Original Article

DOI: 10.5582/irdr.2024.01061

Classification and epidemiologic analysis of 86 diseases in *China's Second List of Rare Diseases*

Junfeng Li[1], Meilin Liu[1,§], Han Li[2,§], Xin Zhang[3,§], Xufei Xiang[1,§], Yiping Wang[1,§], Shuyi Wang[1,§], Jinxiang Han[1,4], Yanqin Lu[1,4,*]

[1] Key Laboratory for Biotech-Drugs of National Health Commission, Key Laboratory for Rare & Uncommon Diseases of Shandong Province, Biomedical Sciences College & Shandong Medicinal Biotechnology Centre, Shandong First Medical University & Shandong Academy of Medical Sciences, Ji'nan, China;
[2] Shandong Experimental High School, Ji'nan, China;
[3] Digestive Diseases Hospital of Shandong First Medical University, Ji'ning, China;
[4] Department of Endocrinology and Metabology, The First Affiliated Hospital of Shandong First Medical University & Shandong Provincial Qianfoshan Hospital, Ji'nan, China.

SUMMARY Following the release of *China's First List of Rare Diseases* in May 2018, the Chinese government officially published *China's Second List of Rare Diseases* in September 2023. To date, there is no unified standard and international consensus for rare diseases, and epidemiologic data for most rare diseases in China are lacking. We investigated 86 rare diseases on the second list using Orphanet and other databases to clarify the classification, nomenclature, and epidemiologic data for these diseases, and we summarized the genotype and phenotype of hereditary diseases. The results showed that most of 86 rare diseases were coded in the database of Unified Medical Language System (UMLS), Orphanet, Medical Subject Headings (MeSH) and International Classification of Diseases, Eleventh Revision (ICD-11). Some rare diseases are composed by group of different disorders, in which multiple identifiers existed. Meanwhile, some rare diseases have different subtypes, which correspond to different identifiers. This increases the actual number of rare diseases in the second list. Over 50% of rare diseases are genetic rare diseases and they are mainly classified into neoplastic diseases, transplant-related disorders and neurological diseases. Epidemiologic data indicated that these rare diseases had a broad prevalence spectrum and over 20 rare diseases had a prevalence of over 1/10,000, these rare diseases in the *China's Second List of Rare Diseases* expanded the number and scope of rare diseases according to the China's official definition of rare diseases.

Keywords *China's Second List of Rare Diseases*, classification, nomenclature, epidemiology, incidence, prevalence

1. Introduction

A rare disease is a health condition with low prevalence and incidence compared with other more prevalent diseases in the general population. Although these diseases are individually rare, over 7,000 conditions have been identified, which affect 3.5%-5.9% of individuals worldwide or an estimated 263-446 million individuals collectively (*1*). Rare diseases are a global public health issue. Patients with rare diseases face challenges regarding diagnosis, treatment, and care. Great advances have been made in the diagnosis, treatment, care, and epidemiology of rare diseases in China, and several polices have been enacted to ensure progress in the area of rare diseases (*2,3*).

Before China's official definition of rare diseases in 2021, five bodies including the National Health Commission of the People's Republic of China and the National Medical Products Administration issued *China's First List of Rare Diseases* in May 2018 (*4,5*). Five years later, *China's Second List of Rare Diseases* was officially released on 21 September 2023 by six bodies. Until now, 207 rare diseases have been included on this list. In September 2021, China defined rare diseases as a condition with at least one of the following three criteria: an incidence among newborns of less than 1/10,000, a prevalence of less than 1/10,000, and an affected population of less than 140,000 (*6*). Thus, the rare diseases list and this definition have now been simultaneously adopted in China.

There is no unified international standard and consensus to date on how rare diseases should generally be named and classified. The nomenclature of rare diseases is recorded using international terminologies in different reference databases. In this study, we investigated the classification, genetic information, incidence, and prevalence of 86 conditions in *China's Second List of Rare Diseases* using Orphanet, OMIM, and other databases.

2. Methods

The nomenclature of 86 conditions in *China's Second List of Rare Diseases* was retrieved using international terminologies in different reference databases, including ICD-10 (*7*) and ICD-11 (*8*) (*https://icd.who.int/en*), Online Mendelian Inheritance in Man (OMIM) (*www.omim.org*), Orphanet (*https://www.orpha.net/consor/cgi-bin/index.php*), UMLS (*9*) (*https://www.nlm.nih.gov/research/umls/index.html*), MeSH (*https://www.nlm.nih.gov/mesh/meshhome.html*), Medical Dictionary for Regulatory Activities Terminology (MedDRA) (*10*) (*https://www.meddra.org/*), and Genetic and Rare Diseases (GARD) (*https://rarediseases.info.nih.gov/*). The classification and epidemiology of 86 rare diseases were analyzed through Orphanet database. The genetic information was summarized using OMIM database.

3. Results

3.1. Nomenclature of 86 diseases in *China's Second List of Rare Diseases*

Rare diseases in *China's Second List of Rare Diseases* were mapped to seven different reference databases (Table 1). UMLS coded the most rare diseases (n = 78, 90.70%) with a unique identifier, followed by 77 diseases (89.53%) in ORPHAcode, 74 diseases (86.05%) in MeSH, 66 diseases (76.74%) in ICD-11, and 64 diseases (74.42%) in ICD-10, 59 diseases (68.60%) in GARD and 57 diseases (66.28%) in MedDRA (Supplemental Table S1, *https://www.irdrjournal.com/action/getSupplementalData.php?ID=220*). In ICD-10, identifier C49.9 encodes both dermatofibrosarcoma protuberans and epithelioid sarcoma, Lennox-Gastaut syndrome and West syndrome were classified into other generalized epilepsy and epileptic syndromes (G40.4). Bardet-Biedl syndrome and IGF1 deficiency were classified into hypopituitarism in ICD-11 (5A61.0). Except MeSH, melanoma and pemphigus were coded by at least two different codes in all other 6 nomenclature systems. PIK3CA related overgrowth syndrome and was multiple coded in 5 different nomenclature systems.

