Archives of Public Health

RESEARCH

Open Access

Prevalence, mutation distribution, and economic burden of thalassemia in China: a systematic review and regional analysis

Aiqi Cai[1,2†], Xiaohu Liu[2†], Qingwen Ma[3,5], Guangyu He[2], Chanchan Jing[2], Jing He[1,2], Fanyi Zeng[3,4,5*] and Baosheng Zhu[1,2,5*]

Abstract

Background Thalassemia poses a significant public health and economic challenge in China. Comprehensive data on its epidemiology, mutation spectrum, and economic impact are critical for improving prevention and treatment.

Methods This systematic review, adhering to PRISMA guidelines, analyzed studiesP{lea published between 1987 and 2024 from PubMed, FMRS, and CNKI. Data from 29 eligible studies, covering 679,697 individuals across 17 regions, were standardized using next-generation sequencing benchmarks.

Results The national carrier rate was 8.95%, with $-^{SEA}$ and $\beta^{CD41-42}$ mutations most prevalent, particularly in Hainan and Guangxi. Thalassemia contributes an annual economic burden of 9.19 billion RMB and approximately 3,590 stillbirths in the absence of prevention programs.

Conclusions While regional screening has reduced incidence since 1987, the disease continues to strain China's healthcare system. These findings highlight the pressing need to expand national prevention programs to mitigate its health and economic impacts. Insights from this review are vital for shaping public health strategies in high-prevalence areas.

Keywords Thalassemia, Incidence, China, Decreasing, Prevention

†Aiqi Cai and Xiaohu Liu contributed equally to this work.

*Correspondence:
Fanyi Zeng
fzeng@vip.163.com
Baosheng Zhu
bszhu@aliyun.com
[1] Faculty of Life Science and Technology, Medical School, Kunming University of Science and Technology, Kunming 650500, People's Republic of China
[2] Department of Medical Genetics, Yunnan Provincial Key Laboratory for Birth Defects and Genetic Diseases, NHC Key Laboratory of Healthy Birth and Birth Defect Prevention in Western China, Yunnan Provincial Clinical Medical Center for Blood Diseases and Thrombosis Prevention and Treatment, the First People'S Hospital of Yunnan Province & Affiliated Hospital of Kunming University of Science and Technology, Kunming 650032, China
[3] Shanghai Children'S Hospital, Shanghai Institute of Medical Genetics, Shanghai Jiao Tong University School of Medicine, 24/1400 West Beijing Road, Shanghai 200040, China
[4] Department of Histo-Embryology, Genetics and Developmental Biology, Shanghai Jiao Tong University School of Medicine, Shanghai, China
[5] NHC Key Laboratory of Medical Embryogenesis and Developmental Molecular Biology, Shanghai Key Laboratory of Embryo and Reproduction Engineering, 24/1400 West Beijing Road, Shanghai 200040, China

Text box 1. Contributions to the literature

This study introduces an innovative approach to thalassemia epidemiological analysis by using standardized with data generated from second- and third-generation sequencing (NGS/TGS) screen, improving the accuracy of carrier rate estimates.We provide the national and regional update on thalassemia carrier rates in China, standardizing conventional screening results to offer more reliable disease burden estimates. Besides, this research highlights the mutation spectrum of thalassemia in China and provides the first comprehensive economic impact analysis, projecting healthcare costs without prevention programs. By comparing results with the 1987 national survey, this study shows progress in prevention while addressing the ongoing challenges in disease control.

Background

Thalassemia, a hemoglobin disorder resulting from mutations in globin genes, disrupts globin synthesis in red blood cells, leading to hemolytic anemia, developmental impairments, and early mortality. Individuals carrying a mutation on one chromosome are classified as thalassemia carriers, while those with mutations on both homologous chromosomes develop thalassemia disease, which can cause anemia, developmental impairments, and increased mortality. The prevalence of thalassemia is often measured using carrier rates, which represent the proportion of individuals in a population carrying a globin gene mutation. Approximately 300,000–500,000 children with thalassemia and other hemoglobinopathies are born each year worldwide, with high prevalence in tropical and subtropical areas, including the Mediterranean, Africa, the Middle East, the Indian Subcontinent, Southeast Asia, and southern China [1, 2]. Most globin mutation carriers are phenotypically healthy and are identified only through hematological testing or gene analysis. Routine blood exams revealing a mean corpuscular volume (MCV) < 80 fL and mean corpuscular hemoglobin (MCH) < 27 pg, followed by hemoglobin electrophoresis, are key indicators.

A national survey on hemoglobin disorders conducted in 1987, involving over 900,000 subjects across nearly all of China, identified α-thalassemia and β-thalassemia as the most prevalent types and highlighted the public health burden in southern China [3]. According to The *Blue Book of Thalassemia in China (2020)*, beyond its clinical impact, thalassemia places a heavy economic and social burden on patients, families, and healthcare systems [4]. The lifelong treatment for transfusion-dependent thalassemia (TDT), including regular blood transfusions, iron chelation therapy, and, in some cases, hematopoietic stem cell transplantation, is costly. Many patients face financial strain, reduced work capacity, and lower quality of life, while caregivers often experience job loss and psychological stress. Nationally, thalassemia significantly increases healthcare costs and productivity losses, underscoring the need for effective prevention strategies.

The latest nationwide epidemiological survey or systematic review on thalassemia prevalence in China was conducted in 2017, offering valuable insights at the time [5]. However, with advances in genetic screening technologies and demographic shifts over the past few years, an updated assessment is needed. Previous studies primarily relied on conventional screening methods, leading to inconsistencies due to variations in sensitivity and specificity. To address these gaps, this study provides the first nationwide update in over five years, systematically analyzing regional thalassemia prevalence and globin mutation spectrums across China. By incorporating data from next-generation sequencing (NGS) and third-generation sequencing (TGS) platforms or normalized conventional screenings, this study ensures a standardized and comparable carrier rate estimation, enhancing data reliability. The findings offer a comprehensive and up-to-date epidemiological assessment of thalassemia in China and may provide valuable insights for other high-prevalence regions worldwide.

Methods

Definition and interpretation of terms in this article

Carrier rate of thalassemia mutations: This is the proportion of mutation carriers within a surveyed population. Given that thalassemia patients carry mutations on both homologous chromosomes, the carrier rate is calculated as:

$$Carrier\ Rate = \frac{(number\ of\ carriers + 2 \times number\ of\ patients)}{(total\ surveyed + number\ of\ patients)}$$

To mitigate potential bias from regional studies in high-prevalence areas, we used a population-weighted national carrier rate rather than weighting based on the number of individuals screened in each region.

Carrier rates of common thalassemia mutations: This includes rates for the six most prevalent α- and β-thalassemia mutations in China. For α-thalassemia, these are $-^{SEA}$, $-\alpha^{3.7}$, $-\alpha^{4.2}$, $\alpha\alpha^{CS}$, $\alpha\alpha^{WS}$, and $\alpha\alpha^{QS}$; for β-thalassemia, they include $\beta^{CD41-42}$, $\beta^{IVS-II-654}$, β^{CD17}, $\beta^{CD26\ (\beta E)}$, β^{-28}, $\beta^{CD71-72}$. To reduce bias, mutation proportions were first calculated and then multiplied by the overall carrier rate of each thalassemia type. When only mutation proportions were available without carrier rates, only proportions were reported.

Normalizations of carrier rate generated by conventional screening methods

Since two comparative studies revealed that conventional screening methods (hematological analysis and

traditional genotypic analysis, Figure S1) failed to detect certain thalassemia mutations [6, 7]. Carrier rates of α- or β-thalassemia obtained from conventional screening methods were adjusted by multiplying coefficients of 1.092 and 1.082, respectively, to align them with the rates derived from NGS-based screening. Due to the lack of studies comparing TGS detection rates with conventional methods in China, we assumed TGS detection rates to be comparable to those from NGS.

