MSUD prevalence estimates from publicly available datasets

Rebecca Ahrens-Nicklas, MD, PhD
Division of Human Genetics, Section of Metabolism
The Children's Hospital of Philadelphia

Children's Hospital of Philadelphia®

How common is MSUD?

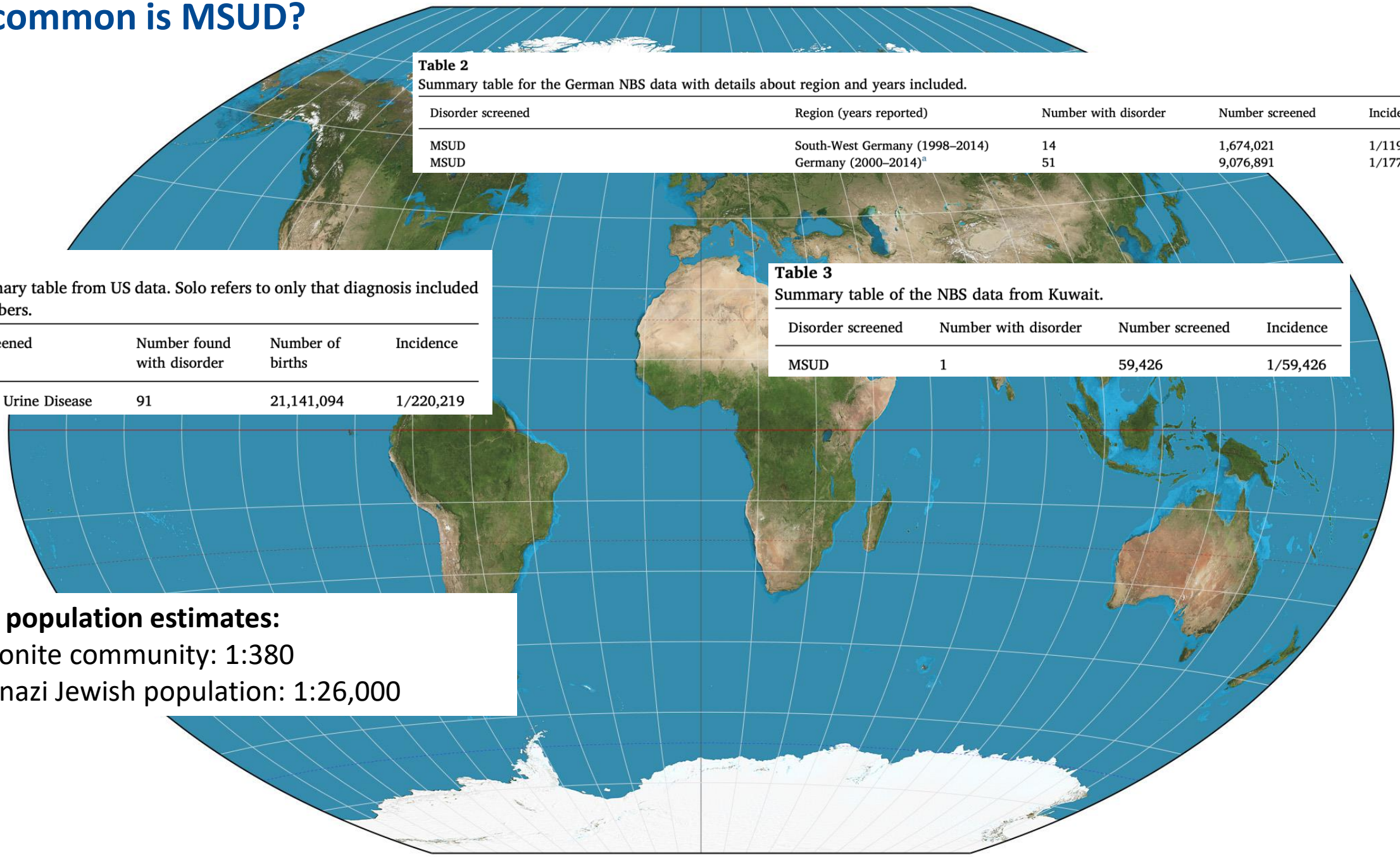


Table 2
Summary table for the German NBS data with details about region and years included.

Disorder screened	Region (years reported)	Number with disorder	Number screened	Incidence
MSUD	South-West Germany (1998–2014)	14	1,674,021	1/119,573
MSUD	Germany (2000–2014)[a]	51	9,076,891	1/177,978

Table 1
Overall summary table from US data. Solo refers to only that diagnosis included in these numbers.

Disorder screened	Number found with disorder	Number of births	Incidence
Maple Syrup Urine Disease	91	21,141,094	1/220,219

Table 3
Summary table of the NBS data from Kuwait.

