







OPEN ACCESS

Original research

# Prevalence of Fabry disease-causing variants in the UK Biobank

Mark Gilchrist <sup>1</sup>, Francesco Casanova,<sup>2</sup> Jess S Tyrrell <sup>2</sup>, Stuart Cannon <sup>1</sup>, Andrew R Wood,<sup>3</sup> Nicole Fife,<sup>1</sup> Katherine Young <sup>1</sup>, Richard A Oram,<sup>4</sup> Michael N Weedon<sup>4</sup>

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/jmg-2022-108523>).

<sup>1</sup>College of Medicine and Health, University of Exeter Medical School, Exeter, UK

<sup>2</sup>College of Medicine and Health, University of Exeter, Exeter, UK

<sup>3</sup>Genetics of Complex Traits, University of Exeter Medical School, Exeter, UK

<sup>4</sup>Institute of Biomedical and Clinical Science, University of Exeter Medical School, Exeter, UK

## Correspondence to

Dr Mark Gilchrist, University of Exeter Medical School, Exeter, UK; [m.gilchrist@exeter.ac.uk](mailto:m.gilchrist@exeter.ac.uk)

Received 8 March 2022

Accepted 15 July 2022

Published Online First 17 August 2022

## ABSTRACT

**Background** Fabry disease is an X-linked lysosomal storage disorder resulting from deficiency of the alpha-galactosidase A enzyme leading to accumulation of globotriaosylceramide in multiple organ sites with prominent cardiovascular and renal involvement. Global prevalence estimates of Fabry disease based on clinical ascertainment range from 1 in 40 000 to 1 in 170 000. We aimed to determine the prevalence of Fabry disease-causing variants in UK Biobank.

**Methods** We sought *GLA* gene variants in exome sequencing data from 200 643 individuals from UK Biobank. We used ACMG/AMP guidelines (American College of Medical Genetics/Association for Molecular Pathology) to classify pathogenicity and compared baseline biomarker data, hospital ICD-10 (International Classification of Diseases version-10) codes, general practitioner records and self-reported health data with those without pathogenic variants.

**Results** We identified 81 *GLA* coding variants. We identified eight likely pathogenic variants on the basis of being rare (<1/10 000 individuals) and either previously reported to cause Fabry disease, or being protein-truncating variants. Thirty-six individuals carried one of these variants. In the UK Biobank, the prevalence of likely pathogenic Fabry disease-causing variants is 1/5732 for late-onset disease-causing variants and 1/200 643 for variants causing classic Fabry disease.

**Conclusion** Fabry disease-causing *GLA* variants are more prevalent in an unselected population sample than the reported prevalence of Fabry disease. These are overwhelmingly variants associated with later onset. It is possible the prevalence of later-onset Fabry disease exceeds current estimates.

## INTRODUCTION

Fabry disease (OMIM 301500) is an X-linked lysosomal storage disease caused by variants in the *GLA* gene. *GLA* encodes the alpha-galactosidase A ( $\alpha$ -Gal A) enzyme responsible for degradation of glycosphingolipids, notably globotriaosylceramide and globotriaosylsphingosine.<sup>1</sup> Deficient activity of  $\alpha$ -Gal A leads to progressive accumulation of these substrates over time in every organ. There is considerable heterogeneity in clinical presentation in terms of pattern of organ involvement and age of onset between variants and even within families with the same variant.

## WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Fabry disease is typically thought of as a rare X-linked disease with a prevalence between 1 in 40 000 and 1 in 140 000. Newborn screening suggests a higher prevalence but relies on dried blood spot testing which is unreliable in affected females.

## WHAT THIS STUDY ADDS

⇒ This is one of the largest Fabry screening efforts ever undertaken and to our knowledge the largest based on exome sequencing and the largest in an unselected adult population including men and women. The prevalence of Fabry disease-causing variants is 1 in 5573 older adults. These are predominantly associated with late-onset disease or may demonstrate incomplete penetrance.

## HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ It is not known how to manage infants identified through newborn screening efforts. Our data offer significant insight into the likely clinical course of children identified with these late-onset variants which can inform clinical monitoring and treatment guidelines.

Though X-linked, Fabry disease affects females through random X-inactivation. The pattern of organ involvement and severity consequently can be comparable with that of classically affected males.<sup>2</sup> Prior to the advent of enzyme-replacement therapy, life expectancy in female ‘carriers’ of Fabry disease-causing variants was 15 years less than the general population.<sup>3</sup>

Global prevalence estimates of Fabry disease based on clinical ascertainment range from 1 in 40 000 to 1 in 170 000.<sup>4</sup> Newborn screening studies examining  $\alpha$ -Gal A activity on dried blood spots with subsequent confirmatory genetic testing have suggested a higher prevalence of Fabry disease, particularly variants associated with late-onset Fabry disease, than might be expected from studies of confirmed Fabry cases. Spada *et al* found a prevalence of known disease-causing pathogenic variants of 1 in 4600 in consecutively screened newborn Italian boys.<sup>5</sup> The prevalence of late onset to classic phenotype-associated pathogenic variants in this cohort was 7:1. Sawada *et al* screened 599 711



© Author(s) (or their employer(s)) 2023. Re-use permitted under CC BY. Published by BMJ.

