OPEN ACCESS

Inherited predisposition to pneumothorax: estimating the frequency of Birt-Hogg-Dubé syndrome from genomics and population cohorts

Bryndis Yngvadottir,[1] Lucy Richman,[2] Avgi Andreou,[1] Jessica Woodley,[3] Anita Luharia,[3] Derek Lim,[4] Genes & Health Research Team,[5] Eamonn R Maher,[1,4,6] Stefan J Marciniak [2,7]

[1]Department of Genomic Medicine, University of Cambridge, Cambridge, UK
[2]Department of Respiratory Medicine, University of Cambridge, Cambridge, UK
[3]Central and South Genomic Laboratory Hub, Birmingham Women's and Children's Hospitals NHS Foundation Trust, Birmingham, UK
[4]Department of Clinical Genetics, Birmingham Women's and Children's Hospitals NHS Foundation Trust, Birmingham, UK
[5]Genes & Health, Queen Mary University of London, London, UK
[6]Aston University, Birmingham, UK
[7]Cambridge Institute for Medical Research (CIMR), University of Cambridge, Cambridge, UK

Correspondence to
Professor Stefan J Marciniak; sjm20@cam.ac.uk

ERM and SJM are joint senior authors.

To cite: Yngvadottir B, Richman L, Andreou A, *et al*. *Thorax* Epub ahead of print: [*please include* Day Month Year]. doi:10.1136/thorax-2024-221738

ABSTRACT

Birt-Hogg-Dubé syndrome (BHDS) is the most common monogenic cause of pneumothorax. Most affected families have pathogenic variants in the *FLCN* gene. Using large genomic registries (UK Biobank (UKB), 100,000 Genomes Project and East London Genes & Health) including >550 000 individuals, we demonstrate that the frequency of clinically validated loss-of-function *FLCN* variants is 1 in 2710 to 4190. While the lifetime risk of pneumothorax in *FLCN* mutation carriers in the UKB and a BHDS clinical cohort was substantial (28.4% and 37.3%, respectively, to age 65 years), the lifetime risk of renal cancer was significantly lower in UKB than in BHDS patients (1% vs 32.1%). These findings highlight the importance of clinical context in managing individuals with *FLCN* mutations.

INTRODUCTION

Birt-Hogg-Dubé syndrome (BHDS) is an autosomal dominant disorder comprising benign skin tumours, cystic lung disease, pneumothorax and kidney cancer.[1–3] The population prevalence of BHDS is uncertain, but frequencies around 1 in 200 000 are commonly quoted (https://www.orpha.net/en/disease/detail/122?name=FLCN&mode=gene). Pulmonary cysts occur in 70–92% of patients but are asymptomatic unless a pneumothorax occurs (30–51%, median age of 30–38 years).[4] Renal cell carcinoma (RCC) occurs in 12–23% of individuals (median age 46–56 years) and can be multifocal.[4–6] Early diagnosis of BHDS enables RCC surveillance. BHDS is caused by pathogenic variants (mostly loss-of-function) in the *FLCN* gene[3] and >190 pathogenic variants have been identified with an estimated lifetime penetrance of 84–95%.[4 7] Cutaneous fibrofolliculomas on the face and upper trunk are often overlooked.[2] To define the prevalence and penetrance of BHDS, we investigated the frequency of *FLCN* mutations in large-scale genomic database research studies.

METHODS

We analysed exomes/genomes of 556 898 individuals recruited to the 100,000 Genomes Project (100kGP),[8] the UK Biobank (UKB)[9] and East London Genes & Health (ELGH).[10] Variants in the *FLCN* gene region were extracted from sequencing data, annotated and filtered to prioritise loss-of-function (predicted to cause a premature stop codon, a frameshift, or abolish a canonical splice site). Variants were then reviewed for pathogenicity and class-categorised according to UK clinical diagnostic standards. Their prevalence was then calculated for each cohort separately. We calculated age-related risks for *FLCN* mutation carriers comparing UKB to a UK clinical series of BHDS patients (128 carriers from 43 families).[6] Further details on the clinical cohort, datasets, data processing and analysis can be found in the online supplemental methods.