Disorder screened	Number with disorder	Number screened	Incidence
MSUD	1	59,426	1/59,426

Other population estimates:
Mennonite community: 1:380
Ashkenazi Jewish population: 1:26,000

Chapman, et al. MGM Reports 15 (2018) 106–109

Why is it important to know incidence? Orphan drug designation

Orphan drug designation qualifies sponsors for incentives including:

- 7-year marketing exclusivity to sponsors of approved orphan products

- 25% federal tax credit for expenses incurred in conducting clinical research within the United States
  - Tax credits may be applied to prior year or applied over as many as 20 years to future taxes

- Waiver of Prescription Drug User Fee Act (PDUFA) fees for orphan drugs
  - A value of approximately $2.9 million in 2021

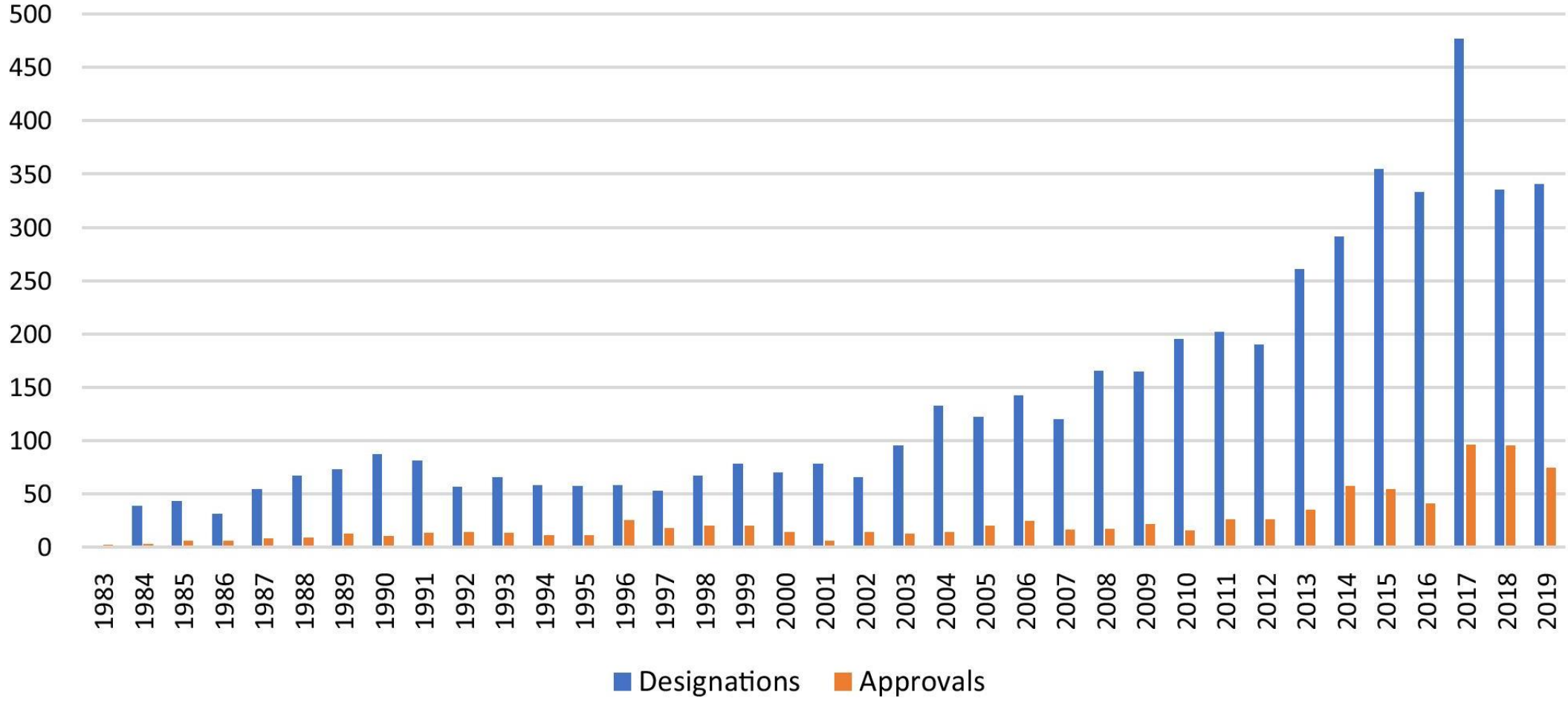
- Ability to qualify to compete for research grants from the Office of Orphan Products Development (OOPD) to support clinical studies for orphan drugs

- Eligibility to receive regulatory assistance and guidance from the FDA in the design of an overall drug development plan