**To cite:** Gilchrist M, Casanova F, Tyrrell JS, *et al*. *J Med Genet* 2023;**60**:391–396.

male and female newborns in western Japan and detected a prevalence of pathogenic variants of 1 in 11 854 following initial screening by measuring  $\alpha$ -Gal A activity on dried blood spot tests.<sup>6</sup> A US-based study arrived at a prevalence of 1 in 21 973.<sup>7</sup> The highest suggested prevalence was from a study of 171 977 Taiwanese newborns which arrived at a prevalence of approximately 1 in 1250. Much of this high prevalence was driven by a single variant, c.936+919G>A (IVS4+919G>A). This deep intronic variant accounted for 86% of those identified.<sup>8</sup> While it will not be known for many years what proportion of those individuals with pathogenic variants go on to develop clinical manifestations of Fabry disease, it is worth noting that 3 of the 57 individuals followed up as a result of Sawada *et al*'s study had developed manifestations of Fabry disease by the age of 5 years and commenced enzyme-replacement therapy by the age of 10 years.<sup>6</sup>

When assessed in cohorts referred for testing with a high clinical suspicion of Fabry disease (positive family history and/or characteristic phenotype),  $\alpha$ -Gal A activity in dried blood spot tests has a sensitivity reported to be 100% in males, though this is halved in females.<sup>9</sup> Sensitivity in lower-risk cohorts is not known. Efforts to identify patients with early manifestations based on a single pathological finding, such as early-onset stroke, reduced estimated glomerular filtration rate (eGFR) or the need for renal-replacement therapy, have notoriously low yields.<sup>10–13</sup>

The UK Biobank offers an unparalleled opportunity to determine the prevalence of Fabry disease-causing pathogenic variants in an unselected adult population. Unselected exome screening for Fabry disease-causing pathogenic variants has not been performed at this scale previously.

Using the UK Biobank, we sought to describe the prevalence of known Fabry disease-causing variants and associated phenotype. Further, we aimed to determine whether participants with characteristic clinical manifestations of later-onset Fabry disease would be predictive of the presence of known Fabry disease-causing variants.

## Methods

The UK Biobank recruited a sample of more than 500 000 individuals between the ages of 37 and 73 years from 2006 to 2010 across the UK.<sup>14</sup> At baseline, participants completed detailed questionnaires on health, sociodemographic and lifestyle factors. Blood and urine samples were obtained for genetic analysis along with a broad range of markers of health and disease.

Based on literature review, we created two cohorts of individuals with combinations of comorbidity which would be suggestive of a potential underlying diagnosis of Fabry disease using Hospital Episodes Statistics (HES) data. The first was men under the age of 60 years with chronic kidney disease (defined as eGFR <60 mL/min or albumin-creatinine ratio (ACR) >2.5 mg/mmol) and cardiac disease (including ischaemic heart disease, cardiomyopathy, dysrhythmias or cardiac failure) and hearing loss. The second group was men under the age of 60 years with chronic kidney disease (defined as eGFR <60 mL/min or ACR >2.5 mg/mmol) and cardiac disease and cerebrovascular disease. For a full list of HES codes used, please see online supplemental file.

Exome sequencing data from 200 643 individuals from the UK Biobank were examined for variants in the *GLA* gene. Likely pathogenic variants were searched for in ClinVar and Human Gene Mutation Database with subsequent review of supporting literature.

Variants are annotated by AlaMut batch software V.1.8 (Interactive Biosoftware, France) using the NM\_000169.2 transcript

for *GLA*. The pathogenic variants can be missense or protein truncating. We defined a protein-truncating variant (PTV) as a variant that is predicted to cause a stop gain, a frameshift or abolish a canonical splice site (–2 or +2bp from exon). In this study, we excluded PTVs in the last exon of each gene.

We reviewed all heterozygous missense and PTV in UK Biobank that were observed at minor allele frequency <0.001 in gnomAD V2 (N=141 456).<sup>15</sup> ACMG/AMP (American College of Medical Genetics/Association for Molecular Pathology) guidelines were used to classify missense variants as pathogenic/likely pathogenic.<sup>16</sup> All PTVs were classified as pathogenic.

Pathogenic variants (missense and PTVs) were reviewed in Integrative Genomics Viewer to remove false positive variants. The variants considered to be excellent quality<sup>17</sup> and deemed to be high confidence by LOFTEE<sup>15</sup> (only for PTV) were included in the analysis.

Data were analysed using Stata V.16 (Statacorp, Texas, USA). Comparisons between the group with Fabry-causing variants and the remainder of the Biobank cohort with exome sequencing data were performed using Fisher's exact test.

HES data for each individual were examined for possible Fabry-associated phenotypes mirroring the risk cohort search. In addition, primary care data were searched for the terms associated with classic as well as late-onset Fabry phenotypes (see online supplemental data).

Self-reported data for cardiovascular disease, hearing loss and family history of cardiovascular disease from questionnaires at the assessment centre were compared in those with and without Fabry disease-causing variants.

Kidney pathology was assessed by CKD-EPI eGFR and urinary ACR.

