Prevalence and disease progression of genetically-confirmed facioscapulohumeral muscular dystrophy type 1 (FSHD1) in China between 2001 and 2020: a nationwide population-based study

Zhiqiang Wang,[a,#] Liangliang Qiu,[a,#] Minting Lin,[a,#] Long Chen,[a] Fuze Zheng,[a] Lin Lin,[a] Feng Lin,[a] Zhixian Ye,[a] Xiaodan Lin,[a] Junjie He,[a] Lili Wang,[a] Xin Lin,[a] Qifang He,[a] Wanjin Chen,[a] Yi Lin,[a,] Ying Fu,[a,*] and Ning Wang[a,*]*

[a]Department of Neurology and Institute of Neurology of First Affiliated Hospital, Institute of Neuroscience, and Fujian Key Laboratory of Molecular Neurology, Fujian Medical University, Fuzhou 350005, China

Summary

Background Facioscapulohumeral muscular dystrophy type 1 (FSHD1) is a rare disease, which is often underdiagnosed due to its heterogeneous presentations and complex molecular genetic basis, leading to a lack of population-based epidemiology data, especially of prevalence and disease progression.

Methods Fujian Neuromedical Centre (FNMC) is a diagnosis centre for clinical-genetic FSHD in China, and the only one employing pulsed-field gel electrophoresis (PFGE)-based Southern blotting for all FSHD1 genetic tests. Three sources distributed across all six spatial zones in China, were used to obtain information regarding FSHD1 events, namely, FNMC, Genetic and Myopathy Group (branches of the Neurology Society of the Chinese Medical Association), and "FSHD-China" (an organization supported by FSHD patients). During 2001-2020, all genetically-confirmed FSHD1 from China were registered in FNMC. Follow-up was conducted in the 20-year period to obtain data on disease progression, which was mainly described in terms of independent ambulation loss.

Findings Of the 1,744 FSHD1 genetic tests (total test number 1,802) included in the analysis, 997 (57.2%) patients from 620 families were diagnosed with FSHD1. The estimated prevalence of genetically-confirmed FSHD1 in China is 0.75 per million (95% confidence interval [CI], 0.70-0.79) during 2001-2020, with 0.78 (95% CI, 0.72-0.85) in males and 0.71 (95% CI, 0.65-0.78) in females. The estimated prevalence increased from 0.22 (95% CI, 0.19-0.26) per million in 2001-2015 to 0.53 (95% CI, 0.49-0.57) per million in 2016-2020 ($p < 0.001$). The prevalence in Fujian province was 7.10 per million, 4.66 per million, and 2.44 per million, during 2001-2020, 2001-2015, and 2016-2020, respectively. Among the 861 symptomatic plus asymptomatic patients of the total 997 patients, the median onset age at first-ever muscle weakness was 16 years of age (range 1-81); the median number of contracted D4Z4 repeats was 5 units (range 1-9); the median 4qA-allele-specific methylation level was 41% (range 14%-69%). Of the 977 symptomatic patients followed-up during 2001-2020, 117 patients (12.0%) lost independent ambulation. The expected duration from onset of first-ever muscle weakness to onset of independent ambulation loss was 40 years. The group with loss of independent ambulation had a smaller number of contracted D4Z4 repeats ($p < 0.001$) and had an earlier onset age of first-ever muscle weakness ($p < 0.001$) compared to the group without loss of independent ambulation.

Interpretation Our research captures the largest genetically-confirmed FSHD1 population worldwide, to calculate its prevalence of 0.75 per million in China from 2001 to 2020. Approximately 12.0% of symptomatic plus asymptomatic patients of FSHD1 will lose independent ambulation in 40 years from onset of first-ever muscle weakness.

*Corresponding authors:
    E-mail addresses: linyi7811@163.com (Y. Lin), fuying@fjmu.edu.cn (Y. Fu), ningwang@fjmu.edu.cn (N. Wang).
# Contributed equally.

Research in context

Evidence before this study

We searched PubMed for articles published before February 25, 2020 on prevalence and disease progression of genetically-confirmed facioscapulohumeral muscular dystrophy type 1 (FSHD1). We used the search string "FSHD AND Prevalence AND Disease progression". The resulting reports were manually reviewed. Only one study reported population-based estimated prevalence of 120 per million, however, this study used a capture-recapture model from three registries, and two of the three registries only diagnosed by clinical features. Our extensive search did not identify any nationwide population-based study of FSHD1.

Added value of this study

Our research captures the largest genetically-confirmed FSHD1 population worldwide, with 997 patients from 620 families. In this study, we estimated a lower-bound but true prevalence of 0.75 per million during 2001-2020. The estimated prevalence increased from 0.22 per million in 2001-2015 to 0.53 per million in 2016-2020. The prevalence in Fujian province was relatively higher at 7.10 per million, 4.66 per million, and 2.44 per million, during 2001-2020, 2001-2015, and 2016-2020, respectively. Of 977 symptomatic patients that were followed-up on during 2001-2020, 117 patients (12.0%) lost independent ambulation. The expected duration from onset of first-ever muscle weakness to onset independent ambulation loss was 40 years. Clinical and genetic features of FSHD1 are also provided.

Implications of all the available evidence

This is the first population-based study to estimate the prevalence within all age groups of genetically confirmed FSHD1 in China. It showed that approximately 12.0% of FSHD1 patients will lose independent ambulation within 40 years of first-ever muscle weakness. This study helps in making important medical and financial decisions concerning FSHD1.

Introduction

Facioscapulohumeral muscular dystrophy type 1 (FSHD1) is one of rare diseases and is one of the most common inherited muscular dystrophies with autosomal dominance.[1] The core phenotype of FSHD1 is characterized by progressive weakness of the facial, scapular, and humeral muscles, with later involvement of the trunk and lower extremities, [2] resulting in disability with loss of independent ambulation or even wheelchair dependence. FSHD1 has a unifying genetic model caused by a contraction to 1-10 D4Z4 macrosatellite repeats on a 4qA-specific FSHD1-permissive haplotype, [3−5] which was sensitively identified by pulsed-field gel electrophoresis (PFGE)-based Southern blotting.[6,7]

In the pre-genomic era (before 1991),[8,9] several studies reported estimated prevalence of FSHD1 that were based on clinical diagnosis, with an equivocal prevalence ranging from 0.21 to 6.80 per 100,000 worldwide. In the post-genomic era, five studies regarding the prevalence of genetically-confirmed FSHD1 were performed, and only one study reported population-based estimate prevalence.[10−12] This population-based study used a capture-recapture methodology from three registries within a 10-year period (419 patients), which reported that the estimated prevalence of FSHD1 in the Netherlands was 12 per 100,000.[10,13] However, two of the three registries were based on clinical diagnosis alone (no genetic test), which might have influenced the above-mentioned results. In Asia, Japan and Israel also have reported FSHD1 prevalence through regional-based estimates.[14,15]

Due to its rare and heterogeneous clinical presentations, as well as complex molecular genetic basis, large-scale data reporting on the epidemiology of FSHD1 is of limited availability.[16] China is a large-scale state accounting for 21.6% of Asia's land area and 18.8% of the world's population (approximately 1.4 billion). Fujian Neuromedical Centre (FNMC) is a diagnosis centre for clinical-genetic FSHD in China, and the only one employing PFGE-based Southern blotting. Over a period of 20 years (2001-2020), this centre has consecutively and non-selectively registered patients/families that have been genetically-confirmed via PFGE-based Southern blotting.[17,18] We aim to provide large-scale patient epidemiologic data from this registry, particularly in terms of prevalence and disease progression.

