screw loosening was approximately 23.19%. The reliability of interobserver and intraobserver measurements of HU value was deemed excellent, as indicated by ICC values of 0.88 and 0.86, respectively. The determination of screw loosening demonstrated high agreement, with kappa values of 0.79 and 0.76, respectively. There were few missing values for the study variables. No statistically significant difference was found in the patient characteristics between the training and test datasets.

3.2. Crucial features

The optimal parameter (lambda) for the LASSO model selection was determined using ten-fold cross-validation. With the optimal lambda, ten features demonstrated non-zero coefficients (Figure 4), encompassing age, disease diagnosis: degenerative scoliosis, number of fused levels, fixation to S1, HU value, preoperative PT,

Table 1. Demographic characteristics and clinical information of the study participants who underwent transpedicular lumbar fixation surgery

Variables	All (*n* = 552)	Loosening group (*n* = 128)	Non-loosening group (*n* = 424)	*p*-Value
Age, median (Q1, Q3)	61 (55, 70)	69 (64, 74)	58 (54, 66)	< 0.001
Sex, *n* %				0.105
Female	315 (57.065)	81 (63.281)	234 (55.189)	
Male	237 (42.935)	47 (36.719)	190 (44.811)	
Height, median (Q1, Q3)	162 (158,170)	161 (158, 170)	163 (158, 170)	0.107
Weight, median (Q1, Q3)	67.5 (60, 75)	65 (60, 72.5)	68 (60, 75)	0.110
BMI, median (Q1, Q3)	24.65 (22.86, 26.74)	24.615 (22.823, 26.725)	24.770 (22.883, 26.823)	0.584
Hypertension, n %	224 (40.580)	59 (46.094)	165 (38.915)	0.147
Diabetes, *n* %	98 (17.754)	29 (22.656)	69 (16.274)	0.098
Alcohol, *n* %	94 (17.029)	22 (17.188)	72 (16.981)	0.957
Smoking, *n* %	35 (6.341)	7 (5.469)	28 (6.604)	0.644
Primary diagnosis				
Lumbar disc herniation, *n* %	185 (33.514)	36 (28.125)	149 (35.142)	0.141
Lumbar degenerative spondylolisthesis, *n* %	145 (26.268)	30 (23.438)	115 (27.123)	0.406
Lumbar spinal stenosis, *n* %	191 (34.601)	47 (36.719)	144 (33.962)	0.566
Degenerative scoliosis, *n* %	31 (5.616)	15 (11.719)	16 (3.774)	< 0.001
Number of fused levels, *n* %				< 0.001
1	199 (36.051)	27 (21.094)	172 (40.566)	
		13.57		
2	188 (34.058)	45 (35.156)	143 (33.726)	
		23.94		
3	103 (18.659)	28 (21.875)	75 (17.689)	
		27.18		
4	62 (11.232)	28 (21.875)	34 (8.019)	
		45.16		
Intraoperative blood loss, median (Q1, Q3)	250.0 (150.0, 400.0)	250.0 (150.0, 462.5)	250.0 (100.0, 350.0)	0.103
Duration of surgery, median (Q1, Q3)	165.0 (133.75, 200.0)	180.0 (148.75, 206.25)	162.5 (130.0, 200.0)	0.052
Fixation to S1, *n* %	268 (48.551)	79 (61.719)	189 (44.575)	< 0.001
Hospitalization time, median (Q1, Q3)	11 (9, 13)	11.5 (9, 14)	11 (9, 13)	0.128

Table 2. The preoperative and postoperative radiographic data for the loosening group and the non-loosening group

Variables	All (*n* = 552)	Loosening group (*n* = 128)	Non-loosening group (*n* = 424)	*p*-Value
The lowest lumbar BMD, median (Q1, Q3)	1.045 (0.947, 1.189)	1.030 (1.028, 1.056)	1.045 (0.929, 1.321)	0.044
The lowest T-score, median (Q1, Q3)	-1.9 (-2.7, -0.6)	-2.45 (-3.2, -0.2)	-1.8 (-2.7, -1)	0.081
HU value, median (Q1, Q3)	138.25 (113.5, 164.0)	96.875 (76, 114.563)	148.250 (126, 169.688)	< 0.001
lumbar instability, *n* %	187 (33.877)	41 (32.031)	146 (34.434)	0.615
Preoperative LL, median (Q1, Q3)	42 (33, 50)	39 (31, 51.25)	42 (35, 49.25)	0.106
Preoperative PI, median (Q1, Q3)	48 (41, 54)	51.7(44, 60.25)	50.6 (44, 63)	0.763
Preoperative PT, median (Q1, Q3)	17 (13, 23)	23 (16, 30)	16 (13, 21)	< 0.001
Preoperative SS, median (Q1, Q3)	30 (25, 36)	30 (23.75, 37)	30 (25, 36)	0.832
Preoperative PI-LL, median (Q1, Q3)	6.0 (3.0, 9.0)	10.5 (3, 22)	5 (2, 8)	< 0.001
Postoperative LL, median (Q1, Q3)	43.0 (36, 50)	46 (38, 53)	42 (36, 49)	< 0.001
Postoperative PI, median (Q1, Q3)	48 (42, 56)	52 (45, 60)	51 (44.75, 58)	0.778
Postoperative PT, median (Q1, Q3)	13 (9, 16)	15 (10, 20)	12 (9, 15)	< 0.001
Postoperative SS, median (Q1, Q3)	36 (30, 41)	31.6 (26.6, 37)	33 (27, 37)	0.384
Postoperative PI-LL, median (Q1, Q3)	6.0 (2.0, 9.0)	10 (2.5, 17)	5 (2, 9)	< 0.001

Abbreviations: BMD, bone mineral density; HU, Hounsfield unit; LL, lumbar lordosis; PI, pelvic incidence; PT, pelvic tilt; SS, sacral slope.

preoperative PI-LL, postoperative LL, postoperative PT, and postoperative PI-LL.

3.3. Model performance

Eight machine learning algorithms were used to construct prediction models for screw loosening, and the predictive performance of each model was evaluated by calculating Brier scores and AUROC. In comparison to the other models, XGBoost demonstrated the lowest Brier score, as illustrated in Figure 5A, which presented the calibration plots for all eight models. Furthermore, the XGBoost model outperformed the others with a higher AUROC, as shown in Figure 5B. Based on the AUROC values of the eight models, a forest plot illustrating the AUC scores for multiple models was generated (Figure 5C). Through ten-fold cross-validation, the XGBoost model achieved a smaller standard deviation of 0.036 for

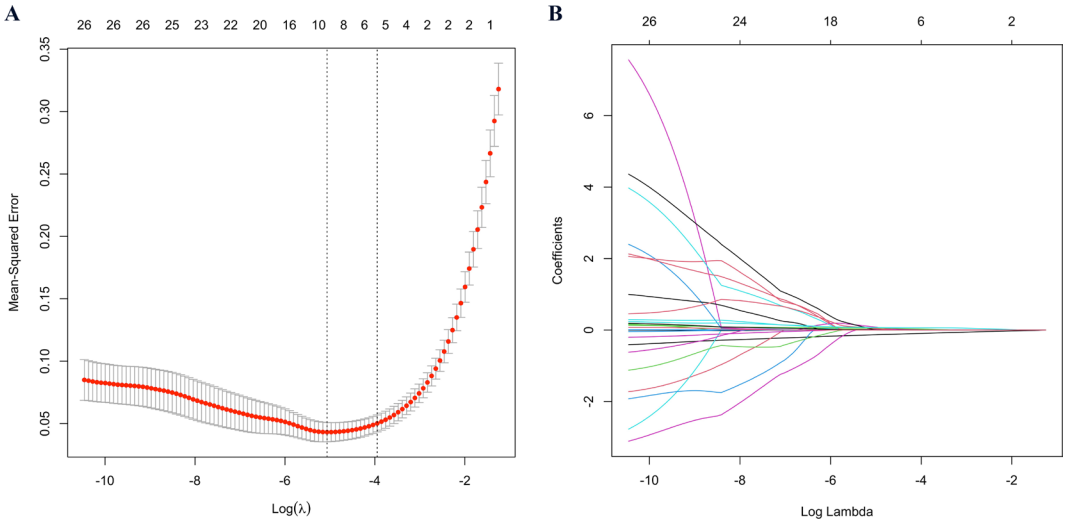

Figure 4. Clinical and radiographic feature selection using the LASSO regression. (A) LASSO coefficient profiles of 32 features. **(B)** Feature selection for the predictive model. Turning parameter (λ) selection used tenfold cross-validation. The vertical axis shows the model misclassification rate, and the horizontal axis shows the log(λ). The two vertical dashed lines represent the minimum value and one standard deviation on one side from the minimum value.

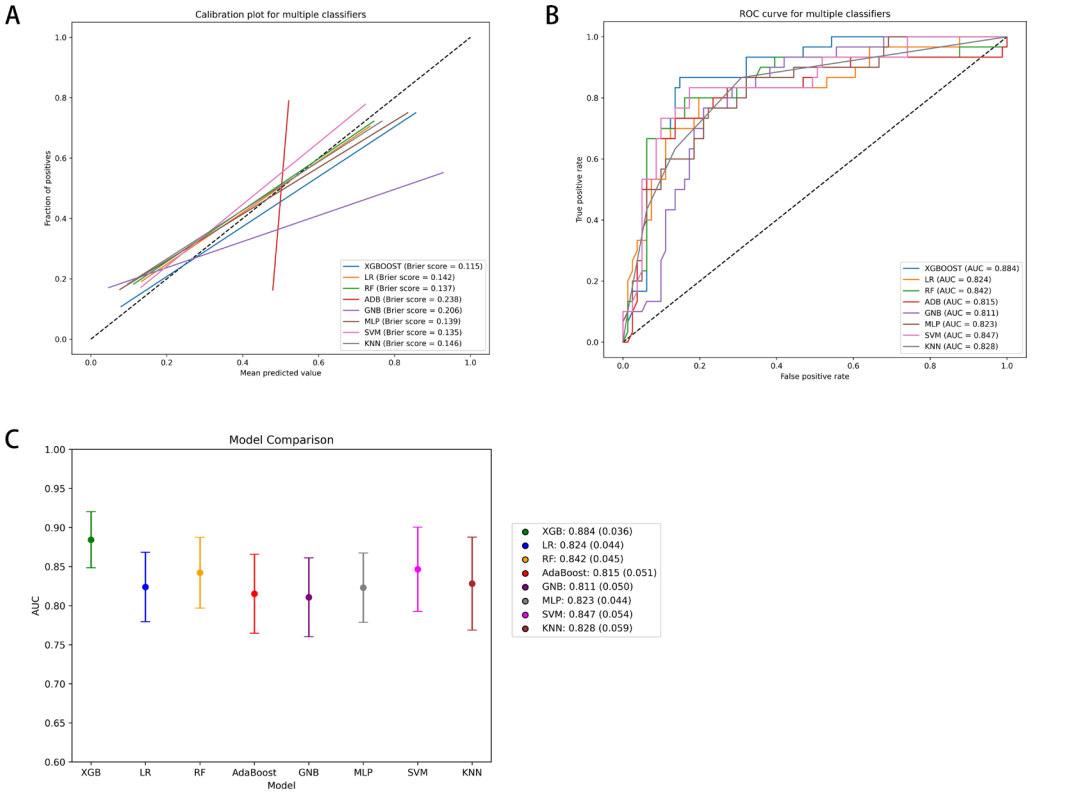

Figure 5. Model performance. (A) Calibration plots of the eight models. **(B)** Receiver-operating characteristic curves for the eight models. **(C)** Forest plot of the AUC score and 95 CI% of the eight models.

its AUC score. This outcome suggested that the XGBoost model exhibited the most stable performance compared to the other seven models. The performance metrics of the eight models in the test dataset are presented in Table 3.

In Figure 6, the XGBoost model was analyzed using the SHAP method. This figure provided a clear understanding of the contribution of each feature to the model output. Additionally, the bar chart illustrated the magnitude of the impact that the feature importance had on the model predictions.

3.4. Application of the model

The SHAP waterfall and force plots for the XGBoost model are shown in Figure 7. Inputting the clinical information of a typical patient into the model, for example, in Figure 7A, the true outcome of the

Table 3. Performance metrics for eight models in the test dataset

Model	Accuracy	Sensitivity	Specificity	Precision	F1 score
XGBoost	0.847	0.600	0.938	0.783	0.679
LR	0.793	0.767	0.802	0.706	0.667
RF	0.829	0.800	0.840	0.714	0.716
AdaBoost	0.782	0.732	0.801	0.729	0.612
GNB	0.775	0.633	0.827	0.576	0.638
MLP	0.802	0.600	0.877	0.643	0.621
SVM	0.829	0.733	0.864	0.667	0.698
KNN	0.802	0.633	0.864	0.633	0.633

Abbreviations: XGBoost, eXtreme Gradient Boosting; LR, Logistic Regression; RF, Random Forest; AdaBoost, Adaptive Boosting; GNB, Gaussian Naive Bayes; MLP, Multi-layer Perceptron; SVM, Support Vector Machine; KNN, K-Nearest Neighbor Machine.

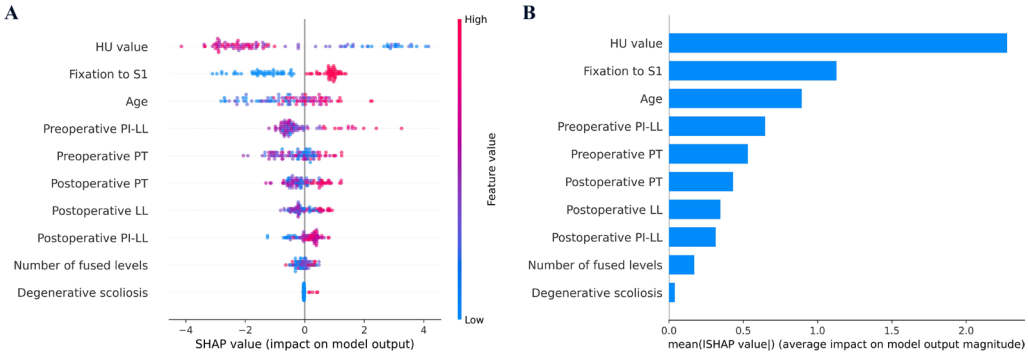

Figure 6. SHAP analysis of the XGBoost model. (A) SHAP summary plot of clinical features, with red indicating higher values and blue indicating lower values. **(B)** Importance matrix plot of the XGBoost model, indicating the importance of each variable in predicting postoperative screw loosening.