## RESULTS

### Pathogenic variant prevalence

We identified 81 coding variants in the *GLA* gene. Eight of these variants were rare (<1 in 10 000 individuals) and either previously reported to cause Fabry disease or PTVs (table 1). Thirty-six individuals, (15 male, 21 female) carried one of these *GLA* variants (see table 2 for variant details). Thirty-three of the 36 individuals were recorded as being of European ancestry. The cardiac predominant phenotype causing c.644A>G (p.N215S) variant (18 of 36) was the most common variant. Six of 36 had the c.1087C>T (R363C) variant, 4 of 36 had c.1067G>A (p.R356Q). The c.335G>A (p.R112H), c.902G>A (p.R301Q) and c.593T>C (p.I198T) variants were each identified in two individuals. There were single individuals with c.695T>C (p.I232T) and c.718\_719del (p.K240fs) variants.

In the UK Biobank, the overall prevalence of Fabry disease-causing variants is 1 in 5573. Only one variant, c.718\_719del (p.K240fs), is consistently associated with the classic phenotype, 1 in 200 643. All other variants reported here are commonly associated with later onset or variable penetrance with a prevalence of 1 in 5732. The c.644A>G (p.N215S) variant alone has a prevalence of 1 in 11 147.

### CLINICAL CHARACTERISTICS

One male individual with c.902G>A (p.R301Q), one male individual with the c.644A>G (p.N215S) variant and one female individual with c.718\_719del (p.K240fs) variants had HES codes compatible with a known diagnosis of Fabry disease. The remaining participants did not have evidence of an existing diagnosis of Fabry disease in either HES or primary care data. No

**Table 1** Disease-causing variants in the UK Biobank with associated disease types and reported ranges of enzyme activity from clinical populations and in vitro studies

Variant	Protein change	N in biobank	gnomAD freq (MAF)	Exon	Disease type	Enzyme activity
						(% WT)
c.644A>G	p.N215S	18 (1/11 147)	1/183 422	5	Late onset	5–39.5 (18–21)
c.1087C>T	p.R363C	6 (1/33 461)	2/183 376	7	Late onset	7.5 (21)
c.1067G>A	p.R356Q	4 (1/50 161)	2/183 292	7	Late onset	12.9–86.9 (8, 19–22)
c.335G>A	p.R112H	2 (1/100 322)	3/183 240	2	Late onset	<5 (19)
c.902G>A	p.R301Q	2 (1/100 322)	Not seen	6	Both described	3–5.6 (20–22)
c.593T>C	p.I198T	2 (1/100 322)	2/18 1907	4	Late onset	10–64.7 (20, 21, 25)
c.718_719del	p.K240fs	1 (1/200 643)	Not seen	5	Classic	0.5–2.5 (24, 26)
c.695T>C	p.I232T	1 (1/200 643)	Not seen	5	Late onset	7.2–15 (19–22, 27)

MAF, minor allele frequency; WT, wild type.

individuals were identified with a HES or primary care code of Fabry disease who did not have a pathogenic variant.

Age at cohort entry did not differ by variant status at 53.1±8.7 and 55.3±8.4 years vs 56.7±8.2 and 56.3±8.0 years for men and women with and without variants, respectively ( $p=0.13$  and  $p=0.60$ , two-sample t-test with unequal variance) (see table 2). The sex split in those with and without Fabry disease-causing variants was not different at 58.3% female and 41.7% in those with variants compared with 55.1% female and 44.9% in the remainder of the biobank cohort ( $p=0.69$ , Pearson's  $\chi^2$ ). Kidney function was not different for men with Fabry disease-causing variants compared with those without with CKD-EPI eGFRs of 93±14 vs 92±13 mL/min/1.73 m<sup>2</sup> ( $p=0.91$ , two-sample t-test with unequal variance), nor for women 89±15 vs 91±13 mL/min/1.73 m<sup>2</sup>, respectively ( $p=0.50$ , two-sample t-test with unequal variance).

The likelihood of microalbuminuria (ACR >3.0 mg/mmol) was higher in men with Fabry disease-causing pathogenic variants (5 of 14) than for the background population (4935 of 82 865, OR 9.33 (95% CI 2.45 to 31.01),  $p=0.0007$ ), but not for women (2 of 19 vs 5185 of 101 593). Participants with Fabry disease-causing variants were more likely to have a documented family history of cardiovascular disease in HES or reported at the assessment centre than the biobank cohort as a whole (13.9% vs 4.7%, OR 3.31 (95% CI 1.004 to 8.59),  $p=0.025$ ). For participants with Fabry disease-causing variants, the prevalence of cardiovascular disease was 27.8% compared with 15.0% for the overall biobank cohort (OR 2.19 (95% CI

0.94 to 4.69),  $p=0.056$ , Fisher's exact test). When split by sex, 40% of men had a history of cardiovascular disease compared with 20% for the overall biobank cohort (OR 2.60 (95% CI 0.76 to 8.19),  $p=0.099$ ), and 19% of women had a history of cardiovascular disease compared with 10% for the overall biobank cohort (OR 2.01 (95% CI 0.49 to 6.17)). There were no differences in the prevalence of either medically recognised or self-reported hearing loss. Acroparasthesia, angiokeratoma, hypohidrosis and corneal verticillata are phenotypic features highly suggestive of classic Fabry disease though do occur in some individuals with late-onset variants.<sup>18</sup> These manifestations were not noted in our cohort in either HES or primary care records.

Considering the c.644A>G (p.N215S) variant in isolation, one individual had a known diagnosis of Fabry disease and characteristic phenotypic features. Three individuals had cardiac pathology (two atherosclerotic heart disease, one aortic stenosis) which may have been related to Fabry disease, but also had presence of traditional cardiovascular risk factors (hypertension and type 2 diabetes). One woman had microalbuminuria in the absence of other relevant pathology. The remaining 14 individuals with the N215S variant did not have comorbidity suggestive of clinically evident Fabry disease.

