BMJ Public Health

Sickle cell allele distribution in southern India: a population-based study in a rural–tribal context

https://bmjpublichealth.bmj.com/content/bmjph/3/2/e003108.full.pdf

Yogish Channa Basappa ,[1] Pooja Aggarwal ,[1,2] Prashanth N Srinivas,[1] Lavanya B Ramegowda,[3,4] Giriraj R Chandak,[5] Deepa Bhat [2]

To cite: Channa Basappa Y, Aggarwal P, Srinivas PN, *et al*. Sickle cell allele distribution in southern India: a population-based study in a rural–tribal context. *BMJ Public Health* 2025;**3**:e003108. doi:10.1136/bmjph-2025-003108

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/bmjph-2025-003108).

YCB and PA are joint first authors.

Received 25 April 2025
Accepted 15 August 2025

[1]Institute of Public Health Bengaluru, Bangalore, Karnataka, India
[2]Department of Anatomy, JSS Medical College, Mysuru, Karnataka, India
[3]Department of Biochemistry, JSS Medical College, Mysuru, Karnataka, India
[4]JSS Academy of Higher Education and Research, Mysuru, Karnataka, India
[5]CSIR - Centre for Cellular and Molecular Biology, Hyderabad, Telangana, India

Correspondence to
Dr Deepa Bhat;
deepabhat@jssuni.edu.in

ABSTRACT

Background and objective Sickle cell disease (SCD) is a genetic disorder of haemoglobin affecting red blood cells with the second highest burden in India. In this study, we examine the epidemiology of SCD in Chamarajanagar district, a largely rural south Indian district in Karnataka with a mixed rural tribal and non-tribal population to generate a more comprehensive understanding of the prevalence of SCD in southern India through the axes of social demographics.

Methods We collected household and individual sociodemographic data and blood samples for determining SCD status using a cross-sectional study design among a subsample of consenting participants from forest-dwelling tribal communities and non-tribal communities. This study is nested in a previous study, 'Towards Health Equity & Transformative Action on Tribal Health', a multisite research study on tribal health inequities.

Results Of the 547 study participants screened, 80 (14%) were found positive for the sickle cell allele and 238 (44%) were found anaemic. The association between selected sociodemographic characteristics and the sickle cell allele was also explored and presented for both scheduled tribes (STs) and non-ST populations.

Interpretation and conclusions We observed a higher prevalence of sickle cell allele among people residing in tribal and remote villages. Although no significant association was found between sickle cell allele and demographics like sex, marital status and literacy, there was a considerable association between sickle cell allele and Socio-Geographical Disadvantage Index among ST and non-ST communities, the implications of which are discussed in the study.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Sickle cell disease (SCD) is a significant public health issue in India, with high prevalence in tribal populations, including in Karnataka. Yet, most research overlooks non-tribal groups living in the same areas and seldom considers social determinants of health.

WHAT THIS STUDY ADDS

⇒ This study reports on SCD epidemiology in Chamarajanagar district, a largely rural south Indian district in Karnataka with a mixed rural tribal and non-tribal population.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ The high frequency of sickle cell allele observed in tribal remote villages highlights the need to establish comprehensive healthcare models suited for communities with social vulnerabilities.

INTRODUCTION

Sickle cell disease (SCD) is caused by a genetic disorder due to a single gene mutation in the beta-haemoglobin gene that leads to a chronic disease with various systemic manifestations along the life course. While a single copy of the sickle gene mutation results in having the trait or becoming a carrier, two copies result in the disease. It has been recognised as one of several neglected diseases globally, particularly in sub-Saharan Africa and other low-income and middle-income countries (LMICs).[1] The ethical imperative for action on over a century's neglect of SCD has been highlighted, especially in LMICs, given recent advancements in diagnostics, treatment, and long-term support for SCD survivors.[2]

There are several gaps in understanding the epidemiology of the disease in the Indian population. Early studies were largely limited to describing its presence and estimating prevalence among tribal populations in southern and central India.[3] In India, scheduled tribes (STs) are constitutionally recognised indigenous communities identified as socially and economically disadvantaged. In contrast, non-STs include other population groups that have greater access to education, employment and health services. These groups are distinct from scheduled castes (SCs), who have historically experienced caste-based discrimination, and from the non-scheduled populations, which include the dominant caste and other social groups.[4] The higher presence of SCD among SC and ST populations has

been attributed to (a) malaria endemicity,[5] (b) higher endogamy rate[6] and (c) competitive exclusion of sickle cell mutation by other Hb variants such as B-thalassaemia.[7] With more population-based screenings, SCD has also been diagnosed in non-tribal groups (including SC groups), although with a lower prevalence (0.84%) compared with the tribal groups (4.05%).[8–11] The most recent modelling-based estimates by Hockham *et al*, based on a pooling of 249 datasets from 2010 to 2017, revealed a state-level prevalence ranging from 0% to 20% in 18 states (out of 28 states and 8 UTs in India).[9] Many datasets included in their analysis (171 out of 249) were largely tribal, revealing the possibly higher focus on tribal populations in SCD research in India and fewer studies from southern India resulting in the low prevalence estimation for this region. Hence, several gaps in geographical coverage of wider population groups, particularly from Karnataka, have limited modelling outcomes yielding conservative estimates.[9 12]

The Indian government recently launched the National Sickle Cell Anaemia Elimination Mission (NSCAEM) with the aim "*To improve care of all Sickle Cell Disease patients for their better future and to lower the prevalence of the disease through multi-faced coordinated approach towards screening and awareness strategies*".[13] With this mission, there is an opportunity now to address the neglect of SCD as well as the inequities faced by tribal and other remote rural communities in accessing diagnosis and treatment options. While the programme's high focus on tribal populations addresses an important equity issue, the need to expand (if not universalise) screening, treatment and follow-up services across a wider population group reflecting the wider burden of SCD is currently unexplored.