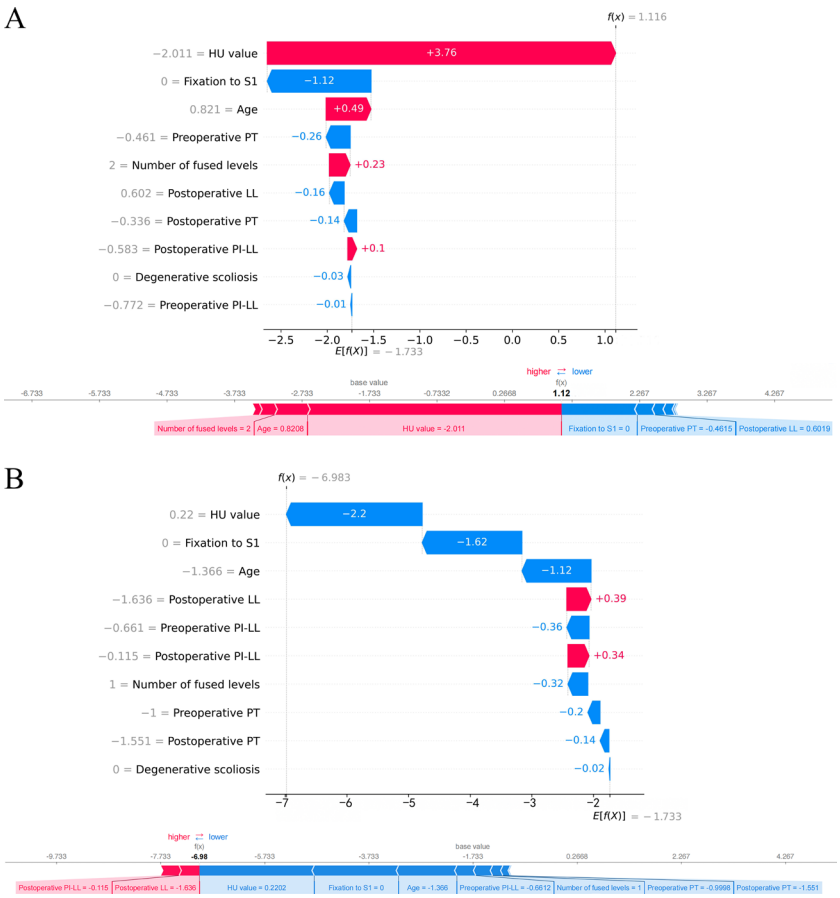

Figure 7. SHAP force plot for patients in the dataset at (A) high or (B) low risk of postoperative screw loosening.

patient was screw loosening, and the predicted value of the model was 1.12, with a predicted probability of 75.40% for screw loosening; in Figure 7B, the true outcome of the patient was no screw loosening, and the predicted value of the model was -6.98, with a predicted probability of 0.09% for screw loosening.

3.5. Prospective validation

A total of 45 patients were enrolled in this study for prospective validation, with 20.0% (9/45) of them experiencing screw loosening. The proposed model achieved an accuracy of 83.32% when tested on the prospective dataset, with a respective sensitivity of 0.543 and specificity of 0.940.

4. Discussion

Few models are available to predict screw loosening after degenerative lumbar fusion surgery in patients with LDD. In this study, we applied machine learning methods to identify risk factors for pedicle screw loosening after degenerative lumbar fusion surgery and to develop risk prediction models for screw loosening. The performance of eight machine learning models was compared. The results showed that the XGBoost model had the highest AUC (88.4%). The calibration of the models was quantitatively compared using Brier scores. The calibration of the XGBoost model showed good agreement between the prediction outcome and the actual observed outcome. The standard deviation of the AUC score of the XGBoost model obtained after using ten cross-validations was 0.036, which was smaller than the other seven models, suggesting that the XGBoost model has the most stable performance. Based on the above aspects, it can be concluded that the XGBoost model exhibited superior performance compared to seven other machine learning models. The SHAP method further explained the predictors and model prediction performance. It provided a simple and robust method for individualized prediction of pedicle screw loosening after degenerative lumbar fusion surgery, which can provide important information for medical decision support.

The rate of screw loosening in the current study was observed to be 23.19%. In a study conducted by Tokuhashi *et al*. (*36*), they reported a screw loosening rate of 26.8% at the 12-month follow-up in patients with LDD who underwent pedicle screw fixation in the lumbar spine. Additionally, Shin *et al*. (*19*) found a screw loosening rate of 22.5%. The occurrence of screw loosening is caused by a variety of factors (*3,7,19,37*). With a total of ten variables included in the XGBoost model analysis, we found that potential risk factors for screw loosening were associated with low BMD, older age, fixation to S1, multi-segment fixation, and sagittal imbalance.

Osteoporosis is the most commonly discussed cause of screw loosening. This is because in patients with osteoporosis, the screw-bone interface has a lower ability to bind the screw, leading to reduced screw pullout strength. A biomechanical study demonstrated that decreased bone density resulted in a decline in screw pullout force, ultimately leading to the occurrence of screw loosening (*38*). Osteoporosis is typically diagnosed using the standard technique of DXA. Previous studies have shown a difference in DXA between patients undergoing lumbar fusion surgery with and without screw loosening (*20,39*), but Kim *et al*. reported no difference in DXA between the two groups (*37*). These contradictory findings can be attributed to the inaccuracies of DXA in evaluating BMD. Degenerative changes in the lumbar spine may lead to overestimation of BMD, particularly in patients with severe lumbar degeneration (*28*). Hence, in this study, we used lumbar CT to measure the HU value of the vertebral body and recorded the T-score and lumbar BMD results of DXA. The results of the study suggested that the screw loosening rate was higher in patients who possessed a low HU value than in those who had a high HU value, but that the T-score and lumbar BMD value of the DXA performed poorly in recognizing screw loosening. Furthermore, since BMD decreases significantly with age, our findings also suggested that screw loosening was more common in older patients. Therefore, we should not concentrate only on the DXA results when making a surgical strategy for lumbar fixation in elderly patients with LDD. We recommend routinely measuring the HU value for surgical planning in LDD patients.

Some studies have illustrated the significance of S1 fixation in the occurrence of screw loosening. Serving as a critical connection between the spine and pelvis, S1 exhibits a greater susceptibility to loosening due to its longer lever arm and larger physiological arc under fixed stress (*20,40*). The anatomical attributes of the S1 pedicle, characterized by a larger diameter and shorter length compared to lumbar pedicles, and the presence of predominantly cancellous bone within the sacrum, collectively point towards a heightened susceptibility of S1 screw loosening (*41*). These anatomical factors likely contribute to the increased incidence of screw loosening in the S1 region.

Multiple-segment screw fixation have consistently identified as a notable risk factor associated with screw loosening (*20,36,42*). According to the study by Zou *et al*. (*43*), the rate of screw loosening in single-level procedures was found to be 4.1%, while it increased to 33.3%, 53.3%, and 78.8% in two-level, three-level, and four-level procedures, respectively. In our own investigation, we observed a similar trend; specifically, an escalating rate of screw loosening was observed with an increasing number of screw fixation levels. This rise in the incidence of screw loosening with the

more segments for screw fixation can be attributed to the amplified cantilever bending moment exerted on the surgical construct (*3,7,20,42*). Notably, screw loosening frequently occurred at the distal end of the screw instrumentation in patients undergoing multi-level fixation (*44*).

The presence of sagittal imbalance can contribute to an elevated risk of screw loosening. In our study, we identified postoperative LL as a predictive factor for screw loosening. Livshits *et al*. (*45*) demonstrated that restoring postoperative LL was associated with a reduced incidence of screw loosening. Kuo *et al*. (*46*) found that, even among patients undergoing dynamic stabilization, the loss of LL postoperatively increased the rate of screw loosening. In this study, the loosening group also exhibited higher preoperative and postoperative PT compared to the non-loosening group. The posterior tilting of the pelvis (increased PT) could be a compensatory response to sagittal malalignment (*47*). Furthermore, studies have shown that PI-LL mismatch is an important indicator of sagittal balance and is associated with adjacent segment degeneration, screw loosening, and disability and quality of life scores (*48,49*). In our study, we revealed that preoperative and postoperative PI-LL mismatch was a significant predictive feature of screw loosening.

In this study, we used prospective data to validate the predictive performance of the model, but there are still some limitations of this study. First, given the small sample size of this study, further research with a larger sample is needed to validate the predictive model. Second, the data in this study was collected from a single large academic medical center. Consequently, the generalizability of this model to other medical institutions may be limited. It is highly probable that recalibration of the model would be essential when implementing it in another institution, as the relative weights of the features may necessitate adjustments. Last, an independent dataset is indispensable to assess the model's extrapolation and generalization. To address this need, our future research endeavors will focus on acquiring an ample number of external validation datasets to further refine and enhance the performance of this model.

5. Conclusion

In this study, we developed eight different prediction models for postoperative screw loosening, among which the XGBoost model demonstrated good discrimination and overall performance. In addition, based on interpretable techniques, this model enables individualized prediction of postoperative screw loosening. We believe that this model is an important tool for identifying the postoperative occurrence of pedicle screw loosening in patients requiring degenerative lumbar fusion surgery.

Funding: None.

Conflict of Interest: The authors have no conflicts of interest to disclose.

References

1. Boos N, Webb JK. Pedicle screw fixation in spinal disorders: a European view. Eur Spine J. 1997; 6:2-18.
2. Hoppe S, Keel MJ. Pedicle screw augmentation in osteoporotic spine: indications, limitations and technical aspects. Eur J Trauma Emerg Surg. 2017; 43:3-8.
3. Galbusera F, Volkheimer D, Reitmaier S, Berger-Roscher N, Kienle A, Wilke HJ. Pedicle screw loosening: a clinically relevant complication? Eur Spine J. 2015; 24:1005-1016.
4. Alanay A, Vyas R, Shamie AN, Sciocia T, Randolph G, Wang JC. Safety and efficacy of implant removal for patients with recurrent back pain after a failed degenerative lumbar spine surgery. J Spinal Disord Tech. 2007; 20:271-277.
5. Berjano P, Bassani R, Casero G, Sinigaglia A, Cecchinato R, Lamartina C. Failures and revisions in surgery for sagittal imbalance: analysis of factors influencing failure. Eur Spine J. 2013; 22 Suppl 6:S853-858.
6. Bredow J, Boese CK, Werner CM, Siewe J, Löhrer L, Zarghooni K, Eysel P, Scheyerer MJ. Predictive validity of preoperative CT scans and the risk of pedicle screw loosening in spinal surgery. Arch Orthop Trauma Surg. 2016; 136:1063-1067.
7. Röllinghoff M, Schlüter-Brust K, Groos D, Sobottke R, Michael JW, Eysel P, Delank KS. Mid-range outcomes in 64 consecutive cases of multilevel fusion for degenerative diseases of the lumbar spine. Orthop Rev (Pavia). 2010; 2:e3.
8. Reis MT, Reyes PM, Bse, Altun I, Newcomb AG, Singh V, Chang SW, Kelly BP, Crawford NR. Biomechanical evaluation of lateral lumbar interbody fusion with secondary augmentation. J Neurosurg Spine. 2016; 25:720-726.
9. Bokov A, Bulkin A, Aleynik A, Kutlaeva M, Mlyavykh S. Pedicle Screws Loosening in Patients With Degenerative Diseases of the Lumbar Spine: Potential Risk Factors and Relative Contribution. Global Spine J. 2019; 9:55-61.
10. Zou D, Sun Z, Zhou S, Zhong W, Li W. Hounsfield units value is a better predictor of pedicle screw loosening than the T-score of DXA in patients with lumbar degenerative diseases. Eur Spine J. 2020; 29:1105-1111.
11. Mikula AL, Puffer RC, Jeor JDS, Bernatz JT, Fogelson JL, Larson AN, Nassr A, Sebastian AS, Freedman BA, Currier BL, Bydon M, Yaszemski MJ, Anderson PA, Elder BD. Teriparatide treatment increases Hounsfield units in the lumbar spine out of proportion to DEXA changes. J Neurosurg Spine. 2019; 1-6.
12. Xie Y, Fu Q, Chen ZQ, Shi ZC, Zhu XD, Wang CF, Li M. Comparison between two pedicle screw augmentation instrumentations in adult degenerative scoliosis with osteoporosis. BMC Musculoskelet Disord. 2011; 12:286.
13. Piñera AR, Duran C, Lopez B, Saez I, Correia E, Alvarez L. Instrumented lumbar arthrodesis in elderly patients: prospective study using cannulated cemented pedicle screw instrumentation. Eur Spine J. 2011; 20 Suppl 3:408-414.
14. Dai F, Liu Y, Zhang F, Sun D, Luo F, Zhang Z, Xu J.

Surgical treatment of the osteoporotic spine with bone cement-injectable cannulated pedicle screw fixation: technical description and preliminary application in 43 patients. Clinics. 2015; 70:114-119.
15. Blake GM, Fogelman I. The role of DXA bone density scans in the diagnosis and treatment of osteoporosis. Postgrad Med J. 2007; 83:509-517.
16. Pickhardt PJ, Pooler BD, Lauder T, del Rio AM, Bruce RJ, Binkley N. Opportunistic screening for osteoporosis using abdominal computed tomography scans obtained for other indications. Ann Intern Med. 2013; 158:588-595.
17. Gausden EB, Nwachukwu BU, Schreiber JJ, Lorich DG, Lane JM. Opportunistic Use of CT Imaging for Osteoporosis Screening and Bone Density Assessment: A Qualitative Systematic Review. J Bone Joint Surg Am. 2017; 99:1580-1590.
18. Pompe E, de Jong PA, de Jong WU, Takx RA, Eikendal AL, Willemink MJ, Oudkerk M, Budde RP, Lammers JW, Mohamed Hoesein FA. Inter-observer and inter-examination variability of manual vertebral bone attenuation measurements on computed tomography. Eur Radiol. 2016; 26:3046-3053.
19. Shin HK, Koo HW, Kim KH, Yoon SW, Sohn MJ, Lee BJ. The Usefulness of Trabecular CT Attenuation Measurement at L4 Level to Predict Screw Loosening After Degenerative Lumbar Fusion Surgery: Consider Number of Fused Levels and Postoperative Sagittal Balance. Spine. 2022; 47:745-753.
20. Kim JB, Park SW, Lee YS, Nam TK, Park YS, Kim YB. The Effects of Spinopelvic Parameters and Paraspinal Muscle Degeneration on S1 Screw Loosening. J Korean Neurosurg Soc. 2015; 58:357-362.
21. Sakai Y, Takenaka S, Matsuo Y, Fujiwara H, Honda H, Makino T, Kaito T. Hounsfield unit of screw trajectory as a predictor of pedicle screw loosening after single level lumbar interbody fusion. J Orthop Sci. 2018; 23:734-738.
22. Erickson BJ. Basic Artificial Intelligence Techniques: Machine Learning and Deep Learning. Radiol Clin North Am. 2021; 59:933-940.
23. Fogel AL, Kvedar JC. Artificial intelligence powers digital medicine. NPJ Digit Med. 2018; 1:5.
24. Rudisill SS, Hornung AL, Barajas JN, *et al*. Artificial intelligence in predicting early-onset adjacent segment degeneration following anterior cervical discectomy and fusion. Eur Spine J. 2022; 31:2104-2114.
25. Chang M, Canseco JA, Nicholson KJ, Patel N, Vaccaro AR. The Role of Machine Learning in Spine Surgery: The Future Is Now. Front Surg. 2020; 7:54.
26. Karhade AV, Thio Q, Ogink PT, Bono CM, Ferrone ML, Oh KS, Saylor PJ, Schoenfeld AJ, Shin JH, Harris MB, Schwab JH. Predicting 90-Day and 1-Year Mortality in Spinal Metastatic Disease: Development and Internal Validation. Neurosurgery. 2019; 85:E671-e681.
27. Joshi RS, Haddad AF, Lau D, Ames CP. Artificial Intelligence for Adult Spinal Deformity. Neurospine. 2019; 16:686-694.
28. Zou D, Li W, Deng C, Du G, Xu N. The use of CT Hounsfield unit values to identify the undiagnosed spinal osteoporosis in patients with lumbar degenerative diseases. Eur Spine J. 2019; 28:1758-1766.
29. Sandén B, Olerud C, Petrén-Mallmin M, Johansson C, Larsson S. The significance of radiolucent zones surrounding pedicle screws. Definition of screw loosening in spinal instrumentation. J Bone Joint Surg Br. 2004; 86:457-461.
30. Spirig JM, Sutter R, Götschi T, Farshad-Amacker NA, Farshad M. Value of standard radiographs, computed tomography, and magnetic resonance imaging of the lumbar spine in detection of intraoperatively confirmed pedicle screw loosening-a prospective clinical trial. Spine J. 2019; 19:461-468.
31. Tibshirani R. Regression shrinkage and selection *via* the lasso. Journal of the Royal Statistical Society, Series B. 1996; 58.
32. Okada A, Hashimoto Y, Goto T, Yamaguchi S, Ono S, Ikeda Kurakawa K, Nangaku M, Yamauchi T, Yasunaga H, Kadowaki T. A Machine Learning-Based Predictive Model to Identify Patients Who Failed to Attend a Follow-up Visit for Diabetes Care After Recommendations From a National Screening Program. Diabetes Care. 2022; 45:1346-1354.
33. Riley RD, Ensor J, Snell KIE, Harrell FE, Jr., Martin GP, Reitsma JB, Moons KGM, Collins G, van Smeden M. Calculating the sample size required for developing a clinical prediction model. Bmj. 2020; 368:m441.
34. Le Glaz A, Haralambous Y, Kim-Dufor DH, Lenca P, Billot R, Ryan TC, Marsh J, DeVylder J, Walter M, Berrouiguet S, Lemey C. Machine Learning and Natural Language Processing in Mental Health: Systematic Review. J Med Internet Res. 2021; 23:e15708.
35. Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, Nair B, Katz R, Himmelfarb J, Bansal N, Lee SI. From Local Explanations to Global Understanding with Explainable AI for Trees. Nat Mach Intell. 2020; 2:56-67.
36. Tokuhashi Y, Matsuzaki H, Oda H, Uei H. Clinical course and significance of the clear zone around the pedicle screws in the lumbar degenerative disease. Spine. 2008; 33:903-908.
37. Kim HJ, Kim SG, Lee HM, Kim HS, Moon ES, Park JO, Seol NH, Moon SH. Risk factors associated with the halo phenomenon after lumbar fusion surgery and its clinical significance. Asian Spine J. 2008; 2:22-26.
38. Zindrick MR, Wiltse LL, Widell EH, Thomas JC, Holland WR, Field BT, Spencer CW. A biomechanical study of intrapeduncular screw fixation in the lumbosacral spine. Clin Orthop Relat Res. 1986; 99-112.
39. Okuyama K, Abe E, Suzuki T, Tamura Y, Chiba M, Sato K. Influence of bone mineral density on pedicle screw fixation: a study of pedicle screw fixation augmenting posterior lumbar interbody fusion in elderly patients. Spine J. 2001; 1:402-407.
40. Luk KD, Chen L, Lu WW. A stronger bicortical sacral pedicle screw fixation through the s1 endplate: an *in vitro* cyclic loading and pull-out force evaluation. Spine. 2005; 30:525-529.
41. McLachlin SD, Al Saleh K, Gurr KR, Bailey SI, Bailey CS, Dunning CE. Comparative assessment of sacral screw loosening augmented with PMMA versus a calcium triglyceride bone cement. Spine. 2011; 36:E699-704.
42. Wu X, Shi J, Wu J, Cheng Y, Peng K, Chen J, Jiang H. Pedicle screw loosening: the value of radiological imagings and the identification of risk factors assessed by extraction torque during screw removal surgery. J Orthop Surg Res. 2019; 14:6.
43. Zou D, Muheremu A, Sun Z, Zhong W, Jiang S, Li W. Computed tomography Hounsfield unit-based prediction of pedicle screw loosening after surgery for degenerative lumbar spine disease. J Neurosurg Spine. 2020; 32:716-721.
44. Ohba T, Ebata S, Oba H, Koyama K, Haro H. Risk Factors