The c.1067G>A (p.R356Q) variant has been associated with classic Fabry disease in some published cases.<sup>19</sup> None of the four individuals (three women, one man) with this variant has a diagnosis of Fabry disease nor clinical features suggestive of Fabry disease.

**Table 2** Baseline characteristics of participants with Fabry disease-causing variants

	Combined	Male (M)	Female (F)	Background*	Significance
n (%)	36 (100)	15 (42)	21 (58)	44.9 M, 55.1 F	$p=0.69$ †
Age (years), mean (SD)	54.8±8.4	53.1±8.7	55.3±8.4	56.7±8.2M, 56.3±8.0 F	$p=0.13$ M, $p=0.60$ F‡
Diagnosed Fabry disease	3	2	1	0	NT
Cardiovascular disease, n (%)	10 (27.8)	6 (40)	4 (19)	18 221 (20) M, 11 557 (10) F	$p=0.099$ M, $p=0.27$ F
Hearing impairment, n (%)	1 (2.8)	1	0	3773 (1.88)	NT
CKD-EPI eGFR mL/min/1.73 m <sup>2</sup> §	90±15	93±14	89±15	92±14 M, 91±13 F	$p=0.91$ M, $p=0.50$ F‡
Microalbuminuria (n=35)¶, n (%)	7 (20)	5 (35.7)	2 (9.5)	4935 (5.6) M, 5185 (4.9) F	$p=0.0007$ M, $p=0.27$ F

P values are for combined men and women with Fabry-causing variants versus the background cohort unless specified otherwise using Fisher's exact test.

\*n=200 596 (194 578 for ACR data).

†Pearson's  $\chi^2$ .

‡Independent samples t-test.

§Creatinine missing for one woman.

¶ACR missing for one man.

\*\*

ACR, albumin-creatinine ratio; CKD-EPI eGFR, Chronic Kidney Disease Epidemiology Collaboration estimated glomerular filtration rate; NT, not tested.



## RISK COHORTS

Using the first phenotype-generated high-risk cohort, we identified 470 individuals meeting the criteria. From the second risk cohort, 488 individuals with exome sequencing data were identified meeting the criteria. Only the individual with a pre-existing diagnosis of Fabry disease with the c.902G>A (R301Q) variant was identified through this approach and was captured in both risk cohorts.

## DISCUSSION

We have shown that the overall prevalence of Fabry disease-causing variants in the UK Biobank is 1 in 5573 with the majority being those associated with a late-onset phenotype. This is similar to and consistent with the 1 in 4600 identified following initial screening by enzyme activity in dried blood spot tests in males in Spada *et al*'s work.<sup>5</sup> Our findings are also consistent with the figure of 1 in 11854 from Sawada *et al*'s cohort once the reduced sensitivity of blood spot testing in women is taken into account. It further reinforces the lack of sensitivity of dried blood spot tests in women.<sup>20</sup>

The c.644A>G (p.N215S) variant alone has a prevalence of 1 in 11147. Evidence for pathogenicity of this variant comes from multiple cohort studies and is associated with later-onset predominantly cardiac manifestations.<sup>18–21</sup> There was considerable heterogeneity in both severity and age of onset of clinically apparent disease in these studies. Our findings support this observation of heterogeneity and highlights that ascertainment impacts the phenotype associated with a variant. We found variation in the clinical phenotype of the individuals with this variant as ascertained from general practitioner (GP) records, self-report and ICD-10 (International Classification of Diseases version-10) codes. It is possible that this reflects poor sensitivity for finding Fabry disease's features when using HES data and GP records, alternatively this may reflect variable penetrance. This variant is an example of why the relatively high frequency of Fabry disease's variants we found, compared with studies based on clinical features of Fabry disease, may include individuals with mild Fabry disease that would have not previously been diagnosed and pathogenic variants with incomplete penetrance. Further evidence for this could be found in future follow-up studies of individuals identified with Fabry disease's variants in screening studies, and in other population biobanks.

We attempted to identify people who were likely to have Fabry-causing pathogenic variants by identifying individuals with an aggregation of comorbidity which would be highly suggestive of underlying Fabry disease. This approach failed to identify any individuals who did not have a prior diagnosis of Fabry disease who did have a likely pathogenic variant. This finding mirrors the difficulties in identifying individuals with Fabry disease from single clinical characteristics.<sup>10–12</sup> It is possible that different criteria would provide greater sensitivity. Our results suggest it will be difficult to apply screening criteria to either hospital or GP records with enough sensitivity or specificity to find undiagnosed patients with Fabry disease. It is possible that more detailed analysis of hospital or GP records, or integration of other imaging or biomarker data may in the future be able to. This further highlights the need to develop robust strategies for the identification of people with Fabry disease from clinical cohorts.

One key strength of the current work is that by using exome sequencing, at-risk females will be identified. Like other X-linked conditions, Fabry is under-recognised in females because of lack of sensitivity of existing clinical tests, and diagnosis is more

likely to occur when there is an affected male family member.<sup>22</sup> Newborn screening studies typically use dried blood spot testing for  $\alpha$ -Gal A activity,<sup>5–7</sup> but this is less sensitive in females<sup>20</sup> and therefore birth studies using this method are likely to underestimate frequency of Fabry disease's variants in a population.

