Using the Genomics England National Genomic Research Library (NGRL) and UK Biobank to investigate the genetic, phenotypic and clinical landscape of thymidine kinase 2 deficiency (TK2d)

Robert McFarland[1,2], **Miruna Carmen Barbu**[3], Kaja Zarakowska[4], Martin Armstrong[5], Kate Witkowska[3], Olga Giannakopoulou[6], James Staley[6], Ella Davyson[3,7], Nour Elkhateeb[3], Chris Odhams[3]

[1]Translational and Clinical Research Institute, University of Newcastle, Newcastle upon Tyne, UK;
[2]NHS Rare Mitochondrial Disorders Service, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK;
[3]Genomics England, London, UK; [4]UCB, Bulle, Switzerland; [5]UCB, Braine-l'Alleud, Belgium; [6]UCB, Slough, UK;
[7]School of Informatics, University of Edinburgh, Edinburgh, UK

16th European Paediatric Neurology Society Congress, Munich, Germany
11 July 2025

Please use this QR code to download a PDF of this presentation

Disclosures and acknowledgements

Robert McFarland receives funding from Action Medical Research, the Leigh Syndrome International Consortium, The Lily Foundation and Wellcome, and has consultancy agreements with Abliva, Pretzel Therapeutics and UCB. **Miruna Carmen Barbu**, **Kate Witkowska**, **Nour Elkhateeb** and **Chris Odhams** are employees of Genomics England. **Kaja Zarakowska**, **Martin Armstrong**, **Olga Giannakopoulou** and **James Staley** are employees of and stockholders in UCB. **Ella Davyson** received funding from Genomics England as part of a 6-month internship, and receives funding from the UK Research and Innovation Centre for Doctoral Training in Biomedical Artificial Intelligence (grant EP/S02431X/1) at The University of Edinburgh School of Informatics

This research was made possible through access to data and findings in the National Genomic Research Library via the Genomics England Research Environment

This research has been conducted using UK Biobank under Application Number 44411

This analysis was funded by UCB

The authors acknowledge Tamsin Chambers PhD of PharmaGenesis Cambridge, Cambridge, UK for writing and editorial assistance, which was funded by UCB, in accordance with Good Publication Practice 2022 (GPP 2022) guidelines (https://www.ismpp.org/gpp-2022)

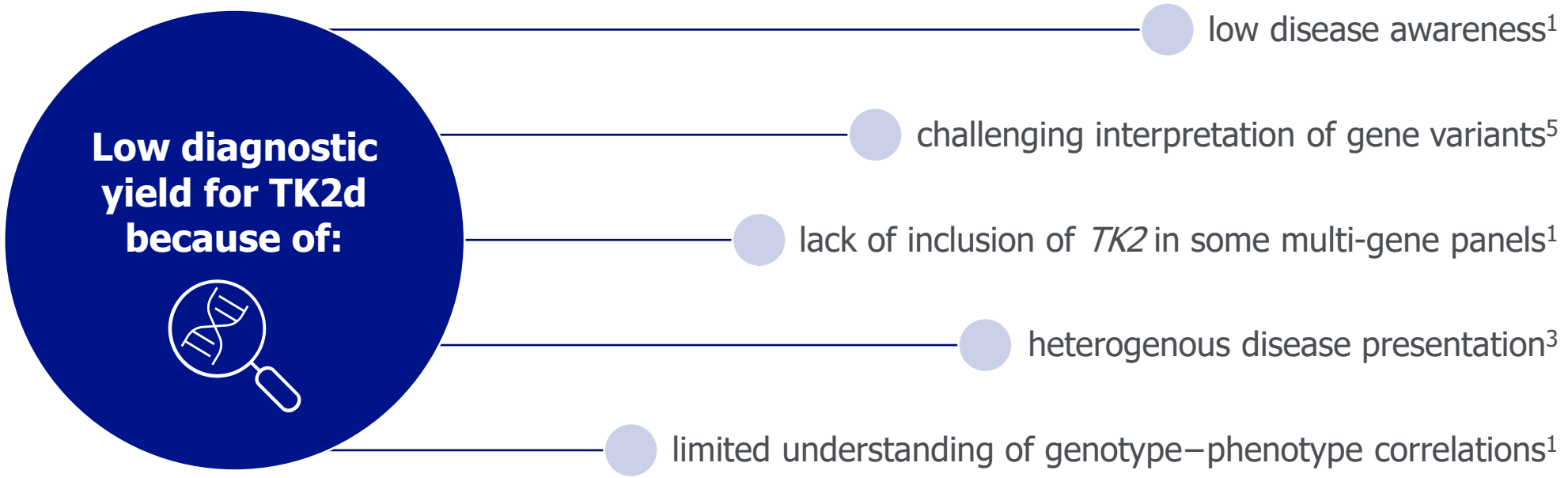
The authors would like to thank Margarita Lens MSci CMPP of UCB for publication coordination