for Clinically Relevant Loosening of Percutaneous Pedicle Screws. Spine Surg Relat Res. 2019; 3:79-85.
45. Livshits G, Ermakov S, Popham M, Macgregor AJ, Sambrook PN, Spector TD, Williams FM. Evidence that bone mineral density plays a role in degenerative disc disease: the UK Twin Spine study. Ann Rheum Dis. 2010; 69:2102-2106.
46. Kuo CH, Chang PY, Tu TH, Fay LY, Chang HK, Wu JC, Huang WC, Cheng H. The Effect of Lumbar Lordosis on Screw Loosening in Dynesys Dynamic Stabilization: Four-Year Follow-Up with Computed Tomography. Biomed Res Int. 2015; 2015:152435.
47. Ochtman AEA, Kruyt MC, Jacobs WCH, Kersten R, le Huec JC, Öner FC, van Gaalen SM. Surgical Restoration of Sagittal Alignment of the Spine: Correlation with Improved Patient-Reported Outcomes: A Systematic Review and Meta-Analysis. JBJS Rev. 2020; 8:e1900100.
48. Banno T, Hasegawa T, Yamato Y, Kobayashi S, Togawa D, Oe S, Mihara Y, Matsuyama Y. Prevalence and Risk Factors of Iliac Screw Loosening After Adult Spinal Deformity Surgery. Spine. 2017; 42:E1024-e1030.
49. Schwab FJ, Blondel B, Bess S, *et al*. Radiographical spinopelvic parameters and disability in the setting of adult spinal deformity: a prospective multicenter analysis. Spine. 2013; 38:E803-812.

Received December 28, 2023; Revised February 14, 2024; Accepted February 21, 2024.

**Address correspondence to:*
Yuntao Wang, Department of Spine Surgery, Southeast University ZhongDa Hospital, Nanjing 210009, Jiangsu, China.
E-mail:220183560@seu.edu.cn; wangyttod@seu.edu.cn

Released online in J-STAGE as advance publication February 27, 2024.

Original Article

DOI: 10.5582/bst.2024.01006

Diabetes mellitus, glycemic traits, SGLT2 inhibition, and risk of pulmonary arterial hypertension: A Mendelian randomization study

Jiang-shan Tan[1], Yanmin Yang[1,*], Jingyang Wang[1], Yimeng Wang[1], Tingting Lv[2], Yuyuan Shu[1], Wei Xu[1], Lingtao Chong[2]

[1] Emergency and Critical Care Center, Fuwai Hospital, National Center for Cardiovascular Diseases of China, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China;
[2] Key Laboratory of Pulmonary Vascular Medicine, Fuwai Hospital, State Key Laboratory of Cardiovascular Disease, National Center for Cardiovascular Diseases, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China.

SUMMARY This study aimed to investigate the causal role of diabetes mellitus (DM), glycemic traits, and sodium-glucose cotransporter 2 (SGLT2) inhibition in pulmonary arterial hypertension (PAH). Utilizing a two-sample two-step Mendelian randomization (MR) approach, we determined the causal influence of DM and glycemic traits (including insulin resistance, glycated hemoglobin, and fasting insulin and glucose) on the risk of PAH. Moreover, we examined the causal effects of SGLT2 inhibition on the risk of PAH. Genetic proxies for SGLT2 inhibition were identified as variants in the *SLC5A2* gene that were associated with both levels of gene expression and hemoglobin A1c. Results showed that genetically inferred DM demonstrated a causal correlation with an increased risk of PAH, exhibiting an odds ratio (OR) of 1.432, with a 95% confidence interval (CI) of 1.040-1.973, and a p-value of 0.028. The multivariate MR analysis revealed comparable outcomes after potential confounders (OR = 1.469, 95%CI = 1.021-2.115, p = 0.038). Moreover, genetically predicted SGLT2 inhibition was causally linked to a reduced risk of PAH (OR = $1.681*10^{-7}$, 95%CI = $7.059*10^{-12}$-0.004, p = 0.002). Therefore, our study identified the suggestively causal effect of DM on the risk of PAH, and SGLT2 inhibition may be a potential therapeutic target in patients with PAH.

Keywords diabetes mellitus, glycemic traits, SGLT2 inhibition, pulmonary arterial hypertension, mendelian randomization

1. Introduction

Diabetes mellitus (DM) is an increasingly common and potentially devastating medical condition, presenting a major public health challenge in the contemporary era (*1*). It is well known that DM results in a range of macro- and microvascular complications (*2*), including coronary heart disease, stroke, or retinopathy (*3*). However, the potential role of diabetes in pulmonary vasculature is underappreciated. Pulmonary arterial hypertension (PAH), as a critical pulmonary vasculature, is characterized by a chronic elevation of pulmonary arterial pressure and the remodeling of pulmonary arteries, which can ultimately lead to heart failure and mortality (*4*). Epidemiologic studies have shown that about 21% to 31% of patients with PAH have diabetes (*5,6*). Besides, patients with type II DM (T2DM) are at a higher risk of PAH independent of traditional risk factors, including smoking, congestive heart failure, coronary artery disease, and hypertension (*7*). However, all these conclusions were drawn from observational studies, which may have been biased by some known and even unknown confounders. Therefore, whether there are causal effects of DM and glycemic traits (including insulin resistance, glycated hemoglobin, fasting insulin, and glucose) on PAH remains unknown.

Sodium-glucose cotransporter 2 (SGLT2) inhibitors represent a category of oral drugs designed to manage type 2 diabetes, with empagliflozin and dapagliflozin being among the most prominent members of this class (*8*). These agents primarily operate by inhibiting insulin resistance, thus modulating glucose metabolism (*8*). Moreover, they offer a range of additional benefits, including anti-inflammatory, antioxidant, and anti-fibrotic effects. Clinical trials examining cardiovascular outcomes of novel antidiabetic drugs, sanctioned by the United States Food and Drug Administration (FDA), have demonstrated that SGLT2 inhibitors

can significantly lower the risk of major adverse cardiovascular events (MACE), such as cardiovascular death, non-fatal stroke, or myocardial infarction (*9-13*). Furthermore, dapagliflozin and empagliflozin have been proven to reduce the likelihood of hospitalization for heart failure in patients with diabetes (*9-11,14*). Empagliflozin was initially approved for use in adult patients with type 2 diabetes in China in 2017. In 2022, the China National Medical Products Administration approved a new indication for empagliflozin in the treatment of adult patients with heart failure, regardless of the presence of diabetes, and reduced ejection fraction.

It is noteworthy that clinical studies examining patients with heart failure specifically excluded those with PAH (*9-11,14*), yet those with pulmonary hypertension (PH) due to left heart disease (PH-LHD) were not excluded from these trials. Consequently, investigators have undertaken additional basic and clinical investigations to explore the potential effects of SGLT2 inhibitors in treating PAH.

2. Research Design and Methods

2.1. Study design

The study design is depicted in Figure 1, which outlines the following steps: 1) identification of genetic variants that act as proxies for the impact of SGLT2 inhibition, 2) selection of DM and glycemic traits (including insulin resistance, glycated hemoglobin, fasting insulin, and glucose) as exposures, 3) designation of PAH as the outcome, 4) execution of a two-sample MR analysis to determine the causal effects of DM, glycemic trait, and SGLT2 inhibition on PAH, and 5) multivariable MR conditioning on potential confounders, including body mass index (BMI), smoking, heart failure (HF), and hypertension. As all the data in this study were acquired from publicly available published databases and were originally approved by their respective ethics committees, with all participants having signed informed consent, this study may be considered exempt from requiring further ethics committee approval.

2.2. Selection and validation for genetic predictors of SGLT2 inhibition

Referring to Min Xu *et al*., we used a similar approach to screen genetic variants associated with SGLT2 inhibitors (*15*). The selection of genetic variants acting as proxies for SGLT2 inhibition involved four steps, as shown in Figure 1A. Firstly, genetic variants associated with *SLC5A2* mRNA expression levels were selected using data from GTEx (*16*) and eQTLGen Consortium (*17*), along with information on the potential functional genes

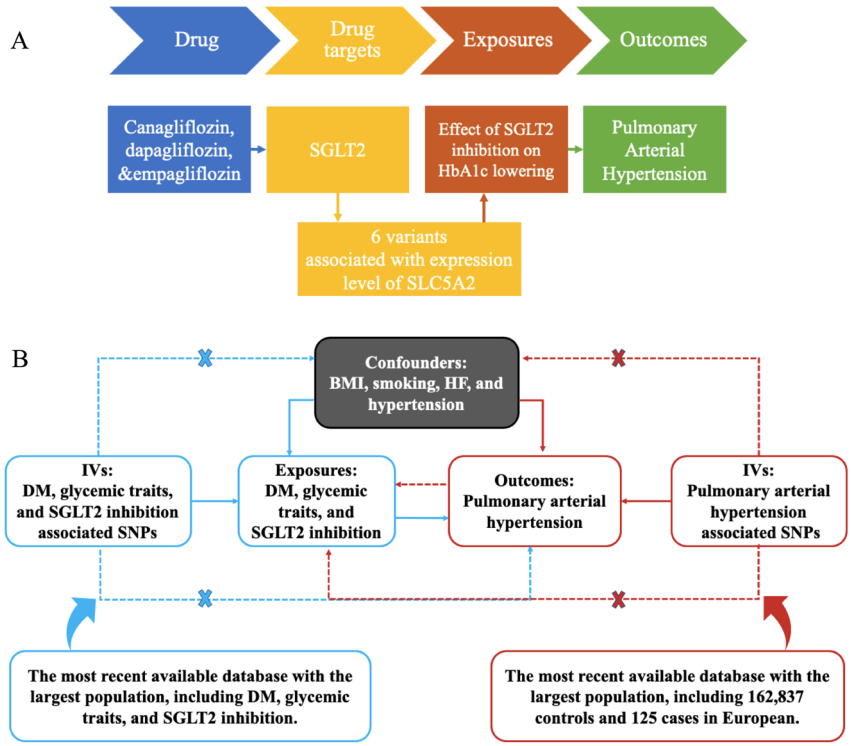

Figure 1. Study design. A: The aim of this study is to explore whether there is a causal effect of DM, glycemic traits, and SGLT2 inhibition (exposures) on PAH (outcomes). The diagram displays the selected drugs and their targets in blue, based on literature evidence. Genetic variants associated with the *SLC5A2* gene, which is the functional gene of SGLT2 inhibitors, are shown in yellow and were obtained using data from GTEx and eQTLGen. The red modules and black arrows represent the MR estimates of the effects of the variants on the HbA1c level, using data from the UK Biobank. Green modules represent selected PAH outcomes. The diagram illustrates the two-step MR models: Step 1 establishes the causal effect of DM, glycemic traits, and SGLT2 inhibition on PAH, while Step 2 establishes the causal effect of PAH on DM, glycemic traits, and SGLT2 inhibition.

of SGLT2 inhibitors. Secondly, variants showing region-wide associations with glycated hemoglobin(HbA1c) were selected using data from a subgroup of unrelated individuals with European ancestry and without diabetes in the UK Biobank (*18*). Thirdly, a genetic colocalization approach was used to verify whether *SLC5A2* and HbA1c share the same causal variant (*19*). Finally, a standard clumping process was carried out to select six genetic variants robustly associated with SGLT2 inhibition *via* HbA1c as genetic predictors for the MR analysis. The strength of the genetic predictors of each tested exposure was estimated using F statistics.

2.3. Selection of DM and glycemic traits

To achieve textual consistency and procure a sufficient quantity of single nucleotide polymorphisms (SNPs), this study has elected to utilize a GWAS significant threshold of $P < 5\times10^{6}$ as instrumental variants (IVs) associated with a given trait. First, we selected 163 variants significantly associated with T2D from a GWAS of 36,219 patients with DM and 182,573 controls (Table 1). The diagnosis of DM was determined through a range of methods, such as physician diagnosis, elevated levels of fasting glucose or HbA1c, self-reported use of DM medications, or ICD coding, either in isolation or in combination, as described in the primary GWAS articles.