In this study, we examine the epidemiology of SCD in Chamarajanagar district, a largely rural south Indian district in Karnataka with a mixed rural tribal and non-tribal population living close to protected forest areas. Tribal populations living in or near forest areas in India and elsewhere have historically faced restrictions on their land ownership and mobility due to strict wildlife conservation measures, which too have contributed to the historical neglect by public services and systems.[14 15] Multiple factors determine the health outcome of people with SCD (PwSCDs), such as access to healthcare, cultural practices, awareness and socioeconomic status.[16] By describing the epidemiology of SCD among tribal and nearby non-tribal populations in and around two large protected areas in Chamarajanagar, we aim to contribute to a more comprehensive understanding of the prevalence of SCD in southern India through the axes of social demographics.[12]

METHODS

This study is nested within the larger Towards Health Equity & Transformative Action on Tribal Health (THETA) study in the Karnataka site, which was conducted in two phases. In phase 1, we collected household and individual sociodemographic information. In phase 2, a subset of participants identified and consented during phase 1 from the households was further contacted, and their SCD status was determined along with their dyslipidaemia and diabetes status.[17 18]

Study setting

Chamarajanagar is the southernmost district in Karnataka, located at the trijunction of Karnataka, Tamil Nadu and Kerala (figure 1). Nearly half (48%) of the district is forested and protected under India's Wildlife Protection Act of 1972. The district has two tiger reserves, the Biligiri Ranganathaswamy Hills (BR Hills) and the Bandipur tiger reserve, and a wildlife sanctuary, Maleya Mahadeshwara Wildlife Sanctuary (MM Hills) (figure 1).[17 18] Chamarajanagar district is ranked 21st out of the 30 districts in Karnataka on the Human Development Index (0.601) and lacks major industries that can provide sustainable livelihoods.[19] The majority of the population lives in rural areas (83%); around 12% (1,20,219) belong to ST communities. Of the 51 ST communities in Karnataka, 12 are forest-dwelling Adivasi groups.[20] Among these, approximately 30 000 adivasis from three major communities—Soliga, Betta Kuruba and Jennu Kuruba—reside in Chamarajanagar.[21]

Sampling

This study is nested within the THETA project, a multisite research study on tribal health inequities. We have sampled forest-dwelling tribal communities (adivasis) and non-tribal communities living in and around the BR Hills and MM Hills, which spread across four talukas of the five in the district, all of which have a higher proportion of forest-associated ST population, viz. Hanuru, Kollegal, Yelandhur and Chamarajanagar (talukas are administrative subdivisions of districts with typical populations ranging from 100 000 to 400 000). These specific talukas were selected because they have a higher proportion of forest-associated ST populations. We excluded the Bandipur Tiger Reserve from our sampling area, as its residents have been relocated as part of conservation efforts. A multistage stratified sampling method was used to recruit the participants across the gradient of the Socio-Geographical Disadvantage Index (SGDI). SGDI was calculated as the aggregate index based on predefined variables such as transport access to healthcare facilities, education etc for both B R Hills and M M Hills protected area, as explained in the previously published THETA protocol paper.[17 18] Online supplemental figure 1 provides a detailed overview of the sampling and how the villages were selected.

Among the samples from ST communities, most samples were from the Soliga adivasi community because they are a significant ST population living in and around MM Hills Wildlife Sanctuary and BR Hills Tiger Reserve Forest.

Sample size

We estimated a sample size of 311 for each site, aiming for a total of 622 participants. This estimation assumes

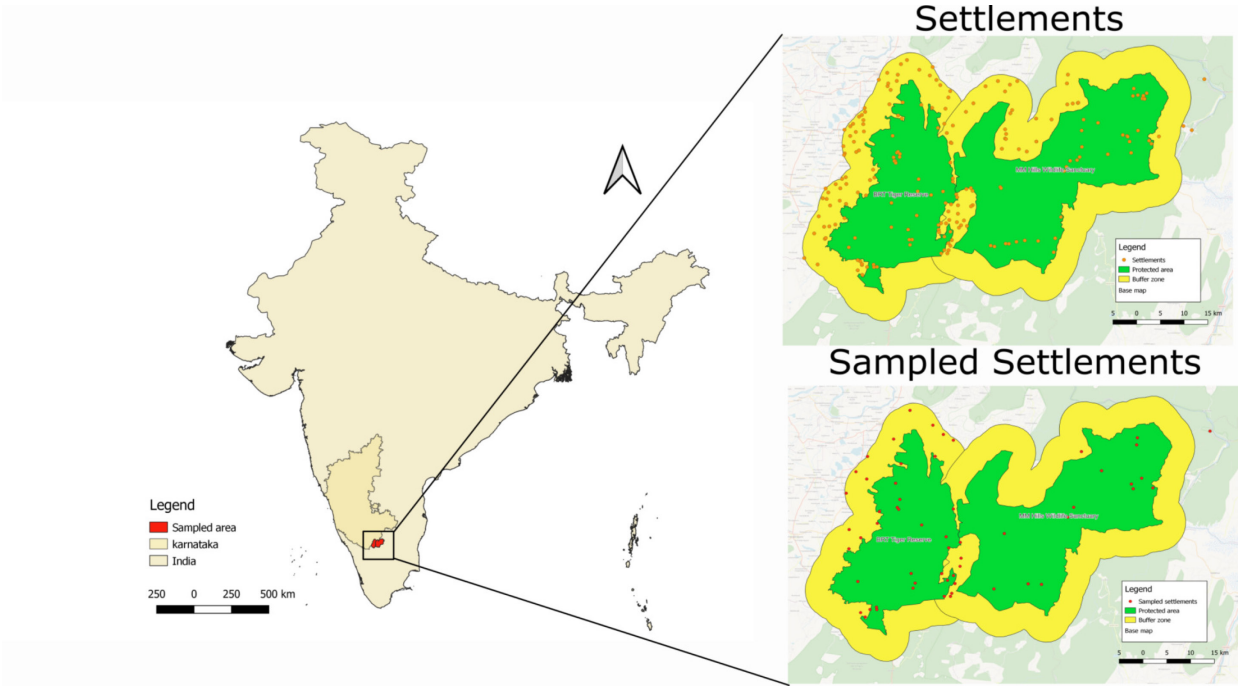

Figure 1 Map of study sites. The Biligiriranga Ranganathaswamy Hills and Maleya Mahadeshwara Wildlife Sanctuary in Karnataka, with hamlets and villages (settlements) marked by red dots. Green indicates the protected forest area, while yellow-brown signifies the buffer zone of Chamarajanagar district. Spatial data for administrative boundaries (national, state and district levels) were obtained from the geoBoundaries database (https://www.geoboundaries.org),[44] and protected area boundaries were retrieved from the World Database on Protected Areas (https://www.protectedplanet.net).[45] Custom buffer zones were generated by the authors using QGIS software.[46]

an average prevalence of 13% for the sickle cell allele (ranging from 0% to 35%),[22] by assuming 95% CI, 80% power, an alpha error of 0.05 and the design effect of 1.8. The standard sample size formula for cross-sectional/cohort study design was applied using OpenEpi V.3.01.[23]

Participants' selection

We employed the Kish grid sampling method to ensure a representative selection of study participants within each household.[24] However, this approach was not consistently successful, as adult male participants often either refused to take part or were unavailable at the time of the visit, despite our efforts to follow-up with them for two to three visits. In such instances, we replaced the adult male participants with other household members, typically adult females. We excluded individuals who were severely ill or bedridden, those unable to participate in the study, and pregnant or lactating women.