Why is it important to know incidence? Orphan drug designation

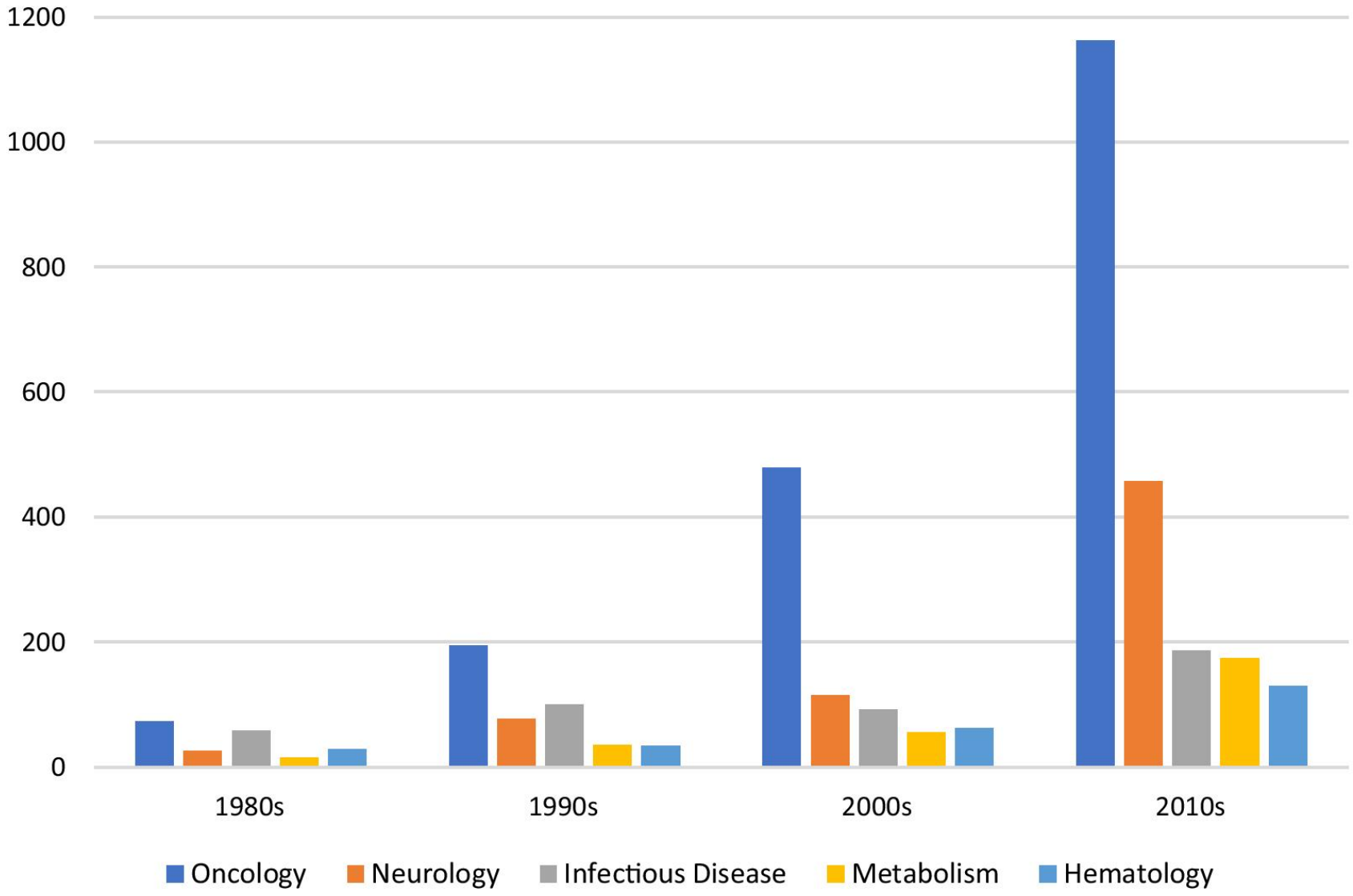
The Orphan Drug Act defines a rare disease as a disease or condition that affects less than 200,000 people in the United States.