Variants associated with insulin resistance were obtained from a GWAS of 37,037 individuals published in 2010 (*20*), which was the latest and largest publicly available GWAS in insulin resistance (Table 1).

Additionally, variants linked to the specific glycemic traits, including fasting insulin, fasting glucose, and HbA1c, were procured from GWAS publications authored by the Meta-Analysis of Glucose and Insulin-related Traits Consortium (MAGIC) investigators (*21*).

2.4. Selection of potential confounders

BMI, smoking, HF, and hypertension were known to be associated with both DM (*22*) and PAH (*4*) and hence, these factors could potentially confound the judgment of the causal relationship between DM and PAH. To reduce the influence of potential confounding factors on the study results, we further explored the causal association between DM and PAH by using multivariate MR analysis, after gradually adjusting for BMI, smoking, HF, and hypertension and simultaneously adjusting for the above factors.

2.5. Study outcomes

Given that MR analyses have explored the causal effects of glycemic traits, and SGLT2 inhibition on DM and recent studies have also revealed the potentially beneficial effects of DM, glycemic traits, and SGLT2 inhibition on PAH, we mainly focused our outcome on PAH. Variants associated with PAH were obtained from a GWAS of 125 cases and 162,837 controls published in 2021, including 16,380,163 SNPs in European populations.

2.6. Single nucleotide polymorphisms selection

The selection of SNPs for instrumental variables followed specific criteria. Firstly, SNPs that showed a significant correlation with our exposures with $P < 5*10^{-6}$ were chosen. Secondly, to ensure the independence of instrumental variables, the selected SNPs were clumped using the 1000 Genomes project European sample data as a reference. A clumping window of 10000 kb and a linkage disequilibrium R^2 threshold of < 0.001 were employed in this process. Lastly, the instrumental strength was evaluated using the F statistic, and weak instruments with an F statistic < 10 were excluded due to bias.

2.7. Statistical analyses

In the realm of MR studies, pleiotropy remains a central consideration, as genetic variants used as instruments may affect the outcome through pathways other than the exposure of interest. If unaccounted for, this phenomenon can introduce bias into the causal inference. Different MR methods have been developed to mitigate this, each based on distinct assumptions about the nature

Table 1. The characteristics of all enrolled traits

Exposures	Cases (N)	Controls (N)	Sample size (N)	Number of SNPs	Year
Diabetes mellitus	36,219	182,573	218,792	16,380,466	2021
Potential Confounders					
Smoking initiation	1,237	359,957	361,194	9,646,741	2018
Body mass index	/	/	51,852	/	2022
Heart failure	1,088	360,106	361,194	9,806,537	2018
Hypertension	1,237	359,957	361,194	9,646,741	2018
Glycemic traits					
Glycated hemoglobin levels	/	/	146,806	30,649,064	2021
Insulin resistance	/	/	37,037	2,435,028	2010
Fasting insulin	/	/	151,013	29,664,438	2021
Fasting glucose	/	/	200,622	31,008,728	2021

and impact of pleiotropy. The inverse variance weighted (IVW) method assumes that pleiotropic effects are balanced and average to zero, an assumption that may not hold in all scenarios.

On the other hand, methods such as MR-Egger allow for the detection and correction of unbalanced pleiotropy by introducing an intercept term, effectively adjusting the estimate if pleiotropic effects are directional. The Weighted Median approach provides a robust median-based estimate as long as at least 50% of the information comes from valid instruments. Furthermore, advanced methods like MR pleiotropy residual sum and outlier (MR-PRESSO) detect and correct for pleiotropic outliers, and the recently developed multivariable MR framework was used to adjust for multiple pleiotropic pathways simultaneously by including multiple confounders. In addition, the Maximum Likelihood Method is employed as an alternative analytical strategy to the IVW approach, particularly in scenarios where the association between the IVs and the exposure is weak, or when there is an overlap of samples within the study (*23*). We employed a multi-step methodology to ascertain the most appropriate and efficacious MR analytical approach for our study. Initially, the best-fitting model was determined by considering heterogeneity through the use of the goodness-of-fit heterogeneity statistic method (*24*). Concurrently, the F statistic was utilized to appraise the strength of the association between IVs and the outcomes of interest, ensuring the selection of robust IVs. Lastly, we leveraged PhenoScanner, an expansive repository containing detailed genotypic and phenotypic association data, to discern whether the chosen SNPs were linked to any potential risk factors, which allowed us to exclude SNPs that exhibited associations with possible confounding factors, thus refining our instrumental variable selection.

A stable causal association was only established if the sensitivity analyses yielded similar results to the IVW method. The leave-one-out analysis was used to evaluate whether a variant drove the correlation between exposure and outcome by removing a single SNP each time. In addition to utilizing PhenoScanner, our analytical approach also incorporated the mr_pleiotropy_test to further ascertain the presence of horizontal pleiotropy. The Cochrane Q statistic was used to evaluate heterogeneity. Should one or more SNPs exert a significant influence on the outcomes, it becomes imperative to exclude these SNPs and reiterate the analysis to arrive at a definitive conclusion (*25-27*). R (4.0.3) and TwoSampleMR version 0.5.6 packages were used for all statistical analyses. The Bonferroni correction was applied in our analysis to account for multiple comparisons, thereby maintaining control over the family-wise error rate by adjusting the significance threshold in accordance with the number of tests conducted. Within the framework of the current MR analysis, a p-value of less than 0.008 (0.05 divided by 6) was deemed significant, whereas a p-value greater than 0.008 but less than 0.05 was indicative of a suggestive association.

3. Results

3.1. Details of the GWAS datasets and selected SNPs for the exposures

As is shown in Figure 1, we explored the causal effect of DM, SGLT2 inhibition, and glycemic traits on the risk of PAH and estimated the verse causal effects of PAH on DM, SGLT2 inhibition, and glycemic traits. The detailed characteristics of all enrolled traits are shown in Table 1. A total of 163 SNPs were significantly associated with DM, with an F value of 21-659 (Table S1, *http://www.biosciencetrends.com/action/getSupplementalData.php?ID=184*), and 6 SNPs were significantly associated with SGLT2 inhibition (with an F value of 13-64 in Table 2), suggesting all of those SNPs were strong instrumental variants. Similarly, 140 SNPs with glycated hemoglobin (F values of 14-1391 in Table S2, *http://www.biosciencetrends.com/action/getSupplementalData.php?ID=184*), 8 SNPs with insulin resistance (F values of 20-28 in Table S3, *http://www.biosciencetrends.com/action/getSupplementalData.php?ID=184*), 88 SNPs with fasting insulin (F values of 13-173 in Table S4, *http://www.biosciencetrends.com/action/getSupplementalData.php?ID=184*), and 120 SNPs with fasting glucose (F values of 14-1650 in Table S5, *http://www.biosciencetrends.com/action/getSupplementalData.php?ID=184*) were enrolled in the present MR analysis of the causal associations between

Table 2. Genetic variants selected for the Instrument Variables of SGLT2 inhibition

rs ID	Chr.	Position	EA	OA	Beta	Se	P	EAF	Included
rs4488457	16	31659189	T	G	-0.013	0.003	2.90E-07	0.712	Yes
rs8057326	16	31524123	T	C	-0.008	0.002	2.80E-04	0.523	Yes
rs11865835	16	31509816	T	C	-0.011	0.003	1.34E-05	0.248	Yes
rs9930811	16	31400360	A	G	-0.016	0.002	8.96E-12	0.365	No*
rs34497199	16	31551332	C	T	-0.012	0.002	5.98E-07	0.475	Yes
rs35445454	16	31699326	C	T	-0.013	0.002	1.24E-07	0.344	Yes

EA, Effect allele; OA, Other alleles. EAF, Effect allele frequency. *rs9930811 was excluded because this SNP showed significant heterogeneity in the leave-one analysis.

glycemic traits and PAH. In the potential confounders, 50 SNPs with BMI (F values of 20-173 in Table S6, *http://www.biosciencetrends.com/action/getSupplementalData.php?ID=184*), 228 SNPs with smoking (F values of 20-145 in Table S7, *http://www.biosciencetrends.com/action/getSupplementalData.php?ID=184*), 19 SNPs with HF (F values of 20-31 in Table S8, *http://www.biosciencetrends.com/action/getSupplementalData.php?ID=184*), and 13 SNPs with hypertension (F values of 21-29 in Table S9, *http://www.biosciencetrends.com/action/getSupplementalData.php?ID=184*) were included in the causal estimation between potential confounders and PAH.

DM, Diabetes mellitus. BMI, Body mass index. HF, Heart failure. SGLT2, Sodium-glucose cotransporter 2. PAH, Pulmonary arterial hypertension.

3.2. Causal estimation of DM on PAH

In the context of IVW analysis, genetically predicted DM demonstrated a causal association with a 43% elevated risk of PAH, denoted by an odds ratio (OR) of 1.432, and a 95% confidence interval (CI) of 1.040-1.973, with a p-value of 0.028 ($0.008 < p < 0.05$, Figure 2). Nevertheless, statistical significance was not reached in MR Egger's sensitivity analysis (OR = 1.059, 95%CI = 0.619-1.812, $p = 0.863$, Figure 2A) nor in the Weighted Median analysis (OR = 0.945, 95% CI = 0.497-1.798, $p = 0.829$, Figure 2A). The causal estimation between DM and PAH was observed to have no significant pleiotropy (egger intercept = 0.039, $p = 0.146$) nor heterogeneity (Q = 173.804, $p = 0.232$ in the IVW MR heterogeneity test, the scatter plot of Figure 3A, the forest plot of each SNP in Figure S1 (*http://www.biosciencetrends.com/action/getSupplementalData.php?ID=182*), and the leave-one-out analysis in Figure S2, *http://www.biosciencetrends.com/action/getSupplementalData.php?ID=183*), implying that the IVW causal estimation was more reliable. Besides, the results of MR-PRESSO revealed similar results (OR = 1.432, 95%CI = 1.040-1.973, $p = 0.029$) as IVW and no outlier-corrected results due to there being no significant heterogeneity and pleiotropy. The results of the Maximum Likelihood Method (OR = 1.436, 95%CI = 1.037-1.988, $p = 0.030$) were also similar to that of IVW due to the strong association between the IVs and the exposure (all F statistics > 10, Table S1, *http://www.biosciencetrends.com/action/getSupplementalData.php?ID=184*) and no overlap of samples within two samples. Therefore, DM was causally associated with a 43% higher risk of PAH, which can also be found in the Radial IVW plot (Figure 4).

In the multivariate MR analysis, comparable results were observed after adjusting for smoking (OR = 1.440, 95%CI = 1.017-2.040, $p = 0.039$), BMI (OR = 1.412, 95%CI = 0.972-2.051, $p = 0.069$), HF (OR = 1.397, 95%CI = 1.011-1.931, $p = 0.043$), hypertension (OR = 1.400, 95%CI = 1.007-1.946, $p = 0.045$), and all potential confounders (OR = 1.469, 95%CI = 1.021-2.115, $p = 0.038$, Figure 2). However, it is noteworthy that MR results were no longer significant after adjusting for BMI, suggesting that further investigation is required to determine whether DM increases the risk of PAH through BMI.

3.3. Causal estimation of glycemic traits and SGLT2 inhibition on PAH

Furthermore, we investigated the potential causal association between glycemic traits and PAH, but our IVW analysis and all sensitivity analyses revealed no significant direct causal relationship between them (all $p > 0.05$, Table 3).

Building upon our investigation of DM and glycemic traits, we proceeded to explore the causal relationship between SGLT2 inhibitors and PAH. When we

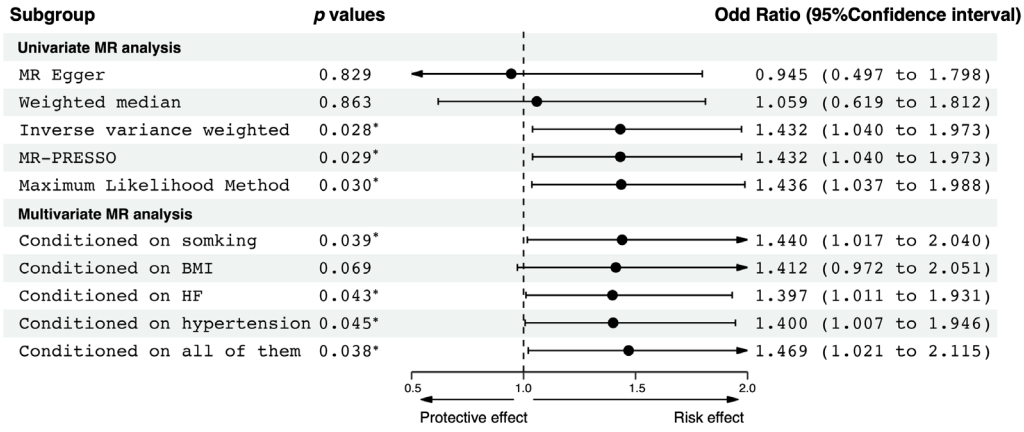

Figure 2. The forest plot for the causal role of diabetes mellitus on PAH. The causal effect of diabetes mellitus on the risk of PAH in different MR analyses and the causal effect of diabetes mellitus on the risk of PAH in the multivariate MR analysis. $^{*}p < 0.05$. The p for Bonferroni adjustment is 0.008 (0.05 divided by 6). MR, Mendelian randomization. PAH, Pulmonary arterial hypertension. MR-PRESSO, Mendelian randomization Pleiotropy RESidual Sum and Outlier.

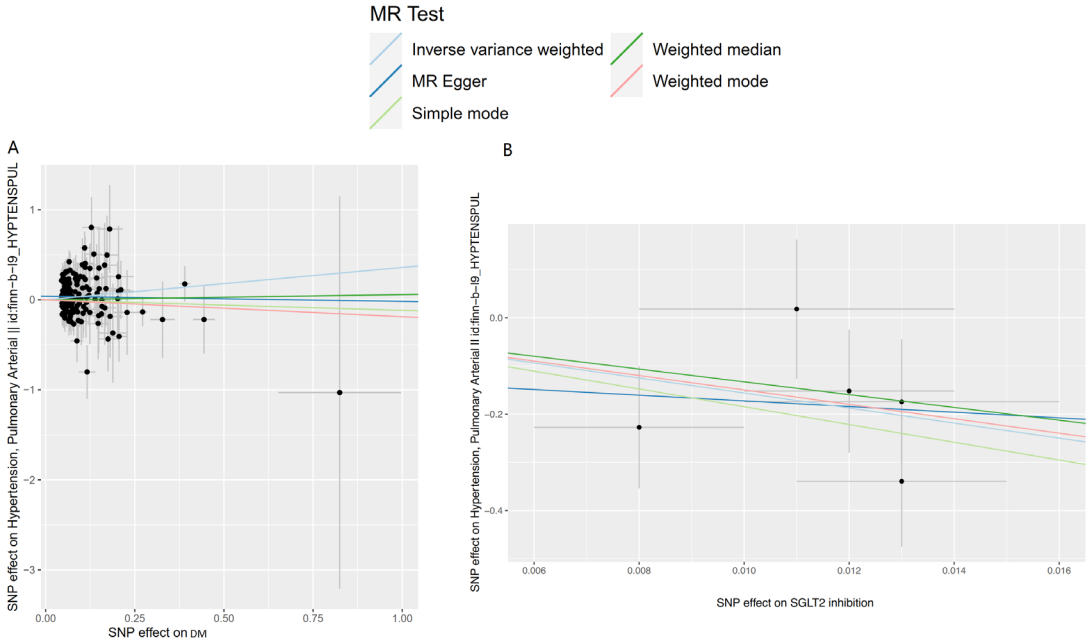

Figure 3. The scatter plot for the causal estimation of diabetes mellitus and SGLT2 inhabitation on PAH. A. the scatter plot for the causal estimation of diabetes mellitus on PAH and B. the scatter plot for the causal estimation of SGLT2 inhabitation on PAH. SGLT2, Sodium-glucose cotransporter 2. PAH, Pulmonary arterial hypertension.