Methods

Anonymized raw data are available upon request from the corresponding author.

Study population

This was a population-based study of the prevalence and disease progression of genetically-confirmed FSHD1 between 2001 and 2020 in the Chinese population. The study population consisted of all genetically-confirmed FSHD1 patients or families (with at least 1 affected member) in China, irrespective of age. These FSHD1 patients presented at least one contracted D4Z4 repeats with 4qA-specific FSHD1-permissive haplotype,[3-5,19] the diagnosis of which was performed at FNMC, the clinical genetic test hospital for FSHD1 in China to employ PFGE-based Southern blotting.[5,17-20]

The total population was 1,332,810,869 persons distributed across China according to the 2010 Chinese Census (http://www.stats.gov.cn/tjsj/pcsj/rkpc/6rp/indexch.htm). Three sources were used to obtain

information regarding FSHD1 events. The first, was FNMC. This was the main resource because it is the only clinical-genetic centre employing PFGE-based Southern blotting for FSHD1 in China. All the inpatients and outpatients of suspected clinical FSHD1 were examined by the same neurologist (Z.Q.W.) and clinical assessments were performed by one trained team, to ensure standardization and reliability of data. Second, more than 50 hospitals are networked within Genetic and Myopathy Group of the Neurology Society of the Chinese Medical Association (https://www.cma.org.cn/), which are distributed throughout all six of the spatial zones in China: north, northeast, east, south, southwest, and west, comprising 31 provinces and municipalities, according to the China Statistical Yearbook (http://www.stats.gov.cn/). Third, a patient advocacy and support organization named "FSHD-China" (http://www.fshd-china.org/), which is a member of the International FSHD Patient Advocacy Organizations. The latter two sources reported initial FSHD1 information to our centre; the neurologist (Z.Q.W.) then contacted patients suspected to have FSHD1 through telephone or remote video for clinical evaluation and assessments; finally, the clinical suspected patient mailed a blood sample to FNMC for genetic test. Detailed information of the latter two sources can be seen in **Appendix Figure 1** and **Appendix Table 1**.

All available data pertaining to demographics and clinical and genetic features of FSHD1 were collected, and yearly follow-ups were made through outpatient service or telephone survey or remote video conferences conducted by the same neurologist (Z.Q.W.).

During follow-up, we defined independent ambulation loss based on the modified Rankin Scale (mRS), a simple 6-point assessment that included reference to limitation in activity, with a grade at 4-5 that individuals were unable to walk without assistance.[21] An inherent advantage of this study is its feature of having been hospital-based which ensured "Quality Assurance" for all genetically confirmed FSHD1 patients. In line with WHO guidelines, this study defined an adult as a person older than 19 years of age, and a child is a person 19 years or younger. This study was approved by the Ethics Committee for Medical Research of the First Affiliated Hospital of Fujian Medical University, and informed consent was obtained from all participants.

Case ascertainment

First, suspected clinical FSHD1 patients must have a distinctive and regional distribution of muscle weakness of the facial, scapular, and humeral muscles (with complaints and indications). Then, genetic tests of PFGE-based Southern blotting were performed for these clinical suspected FSHD1 patients on the basis of the family (if available) as a whole. Finally, eligible participants were genetically confirmed patients who presented a contraction to 1-10 D4Z4 repeats with a 4qA-specific FSHD1-permissive haplotype.

Laboratory protocols for PFGE-based Southern blotting were utilized as previously described.[17,18] Briefly, human genomic DNA was gently extracted from peripheral blood samples, and was embedded in agarose plugs. Five micrograms of DNA were digested with enzymes of *Eco*RI/*Hind*III (or *Eco*RI) and *Eco*RI/*Hind*III/*Bln*I (or *Eco*RI/*Bln*I) for array sizing; for analysis of the allelic variation on 4qter, DNA was digested with *Hind*III only (*Eco*RI and *Hind*III were purchased from New England Biolabs, and *Bln*I was purchased from Takara Bio). DNA was separated by PFGE. After blotting to an Amersham Hybond-N+ Membrane (GE Healthcare), the DNA was hybridized with probes p13E-11 (D4F104S1), 4qA (9B6A), and 4qB. After exposure, the *Eco*RI/*Hind*III (or *Eco*RI) array sizes were measured according to the MidRange PFG Marker (New England Biolabs) or CHEF DNA Size Standard (Bio-Rad Laboratories). The arrays were assigned to their respective chromosomes based on their *Bln*I sensitivity. The D4Z4 repeats number was calculated from the *Eco*RI fragment size using the formula: (D4Z4-containing fragment length (in kb, following *Eco*RI/*Hind*III)−4.7)/3.3 or (D4Z4-containing fragment length (in kb, following *Eco*RI digestion)−6.8)/3.3.[18] Additionally, we used the 4qA-allele-specific FasPAS primer to investigate CpGs hypomethylation in a region distal to the DRA with sodium bisulphite sequencing (conducted by Genesky Biotechnologies, Shanghai, China).[18,22]

Clinical assessments

Clinical outcome measures were structured according to our previous study.[18] Clinical phenotypes were assessed by the 2016 Comprehensive Clinical Evaluation Form (CCEF) for FSHD.[23] Muscle strength was assessed bilaterally by manual muscle testing in 14 muscle pairs (deltoid, triceps, biceps, wrist extensor, wrist flexor, finger extensor, finger flexor, gluteus maximus, gluteus medius, iliopsoas, hamstring, quadriceps femoris, tibialis anterior, and gastrocnemius muscle). We calculated the Medical Research Council (MRC) scale sum score by summing all of the average scores of each pair of muscles.[24] Muscle strength was also evaluated by FSHD clinical score.[25] Both the clinical severity score (CSS) and the age-corrected clinical severity score (ACSS) were adopted to determine disease severity; the ACSS was adjusted for the patient's age at diagnosis: $((CSS \times 2)/\text{age at diagnosis}) \times 1000$.[26,27]

Statistical analysis

The awareness of FSHD1 in China is relatively low, but has slowly gained momentum since 2016; hence, FSHD1 prevalence was analysed over two separate time periods—2001-2015 and 2016-2020—from 2001 to

2020. Prevalence was calculated separately using the number of FSHD1 mutation carriers who were genetically-confirmed in 2001-2015, 2016-2020, and 2001-2020 as the numerator. The denominator was always set to the 2010 population census for China. We analysed the trends in prevalence of FSHD1 by comparing the two study periods (2001-2015 and 2016-2020) overall, because of the lower percentage of genetic diagnoses in the first study period. In addition, prevalence was stratified by sex, spatial zones, and provinces and municipalities per million at risk according to the 2010 population census of China, which provides precise and in-depth information on the population of China.

The proportion of patients who suffered independent ambulation loss was defined as the total number of cases in which independent ambulation was lost during 2001-2020 as the numerator, divided by the total followed-up FSHD1 cases as the denominator. Duration of independent ambulation loss was derived by deduct the onset age at first-ever muscle weakness from the onset age at onset of independent ambulation loss. The Kolmogorov-Smirnov tests or Shapiro-Wilk's test was used to confirm that most of the variables in the study were not normally distributed. Thus, continuous variables were presented as median (range) and categorical data as number (proportion). 95% confidence intervals (CI) were calculated using the Wald formula (95% CI = p ± $1.96\sqrt{\frac{p(1-p)}{n}}$) given the large sample size in this study. The Mann-Whitney U test was used to compare numerical variables; the χ^2 test was used to compare categorical variables; and correlations were measured by the Spearman rank correlation coefficient. Cumulative disability curves were estimated using the Kaplan-Meier method to assess disease progression with time span as loss of independent ambulation since first-ever muscle weakness, and the log-rank test was used to compare differences between males and females. All the statistical analyses were performed using SPSS (version 25; IBM, USA). A p value < 0.05 was considered statistically significant.