Why is it important to know incidence? Orphan drug designation



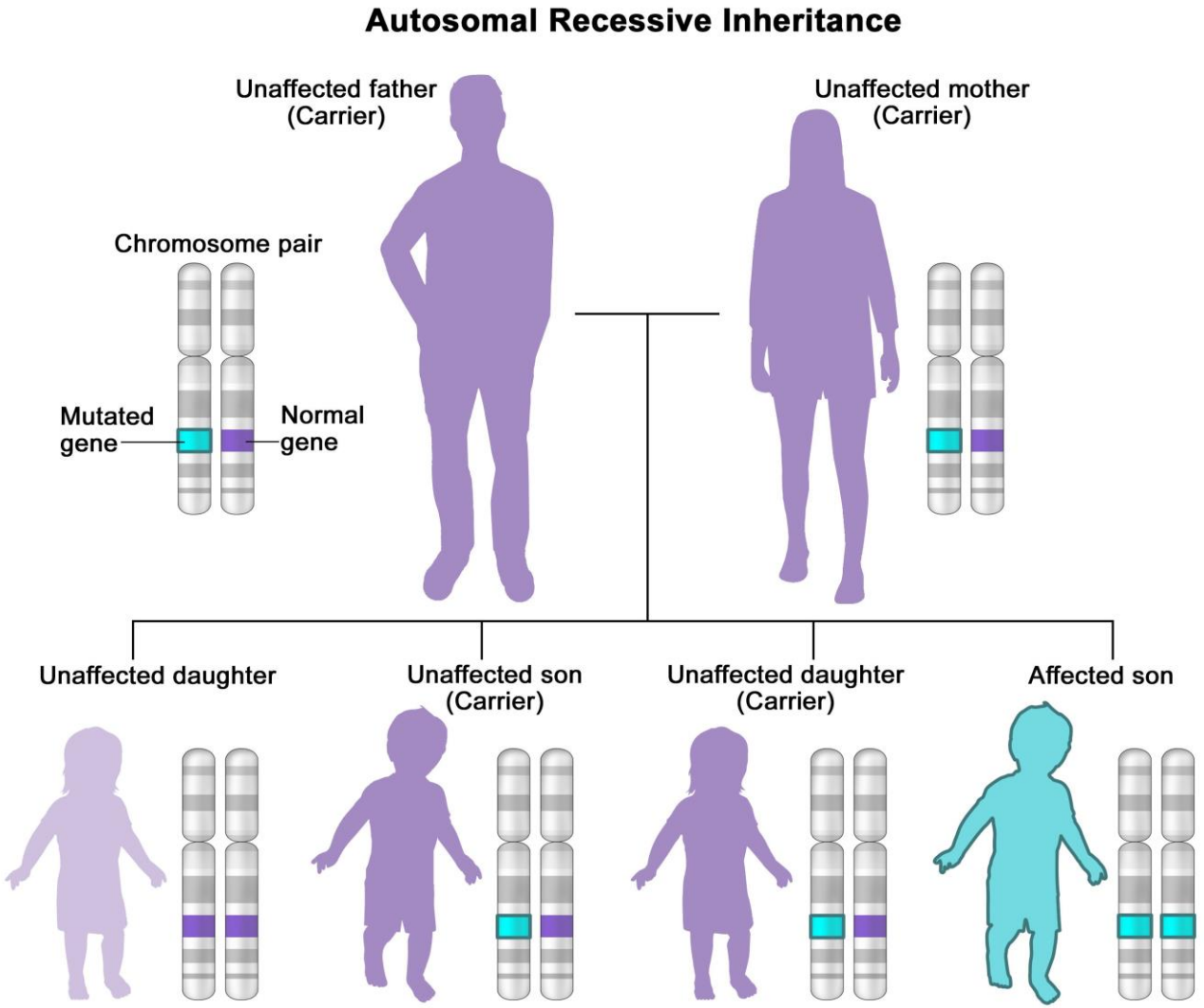
Miller et al. *Orphanet J Rare Dis* **16**, 265 (2021).