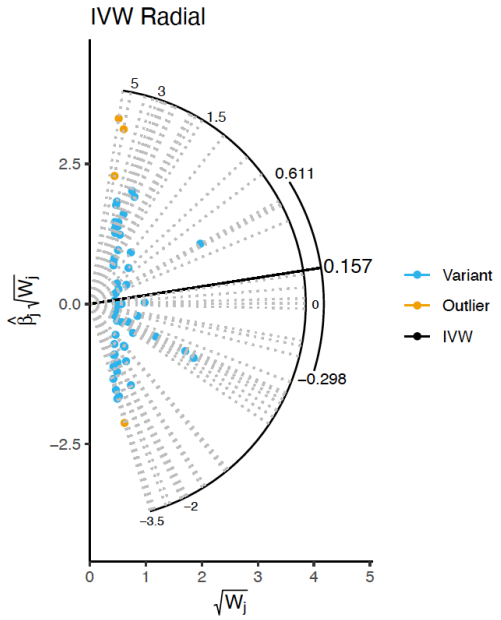

Figure 4. The plots of Radial IVW in exploring the causal associations between diabetes mellitus and pulmonary arterial hypertension.

enrolled all 6 SNPs in the MR analysis, no significant causal association was observed in the IVW analysis (Table S10, *http://www.biosciencetrends.com/action/getSupplementalData.php?ID=184*). However, Figure 5A illustrates that upon the exclusion of rs9930811, the forest plots depicting the effects on PAH remained entirely on the left side of the zero line, without crossing it. Conversely, in the leave-one-out analysis for all other SNPs, the forest plots for PAH effects did intersect with the zero line. This pattern suggests that rs9930811 may exhibit heterogeneity in comparison to the other SNPs assessed (Figure 5A).

Therefore, in the formal analysis, we excluded rs9930811, resulting in the inclusion of a total of 5 SNPs in the MR analysis. In the IVW analysis, genetically predicted SGLT2 inhibition was causally associated with a lower risk of PAH (OR = $1.681*10^{-7}$, 95%CI = $7.059*10^{-12}$-0.004, $p = 0.002 < 0.008$, Table 3). In the sensitivity analysis, both MR Egger and Weighted median revealed similar estimations (Table 3). The leave-one-out analysis (Figure 5B), scatter plot (Figure 3B), forest plot (Figure 6A), funnel plot (Figure 6B), and the MR heterogeneity test revealed no significant heterogeneity. Both MR-PRESSO (OR = $1.432*10^{-7}$, 95%CI = $6.222*10^{-12}$-0.003, $p = 0.037 < 0.050$) and the Maximum Likelihood Method (OR = $6.908*10^{-8}$, 95%CI = $9.810*10^{-13}$-0.005, $p = 0.004 < 0.008$) also yielded results consistent with those obtained *via* the IVW analysis in determining this causality. Besides, no pleiotropy was observed in the MR pleiotropy test, suggesting that the results of IVW analysis are robust and reliable.

3.4. Causal estimation of PAH on the risk of DM and glycemic traits

In further analysis, we explored the causal association between PAH and the risk of DM or the levels of glycemic traits. As is shown in Figure 7, no causal associations were observed in the estimation between PAH and the risk of DM or the levels of glycemic traits (all $p > 0.05$).

Table 3. The causal estimation of glycemic traits and SGLT2 inhibition on PAH

Exposure	method	Nsnp	OR (95%CI)	*p* values
Insulin resistance	MR Egger	8	0(0-859.940)	0.283
	Weighted median	8	0.016(0-4.655)	0.147
	Inverse variance weighted	8	0.026(0-2.529)	0.118
	MR-PRESSO	8	0.026(0.002-0.431)	0.038*
	Maximum Likelihood Method	8	0.026(0-2.743)	0.125
Glycated hemoglobin	MR Egger	141	49.577(1.199-2050.744)	0.042*
	Weighted median	141	1.028(0.034-31.423)	0.987
	Inverse variance weighted	141	0.591(0.077-4.548)	0.614
	MR-PRESSO	141	0.591(0.077-4.548)	0.614
	Maximum Likelihood Method	141	0.587(0.075-4.609)	0.613
Fasting insulin	MR Egger	88	1.215(0.001-1061.085)	0.948
	Weighted median	88	0.31(0.009-10.772)	0.515
	Inverse variance weighted	88	0.458(0.046-4.607)	0.508
	MR-PRESSO	88	0.458(0.046-4.607)	0.509
	Maximum Likelihood Method	88	0.461(0.044-4.803)	0.517
Fasting glucose	MR Egger	120	12.796(1.044-156.863)	0.049*
	Weighted median	120	1.633(0.144-18.535)	0.698
	Inverse variance weighted	120	1.323(0.317-5.525)	0.701
	MR-PRESSO	120	1.282(0.308-5.334)	0.733
	Maximum Likelihood Method	120	1.283(0.305-5.396)	0.734
SGLT2 inhibition	MR Egger	5	$0.003(0\text{-}4.108*10^{25})$	0.870
	Weighted median	5	1.681e-07 (0-1.148)	0.052
	Inverse variance weighted	5	$1.681*10^{-7}(7.059*10^{-12}\text{-}0.004)$	0.002*#
	MR-PRESSO	5	$1.432*10^{-7}(6.222*10^{-12}\text{-}0.003)$	0.037*
	Maximum Likelihood Method	5	$6.908*10^{-8}(9.810*10^{-13}\text{-}0.005)$	0.004#

Glycemic traits, including insulin resistance, glycated hemoglobin levels, fasting insulin, and fasting glucose. PAH, pulmonary arterial hypertension. OR, odd ratio. CI: Confidence interval. SGLT2: Sodium-glucose cotransporter 2. *$p < 0.05$. #$p < 0.008$ (0.05 divided by 6 in the Bonferroni adjustment). MR-PRESSO, Mendelian randomization Pleiotropy RESidual Sum and Outlier.

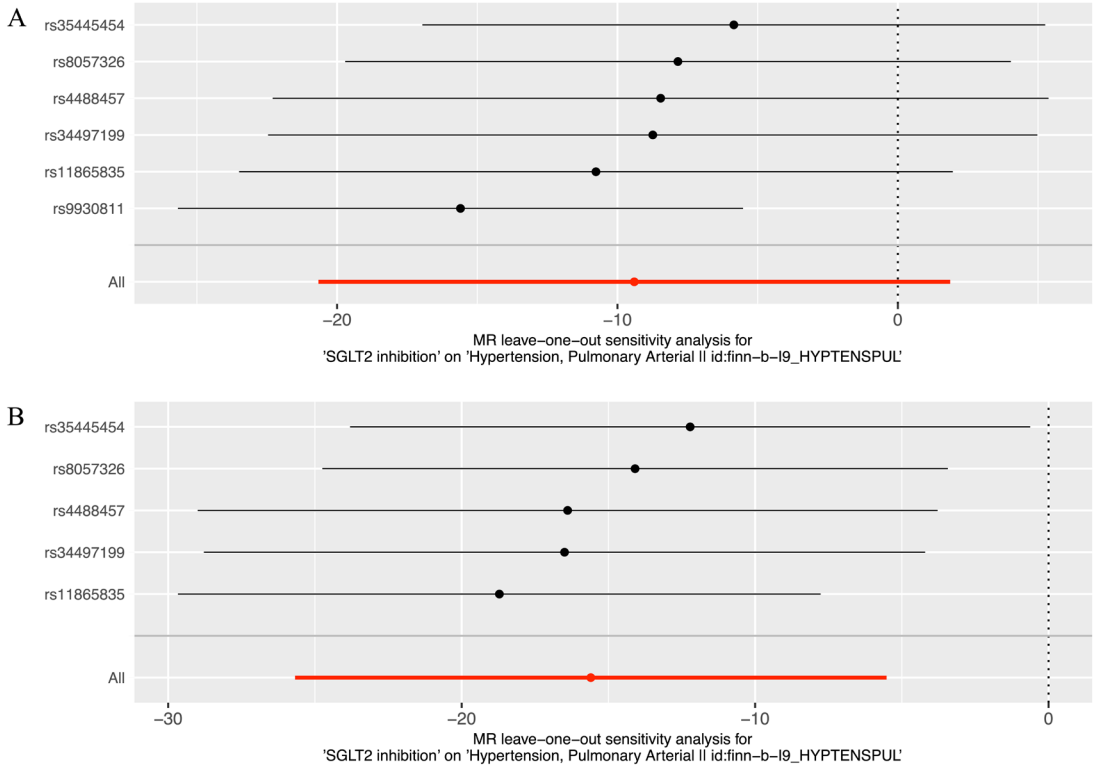

Figure 5. Mendelian randomization leave-one-out sensitivity analysis for SGLT2 inhibition on pulmonary arterial hypertension. A. The leave-one-out sensitivity analysis of all 6 variants showed significant heterogeneity when leaving the rs9930811. B. When rs9930811 was excluded, sensitivity analysis was performed again and there was no significant heterogeneity between SNPs.

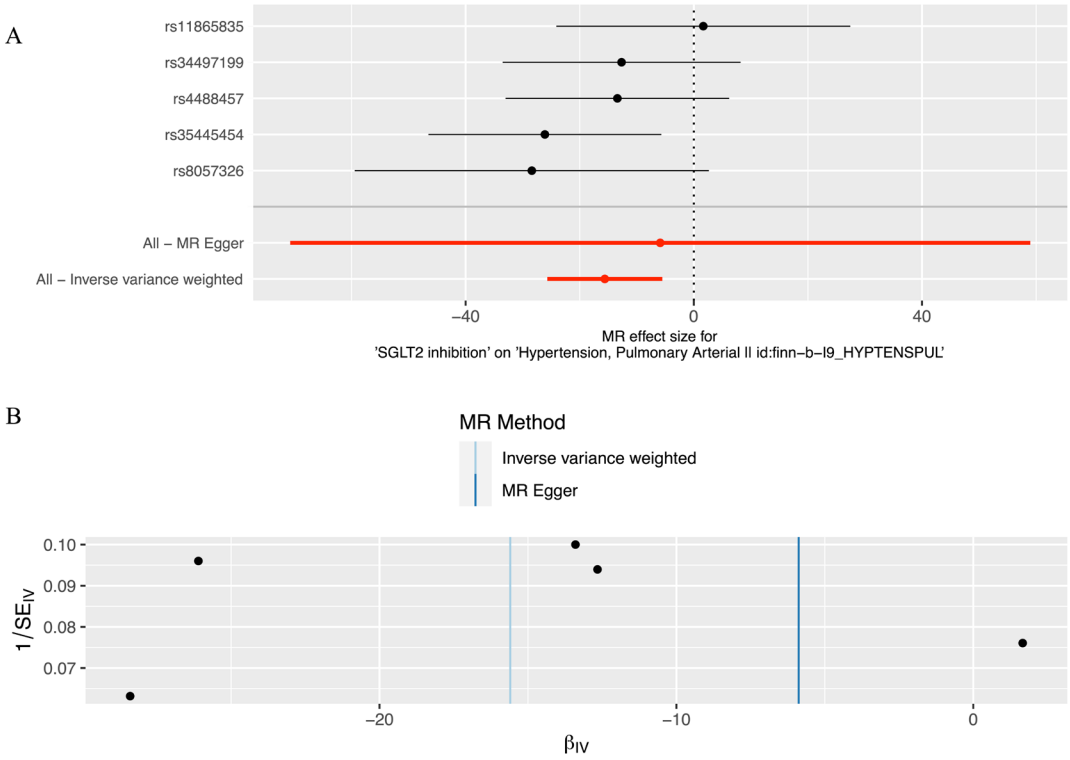

Figure 6. The sensitivity analysis for the causal association between SGLT2 inhibition and the risk of PAH. A. the forest plot for the causal estimation of SGLT2 inhibition on PAH, B. the funnel plots for the causal effect of SGLT2 inhibition on PAH. SGLT2, Sodium-glucose cotransporter 2. PAH, Pulmonary arterial hypertension.

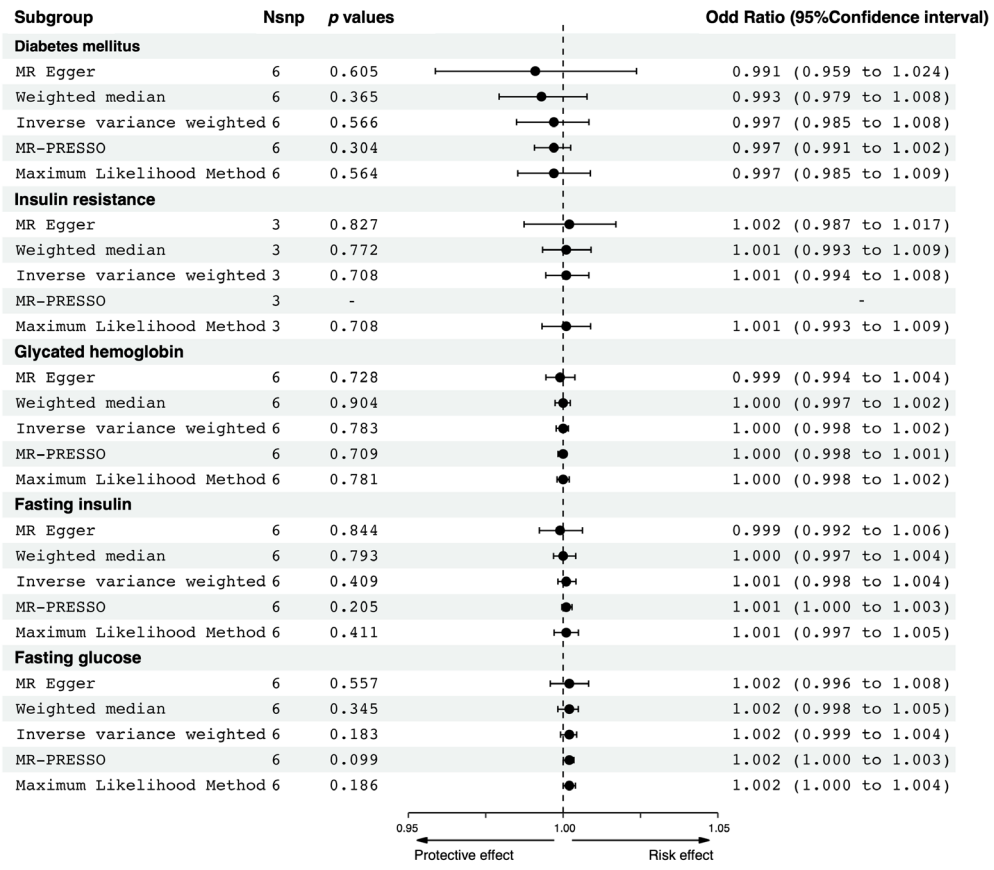

Figure 7. The MR analysis for the causal effect of PAH on the risk of DM and the levels of glycemic traits. PAH, Pulmonary arterial hypertension. MR, Mendelian randomization. The *p* for Bonferroni adjustment is 0.008 (0.05 divided by 6).