Searches

We searched PubMed, FMRS, and CNKI (Chinese) databases for original epidemiological studies on thalassemia in mainland China, Hong Kong, Macao, and Taiwan, published between January 1, 1987, and July 9, 2024. Searches were conducted in English and Chinese, without language restrictions, to ensure comprehensive coverage. Additionally, references from relevant articles were manually screened for potential studies not indexed in these databases. No grey literature sources (e.g., government reports, conference proceedings) were included, as they do not typically provide structured epidemiological data. The search strategy incorporated free-text keywords. The Boolean search strings included: ((thalassemia) OR (thalassaemia)) AND ((prevalence) OR (morbidity) OR (carrier rate) OR (carrier screening) OR (epidemiology)).

Study inclusion and exclusion criteria

Study selection was conducted in accordance with the PRISMA guidelines [8]. A PRISMA flow diagram summarizing the selection process is provided in Fig. 1. The full PRISMA Checklist is available in supplementary. This systematic review was prospectively registered on the Open Science Framework (OSF) under the registration

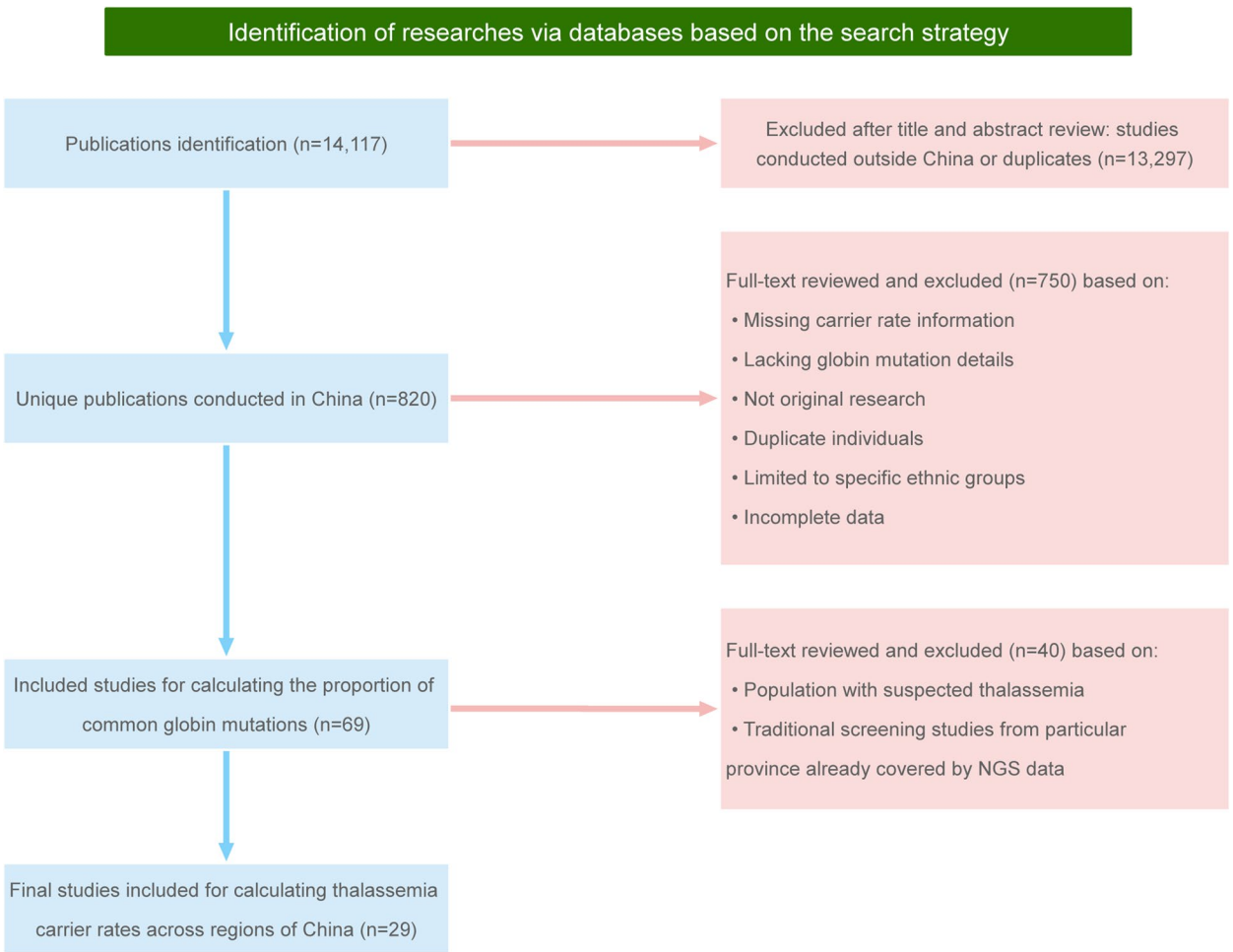

Fig. 1 PRISMA flow diagram for thalassemia data source search. Figure shows our search strategy which is described in the main text. Globin mutations refer specifically to thalassemia-related mutation types. Since most excluded studies met multiple exclusion criteria, we did not report the exact number for each criterion separately; instead, we provided the total number of exclusions at each step

DOI (https://doi.org/https://doi.org/10.17605/OSF.IO/Q329A). The full protocol, including details on study selection, data extraction, and synthesis methods, can be accessed at this link.

Inclusion criteria

Studies were included if they met all of the following criteria: (i) Cross-sectional studies conducted in China (including Hong Kong, Macau, and Taiwan); (ii) Studies that reported thalassemia prevalence or provided sufficient data (e.g., sample size and number of thalassemia patients) to allow prevalence calculation; (iii) Studies based on epidemiological surveys in general populations; (iv) Studies that detailed globin gene mutations relevant to prevalence estimates; (v) Studies published in Chinese and/or English.

Exclusion criteria

Studies were excluded if they met any of the following criteria: (i) Studies that lacked necessary details for calculating thalassemia carrier rates, mutation proportions, or individual carrier rates of common mutations; (ii) Studies based on special populations (e.g., the elderly, pregnant women, or occupational groups) or non-representative areas (e.g., schools, factories, or disaster-affected regions); (iii) Studies using overlapping or duplicate samples; (iv) In provinces with both traditional and NGS screening data, only NGS-based studies were included to ensure methodological consistency; (v) Review articles, case reports, and systematic analyses were excluded; (vi) For studies reporting mutation proportions, all were included regardless of population selection or testing methods, as these factors do not impact mutation proportion estimates. This approach helped expand sample size and minimize selection bias.

Data extraction strategy

The following data were extracted from each study: title, first author, year of research, study location, category of population, age of population, sampling method, screening method, sample size, reported carrier size, and mutation proportions. Data extraction was conducted using a standardized data collection form. Two co-first authors independently screened titles and abstracts. After that, full-text articles of potentially eligible studies were retrieved and independently reviewed by two co-first authors. Each study was assessed based on the predefined inclusion and exclusion criteria to determine its eligibility for final inclusion. The key factors evaluated during full-text screening included study design (cross-sectional epidemiological surveys), population representativeness (general population vs. selected subgroups), data completeness (availability of carrier rate calculations and mutation profiles), and methodological consistency (screening method used, sample size, and genetic analysis techniques). Discrepancies in full-text screening decisions were discussed and resolved through consensus with two corresponding authors with expertise in epidemiology and systematic reviews. This rigorous screening process ensured consistency in study selection and minimized selection bias.

Potential effect modifiers and reasons for heterogeneity

Potential sources of heterogeneity include differences in sampling methods, screening technologies, geographic regions, and population characteristics. These factors were assessed in subgroup analyses where applicable. Additionally, variability in carrier rate estimates due to screening method differences was accounted for by applying normalization factors (see "Normalizations of Carrier Rate" section).

Study quality assessment

The quality of included studies was assessed using the Quality Rating Scheme for Studies and Other Evidence provided by JAMA Network Open (https://jamanetwork.com/journals/jamanetworkopen/pages/instructions-for-authors#SecRatingsofQuality). This rating system evaluates key methodological aspects, including study design, sampling methods, measurement reliability, potential selection bias, and statistical validity. Two authors assessed each study, and discrepancies were resolved by discussion. The detailed quality ratings are presented in Supplementary Table S1.