Role of the funding source
The funding sources of this study did not influence nor participate in study design, data collection, data analysis, data interpretation, or drafting of the report. The corresponding author had full access to all the data in the study as well as final responsibility for the decision to submit for publication.

Results

Nationwide genetically-confirmed FSHD1 families
Nationwide, a total of 620 families were identified with FSHD1 mutation carriers by PFGE-based Southern blotting between January 2001 and December 2020.

Regionally, 60 (9.7%) were located in the north, 30 (4.8%) were located in the northeast, 364 (58.7%) were located in the east, 104 (16.8%) were located in the south, 36 (5.8%) were located in the southwest, and 21 (3.4%) were located in the west (the information from five [0.8%] families were missing) (Figure 1). The number of FSHD1 families in the context of geographic distribution varied from 0 in Tibet and Qinghai (both of which are located in west China) to 176 (28.4%) in Fujian (located in east China; the location of genetic test centre). Shandong and Zhejiang (both hosting more than 50 families) were also high-identification locations along with Fujian. Approximately 71.6% were located in non-Fujian areas. Ningxia, Hainan, Xinjiang, Guangxi, Tianjin, and Shanghai were low-identification locations (< 5 families).

The proportion of genetically confirmed FSHD1 families increased from 32.6% (202 families) in 2001-2015 to 67.4% (418 families) in 2016-2020. The proportion of genetically confirmed FSHD1 families located in non-Fujian areas increased from 38.1% (77 families) in 2001-2015 to 87.8% (367 families) in 2016-2020 (χ^2 [1, N=620] = 165.3, p < 0.001). Identified FSHD1 families were mainly confined to Fujian (61.9%) during 2001-2015, however, the geographic distribution of identified FSHD1 families during 2016-2020 was similar to that during 2001-2020 (Table 1).

Prevalence of genetically-confirmed FSHD1
A total of 1,802 FSHD1 genetic tests were ordered during 2001-2020. Of these, 58 (3.2%) tests were excluded due to incomplete testing that prevented a definitive diagnostic interpretation. This resulted in a sample size of 1,744 tests: 997 cases (57.2%) met the criteria for a genetic diagnosis of FSHD1, five cases (0.3%) met the criteria for a genetic diagnosis of FSHD2 (permissive 4qA haplotype and hypomethylation and *SMCHD1* variant [c.1 A>G]), and 742 cases (42.5%) failed to meet criteria for FSHD1 or FSHD2 (FSHD1-, 2-negative). Of these 742 negative cases, 591 cases were parents and/or relatives of genetically confirmed FSHD1.

The overall prevalence of genetically-confirmed FSHD1 in China was estimated to be 0.75 (95% confidence interval [CI], 0.70-0.79) per million in 2001-2020. A male predominance was observed, where the male to female ratio was 1.15, with an estimated prevalence of 0.78 (95% CI, 0.72-0.85) per million in males and 0.71 (95% CI, 0.65-0.78) in females (Table 1). Of the 997 positive genetic tests, 111 (11.1%) were *de novo* pathogenic alleles with an estimated prevalence of 0.08 (95% CI, 0.07-0.10) per million; 326 (32.7%) were inherited pathogenic alleles with an estimated prevalence of 0.24 (95% CI, 0.22-0.27) per million; and the remaining 560 (56.2%) had unknown results for parental genetic tests. Generally, the estimated prevalence was high in east China (prevalence, 1.41 [95% CI, 1.29-1.53]). In contrast, southwest China (prevalence,
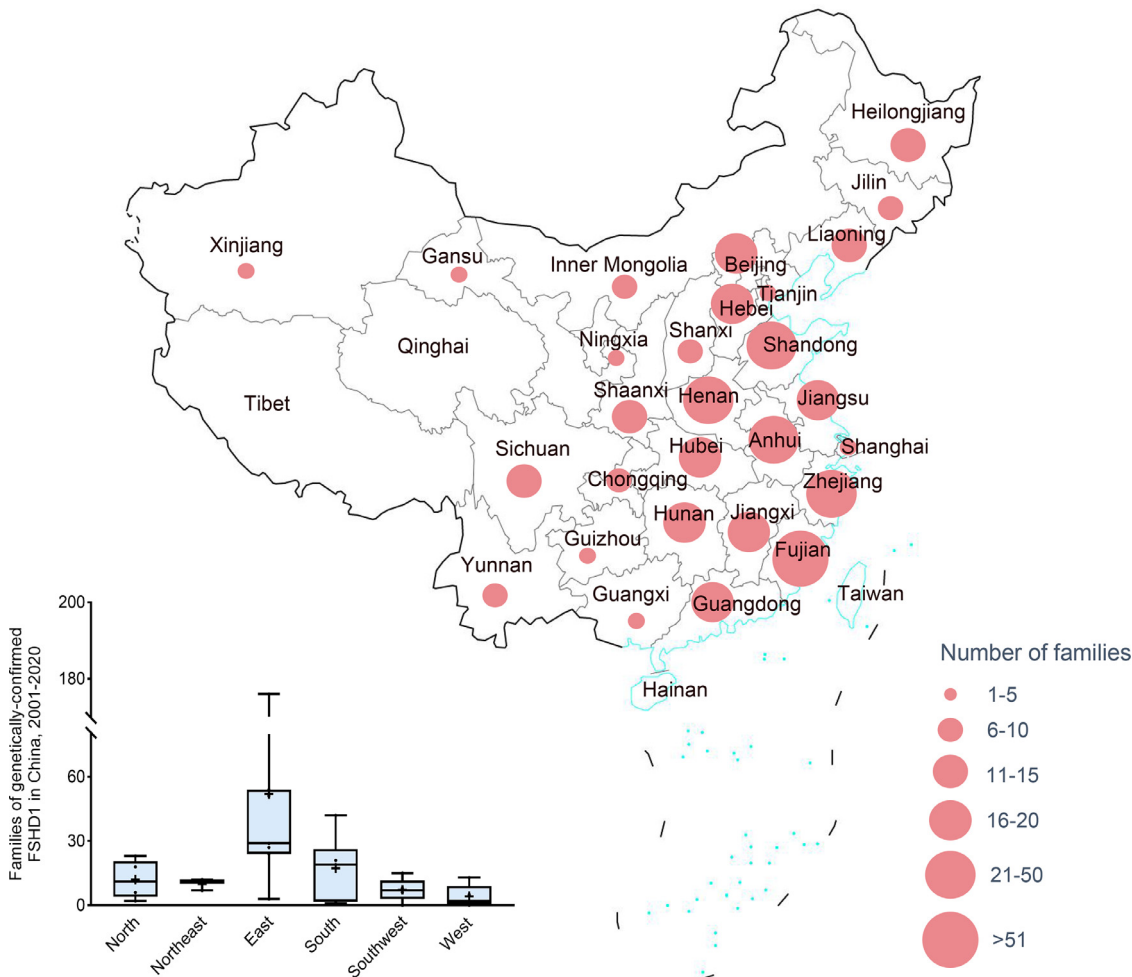
Figure 1. Distribution of families with genetically-confirmed FSHD1 in China, 2001-2020.