The majority of participants in whom we identified Fabry-causing variants had neither a known diagnosis of Fabry disease nor pathology highly suggestive of underlying Fabry disease. There are several possible explanations for this: 58% of individuals identified with a Fabry disease-causing pathogenic variant are female in this cohort and females are more likely to have delays or missed diagnoses of Fabry where it is present.<sup>22</sup> It is also likely that females will have a milder or later-onset phenotype for a given pathogenic variant,<sup>2</sup> although we were not able to test for differences in phenotype between males and females due to the number of variants we identified in this study. Additionally, the age of UK Biobank participants is younger than the typical age at first presentation with a Fabry indicator presentation, particularly for the variants associated with late onset, that make up the majority of variants we found.<sup>18</sup> Primary care data are available for only half of the cohort, thus some individuals with pathology attributable to Fabry disease may not be identified. While we have used ACMG/AMP guidelines for classification of pathogenicity of variants,<sup>17</sup> it remains possible that one or more of our variants are not truly pathogenic.

Previous analyses of UK Biobank data have shown that disease penetrance from a given pathogenic variant may be substantially lower than estimated from studies of cohorts with a known genetic diagnosis.<sup>23</sup> A key strength of the current work compared with newborn screening studies is that our cohort has had a median of 54 years in which to develop a Fabry-associated phenotype. From this, we see that the majority of those who have a late-onset associated variant will not develop Fabry-associated pathology by this time. Nevertheless, the difference in ACR in males, family history of cardiovascular disease and numerical differences in cardiovascular disease compared with the background biobank cohort suggests the possibility of undiagnosed Fabry disease in some of our individuals with variants.

The variants identified here were predominantly those which are known to be later-onset variants or for which there is conflicting evidence supporting classification as classic or late onset. Participants in the UK Biobank were between 40 and 69 years of age at entry into the cohort.<sup>14</sup> The age of participants favours later-onset variants and less severe phenotypes as individuals with classic disease-causing variants being less likely to survive to an age which would have permitted study entry. It is also worth noting that UK Biobank participants have a lower prevalence of cardiovascular disease, renal disease and a longer life expectancy than the background population.<sup>24</sup> As such, those with disease manifestations from Fabry disease including those with later-onset variants may have been less likely to enter the study. If this were the case, it would only be likely to result in an underestimation of Fabry disease's variant frequency, and therefore our estimate of frequency could be considered conservatively.

With the difference in the prevalence of family history of cardiovascular disease between those with and without Fabry disease-causing variants, the possibility of clinically evident Fabry disease in the families of the individuals identified here with Fabry disease-causing variants must be considered. Similarly, the higher likelihood of microalbuminuria reported here in men may be an early marker of renal involvement in Fabry disease.

Earlier work has produced conflicting assessments of the phenotypes associated with some variants in terms of late-onset or classic disease. The c.1067G>A (p.R356Q) variant was associated with the classic phenotype in the finding of an Italian case study.<sup>19</sup> With a reported  $\alpha$ -Gal A enzyme activity of 15% of wild type<sup>25</sup> and the absence of confirmed Fabry or Fabry phenotype in the present work, this would favour classification as a late-onset or reduced penetrance phenotype.

The principle limitation to the current work is that participants in the UK Biobank cannot be recontacted for further study. Consequently, it is not possible to confirm a diagnosis of Fabry disease with measurement of enzyme activity or relevant substrates. Our findings may not be generalisable outside the populations of predominantly European descent. The c.644A>G (p.N215S) variant is known to be the most common late-onset variant in populations of European descent, though it has been described in people of Japanese and Chinese descent.<sup>6,26</sup> In addition, we cannot exclude the presence of intronic variants such as the c.936+919G>A variant, which has been frequently found in Asian populations.<sup>8</sup>

The optimum method for screening for Fabry disease at population level is not known. Screening based on clinical characteristics has previously been shown to have a low diagnostic yield.<sup>10–12</sup> We were unable to identify new cases by such an approach in the UK Biobank cohort. Genetic screening of populations such as presented here may identify people with Fabry disease-causing pathogenic variants who do not go on to develop the disease. In the work by Spada *et al* and Sawada *et al*, enzyme activity was assayed on a dried blood spot test before confirmatory genetic testing.<sup>5,6</sup> While this may increase the likelihood that those identified will develop the disease, it is of limited utility for females.

The clinical consequences of Fabry disease can be devastating with substantial reductions in quality of life and life expectancy as a result of major cardiovascular events, neurological involvement and kidney disease. With the interval between identification of those at risk and onset of clinical manifestations, where they occur, typically being several decades later, significant questions remain around who to screen and when to treat to avoid these consequences.

Fabry disease-causing *GLA* variants are far more prevalent in an unselected population sample than would be expected from the reported prevalence of Fabry disease. This may be because some of the reported variants have reduced pathogenicity or penetrance, or are later-onset causing variants with variable penetrance. Our work suggests that the true prevalence of later-onset Fabry disease is substantially higher than current estimates. As genetic screening becomes more widely used, it will be essential to understand the penetrance of these variants to inform clinical monitoring strategies and identify those who will benefit from treatment.