Data first presented at Genomics of Rare Disease; Hinxton, Cambridge, UK; 9–11 April 2025

NHS, National Health Service.

TK2d is an ultra-rare, autosomal recessive, progressive and often life-threatening mitochondrial myopathy[1–3]

- Pathogenic *TK2* variants lead to impaired activity of TK2, an enzyme essential for mtDNA maintenance[1,2]
- Early diagnosis can facilitate appropriate disease management, as well as access to emerging treatments[1,4]



mtDNA, mitochondrial DNA; *TK2*, thymidine kinase 2 gene; TK2, thymidine kinase 2 enzyme; TK2d, thymidine kinase 2 deficiency.

1. Berardo A, *et al. J Neuromuscul Dis* 2022;9:225–35; 2. Wang J, *et al. Mol Genet Metabol* 2018;124:124–30; 3. Garone C, *et al. J Med Genet* 2018;55:515–21; 4. Stenton SL, Prokisch H. *EBioMedicine* 2020;56:102784; 5. Schlieben LD, Prokisch H. *Front Cell Dev Biol* 2020;8:600079.

Large cohort datasets could provide deeper insights into the genetic, phenotypic and clinical landscape of TK2d

NGRL[1]

- Managed by Genomics England
- De-identified whole-genome sequencing data, with deep phenotyping at recruitment
- Family information available for enrolled participants
- Cohorts enriched for rare diseases

100,000 Genomes Project

The UK NHS Genomic Medicine Service

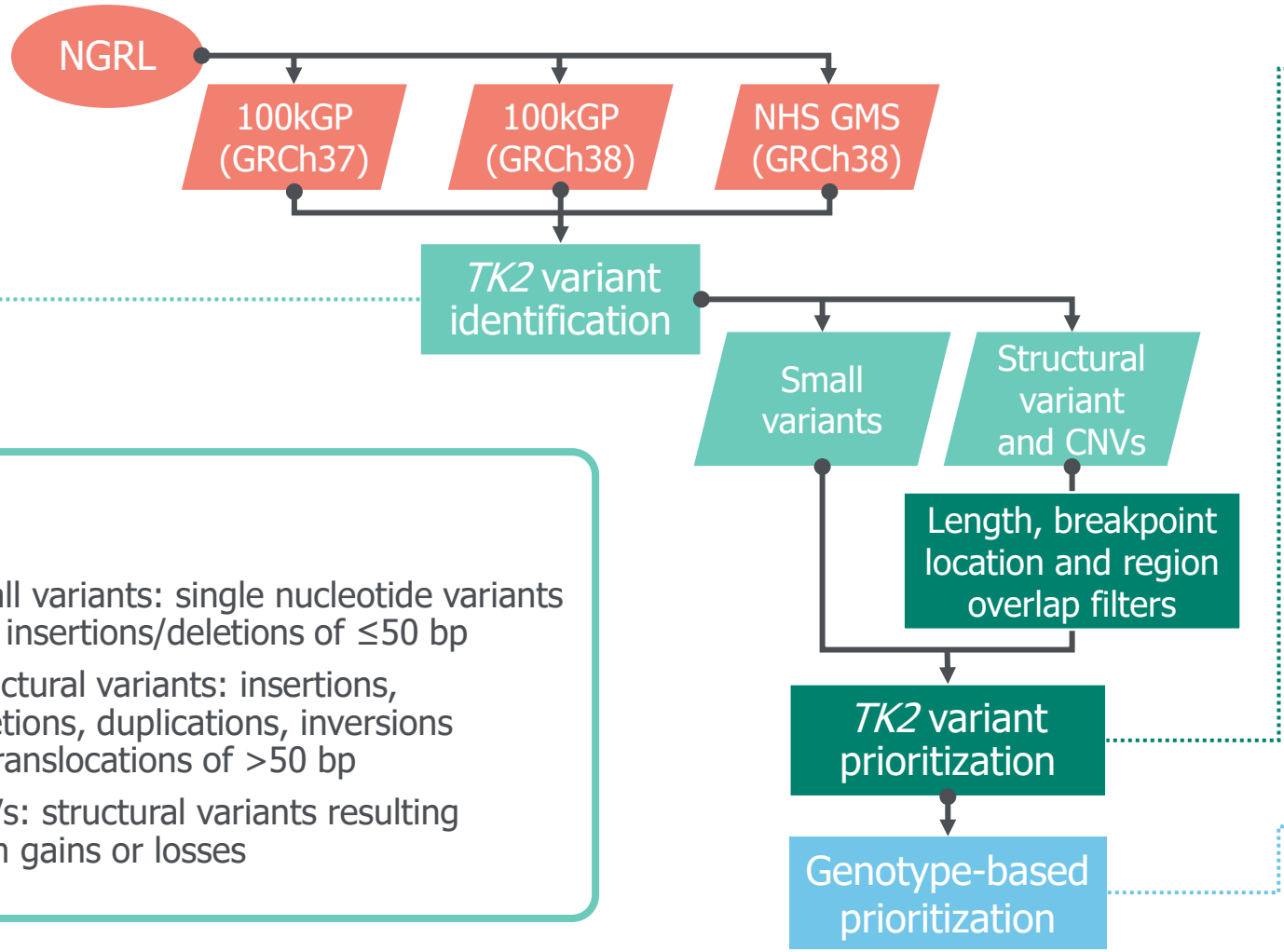
UK Biobank[2]