	The 2010 population census of China, n	2001-2015		2016-2020		2001-2020	
		Genetically-confirmed cases, n (%)	Prevalence rate, per million (95% CI)	Genetically-confirmed cases, n (%)	Prevalence rate, per million (95% CI)	Genetically-confirmed cases, n (%)	Prevalence rate, per million (95% CI)
Male	682329104 (51.2)	151 (52.6)	0.22 (0.19-0.26)	383 (53.9)	0.56 (0.51-0.62)	534 (53.6)	0.78 (0.72-0.85)
Female	650481765 (48.8)	136 (47.4)	0.21 (0.17-0.24)	327 (46.1)	0.50 (0.45-0.56)	463 (46.4)	0.71 (0.65-0.78)
Total	1,33,28,10,869	287	0.22 (0.19-0.24)	710	0.53 (0.49-0.57)	997	0.75 (0.70-0.79)

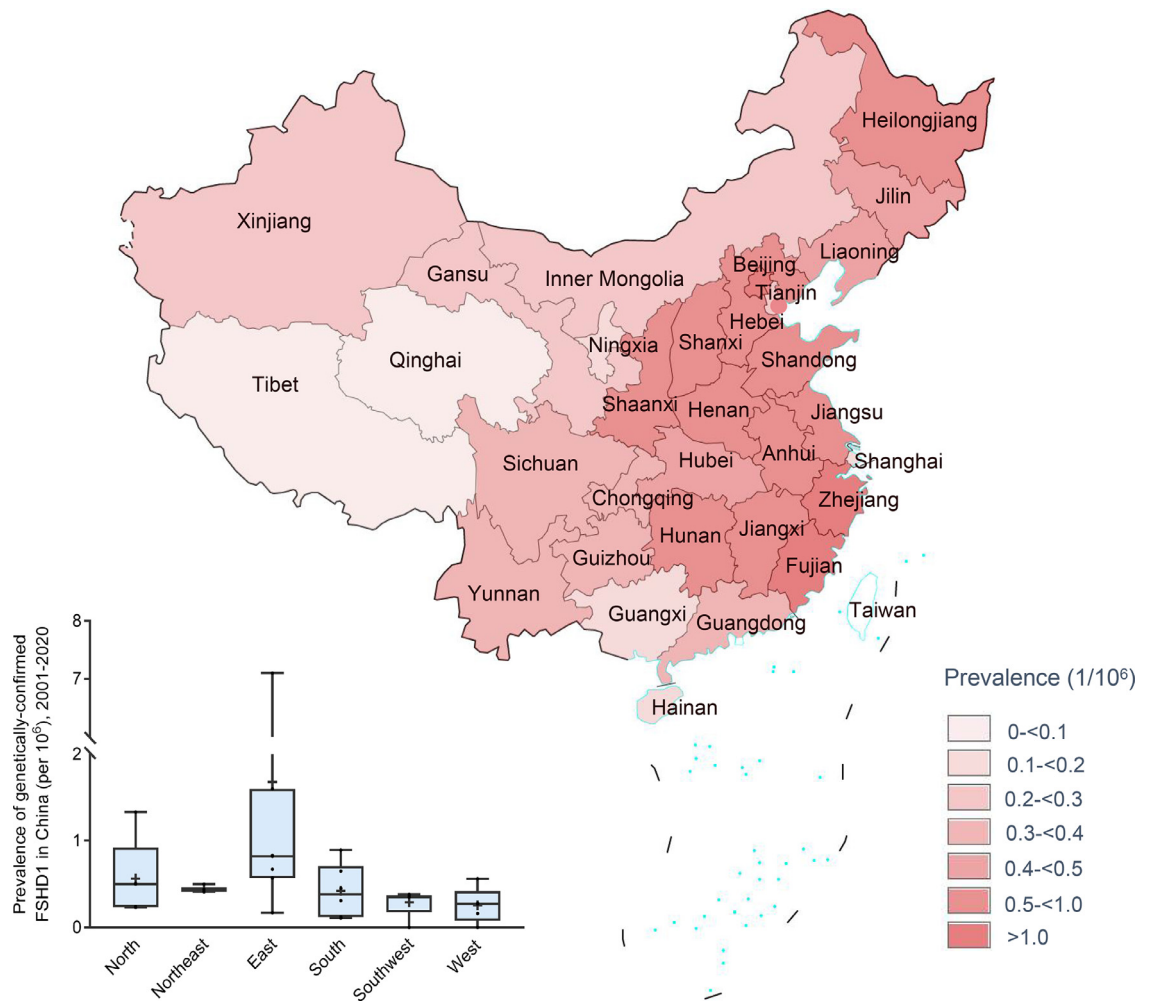*Table 1*: **Prevalence of genetically-confirmed FSHD1 in China.**
Abbreviation: FSHD1 = Facioscapulohumeral muscular dystrophy type 1.

0.35 [95% CI, 0.27-0.44]) and west China (prevalence, 0.36 [95% CI, 0.24-0.48]) had low estimated prevalence. A positive correlation ($\rho = 0.89$, $p = 0.03$) was found between number of hospitals and number of genetically-confirmed FSHD1 cases among different spatial zones (**Appendix Figure 2**). The estimated crude prevalence in the context of geographic distribution varied from 0 in Tibet and Qinghai (both are located in west

China) to 7.10 (95% CI, 6.24-7.96) in Fujian (located in east China). Beijing (located in north China; prevalence, 1.33 [0.82-1.84]) and Zhejiang (located in east China; prevalence, 1.60 [1.26-1.93]) were also high-prevalence gathering place next to Fujian (**Figure 2**; **Appendix Table 2**).

In 2001-2015, the estimated prevalence of genetically-confirmed FSHD1 in China was limited, with an

Figure 2. Prevalence of genetically-confirmed FSHD1 in China, 2001−2020.

overall prevalence of 0.22 (95% CI, 0.19-0.24) per million. Fujian province was a high prevalence province with an estimated prevalence of 4.66 (95% CI, 3.97-5.36) per million. Zhejiang, Beijing, Anhui, Jiangxi, Henan, Hubei, Liaoning, and Shaanxi were low provinces with estimated FSHD1 prevalence between 0.1 and 1.0 per million. The remaining provinces and municipalities had very low prevalence values of between 0 and 0.1 per million. In 2016-2020, the overall estimated prevalence of genetically-confirmed FSHD1 in China was 0.53 (95% CI, 0.49-0.57) per million. We observed significantly increasing prevalence in areas outside of Fujian, ranging from 0.09 (95% CI, 0.07-0.10) per million in 2001-2015 to 0.48 (95% CI, 0.44-0.52) per million in 2016-2020 (χ^2 [1, N=997] =235.6; p < 0.001); the latter increased to 0.57 (95% CI, 0.53-0.61) per million in the period from 2001-2020 (χ^2 [1, N=1284] = 112.8; p = < 0.001). The geographic distribution of the estimated prevalence in 2016-2020 was similar to that in 2001-2020, which also varied from 0 in

Tibet and Qinghai to 2.44 (95% CI, 1.94-2.94) in Fujian. Beijing (prevalence, 1.07 per million; 95% CI, 0.61-1.53) and Zhejiang (prevalence, 1.27 per million; 95% CI, 0.97-1.57) also had high-prevalence along with Fujian province (**Appendix Figure 3-4; Appendix Table 2**).