In Orphanet database, acquired hemophilia, congenital biliary atresia, malignant hyperthermia, melanoma, pemphigus and thalassemia major were coded using more than two different codes. ORPHA163596 and 231214 coded alpha-thalassemia major and beta-thalassemia major, respectively. Syndrome with alpha-thalassemia as a major feature (ORPHA 232288) was a group of disorders consisting of alpha-thalassemia-intellectual disability syndrome linked to chromosome 16 (ORPHA 98791), alpha-thalassemia-myelodysplastic syndrome (ORPHA 231401) and X-linked alpha-thalassemia-intellectual disability syndrome (ORPHA 847). Code 182095 represents a group of disorders of interstitial lung disease (ILD), including interstitial lung disease in childhood and adulthood ORPHA (264757), interstitial lung disease specific to adulthood (ORPHA 264735) and interstitial lung disease specific to childhood (ORPHA 264656).

In ICD-10, nine different diseases were non-coded. Cutaneous neuroendocrine carcinoma, melanoma, pemphigus, pheochromocytoma, PIK3CA related overgrowth syndrome and thalassemia major were coded using more than three different codes. In ICD-11, five diseases were non-coded. Thirteen types of rare tumor were coded using at least two different identifiers. A total of 26 diseases were coded using more than two different identifiers and 13 of them were tumor diseases. In UMLS, gastroenteropancreatic neuroendocrine neoplasm was non-coded. In the databases of MeSH, GARD and MeDRA, a total of 11, 24 and 26 rare diseases were non-coded, respectively (Table 1, Supplemental Table S1, *https://www.irdrjournal.com/action/getSupplementalData.php?ID=220*).

3.2. Classification of 86 diseases in *China's Second List of Rare Diseases*

In the Orphanet classification system of rare diseases, 83 diseases with ORPHAcodes in *China's Second List of Rare Diseases* were mapped (Supplemental Table S2, *https://www.irdrjournal.com/action/getSupplementalData.php?ID=221*). Of these, 45 are genetic rare diseases, accounting for 53.49%. These were followed by 28 (32.56%) neoplastic diseases and transplant-related disorders, respectively; 24 neurological diseases (27.91%), 15 skin diseases (17.44%), 13 renal diseases (15.12%), 12 developmental anomalies during embryogenesis, hepatic diseases, hematological diseases, endocrine diseases, ophthalmic disorders and systemic and rheumatological diseases, each accounting for 13.95% (Table 2).

3.3. Group of rare disorders and genotypic heterozygosity of rare diseases

Some rare diseases are groups of disorders and others have different subtypes according to classification, leading to a higher number of rare diseases in *China's Second List of Rare Diseases* than the actual number (Supplemental Table S3, *https://www.irdrjournal.com/action/getSupplementalData.php?ID=222*). As an

Table 1. Identifiers of 86 rare diseases in *China's Second List of Rare Diseases* encoded by seven different reference databases

#	Rare disease	ORPHA	ICD-10	ICD-11	UMLS	MeSH	GARD	MedDRA
1	Achondroplasia	15	Q77.4	LD24.00	C0001080	D000130	8173	C0001080
2	Acquired hemophilia	599480	D66.	3B22	C0272325	C536392	6405	10082745
		955485	X01		C0398609			10082747
3	Acromegaly	963	E22.0	5A60.0	C0001206	D000172	5725	10000599
4	Adult-onset Still disease	829	M06.1	FA23	C0085253	D016706	436	10064056
5	Alagille syndrome	52	Q44.7	LB20.0Y	C0085280	D016738	804	10053870
6	Alpha-1-antitrypsin deficiency	60	E88.0	5C5A	C0221757	D019896	5784	10001806
7	ANCA-associated vasculitis	156152	-	4A44.A	C2717865	D056648	-	-
8	Bardet-Biedl syndrome	110	Q87.8	5A61.0	C0752166	D020788	6866	10056715
9	Behçet's disease	117	M35.2	4A62	C0004943	D001528	848	10004213
10	Blue rubber bleb nevus	1059	Q27.8	LC51	C0346072	C536240	5940	-
11	CDKL5-deficiency disorder	505652	-	8A62.Y	C4750718	C564064	-	10083005
12	Choroideremia	180	H31.2	9B61	C0008525	D015794	6061	10008791
13	Chronic inflammatory demyelinating polyneuropathy	2932	G61.8	8C01.3	C0393819	D020277	6102	10057645
14	Clear cell sarcoma of kidney	457246	C64	XH0765	C0334488	-	-	10009253
15	Cold agglutinin disease	56425	D59.1	3A20.1	C0175816	D000744	6130	-
16	Congenital biliary atresia	498345	Q44.2	LB20.21	C5680082	D001656	-	-
		30391						
17	Congenital factor VII deficiency	327	D68.2	3B14.7	C0015503	D005168	2238	10016079
18	Cryopyrin associated periodic syndrome/ NLRP3-associated systemic autoinflammatory disease	208650	-	4A60.1	C2316212	D056587	10927	10068850
19	Cutaneous neuroendocrine carcinoma (Merkel cell carcinoma)	79140	C44.3	2C34	C0007129	D015266	9266	-
			C44.6	XH8IN8				
			C44.7					
20	Cutaneous T-cell lymphomas	178551	C84.8	XH1951	C5680497	D016410	-	-
			C86.3	2B0Y				
			C86.6	2B0Z				
				2B03.0				
				XH84A5				
				XH5SC3				
				XH7S84				
				XH7EL2				
				XH2513				
21	Cystinosis	213	E72.0	5C60.1	C4316899	D003554	6236	10011777
22	Dermatofibrosarcoma protuberans	31112	C49.9	2B53.Y	C0392784	D018223	9569	10057070
				XH4QZ8				
				XH5CT4				
				XH9V92				
23	Eosinophilic gastroenteritis	2070	K52.8	DA94.21	C1262481	C535952	-	10017902

Note: "- "indicates no records are available.

