

## Tutorial for using interactive website for rare diseases in the UK Biobank

1. Select a disease from the rare disease table. The table consists of four columns: Disease Name (from OrphaNet), ORPHA Number (from OrphaNet), ICD10 Codes (from our consensus ICD-10 to ORPHA number mapping), UKB Count (the number of individuals in the UK Biobank indicated as having the rare disease by our mapping) and Group (determined from the ICD-10 chapters).

Show 10 entries Search:

Disease Name	ORPHA Number	ICD10 Codes	UKB Count	Group
22q11.2 deletion syndrome	567	D82.1	18	Other Immune
45, X/46, XY mixed gonadal dysgenesis	1772	Q98.7	1	Congenital
AA amyloidosis	85445	E85.3	35	Endocrine/metabolic
Absence of uterine body	180142	Q51.0	3	Congenital

2. You can change the number of entries shown in a single screen by selecting a different value from the “Show XX entries” drop-down box. It is also possible to search for specific rare diseases using the “Search” box or sort by the values in each column by clicking on the up/down arrows next to the column header. At the bottom of the table, is a navigation tool, which can be used to move to subsequent rare diseases, either by clicking on the screen number, or the “Next” button.

Acheiria	294983	Q71.3	4	Congenital
Achondroplasia	15	Q77.4	15	Congenital
Acquired ichthyosis	454	L85.0	6	Skin/Subcutaneous
Acquired methemoglobinemia	464453	D74.8	1	Blood
Acromegaly	963	E22.0	134	Endocrine/metabolic
Actinomycosis	457095	A42.0, A42.1, A42.2, A42.7, A42.8	17	Infectious/parasitic
Disease Name	ORPHA Number	ICD10 Codes	UKB Count	Group

Showing 1 to 10 of 420 entries Previous **1** 2 3 4 5 ... 42 Next

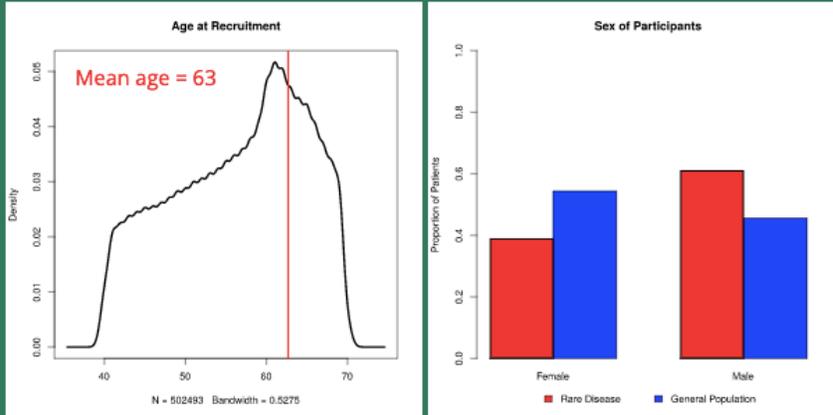
3. When hovering over a rare disease in the table, its background color will change to a darker grey, indicating it is clickable. Then, when clicking on a disease in the table, its

background color will become yellow and information about the disease will be displayed above.

The screenshot shows a web interface with a dark green header containing navigation tabs: 'About', 'Disease Summary' (selected), 'UK Biobank Demographics', and 'UK Biobank Comorbidities'. Below the header is a card for '22q11.2 deletion syndrome'. On the left side of the card, there are details: 'Prevalence: 1-5 / 10 000', 'Age of onset: All ages', 'Age of death: any age', 'Inheritance: Autosomal dominant', and 'Genes: ARVCF, COMT, GP1BB, HIRA'. On the right side, there is a table with two columns: 'Phenotype' and 'Frequency'. The phenotypes listed are 'Abnormal aortic arch morphology', 'Abnormal facial shape', 'Abnormality of the pharynx', 'Abnormal pulmonary valve morphology', 'Atrial septal defect', and 'Bulbous nose', all with a frequency of 'Very frequent (99-80%)'. Below the card, there is a search bar and a table with the following data:

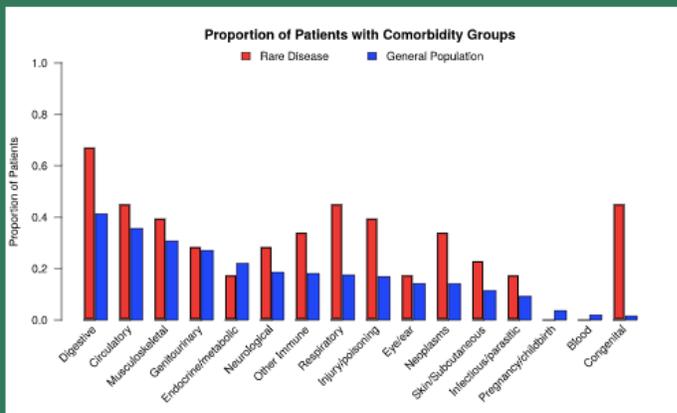
Disease Name	ORPHA Number	ICD10 Codes	UKB Count	Group
22q11.2 deletion syndrome	567	D82.1	18	Other Immune

- The first tab to be automatically selected at the top of the screen is “Disease Summary”. This shows information about the rare disease extracted and summarized from OrphaNet. On the left-hand side are specific details, including the disease prevalence, age of onset, age of death, inheritance mode, associated genes (according to OrphaNet) and disease categories, while on the right-hand side are the phenotypes known to occur with the disease, such as its symptoms. While this information is displayed in the most compact way possible, it is straightforward to access further information on the rare disease by clicking on its name in the left-hand side (again, when hovered over, it will change color, indicating it is clickable). Clicking on the name of the rare disease will open a new tab containing the OrphaNet website for the rare disease.
- Clicking on ‘UK Biobank Demographics’ will navigate to the next tab, showing the age at recruitment and sex of participants for the rare disease, from our analysis of the UK Biobank. Both data points are shown in the context of the general population in the UK Biobank. The left plot shows the age distribution for the general population in the UK Biobank as a black line, while the mean age for the rare disease is indicated by a red line (with text overlaid in red to show the mean age of recruitment for the rare disease). The right plot shows the sex proportion for the general population in the UK Biobank using blue bars while the sex proportion for the rare disease is shown using red bars. So as to abide by the privacy restrictions of the UK Biobank, no age/sex information is shown for diseases with fewer than 5 individuals.



No age/sex information is shown for diseases with fewer than 5 individuals in the UK Biobank for privacy reasons.

- Clicking on 'UK Biobank Comorbidities' will navigate to the final tab. This shows a plot of the proportion of participants with different comorbidity groups (defined according to the ICD-10 chapters) compared to the general population in the UK Biobank. The proportions for the general population are shown in blue, while the proportions for the rare disease are shown in red. Along the y-axis is the proportion of participants (either from the background population or the specific disease) that have at least one comorbidity from each group. Again, so as to abide by the privacy restrictions of the UK Biobank, no age/sex information is shown for diseases with fewer than 5 individuals.



No comorbidity information is shown for diseases with fewer than 5 individuals in the UK Biobank for privacy reasons.

If you have any questions or comments regarding the website, please do not hesitate to contact Matthew Patrick ([mattpat@med.umich.edu](mailto:mattpat@med.umich.edu)).