Data collected

All questionnaires were administered by locally recruited data enumerators from the Soliga and local community members from the Chamarajanagar district, who were trained to collect data in the local dialect language. After identifying the household, all the adult participants aged 18 years and above were invited to participate in the study. Informed verbal consent was obtained from the head of the household and each participant. The verbal consent process was fully recorded in the data collection app from the initiation of the conversation till the obtaining of the consent. We have used the tools published in the THETA protocol to collect information on sociodemographic characteristics and indicators necessary to measure health inequities. Based on the type and frequency of physical activity, participants were classified into low, moderate and vigorous activity according to the International Physical Activity Questionnaire.[25 26] Additionally, anthropometric measurements, including body mass index (BMI), were calculated based on the National Programme for Prevention and Control of Cancer, Diabetes, Cardiovascular Diseases and Stroke guidelines for Indian communities.[27] Participants for the B R Hills sites were recruited from 2 December 2018 to 24h June 2019. For the M M Hills site, recruitment took place from 14 December 2019 to 20 March 2020.

Sample collection and analysis

Since this study was nested within the THETA study, we have collected informed consent to store the DNA samples and test for SCD. Approximately 2 mL of venous blood was collected from each participant at their village/settlement into tubes containing EDTA to prevent coagulation. Subsequently, haemoglobin concentration was measured on-site using the HemoCue Hb 301 analyser.[28] Samples were maintained at 4 °C during transportation from the collection site to the JSS Hospital laboratory, ensuring delivery within 24 hours. DNA was extracted using a non-enzymatic salting-out method and quantified with a Nanodrop spectrophotometer.[29] Extracted DNA samples were stored at −20°C until

further analysis. Thawed DNA samples underwent PCR amplification to detect the sickle cell mutation, employing the Amplification-Refractory Mutation System (ARMS) PCR technique at the Centre for Cellular & Molecular Biology Laboratory in Hyderabad, India. Samples were classified as 'trait' or 'diseased' based on band patterns observed during agarose gel electrophoresis. To validate our findings, 25% of the samples were re-evaluated using a point-of-care testing method approved by the Indian Council of Medical Research.[30]

Operational definition

Anaemia was defined according to WHO criteria: haemoglobin concentration <120 g/L for non-pregnant women and <130 g/L for men.[31]

Statistical analysis

Sociodemographic characteristics, wealth index, tobacco use, alcohol use, physical activity and nutritional status were summarised using proportions, 95% CIs and a χ^2 test was used to assess the differences in proportions between tribal and non-tribal groups. Multivariable logistic regression models were adjusted for age, caste, occupation, marital status, anaemia, site, SGDI, wealth index, homogeneity of the community, place of residence, tobacco use, alcohol use, activity and nutritional status and were used to explore the association with the presence of sickle cell allele. The variables that showed significance at 0.2 in the univariate analysis were included in the multivariable model. A p value of less than 0.05 was considered statistically significant. Analysis was conducted using R software V.4.1, released on 14 June 2024, along with the packages dplyr, table summary and ggplot2.[32] The sickle cell allele proportion was determined at the village level for geospatial analysis. A univariate local Moran's I test was conducted to examine spatial correlation, using 999 permutations, a significance threshold set at a p value of 0.05 and a contiguity based on spatial proximity weights matrix using GeoDa software.[33]

Study information was conveyed through a recorded video in the local dialect, and the participants received a written participation sheet that detailed the study's purpose, risks and benefits. After discussing the information sheet, oral consent was obtained, documented and securely stored in the Fulcrum offline data collection app.[34] The consent process was observed by either family members or community leaders.

While piloting the THETA study, we found that verbal consent was the most effective method. It was perceived as less intimidating than written consent and was more culturally accepted.[17] We informed the institutional ethics committees about the potential risks of the oral consent approach, particularly the risk of excluding certain participants if written consent was required. To keep the community informed and adhere to best research practices, we collaborated with the local community-based organisation Zilla Budakattu Girijana Abhivruddhi Sangha as gatekeepers. Implementing recorded verbal consent in the presence of a witness effectively secured informed consent.[35]

Table 1 Distribution of study sample across ST, non-ST and mixed villages

Nature of village	Number of villages	Number of villages SCD prevalence
Tribal	23 (39%)	17 (53%)
Non-tribal	14 (24%)	2 (6%)
Mixed	22 (37%)	13 (41%)
Total	59 (100%)	32 (100%)

SCD, sickle cell disease; ST, scheduled tribe.

RESULTS

The sociodemographic characteristics of both ST and non-ST populations in our study are described, followed by the prevalence of anaemia and the sickle cell allele in both populations. The association between selected sociodemographic characteristics and the sickle cell allele was explored. An analysis to examine the geo-spatial distribution of the sickle cell allele was also conducted, and the results are presented below.

Out of the 61 surveyed villages, participants with the sickle cell allele were identified in 34 villages. Of those with the sickle cell allele, 53.7% (n=43) resided exclusively in tribal villages, 42.5% (n=34) lived in mixed villages containing both tribal and non-tribal communities, and 3.8% (n=3) resided in non-tribal villages (see table 1).

Population characteristics

Due to the COVID-19 pandemic, data collection was halted during the second wave after achieving 88% of the proposed sample size (n=547/622). Of the 547 participants, 61% (n=336) were from B R Hills, representing 40 villages, with 58% (n=318) belonging to ST communities. In the M M Hills site, data were collected from 18 villages, where 64% (n=135) of participants were from ST communities. Overall, 59% of the sample comprised ST individuals.

Among the participants, 49% (n=268) were literate, including 47% (n=150) of ST individuals and 52% (n=118) of non-ST individuals. Women accounted for 69% (n=378) of total participants, with 59% (n=222) from ST communities and 41% (n=156) from non-ST communities. Regarding socioeconomic status, 21% of participants belonged to the lowest wealth quintile, of whom 65% were from ST communities. In contrast, 16% (n=87) of participants belonged to the highest wealth quintile, with 78% (n=68) from non-ST communities. Participants were relatively evenly distributed across the three remoteness gradients: 39% from remote areas, 31% from somewhat remote areas and 30% from non-remote areas. Among ST participants, the highest proportion was from remote villages (37%, n=118), followed by non-remote areas (36%, n=113). Among non-ST participants, 41% (n=95) were from remote areas, followed by 36% (n=82) from somewhat remote villages (see table 2).