Data synthesis and presentation

Descriptive statistics were used to summarize thalassemia carrier rates and mutation proportions. To estimate regional carrier rates, we calculated a pooled carrier rate by summing the total number of identified carriers across all included studies and dividing it by the total sample size of these studies. This method accounts for differences in study sample sizes and provides a more representative estimate than a simple arithmetic mean.

Given the inherent geographic variation in thalassemia carrier rates and the fact that no single study provides a complete census of a region's population, heterogeneity between studies is expected and does not indicate methodological inconsistency. Therefore, statistical heterogeneity metrics such as I^2 were not calculated, as they may not be meaningful in this context. Instead, we focused on ensuring that included studies met predefined methodological quality criteria to enhance the reliability of the reported estimates. Publication bias was not formally assessed, as the primary outcome (carrier rate) is

an epidemiological measure rather than a comparative effect size.

Figure and supplementary material generation

All figures and supplementary materials were generated based on data extracted from the included studies. Figure 1 (PRISMA flow diagram) illustrates the study selection process, detailing the number of records screened, excluded, and finally included. Figure 2 and Fig. 3 presents regional thalassemia carrier rates, calculated using a pooled weighted approach based on sample size. Figures were created using ArcMap software for geographic visualization.

Results

Identification and selection of literature and data

From an initial pool of 820 relevant studies published between 1987 and 2024, 29 studies were selected for this systematic analysis of thalassemia prevalence in China. These studies were identified from 337 publications in PubMed, 266 in FMRS, and 217 in CNKI (Fig. 1) [6, 7, 9–35]. Epidemiological data from 15 studies across eight regions, initially based on conventional screening methods, were standardized to calculate carrier rates, while data from other regions were obtained through NGS or TGS methods (Table S1).

Thalassemia carrier rate and incidence in partial areas of China

Thalassemia prevalence data were extracted from primary studies across 17 regions in China, encompassing a total of 679,697 individuals tested (Table 2). National average carrier rates were calculated using a population-weighted approach. Across these 17 regions, prevalence rates for α-thalassemia, β-thalassemia, and α/β-thalassemia were 6.32% (adjusted to 6.38%), 2.34% (adjusted to 2.37%), and 0.20%, respectively.

The surveys were primarily conducted in southern China, with limited data from northern regions. Results show a strong inverse correlation between thalassemia carrier rates and latitude, with higher rates in southern regions that decline progressively northward (Fig. 2). Hainan has the highest overall carrier rate, particularly for α-thalassemia, followed by Guangxi, where both α- and β-thalassemia carrier rates exceed 20%. While Hainan leads in α-thalassemia prevalence, Guangxi has the highest β-thalassemia rates (Table 2). Despite similar latitudes, Guangxi, Guangdong, and Yunnan show significant variation.

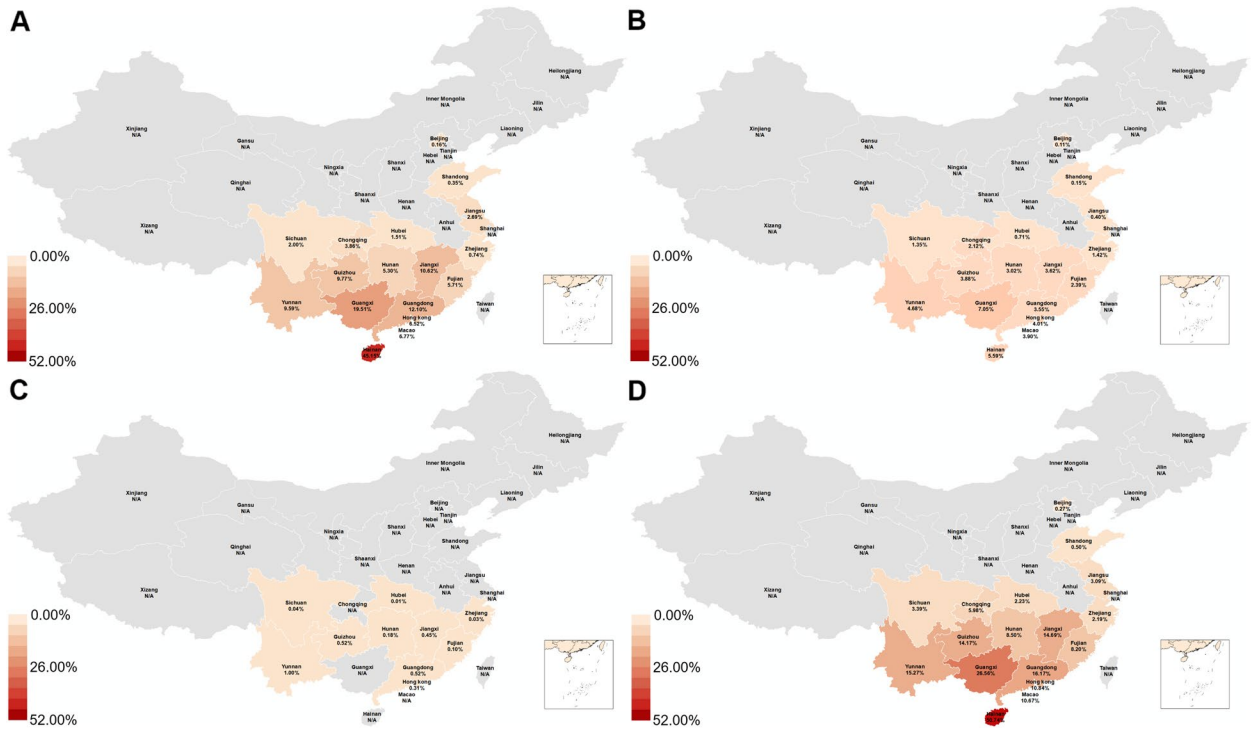

Fig. 2 Geographic distribution of thalassemia carrier rates in regions of China. Figures (**A**-**D**) present the distribution of thalassemia carrier rates across Chinese regions: (**A**) Carrier rates of α-thalassemia, (**B**) Carrier rates of β-thalassemia, (**C**) Carrier rates of α- combine β-thalassemia, (**D**) Total thalassemia carrier rates. Color intensity in each map reflects carrier rates or proportions, with darker shades indicating higher values. Regions lacking carrier rate data are shown in gray. Data were extracted from included studies and plotted using ArcMap GIS software