Table 1. Identifiers of 86 rare diseases in *China's Second List of Rare Diseases* encoded by seven different reference databases (continued)

#	Rare disease	ORPHA	ICD-10	ICD-11	UMLS	MeSH	GARD	MedDRA
24	Epithelioid sarcoma	293202	C49.9	2B5F.2 XH4F96 XH92Y0 XH4BT2 XH13Z5	C0205944	D012509	10181	10015099
25	Facioscapulohumeral muscular dystrophy	269	G71.0	8C70.3	C0238288	D020391	9941	10064087
26	Familial hemophagocytic lymphohistiocytosis	540	D76.1	4A01.23	C0272199	-	6589	10070904
27	Familial adenomatous polyposis	733	D12.6	2B90.Y	C0032580	D011125	6408	10056981
28	Fibrodysplasia ossificans progressiva	337	M61.1	FB31.1	C0016037	D009221	6445	10068715
29	Fragile X syndrome	908	Q99.2	LD55	C0016667	D005600	6464	10017324
30	Gangliosidosis	309144	E75.1 E75.0	5C56.00	C0017083	D005733	12510	-
31	Gastroenteropancreatic neuroendocrine neoplasm	100092	-	-	-	-	2437	-
32	Gastrointestinal stromal tumor	44890	C26.9	2B5B 2B5B.0 2B5B.1 2B5B.Y 2B5B.Z 2E87 XH9HQ1	C0238198	D046152	8598	10051066
33	Generalized pustular psoriasis	247353	L40.1	EA90.40	C0343055	-	12819	-
34	Genetic hypoparathyroidism	208593	-	-	C5680825	-	-	-
35	Giant cell arteritis	397	M31.6	4A44.2	C0039483	D013700	9615	10018250
36	Giant cell tumor of bone	363976	D48.0	2F7B 2F9B XH0492 XH4TC2	C0206638	D018212	-	-
37	Glanzmann thrombasthenia	849	D69.1	3B62.0Y	C0040015	D013915	2478	-
38	Glioblastoma	360	C71.9	2A00.00 XH0MB1 XH17J4 XH2BA5 XH49K9 XH4FN3 XH5571 XH7F82 XH8UC5 2A02.00	C1621958	D005909	2491	10018336
39	Gorlin syndrome	377	C44.9	LD2D.4	C0004779	D001478	7166	10062804
40	Hidradenitis suppurativa	-	L73.2	ED92.0	C0162836	D017497	-	-
41	Hutchinson-Gilford progeria syndrome	740	E34.8	LD2B	C0033300	D011371	7467	10036794

Note: "- "indicates no records are available.

Table 1. Identifiers of 86 rare diseases in *China's Second List of Rare Diseases* encoded by seven different reference databases (continued)

#	Rare disease	ORPHA	ICD-10	ICD-11	UMLS	MeSH	GARD	MedDRA
42	Inflammatory myofibroblastic tumor	178342	D48.7	2E92.1 2F30.Y XH66Z0	C0334121	-	7146	10067917
43	Leber congenital amaurosis	65	H35.5	9B70	C0339527	D057130	634	10070667
44	Lennox-Gastaut syndrome	2382	G40.4	8A62.1	C0238111	D065768	9912	10048816
45	Limbal stem cell deficiency	171673	H18.7	-	C1561989	D000092423	-	10072138
46	Malignant hyperthermia	423 46650	T88.3	8C78 NE86	C0024591 C2930828	D008305	-	-
47	Malignant pleural mesothelioma	50251	C45.0	2C26.0	C1377913	D000086002	7026	10059518
48	Melanoma	617910 252031 404560 618 293822 168999 97338 252050 39044	C43.001 C43.101 C43.151 C43.201 C43.251 C43.301 C43.302 C43.351 C43.352 C43.401 C43.402 C43.501 C43.551 C43.552 C43.553 C43.601 C43.602 C43.651 C43.701 C43.751 C43.752 C43.851 C43.901 C43.902	XH4846 XH25M1 2C00.1 2C22.3 2C30 2C30.0 2C30.1 2C30.2 2C30.3 2C30.Y 2C30.Z 2C70.1 2C71.1 2C81.1 2D00.0 2D01.0 2E63 2E63.0 2E63.0Z 2E63.1 2E63.Y 2E63.Z 9B71.40	C0206651 C0220633 C0346360 C1512419 C2314896 C4749348 C4749577 C5191057	D008545	8621 120161	10061252 10066384
49	Metachromatic leukodystrophy	512	E75.2	5C56.02	C0023522	D007966	3230	10067609
50	Monogenic non-syndromic obesity-Genetic non-syndromic obesity	98267	E66.8	5B81.Y	C5680229	-	-	-

Note: "- "indicates no records are available.

Table 1. Identifiers of 86 rare diseases in *China's Second List of Rare Diseases* encoded by seven different reference databases (continued)

#	Rare disease	ORPHA	ICD-10	ICD-11	UMLS	MeSH	GARD	MedDRA
51	Multiple endocrine neoplasia	276161	D44.8	2F7A.Y	C0027662	D009377	-	10061299
52	Narcolepsy	619284	-	7A20	C0027404	D009290	-	10028713
				7A20.0				
				7A20.1				
				7A20.Z				
				VV01				
53	Neuroblastoma	635	C74.9	2A00.11	C0027819	D009447	7185	10029260
				XH85Z0				
54	Neurofibromatosis	634518	Q85.001	LD2D.1	C5816781	D017253	-	-
55	Neuronal ceroid lipofuscinosis	216	E75.4	5C56.1	C0027877	D009472	10739	10074607
56	Neurotrophic keratitis	137596	H16.2	1F00.10	C0339296	-	-	10069732
57	Osteosarcoma	668	C41.9	2B50	C0029463	D012516	7284	10031291
				2B50.0				
				2B50.1				
				2B50.2				
				2B50.Y				
				2B50.Z				
				XH06W9				
				XH0Y34				
				XH1S32				
				XH1XF3				
				XH1Y90				
				XH23T4				
				XH29N8				
				XH2CD6				
				XH3T03				
				XH48A9				
				XH4EZ4				
				XH5CL5				
				XH5FH4				
				XH6E77				
				XH6LT5				
				XH6TL0				
				XH7N84				
				XH7XB9				
				XH8HG5				
				XH8J23				
				XH8X47				
				XH9119				
				XH9344				
				XH9LS2				

Note: "- "indicates no records are available.