- Prospective cohort study of ~500,000 adults, aged 40–69 years at recruitment, from the UK[3]
- De-identified whole-exome data, with deep phenotyping at recruitment and longitudinal medical and hospital admission data[2,3]
- Nominally healthy cohort

Objectives: to use large cohort datasets to identify and characterize *TK2* variants, and to phenotypically characterize carriers of *TK2* variants

This research has been conducted using UK Biobank under Application Number 44411.
NGRL, National Genomic Research Library; NHS, National Health Service; *TK2*, thymidine kinase 2 gene; TK2d, thymidine kinase 2 deficiency.
1. The National Genomic Research Library v5.1, Genomics England. https://doi.org/10.6084/m9.figshare.4530893.v7; 2. UK Biobank [Internet]. c2025. Available from: https://www.ukbiobank.ac.uk/ (Accessed 13 June 2025);
3. Bycroft C, *et al. Nature* 2018;562:203–9.

A comprehensive variant annotation strategy was applied to prioritize *TK2* variants identified in the NGRL



- Small variants: single nucleotide variants and insertions/deletions of ≤50 bp
- Structural variants: insertions, deletions, duplications, inversions or translocations of >50 bp
- CNVs: structural variants resulting from gains or losses

Variant annotation strategy

- Allele frequencies ≤1% in both the NGRL and gnomAD
- Confirmed pathogenicity in publicly available databases (Genomics England Rare Disease Interpretation and Diagnostic Pipeline, ClinVar, VarSome)
- *In silico* predicted deleteriousness (VEP, CADD/CADD-SV, REVEL, LOFTEE, SpliceAI)
- Custom list provided by the study sponsor

Variants carried as alternate homozygous or compound heterozygous in ≥1 participant

100kGP, 100,000 Genomes Project; bp, base pairs; CADD, Combined Annotation Dependent Depletion; CADD-SV, CADD Structural Variant; CNV, copy number variant; gnomAD, Genome Aggregation Database; GRCh37, Genome Reference Consortium human build 37; GRCh38, Genome Reference Consortium human build 38; LOFTEE, Loss-Of-Function Transcript Effect Estimator; NGRL, National Genomic Research Library; NHS GMS, National Health Service Genomic Medicine Service; REVEL, Rare Exome Variant Ensemble Learner; *TK2*, thymidine kinase 2 gene; VEP, Variant Effect Predictor.