4. Discussion

In the present MR study, we found potential evidence of the causal roles of genetically predicted DM and SGLT2 inhibition in the risk of PAH. For the association between DM and PAH, the relationship attenuated when we included BMI in the analyses, suggesting that the causal effects, if real, might be partially mediated by BMI. We did not find any association between genetic determinants of IR, fasting insulin, fasting glucose, or HbA1c and PAH. These findings shed additional light on the relationships between DM, SGLT2 inhibition and the risk of PAH, supporting the potential implication of SGLT2 inhibition in patients with DM to prevent the development of PAH, even in PAH without DM to treat them.

Diabetes was a common comorbidity in patients with PAH, accounting for 21% to 31% (*5,6*). Besides, patients with chronic obstructive pulmonary disease (COPD) and diabetes exhibit more severe PH than patients with COPD alone (*28*). All of these observational studies suggested potential causal associations between DM and PAH. However, many known or unknown confounders may bias the conclusion of observational studies. Therefore, the causal association between DM and PAH remains unknown. As far as I am concerned, this is the first study to systematically investigate the causal effect of DM, glycemic traits, and SGLT2 inhibition on the risk of PAH.

This MR study revealed that patients with DM were causally associated with a 43% higher risk of PAH, which was consistent with observational studies (*5,6*). Besides its impact on morbidity, DM can exacerbate the severity of PAH and adversely affect the prognosis of patients (*6,28*) . Although managed similarly, survival in patients with DM was significantly worse than those with PAH alone and persisted as such even after statistical adjustment (*6*). Animal studies have demonstrated that diabetes contributes to PH not only by inducing left-heart dysfunction but also through direct effects of hyperglycemia on the pulmonary vasculature, which may also elucidate why patients suffering from COPD and diabetes exhibit more severe PH than patients with COPD alone (*28*). Notably, the causal association between DM and PAH was not significant after adjusting the BMI. However, the role of obesity in PAH has not yet reached a unified consensus. A cohort study of adults with PAH from the Pulmonary Hypertension Association Registry revealed that patients with overweight or obese had a worse health-related quality of life despite better transplant-free survival compared with normal-weight patients (*29,30*). Therefore, the relationship between BMI and PAH and whether the causal effect of DM on the risk of PAH was mediated by BMI needs further investigation.

Similarly, IR has also been associated with PAH in observational studies (*31*). Evidence suggests that IR is more prevalent among females with PAH compared to the general population, suggesting that insulin resistance may serve as a novel risk factor or disease modifier that could impact the survival of patients with PAH (*31*). Besides, studies have suggested that ET-1 antagonists may exert some of their effects on the pulmonary vasculature *via* IR pathways (*32,33*). Therefore, we also explored the causal effect of glycemic traits on the risk of PAH. Since glycemic traits are continuous variables, positive results were not observed as in DM, indicating that a single standard deviation change in glycemic traits (such as IR, fasting insulin, fasting glucose, and HbA1c) is not significant enough to increase the risk of PAH. However, DM is a categorical variable and has been found to significantly increase the risk of PAH.

SGLT2 inhibitors are antihyperglycemic agents that have been approved for use in patients with diabetes to lower blood glucose levels and reduce cardiovascular risk. In this study, we also found a causal effect of SGLT2 inhibitors on a lower risk of PAH, implying that SGLT2 inhibitors may contribute to reducing the risk of PAH. An *in vitro* analysis of pulmonary and coronary arteries in diabetic mice revealed that inhibition of the SGLT enzyme with either a nonspecific SGLT inhibitor or the SGLT2-specific inhibitor canagliflozin resulted in targeted dilation of pulmonary arteries (*34*). Biswajit Chowdhury *et al*. investigated the potential therapeutic impact of empagliflozin on rats with PAH by constructing a rat model of monocrotaline (MCT)-induced PAH (*35*). The results demonstrated that treatment with empagliflozin could considerably reduce mean pulmonary artery pressure and right ventricular systolic pressure in the MCT-induced PAH rats regarding hemodynamics (*35*). Pathologically, it could decrease pulmonary vascular remodeling, right ventricular hypertrophy, and fibrosis, thereby significantly decreasing the mortality rate of the PAH rats (*35*).

In addition to fundamental research, clinical investigators have delved into the impact of empagliflozin on pulmonary arterial stress. In Japan, a randomized, controlled, open-label trial was conducted to assess the effects of dapagliflozin treatment on the development of exercise-induced PH (post-exercise echocardiographic right ventricular (RV) systolic pressure (RVSP) of 50 mmHg) in 78 patients with type 2 diabetes, hypertension, and/or stable ischemia with normal EF at baseline. The trial excluded patients with advanced heart failure (New York Heart Association class 3 or 4 or any previous hospitalization for heart failure) or patients with resting RVSP >50 mmHg. An investigator-initiated, randomized, multicenter, double-blind, placebo-controlled trial named EMBRACE-HF (Empagliflozin assessed by measuring the effect on hemodynamics in patients with heart failure) (*11*) demonstrated that empagliflozin (10 mg/day) significantly reduced PA diastolic blood pressure using an implanted pulmonary artery (PA) pressure transducer (CardioMEMS), with effects beginning at Week 1 and amplifying over time. This result was consistent with PA systolic pressure and PA mean pressure. Additionally, a study by Klara Kirschbaum *et al*. further confirmed the

reliability and repeatability of the above study findings through retrospective analysis of 17 patients with heart failure or diabetes (*9*). After ten weeks of SGLT2 inhibitor treatment (empagliflozin or dapagliflozin), PA systolic blood pressure (-3.59 ± 1.55 mmHg; P = 0.034), mean pressure (-3.06 ± 1.22 mmHg; P = 0.014), and diastolic blood pressure (-2.65 ± 0.98mmHg; P = 0.008), and the change in PAP were already present after three weeks of treatment and increased over time. It is noteworthy that the use of diuretics and other heart failure medications did not change significantly during the observation period. Although these three human-based clinical studies have some methodological limitations, they suggest that SGLT2 inhibitors may be effective in addressing myocardial dysfunction and pulmonary vascular disease in patients with PAH and PH-LHD.

Both fundamental and clinical studies corroborate the potential of SGLT2 inhibitors in reducing PA pressure. However, the aforementioned clinical trials were conducted in patients with heart failure or diabetes, and there is a dearth of effective data regarding the direct application of SGLT2 inhibitors in patients with PAH. Therefore, this study was designed to evaluate the causal effect of SGLT-2 inhibitors in PAH patients using MR. The conclusion of this MR study is in line with what was drawn from the aforementioned clinical and fundamental studies, indicating that SGLT-2 inhibitors are expected to emerge as a new treatment alternative for PAH patients and offer new insights for research into the mechanisms of PAH.

MR is a widely used method in epidemiological research to investigate the causal relationship between exposure and outcome by exploiting genetic variants as instrumental variables. While MR has several advantages over traditional observational studies, such as reducing confounding and reverse causation, it also has limitations. One of the main limitations of MR analysis is the assumption that genetic variants used as IVs only affect the outcome through the exposure of interest and not through other pathways. Violation of this assumption, known as pleiotropy, can result in biased estimates of the causal effect. In this study, we have conducted several sensitivity analyses to confirm no pleiotropy exists. Additionally, in our investigation, both MR Egger's sensitivity analysis and the weighted median method indicated a trend congruent with the IVW analysis, yet they fell short of achieving statistical significance. Consequently, we employed a variety of methodologies to ascertain the most appropriate MR analysis technique, ultimately determining that IVW was the superior one. Furthermore, the results obtained from IVW were comparable to those derived from MR-PRESSO and the Maximum Likelihood Method, which bolsters the credibility of our findings. The sample size and the strength of the genetic instruments used are very important in the MR analysis. Therefore, the sample size of included trials may be a limiting factor. Future studies should include larger sample sizes and use the application of novel MR methodologies that are emerging in the field to further validate our conclusions. Lastly, it is notable that this MR was finished based on European populations and whether the conclusion can be generalized to other populations needs further study (*36*).

In summary, this is the first study to systematically explore the causal associations between DM, glycemic traits, SGLT2 inhibition, and the risk of PAH. Our results suggested that DM is suggestively associated with PAH, and SGLT2 inhibition may be a potential therapeutic target in patients with PAH.

Acknowledgements

We thank the researchers who shared the GWAS data and to all the participants in these studies.

Funding: National Clinical Medical Research Center for Cardiovascular Diseases (ID: NCRC2020015) and High-Level Hospital Clinical Research Funding (ID: 2022-GSP-GG-26).

Conflict of Interest: The authors have no conflicts of interest to disclose.

References

1. Sacks LJ, Pham CT, Fleming N, Neoh SL, Ekinci EI. Considerations for people with diabetes during the Coronavirus Disease (COVID-19) pandemic. Diabetes Res Clin Pract. 2020;166:108296.
2. Zimmet P, Alberti KG, Shaw J. Global and societal implications of the diabetes epidemic. Nature. 2001; 414:782-787.
3. Tomic D, Shaw JE, Magliano DJ. The burden and risks of emerging complications of diabetes mellitus. Nat Rev Endocrinol. 2022; 18:525-539.
4. Humbert M, Kovacs G, Hoeper MM, *et al*. 2022 ESC/ERS Guidelines for the diagnosis and treatment of pulmonary hypertension. Eur Respir J. 2023; 61:2200879.
5. Rosenkranz S, Pausch C, Coghlan JG, *et al*. Risk stratification and response to therapy in patients with pulmonary arterial hypertension and comorbidities: A COMPERA analysis. J Heart Lung Transplant. 2023; 42:102-114.
6. Abernethy AD, Stackhouse K, Hart S, Devendra G, Bashore TM, Dweik R, Krasuski RA. Impact of diabetes in patients with pulmonary hypertension. Pulm Circ. 2015; 5:117-123.
7. Movahed MR, Hashemzadeh M, Jamal MM. The prevalence of pulmonary embolism and pulmonary hypertension in patients with type II diabetes mellitus. Chest. 2005; 128:3568-3571.
8. Heerspink HJ, Perkins BA, Fitchett DH, Husain M, Cherney DZ. Sodium Glucose Cotransporter 2 Inhibitors in the Treatment of Diabetes Mellitus: Cardiovascular and Kidney Effects, Potential Mechanisms, and Clinical Applications. Circulation. 2016; 134:752-772.
9. Zinman B, Wanner C, Lachin JM, Fitchett D, Bluhmki E, Hantel S, Mattheus M, Devins T, Johansen OE, Woerle HJ, Broedl UC, Inzucchi SE. Empagliflozin, Cardiovascular

Outcomes, and Mortality in Type 2 Diabetes. N Engl J Med. 2015; 373:2117-2128.
10. Neal B, Perkovic V, Mahaffey KW, de Zeeuw D, Fulcher G, Erondu N, Shaw W, Law G, Desai M, Matthews DR. Canagliflozin and Cardiovascular and Renal Events in Type 2 Diabetes. N Engl J Med. 2017; 377:644-657.
11. Wiviott SD, Raz I, Bonaca MP, *et al*. Dapagliflozin and Cardiovascular Outcomes in Type 2 Diabetes. N Engl J Med. 2019; 380:347-357.
12. Cannon CP, Pratley R, Dagogo-Jack S, *et al*. Cardiovascular Outcomes with Ertugliflozin in Type 2 Diabetes. N Engl J Med. 2020; 383:1425-1435.
13. McMurray JJV, Solomon SD, Inzucchi SE, *et al*. Dapagliflozin in Patients with Heart Failure and Reduced Ejection Fraction. N Engl J Med. 2019; 381:1995-2008.
14. FDA Office of Communications, Center for Drug Evaluation and Research. Guidance for Industry Diabetes Mellitus-Evaluating Cardiovascular Risk in New Antidiabetic Therapies to Treat Type 2 Diabetes; Availability. 2008. 73 FR 77724, E8-30086. *https://www.federalregister.gov/documents/2008/12/19/E8-30086/guidance-for-industry-on-diabetes-mellitus-evaluating-cardiovascular-risk-in-new-antidiabetic* (accessed January 10, 2024)
15. Xu M, Zheng J, Hou T, Lin H, Wang T, Wang S, Lu J, Zhao Z, Li M, Xu Y, Ning G, Bi Y, Wang W. SGLT2 Inhibition, Choline Metabolites, and Cardiometabolic Diseases: A Mediation Mendelian Randomization Study. Diabetes Care. 2022; 45:2718-2728.
16. The GTEx Consortium atlas of genetic regulatory effects across human tissues. Science. 2020; 369:1318-1330.
17. Võsa U, Claringbould A, Westra HJ, *et al*. Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. Nature Genet. 2021; 53:1300-1310.
18. Elsworth B, Lyon M, Alexander T, Liu Y, Matthews P, Hallett J, Bates P, Palmer T, Haberland V, Smith GD, Zheng J, Haycock P, Gaunt TR, Hemani G. The MRC IEU OpenGWAS data infrastructure. 2020; *https://www.biorxiv.org/content/10.1101/2020.08.10.244293v1*
19. Walker VM, Kehoe PG, Martin RM, Davies NM. Repurposing antihypertensive drugs for the prevention of Alzheimer's disease: a Mendelian randomization study. Int J Epidemiol. 2020; 49:1132-1140.
20. Dupuis J, Langenberg C, Prokopenko I, *et al*. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. Nature Genet. 2010; 42:105-116.
21. Chen J, Spracklen CN, Marenne G, *et al*. The trans-ancestral genomic architecture of glycemic traits. Nature Genet. 2021; 53:840-860.
22. Kelsey MD, Nelson AJ, Green JB, Granger CB, Peterson ED, McGuire DK, Pagidipati NJ. Guidelines for Cardiovascular Risk Reduction in Patients With Type 2 Diabetes: JACC Guideline Comparison. J Am Coll Cardiol. 2022; 79:1849-1857.
23. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol. 2013; 37:658-665.
24. Bowden J, Del Greco MF, Minelli C, Davey Smith G, Sheehan N, Thompson J. A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization. Stat Med. 2017; 36:1783-1802.
25. Zhu Z, Zheng Z, Zhang F, Wu Y, Trzaskowski M, Maier R, Robinson MR, McGrath JJ, Visscher PM, Wray NR, Yang J. Causal associations between risk factors and common diseases inferred from GWAS summary data. Nat Commun. 2018; 9:224.
26. Fang Y, Si X, Wang J, Wang Z, Chen Y, Liu Y, Yan Y, Tian J, Zhang B, Pu J. Alzheimer Disease and Epilepsy: A Mendelian Randomization Study. Neurology. 2023; 101:e399-e409.
27. Wu F, Huang Y, Hu J, Shao Z. Mendelian randomization study of inflammatory bowel disease and bone mineral density. BMC Med. 2020; 18:312.
28. Makarevich AE, Valevich VE, Pochtavtsev AU. Evaluation of pulmonary hypertension in COPD patients with diabetes. Adv Med Sci. 2007; 52:265-272.
29. Hu S, Tan JS, Liu S, Guo TT, Song W, Peng FH, Wu Y, Gao X, Hua L. Development and Validation of a Nomogram for Predicting the Long-Term Survival in Patients With Chronic Thromboembolic Pulmonary Hypertension. Am J Cardiol. 2022; 163:109-116.
30. Min J, Feng R, Badesch D, *et al*. Obesity in Pulmonary Arterial Hypertension (PAH): The Pulmonary Hypertension Association Registry (PHAR). Ann Am Thorac Soc. 2020; 18:229-237.
31. Zamanian RT, Hansmann G, Snook S, Lilienfeld D, Rappaport KM, Reaven GM, Rabinovitch M, Doyle RL. Insulin resistance in pulmonary arterial hypertension. Eur Respir J. 2009; 33:318-324.
32. Said SA, Ammar el SM, Suddek GM. Effect of bosentan (ETA/ETB receptor antagonist) on metabolic changes during stress and diabetes. Pharmacol Res. 2005; 51:107-115.
33. Verma S, Arikawa E, McNeill JH. Long-term endothelin receptor blockade improves cardiovascular function in diabetes. Am J Hypertens. 2001; 14:679-687.
34. Han Y, Cho YE, Ayon R, Guo R, Youssef KD, Pan M, Dai A, Yuan JX, Makino A. SGLT inhibitors attenuate NO-dependent vascular relaxation in the pulmonary artery but not in the coronary artery. Am J Physiol Lung Cell Mol Physiol. 2015; 309:L1027-1036.
35. Chowdhury B, Luu AZ, Luu VZ, Kabir MG, Pan Y, Teoh H, Quan A, Sabongui S, Al-Omran M, Bhatt DL, Mazer CD, Connelly KA, Verma S, Hess DA. The SGLT2 inhibitor empagliflozin reduces mortality and prevents progression in experimental pulmonary hypertension. Biochem Biophys Res Commun. 2020; 524:50-56.
36. Tan JS, Yan XX, Wu Y, Gao X, Xu XQ, Jiang X, Jia L, Hu S, Hua L, Wang XJ. Rare variants in MTHFR predispose to occurrence and recurrence of pulmonary embolism. Int J Cardiol. 2021; 331:236-242.