Table 1. Identifiers of 86 rare diseases in *China's Second List of Rare Diseases* encoded by seven different reference databases (continued)

#	Rare disease	ORPHA	ICD-10	ICD-11	UMLS	MeSH	GARD	MedDRA
58	Pemphigus	704	L10.0	EB40.0	C0030809	D010392	6559	10052802
		2841	L10.1	EB40.0Y	C0085106		7354	10057053
		46485	L10.2	EB40.1	C0263312		7355	10057056
		63455	L10.3	EB40.1	C0263313			10057069
		79479	L10.4	EB40.1	C0263314			10058917
		79480	L10.8	EB40.2	C0263316			
		79481	L10.8	EB40.Y	C1112570			
		208524	L10.8	EC20.2	C1274167			
		555905	Q82.8		C4749730			
		636955			C5681323			
59	Persistent pulmonary hypertension of the newborn	-	P29.3	KB42	C0031190	D010547	-	-
60	Pheochromocytoma	-	C74.101	5A75	C0031511	D010673	-	-
			D35.051	XH3854	C0334419			
				XH9K97	C1302282			
				2D11.1				
61	PIK3CA related overgrowth syndrome	530313	Q04.5	EF02.1	C0431391	C536142	6950	10081236
			Q74.0	LA05.1	C1865285	C567763	10939	
			Q74.2	LB97.1	C2751313	C567863	2637	
			Q87.3	LD2C	C2752042	D065705		
				LD2F.1Y	C4749904			
				Q74.0	C5192432			
					C5679987			
					C5679988			
					C5680341			
62	Polycythaemia vera	729	D45	2A20.4	C0032463	D011087	7422	10036057
63	Primary biliary cholangitis	186	K74.3	DB96.1	C0008312	D008105	7459	10080429
				DB96.10				
				DB96.1Y				
				DB96.1Z				
64	Primary ciliary dyskinesia	244	Q34.8	LA75.Y	C4551720	D002925	4484	10069713
65	Primary IGF1 deficiency	73272	E34.3	5A61.0	C1837475	C563867	10627	-
66	Primary immunodeficiency	101997	-	L1-4A0	C0398686	D000081207	-	10064859
67	Primary myelofibrosis	824	D47.4	2A20.2	C0001815	D055728	8618	10077161
				XH7GG7				
68	Primary sclerosing cholangitis	171	K83.0	DB96.20	C0566602	D015209	1280	10036732
				DB96.2Y				
				DB96.2Z				

Note: "- "indicates no records are available.

Table 1. Identifiers of 86 rare diseases in *China's Second List of Rare Diseases* encoded by seven different reference databases (continued)

#	Rare disease	ORPHA	ICD-10	ICD-11	UMLS	MeSH	GARD	MedDRA
69	Progressive fibrosing interstitial lung disease	182095	J84.9 J84.1	CB03.6 CB04.Y CB04.Z CB05.1 CB05.2 CB05.3 CB05.4 CB05.Y CB05.Z	C0206062	D017563	-	10022611
70	Recurrent pericarditis	251307	I09.2	-	C4707790	-	-	-
71	Retinopathy of prematurity	90050	H35.1	9B71.3	C0035344	D012178	5695	10038933
72	Rett syndrome	778	F84.2	LD90.4	C0035372	D015518	5696	10039000
73	Short bowel syndrome	104008	-	DA96.04 KB89.1	C0036992	D012778	1502	10049416
74	Systemic juvenile idiopathic arthritis	85414	M08.2	FA24.4	C0087031	-	10966	10042061
75	Systemic mastocytosis	2467	C96.2	2A21.0 2A21.0Y 2A21.0Z XH10N1 XH1H01 XH2Y59 XH5191	C0221013	D034721	8616	10042949
76	Takayasu arteritis	3287	M31.4	4A44.1	C0039263	D013625	7730	10043097
77	Tenosynovial giant cell tumor/Pigmented villonodular synovitis	66627	M12.2	XH6911 XH0HZ1 XH52J9 XH5AQ9	C1318543	D000070779	7396	-
78	Thalassemia major	163596 231214 98791 231401 847	D46.7 D56.0 D56.1	3A50.03 3A50.0Y 3A50.1 3A50.2 D56.0	C5680928 C0002875 C0272005 C0585216 C0795917 C1845055	D017086	5864	-
79	Thrombotic thrombocytopenic purpura	54057	M31.1	3B64.14	C0034155	D011697	-	10043648
80	Transthyretin amyloidosis	271861	-	5D00.20 BC43.20	C5679761	C567782	-	-
81	Tumor necrosis factor receptor associated periodic syndrome	32960	E85.0	4A60.2	C1275126	-	8457	-
82	Tumor-induced osteomalacia	352540	M83.8	-	C1274103	C537751	9652	-
83	Von Hippel-Lindau syndrome	892	Q85.8	5A75	C0019562	D006623	7855	10047716
84	Von Willebrand disease type3	166096	D68.0	3B12	C1264041	D056729	-	-

Note: "- "indicates no records are available.