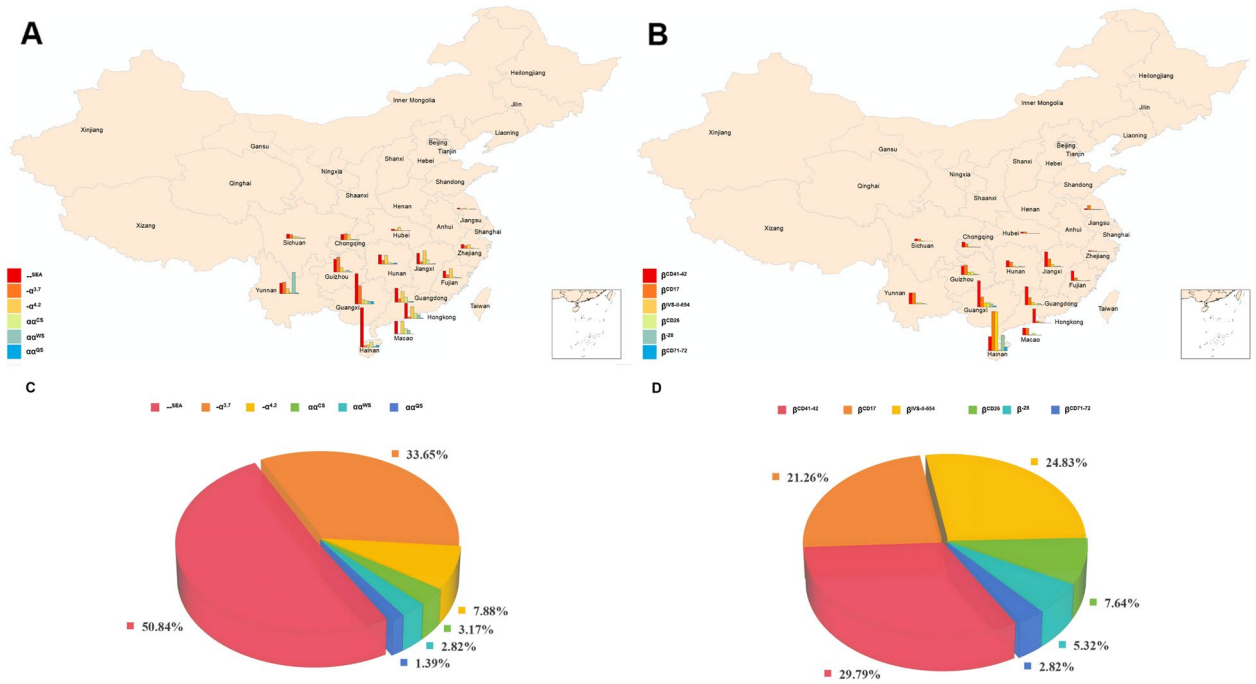

Fig. 3 Geographical distribution and composition of six most common α-thalassemia and β-thalassemia mutations in regions of China. **A** illustrates the composition of six common α-thalassemia globin mutations across various provinces in China. **B** illustrates the composition of six common β-Thalassemia globin mutations across various provinces in China. Each vertical bar represents the proportional contribution of different thalassemia globin deletions and mutations within each region. **C** indicates to the overall proportion of the six most common α-globin gene mutations in China. **D** indicates to the overall proportion of the six most common β-Thalassemia gene mutations in China. The proportion was calculated by assigning weights according to the population size of each region. Data were extracted from included studies and plotted using ArcMap GIS software

The six most common mutations of α-thalassemia and β-thalassemia

The top six most common mutations of α- thalassemia in China were ranked by carrier rates based on data from 1,506,412 individuals across 18 regions between 1997 to 2024 (Table S2). The $--^{SEA}$ mutation predominates in nearly all provinces, confirming its status as the most common prevalent form of α-thalassemia in China (Fig. 3A, C). The $-\alpha^{3.7}$ mutation follows closely as the second most common, exhibiting similar proportion across most regions. However, the proportion of the $-\alpha^{3.7}$ mutation is significantly higher than that of the $--^{SEA}$ mutation in Guizhou, Jiangsu, and Hainan. Notably, in Hainan, the $-\alpha^{4.2}$ and $\alpha\alpha^{WS}$ mutations also occur at higher rates than the $--^{SEA}$ mutation, creating a stark contrast with other regions of China.

The top six most common mutations of β-thalassemia were identified from studies involving 1,494,568 individuals across 21 regions of China, spanning the years 1997 to 2024 (Table S3). Contrasts with the distribution of α-thalassemia, mutation spectrum of β-thalassemia is notably more diverse (Fig. 3B). The $\beta^{CD41-42}$ mutation is prevalent across multiple regions, making it the most common β-thalassemia mutation in China (Fig. 3D). Additionally, the $\beta^{IVS-II-654}$ mutation is prominent in southeastern coastal and central regions, particularly in Hunan, Hubei, and Fujian. In contrast, the β^{CD17} mutation predominates in southwestern regions, particularly in Guizhou. Notably, contrast with other southwestern regions, the most prevalent mutation in Yunnan is β^{CD26}. Furthermore, the β^{-28} mutation constitutes a significant proportion of the mutation spectrum in southeastern coastal areas. These findings underscore the pronounced regional differences in the genetic landscape of β-thalassemia across China.

Economic burden of thalassemia in China

Using regional thalassemia carrier rates, we calculated prevalence across 17 regions and converted these rates into estimated patient numbers to assess the annual economic burden of thalassemia in the absence of intervention (Table S4). Based on data from the *National Bureau of Statistics* of China (https://www.stats.gov.cn/sj/), an estimated 6.20 million newborns are expected in these regions next year. Without effective thalassemia prevention programs, at least 18,818 newborns would develop thalassemia, including approximately 1,117 cases of TDT and 3,590 cases of Bart's hydrops fetalis. TDT

Table 2 Carrier Rates and Incidences of Thalassemia in 17 Regions of China

province	Cases	α- thalassemia		β- thalassemia		Carrier rate of α- and β- thalassemia	Total incidence (‰)
		Carrier rates	Adjusted carrier rate	Carrier rates	Adjusted carrier rate		
Beijing [9]	11,766	0.14%	0.16%	0.10%	0.11%	N/A	< 0.01‰
Chongqing [10]	1841	3.86%	3.86%	2.12%	2.12%	N/A	0.48‰
Fujian [11]	189,414	5.23%	5.71%	2.21%	2.39%	0.10%	1.00‰
Guangdong [7, 12, 13]	31,568	12.10%	12.10%	3.55%	3.55%	0.52%	4.39‰
Guangxi [7]	4834	19.51%	19.51%	7.05%	7.05%	N/A	10.76‰
Guizhou [7, 14, 15]	31,353	9.77%	9.77%	3.88%	3.88%	0.52%	3.13‰
Hainan [7]	2720	45.15%	45.15%	5.59%	5.59%	N/A	51.74‰
Hubei [16–20]	97,689	0.91%	0.99%	0.36%	0.38%	0.01%	0.07‰
Hunan [21, 22]	41,753	5.30%	5.30%	3.02%	3.02%	0.18%	1.01‰
Jiangsu [23]	4276	2.69%	2.69%	0.40%	0.40%	N/A	0.19‰
Jiangxi [24]	136,312	10.62%	10.62%	3.62%	3.62%	0.45%	3.48‰
Shandong [25]	16,098	0.32%	0.35%	0.14%	0.15%	N/A	0.04‰
Sichuan [26–28]	58,675	1.83%	2.00%	1.25%	1.35%	0.04%	0.15‰
Yunnan [6, 7, 29, 30]	11,133	9.59%	9.59%	4.68%	4.68%	1.00%	3.61‰
Zhejiang [31, 32]	32,581	0.68%	0.74%	1.31%	1.42%	0.03%	0.07‰
Hong Kong [33, 34]	1943	5.97%	6.52%	3.71%	4.01%	0.31%	1.63‰
Macao [35]	5741	6.20%	6.77%	3.60%	3.90%	N/A	1.53‰

Both carrier rates of α-thalassemia and β-thalassemia have included the carrier rate of α- and β-thalassemia, respectively; Adjusted value for α-thalassemia carrier rate = α-thalassemia carrier rate × 1.092, Adjusted value for β-thalassemia carrier rate = β-thalassemia carrier rate × 1.082, among the 15 provinces and cities, the thalassemia carrier rate in Hainan, Guangxi, Guangdong, Guizhou, Yunnan and Hunan are all data for NGS test sampled by ethnic population proportion, and there is no need for correction

cases include β-thalassemia major, and about one-third of β-thalassemia intermedia and Hemoglobin H cases. Due to uncertainties in low-frequency variants, TDT estimates are based on the six most prevalent α- and β-thalassemia gene mutations.

According to the *Blue Book of Thalassemia in China (2020)* [4], TDT patients currently have two treatment options in China: transfusion with iron chelation therapy or hematopoietic stem cell transplantation, while gene therapy remains experimental. Approximately 38.17% of TDT patients plan to pursue stem cell transplantation at a cost of 400,000–500,000 RMB, typically between ages 5 and 15, with ongoing transfusions and iron chelation therapy until transplantation is complete. However, only 48.27% find a compatible donor, leaving transfusion and iron chelation therapy as the sole option for most. Effective transfusions treatment extends TDT patients'life expectancy to 40–50 years but with a lifetime cost of 4–7 million RMB. Without treatment, life expectancy drops to 5–15 years, with an average medical expense of approximately 512,000 RMB. Unfortunately, 48.67% of patients do not receive adequate transfusions (Fig. 4A).