Why is it important to know incidence? Orphan drug designation

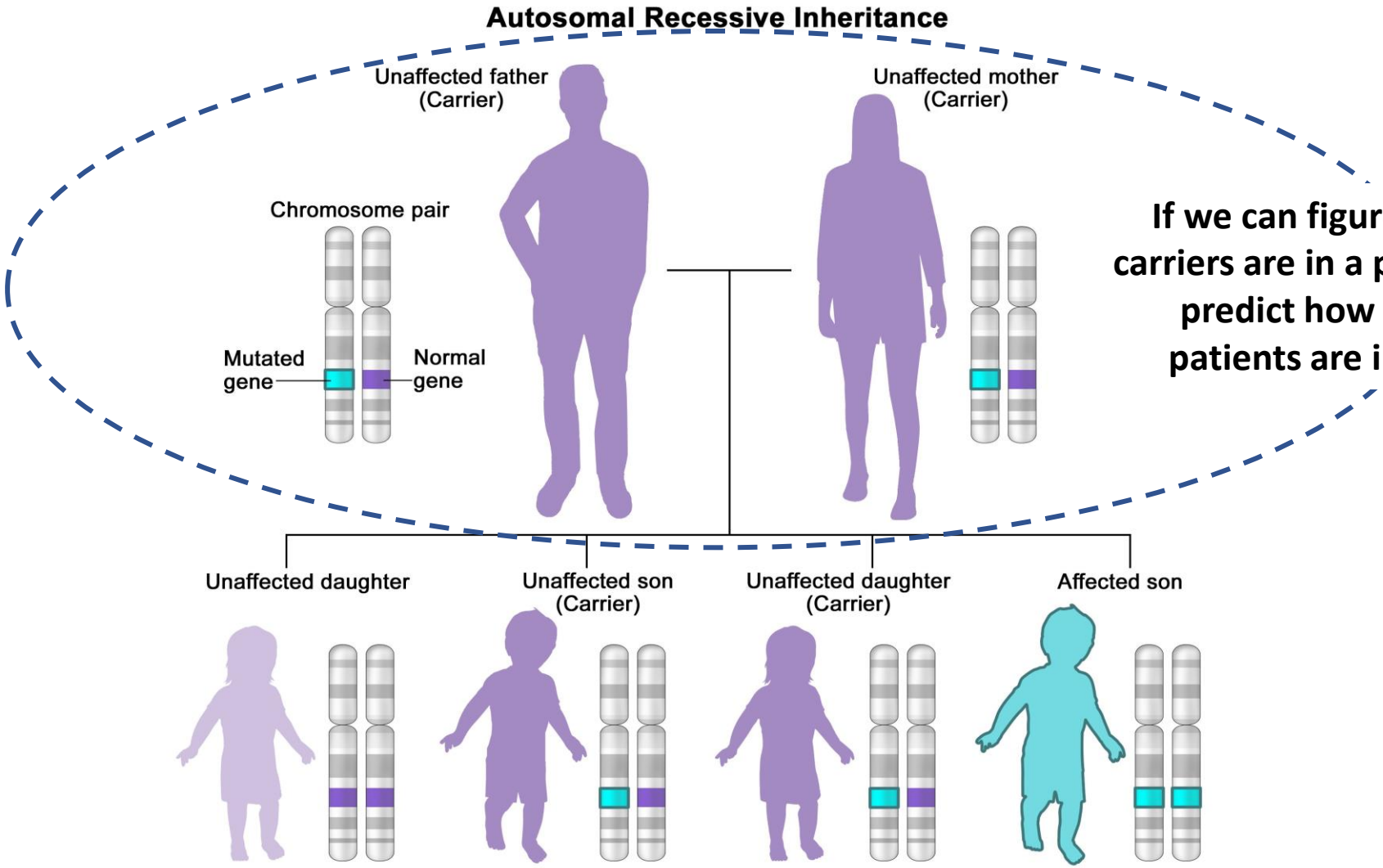


Miller et al. *Orphanet J Rare Dis* **16**, 265 (2021).

Is there another way to estimate incidence? Can we use basic genetic principles?

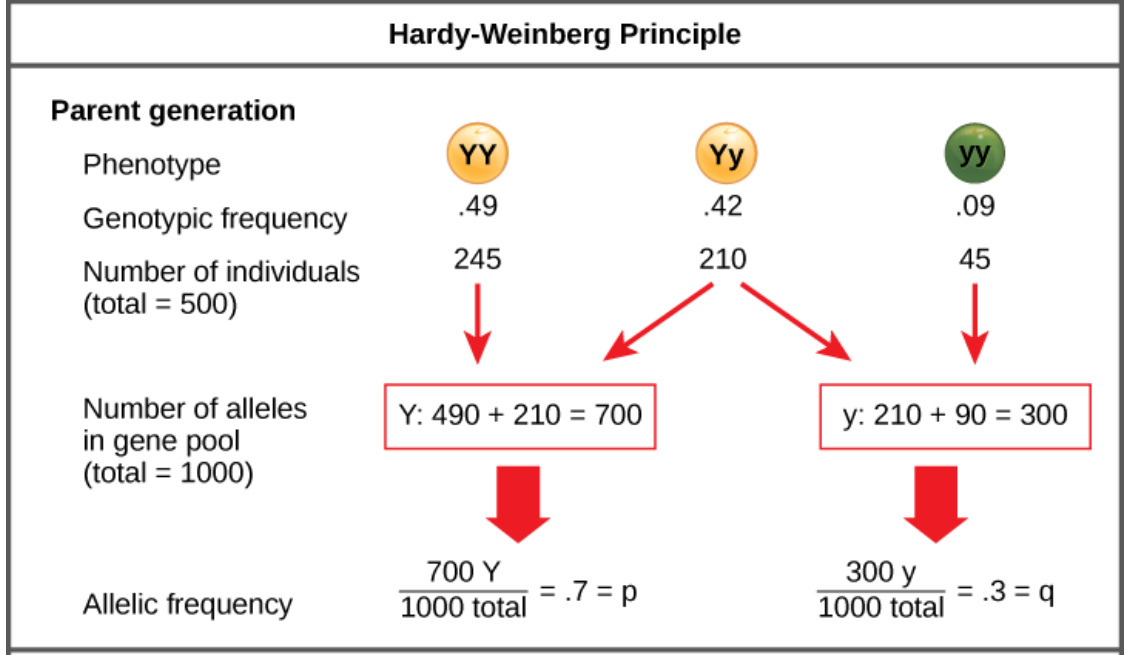


Is there another way to estimate incidence? Can we use basic genetic principles?