Table 1. Identifiers of 86 rare diseases in *China's Second List of Rare Diseases* encoded by seven different reference databases (continued)

#	Rare disease	ORPHA	ICD-10	ICD-11	UMLS	MeSH	GARD	MedDRA
85	Waldenström macroglobulinemia/ Lymphoplasmacytic lymphoma	33226	C88.0	XH8GW4 2A85.4 XH0QZ9	C0024419	D008258	7872	10047801
86	West syndrome/Infantile spasms syndrome	3451	G40.4	8A62.0	C0037769	D013036	7887	10021750

Note: "-" indicates no records are available.

Table 2. The number and percentage of 86 rare diseases in *China's Second List of Rare* according to classification system in Orphanet database

Orphanet classification	number	Percentage (%)
genetic diseases	45	52.35
neoplastic diseases	28	32.56
transplant-related disorders	28	32.56
neurological diseases	24	27.91
skin diseases	15	17.44
renal diseases	13	15.12
developmental anomalies during embryogenesis	12	13.95
hematological diseases	12	13.95
endocrine diseases	12	13.95
ophthalmic disorders	12	13.95
systemic and rheumatological diseases	12	13.95
bone diseases	7	8.14
respiratory diseases	7	8.14
systemic or rheumatologic diseases of childhood	7	8.14
hepatic diseases	5	5.81
circulatory system diseases	5	5.81
inborn errors of metabolism	5	5.81
gastroenterological diseases	5	5.81
immunological diseases	4	4.65
infertility disorders	3	3.45
abdominal surgical diseases	2	2.33
cardiac malformations	1	1.16
gynecological and obstetric diseases	1	1.16

example, primary immunodeficiency (ORPHA 101997) includes primary immunodeficiency owing to a defect in adaptive immunity (ORPHA 179006) and primary immunodeficiency owing to a defect in innate immunity (ORPHA 101988). The former group includes combined T- and B-cell immunodeficiency (ORPHA 101972), immune dysregulation disease with immunodeficiency (ORPHA 169361), immunodeficiency predominantly affecting antibody production (ORPHA 101977), and combined immunodeficiency (ORPHA 331217). The latter group includes autoinflammatory syndrome with immune deficiency (ORPHA 290839), genetic susceptibility to infections owing to particular pathogens (ORPHA 183710) , immunodeficiency owing to a complement cascade protein anomaly (ORPHA 101992), other immunodeficiency syndromes owing to defects in innate immunity (ORPHA 331193), primary immunodeficiency with predisposition to severe viral infection (ORPHA 431156), and quantitative and/ or qualitative congenital phagocyte defect (ORPHA 101985).

Genetic rare diseases in the second list are summarized in Supplemental Table S4 (*https://www.irdrjournal.com/action/getSupplementalData.php?ID=223*). Inheritance, phenotype, and pathogenic genes are included in the table as data collected from the OMIM database. Some rare diseases are caused by mutation of single genes and others are caused by mutation in multiple genes. A total of 53 different phenotypes with 49 pathogenic genes exist in primary ciliary dyskinesia, in both the autosomal recessive

(AR), autosomal dominant (AD) or X-linked recessive inheritance format. Bardet–Biedl syndrome (BBS) is inherited mainly in the form of AR with 22 different phenotypes and 24 genes. Phenotype BBS1 corresponds to three different pathogenic genes *CCDC28B*, *ARL6*, and *BBS1*, with AD or AR inheritance patterns. The phenotype BBS14 corresponds to the *TMEM67* and *CEP290* genes inherited in AR form. The *ARL6* gene corresponds to phenotypes of both BBS1 and BBS3. There are 20 different genes and a corresponding 18 phenotypes in Leber congenital amaurosis. For neuronal ceroid lipofuscinosis, a total of 12 genes corresponds to 14 different subtypes and phenotypes. For Alagille syndrome, three subtypes are classified according to different pathogenic genes with microdeletion in 20p12, *JAG1*, and *NOTCH2*. In summary, genotypic and phenotypic heterozygosity exist among these rare diseases.

3.4. Epidemiology of diseases i*n China's Second List of Rare Diseases*

According to data from Orphanet, the epidemiology of the 86 rare diseases is described in Table 3. Data on the newborn incidence, incidence, and prevalence of 71 rare diseases (82.56%) in *China's Second List of Rare Diseases* are available, including data on the newborn incidence of 7 (8.14%), incidence of 20 rare diseases (23.26%) and the prevalence of 64 (74.42%) rare diseases. As we mentioned above (also see Supplementarry Table S3 (*https://www.irdrjournal.com/action/getSupplementalData.php?ID=222*), some rare disorders are groups composed of different disoders with some correlations, hence, only disorders in the group with epidemiological data were collected. Newborn incidence of 0.6/100,000 in the pediatric population and incidence of 1.6-23/100,000 per year in children were reported in systemic juvenile idiopathic arthritis. The birth incidence of familial adenomatous polyposis was as high as 12.05/100,000. In all, an annual incidence of 20 rare diseases available in Orphanet, a relatively broad newborn incidence was observed in malignant hyperthermia of anesthesia, systemic juvenile idiopathic arthritis, primary biliary cholangitis, polycythaemia vera, Lennox-Gastaut syndrome and cutaneous neuroendocrine carcinoma (Merkel cell carcinoma). The annual incidence of beta-thalassemia major is estimated at 1/100,000 worldwide and 1/10,000 in the European population.