Consequently, in the absence of interventions, China's annual medical burden from TDT patients is projected to increase by around 3.03 billion RMB. Besides, since thalassemia manifests at birth, the indirect economic burden is also substantial. Based on average life expectancy (78 years) and an annual per capita GNP of 89,358 RMB, untreated TDT patients lose an average of 68 years of life, while treated patients lose approximately 33 years. Productivity losses are significant: around 50% of TDT patients are unable to work, and 60% require caregiving. Due to high caregiving costs, most families assume caregiving roles, with 70% of caregivers leaving their jobs, further amplifying economic losses. Therefore, the indirect economic burden due to TDT is projected to rise by about 6.16 billion RMB annually. Without a national thalassemia prevention program, the combined direct and indirect economic burden could reach 9.19 billion RMB next year (Fig. 4B).

This estimate considers only TDT patients. Estimating direct medical costs and indirect burdens for non-transfusion-dependent thalassemia (NTDT) is complex due to its variable severity. Nonetheless, without an effective prevention program, an estimated 14,111 new NTDT cases are expected next year, with many requiring transfusions and experiencing reduced work capacity, further contributing to the economic burden. Additionally, approximately 3,590 stillbirths due to Bart's hydrops fetalis are anticipated annually in the absence of preventive measures.

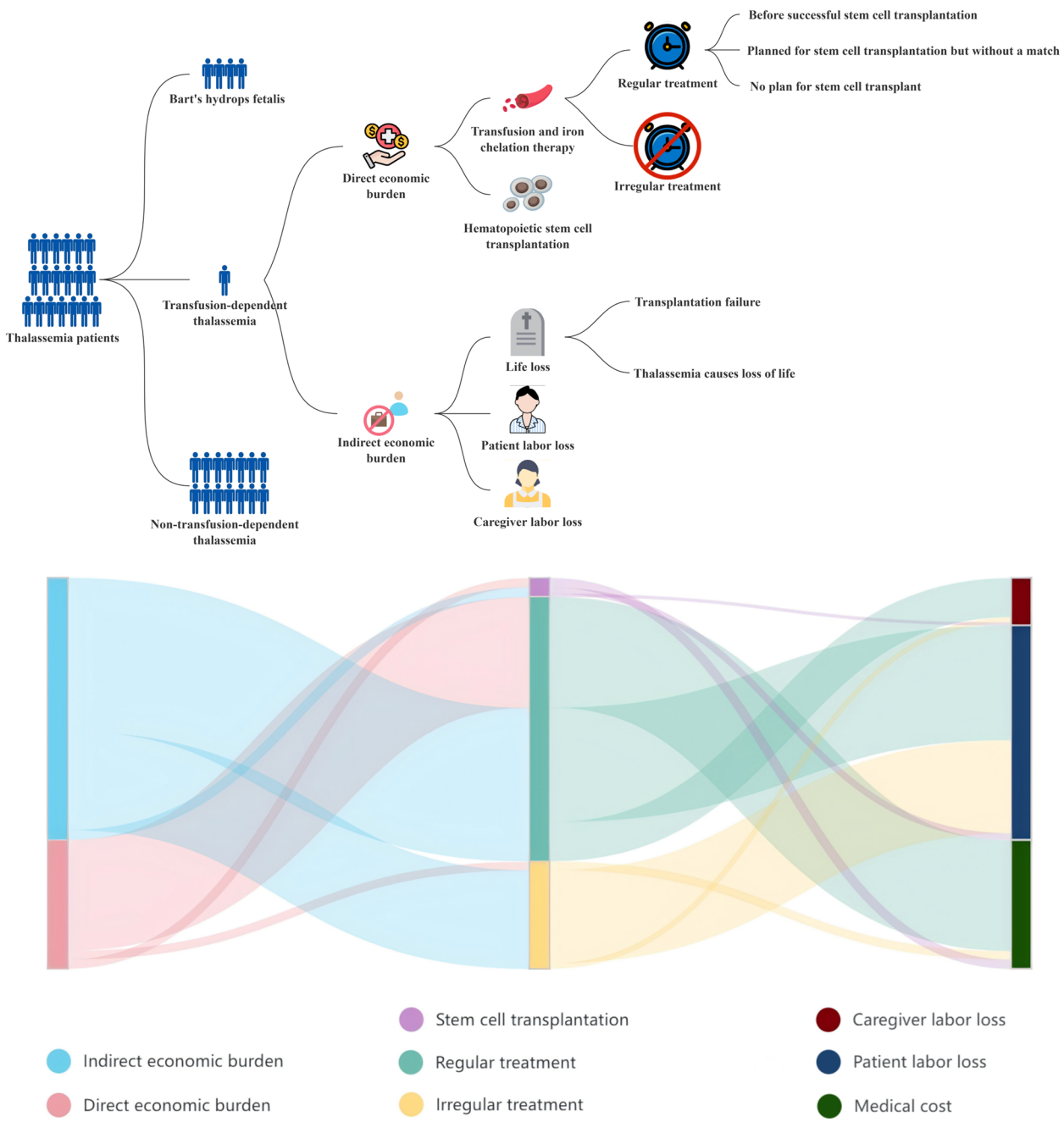

Fig. 4 Economic Burden of Thalassemia Treatment in China. This figure illustrates the economic burden associated with different treatment pathways for thalassemia patients in China. **A** is a tree diagram detailing the economic burden distribution by treatment type, distinguishing between direct (red) and indirect (blue) costs. **B** is a Sankey diagram mapping the flow of economic burdens from initial treatment choices to final financial impacts. On the left, blue denotes indirect economic burden, while red represents direct economic burden. The middle section depicts three primary treatment options: purple for stem cell transplantation, green for regular treatment (timely blood transfusions and iron chelation therapy, including patients who could not find compatible donors), and yellow for irregular treatment (patients without consistent transfusions and chelation therapy). On the right, dark green indicates medical costs, dark blue represents patient labor loss (including premature death and decreased productivity), and dark red shows caregiver labor loss due to family members reducing work to provide care. This figure was produced by the authors based on information from The *Blue Book of Thalassemia in China (2020)*

Discussion

Changes in thalassemia incidence in southern China over the past 37 years

In the late 1970 s, a national survey on hemoglobin disorders in China reported thalassemia incidences of 2.64% for α-thalassemia and 0.66% for β-thalassemia across 12 provinces, establishing a preliminary understanding of thalassemia prevalence [3]. The present study systematically analyzes updated data from 20 regions, focusing primarily on globin gene mutations. Comparisons between current incidences in 10 provinces and data from Zeng et al. in 1987 reveal substantial reductions (Table S5, Fig. 5). Notably, α-thalassemia in Guangxi decreased from 14.95% to 0.95%, with provinces like Sichuan and Zhejiang now reporting minimal rates. Similarly, β-thalassemia incidence dropped from 2.21% to 0.05% in Guizhou and from 1.52% to 0.12% in Guangxi, underscoring effective disease management over time.

The current national average incidences for α- and β-thalassemia have fallen to 0.24% and 0.04%, respectively, marking significant progress. This decline likely reflects the success of the thalassemia prevention program launched in southern China in 2012, which emphasizes carrier screening before marriage and pregnancy, along with prenatal diagnostics for high-risk families. These measures have led to a substantial reduction in TDT births, achieving"zero TDT births"in some regions.