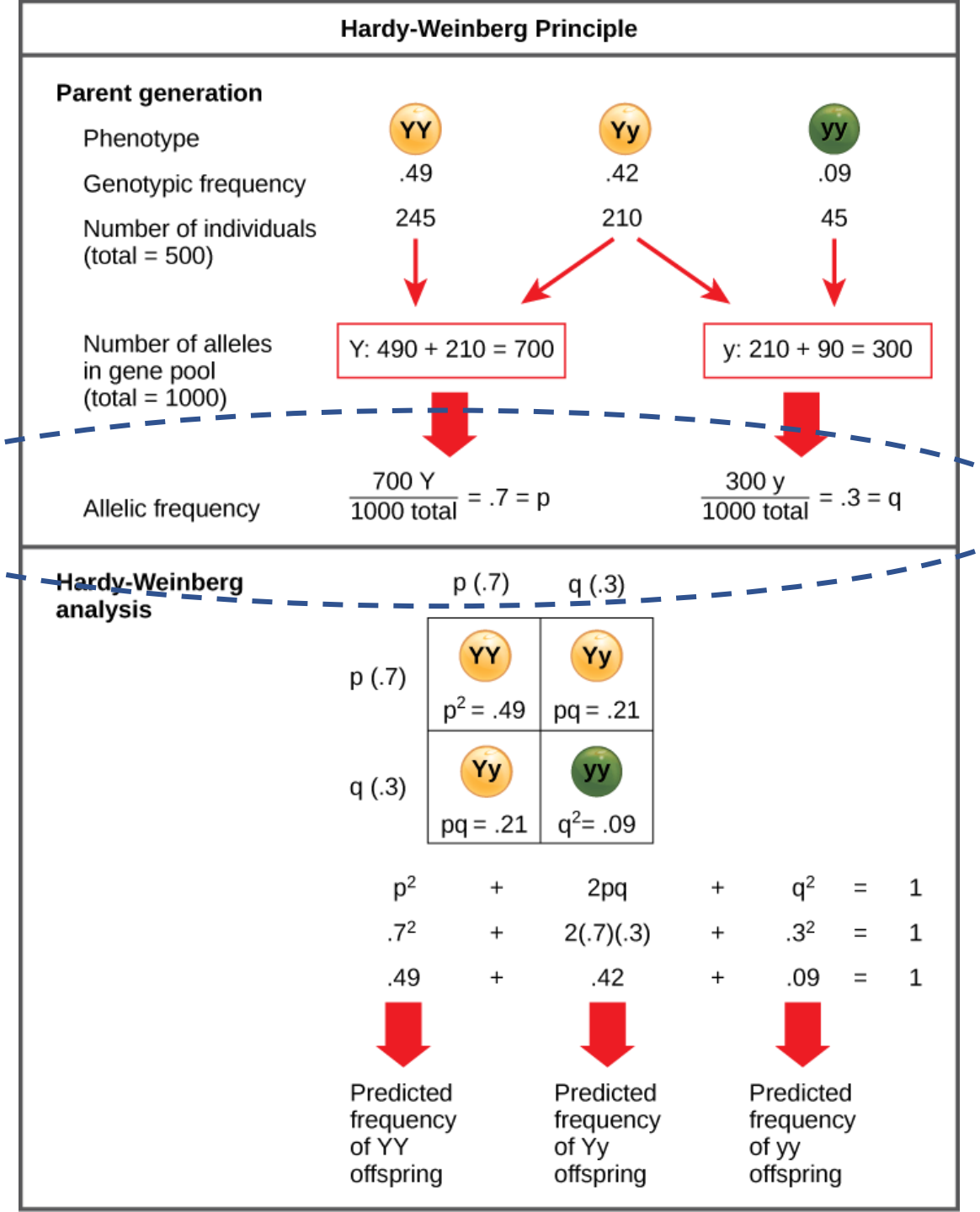
If we can figure out how many carriers are in a population, we can predict how many affected patients are in a population.

Is there another way to estimate incidence?
Can we use basic genetic principles?

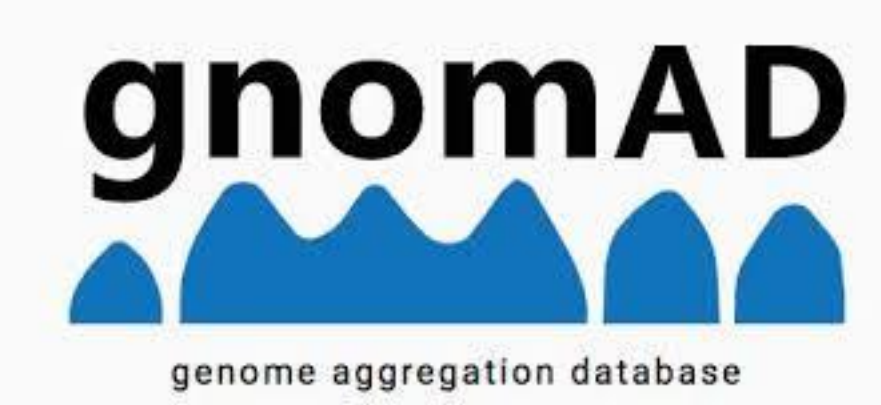


Is there another way to estimate incidence?
Can we use basic genetic principles?