A total of 23 rare diseases or its subtypes showed a prevalence of higher than 1/10,000 including alpha-1-antitrypsin deficiency, ANCA-associated vasculitis, dermatofibrosarcoma protuberans, Fragile X syndrome, neuroendocrine tumor of pancreas, gastrointestinal stromal tumor, giant cell arteritis, Lennox-Gastaut syndrome, Limbal stem cell deficiency, narcolepsy type 1, neuroblastoma, neurotrophic keratitis, pemphigus vulgaris, polycythaemia vera, primary biliary cholangitis, retinopathy of prematurity, gastroschisis, systemic mastocytosis, tenosynovial giant cell tumor (pigmented villonodular synovitis), Hb Bart's hydrops fetalis and thrombotic thrombocytopenic purpura. For thrombotic thrombocytopenic purpura (TTP), the prevalence of iTTP was 14.29/100,000, which is higher than that of the cTTP type. Thirty-five different rare diseases had a prevalence of less than 10/100,000 and greater than or equal to 1/100,000, these included acromegaly, adult-onset Still disease, Bardet-Biedl syndrome, Behçet's disease, CDKL5-deficiency disorder, choroideremia, chronic inflammatory demyelinating polyneuropathy, biliary atresia, cutaneous neuroendocrine carcinoma, cystinosis, eosinophilic gastroenteritis, facioscapulohumeral muscular dystrophy, familial adenomatous polyposis, glioblastoma, Gorlin syndrome, Leber congenital amaurosis, malignant pleural mesothelioma, multiple endocrine neoplasia, and so on. Epithelioid sarcoma, Hutchinson-Gilford progeria syndrome, primary IGF1 deficiency, fibrodysplasia ossificans progressiva, hypoplastic pancreas-intestinal atresia-hypoplastic gallbladder syndrome, cold agglutinin disease, congenital factor VII deficiency and another 13 rare diseases had relatively low prevalence at less than 1/100,000.

4. Discussion

The study nominated the 86 diseases in *China's Second List of Rare Diseases* using different reference systems. The results showed that most rare diseases have unique identifiers, especially in the Orphanet database. Orphanet nomenclature is a powerful classification tool based on the multidimensional nature of rare diseases, which provides a specific terminology for rare diseases. Each clinical entity is assigned a unique and time-stable ORPHAcode. This includes all disorders, subtypes of disorders, and groups of disorders. Over 6000 rare diseases are coded using ORPHAcodes (*11*). ORPHAcodes have high sensitivity and accuracy in describing rare diseases. More than half of ORPHAcodes identify diseases with very low prevalence (less than 1 case per million) (*12*). The International Classification of Diseases is a medical classification list produced by the World Health Organization that is predominantly used in health care systems worldwide. However, less rare diseases are listed in versions older than ICD-10. One ICD code can correspond to different rare entities or to both rare and non-rare entities. This is often not clearly distinguished in health information systems (*13,14*). ICD-11 was adopted by the World Health Assembly in May 2019 and came into effect in January 2022. This version includes nearly 5,500 rare diseases and each has a unique identifier. Rare diseases in ICD-11 are easily available in health information systems and are continually updated (*5,14*). It has a unique identifier for different subtypes of one disease and different conditions

Table 3. Incidence and prevalence of 86 rare diseases in *China's Second List of Rare Diseases* recorded in Orphanet database

#	Rare diseases	Newborn Incidence / 100,000 persons	Incidence / 100,000 persons	Prevalence / 100,000 persons
1	Achondroplasia	4	-	-
2	Acquired hemophilia A	-	-	-
	Acquired hemophilia B	-	-	-
	hemophilia C	-	-	0.1-0.9
3	Acromegaly	-	0.19-1.1 (annual)	1-9
4	Adult-onset Still disease	-	-	1-9
5	Alagille syndrome	-	-	1.43
6	Alpha-1-antitrypsin deficiency		-	10-50
7	ANCA-associated vasculitis	-	-	10-50
	Eosinophilic granulomatosis with polyangiitis	-	-	1-9
	Granulomatosis with polyangiitis	-	0.21-1.19	1-9
	Microscopic polyangiitis	-	-	1-9
8	Bardet-Biedl syndrome	-	-	1 (USA) 1.69 (Denmark) 1.52-2.22 (France)
9	Behçet's disease	-	-	1-9
10	Blue rubber bleb nevus	-	-	-
11	CDKL5-deficiency disorder		-	2.36 (UK (Scotland) birth prevalence
12	Choroideremia	-	-	1-9
13	Chronic inflammatory demyelinating polyneuropathy	-	-	1-9
14	Clear cell sarcoma of kidney	-	-	-
15	Cold agglutinin disease	-	-	0.1-0.9
16	biliary atresia		-	-
	Isolated biliary atresia		-	1-9
	Biliary atresia with splenic malformation syndrome	-	-	-
	Hypoplastic pancreas-intestinal atresia-hypoplastic gallbladder syndrome	-	-	< 0.1
17	Congenital factor VII deficiency	-	-	0.1-0.9
18	Cryopyrin associated periodic syndrome- NLRP3-associated systemic autoinflammatory disease	-	-	0.28 (France)
	CINCA syndrome	-	-	0.28 (the whole spectrum of CAPS)
	Familial cold urticaria	-	-	-
	Muckle-Wells syndrome	-	-	0.28 (France)
19	Cutaneous neuroendocrine carcinoma (Merkel cell carcinoma)	-	0.2-0.4 (annual, white population)	1-9
20	Cutaneous T-cell lymphomas	-	-	-
	Mycosis fungoides and variants		0.29-0.91	-
	Primary cutaneous CD30+ T-cell lymphoproliferative disease	-	-	-
	Adult T-cell leukemia-lymphoma	-	-	1-9
21	Cystinosis		-	1-9
22	Dermatofibrosarcoma protuberans	-	0.5	10-50
23	Eosinophilic gastroenteritis	-	-	-
24	Epithelioid sarcoma	-	-	< 0.1
25	Facioscapulohumeral muscular dystrophy	-	-	1-9
26	Familial hemophagocytic lymphohistiocytosis	-	-	-
27	Familial adenomatous polyposis	12.05	-	1-9
28	Fibrodysplasia ossificans progressiva	-	-	< 0.1
29	Fragile X syndrome	-	-	10-50
30	Gangliosidosis	-	-	-
	GM1 gangliosidosis		-	0.5-1 (live births)
	GM1 gangliosidosis type 1	-	-	0.5-1(live births)
	GM1 gangliosidosis type 2	-	-	< 0.1
	GM1 gangliosidosis type 3	-	-	< 0.1
	GM2 gangliosidosis	-	-	1-9
	GM2 gangliosidosis, AB variant	-	-	< 0.1
	Sandhoff disease	-	-	0.1-0.9
	Tay-Sachs disease	-	-	-
31	Gastroenteropancreatic neuroendocrine neoplasm	-	-	-
	Neuroendocrine tumor of stomach	-	-	1-9
	Neuroendocrine carcinoma of pancreas	-	-	-
	Neuroendocrine tumor of pancreas	-	-	10-50
32	Gastrointestinal stromal tumor	-	-	10-50
33	Generalized pustular psoriasis	-	-	0.1-0.9