Baseline characteristics

Genetic analyses identified 997 FSHD1 mutation carriers during 2001-2020, the median age was 36 years old (range 3-87). Of these mutation carriers, 861 (86.4%) were symptomatic plus asymptomatic patients (asymptomatic means mutation carriers present minor signs without motor impairment); 124 (12.4%) were nonpenetrant FSHD1 mutation carriers with completely normal neurologic examinations (23 [18.7%] in 1-3 D4Z4 units; 71 [57.7%] in 4-6 D4Z4 units; 29 [23.6%] in 7-10 D4Z4 units);[23] and 12 (1.2%) did not have related data. Of the 861 symptomatic plus asymptomatic

patients, 526 (61.1%) were paediatric patients (0-19 years old) with first-ever muscle weakness at a median onset age of 14 years old, of which 193 (36.7%) carried 1-3 units, 297 (56.5%) carried 4-6 units, and 36 (6.8%) carried 7-10 units; the distribution of paediatric patients in the different provinces and municipalities is presented in **Appendix Table 2**. 310 (36.0%) were adult patients (> 19 years old) with a median onset age of 25 years old (range 20-81). Twenty-five (2.9%) had an unknown age of onset. Potential nonpenetrant FSHD1 mutation carriers were found in all age classes up to 69 years old. None of the 124 nonpenetrant FSHD1 mutation carriers presented symptoms during follow-ups in 2001-2020. Although 53 (42.7%) were still under the age of 30, the other 71 (57.3%) were over the age of 30, which may be exhibiting real nonpenetrance.[28]

Among the 655 participants for whom we acquired data about CCEF classification, 458 (69.9%) were of category A, with typical penetrance of both facial and upper limb muscle weakness (49 A1, 351 A2, 58 A3); 43 (6.6%) were of category B (31 B1, 12 B2); 136 (20.8%) were of category C (12 C1, 124 C2; the 124 C2 were the same as the 124 nonpenetrant FSHD1 mutation carriers described above); and 18 (2.7%) were D (18 D1, 0 D2). In particular, the 31 patients who were categorized as B1 subtype presented with muscle weakness limited to the scapular girdle, namely, facial sparing FSHD (SHD; 21 of the 31 SHD were reported in our previous study).[20] We also obtained other clinical data for subsets of our large scale FSHD1 cohort. Among the 548 participants for whom we acquired data for the MRC sum score, the median was 117.8 points ranging from 59.9 to 140.0. We obtained FSHD CS data for 510 participants, and the median was 7 points ranging from 0 to 13. Finally, for 518 participants, the median CSS and ACSS scores were 3 points (range 0-5) and 181.9 points (range 0-1285.7), respectively. Notably, all 124 nonpenetrant FSHD1 mutation carriers got an MRC sum score, FSHD CS score, and CSS score at 140.0 points, 0 points, and 0 points, respectively.

Data showed that the median length of contracted D4Z4 repeats of the 997 genetically confirmed FSHD1 mutation carriers was 23.5kB (range 10.0-36.5), which was 5 units (range 1-9) after formula conversion, of which 267 (26.8%) carried 1-3 units, 594 (59.6%) carried 4-6 units, and 136 (13.6%) carried 7-10 units. The median number of contracted D4Z4 units was statistically lower in the 526 paediatric patients in comparison with that in the 310 adult patients (median 4 vs 5; Mann-Whitney U, 52761; $p < 0.001$). Somatic mosaicism of the D4Z4 repeats contraction was found in 40 (4.0%) of the FSHD1 mutation carriers. 4qA-allele-specific methylation was assessed in 690 FSHD1 mutation carriers, with a median of 41% (range 14% - 69%). The median of 4qA-allele-specific methylation level was significantly lower than in healthy controls (median 41% vs

63%; Mann-Whitney U, 2267; $p < 0.001$). Compared to the adult FSHD1 patients (n = 310), paediatric FSHD1 patients (n = 526) had increased 4qA-allele-specific hypomethylation (median 38% vs 43%; Mann-Whitney U, 27112; $p < 0.001$). Symptomatic plus asymptomatic FSHD1 mutation carriers also had increased 4qA-allele-specific hypomethylation compared with nonpenetrant FSHD1 mutation carriers (median 40% vs 46%; Mann-Whitney U, 17130, $p < 0.001$) (Table 2).

Independent ambulation loss in genetically-confirmed FSHD1

977 (98.0%) of the 997 patients with genetically-confirmed FSHD1 were followed-up in the period of 2001-2020. Among the 977 patients, 117 (12.0%) lost independent ambulation (44 males; 73 females). Of the 251 follow-up cases of older patients (age > 50 years), 51 (20.3%) lost independent ambulation, and 34 (13.5%) were wheelchair dependent.

The median onset age of loss of independent ambulation for all 117 cases was 38 years old (range 9-87). Of those, 81 (69.2%) were paediatric patients, 32 (27.4%) were adult patients, and 4 (3.4%) were missing data about onset age of first-ever muscle weakness. The median time from first-ever muscle weakness to onset of independent ambulation loss was 20 years (range 0-69). 87 (8.9%) of the 977 patients were wheelchair dependent and two died of unknown causes (Table 2). Cumulative disability curves without covariate adjustment showed that FSHD1 patients will be loss independent ambulation about 40 years later since first-ever muscle weakness (Figure 3). No significant differences were observed between males and females (Log-rank (Mantel-Cox) test, χ^2 [1, N=117] = 0.036, $p > 0.05$).

In order to investigate risk factors related to disease progression (described with independent ambulation loss), we divided the 977 follow-up patients into two groups: the independent ambulation loss group (117 cases) and the non-loss of independent ambulation group (860 cases). The former group had an earlier onset age of first-ever muscle weakness (15 versus 17 years old; Mann-Whitney U, 31388; $p < 0.001$), a smaller number of calculated contracted D4Z4 repeats (4.0 versus 5.0 units; Mann-Whitney U, 38791; $p < 0.001$), and a higher proportion of females (62.4% versus 44.7%; χ^2 [1, N=977] =13.0; $p < 0.001$). Furthermore, we found that the period of time from first-ever muscle weakness to the onset of independent ambulation loss was both slightly positively correlated with the onset age of first-ever muscle weakness ($\rho = 0.22$, $p = 0.02$), and with a lower number of calculated contracted D4Z4 repeats ($\rho = 0.31$, $p = 0.001$).

Discussion

A large clinical sample of patients from across China who underwent test for FSHD in our centre provides

	Total = 997
Baseline characteristics	
Demographic	
Sex, M/F, n (%)	534 (53.4)/463 (46.4)
Age, y, median (range)	36 (3-87)
Clinical	
Onset age at first-ever muscle weakness (patient-reported; symptomatic plus asymptomatic patients, n = 842), y, median (range)	16 (1-81)
Symptomatic plus asymptomatic mutation carriers, n (%)	861 (86.4%)
Nonpenetrant mutation carriers, n (%)	124 (12.4%)
Phenotypic classification (CCEF) (evaluated participants = 655)	
Category A	458 (69.9%)
Category B	43 (6.6%)
Category C	136 (20.8%)
Category D	18 (2.7%)
Assessments (symptomatic plus asymptomatic patients), median (range)	
MRC sum score (0−140) (evaluated participants = 548)	117.8 (59.5-140.0)
FSHD clinical score (0−15) (evaluated participants = 510)	7 (0-13)
Clinical severity scale (0−5) (evaluated participants = 518)	3 (0-5)
Age-corrected clinical severity score (0−10 000)	181.9 (0-1285.7)
Genetic/epigenetic	
Length of contracted D4Z4 repeat array, kB, median (range)	23.5 (10.0-36.5)
Size of contracted D4Z4 repeats, units, median (rang)	5 (1-9)
1-3 D4Z4 repeats, n (%)	267 (26.8)
4-6 D4Z5 repeats, n (%)	594 (59.6)
7-9 D4Z6 repeats, n (%)	136 (13.6)
Mosaic mutation, n (%)	40 (4.0)
D4Z4 Methylation level (tested participants = 690), percentage, median (range)	41 (14-69)
Disease progression	
Proportion of independent ambulation loss (n/followed-up cases)	12.0 (117/977)
Onset age at onset of independent ambulation loss, y, median (range)	38 (9-87)
Duration[a], mean (SD), y, median (range)	20 (0-69)
Proportion of wheelchair dependency (n/followed-up cases)	8.9 (87/977)

*Table 2*: Characteristics of genetically-confirmed FSHD1 in China, 2001 to 2020.