RESULTS

Across the three studies, we identified 155 individuals from 556 898 genomes with 45 different pathogenic loss-of-function *FLCN* variants (online supplemental Table S1:figure S1). After correcting for potential causes of ascertainment bias (online supplemental methods), the prevalence of unrelated individuals with a loss-of-function *FLCN* variant in the rare disease arm of the 100kGP was 1 in 2710 (95% CI 1650 to 4480) (table 1). In the UKB cohort, 117 people (78 unrelated) had pathogenic loss-of-function *FLCN* variants giving the estimated prevalence of 1 in 4190 (95% CI 3360 to 5230) (table 1). The ELGH cohort gave a prevalence estimate of 1 in 1490 (95% CI 680 to 3240; six individuals with three loss-of-function variants) (table 1).

We next investigated the frequency of BHDS-related manifestations (pneumothorax and RCC) in individuals with a pathogenic loss-of-function *FLCN* mutation in 100kGP and UKB cohorts (no clinical information available for ELGH). In 100kGP, 3.1% (1/32) had a history of pneumothorax (when aged <28). In the UK Biobank, 25.6% (30/117) had a pneumothorax (median 47 years, range 23–83). The frequency of RCC in *FLCN* mutation carriers was 15.6% (5/32, median age 61, range 25–77 years) and 5.1% (6/117, median age 72, range 46–80 years) in the 100kGP and UKB respectively. In UKB, two individuals with a pathogenic *FLCN* mutation had both a pneumothorax and an RCC. Age-related risks of pneumothorax and RCC in the UKB cohort were calculated and compared with those from a UK clinical series of BHDS patients (online supplemental Table S2 and S3).[6] Though the age-related risk of pneumothorax was higher in BHDS patients than in UKB participants to age 65 years (37.3% and 28.4%, respectively), the difference was not significant (p=0.2154) (figure 1A). However, the lifetime risk for RCC in *FLCN* mutation-carrying individuals was significantly lower in the UKB cohort (1%, 95%CI 0% to 2.8%) than in the BHDS patient cohort (32.1%, 95%CI 18.6% to 43.4%) at age 65 years (p=0.0005) (figure 1B). Age-related risks were not available for 100kGP participants.

Table 1 LoF *FLCN* variants in the three studied cohorts. Numbers were rounded up to three significant figures

	100kGP	UKB	ELGH
Prevalence of LoF *FLCN* variants	1 in 2710* (95% CI 1640 to 4480)	1 in 4190† (95% CI 3360 to 5230)	1 in 1490 (95% CI 680 to 3240)
Frequency of LoF *FLCN* variants	0.0368% (95% CI 0.0223% to 0.0608%)	0.0239% (95% CI 0.0191% to 0.0298%)	0.0673% (95% CI 0.0308% to 0.1467%)
Total number of individuals with LoF *FLCN* variants	32	117	6
Total number of unrelated individuals with LoF *FLCN* variants	15	78	6
% *FLCN* mutation carriers with pneumothorax	3.1% (1/32)	25.6% (30/117)	N/A
% *FLCN* mutation carriers with RCC	15.6% (5/32)	5.1% (6/117)	N/A

N/A: we do not have access to phenotypic information in the ELGH cohort.
*Removed related participants and those recruited to 100kGP for pneumothorax and/or cancer.
†Removed related participants.
ELGH, East London Genes & Health; 100kGP, 100,000 Genomes Project; LoF, loss-of-function; UKB, UK Biobank.

DISCUSSION

Although BHDS has been estimated to affect only 1 in 200 000 people (https://www.orpha.net/en/disease/detail/122?name=FLCN&mode=gene), our analysis of genomic data from 556 898 individuals suggests that pathogenic *FLCN* variants are far more common (1 in 2710 to 4190). These estimates are conservative as we excluded related individuals, those that might have been recruited because of a BHDS-related complication (100kGP) and focused exclusively on loss-of-function *FLCN* variants. Our clinical cohort data (100kGP) are consistent with a smaller genomic study of 135 990 individuals from a healthcare cohort in the USA in which truncating variants in *FLCN* were detected in 1 in 3234 individuals.[11] Importantly, the UKB cohort may better represent the prevalence in a general population.

Together, the results of genomic studies in healthcare (100kGP and Savatt *et al*[11]) or population-based cohorts (UKB) reveal that pathogenic *FLCN* variants are far more common than generally appreciated. Other studies have found that clinically ascertained cohorts have higher penetrance than population-based studies.[12 13] However, while we found some evidence of reduced penetrance for pneumothorax in population-ascertained *FLCN* mutation carriers than in a clinical BHDS patient cohort, this did not reach statistical significance for lifetime risk. In contrast, there was a significant difference in lifetime risks of RCC between BHDS patients and UKB participants (32.1% and 1% respectively; p=0.0005). These findings would be consistent with the hypothesis that environmental or genetic modifiers in families ascertained with clinically diagnosed BHDS result in a higher penetrance for RCC.