Received January 11, 2024; Revised January 31, 2024; Accepted February 5, 2024.

**Address correspondence to:*
Yanmin Yang, Emergency and Intensive Care Center, State Key Laboratory of Cardiovascular Disease, Fuwai Hospital, National Center for Cardiovascular Diseases, Chinese Academy of Medical Sciences and Peking Union Medical College, 167 Beilishilu, Xicheng District, Beijing 100037, China.
E-mail: yymfuwai@163.com

Released online in J-STAGE as advance publication February 8, 2024.

Letter to the Editor

DOI: 10.5582/bst.2024.01000

New mechanisms: From lactate to lactylation to rescue heart failure

Linfeng Yi[1,2,3,4], Dan Tang[1,2,3,4], Xing Xiang[1,2,3,4], Chungang Xiao[5], Huifang Tang[2,3,4,6,*], Hong Huang[2,3,4,*]

[1] Department of Clinical Laboratory Medicine, Institution of microbiology and infectious diseases, The First Affiliated Hospital, Hengyang Medical School, University of South China, Hengyang, Hunan, China;
[2] Clinical Research Center for Myocardial Injury in Hunan Province, Hengyang, Hunan, China;
[3] Institute of Cardiovascular Disease, The First Affiliated Hospital, Hengyang Medical School, University of South China, Hengyang, Hunan, China;
[4] Hunan Provincial Key Laboratory of Multi-omics And Artificial Intelligence of Cardiovascular Diseases, University of South China, Hengyang, Hunan, China;
[5] Department of Cardiology, Hunan University of Medicine General Hospital, Huaihua, Hunan, China;
[6] Department of Cardiology, The First Affiliated Hospital, Hengyang Medical School, University of South China, Hengyang, Hunan, China.

SUMMARY Lactylation of α-myosin heavy chain (α-MHC) has recently been reported to preserve sarcomeric structure and function and attenuate the development of heart failure. Specifically, lactylation enhanced the interaction of α-MHC with the sarcomeric protein Titin, thereby maintaining normal sarcomeric structure and myocardial contractile function. Furthermore, the administration of lactate or inhibition of lactate efflux potentially treats heart failure by restoring lactylation of α-MHC and the interaction of α-MHC with Titin. This finding highlights the significant role of α-MHC lactylation in myocardial diseases and presents a new therapeutic target for the treatment of heart failure.

Keywords α-MHC, lactylation, sarcomeric structure and function, heart failure, Titin

To the Editor,

The heart acts as a pump, facilitating the transportation of blood throughout the human body through the uninterrupted series of contractions and relaxations of its muscular walls (*1*). Heart failure (HF) was charactered by cardiac hypertrophy, fibrosis, abnormal Ca^{2+}-handling. It is reported that sarcomeric proteins Titin plays vital role in regulating the contraction response to cardiac stiffness, as well as a significant therapy target especially in HF with preserved Ejection Fraction (HFpEF) (*2*). In physiological contexts, α-myosin heavy chain's (α-MHC) tail associates with Titin to formulate thick myofilament. A stable structure of thick myofilaments is crucial to maintain the normal structure and contractile function of the heart. However, the key factors that determine the binding between myosin and Titin remain unclear.

In 2019, Professor Zhao firstly demonstrated that lactylation, a type of lactate-mediated protein post-translational modification, plays a significant role in cancer metabolism and immune cells (*3*). Lactylation also has a strong correlation with vascular function, neuroregulation, hypoxia, glycolysis and cell metabolism. While lactate was once considered as a by-product of metabolism, it now has a crucial role as an energy source for the heart. Increasing evidence supports this crucial role in cardiac hypertrophy, injury and HF. Nonetheless, as crucial energy source, there is limited knowledge regarding the physiological and pathological significance of lactate in cardiomyocytes.

A study published in Cell Research by Sun *et al*. has identified the lactylation landscape of heart in mice with HF using lactylation modification-omics (*4*). Their study revealed a significant reduction in lactate concentration in cardiomyocytes during HF. This reduction led to a subsequent decrease in the level of lactylation of α-MHCK1897 and a significant drop in the interaction between α-MHC and Titin, ultimately resulting in HF. Firstly, the authors conducted lactylation modification-omics screening and identified a significant decline in the lactylation level of the α-MHCK1897 site in HF. Subsequently, the α-MHCK1897-specific site modification antibody was developed and confirmed a marked reduction in α-MHCK1897 lactylation levels in both HF patients and mice. Then α-MHCK1897R (a lysine (K) to arginine (R) substitution at position 1897) mutant mice was generated to mimic lactylation inactivation to investigate the effect of α-MHCK1897 in HF. The findings revealed that α-MHCK1897R mutant mice considerably reduced the level of α-MHC

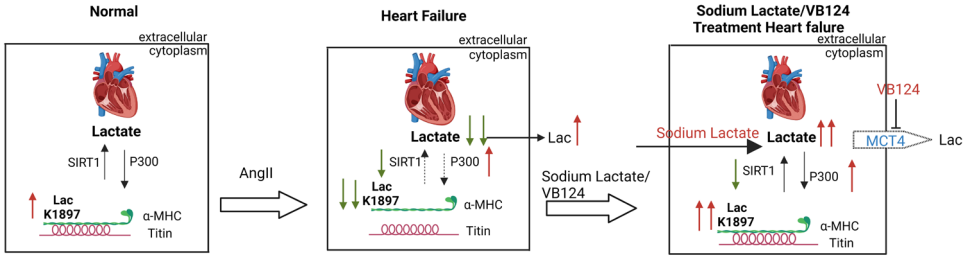

Figure 1. Schematic representation of the proposed mechanistic model of α-MHC lactylation. Under physiological conditions, lactylation of α-MHC preserves its interaction with Titin, maintaining sarcomeric structure and function. However, under pathological stress stimulation, a decrease in lactate concentration in cardiomyocytes leads to a reduction in α-MHC lactylation and α-MHC-Titin interaction, impairing cardiac structure and function. Administering sodium lactate or inhibiting MCT4 in cardiomyocytes can increase lactate concentration, promoting α-MHC lactylation and α-MHC-Titin interaction, which alleviates heart failure.

lactylation, disrupted the interaction between α-MHC and Titin, and worsened HF symptoms. Furthermore, p300 and SIRT1 were identified as acyltransferase and delactylase of α-MHC, respectively. Notably, the primary cause for the reduction of α-MHCK1897 lactylation level is the decrease in lactate concentration in cardiomyocytes, resulted from the excessive efflux and consumption of lactate in cardiomyocytes. Last but not least, sodium lactate (NALA) and VB124, a monocarboxylate transporter 4 (MCT4) inhibitor, were administered to enhance lactate concentration in cardiomyocytes, increase the level of α-MHCK1897 lactylation and significantly alleviate HF (Figure 1).

Lactylation is infrequently observed in HF, giving rise to several intriguing questions. The study conducted various experiments using a mouse model, but disparities in cardiac structure and function exist between mice and humans. The universality and reliability of the results may be limited by subjective evaluations. Therefore, further verification is required to establish the applicability of the research findings to humans. The study indicated that lactylation of α-MHC can preserve the structure and function of sarcomere, thereby decreasing the incidence of HF. Nonetheless, the specific mechanism of lactylation leading to these effects is confined. Though the influence of lactylation on cardiac structure and function were identified, further long-term observations would enable a more nuanced understanding of lactylation's role in heart disease. Although the findings are highly significant for comprehending heart disease, an evaluation of lactate as a possible therapeutic target for clinical application has not yet been conducted. Further research is imperative to assess the practicality and safety of regulating system lactate.

Additionally, there are some studies worth exploring to gain a better understanding of the lactylation mechanism in heart disease. Firstly, researching the roles of lactate and delactate enzymes and their interaction with other modification methods, such as acetylation and methylation (*5,6*). Secondly, the impact of lactylation on the progression of cardiovascular disease, in particular cardiomyocyte injury, myocardial fibrosis and HF, should be investigated. Thirdly, studying the correlation between lactylation and other modifications linked to heart disease and its regulatory mechanism would enhance understanding of the etiology of the disease. Finally, investigating the involvement of lactylation in various tissues and diseases can enhance our comprehension and potential utilisation of this process(*7,8*).

In summary, Sun and his colleagues' study offered a novel approach towards examining the mechanism and treatment of HF. The study uncovered the crucial part of lactylation in regulating myocardial sarcomere structure and functionality, which generating new insights into the pathogenesis of HF. Overall, the lactylation of α-MHCK1897 could be a potentially effective therapeutic target for treating HF, with great translational potential, particularly for patients who have HF and low lactate levels.

Funding: This work was supported by grants from the Natural Science Foundation of Hunan Province (No. 2023JJ30547, 2023JJ60051, 2021JJ70042), Major Project of Hunan Provincial Health Commission (No. W20243224, W20241008).

Conflict of Interest: The authors have no conflicts of interest to disclose.

References

1. Baman JR, Ahmad FS. Heart Failure. JAMA. 2020; 324:1015.
2. Soetkamp D, Gallet R, Parker SJ, Holewinski R, Venkatraman V, Peck K, Goldhaber JI, Marbán E, Van Eyk JE. Myofilament Phosphorylation in Stem Cell Treated Diastolic Heart Failure. Circ Res. 2021; 129:1125-1140
3. Zhang D, Tang Z, Huang H, *et al*. Metabolic regulation of gene expression by histone lactylation. Nature. 2019; 574:575-580.
4. Zhang N, Zhang Y, Xu J, *et al*. α-myosin heavy chain lactylation maintains sarcomeric structure and function and alleviates the development of heart failure. Cell Res. 2023; 33:679-698.
5. An S, Yao Y, Hu H, Wu J, Li J, Li L, Wu J, Sun M, Deng Z, Zhang Y, Gong S, Huang Q, Chen Z, Zeng Z.

PDHA1 hyperacetylation-mediated lactate overproduction promotes sepsis-induced acute kidney injury *via* Fis1 lactylation. Cell Death Dis. 2023; 14:457.
6. Moreno-Yruela C, Zhang D, Wei W, Bæk M, Liu W, Gao J, Danková D, Nielsen AL, Bolding JE, Yang L, Jameson ST, Wong J, Olsen CA, Zhao Y. Class I histone deacetylases (HDAC1-3) are histone lysine delactylases. Sci Adv. 2022; 8:eabi6696.
7. Jin J, Bai L, Wang D, Ding W, Cao Z, Yan P, Li Y, Xi L, Wang Y, Zheng X, Wei H, Ding C, Wang Y. SIRT3-dependent delactylation of cyclin E2 prevents hepatocellular carcinoma growth. EMBO Rep. 2023; 24:e56052.
8. Wang X, Fan W, Li N, Ma Y, Yao M, Wang G, He S, Li W, Tan J, Lu Q, Hou S. YY1 lactylation in microglia promotes angiogenesis through transcription activation-mediated upregulation of FGF2. Genome Biol. 2023; 24:87.

Received January 3, 2024; Revised January 29, 2024; Accepted January 31, 2024.

**Address correspondence to:*
Hong Huang and Huifang Tang, Hunan Provincial Key Laboratory of Multi-omics And Artificial Intelligence of Cardiovascular Diseases, University of South China, Hengyang, Hunan 421001, China.
E-mail: trave1@126.com (HH), tanghuifang999@163.com (TH)

Released online in J-STAGE as advance publication February 8, 2024.

BST
BioScience Trends

Guide for Authors

1. Scope of Articles

BioScience Trends (Print ISSN 1881-7815, Online ISSN 1881-7823) is an international peer-reviewed journal. *BioScience Trends* devotes to publishing the latest and most exciting advances in scientific research. Articles cover fields of life science such as biochemistry, molecular biology, clinical research, public health, medical care system, and social science in order to encourage cooperation and exchange among scientists and clinical researchers.

2. Submission Types

Original Articles should be well-documented, novel, and significant to the field as a whole. An Original Article should be arranged into the following sections: Title page, Abstract, Introduction, Materials and Methods, Results, Discussion, Acknowledgments, and References. Original articles should not exceed 5,000 words in length (excluding references) and should be limited to a maximum of 50 references. Articles may contain a maximum of 10 figures and/or tables. Supplementary Data are permitted but should be limited to information that is not essential to the general understanding of the research presented in the main text, such as unaltered blots and source data as well as other file types.

Brief Reports definitively documenting either experimental results or informative clinical observations will be considered for publication in this category. Brief Reports are not intended for publication of incomplete or preliminary findings. Brief Reports should not exceed 3,000 words in length (excluding references) and should be limited to a maximum of 4 figures and/or tables and 30 references. A Brief Report contains the same sections as an Original Article, but the Results and Discussion sections should be combined.

Reviews should present a full and up-to-date account of recent developments within an area of research. Normally, reviews should not exceed 8,000 words in length (excluding references) and should be limited to a maximum of 10 figures and/or tables and 100 references. Mini reviews are also accepted, which should not exceed 4,000 words in length (excluding references) and should be limited to a maximum of 5 figures and/or tables and 50 references.

Policy Forum articles discuss research and policy issues in areas related to life science such as public health, the medical care system, and social science and may address governmental issues at district, national, and international levels of discourse. Policy Forum articles should not exceed 3,000 words in length (excluding references) and should be limited to a maximum of 5 figures and/or tables and 30 references.

Communications are short, timely pieces that spotlight new research findings or policy issues of interest to the field of global health and medical practice that are of immediate importance. Depending on their content, Communications will be published as "Comments" or "Correspondence". Communications should not exceed 1,500 words in length (excluding references) and should be limited to a maximum of 2 figures and/or tables and 20 references.

Editorials are short, invited opinion pieces that discuss an issue of immediate importance to the fields of global health, medical practice, and basic science oriented for clinical application. Editorials should not exceed 1,000 words in length (excluding references) and should be limited to a maximum of 10 references. Editorials may contain one figure or table.

News articles should report the latest events in health sciences and medical research from around the world. News should not exceed 500 words in length.

Letters should present considered opinions in response to articles published in *BioScience Trends* in the last 6 months or issues of general interest. Letters should not exceed 800 words in length and may contain a maximum of 10 references. Letters may contain one figure or table.