**Twitter** Mark Gilchrist @gilchristN03

**Acknowledgements** This research has been conducted using the UK Biobank resource. This work was carried out under UK Biobank project number 49847.

**Contributors** Conceptualisation—MG, RAO, MNW. Data curation—ARW, SC, FC, JST, MNW. Formal analysis—MG, SC, NF, KY, ARW, FC. Supervision—MNW. Writing (original draft)—MG. Writing (review and editing)—all authors. MG acts as guarantor.

**Funding** This work was supported by the MRC (grant no. MR/T00200X/1). RAO is funded by a Diabetes UK Harry Keen Fellowship (16/0005529).

**Competing interests** MG has received support for conference attendance from Sanofi Genzyme. He has also provided professional services for and been compensated by AstraZeneca, Lilly and Napp Pharmaceuticals. No other authors have relevant conflicts of interest to declare.

**Patient consent for publication** Not required.

**Ethics approval statement:** The UK Biobank resource was approved by the UK Biobank Research Ethics Committee and received ethical approval from the UK National Health Service's National Research Ethics Service (ref 11/NW/0382). All participants provided written informed consent to participate.<sup>14</sup>

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data may be obtained from a third party and are not publicly available. Data are available through the UK Biobank via <https://www.ukbiobank.ac.uk/enable-your-research/apply-for-access>.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: <https://creativecommons.org/licenses/by/4.0/>.

#### ORCID iDs

Mark Gilchrist <http://orcid.org/0000-0002-0284-6048>

Jess S Tyrrell <http://orcid.org/0000-0002-9256-6065>

Stuart Cannon <http://orcid.org/0000-0002-5108-7615>

Katherine Young <http://orcid.org/0000-0003-2570-3864>

#### REFERENCES

- 1 Germain DP. Fabry disease. *Orphanet J Rare Dis* 2010;5:30. doi:10.1186/1750-1172-5-30
- 2 Guffon N. Clinical presentation in female patients with Fabry disease. *J Med Genet* 2003;40:38e–38. doi:10.1136/jmg.40.4.e38
- 3 MacDermot KD, Holmes A, Miners AH. Anderson-Fabry disease: clinical manifestations and impact of disease in a cohort of 60 obligate carrier females. *J Med Genet* 2001;38:769–75. doi:10.1136/jmg.38.11.769
- 4 Biegstraaten M, Arngrimsson R, Barbey F, Boks L, Cecchi F, Deegan PB, Feldt-Rasmussen U, Geberhiwot T, Germain DP, Hendriks C, Hughes DA, Kantola I, Karabul N, Lavery C, Linthorst GE, Mehta A, van de Mheen E, Oliveira JP, Parini R, Ramaswami U, Rudnicki M, Serra A, Sommer C, Sunder-Plassmann G, Svarstad E, Sweeb A, Terryn W, Tylki-Szymanska A, Tøndel C, Vujkovic B, Weidemann F, Wijburg FA, Woolfson P, Hollak CEM. Recommendations for initiation and cessation of enzyme replacement therapy in patients with Fabry disease: the European Fabry Working group consensus document. *Orphanet J Rare Dis* 2015;10:36. doi:10.1186/s13023-015-0253-6
- 5 Spada M, Pagliardini S, Yasuda M, Tükel T, Thiagarajan G, Sakuraba H, Ponzone A, Desnick RJ. High incidence of later-onset Fabry disease revealed by newborn screening. *Am J Hum Genet* 2006;79:31–40. doi:10.1086/504601
- 6 Sawada T, Kido J, Yoshida S, Sugawara K, Momosaki K, Inoue T, Tajima G, Sawada H, Mastumoto S, Endo F, Hirose S, Nakamura K. Newborn screening for Fabry disease in the Western region of Japan. *Mol Genet Metab Rep* 2020;22. doi:10.1016/j.ymgmr.2019.100562
- 7 Burton BK, Charrow J, Hoganson GE, Waggoner D, Tinkle B, Braddock SR, Schneider M, Grange DK, Nash C, Shryock H, Barnett R, Shao R, Basheeruddin K, Dizikes G. Newborn screening for lysosomal storage disorders in Illinois: the initial 15-month experience. *J Pediatr* 2017;190:130–5. doi:10.1016/j.jpeds.2017.06.048
- 8 Hwu W-L, Chien Y-H, Lee N-C, Chiang S-C, Dobrovolsky R, Huang A-C, Yeh H-Y, Chao M-C, Lin S-J, Kitagawa T, Desnick RJ, Hsu L-W. Newborn screening for Fabry disease in Taiwan reveals a high incidence of the later-onset *GLA* mutation c.936+919G>A (IVS4+919G>A). *Hum Mutat* 2009;30:1397–405. doi:10.1002/humu.21074
- 9 Stiles AR, Zhang H, Dai J, McCaw P, Beasley J, Rehder C, Koerber DD, McDonald M, Bali DS, Young SP. A comprehensive testing algorithm for the diagnosis of Fabry disease in males and females. *Mol Genet Metab* 2020;130:209–14. doi:10.1016/j.ymgme.2020.04.006
- 10 Reynolds TM, Tylee KL, Booth KL, Wierzbicki AS, PATHFINDER Project Collaboration group. Identification of patients with Fabry disease using routine pathology results: PATHFINDER (eGFR) study. *Int J Clin Pract* 2021;75:e13672. doi:10.1111/ijcp.13672
- 11 Vigneau C, Germain DP, Larmet D, Jabbour F, Hourmant M, SNOUFY Investigators Group. Screening for Fabry disease in male patients with end-stage renal disease in Western France. *Nephrol Ther* 2021;17:180–4. doi:10.1016/j.nephro.2021.03.002
- 12 Kinoshita N, Hosomi N, Matsushima H, Nakamori M, Yagita Y, Yamawaki T, Torii T, Kitamura T, Sueda Y, Shimomura R, Araki M, Nezu T, Aoki S, Ishii S, Maruyama H, Matsumoto M, Maruyama H. Screening for Fabry disease in Japanese patients with