There are some limitations to our analysis. We were unable to assess the frequency of skin lesions or lung cysts in participants with *FLCN* mutations. The absence of renal surveillance in the population cohort might contribute, in part, to the lower ascertainment/later diagnosis of renal cancer in these individuals. The international statistical classification of diseases and related health problems, 10th revision (ICD10) code 'C64 - malignant neoplasm of kidney, except renal pelvis' does not include cases of oncocytoma and we are unable to distinguish between the different histological subtypes of RCC. While the ICD10 code 'D30.0 Benign neoplasm: Kidney' would capture oncocytomas, this code was only reported for 3 out of 502 369 UKB participants, none of which had pathogenic *FLCN* mutations. We further note that in our clinical series of BHDS patients, 93% of patients with a renal tumour presented with RCC, indicating that our results are unlikely to be affected by large numbers of individuals with oncocytomas not being identified in UKB. UKB participants were aged >40 years and early onset, fatal cases of RCC would have been excluded from participation (though most RCCs in BHDS occur after age 50).[2 6] The finding that *FLCN* mutation carriers are much more common than previously supposed should encourage the application of genetic testing

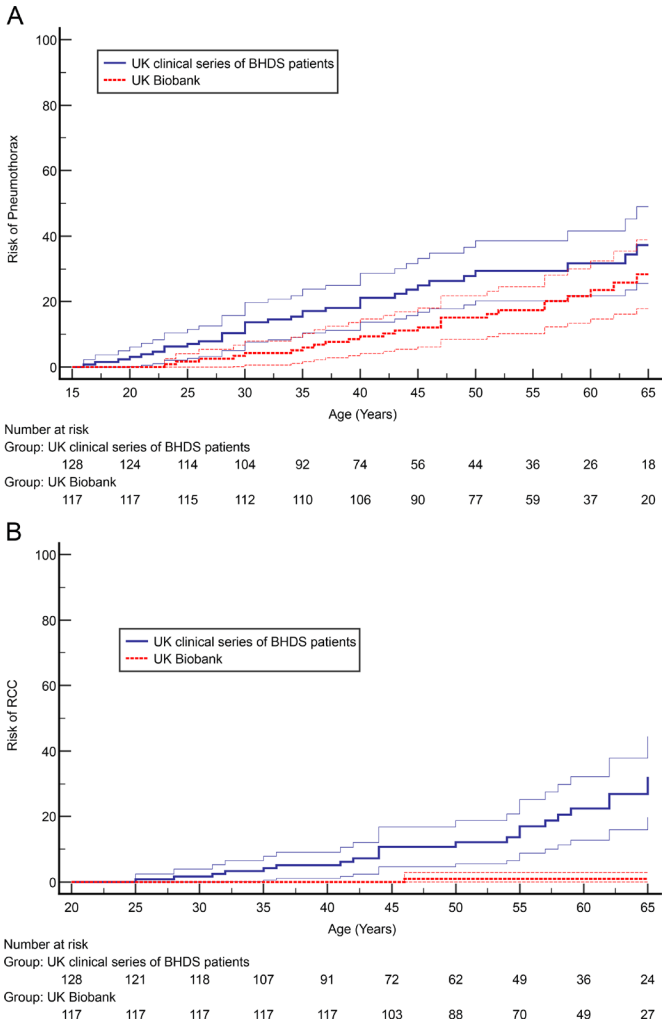

Figure 1 Age-related manifestations of Birt-Hogg-Dubé syndrome (BHDS). Kaplan-Meier survival curves (with 95% CI) for age-related risk of developing (A) pneumothorax and (B) renal cell carcinoma (RCC) in patients with BHDS with pathogenic loss-of-function *FLCN* mutations (blue line) versus UK Biobank participants with pathogenic loss-of-function *FLCN* mutations (red line).

for BHDS in individuals with familial or recurrent pneumothorax or familial or multiple RCC, even if a family history or other features of BHDS are absent. However, when a *FLCN* pathogenic variant is detected as an incidental/secondary finding, while the risks of pneumothorax are appreciable, the application of screening protocols for RCC (eg, annual renal MRI) based on penetrance estimates from clinical BHDS cohorts might be less cost-effective than less intense screening (eg, biennial renal ultrasonography). Prospective follow-up to document the complication risks in individuals with an incidental germline pathogenic *FLCN* mutation identified by genomic analysis is required to delineate the optimal clinical management of such individuals.