Table 2 Key demographic characteristics of the study population

Characteristic	Overall N=547*	Non-ST N=229*	ST N=318*	P value†
PCR status				<0.001
Non-sickle	467 (85%)	209 (91%)	258 (81%)	
Sickle	80 (14.6%)	20 (8.7%)	60 (19%)	
Age				0.3
<30	185 (34%)	71 (31%)	114 (36%)	
30–40	172 (31%)	70 (31%)	102 (32%)	
>40	190 (35%)	88 (38%)	102 (32%)	
Anaemia				0.4
Anaemic	238 (44%)	95 (41%)	143 (45%)	
Non-anaemic	309 (56%)	134 (59%)	175 (55%)	
Sex				0.7
Female	378 (69%)	156 (68%)	222 (70%)	
Male	169 (31%)	73 (32%)	96 (30%)	
Marital status				0.2
Never married	44 (8.0%)	13 (5.7%)	31 (9.7%)	
Married	459 (84%)	198 (86%)	261 (82%)	
Divorced/separated/widowed	44 (8.0%)	18 (7.9%)	26 (8.2%)	
Literacy				0.3
Illiterate	279 (51%)	111 (48%)	168 (53%)	
Literate	268 (49%)	118 (52%)	150 (47%)	
Occupation				0.4
Earning	360 (66%)	146 (64%)	214 (67%)	
No income	187 (34%)	83 (36%)	104 (33%)	
Site				0.028
BR hills	336 (61%)	153 (67%)	183 (58%)	
MM hills	211 (39%)	76 (33%)	135 (42%)	
SGID				0.004
Non-remote	165 (30%)	52 (23%)	113 (36%)	
Remote	213 (39%)	95 (41%)	118 (37%)	
Somewhat remote	169 (31%)	82 (36%)	87 (27%)	
Wealth index				<0.001
Poorest	117 (21%)	41 (18%)	76 (24%)	
Poorer	119 (22%)	24 (10%)	95 (30%)	
Middle	121 (22%)	40 (17%)	81 (25%)	
Richer	103 (19%)	56 (24%)	47 (15%)	
Richest	87 (16%)	68 (30%)	19 (6.0%)	
Communities lived				<0.001
Mixed	296 (54%)	153 (67%)	143 (45%)	
Non-tribal	87 (16%)	75 (33%)	12 (3.8%)	
Tribal	164 (30%)	1 (0.4%)	163 (51%)	
Village located				<0.001
Inside	193 (35%)	56 (24%)	137 (43%)	
Outside	354 (65%)	173 (76%)	181 (57%)	
Smoking tobacco	99 (18%)	34 (15%)	65 (20%)	0.094
Passive smoking	228 (42%)	83 (36%)	145 (46%)	0.029

Continued

Table 2 Continued

Characteristic	Overall N=547*	Non-ST N=229*	ST N=318*	P value†
Smokeless tobacco	64 (12%)	17 (7.4%)	47 (15%)	0.008
Alcohol use	62 (11%)	24 (10%)	38 (12%)	0.6
Vigorous activity	328 (60%)	121 (53%)	207 (65%)	0.004
Moderate activity	513 (94%)	211 (92%)	302 (95%)	0.2
Walking/cycling	482 (88%)	191 (83%)	291 (92%)	0.004
BMI				<0.001
Underweight	235 (43%)	77 (34%)	158 (50%)	
Healthy weight	196 (36%)	76 (33%)	120 (38%)	
Overweight	49 (9.0%)	26 (11%)	23 (7.2%)	
Obesity	67 (12%)	50 (22%)	17 (5.3%)	

*n (%).
†Pearson's χ^2 test.
BMI, body mass index; BR Hills, Biligiri Ranganathaswamy Hills; MM Hills, Maleya Mahadeshwara Wildlife Sanctuary; SGDI, Socio-Geographical Disadvantage Index.

Among the participants, 18% (n=99) reported tobacco smoking. Stratification by (define ST) status revealed higher usage among ST participants at 20% (n=65; 95% CI (16% to 25%)) compared with non-ST participants at 15% (n=34; 95% CI (11% to 20%)). Secondhand smoke exposure (passive smoking) was reported by 42% (n=228; 95% CI (38% to 46%)) of participants, with a higher occurrence among ST individuals at 46% (n=145; 95% CI (40% to 51%)) compared with non-ST individuals at 36% (n=83; 95% CI (30% to 43%)). Alcohol use was comparable between ST and non-ST groups, reported by 12% (n=38; 95% CI (8.7% to 16%)) and 10% (n=24; 95% CI (7% to 15%)) of participants, respectively. Regarding physical activity, 65% (n=207; 95% CI (60% to 70%)) of tribal community members engaged in vigorous activity, compared with 53% (n=121; 95% CI (46% to 59%)) of non-tribal community members (see table 2 and online supplemental table 1).

Anaemia and sickle cell gene prevalence by age and demographic characteristics

The prevalence of the sickle cell allele in our study population was 14.6% (n=80/547), with 1% (n=5) confirmed to have SCD. Among the five confirmed SCD cases, four were from Adivasi communities. The proportion of individuals with sickle cell trait was 13% (n=75/547) (see table 2). The median age of individuals with sickle cell allele and healthy subjects was comparable at 32 years (IQR 25–45) and 35 years (IQR 27–42), respectively. The overall prevalence of anaemia was 44%, with a similar prevalence among ST individuals at 45% (95% CI (39% to 51%)) and non-ST individuals at 41% (95% CI (35% to 48%)). In B R Hills, the prevalence of anaemia was 54% (n=180; 95% CI (41% to 52%)), whereas in M M Hills, it was 27% (n=58; 95% CI (22% to 34%)) (see table 2 and online supplemental table 1).

Exploration of the association of key demographic and other attributes with sickle cell allele

Sex, marital status, earning status, literacy, tobacco smoking and alcohol use showed no statistically significant differences between tribal and non-tribal communities. In contrast, sickle haemoglobin (disease and trait), study site (B R Hills or M M Hills), SGDI, wealth index, engagement in vigorous activity and walking or cycling exhibited statistically significant differences between the two groups. No significant statistical differences were observed between tribal and non-tribal individuals carrying the sickle cell allele for age, marital status, literacy, income, wealth index, tobacco smoking, alcohol consumption or physical activity levels. However, among non-tribal participants, sex was significantly associated with the presence of the sickle cell mutation, with males exhibiting a lower prevalence at 15% (n=3/20) compared with females at 85% (n=17/20) (see table 3 and online supplemental table 1).