- young-onset stroke by measuring  $\alpha$ -galactosidase A and globotriaosylsphingosine. *Journal of Stroke and Cerebrovascular Diseases* 2018;27:3563–9. doi:10.1016/j.jstrokecerebrovasdis.2018.08.025
- 13 Rolfes A, Fazekas F, Grittner U, Dichgans M, Martus P, Holzhausen M, Böttcher T, Heuschmann PU, Tatlisumak T, Tanislav C, Jungehulsing GJ, Giese A-K, Putaala J, Huber R, Bodechtel U, Lichy C, Enzinger C, Schmidt R, Hennerici MG, Kaps M, Kessler C, Lackner K, Paschke E, Meyer W, Mascher H, Riess O, Kolodny E, Norrving B, Stroke in Young Fabry Patients (sifap) Investigators. Acute cerebrovascular disease in the young: the stroke in young Fabry patients study. *Stroke* 2013;44–9. doi:10.1161/STROKEAHA.112.663708
  - 14 Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A, Vukcevic D, Delaneau O, O'Connell J, Cortes A, Welsh S, Young A, Effingham M, McVean G, Leslie S, Allen N, Donnelly P, Marchini J. The UK Biobank resource with deep phenotyping and genomic data. *Nature* 2018;562:203–9. doi:10.1038/s41586-018-0579-z
  - 15 Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alfoldi J, Wang Q, Collins RL, Laricchia KM, Ganna A, Birnbaum DP, Gauthier LD, Brand H, Solomonson M, Watts NA, Rhodes D, Singer-Berk M, England EM, Seaby EG, Kosmicki JA, Walters RK, Tashman K, Farjoun Y, Banks E, Poterba T, Wang A, Seed C, Whiffin N, Chong JX, Samocha KE, Pierce-Hoffman E, Zappala Z, O'Donnell-Luria AH, Minikel EV, Weisburd B, Lek M, Ware JS, Vittal C, Armean IM, Bergelson L, Cibulskis K, Connolly KM, Covarrubias M, Donnelly S, Ferriera S, Gabriel S, Gentry J, Gupta N, Jeandet T, Kaplan D, Llanwarne C, Munshi R, Novod S, Petrillo N, Roazen D, Ruano-Rubio V, Saltzman A, Schleicher M, Soto J, Tibbetts K, Tolonen C, Wade G, Talkowski ME. Genome aggregation database C, Neale BM, Daly MJ, MacArthur dG. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* 2020:581–43.
  - 16 Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL, ACMG Laboratory Quality Assurance Committee. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of medical genetics and genomics and the association for molecular pathology. *Genet Med* 2015;17:405–24. doi:10.1038/gim.2015.30
  - 17 Robinson JT, Thorvaldsdóttir H, Wenger AM, Zehir A, Mesirov JP. Variant review with the integrative genomics Viewer. *Cancer Res* 2017;77:e31–4. doi:10.1158/0008-5472.CAN-17-0337
  - 18 Lavalle L, Thomas AS, Beaton B, Ebrahim H, Reed M, Ramaswami U, Elliott P, Mehta AB, Hughes DA. Phenotype and biochemical heterogeneity in late onset Fabry disease defined by N215S mutation. *PLoS One* 2018;13:e0193550. doi:10.1371/journal.pone.0193550
  - 19 Duro G, Zizzo C, Cammarata G, Burlina A, Burlina A, Polo G, Scalia S, Oliveri R, Sciarrino S, Francofonte D, Alessandro R, Pisani A, Palladino G, Napoletano R, Tenuta M, Masarone D, Limongelli G, Riccio E, Frustaci A, Chimenti C, Ferri C, Pieruzzi F, Pieroni M, Spada M, Castana C, Caserta M, Monte I, Rodolico MS, Feriozzi S, Battaglia Y, Amico L, Losi MA, Autore C, Lombardi M, Zoccali C, Testa A, Postorino M, Mignani R, Zachara E, Giordano A, Colomba P. Mutations in the *GLA* gene and LysoGb3: is it really Anderson-Fabry disease? *Int J Mol Sci* 2018;19:3726. doi:10.3390/ijms19123726
  - 20 Balendran S, Oliva P, Sansen S, Mechtler TP, Streubel B, Cobos PN, Lukacs Z, Kasper DC. Diagnostic strategy for females suspected of Fabry disease. *Clin Genet* 2020;97:655–60. doi:10.1111/cge.13694
  - 21 Germain DP, Brand E, Burlina A, Cecchi F, Garman SC, Kempf J, Laney DA, Linhart A, Maródi L, Nicholls K, Ortiz A, Pieruzzi F, Shankar SP, Waldek S, Wanner C, Jovanovic A. Phenotypic characteristics of the p.Asn215Ser (p.N215S) *GLA* mutation in male and female patients with Fabry disease: A multicenter Fabry Registry study. *Mol Genet Genomic Med* 2018;6:492–503. doi:10.1002/mgg3.389
  - 22 Eng CM, Fletcher J, Wilcox WR, Waldek S, Scott CR, Sillence DO, Breunig F, Charrow J, Germain DP, Nicholls K, Banikazemi M. Fabry disease: baseline medical characteristics of a cohort of 1765 males and females in the Fabry registry. *J Inher Metab Dis* 2007;30:184–92. doi:10.1007/s10545-007-0521-2
  - 23 Wright CF, West B, Tuke M, Jones SE, Patel K, Laver TW, Beaumont RN, Tyrrell J, Wood AR, Frayling TM, Hattersley AT, Weedon MN. Assessing the pathogenicity, penetrance, and expressivity of putative disease-causing variants in a population setting. *Am J Hum Genet* 2019;104:275–86. doi:10.1016/j.ajhg.2018.12.015
  - 24 Fry A, Littlejohns TJ, Sudlow C, Doherty N, Adamska L, Sprosen T, Collins R, Allen NE, Sociodemographic Cof. Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. *Am J Epidemiol* 2017;186:1026–34. doi:10.1093/aje/kwx246
  - 25 Wu X, Katz E, Della Valle MC, Mascioli K, Flanagan JJ, Castelli JP, Schiffmann R, Boudes P, Lockhart DJ, Valenzano KJ, Benjamin ER. A pharmacogenetic approach to identify mutant forms of  $\alpha$ -galactosidase a that respond to a pharmacological chaperone for Fabry disease. *Hum Mutat* 2011;32:965–77. doi:10.1002/humu.21530
  - 26 Sheng B, Yim KF, Lau LK, Lee HCH, Fung KSS, Ma KFJ, Chak WL, HCH L, Ma KFJ CWL. Two related Chinese Fabry disease patients with a p.N215S pathological variant who presented with nephropathy. *Mol Genet Metab Rep* 2020;24. doi:10.1016/j.ygm.2020.100596