Large genome sequencing datasets can help us figure out allele frequencies



Large databases of genomic data



Population	Description	Genomes
afr	African/African American	21,042
ami	Amish	450
amr	Latino/Admixed American	6,835
asj	Ashkenazi Jewish	1,662
eas	East Asian	1,567
fin	Finnish	5,244
nfe	Non-Finnish European	32,299
sas	South Asian	1,526
oth	Other (population not assigned)	1,077
Total		**71,702**

https://gnomad.broadinstitute.org

Examples of *DBT* variants in gnomAD

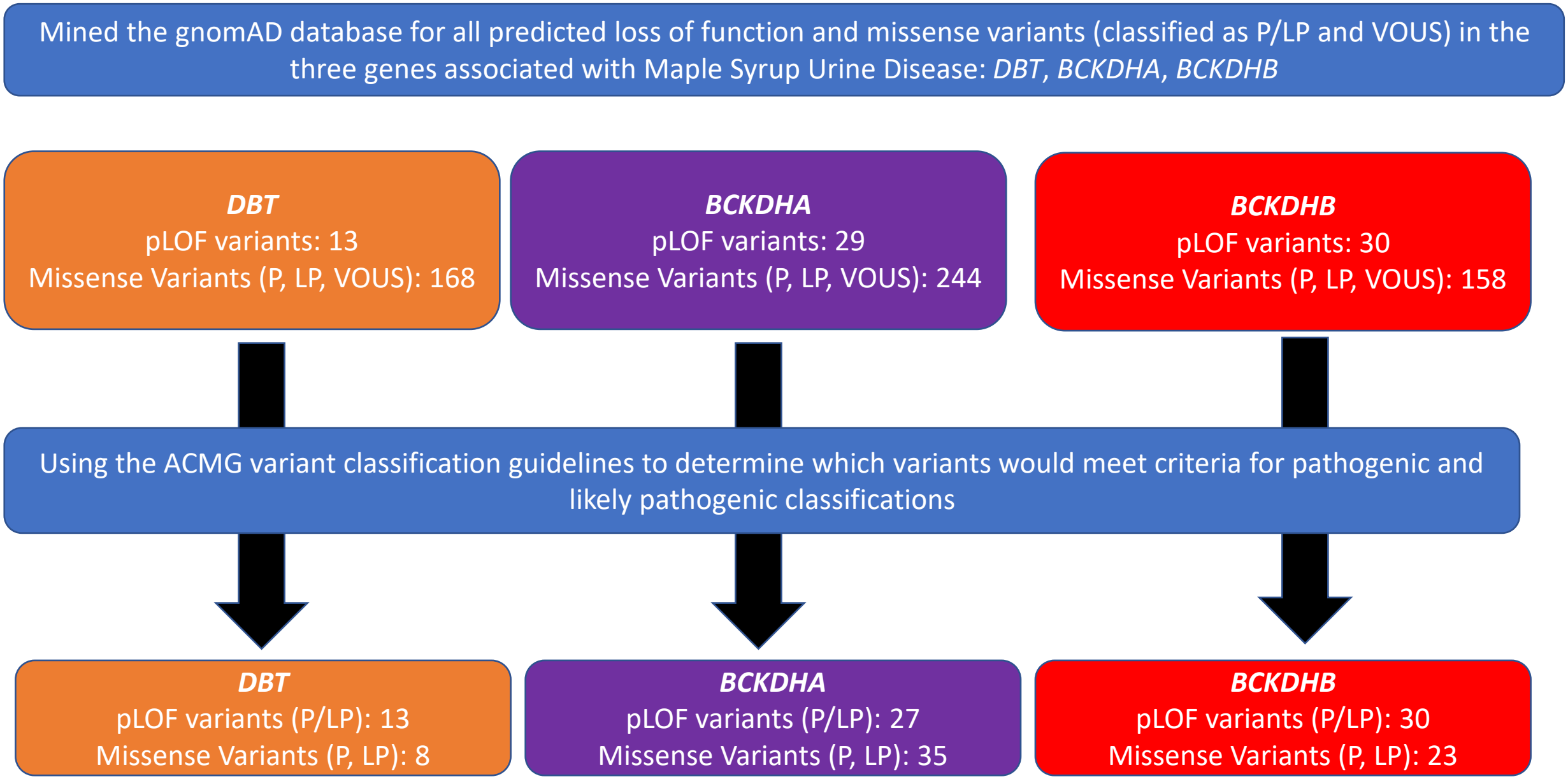
Variant ID	Clinical Significance	Allele Count	Allele Number	Allele Frequency	VEP Annotation
1-100196389-A-G		2	136152	1.47e-5	missense
1-100196395-C-G		1	125028	8.00e-6	missense
1-100196395-C-T	Conflicting interpretations of pathog...	2	125028	1.60e-5	missense
1-100196405-G-C		1	110086	9.08e-6	missense
1-100196410-A-G		1	130224	7.68e-6	missense
1-100196413-G-A	Pathogenic/Likely pathogenic	2	106800	1.87e-5	stop gained
1-100196425-AGAAAT-A		5	59838	8.36e-5	splice region