In multiple logistic regression analysis, adjustments were made for age category, anaemia, occupation, marital status, data collection site, SGDI, wealth index, community homogeneity, residence inside or outside the protected area, tobacco smoking, smokeless tobacco use, exposure to passive smoking, alcohol consumption, physical activity and nutritional BMI. Univariate analysis of the sickle cell allele showed associations with caste category, study site, SGDI, wealth index, place of residence, passive smoking and BMI. After adjusting for these factors, participants residing in exclusive tribal villages had 1.75 times the odds (95% CI (0.75 to 4.15)) of having the disease compared with those living in mixed communities. Additionally, participants exposed to passive smoking had 1.71 times the odds (95% CI (1.02 to 2.88)) of having the disease (see table 3). The Moran plot revealed no significant clustering pattern of these settlements, with a linear

Table 3 Exploration of the relationship between sickle cell allele prevalence and sociodemographic characteristics including remoteness, wealth index, substance use, physical activity and BMI using multivariate logistic regression

	Unadjusted analysis			Adjusted analysis		
Characteristic	**OR**	**95% CI**	**P value**	**OR**	**95% CI**	**P value**
Caste			**<0.001**			0.9
Non-ST	—	—		—	—	
ST	2.43	1.44 to 4.25		0.94	0.42 to 2.09	
Age			0.4			
<30	—	—				
30–40	0.70	0.39 to 1.26				
>40	0.76	0.43 to 1.33				
Anaemia			0.3			
Non-anaemic	—	—				
Anaemic	1.28	0.79 to 2.06				
Occupation			>0.9			
Earning	—	—				
No income	0.98	0.59 to 1.60				
Marital status			>0.9			
Never married	—	—				
Married	1.10	0.48 to 2.99				
Divorced/separated/widowed	1.00	0.29 to 3.46				
Literacy			0.6			
Illiterate	—	—				
Literate	0.88	0.54 to 1.41				
Site			**0.025**			0.15
BR Hills	—	—		—	—	
MM Hills	0.56	0.32 to 0.93		1.74	0.83 to 3.77	
SGDI			**0.005**			**0.016**
Non-remote	—	—		—	—	
Remote	1.18	0.69 to 2.04		0.91	0.49 to 1.69	
Somewhat remote	0.43	0.21 to 0.84		0.38	0.18 to 0.79	
Wealth			**0.007**			0.5
Poorest	—	—		—	—	
Poorer	1.44	0.75 to 2.81		1.52	0.75 to 3.14	
Middle	0.90	0.44 to 1.82		1.16	0.54 to 2.49	
Richer	0.75	0.34 to 1.58		1.09	0.46 to 2.51	
Richest	0.25	0.07 to 0.69		0.55	0.14 to 1.77	
Homogeneity of the community			<0.001			**0.038**
Mixed	—	—		—	—	
Non-tribal	0.28	0.07 to 0.79		0.31	0.07 to 0.99	
Tribal	2.74	1.67 to 4.53		1.75	0.75 to 4.15	
Place of residence			**0.029**			0.2
Inside the protected area	—	—		—	—	
Outside the protected area	0.58	0.36 to 0.94		0.69	0.38 to 1.25	
Smoking tobacco			0.9			
No	—	—				
Yes	1.05	0.55 to 1.89				

Continued

Table 3 Continued

Characteristic	Unadjusted analysis			Adjusted analysis		
	OR	95% CI	P value	OR	95% CI	P value
Passive smoking			**0.035**			**0.040**
No	—	—		—	—	
Yes	1.67	1.04 to 2.70		**1.71**	1.02 to 2.88	
Smokeless tobacco			0.8			
No	—	—				
Yes	1.09	0.50 to 2.16				
Alcohol use			0.7			
No	—	—				
Yes	0.85	0.36 to 1.77				
Vigorous activity			0.079			
No	—	—				
Yes	1.56	0.95 to 2.63				
Moderate activity			0.6			
No	—	—				
Yes	1.30	0.50 to 4.48				
Walking/cycling			0.3			
No	—	—				
Yes	1.48	0.69 to 3.67				
BMI			**0.010**			0.6
Underweight	—	—		—	—	
Healthy weight	0.62	0.36 to 1.04		0.78	0.44 to 1.36	
Overweight	0.70	0.27 to 1.58		1.09	0.39 to 2.72	
Obesity	0.20	0.05 to 0.57		0.49	0.11 to 1.64	

Bold values indicate P value < 0.05.
BMI, body mass index; BR Hills, Biligiri Ranganathaswamy Hills; MM Hills, Maleya Mahadeshwara Wildlife Sanctuary; SGDI, Socio-Geographical Disadvantage Index; ST, scheduled tribe.

fit showing a Local Moran's I value of 0.23 for the sickle cell allele (see figure 2). This suggests that the distribution of the sickle cell allele among study participants was uniform across the different sites.

DISCUSSION

India has been ranked second in terms of predicted SCD births after sub-Saharan Africa.[36] In India, the sickle mutation has historically been reported from isolated ST populations, with the allele frequency reaching up to 35% in some communities.[22] However, data from cohabitating ST and non-ST communities from rural Karnataka are scarce. We found a total prevalence of 14% for sickle cell allele among our study participants, the majority of them (96%) staying in tribal villages. This is comparable to prevalence among the ST population of another district (Mysuru) in Karnataka.[37] In ST communities with larger populations, recessive mutations such as SCD could dilute over time despite some degree of endogamy. However, in small endogamous communities such as the Soliga, this dilution may not occur and the prevalence of the sickle cell allele may be higher than in other larger populations. This could be partly explained by Nakatsuka *et al* who examined population-specific disease-associated genes in South Asia and highlighted that while marriages between close relatives are a significant driver of rare recessive diseases, founder events also play a substantial role in contributing to their prevalence in South Asian populations.[38] This highlights the importance of consanguinity as a factor that may continue driving the persistence of relatively higher frequency of sickle cell allele or other similar autosomal recessive genetic diseases with implications for the design of elimination programmes (see for instance the NSCAEM).