Unique mutation profiles of thalassemia in China

Although Hainan has the highest thalassemia carrier rate in China, its mutation profile is distinct from other regions. The most common α-thalassemia mutation in Hainan is $-\alpha^{3.7}$, followed by $-\alpha^{4.2}$, with the typically prevalent $--^{SEA}$ mutation only third. This unusual distribution may be linked to the Li population, Hainan's largest ethnic minority, where $-\alpha^{3.7}$ and $-\alpha^{4.2}$ mutations predominate, and $--^{SEA}$ accounts for less than 5%. In contrast, approximately 20% of Hainan's Han Chinese population carries the $--^{SEA}$ mutation, a proportion notably different from both the Li population and Han populations in other regions [36, 37]. A similar pattern is observed

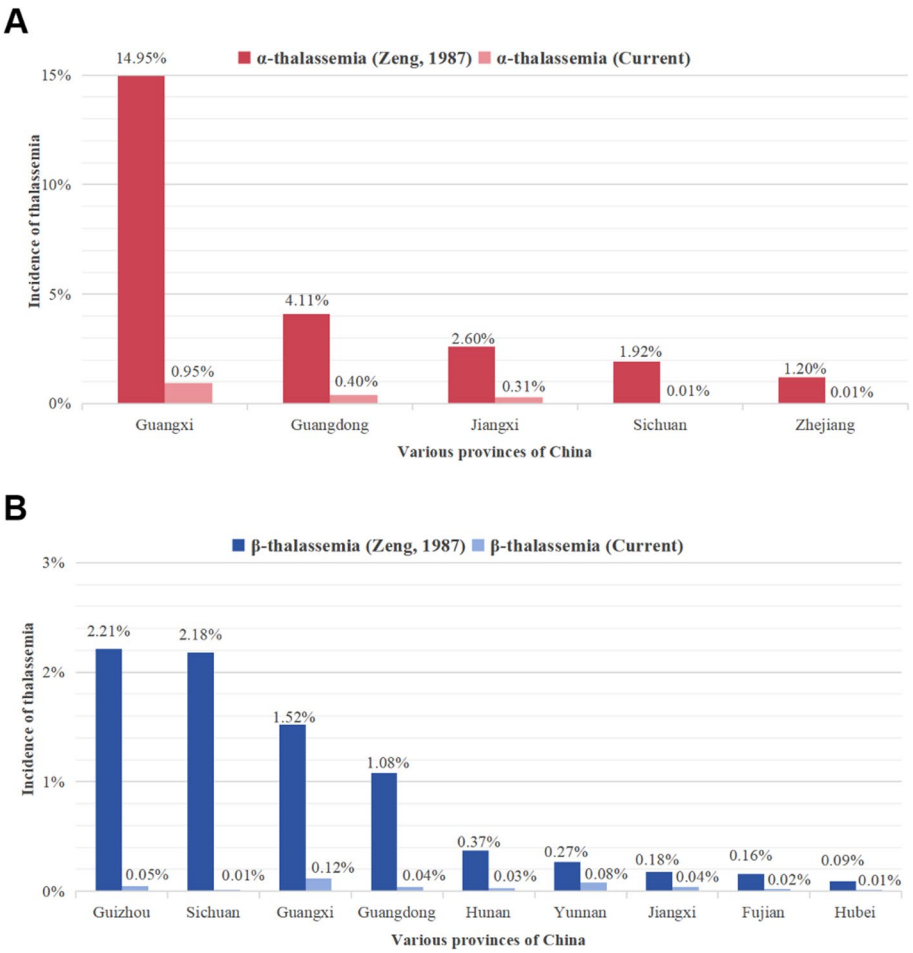

Fig. 5 Change trend of incidence of thalassemia over 37 Years. **A** represents the changes in α-thalassemia incidence, comparing the data from report of Zeng et al. in 1987 with the current rates. **B** represents a similar comparison for β-thalassemia

in Yunnan—another ethnically diverse province—where the most common β-thalassemia mutations are β^{CD26} and β^{CD17}, with $\beta^{CD41-42}$, usually dominant in other regions, ranking third [6, 38, 39]. Research indicates that areas like Yunnan and Hainan exhibit mutation profiles distinct from other parts of China, seen not only among local ethnic minorities but also within regional Han populations. The geographic distribution of these mutations likely reflects historical migration patterns, population genetics, and varying effects of natural selection, particularly in response to malaria prevalence.

Thalassemia remains a significant disease burden in China

The *Blue Book of Thalassemia in China (2020)* reports that approximately 15,000 patients with thalassemia major incur a collective treatment cost of 72 billion RMB [4]. Nationwide, an estimated 300,000 individuals have thalassemia major or intermedia, with many intermedia patients also requiring medical care. Thalassemia patients often face disruptions in education and employment due to frail health and frequent hospital visits, leading to lower educational attainment, limited job skills, and reduced employment prospects—further intensifying the economic burden beyond direct medical expenses.

In this study, we focused on estimating the economic burden for TDT patients, but we did not include intermediate or mild thalassemia cases. Although TDT patients represent a smaller portion of the total burden, intermediate and mild cases, being more common and having longer life expectancy, would likely contribute to a much higher economic burden. Unfortunately, current data and methods do not allow us to estimate the burden for these groups, so we cannot provide a clear range for the total economic burden. As a result, our estimates only reflect the burden from TDT patients, which is likely an underestimation of the true economic impact. Given the lack of reliable data for intermediate and mild cases, we cannot conduct a sensitivity analysis. Therefore, these estimates should be viewed as indicative rather than definitive.

The China Birth Defects Intervention and Assistance Foundation has contributed over 10 million RMB to support 1,500 children with thalassemia major in provinces such as Guizhou, Fujian, Guangxi, and Hainan. Despite these efforts, significant unmet needs remain. A comprehensive understanding of the disease burden is hindered by the lack of a national thalassemia registry.

Efficiency of NGS and TGS screening in high-prevalence thalassemia regions

NGS and TGS have significantly improved thalassemia screening in high-prevalence areas. Compared to conventional methods, NGS identifies a broader range of rare hemoglobin variants (e.g., Hb New York, Hb Hekinan, Hb J-Bangkok), markedly enhancing detection rates [7]. He et al. found that NGS identified 22.9% more carriers than hematology screening alone [6], with additional studies supporting its efficacy [22, 34, 40].

Despite these benefits, NGS has limitations with complex α- and β-globin gene mutations. Its short-read length, though ideal for high-throughput analysis of point mutations and small indels, is less effective for detecting large structural variations (e.g., inversions, translocations, tandem duplications), homologous or repetitive region mutations (e.g., HBA1/HBA2 or HBB/HBD), repeat amplifications (e.g., ααααanti3.7 and ααααanti4.2), de novo mutations in high-GC content areas, and epigenetic modifications. Additionally, NGS cannot determine cis/trans configurations of mutations [41, 42].

TGS, with its long-read capabilities, overcomes many of NGS's limitations by covering extensive gene regions without PCR, providing faster and more accurate genotyping [43–45]. However, TGS has challenges: higher error rates, specialized library requirements, increased costs, and more intensive data processing and storage demands [46, 47].While TGS's cost and complexity make it less viable for large-scale screening [48], it is particularly useful for severe phenotypes unexplained by conventional methods. In such cases, TGS, combined with genome assembly, allows for rapid and precise identification of rare mutations and gene modifications, contributing valuable insights into disease pathology [47].

Current status of thalassemia prevention and control in China

Currently, limited awareness and screening motivation outside major healthcare hubs, combined with restricted access in some regions, highlight the need for targeted public health outreach. In 2023, China plans to establish a national network of 101 medical institutions across 10 provinces to improve access to high-quality medical resources. This coordinated network aims to advance thalassemia prevention and care by fostering collaboration among institutions, maternal and child healthcare facilities, and primary care centers, while streamlining resources and enhancing service models from prevention through treatment.

For high-risk families, current treatment options have significant limitations. Blood transfusions and iron chelation impose considerable medical and indirect economic burdens, while hematopoietic stem cell transplantation and gene therapy face unresolved challenges in preserving fertility before adolescence [49]. Preconditioning for these therapies can impair fertility, with optimal transplantation timing falling prior to adolescence [50, 51].

Given these challenges, preimplantation genetic testing (PGT) could offer a viable preventive alternative [52].

Compared to aforementioned treatments, PGT may substantially reduce families'direct and indirect financial burdens, although further studies are needed to confirm its long-term efficacy and cost-effectiveness. Expanding preventive measures like PGT, along with enhanced education and healthcare resources, could improve outcomes and alleviate the socioeconomic impact of thalassemia on families and society.