[a] Duration was derived by deduct the onset age at first-ever muscle weakness from the onset age at onset of independent ambulation loss.

unique insights into the genetic epidemiology of this complex disease. It is the first population-based study to estimate the prevalence for all age groups of genetically confirmed FSHD1 in China (approximately 1.4 billion population). With a 20-year follow up (2001-2020), the longest longitudinal genetic study for FSHD1, disease progression concerning independent ambulation loss is also reported in this study. As clinical studies on FSHD interventions can be expected in the near future, accurate data on FSHD epidemiology are needed for trial readiness. So, this study not only fills a void in epidemiologic data and enriches the global outlook for this rare disease of FSHD1, but also provides urgent data for clinical trial readiness.

In this largest cohort consisting of symptomatic patients plus asymptomatic mutation carriers and nonpenetrant mutation carriers, we found a lower-bound estimated prevalence of 0.70 per million during 2001-2020. It is lower than the data has been reported (2.1 and 120 per million).[10,11] The data of prevalence in the Netherlands was based on capture-recapture model from three registries, and two of the three registries only diagnosis by clinical features. Based on 419 FSHD patients diagnosed within a 10-year period, the prevalence rate was 120 per million. Through mapping FSHD1 prevalence in 31 provinces and municipalities in China, we found notably higher prevalence of FSHD1 in three provinces: Fujian (genetic test centre at First Affiliated Hospital), Zhejiang (a historical hospital of genetic research), and Beijing (the capital, which has the largest number of hospitals from Genetic and Myopathy Group of the Neurology Society of the Chinese Medical Association). It was similar to FSHD1 research status in the Netherlands with higher prevalence rates in Leiden (genetic test centre), Amsterdam (the capital), and somewhat higher in Nijmegen (a historical region of genetic research).[13] These suggested that medical resources (such as trained professionals and equipment) play a role in disease detection.
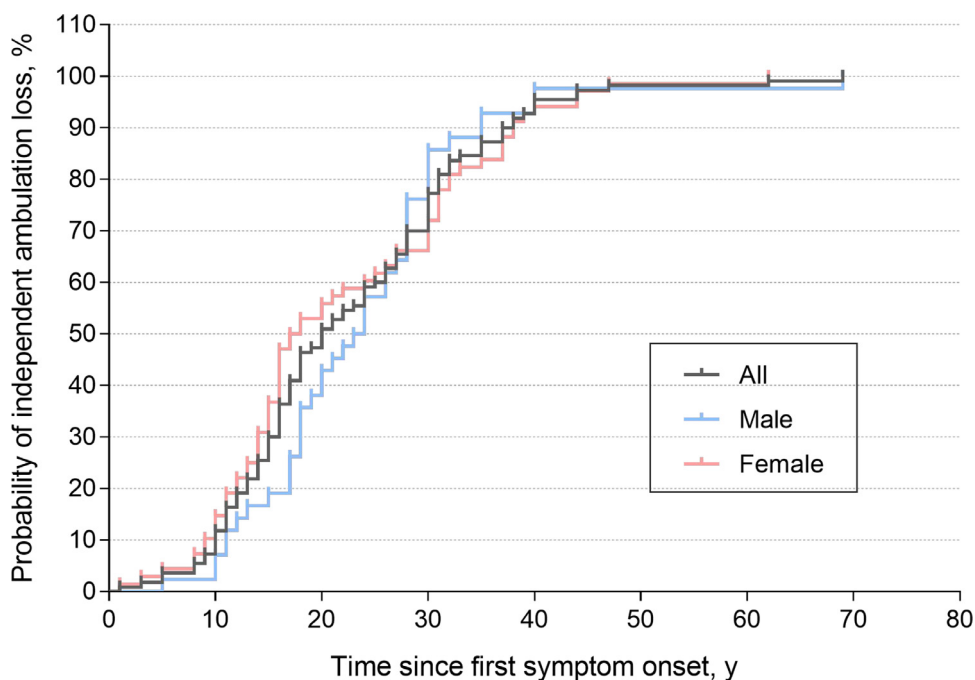
Figure 3. Prediction of independent ambulation loss among patients with genetically-confirmed FSHD1. Curves are unadjusted for covariates. End point was defined as losing independent ambulation based on the modified Rankin Scale (mRS) with a grade of 4-5.

Of the three provinces of Fujian, Zhejiang, and Beijing with obviously higher prevalence rates of genetically confirmed FSHD1 in our study. All the three provinces were located in economically eastern China among the three zones: eastern, central, and western.[29] The highest was in Fujian province (with a population of 36,894,217 persons) in 7.10 per million, 4.66 per million, and 2.44 per million, during 2001-2020, 2001-2015, and 2016-2020, respectively. All three prevalence data from Fujian were lower than the Netherlands prevalence (120 per million). This is probably explained by the fact that Fujian has been the single clinical genetic test centre for FSHD1, which can avoid registering overlap data. In addition, capture-recapture methodology is based on several assumptions and maybe lead to overestimation of the Netherlands prevalence. It's also likely that the prevalence of China is truly lower than the Netherlands. However, the prevalence rate of Fujian was comparable to that found in several other regions around the world relative true. The prevalence in Fujian province during 2001-2020 was like Italian data (6.1 per million),[30] and was higher than that in several reports from Asian countries, such as Shimane, Japan (3.9 per million),[31] Kumamoto district, Japan (3.0 per million),[32] Israel (5.3 per million).[15]

The prevalence of FSHD1 in many non-Fujian areas was likely underestimated, especially those areas in the central economic zone of China (prevalence, 0.66 per million) and that of western China (prevalence, 0.32 per million), comparing to economically eastern China (prevalence, 1.08 per million). Encouragingly, the reported prevalence of FSHD1 in areas outside of Fujian has been slowly gaining momentum over the last five years, rising from 0.09 per million in 2001-2015 to 0.48 per million in 2016-2020. This was largely due to active promotion by the government and various social forces, such as the the Chinese Medical Association and the patient advocacy groups (such as "FSHD-China") since 2016. Hence, the training of more professional physicians, establishing additional and higher quality medical facilities, raising the proportion of total health expenditure by expanding basic medical insurance, and strengthening awareness of FSHD1—especially in areas beyond Fujian, Zhejiang, and Beijing provinces—may all be contributing to the increased diagnostic rate of FSHD1.