Under large population-based screening programmes in Chhattisgarh from 2007 to 2017, 1.72 million people were screened for SCD, with 10% found to have sickle cell allele and 0.47% homozygous. Despite tribal populations being the most affected, non-tribal groups like OBCs, SCs and general castes also showed significant prevalence in this study.[10] We also observed considerable non-ST prevalence of sickle cell allele in our study cohort

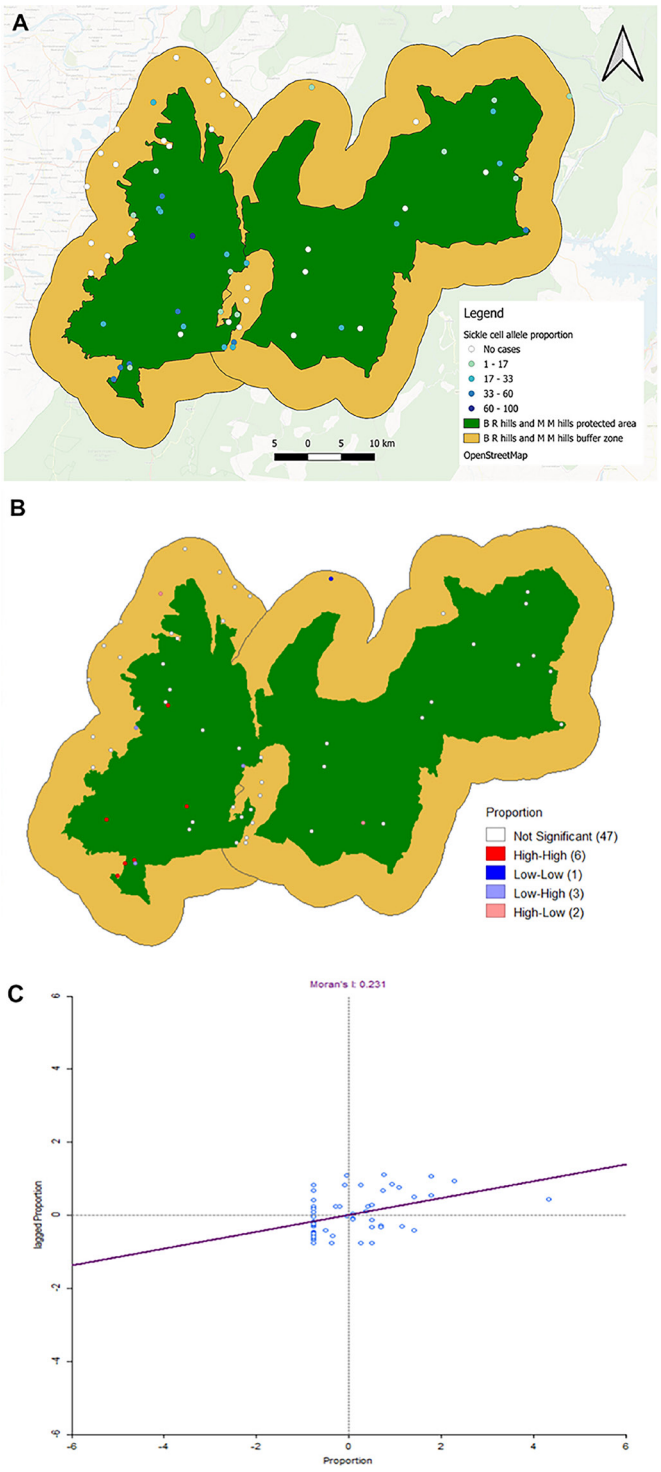

Figure 2 Map of the sickle cell allele distribution. (A) A map of the study area in Chamarajanagar district, Karnataka, illustrating the villages where sampling occurred, the forest boundary, and the gradient of sickle cell allele prevalence at the settlement level. (B) Univariate LISA maps depicting the clustering of hotspots and cold spots for sickle cell allele among settlements. (C) Moran scatter plot displaying the distribution of the sickle cell allele. Protected area boundaries were retrieved from the World Database on Protected Areas (https://www.protectedplanet.net).[45] Custom buffer zones were generated by the authors using QGIS software and LISA maps were created using GeoDa software.[33 46]

(8.7%, see table 2). This advocates for the universalisation of the current NSCAEM which has a focus on the ST population. While the focus on ST is well justified given the several decades of historical neglect of health problems in ST populations, the continuing identification of the allele as a 'tribal problem/disease' could hinder universal access to treatment for all populations as well as contribute to stigma. Policy-makers and implementers should consider universalising the screening and treatment along with a special focus in geographies and populations with higher prevalence. In addition, wherever specific sociocultural aspects that require adaptations in screening, treatment and care delivery are needed, these could be guided by contextually relevant research using participatory action research and other approaches that allow communities to engage with the design of targeted interventions.[16 39]

Fewer men are observed having sickle cell mutation among non-ST. This could be due to higher comorbidities observed among men compared with women due to health harm behaviours such as smoking, alcohol, etc as reported earlier from the same study population.[18] However, this could also be due to the sampling bias, as women were more represented than men in our household surveys. Additionally, a comparatively higher mortality rate among men with SCD could contribute to this, but this needs to be corroborated through longitudinal or population registry-based studies which are scarce in India.[40] In settings where more robust data are available (like from the USA), the average life expectancy of publicly insured patients living with SCD is 52.6 years (compared with 73.5 years for men and 79.3 years for women in the general population).[41 42]

Our study indicated that 44% (with 45% in ST and 41% in non-ST groups) experienced anaemia, which may not necessarily be due to nutritional factors. Nutritional anaemia, typically resulting from iron, folate or vitamin B_{12} deficiencies, is commonly addressed through iron-folate supplementation in primary health centres across India. However, iron deficiency accounts for only one-third of anaemia cases in the country, highlighting the need to move away from a one-size-fits-all strategy and to implement a more customised approach for tackling anaemia.[43] Treating undiagnosed SCD patients with iron supplements can be counterproductive, increasing the risk of iron overload and associated organ damage. Therefore, routine screening for SCD is crucial to differentiate anaemia types and provide appropriate, safe treatment strategies.

While undertaking an evidence-based appraisal of SCD from a health systems lens, Raman *et al* highlighted the importance of integrating various social disparities into SCD investigations.[12] These disparities, including economic background, sex, location and other susceptibilities, are currently underexplored in the literature, especially from among multiple remote, rural and ST communities. By focusing on these, this study contributes to developing a more comprehensive understanding of

SCD epidemiology in a remote, rural setting in southern Karnataka.