X Stefan J Marciniak @Prof_Marciniak

Acknowledgements We thank the participants in all cohorts for their involvement. This research was supported by the Myrovlytis Trust (MT21_1) and NIHR Cambridge Biomedical Research Centre (BRC-1215-20014 and NIHR203312). The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care. The authors would like to thank Professor Paul Pharoah (UCLA) for his advice on the analytical approaches adopted in the manuscript.100kGP: This research was made possible through access to data in the National Genomic Research Library, which is managed by Genomics England Limited (a wholly owned company of the Department of Health and Social Care). The National Genomic Research Library holds data provided by patients and collected by the NHS as part of their care and data collected as part of their participation in research. The National Genomic Research Library is funded by the National Institute for Health Research and NHS England. The Wellcome Trust, Cancer Research UK and the Medical Research Council have also funded research infrastructure. UKB: This research has been conducted using the UK Biobank Resource under Application Number 95547. ELGH: Genes & Health is/has recently been core-funded by Wellcome (WT102627, WT210561), the Medical Research Council (UK) (M009017, MR/X009777/1, MR/X009920/1), Higher Education Funding Council for England Catalyst, Barts Charity (845/1796), Health Data Research UK (for London substantive site), and research delivery support from the NHS National Institute for Health Research Clinical Research Network (North Thames). Genes & Health is/has recently been funded by Alnylam Pharmaceuticals, Genomics PLC; and a Life Sciences Industry Consortium of Astra Zeneca PLC, Bristol-Myers Squibb Company, GlaxoSmithKline Research and Development Limited, Maze Therapeutics Inc, Merck Sharp & Dohme LLC, Novo Nordisk A/S, Pfizer Inc, Takeda Development Centre Americas Inc.We thank Social Action for Health, Centre of The Cell, members of our Community Advisory Group, and staff who have recruited and collected data from volunteers. We thank the NIHR National Biosample Centre (UK Biocentre), the Social Genetic & Developmental Psychiatry Centre (King's College London), Wellcome Sanger Institute, and Broad Institute for sample processing, genotyping, sequencing and variant annotation.We thank: Barts Health NHS Trust, NHS Clinical Commissioning Groups (City and Hackney, Waltham Forest, Tower Hamlets, Newham, Redbridge, Havering, Barking and Dagenham), East London NHS Foundation Trust, Bradford Teaching Hospitals NHS Foundation Trust, Public Health England (especially David Wyllie), Discovery Data Service/Endeavour Health Charitable Trust (especially David Stables), Voror Health Technologies Ltd (especially Sophie Don), NHS England (for what was NHS Digital) - for GDPR-compliant data sharing backed by individual written informed consent. Most of all we thank all of the volunteers participating in Genes & Health.

Collaborators Genes & Health Research Team The current members from the Genes & Health Research Team (in alphabetical order by surname) are:Shaheen Akhtar, Mohammad Anwar, Elena Arciero, Omar Asgar, Samina Ashraf, Saeed Bidi, Gerome Breen, James Broster, Raymond Chung, David Collier, Charles J Curtis, Shabana Chaudhary, Megan Clinch, Grainne Colligan, Panos Deloukas, Ceri Durham, Faiza Durrani, Fabiola Eto, Sarah Finer, Joseph Gafton, Ana Angel, Chris Griffiths,Joanne Harvey, Teng Heng, Sam Hodgson, Qin Qin Huang, Matt Hurles, Karen A Hunt, Shapna Hussain, Kamrul Islam, Vivek Iyer, Benjamin M Jacobs, Ahsan Khan, Claudia Langenberg, Cath Lavery, Sang Hyuck Lee, Daniel MacArthur, Sidra Malik, Daniel Malawsky, Hilary Martin, Dan Mason, Rohini Mathur, Mohammed Bodrul Mazid,John McDermott, Caroline Morton, Bill Newman, Elizabeth Owor, Asma Qureshi, Shwetha Ramachandrappa, Mehru Raza, Jessry Russell, Nishat Safa, Miriam Samuel, Michael Simpson, John Solly, Marie Spreckley. Daniel Stow, Michael Taylor, Richard C Trembath, Karen Tricker, David A van Heel, Klaudia Walter, Caroline Winckley, Suzanne Wood, John Wright, Ishevanhu Zengeya, Julia Zöllner.