The next step is to determine which variants are truly disease-causing

	Benign		Pathogenic			
	Strong	Supporting	Supporting	Moderate	Strong	Very strong
Population data	MAF is too high for disorder BA1/BS1 **OR** observation in controls inconsistent with disease penetrance BS2			Absent in population databases PM2	Prevalence in affecteds statistically increased over controls PS4	
Computational and predictive data		Multiple lines of computational evidence suggest no impact on gene /gene product BP4 Missense in gene where only truncating cause disease BP1 Silent variant with non predicted splice impact BP7 In-frame indels in repeat w/out known function BP3	Multiple lines of computational evidence support a deleterious effect on the gene /gene product PP3	Novel missense change at an amino acid residue where a different pathogenic missense change has been seen before PM5 Protein length changing variant PM4	Same amino acid change as an established pathogenic variant PS1	Predicted null variant in a gene where LOF is a known mechanism of disease PVS1
Functional data	Well-established functional studies show no deleterious effect BS3		Missense in gene with low rate of benign missense variants and path. missenses common PP2	Mutational hot spot or well-studied functional domain without benign variation PM1	Well-established functional studies show a deleterious effect PS3	
Segregation data	Nonsegregation with disease BS4		Cosegregation with disease in multiple affected family members PP1	Increased segregation data →		
De novo data				De novo (without paternity & maternity confirmed) PM6	De novo (paternity and maternity confirmed) PS2	
Allelic data		Observed in *trans* with a dominant variant BP2 Observed in *cis* with a pathogenic variant BP2		For recessive disorders, detected in trans with a pathogenic variant PM3		
Other database		Reputable source w/out shared data = benign BP6	Reputable source = pathogenic PP5			
Other data		Found in case with an alternate cause BP5	Patient's phenotype or FH highly specific for gene PP4			

Diego Quintero, MS CGC

Kierstin Keller, MS CGC


Mined the gnomAD database for all predicted loss of function and missense variants (classified as P/LP and VOUS) in the three genes associated with Maple Syrup Urine Disease: *DBT*, *BCKDHA*, *BCKDHB*
DBT
pLOF variants: 13
Missense Variants (P, LP, VOUS): 168
BCKDHA
pLOF variants: 29
Missense Variants (P, LP, VOUS): 244
BCKDHB
pLOF variants: 30
Missense Variants (P, LP, VOUS): 158
Using the ACMG variant classification guidelines to determine which variants would meet criteria for pathogenic and likely pathogenic classifications
DBT
pLOF variants (P/LP): 13
Missense Variants (P, LP): 8
BCKDHA
pLOF variants (P/LP): 27
Missense Variants (P, LP): 35
BCKDHB
pLOF variants (P/LP): 30
Missense Variants (P, LP): 23

Determined variant carrier rate (VCR) of each likely pathogenic and pathogenic variant

VCR = (allele count – homozygote count) / (0.5 * allele number)

DBT variants: 21

BCKDHA variants: 62

BCKDHB variants: 53

Sum VCR to determine Gene Carrier Rate (GCR)

DBT **carrier rate**: 0.00191775

BCKDHA **carrier rate** : 0.001919908

BCKDHB **carrier rate** : 0.00288119

Calculate gene specific incidence using The Hardy-Weinberg Equation

DBT **Incidence**: 1 in 271,905

BCKDHA **Incidence:** 1 in 271,294

BCKDHB **Incidence:** 1 in 120,464

Combine to determine MSUD incidence

1 in 63, 837

AHRENS-NICKLAS LAB

A special thank you to all of the families, patients, and foundations that support our work!

Key Collaborators:

Brian White
Laura Adang
Lars Schlotawa
Kiran Musunuru
Bill Peranteau
Xiao Wang
Mohamad-Gabriel Alameh
Lindsey George
Eric Marsh
Beverly Davidson
Luis Tecedor
Elizabeth Bhoj
Stefano Rivella
Lucas Tricoli
Adeline Vanderver

MSUD FAMILY SUPPORT GROUP

NIH National Institute of Neurological Disorders and Stroke

Chan Zuckerberg Initiative

CGE Somatic Cell Genome Editing

the orphan disease center

BGTC AMP® Bespoke Gene Therapy Consortium FNIH

Children's Hospital of Philadelphia RESEARCH INSTITUTE | Intellectual and Developmental Disabilities Research Center

NIH Eunice Kennedy Shriver National Institute of Child Health and Human Development

MSD ACTION FOUNDATION

HRB Health Research Board