## Prevalence and penetrance of Fabry Disease causing variants in the UK population

### Primary care data search

Acroparasthesia

Angiokeratoma

Sensorineural hearing loss

Hypohidrosis

Anhydrosis

LVH (ventricular)

Congestive heart failure

Cardiomyopathy

Dysrhythmia

Conduction

Stroke

Cerebrovascular

TIA (Transient Ischaemic Attack)

Neuropathy

Ischaemic

The terms above were searched for. Where a word is underlined, the underlined section was searched for in addition to the whole phrase.

### Fabry phenotype risk cohorts:

#### Risk Cohort 1

#### ICD 10 codes

H903 H904 H905 H906 H907 H908 H912 H918 H919 H933 H938 H939 I200 I201 I208 I209  
I210 I211 I212 I213 I214 I219 I248 I249 I250 I251 I252 I255 I256 I258 I259 I250 I251 I252  
I255 I256 I258 I259 I421 I422 I425 I428 I429 I431 I438 I440 I441 I442 I443 I444 I446 I447

I450 I451 I452 I453 I454 I455 I456 I458 I459 I460 I461 I469 I460 I461 I469 I480 I481 I482  
I483 I484 I489 I490 I491 I492 I493 I494 I498 I499 I500 I501 I509 I517 I518 I519  
ICD 9 codes

3891 3892 3899 4109 4119 4129 4139 4140 4148 4149 4254 4260 4263 4264 4266 4267  
4269 4271 4272 4273 4274 4275 4278 4279 4280 4281 4292

Risk Cohort 2

ICD 10 codes

I630 I631 I632 I633 I634 I635 I636 I638 I639 I64 I660 I661 I662 I663 I664 I668 I669 I672 I678  
I679 I200 I201 I208 I209 I210 I211 I212 I213 I214 I219 I248 I249 I250 I251 I252 I255 I256  
I258 I259 I250 I251 I252 I255 I256 I258 I259 I421 I422 I425 I428 I429 I431 I438 I440 I441  
I442 I443 I444 I446 I447 I450 I451 I452 I453 I454 I455 I456 I458 I459 I460 I461 I469 I460  
I461 I469 I480 I481 I482 I483 I484 I489 I490 I491 I492 I493 I494 I498 I499 I500 I501 I509  
I517 I518 I519

ICD 9 Codes

4109 4119 4129 4139 4140 4148 4149 4254 4260 4263 4264 4266 4267 4269 4271 4272  
4273 4274 4275 4278 4279 4280 4281 4292 4349 4359 4369 4371 4379



	N215S		Non-N215S	
	Male	Female	Male	Female
n(%)	5(14)	13(36)	10(28)	8(22)
Age (years), mean(SD)	52.4±11	55.3±8	53.5±7.8	55.25±9.6
Diagnosed Fabry Disease	1	0	1	1
Cardiovascular disease, n(%)	2	3	4	1
Hearing impairment, n(%)	0	0	0	1
CKD-EPI eGFR mls/min/1.73m <sup>2</sup> *	98±12	91±12	90±15	85±18
Microalbuminuria #	2	3	3	0

Supplementary table 1: Phenotype data when analysed by presence of N215S data versus non-N215S variants. (\*creatinine missing for 1 female, #ACR missing for 1 male).