Contributors BY, ERM and SJM were involved in the conception and design, data analysis and interpretation. BY performed the data analysis.ERM, LR, JW, AL and DL performed clinical and/or genetic data analysis. All authors were involved in drafting the submitted article. SJM is the guarantor of the content of the manuscript.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by REC 11/NW/0382REC 14/EE/1112 Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement 100kGP: Research on the de-identified patient data used in this publication can be carried out in the Genomics England Research Environment subject to a collaborative agreement that adheres to patient led governance. All interested readers will be able to access the data in the same manner that the authors accessed the data. For more information about accessing the data, interested readers may contact research-network@genomicsengland.co.uk or access the relevant information on the Genomics England website: https://www.genomicsengland.co.uk/research. UKB: Data from UK Biobank is available on application from https://www.ukbiobank.ac.uk/ All researchers who wish to access the research resource must register with UK Biobank by completing the registration form in the Access Management System (AMS). https://www.ukbiobank.ac.uk/ enable-your-research/register ELGH: The VCF data (accession number EGAD00001005469) is available from the EBI-EGA genotype phenotype archive (www.ebi.ac.uk/ega). Users need to complete the standard Wellcome Sanger Institute Data Access Agreement.

ORCID iD

Stefan J Marciniak http://orcid.org/0000-0001-8472-7183

REFERENCES

1 Birt AR, Hogg GR, Dubé WJ. Hereditary multiple fibrofolliculomas with trichodiscomas and acrochordons. *Arch Dermatol* 1977;113:1674–7.
2 Menko FH, van Steensel MAM, Giraud S, *et al*. Birt-Hogg-Dubé syndrome: diagnosis and management. *Lancet Oncol* 2009;10:1199–206.
3 Nickerson ML, Warren MB, Toro JR, *et al*. Mutations in a novel gene lead to kidney tumors, lung wall defects, and benign tumors of the hair follicle in patients with the Birt-Hogg-Dubé syndrome. *Cancer Cell* 2002;2:157–64.
4 Matsumoto K, Lim D, Pharoah PD, *et al*. A systematic review assessing the existence of pneumothorax-only variants of FLCN. Implications for lifelong surveillance of renal tumours. *Eur J Hum Genet* 2021;29:1595–600.
5 Pavlovich CP, Grubb RL, Hurley K, *et al*. EVALUATION AND MANAGEMENT OF RENAL TUMORS IN THE BIRT-HOGG-DUBÉ SYNDROME. *Journal of Urology* 2005;173:1482–6.
6 Bruinsma FJ, Dowty JG, Win AK, *et al*. Update of penetrance estimates in Birt-Hogg-Dubé syndrome. *J Med Genet* 2023;60:317–26.
7 Lim DHK, Rehal PK, Nahorski MS, *et al*. A new locus-specific database (LSDB) for mutations in the folliculin (FLCN) gene. *Hum Mutat* 2010;31:E1043–51.
8 Caulfield M, Davies J, Dennys M, *et al*. The national genomics research library. 2017. Available: https://doi.org/10.6084/m9.figshare.4530893.v7
9 Bycroft C, Freeman C, Petkova D, *et al*. The UK Biobank resource with deep phenotyping and genomic data. *Nature New Biol* 2018;562:203–9.
10 Finer S, Martin HC, Khan A, *et al*. Cohort Profile: East London Genes & Health (ELGH), a community-based population genomics and health study in British Bangladeshi and British Pakistani people. *Int J Epidemiol* 2020;49:20–21i.
11 Savatt JM, Shimelis H, Moreno-De-Luca A, *et al*. Frequency of truncating FLCN variants and Birt-Hogg-Dubé-associated phenotypes in a health care system population. *Genet Med* 2022;24:1857–66.
12 Kingdom R, Tuke M, Wood A, *et al*. Rare genetic variants in genes and loci linked to dominant monogenic developmental disorders cause milder related phenotypes in the general population. *Am J Hum Genet* 2022;109:1308–16.
13 Jackson L, Weedon MN, Harrison JW, *et al*. Influence of family history on penetrance of hereditary cancers in a population setting. *Genetic and Genomic Medicine* [Preprint] 2023.