Limitations and future directions

This study is the first to evaluate thalassemia carrier rates across China using data exclusively from NGS or TGS platforms or from conventional screenings normalized to NGS standards, ensuring standardized carrier rate estimates across different detection methods. These methodological advancements highlight the unique contribution of our study to the field of thalassemia epidemiology. However, underrepresentation of data from regions (northern provinces) with lower carrier rates may lead to an overestimation of the national carrier rate, and geographic and ethnic differences across provinces could influence the accuracy of the estimates. These factors may limit the applicability of our conclusions to all regions of China. Further studies with more comprehensive data coverage to strengthen the validity of national estimates will be needed. Nonetheless, our economic burden estimate remains conservative, focusing on TDT cases from the 12 most prevalent mutations, indicating that the true burden could be greater. Selection bias is another concern, as China's geographic and ethnic diversity contribute to substantial regional variations in carrier rates, even within the same province or city. Few studies account for urban or ethnic demographics, and variability among laboratories may further introduce bias. Additionally, our literature search was limited to PubMed, FMRS and CNKI, without incorporating more databases. While this approach ensured a focused selection of high specificity studies, it may have excluded some relevant publications. However, we mitigated this limitation by applying strict inclusion criteria, full-text screening, and adherence to PRISMA guidelines, reducing the likelihood of omitting key studies.

Establishing a national thalassemia registry that incorporates big data analytics is crucial for obtaining precise epidemiological data and optimizing resource allocation for effective prevention and treatment. In the absence of a standardized national survey, this study offers an initial assessment of China's thalassemia burden, drawing on available carrier rates and mutation profiles. Observed trends over the last three decades, along with current prevention efforts, may serve as useful references for other regions with high prevalence of thalassemia and hemoglobinopathies.

Conclusion

In summary, China, with its long-standing history of thalassemia, has achieved substantial progress in disease control. This study systematically analyzed epidemiological data from 17 regions across the country, showing notable improvements attributed to government-led prevention policies and the efforts of healthcare professionals. Although incidence has decreased under these programs, thalassemia remains a significant health burden. Further prevention strategies and ongoing planning are essential in high-prevalence regions.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13690-025-01575-7.

Supplementary Material 1.

Supplementary Material 2.

Acknowledgements
Not applicable.

Authors' contributions
B.Z. and F.Z. are responsible for ensuring the integrity of the data and the accuracy of the analysis. Both authors contributed equally to this work. Concept and design: B.Z., F.Z., A.C., Q.M. and X.L.. Acquisition, analysis, or interpretation of data: A.C., Q.M. and X.L.. Drafting of the manuscript: A.C. and X.L.. Critical review of the manuscript for important intellectual content: All authors. Statistical analysis: A.C., Q.M., X.L. and G.H.. Funding acquisition: B.Z. and A.C.. Supervision: B.Z., F.Z. and J.H..

Funding
This study was supported by grant 81760037 from the National Natural Science Foundation of China (NSFC), Yunling Scholar Project of Yunnan Province (YNWR-YLXZ- 2019–0005) and Yunnan Provincial Clinical Medical Center for Blood Diseases and Thrombosis Prevention and Treatment.

Data availability
No datasets were generated or analysed during the current study.

Declarations

Ethics of this study was approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 24 December 2024 Accepted: 24 March 2025
Published online: 05 April 2025

References

1. Taher, A. T., Weatherall, D. J. & Cappellini, M. D. Thalassaemia. Lancet 391, 155–167 (2018). https://doi.org/10.1016/s0140-6736(17)31822-6