Southern blotting analysis of restriction enzyme digests separated by pulsed field gel electrophoresis (PFGE) was a traditional method for FSHD1 genetic test. However, this method needs high level of technical capability and quality of personnel, which limited epidemic screening for FSHD1. Other testing methods in our lab have been explored to measure truncated 4q alleles in recent years, e.g., molecular combing [33] and BioNano single-molecule optical mapping [34]. Based on PFGE-based Southern blotting, this study identified a total of 997 patients (620 families) from 1,744 genetic tests, reaching a rate of genetic diagnosis of FSHD1 of 57.2%. 747 FSHD1-negative cases were mostly (591 cases) from family members. Five cases were confirmed to have FSHD2. In the past, a largest clinical sample of patients from across the United States who underwent testing for FSHD with same method, which

reported that of the 1,594 patients with FSHD tests, 664 (41.7%) were diagnosed with FSHD1.[35] The proportion of genetic diagnoses was slightly higher in this study, which could have been caused by the increased family recruitment. Our genetic data summarized in Table 2 by size of contracted D4Z4 repeats was significantly different from the data collected at the Iowa Molecular Pathology Laboratory: 1-3 D4Z4 repeats (China 26.8% versus Iowa 16.4%), 4-6 D4Z4 repeats (China 59.6% versus Iowa 48. 2%), and 7-10 repeats (China 13.6% versus Iowa 35.4%).[35]

The size of the contracted D4Z4 repeats in both our paediatric patients (mean 3.6 units) and adult cohort (mean 4.7 units) from China tended to be smaller than that reported in the Dutch child cohort (mean 5.2 units) and adult cohort (mean 5.8 units).[36] D4Z4 PAS-specific methylation in Chinese population was similar to that reported in a 2016 Dutch study (mean 41% in our population versus about 40% in the 2016 Dutch study [22]).

Additionally, whereas the frequency of nonpenetrant FSHD mutation carriers tended to be higher in a recent family-based study (17%),[28] our population showed a lower frequency (12%). Regarding ethnicity-specific trends, the age at onset tended to be earlier in China (median 16 years; mean 18 years) compared to European countries, such as the Netherlands (median 20 years) and Italy (mean 33 years); [37]. However, the age of onset is older in China compared to what has been reported in Korea (median 13 years).[38] An international, collaborative and registered study with unified protocols will be needed to further clarify these findings and better understand if there are ethnicity differences that contribute to clinical presentation and functional outcomes.

In the present study, we also firstly reported the proportion of independent ambulation loss for genetically-confirmed FSHD1, which was estimated at 12.0% in the 20-year follow-up. The simple 6-point mRS assessment was used to measure independent ambulation for our FSHD1 patients because of its satisfactory reliability, only explicit criterion of walking, and idiosyncratic criteria to raters.[21] We found that symptomatic plus asymptomatic patients were highly likely to lose independent ambulation within 40 years of their first disease-related experience of muscle weakness. Fifty-three nonpenetrant FSHD1 mutation carriers were in pre-clinical stage before the age of 30. The proportion of wheelchair-dependent FSHD1 at 13.5%, among those patients over 50 years old, which is lower than the previously reported 20%.[21,39,40] During 2001-2020, only two patients died, and those deaths were due to unknown causes. This is consistent with previous opinion that FSHD1 typically progresses relatively slowly and generally does not affect mortality.[1,39]

Several studies have indicated that both the size of the contracted D4Z4 repeats and age at first muscle weakness contribute to the severity of FSHD1.[6,28,41] Consistently, we observed that both had an impact on the rate of disease progression. The short size of D4Z4 repeats and the young age at onset of muscle weakness respectively shorten experienced loss of independent ambulation. In addition, we observed slight but significant inverse correlations ($p < 0.05$) between the number of contracted D4Z4 repeats and the span of time from first muscle weakness to first loss of independent ambulation, as well as between the age of first muscle weakness and the span of time until the first loss of independent ambulation. We also found that gender had an impact on disease progression, in that the incidence of loss of independent ambulation was more pronounced in females than males ($p < 0.001$) in this study.

The main limitation of this study is the biased population. At follow-up, a limited chart review was conducted as well as a telephone survey and remote video conferences by the same neurologists. This limited our centre's ability to systematically collect and analyse all data about clinical assessments. More, we did not systematically assess or collect data about quality of life (e.g., psychological well-being, marital relations and home life, financial resources, social support acceptance, etc.) in patients who lost independent ambulation or who were wheelchair dependent. To access accurate epidemiological data for improving prevention and therapy in the future, it is necessary to strengthen the management strategy for this disease. First, train medical professionals in the area of FSHD1 and increase genetic test equipment for strengthening and accelerating accurate diagnosis; second, construct a Nationwide Mandatory Registry System for FSHD1 to gather nationwide epidemiological data in a up-to-date way and to collect natural history data systematically; third, popularize the knowledge and improve the social awareness of FSHD1 by such means as regularly holding FSHD1 medical seminars, patient exchange meetings, and disease publicity activities.

We calculated a lower-bound prevalence underestimate of 0.70 per million in China form the largest study of FSHD1 patients to date. Our findings show that considerable numbers of affected individuals remain undiagnosed in many non-Fujian areas of china. The management strategy for this disease need strengthen. Additionally, our study came to the conclusion that approximately 12.0% of FSHD1 patients lose independent ambulation within 40 years of the first experience of muscle weakness, which is a precious resources for clinical trial readies. This study also indicated that the contracted D4Z4 repeats contribute to the disease progress, consistent with a previous report showing that D4Z4 acts as a transcriptional repressor organizing long distance interactions between the telomeres to inhibit toxic protein expression resulting in muscle weakness.[42] Hence, further researches concerning the locus of the 4q telomere at the nuclear envelope and the long distance interactions of D4Z4 is needed to to slow progression of this disease.[43]

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation	✓(N/A)
Title and abstract			
	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	✓
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	✓
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	✓
Objectives	3	State specific objectives, including any prespecified hypotheses	✓
Methods			
Study design	4	Present key elements of study design early in the paper	✓
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	✓
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	✓
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	N/A
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	✓
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	✓
Bias	9	Describe any efforts to address potential sources of bias	✓
Study size	10	Explain how the study size was arrived at	✓
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	✓
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	✓
		(*b*) Describe any methods used to examine subgroups and interactions	N/A
		(*c*) Explain how missing data were addressed	✓
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	✓
		(*e*) Describe any sensitivity analyses	N/A
Results			a (N/A)
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	✓
		(b) Give reasons for non-participation at each stage	✓
		(c) Use a flow diagram and include the figure number (preferably figure 1) or page number	(N/A)
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	✓
		(b) Indicate number of participants with missing data for each variable of interest	✓
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	✓
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	✓
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	N/A

(continued)

	Item No	Recommendation	√(N/A)
		(Continued)	
		Cross-sectional study—Report numbers of outcome events or summary measures	N/A
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	√
		(*b*) Report category boundaries when continuous variables were categorized	√
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	N/A
Discussion			
Key results	18	Summarise key results with reference to study objectives	√
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	√
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	√
Generalisability	21	Discuss the generalisability (external validity) of the study results	√
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	√

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

*N/A stands for not applicable and may be a reasonable choice depending on the type of study performed

Contributors
NW, YF, and ZQW formulated the study concept, acquired funding and supervised execution of the study. YF and ZQW designed the study. ZQW, LLQ, LC, FZZ, FL, ZXY, XDL, JJH and LLW performed Southern blotting analysis. LLQ, MTL, LL, XL, QFH, WJC, and YL collected data. YF and LLQ wrote the manuscript. NW critically revised and gave final approval for publication of the paper.

Data sharing statement
A fully anonymized version of the dataset used for analysis with individual participant data and a data dictionary will be available for other researchers to apply to use, via https://168.2.5.202:9110. Written proposals will be assessed by corresponding author of this study (ningwang@fjmu.edu.cn) and a decision made about the appropriateness of the use of data.

Fujian medical university school of public health have received consultant fees as statistician to perform the data analysis.

References
1 Wang LH, Tawil R. Facioscapulohumeral Dystrophy. *Curr Neurol Neurosci Rep* 2016;**16**(7):66.

2 Tawil R, Van Der Maarel SM. Facioscapulohumeral muscular dystrophy. *Muscle Nerve* 2006;**34**(1):1–15.

3 Lemmers RJ, et al. Facioscapulohumeral muscular dystrophy is uniquely associated with one of the two variants of the 4q subtelomere. *Nat Genet* 2002;**32**(2):235–6.