Our study has some limitations. The Kish method was partially implemented, resulting in an over-representation of females, even though the ratio in Chamarajanagar district is 993 females for every 1000 males.[21] The COVID-19 pandemic hindered the completion of surveys at M M Hills, which limited the ability to generalise the findings. Nevertheless, we integrated geospatial data from both B R Hills and M M Hills, even though these areas are administratively divided for forestry activities. Interestingly, we observed participants exposed to passive smoking had higher odds of having the sickle cell allele. However, as SCD is a genetic disorder, this association is likely incidental and reflects shared social or environmental factors rather than causation. Further research is needed to determine if this finding results from confounding variables or statistical variation. Further, the use of ARMS PCR while effective for identifying the presence or absence of the sickle allele does not distinguish between sickle cell trait and compound heterozygous states such as sickle-beta thalassaemia. As a result, individuals with sickle-beta thalassaemia may have been misclassified as having sickle cell trait, potentially leading to a slight overestimation of trait prevalence. However, this limitation does not significantly affect the study's primary aim of estimating the overall prevalence of the sickle cell gene.

Given the poorly performing health systems, especially so in ST and other socially disadvantaged populations, a much higher focus on general health systems strengthening in addition to focusing on universal screening and detection (which is the current focus of the NSCAEM programme) of SCD in India is required to eliminate this public health problem by 2040 as envisioned by Indian policy-makers.

Acknowledgements The authors thank Tanya Seshadri, Rahul Amruthapuri and Arun Mitra for their helpful insights that significantly influenced the manuscript. We thank the District Health Officer, the Department of Health and Family Welfare in Chamarajanagar, the Department of Tribal Welfare and the Chamarajanagar Forest Department for their operational support in data gathering and facilitating access to the hamlets. We also appreciate the contribution of various health workers within Chamarajanagar district, which was crucial in finalising this research. Special thanks to Zilla Budakattu Girijana Abhivruddhi Sangha, the organisation representing the Soliga people, without whom this initiative would not have been feasible.

Contributors YCB and PA contributed equally to the planning and execution of the work. LBR and GRC helped with data collection, PNS and DB secured funds, conceptualised and are guarantors and accept full responsibility for the work and/or the conduct of the study, had access to the data and controlled the decision to publish. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted. The corresponding author also requests equal 'first authorship' for YCB and PA. We have used ChatGPT to improve the language and grammar of about 15% text of the Introduction and Discussion sections, mainly to improve sentence readability.

Funding This work was supported by DBT/Wellcome Trust India Alliance fellowship to PNS (IA/CPHI/16/1/502648) (Supports to PNS and YCB) and Centre for Training, Research and Innovation in Tribal Health (CTRITH) IA/CRC/20/1/600007 awarded to PNS, Suresh Shapeti, Deepa Bhat and Upendra Bhojani (Supports to PNS, YCB and PA).

Disclaimer The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Map disclaimer The depiction of boundaries on this map does not imply the expression of any opinion whatsoever on the part of BMJ (or any member of its group) concerning the legal status of any country, territory, jurisdiction or area or of its authorities. This map is provided without any warranty of any kind, either express or implied.

Competing interests None declared.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants. The study received ethics permission from the Institutional Review Board of the Institute of Public Health, Bengaluru (Study ID IEC-FR/03/2018, IEC letter number IPH/18-19/E/226 and IEC IPH/19-20/E/183). For DNA analysis for sickle cell allele, ethics committee permission was obtained from JSS Medical College (JSSMC/IEC/05012022/02NCT/22021-22). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request. All data relevant to the study are included in the article or uploaded as supplementary information.

ORCID iDs
Yogish Channa Basappa http://orcid.org/0000-0002-5045-5121
Pooja Aggarwal http://orcid.org/0009-0009-4851-2949
Deepa Bhat http://orcid.org/0000-0002-4928-1718

REFERENCES

1 Piel FB, Steinberg MH, Rees DC. Sickle Cell Disease. *N Engl J Med* 2017;376:1561–73.
2 Osei MA, McGann PT. Sickle cell disease: time to act on the most neglected global health problem. *Lancet Haematol* 2023;10:e558–9.
3 Saha N, Banerjee B. Haemoglobinopathies in the Indian subcontinent. A review of literature. *Acta Genet Med Gemellol (Roma)* 1973;22:117–38.
4 Xaxa V. Tribes as Indigenous People of India. *Econ Polit Wkly* 1999;34:3589–95.
5 Sharma RK, Thakor HG, Saha KB, *et al*. Malaria situation in India with special reference to tribal areas. *Indian J Med Res* 2015;141:537–45.
6 Dixit S, Sahu P, Kar SK, *et al*. Identification of the hot-spot areas for sickle cell disease using cord blood screening at a district hospital: an Indian perspective. *J Community Genet* 2015;6:383–7.
7 Ghosh K, Colah RB, Mukherjee MB. Haemoglobinopathies in tribal populations of India. *Indian J Med Res* 2015;141:505–8.
8 Colah R, Mukherjee M, Ghosh K. Sickle cell disease in India. *Curr Opin Hematol* 2014;21:215–23.
9 Hockham C, Bhatt S, Colah R, *et al*. The spatial epidemiology of sickle-cell anaemia in India. *Sci Rep* 2018;8:17685.
10 Lad H, Chandak G, Patra P. P-066: BURDEN AND DISTRIBUTION OF SICKLE CELL ANEMIA IN TRIBAL AND NON-TRIBAL POPULATION OF STATE CHHATTISGARH, INDIA. *Hemasphere* 2022;6(Suppl):48.