Prioritized variant carriers and family members were screened for TK2d-related phenotypes and segregation was assessed

Disease terms (recruited rare disease; HPO term; ICD-10 term) relevant to TK2d were categorized based on the following disease domains, created in collaboration with expert clinical review:

- Muscle weakness
- Impaired mobility
- Abnormal creatine kinase level
- Impaired development
- Anaemia
- Epilepsy
- Hearing loss or impairment

Group 1

- Alternate homozygous only
- TK2d-related phenotype present
- Complete penetrance

Group 2

- Alternate homozygous or heterozygous
- TK2d-related phenotype present
- Reduced penetrance

Group 3

- Alternate homozygous or heterozygous
- TK2d-related phenotype absent

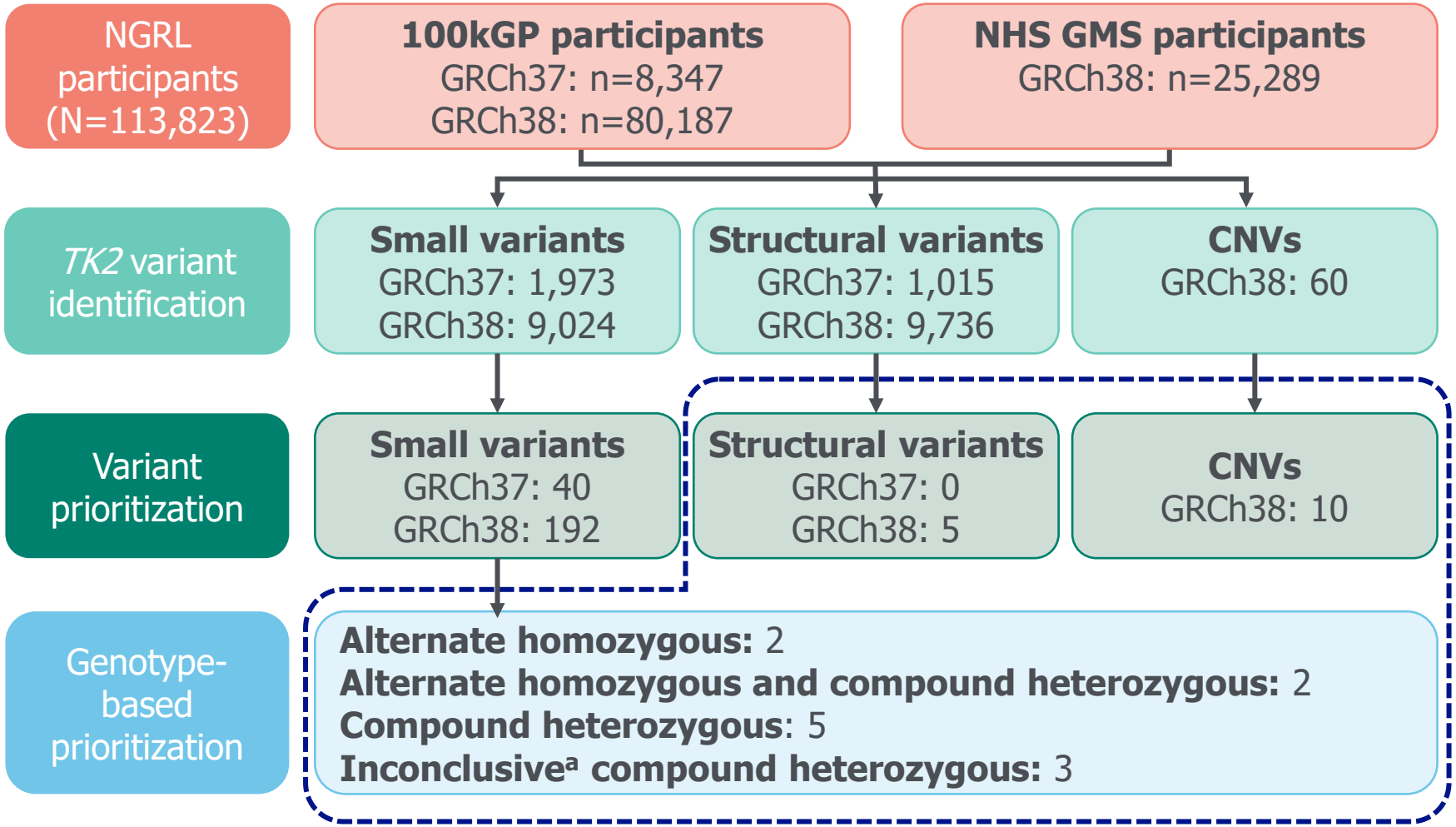
Segregation analysis

UK Biobank analysis

Prioritized alternate homozygous small *TK2* variants identified in the NGRL were screened for in UK Biobank to investigate their pathogenic significance further

HPO, Human Phenotype Ontology; ICD-10, International Statistical Classification of Diseases and Related Health Problems 10th Revision; NGRL, National Genomic Research Library; *TK2*, thymidine kinase 2 gene; TK2d, thymidine kinase 2 deficiency.