2. Modell, B. & Darlison, M. Global epidemiology of haemoglobin disorders and derived service indicators. Bull World Health Organ 86, 480–487 (2008). https://doi.org/10.2471/blt.06.036673
3. Zeng, Y. T. & Huang, S. Z. Disorders of haemoglobin in China. J Med Genet 24, 578–583 (1987). https://doi.org/10.1136/jmg.24.10.578
4. Foundation, B. A. C. & Institute, C. P. R. Blue Book of Thalassemia in China (2020). 1 edn, (China Society Press, 2021).
5. Lai, K., Huang, G., Su, L. & He, Y. The prevalence of thalassemia in mainland China: evidence from epidemiological surveys. Sci Rep 7, 920 (2017). https://doi.org/10.1038/s41598-017-00967-2
6. He, J. et al. Next-generation sequencing improves thalassemia carrier screening among premarital adults in a high prevalence population: the Dai nationality, China. Genet Med 19, 1022–1031 (2017). https://doi.org/10.1038/gim.2016.218
7. Shang, X. et al. Rapid Targeted Next-Generation Sequencing Platform for Molecular Screening and Clinical Genotyping in Subjects with Hemoglobinopathies. EBioMedicine 23, 150–159 (2017). https://doi.org/10.1016/j.ebiom.2017.08.015
8. Page, M. J. et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. J Clin Epidemiol 134, 178–189 (2021). https://doi.org/10.1016/j.jclinepi.2021.03.001
9. Wang, X. Y. & Liu, J. H. [An Analysis of Thalassemia Genes in Beijing Area]. Labeled Immunoassays and Clinical Medicine 29, 221–223 (2022). https://doi.org/10.11748/bjmy.issn.1006-1703.2022.02.009
10. Zuo XZ, et al. Application of Next-generation Sequencing Technology in Screening for Thalassemia. Journal of Practical Obstetrics and Gynecology. 2022;38:146–9.
11. Huang, H. et al. Molecular characterization of thalassemia and hemoglobinopathy in Southeastern China. Sci Rep 9, 3493 (2019). https://doi.org/10.1038/s41598-019-40089-5
12. LI, W. Y. et al. [Application of Next Generation Sequencing to Screen the Thalassemia in Pregnant Women]. Shenzhen Journal of Integrated Traditional Chinese and Western Medicine 34, 30–34 (2024). https://doi.org/10.16458/j.cnki.1007-0893.2024.07.008
13. Chen, Y. et al. Carrier rate of thalassemia among 25,910 high school students in Shaoguan area, China. J Med Screen 31, 53–57 (2024). https://doi.org/10.1177/09691413231188069
14. Wu, L. S. et al. Prevalence of thalassemia-carrier couples and fertility risk assessment. Int J Hematol 119, 374–382 (2024). https://doi.org/10.1007/s12185-024-03722-2
15. Tan, M. et al. Early genetic screening uncovered a high prevalence of thalassemia among 18 309 neonates in Guizhou, China. Clin Genet 99, 704–712 (2021). https://doi.org/10.1111/cge.13923
16. Zhang, W. et al. [Analysis of Neonates'Thalassemia Gene Detection in Huangshi]. Journal of Hubei Polytechnic University 35, 58–61 (2019). https://doi.org/10.3969/j.issn.2095-4565.2019.06.012
17. Yuan, Y. H., Wang, L. L., Deng, W. & Xu, Y. [Analysis of screening results of 28478 newborns with thalassemia in Yichang]. Chin J Fam Plann 28, 1711–1712 (2020). https://doi.org/10.3969/j.issn.1004-8189.2020.10.051
18. Cheng, Y. et al. The prevalence and outcomes of α- and β-thalassemia among pregnant women in Hubei Province, Central China: An observational study. Medicine (Baltimore) 101, e28790 (2022). https://doi.org/10.1097/md.0000000000028790
19. Wang, D., Zhang, D., Yan, X. S. & Liu, Y. M. [The Analysis of Thalassemia Gene Detection Results and Hematological Screening Indicators in the Population of Wuhan Economic and Technological Development Zone]. Labeled Immunoassays and Clinical Medicine 29, 1369–1372+1440 (2022). https://doi.org/10.11748/bjmy.issn.1006-1703.2022.08.023
20. Deng, Y. J., Li, Y., Yang, Y., Gong, Q. & Zhang, Z. J. [Analysis of prenatal screening results of 4 629 pregnant women with thalassemia in Shiyan area of Hubei Province]. Journal of Dalian Medical University 45, 509–512 (2023). https://doi.org/10.11724/jdmu.2023.06.06
21. Xi, H. et al. Epidemiological Survey of Hemoglobinopathies Based on Next-Generation Sequencing Platform in Hunan Province, China. Biomed Environ Sci 36, 127–134 (2023). https://doi.org/10.3967/bes2023.016
22. Zhang, H. et al. Next-generation sequencing improves molecular epidemiological characterization of thalassemia in Chenzhou Region, P.R. China. J Clin Lab Anal 33, e22845 (2019). https://doi.org/10.1002/jcla.22845
23. Shao, B. et al. Mutation spectrum of thalassemia among pre-pregnant adults in the Jiangsu Province by capillary electrophoresis-based multiplex PCR assay. Mol Genet Genomic Med 12, e2344 (2024). https://doi.org/10.1002/mgg3.2344
24. Yang, T. et al. Next-generation sequencing analysis of the molecular spectrum of thalassemia in Southern Jiangxi, China. Hum Genomics 17, 77 (2023). https://doi.org/10.1186/s40246-023-00520-5
25. Liu, R., Zhu, N., Luo, M., Zhao, D. & Chu, Z. H. [Screening and genetic diagnosis of infant thalassemia in Shandong province]. Chinese Journal of Clinical Medicine 27, 1043–1045 (2020). https://doi.org/10.12025/j.issn.1008-6358.2020.20201075
26. Yu, X. et al. Genetic investigation of haemoglobinopathies in a large cohort of asymptomatic individuals reveals a higher carrier rate for β-thalassaemia in Sichuan Province (Southwestern China). Genes Dis 8, 224–231 (2021). https://doi.org/10.1016/j.gendis.2019.11.001
27. Yu, X. et al. Molecular Epidemiological Investigation of Thalassemia in the Chengdu Region, Sichuan Province, Southwest China. Hemoglobin 39, 393–397 (2015). https://doi.org/10.3109/03630269.2015.1070733
28. Li, B., Han, X., Ma, J. & Yang, D. Mutation spectrum and erythrocyte indices characterisation of α-thalassaemia and β-thalassaemia in Sichuan women in China: a thalassaemia screening survey of 42 155 women. J Clin Pathol 74, 182–186 (2021). https://doi.org/10.1136/jclinpath-2020-206588
29. Jiang, T. et al. [Thalassemia Screening and Genetic Diagnosis Results Analysis in Wenshan Zhuang-Miao Autonomous Prefecture,Yunnan Province]. Journal of Kunming University of Science and Technology(Natural Science Edition) 48, 162–167 (2023). https://doi.org/10.16112/j.cnki.53-1223/n.2023.05.253
30. He, J. P. et al. [Study on Screening of Thalassemia Genes by Next-generation Sequencing in Kunming Area]. Journal of Modern Laboratory Medicine 37, 6–8+49 (2022). https://doi.org/10.3969/j.issn.1671-7414.2022.05.002
31. Jiang, H., Zhou, C., Yang, L., Li, S. & Liu, X. [Carrier screening and prenatal diagnosis for thalassemia-associated mutations in Jiaxing area of Zhejiang]. Zhonghua Yi Xue Yi Chuan Xue Za Zhi 40, 295–300 (2023). https://doi.org/10.3760/cma.j.cn511374-20220121-00054
32. Ding, Z. Y., Shen, G. S., Zhang, S. & He, P. Y. Epidemiology of Hemoglobinopathies in the Huzhou Region, Zhejiang Province, Southeast China. Hemoglobin 40, 304–309 (2016). https://doi.org/10.1080/03630269.2016.1200988
33. Lau, Y. L. et al. Prevalence and genotypes of alpha- and beta-thalassemia carriers in Hong Kong -- implications for population screening. N Engl J Med 336, 1298–1301 (1997). https://doi.org/10.1056/nejm199705013361805
34. Chan, O. Y. M. et al. Expanded carrier screening using next-generation sequencing of 123 Hong Kong Chinese families: a pilot study. Hong Kong Med J 27, 177–183 (2021). https://doi.org/10.12809/hkmj208486
35. Ribeiro, M. L. & Tamagnini, G. P. Hemoglobin disorders in Macao. Hemoglobin 21, 271–279 (1997). https://doi.org/10.3109/03630269708997388
36. Yao, H. et al. The spectrum of α- and β-thalassemia mutations of the Li people in Hainan Province of China. Blood Cells Mol Dis 53, 16–20 (2014). https://doi.org/10.1016/j.bcmd.2014.01.003
37. Wang, M., Zhang, X., Zhang, Y. & Xiao, M. Prevalence and Genetic Analysis of Thalassemia and Hemoglobinopathy in Different Ethnic Groups and Regions in Hainan Island, Southeast China. Front Genet 13, 874624 (2022). https://doi.org/10.3389/fgene.2022.874624
38. Zhang, J. et al. The spectrum of α- and β-thalassemia mutations in Yunnan Province of Southwestern China. Hemoglobin 36, 464–473 (2012). https://doi.org/10.3109/03630269.2012.717327
39. Zhang, J. et al. Genetic heterogeneity of the β-globin gene in various geographic populations of Yunnan in southwestern China. PLoS One 10, e0122956 (2015). https://doi.org/10.1371/journal.pone.0122956
40. Zhao, J., Li, J., Lai, Q. & Yu, Y. Combined use of gap-PCR and next-generation sequencing improves thalassaemia carrier screening among premarital adults in China. J Clin Pathol 73, 488–492 (2020). https://doi.org/10.1136/jclinpath-2019-206339
41. Hassan, S. et al. Next-Generation Sequencing (NGS) and Third-Generation Sequencing (TGS) for the Diagnosis of Thalassemia. Diagnostics (Basel) 13 (2023). https://doi.org/10.3390/diagnostics13030373
42. Huang, R. et al. Back-to-Back Comparison of Third-Generation Sequencing and Next-Generation Sequencing in Carrier Screening of Thalassemia. Arch Pathol Lab Med 148, 797–804 (2024). https://doi.org/10.5858/arpa.2022-0168-OA

43. Toledo, D. M. & Lafferty, K. A. Clinical Perspective on Use of Long-Read Sequencing in Prenatal Diagnosis of Thalassemia. Clin Chem 69, 211–212 (2023). https://doi.org/10.1093/clinchem/hvac223
44. Ren, Z. M. et al. Detecting rare thalassemia in children with anemia using third-generation sequencing. Hematology 28, 2241226 (2023). https://doi.org/10.1080/16078454.2023.2241226
45. Huang, W. et al. Nanopore Third-Generation Sequencing for Comprehensive Analysis of Hemoglobinopathy Variants. Clin Chem 69, 1062–1071 (2023). https://doi.org/10.1093/clinchem/hvad073
46. van Dijk, E. L. et al. Genomics in the long-read sequencing era. Trends Genet 39, 649–671 (2023). https://doi.org/10.1016/j.tig.2023.04.006
47. Zhang, Q. et al. An unusual case of thalassemia intermedia with inheritable complex repeats detected by single-molecule optical mapping. Haematologica 109, 1000–1006 (2024). https://doi.org/10.3324/haematol.2023.282902
48. Chen, Y. J. & Xu, Z. H. [Progress in the Genetic Detection of Thalassemia Based on Third-Generation Gene Sequencing --Review]. Zhongguo Shi Yan Xue Ye Xue Za Zhi 32, 634–638 (2024). https://doi.org/10.19746/j.cnki.issn.1009-2137.2024.02.047
49. Hwee, T., Bergen, K., Leppke, S., Silver, A. & Loren, A. Hematopoietic Cell Transplantation and Utilization of Fertility Preservation Services. Biol Blood Marrow Transplant 25, 989–994 (2019). https://doi.org/10.1016/j.bbmt.2019.01.005
50. Yang, E. H., Strohl, H. B. & Su, H. I. Fertility preservation before and after cancer treatment in children, adolescents, and young adults. Cancer 130, 344–355 (2024). https://doi.org/10.1002/cncr.35108
51. Chen, L., Dong, Z. & Chen, X. Fertility preservation in pediatric healthcare: a review. Front Endocrinol (Lausanne) 14, 1147898 (2023). https://doi.org/10.3389/fendo.2023.1147898
52. Mamas, T., Kakourou, G., Vrettou, C. & Traeger-Synodinos, J. Hemoglobinopathies and preimplantation diagnostics. Int J Lab Hematol 44 Suppl 1, 21–27 (2022). https://doi.org/10.1111/ijlh.13851

Publisher's Note