11 Rao P, Raj EA, Natesan S, *et al*. Prevalence of Sickle cell disease, Sickle cell trait and HBS-beta-thalassemia in India: A systematic review and Meta-analysis. *Clin Epidemiol Glob Health* 2024;28:101678.
12 Raman V, Seshadri T, Joice SV, *et al*. Sickle cell disease in India: a scoping review from a health systems perspective to identify an agenda for research and action. *BMJ Glob Health* 2021;6:e004322.
13 Piel FB, Colah R, Jain DL. Casting light on the national mission to eliminate sickle cell disease in India. *Hemasphere* 2024;8:e70033.
14 Rai ND, Benjaminsen TA, Krishnan S, *et al*. Political ecology of tiger conservation in India: Adverse effects of banning customary practices in a protected area. *Singap J Trop Geogr* 2019;40:124–39.
15 Rai ND, Devy MS, Ganesh T, *et al*. Beyond fortress conservation: The long-term integration of natural and social science research for an inclusive conservation practice in India. *Biol Conserv* 2021;254:108888.
16 Cáceres ÁL, Ramesh RM, Newmai P, *et al*. Perceptions, health seeking behavior and utilization of maternal and newborn health services among an indigenous tribal community in Northeast India-a community-based mixed methods study. *Front Public Health* 2023;11:1139334.
17 Srinivas PN, Seshadri T, Velho N, *et al*. Towards Health Equity and Transformative Action on tribal health (THETA) study to describe, explain and act on tribal health inequities in India: A health systems research study protocol. *Wellcome Open Res* 2019;4:202.
18 Mallikarjuna Majgi S, Channa Basappa Y, Belagihalli Manjegowda S, *et al*. Prevalence of dyslipidemia, hypertension and diabetes among tribal and rural population in a south Indian forested region. *PLOS Glob Public Health* 2024;4:e0002807.
19 Government of Karnataka. Human development: performance of districts, taluks and urban local bodies in Karnataka, 2014-snapshot. 2014.
20 Roy S, Hegde HV, Bhattacharya D, *et al*. Tribes in Karnataka: Status of health research. *Indian J Med Res* 2015;141:673–87.
21 India - census of India 2001 - series 12 - Karnataka - district census handbook, part XII. Available: https://censusindia.gov.in/nada/index.php/catalog/43804 [Accessed 06 Mar 2025].
22 Serjeant GR, Ghosh K, Patel J. Sickle cell disease in India: A perspective. *Indian J Med Res* 2016;143:21–4.
23 Dean AG, Sullivan K. OpenEpi: open source epidemiologic statistics for public health. 2019. Available: http://www.openepi.com/Menu/OE_Menu.htm [Accessed 14 Aug 2019].
24 Kish L. A Procedure for Objective Respondent Selection within the Household. *J Am Stat Assoc* 1949;44:380.
25 Sjostrom M, Ainsworth BE, Bauman A, *et al*. Guidelines for data processing analysis of the international physical activity questionnaire (IPAQ) - short and long forms. 2005. Available: https://api.semanticscholar.org/CorpusID:79242415
26 National Institute of Medical Statistics and Division of Non-Communicable Diseases, Indian Council of Medical Researchcc. Integrated disease surveillance project (IDSP), non-communicable disease risk factors survey, phase-I states of India, 2007-08. 2009. Available: https://main.icmr.nic.in/sites/default/files/reports/Phase-1_States_of_India.pdf
27 Department of health and family welfare. National programme for prevention and control of cancer, diabetes, cardiovascular diseases and stroke (NPCDCS). New Delhi Department of health and family welfare, Government of India; 2013. Available: http://www.dhsodisha.nic.in/?q=node/264 [accessed 16 Mar 2019]
28 HemoCue AB. HemoCue® Hb 301 system. Ängelholm, Sweden HemoCue AB; 2024. Available: https://hemocue.com/in/products/hematology/hemocue-hb-301-system/
29 Suguna SA, Nandal DH, Kamble SU, *et al*. Genomic DNA isolation from human whole blood samples by non enzymatic salting out method. *Int J Pharm Pharm Sci* 2014;6:198–9.
30 Kanter J, Telen MJ, Hoppe C, *et al*. Validation of a novel point of care testing device for sickle cell disease. *BMC Med* 2015;13:225.
31 World Health Organization. Haemoglobin concentrations for the diagnosis of anaemia and assessment of severity. Geneva WHO; 2011. Available: https://www.who.int/publications/i/item/WHO-NMH-NHD-MNM-11.1 [accessed 31 Jul 2025]
32 R Core Team. R: a language and environment for statistical computing. Vienna, Austria R Foundation for Statistical Computing; 2024. Available: https://www.R-project.org/
33 Anselin L. Local spatial autocorrelation (1). LISA and Local Moran; 2020. Available: https://geodacenter.github.io/workbook/6a_local_auto/lab6a.html [Accessed 17 Jan 2025].
34 Spatial Networks, Inc. Fulcrum: mobile data collection and workflow platform. San Francisco (CA) Spatial Networks, Inc. c2024–25. Available: https://www.fulcrumapp.complay.google.com+13 [accessed 31 Jul 2025]
35 N Srinivas P, Seshadri T, Velho N, *et al*. Response to correspondence article on the research protocol titled *Towards Health Equity and Transformative Action on tribal health (THETA) studyto describe, explain and act on tribal health inequities in India: A health systems research study protocol*. *Wellcome Open Res* 2023;8:155.
36 Piel FB, Patil AP, Howes RE, *et al*. Global epidemiology of sickle haemoglobin in neonates: a contemporary geostatistical model-based map and population estimates. *The Lancet* 2013;381:142–51.
37 Babu BV, Sharma Y, Sridevi P, *et al*. Feasibility of population-based screening of sickle cell disease through the primary health care system in tribal areas of India. *J Med Screen* 2023;30:28–35.
38 Nakatsuka N, Moorjani P, Rai N, *et al*. The promise of discovering population-specific disease-associated genes in South Asia. *Nat Genet* 2017;49:1403–7.
39 Pratt B, Seshadri T, Srinivas PN. Overcoming structural barriers to sharing power with communities in global health research priority-setting: Lessons from the Participation for Local Action project in Karnataka, India. *Glob Public Health* 2022;17:3334–52.
40 Khan H, Krull M, Hankins JS, *et al*. Sickle cell disease and social determinants of health: A scoping review. *Pediatr Blood Cancer* 2023;70:e30089.
41 Platt OS, Brambilla DJ, Rosse WF, *et al*. Mortality in sickle cell disease. Life expectancy and risk factors for early death. *N Engl J Med* 1994;330:1639–44.
42 Ballas SK, Pulte ED, Lobo C, *et al*. Case series of octogenarians with sickle cell disease. *Blood* 2016;128:2367–9.
43 Ghosh S, Laxmaiah A, Chandak GR, *et al*. Anaemia and iron deficiency in India: a venous blood-based survey of adolescents, adults, and the elderly in eight states. *Eur J Clin Nutr* 2025;79:443–51.
44 Runfola D, Anderson A, Baier H, *et al*. geoBoundaries: A global database of political administrative boundaries. *PLoS One* 2020;15:e0231866.
45 UNEP-WCMC and IUCN. Protected planet: the world database on protected areas (WDPA). Cambridge, UK UNEP-WCMC; 2024. Available: https://www.protectedplanet.net/ [accessed 31 Jul 2025]
46 QGIS Development Team. QGIS geographic information system. Open source geospatial foundation project. 2024. Available: https://qgis.org [Accessed 31 Jul 2025].