Prioritized for further investigation were 12 small variants, 5 structural variants and 10 CNVs



Carriers of prioritized variants and their family members screened for TK2d-relevant phenotypes

[a]Inconclusive compound heterozygous variants are those that could not be confirmed through family structure or phasing information.
100kGP, 100,000 Genomes Project; CNV, copy number variant; GRCh37, Genome Reference Consortium human build 37; GRCh38, Genome Reference Consortium human build 38; NGRL, National Genomic Research Library; NHS GMS, National Health Service Genomic Medicine Service; *TK2*, thymidine kinase 2 gene; TK2d, thymidine kinase 2 deficiency.

Small missense variant chr16:66531432_G_A (p.Thr108Met in exon 5) was identified as a key contributor to TK2d

Allele frequency

- 3.74×10^{-5}

Databases

- Previously identified through the Genomics England Rare Disease Interpretation and Diagnostic Pipeline
- ClinVar/VarSome: pathogenic

In silico tools

- CADD score: ≥15
- VEP: moderate consequence

Carried as alternate homozygous by <5 participants categorized in **group 1**, with TK2d-related phenotypes including muscle weakness, impaired mobility and abnormal creatine kinase levels

Variant ID	Variant-specific counts across family structures									
	Exhibits TK2d-related phenotypes					Does not exhibit TK2d-related phenotypes				
	Singleton	Sibling-pair	Duo	Trio	Quintet	Singleton	Sibling-pair	Duo	Trio	Quintet
chr16:66531432_G_A	<5									

UK Biobank analysis

No UK Biobank participants were alternate homozygous carriers of p.Thr108Met

CADD, Combined Annotation Dependent Depletion; TK2d, thymidine kinase 2 deficiency; VEP, Variant Effect Predictor.

Variants carried by participants categorized in group 2 were considered likely to be benign in relation to TK2d

chr16:66583871_G_A (GRCh37)/ chr16:66549968_G_A (GRCh38)
(p.Arg32Trp in exon 1)

- Allele frequency: $\leq 6.01\times10^{-3}$
- ClinVar/VarSome: benign or likely benign
- CADD score: ≥15; VEP: moderate consequence
- Carried as alternate homozygous by 6 participants and family members, and heterozygous by 15 participants and family members

chr16:66543128_66546738_C_<DEL>
(feature truncating intronic variant 3,611 bp in length spanning intron 2)

- Allele frequency: 1.25×10^{-5}
- VarSome: uncertain significance
- VEP: high consequence
- Carried as heterozygous by <5 participants and family members

Variant ID	Variant-specific counts across family structures									
	Exhibits TK2d-related phenotypes					Does not exhibit TK2d-related phenotypes				
	Singleton	Sibling-pair	Duo	Trio	Quintet	Singleton	Sibling-pair	Duo	Trio	Quintet
chr16:66583871_G_A (GRCh37)/ chr16:66549968_G_A (GRCh38)		<5		<5	<5			<5	<5	
chr16:66543128_66546738_C_<DEL>	<5					<5				

UK Biobank analysis

25 participants were alternate homozygous carriers of p.Arg32Trp; <5 of these participants had HES records for anaemia and other non-TK2d-related morbidities

bp, base pairs; CADD, Combined Annotation Dependent Depletion; GRCh37, Genome Reference Consortium human build 37; GRCh38, Genome Reference Consortium human build 38; HES, Hospital Episode Statistics; TK2d, thymidine kinase 2 deficiency; VEP, Variant Effect Predictor.

Conclusions and outlook

- This study demonstrates how large sequencing datasets and deep phenotyping can be used to study ultra-rare diseases such as TK2d
- Further downstream analyses could include integrating multi-omics data and cross-referencing with TK2d-focused datasets

TK2d, thymidine kinase 2 deficiency.

The authors thank the participants and their caregivers, in addition to the investigators and their teams who contributed to this study

Thank you for your attention

Do you have any questions?