Table 3. Incidence and prevalence of 86 rare diseases in *China's Second List of Rare Diseases* recorded in Orphanet database (continued)

#	Rare diseases	Newborn Incidence / 100,000 persons	Incidence / 100,000 persons	Prevalence / 100,000 persons
34	Genetic hypoparathyroidism	-	-	-
	Autoimmune polyendocrinopathy type 1	-	-	0.1-0.9
	Familial isolated hypoparathyroidism	-	-	< 0.1
	Pseudohypoparathyroidism	-	-	0.1-0.9
35	Giant cell arteritis	-	5.88-20 (annual, adults over 50 years old)	10-50
36	Giant cell tumor of bone	-	-	-
37	Glanzmann thrombasthenia	-	-	-
38	Glioblastoma	-	3	1-9
39	Gorlin syndrome	-	-	1-9
40	Hidradenitis suppurativa	-	-	-
41	Hutchinson-Gilford progeria syndrome	-	-	< 0.1
42	Inflammatory myofibroblastic tumor	-	-	-
43	Leber congenital amaurosis		-	1-9
44	Lennox-Gastaut syndrome	-	0.1-0.28	10-50
45	Limbal stem cell deficiency	-	-	10-50
46	Malignant hyperthermia	-	-	-
	Malignant hyperthermia of anesthesia	-	45342	-
47	Malignant pleural mesothelioma	-	-	-
48	Melanoma	-	-	-
	Conjunctival malignant melanoma	-	-	< 0.1
	Malignant melanoma of the mucosa	-	-	-
	Uveal melanoma	-	-	1-9
49	Metachromatic leukodystrophy	-	-	0.1-0.9
50	Monogenic non-syndromic obesity-Genetic non-syndromic obesity	-	-	-
51	Multiple endocrine neoplasia	-	-	3.33-10 (MEN1) 2.86 (MEN2)
52	Narcolepsy	-	-	-
	Narcolepsy type 1	-	-	10-50
	Narcolepsy type 2	-	-	-
53	Neuroblastoma	-	1.43 (annual, 15 years old)	10-50
54	Neurofibromatosis	-	-	-
	Full NF2-related schwannomatosis	-	-	1-9
	Full schwannomatosis	-	-	-
	Neurofibromatosis type 1		-	10-50
55	Neuronal ceroid lipofuscinosis	-	-	1-9
	ATP13A2-related juvenile neuronal ceroid lipofuscinosis	-	-	< 0.1
	Congenital neuronal ceroid lipofuscinosis	-	-	-
	Infantile neuronal ceroid lipofuscinosis	5 (Finland)	-	-
	Juvenile neuronal ceroid lipofuscinosis	2.22 (Sweden) 6.99 (Germany)		0.46 (Sweden)
	Late infantile neuronal ceroid lipofuscinosis	-		0.3 (Sweden)
56	Neurotrophic keratitis	-	-	42.02 (Europe)
57	Osteosarcoma	-	0.3 (annual)	1-9
58	Pemphigus	-	-	/
	Endemic pemphigus foliaceus	-	-	-
	Pemphigus foliaceus	-	-	-
	Pemphigus vulgaris	-	0.07-0.7 (annual)	10-50
59	Persistent pulmonary hypertension of the newborn	-	-	-
60	Pheochromocytoma	-	-	-
61	PIK3CA related overgrowth syndrome	-	-	-
	CLAPO syndrome	-	-	< 0.1
	CLOVES syndrome	-	-	< 0.1
	Congenital infiltrating lipomatosis of the face	-	-	< 0.1
	Hemihyperplasia-multiple lipomatosis syndrome	-	-	< 0.1
	Megalencephaly-capillary malformation-polymicrogyria syndrome	-	-	< 0.1
	Segmental progressive overgrowth syndrome with fibroadipose hyperplasia	-	-	< 0.1
62	Polycythaemia vera	-	1-2.78	10-50
63	Primary biliary cholangitis	-	0.33-5.8(annual)	10-50
64	Primary ciliary dyskinesia	3.33-6.67	-	-

Table 3. Incidence and prevalence of 86 rare diseases in *China's Second List of Rare Diseases* recorded in Orphanet database (continued)