4 Lemmers RJ, et al. A unifying genetic model for facioscapulohumeral muscular dystrophy. *Science* 2010;**329**(5999):1650–3.

5 Wang ZQ, et al. Distinguishing the 4qA and 4qB variants is essential for the diagnosis of facioscapulohumeral muscular dystrophy in the Chinese population. *Eur J Hum Genet* 2011;**19**(1):64–9.

6 Tawil R, et al. Evidence-based guideline summary: Evaluation, diagnosis, and management of facioscapulohumeral muscular dystrophy: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology and the Practice Issues Review Panel of the American Association of Neuromuscular & Electrodiagnostic Medicine. *Neurology* 2015;**85**(4):357–64.

7 DeSimone AM, et al. Facioscapulohumeral Muscular Dystrophy. *Compr Physiol* 2017;**7**(4):1229–79.

8 Emery AE. Population frequencies of inherited neuromuscular diseases—a world survey. *Neuromuscul Disord* 1991;**1**(1):19–29.

9 Deenen JC, et al. The Epidemiology of Neuromuscular Disorders: A Comprehensive Overview of the Literature. *J Neuromuscul Dis* 2015;**2**(1):73–85.

10 Deenen JC, et al. Population-based incidence and prevalence of facioscapulohumeral dystrophy. *Neurology* 2014;**83**(12):1056–9.

11 Thompson MW, Ludvigsen B, Monckton G. Some problems in genetics of muscular dystrophy. *Rev Can Biol* 1962;**21**:543–50.

12 Flanigan KM, et al. Genetic characterization of a large, historically significant Utah kindred with facioscapulohumeral dystrophy. *Neuromuscul Disord* 2001;**11**(6-7):525–9.

13 van Engelen BG, et al. The Dutch neuromuscular database CRAMP (Computer Registry of All Myopathies and Polyneuropathies): development and preliminary data. *Neuromuscul Disord* 2007;**17**(1):33–7.

14 Nakagawa M, et al. Epidemiology of progressive muscular dystrophy in Okinawa, Japan. Classification with molecular biological techniques. *Neuroepidemiology* 1991;**10**(4):185–91.

15 Kott E, et al. Muscular dystrophy: the relative frequency in the different ethnic groups in Israel. *Confin Neurol* 1973;**35**(3):177–85.

16 Rath A, et al. A systematic literature review of evidence-based clinical practice for rare diseases: what are the perceived and real barriers for improving the evidence and how can they be overcome? 2017;**18**(1):556.

17 Wu ZY, et al. FSHD in Chinese population: characteristics of translocation and genotype-phenotype correlation. *Neurology* 2004;**63**(3):581–3.

18 Qiu L, et al. *Clinical and genetic features of somatic mosaicism in facioscapulohumeral dystrophy* 2020;**57**(11):777–85.

19 Lin F, et al. New Insights into Genotype-phenotype Correlations in Chinese Facioscapulohumeral Muscular Dystrophy: A Retrospective Analysis of 178 Patients. *Chin Med J (Engl)* 2015;**128**(13):1707–13.

20 He JJ, et al. Clinical and genetic features of patients with facial-sparing facioscapulohumeral muscular dystrophy. *Eur J Neurol* 2018;**25**(2):356–64.

21 Wilson JT, et al. Improving the assessment of outcomes in stroke: use of a structured interview to assign grades on the modified Rankin Scale. *Stroke* 2002;**33**(9):2243–6.

22 Calandra P, et al. Allele-specific DNA hypomethylation characterises FSHD1 and FSHD2. *J Med Genet* 2016;**53**(5):348–55.

23 Ricci G, et al. A novel clinical tool to classify facioscapulohumeral muscular dystrophy phenotypes. *J Neurol* 2016;**263**(6):1204–14.

24 Mah JK, et al. A multinational study on motor function in early-onset FSHD. *Neurology* 2018;**90**(15):e1333–8.

25 Lamperti C, et al. A standardized clinical evaluation of patients affected by facioscapulohumeral muscular dystrophy: The FSHD clinical score. *Muscle Nerve* 2010;**42**(2):213–7.

26 Ricci E, et al. Progress in the molecular diagnosis of facioscapulohumeral muscular dystrophy and correlation between the number of KpnI repeats at the 4q35 locus and clinical phenotype. *Ann Neurol* 1999;**45**(6):751–7.

27 van Overveld PG, et al. Variable hypomethylation of D4Z4 in facioscapulohumeral muscular dystrophy. *Ann Neurol* 2005;**58**(4):569–76.

28 Wohlgemuth M, et al. A family-based study into penetrance in facioscapulohumeral muscular dystrophy type 1. *Neurology* 2018;**91**(5):e444–54.

29 Zhang Z, Liu A, Yao S. Convergence of China's regional incomes: 1952−1997. *China Economic Review* 2001;**12**(2-3):243–58.

30 Danieli GA, Vecchi C, Angelini C. Geographic distribution of hereditary myopathies in northeast Italy. *Soc Biol* 1974;**21**(3):235–41.

31 Takeshita K, et al. Survey of Duchenne type and congenital type of muscular dystrophy in Shimane, Japan. *Jinrui Idengaku Zasshi* 1977;**22**(1):43–7.

32 Araki S, Uchino M, Kumamoto T. Prevalence studies of multiple sclerosis, myasthenia gravis, and myopathies in Kumamoto district, Japan. *Neuroepidemiology* 1987;**6**(3):120–9.

33 Zhou YH, et al. The Molecular Diagnosis of Facioscapulohumeral Muscular Dystrophy Based on Single-molecular Fluorescence in Situ Hybridization. *Chin J Clin Neurosci* 2018.

34 Dai Y, et al. Single-molecule optical mapping enables quantitative measurement of D4Z4 repeats in facioscapulohumeral muscular dystrophy (FSHD). *J Med Genet* 2020;**57**(2):109–20.

35 Rieken A, et al. CLIA Laboratory Testing for Facioscapulohumeral Dystrophy: A Retrospective Analysis 2021;**96**(7):e1054–62.

36 Lemmers RJ, et al. Rapid and accurate diagnosis of facioscapulohumeral muscular dystrophy. *Neuromuscul Disord* 2006;**16**(9-10):615–7. author reply 617-8.

37 Ricci G, et al. Large scale genotype-phenotype analyses indicate that novel prognostic tools are required for families with facioscapulohumeral muscular dystrophy. *Brain* 2013;**136**(Pt 11):3408–17.

38 Park HJ, et al. Low D4Z4 copy number and gender difference in Korean patients with facioscapulohumeral muscular dystrophy type 1. *Neuromuscul Disord* 2015;**25**(11):859–64.

39 Statland JM, Tawil R. Risk of functional impairment in Facioscapulohumeral muscular dystrophy. *Muscle Nerve* 2014;**49**(4):520–7.

40 Padberg GW, et al. Facioscapulohumeral muscular dystrophy in the Dutch population. *Muscle Nerve Suppl* 1995(2):S81–4.

41 Goselink RJM, Mul K. Early onset as a marker for disease severity in facioscapulohumeral muscular dystrophy. *Neurology* 2019;**92**(4):e378–85.

42 Mariot V, et al. Correlation between low FAT1 expression and early affected muscle in facioscapulohumeral muscular dystrophy. *Ann Neurol* 2015;**78**(3):387–400.

43 Schätzl T, Kaiser L, Deigner HP. Facioscapulohumeral muscular dystrophy: genetics, gene activation and downstream signalling with regard to recent therapeutic approaches: an update 2021;**16**(1):129.