3. Editorial Policies

For publishing and ethical standards, *BioScience Trends* follows the Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals issued by the International Committee of Medical Journal Editors (ICMJE, *https://icmje.org/recommendations*), and the Principles of Transparency and Best Practice in Scholarly Publishing jointly issued by the Committee on Publication Ethics (COPE, *https://publicationethics.org/resources/guidelines-new/principles-transparency-and-best-practice-scholarly-publishing*), the Directory of Open Access Journals (DOAJ, *https://doaj.org/apply/transparency*), the Open Access Scholarly Publishers Association (OASPA, *https://oaspa.org/principles-of-transparency-and-best-practice-in-scholarly-publishing-4*), and the World Association of Medical Editors (WAME, *https://wame.org/principles-of-transparency-and-best-practice-in-scholarly-publishing*).

BioScience Trends will perform an especially prompt review to encourage innovative work. All original research will be subjected to a rigorous standard of peer review and will be edited by experienced copy editors to the highest standards.

Ethical Approval of Studies and Informed Consent: For all manuscripts reporting data from studies involving human participants or animals, formal review and approval, or formal review and waiver, by an appropriate institutional review board or ethics committee is required and should be described in the Methods section. When your manuscript contains any case details, personal information and/or images of patients or other individuals, authors must obtain appropriate written consent, permission and release in order to comply with all applicable laws and regulations concerning privacy and/or security of personal information. The consent form needs to comply with the relevant legal requirements of your particular jurisdiction, and please do not send signed consent form to *BioScience Trends* to respect your patient's and any other individual's privacy. Please instead describe the information clearly in the Methods (patient consent) section of your manuscript while retaining copies of the signed forms in the event they should be needed. Authors should also state that the study conformed to the provisions of the Declaration of Helsinki (as revised in 2013, *https://wma.net/what-we-do/medical-ethics/declaration-of-helsinki*). When reporting experiments on animals, authors should indicate whether the institutional and national guide for the care and use of laboratory animals was followed.

Reporting Clinical Trials: The ICMJE (*https:// icmje.org/recommendations/browse/publishing-and-editorial-issues/clinical-trial-registration.html*) defines a clinical trial as any research project that prospectively assigns people or a group of people to an intervention, with or without concurrent comparison or control groups, to study the relationship between a health-related intervention and a health outcome. Registration of clinical trials in a public trial registry at or before the time of first patient enrollment is a condition of consideration for publication in *BioScience Trends*, and the trial registration number will be published at the end of the Abstract. The registry must be independent of for-profit interest and publicly accessible. Reports of trials must conform to CONSORT 2010 guidelines (*https://consort-statement.org/consort-2010*). Articles reporting the results of randomized trials must include the CONSORT flow diagram showing the progress of patients throughout the trial.

Conflict of Interest: All authors are required to disclose any actual or potential conflict of interest including financial interests or relationships with other people or organizations that might raise questions of bias

in the work reported. If no conflict of interest exists for each author, please state "There is no conflict of interest to disclose".

Submission Declaration: When a manuscript is considered for submission to *BioScience Trends*, the authors should confirm that 1) no part of this manuscript is currently under consideration for publication elsewhere; 2) this manuscript does not contain the same information in whole or in part as manuscripts that have been published, accepted, or are under review elsewhere, except in the form of an abstract, a letter to the editor, or part of a published lecture or academic thesis; 3) authorization for publication has been obtained from the authors' employer or institution; and 4) all contributing authors have agreed to submit this manuscript.

Initial Editorial Check: Immediately after submission, the journal's managing editor will perform an initial check of the manuscript. A suitable academic editor will be notified of the submission and invited to check the manuscript and recommend reviewers. Academic editors will check for plagiarism and duplicate publication at this stage. The journal has a formal recusal process in place to help manage potential conflicts of interest of editors. In the event that an editor has a conflict of interest with a submitted manuscript or with the authors, the manuscript, review, and editorial decisions are managed by another designated editor without a conflict of interest related to the manuscript.

Peer Review: *BioScience Trends* operates a single-anonymized review process, which means that reviewers know the names of the authors, but the authors do not know who reviewed their manuscript. All articles are evaluated objectively based on academic content. External peer review of research articles is performed by at least two reviewers, and sometimes the opinions of more reviewers are sought. Peer reviewers are selected based on their expertise and ability to provide quality, constructive, and fair reviews. For research manuscripts, the editors may, in addition, seek the opinion of a statistical reviewer. Every reviewer is expected to evaluate the manuscript in a timely, transparent, and ethical manner, following the COPE guidelines (*https://publicationethics.org/files/cope-ethical-guidelines-peer-reviewers-v2_0.pdf*). We ask authors for sufficient revisions (with a second round of peer review, when necessary) before a final decision is made. Consideration for publication is based on the article's originality, novelty, and scientific soundness, and the appropriateness of its analysis.

Suggested Reviewers: A list of up to 3 reviewers who are qualified to assess the scientific merit of the study is welcomed. Reviewer information including names, affiliations, addresses, and e-mail should be provided at the same time the manuscript is submitted online. Please do not suggest reviewers with known conflicts of interest, including participants or anyone with a stake in the proposed research; anyone from the same institution; former students, advisors, or research collaborators (within the last three years); or close personal contacts. Please note that the Editor-in-Chief may accept one or more of the proposed reviewers or may request a review by other qualified persons.

Language Editing: Manuscripts prepared by authors whose native language is not English should have their work proofread by a native English speaker before submission. If not, this might delay the publication of your manuscript in *BioScience Trends*.

The Editing Support Organization can provide English proofreading, Japanese-English translation, and Chinese-English translation services to authors who want to publish in *BioScience Trends* and need assistance before submitting a manuscript. Authors can visit this organization directly at *https://www.iacmhr.com/iac-eso/support.php?lang=en*. IAC-ESO was established to facilitate manuscript preparation by researchers whose native language is not English and to help edit works intended for international academic journals.

Copyright and Reuse: Before a manuscript is accepted for publication in *BioScience Trends*, authors will be asked to sign a transfer of copyright agreement, which recognizes the common interest that both the journal and author(s) have in the protection of copyright. We accept that some authors (*e.g.*, government employees in some countries) are unable to transfer copyright. A JOURNAL PUBLISHING AGREEMENT (JPA) form will be e-mailed to the authors by the Editorial Office and must be returned by the authors by mail, fax, or as a scan. Only forms with a hand-written signature from the corresponding author are accepted. This copyright will ensure the widest possible dissemination of information. Please note that the manuscript will not proceed to the next step in publication until the JPA Form is received. In addition, if excerpts from other copyrighted works are included, the author(s) must obtain written permission from the copyright owners and credit the source(s) in the article.

4. Cover Letter

The manuscript must be accompanied by a cover letter prepared by the corresponding author on behalf of all authors. The letter should indicate the basic findings of the work and their significance. The letter should also include a statement affirming that all authors concur with the submission and that the material submitted for publication has not been published previously or is not under consideration for publication elsewhere. The cover letter should be submitted in PDF format. For an example of Cover Letter, please visit: *https://www.biosciencetrends.com/downcentre* (Download Centre).

5. Submission Checklist

The Submission Checklist should be submitted when submitting a manuscript through the Online Submission System. Please visit Download Centre (*https://www.biosciencetrends.com/downcentre*) and download the Submission Checklist file. We recommend that authors use this checklist when preparing your manuscript to check that all the necessary information is included in your article (if applicable), especially with regard to Ethics Statements.

6. Manuscript Preparation

Manuscripts are suggested to be prepared in accordance with the "Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals", as presented at *https://www.ICMJE.org*.

Manuscripts should be written in clear, grammatically correct English and submitted as a Microsoft Word file in a single-column format. Manuscripts must be paginated and typed in 12-point Times New Roman font with 24-point line spacing. Please do not embed figures in the text. Abbreviations should be used as little as possible and should be explained at first mention unless the term is a well-known abbreviation (*e.g.* DNA). Single words should not be abbreviated.

Title page: The title page must include 1) the title of the paper (Please note the title should be short, informative, and contain the major key words); 2) full name(s) and affiliation(s) of the author(s), 3) abbreviated names of the author(s), 4) full name, mailing address, telephone/fax numbers, and e-mail address of the corresponding author; 5) author contribution statements to specify the individual contributions of all authors to this manuscript, and 6) conflicts of interest (if you have an actual or potential conflict of interest to disclose, it must be included as a footnote on the title page of the manuscript; if no conflict of interest exists for each author, please state "There is no conflict of interest to disclose").

Abstract: The abstract should briefly state the purpose of the study, methods, main findings, and conclusions. For articles that are Original Articles, Brief Reports, Reviews, or Policy Forum articles, a one-paragraph abstract consisting of no more than 250 words must be included in the manuscript. For Communications, Editorials, News, or Letters, a brief summary of main content in 150 words or fewer should be included in the manuscript. For articles reporting clinical trials, the trial registration number should be stated at the end of the Abstract. Abbreviations must be kept to a minimum and non-standard

abbreviations explained in brackets at first mention. References should be avoided in the abstract. Three to six key words or phrases that do not occur in the title should be included in the Abstract page.

Introduction: The introduction should provide sufficient background information to make the article intelligible to readers in other disciplines and sufficient context clarifying the significance of the experimental findings

Materials/Patients and Methods: The description should be brief but with sufficient detail to enable others to reproduce the experiments. Procedures that have been published previously should not be described in detail but appropriate references should simply be cited. Only new and significant modifications of previously published procedures require complete description. Names of products and manufacturers with their locations (city and state/country) should be given and sources of animals and cell lines should always be indicated. All clinical investigations must have been conducted in accordance Materials/ Patients and Methods.

Results: The description of the experimental results should be succinct but in sufficient detail to allow the experiments to be analyzed and interpreted by an independent reader. If necessary, subheadings may be used for an orderly presentation. All Figures and Tables should be referred to in the text in order, including those in the Supplementary Data.

Discussion: The data should be interpreted concisely without repeating material already presented in the Results section. Speculation is permissible, but it must be well-founded, and discussion of the wider implications of the findings is encouraged. Conclusions derived from the study should be included in this section.

Acknowledgments: All funding sources (including grant identification) should be credited in the Acknowledgments section. Authors should also describe the role of the study sponsor(s), if any, in study design; in the collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the paper for publication. If the funding source had no such involvement, the authors should so state.

In addition, people who contributed to the work but who do not meet the criteria for authors should be listed along with their contributions.

References: References should be numbered in the order in which they appear in the text. Citing of unpublished results, personal communications, conference abstracts, and theses in the reference list is not recommended but these sources may be mentioned in the text. In the reference list, cite the names of all authors when there are fifteen or fewer authors; if there are sixteen or more authors, list the first three followed by *et al*. Names of journals should be abbreviated in the style used in PubMed. Authors are responsible for the accuracy of the references. The EndNote Style of *BioScience Trends* could be downloaded at **EndNote** (*https://ircabssagroup.com/examples/ BioScience_Trends.ens*).

Examples are given below:

Example 1 (Sample journal reference):

Inagaki Y, Tang W, Zhang L, Du GH, Xu WF, Kokudo N. Novel aminopeptidase N (APN/CD13) inhibitor 24F can suppress invasion of hepatocellular carcinoma cells as well as angiogenesis. Biosci Trends. 2010; 4:56-60.

Example 2 (Sample journal reference with more than 15 authors):

Darby S, Hill D, Auvinen A, *et al*. Radon in homes and risk of lung cancer: Collaborative analysis of individual data from 13 European case-control studies. BMJ. 2005; 330:223.

Example 3 (Sample book reference):

Shalev AY. Post-traumatic stress disorder: Diagnosis, history and life course. In: Post-traumatic Stress Disorder, Diagnosis, Management and Treatment (Nutt DJ, Davidson JR, Zohar J, eds.). Martin Dunitz, London, UK, 2000; pp. 1-15.

Example 4 (Sample web page reference):

World Health Organization. The World Health Report 2008 – primary health care: Now more than ever. *http://www.who.int/whr/2008/whr08_ en.pdf* (accessed September 23, 2022).

Tables: All tables should be prepared in Microsoft Word or Excel and should be arranged at the end of the manuscript after the References section. Please note that tables should not in image format. All tables should have a concise title and should be numbered consecutively with Arabic numerals. If necessary, additional information should be given below the table.

Figure Legend: The figure legend should be typed on a separate page of the main manuscript and should include a short title and explanation. The legend should be concise but comprehensive and should be understood without referring to the text. Symbols used in figures must be explained. Any individually labeled figure parts or panels (A, B, *etc*.) should be specifically described by part name within the legend.

Figure Preparation: All figures should be clear and cited in numerical order in the text. Figures must fit a one- or two-column format on the journal page: 8.3 cm (3.3 in.) wide for a single column, 17.3 cm (6.8 in.) wide for a double column; maximum height: 24.0 cm (9.5 in.). Please make sure that the symbols and numbers appeared in the figures should be clear. Please make sure that artwork files are in an acceptable format (TIFF or JPEG) at minimum resolution (600 dpi for illustrations, graphs, and annotated artwork, and 300 dpi for micrographs and photographs). Please provide all figures as separate files. Please note that low-resolution images are one of the leading causes of article resubmission and schedule delays.

Units and Symbols: Units and symbols conforming to the International System of Units (SI) should be used for physicochemical quantities. Solidus notation (*e.g.* mg/kg, mg/mL, mol/mm^2/min) should be used. Please refer to the SI Guide *www.bipm.org/en/si/* for standard units.

Supplemental data: Supplemental data might be useful for supporting and enhancing your scientific research and *BioScience Trends* accepts the submission of these materials which will be only published online alongside the electronic version of your article. Supplemental files (figures, tables, and other text materials) should be prepared according to the above guidelines, numbered in Arabic numerals (*e.g.*, Figure S1, Figure S2, and Table S1, Table S2) and referred to in the text. All figures and tables should have titles and legends. All figure legends, tables and supplemental text materials should be placed at the end of the paper. Please note all of these supplemental data should be provided at the time of initial submission and note that the editors reserve the right to limit the size and length of Supplemental Data.

5. Submission Checklist

The Submission Checklist will be useful during the final checking of a manuscript prior to sending it to *BioScience Trends* for review. Please visit Download Centre and download the Submission Checklist file.

6. Online Submission

Manuscripts should be submitted to *BioScience Trends* online at *https://www.biosciencetrends.com/login*. Receipt of your manuscripts submitted online will be acknowledged by an e-mail from Editorial Office containing a reference number, which should be used in all future communications. If for any reason you are unable to submit a file online, please contact the Editorial Office by e-mail at *office@ biosciencetrends.com*

8. Accepted Manuscripts

Page Charge: Page charges will be levied on all manuscripts accepted for publication in *BioScience Trends* (Original Articles / Brief Reports / Reviews / Policy Forum / Communications: $140 per page for black white pages, $340 per page for color pages; News / Letters: a total cost of $600). Under exceptional circumstances, the author(s) may apply to the editorial office for a waiver of the publication charges by stating the reason in the Cover Letter when the manuscript online.

Misconduct: *BioScience Trends* takes seriously all allegations of potential misconduct and adhere to the ICMJE Guideline (*https://icmje.org/recommendations*) and COPE Guideline (*https://publicationethics.org/files/Code_of_conduct_for_journal_editors.pdf*). In cases of suspected research or publication misconduct, it may be necessary for the Editor or Publisher to contact and share submission details with third parties including authors' institutions and ethics committees. The corrections, retractions, or editorial expressions of concern will be performed in line with above guidelines.

(As of December 2022)

BioScience Trends
Editorial and Head Office
Pearl City Koishikawa 603,
2-4-5 Kasuga, Bunkyo-ku,
Tokyo 112-0003, Japan.
E-mail: office@biosciencetrends.com