#	Rare diseases	Newborn Incidence / 100,000 persons	Incidence / 100,000 persons	Prevalence / 100,000 persons
65	Primary IGF1 deficiency	-	-	< 0.1
66	Primary immunodeficiency	-	-	1-9
67	Primary myelofibrosis	-	1 (annual)	1-9
68	Primary sclerosing cholangitis	-	-	1-9
69	Interstitial lung disease	-	-	-
70	Recurrent pericarditis	-	-	-
71	Retinopathy of prematurity	-	-	10-50
72	Rett syndrome	-	-	1-9
73	Short bowel syndrome	-	-	1-9
	Gastroschisis		-	10-50
	Small bowel atresia		-	4-14.29 (live births, Europe)
74	Systemic juvenile idiopathic arthritis	0.6 (prediatric population)	1.6-23 (annual)	1-9
75	Systemic mastocytosis	-	-	10-50
	Aggressive systemic mastocytosis	-	-	0.1-0.9
	Indolent systemic mastocytosis	-	-	10-50
	Systemic mastocytosis with associated hematologic neoplasm	-	-	1-9
	Mast cell leukemia	-	-	< 0.1
	Acute mast cell leukemia	-	-	-
76	Takayasu arteritis	-	-	1-9
77	Tenosynovial giant cell tumor-Pigmented villonodular synovitis	-	-	10-50
78	Thalassemia major	-	-	-
	Beta-thalassemia major	-	1 (worldwide) 10 (EU)	-
	Hb Bart's hydrops fetalis		-	50-500 (live births, Southeast Asia)
79	Thrombotic thrombocytopenic purpura	0.1-0.61(iTTP)	-	1.29 (iTTP, France) 0.04-1.67 (cTTP)
80	Transthyretin amyloidosis	-	-	0.1-0.9
	ATTRV30M amyloidosis	-	0.87 (annual, Portugal)	22.94 (adults)
81	Tumor necrosis factor receptor associated periodic syndrome	-	0.57 (annual, children under 16, Germany)	
82	Tumor-induced osteomalacia	-	-	-
83	Von Hippel-Lindau syndrome	2.78	-	1-9
84	Von Willebrand disease type3	-	-	0.1-0.9
85	Waldenström macroglobulinemia- Lymphoplasmacytic lymphoma	-	0.38 (annual, USA)	0.99 (Europe)
86	West syndrome-Infantile spasms syndrome		-	1-9

in a group. ICD-11, Orphanet, OMIM, MalaCards and other databases all cross-reference each other.

According to global epidemiologic data and the 2021 definition of rare diseases, conditions with a relatively high newborn incidence and/or prevalence (1/10,000) are recorded in *China's Second List of Rare Diseases*. Data on these registered rare diseases are readily available to advance research and development of orphan drugs and treatments for rare diseases.

In conclusion, in this study, we reviewed the classification, nomenclature, and epidemiology of 86 rare diseases in *China's Second List of Rare Diseases*. Including the 121 rare diseases in the first list, a total of 207 rare diseases have been described by the Chinese government. Administrative policies regarding rare diseases have been enacted, which refer to research, medical insurance, orphan drugs, and standards for diagnosis and treatment. These policies make important contributions to progress in the area of rare diseases in China. The diseases on China's second list are important research topics and focusing on these will help China become a healthcare model with respect to rare diseases. The inclusion of conditions with relatively higher prevalence than those included in the current definitions will benefit more patients with rare diseases. With further development of national registries, rare diseases with very low prevalence will likely emerge.

Funding: This work was supported by a grant from the Natural Science Foundation of Shandong Province (General program ZR2023MH276) and Academic Promotion Program of Shandong First Medical University (LJ001).

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

References

1. Shourick J, Wack M, Jannot AS. Assessing rare diseases prevalence using literature quantification. Orphanet J Rare Dis. 2021; 16:139.
2. Zhi W, Liu M, Yang D, Zhang S, Lu Y, Han J. Analysis of marketed orphan drugs in China. Intractable Rare Dis Res. 2023; 12:132-140.
3. Li X, Wu L, Yu L, He Y, Wang M, Mu Y. Policy analysis in the field of rare diseases in China: A combined study of content analysis and Bibliometrics analysis. Front Med (Lausanne). 2023; 10:1180550.
4. He J, Kang Q, Hu J, Song P, Jin C. China has officially released its first national list of rare diseases. Intractable Rare Dis Res. 2018; 7:145-147.
5. World Health Organization. Rare diseases. *https://www.who.int/standards/classifications/frequently-asked-questions/rare-diseases* (accessed October 2, 2024).
6. Lu Y, Han J. The definition of rare disease in China and its prospects. Intractable Rare Dis Res. 2022; 11:29-30.
7. World Health Organization. ICD-10 Version:2019. *https://icd.who.int/browse10/2019/en* (accessed October 4, 2024).
8. World Health Organization. International Statistical Classification of Diseases and Related Health Problems (ICD). *https://www.who.int/standards/classifications/classification-of-diseases* (accessed October 4, 2024).
9. Bodenreider O. The Unified Medical Language System (UMLS): Integrating biomedical terminology. Nucleic Acids Res. 2004; 32:D267-D270.
10. Brown EG, Wood L, Wood S. The medical dictionary for regulatory activities (MedDRA). Drug Saf. 1999; 20:109-117.
11. Nguengang Wakap S, Lambert DM, Olry A, Rodwell C, Gueydan C, Lanneau V, Murphy D, Le Cam Y, Rath A. Estimating cumulative point prevalence of rare diseases: analysis of the Orphanet database. Eur J Hum Genet. 2020; 28:165-173.
12. Mazzucato M, Pozza LVD, Facchin P, *et al.* ORPHAcodes use for the coding of rare diseases: Comparison of the accuracy and cross country comparability. Orphanet J Rare Dis. 2023; 18:267.
13. Rath A, Aymé S, Bellet B. Classification of rare diseases: A worldwide effort to contribute to the International Classification of Diseases. Orphanet J Rare Dis. 2010; 5:O21.
14. Aymé S, Bellet B, Rath A. Rare diseases in ICD11: Making rare diseases visible in health information systems through appropriate coding. Orphanet J Rare Dis. 2015; 10:35.
15. Jia J, An Z, Ming Y, Guo Y, Li W, Liang Y, Guo D, Li X, Tai J, Chen G, Jin Y, Liu Z, Ni X, Shi T. eRAM: Encyclopedia of rare disease annotations for precision medicine. Nucleic Acids Res. 2018; 46:D937-D943.

Received November 2, 2024; Revised November 21, 2024; Accepted November 25, 2024.

[§]These authors contributed equally to this work.
**Address correspondence to:*
Yanqin Lu, The First Affiliated Hospital of Shandong First Medical University & Shandong Provincial Qianfoshan Hospital, No. 16766 Jingshi Road, Ji'nan, Shandong 250013, China.
E-mail: yqlu@sdfmu.edu.cn

Released online in J-STAGE as advance publication November 29, 2024.