Journal of Cystic Fibrosis 22 (2023) 499–504

Contents lists available at ScienceDirect

Journal of Cystic Fibrosis

journal homepage: www.elsevier.com/locate/jcf

Original Article

The risk of colorectal cancer in individuals with mutations of the cystic fibrosis transmembrane conductance regulator (CFTR) gene: An English population-based study

Rebecca J. Birch[a,b,*], Daniel Peckham[a,c], Henry M. Wood[a], Philip Quirke[a], Rob Konstant-Hambling[d], Keith Brownlee[e], Rebecca Cosgriff[f], Genomics England Research Consortium[5,#], Nicholas Burr[g], Amy Downing[a,b]

[a] Leeds Institute of Medical Research at St James's, University of Leeds
[b] Leeds Institute for Data Analytics, University of Leeds
[c] Leeds Teaching Hospitals NHS Trust
[d] NHS England
[e] Cystic Fibrosis Trust, London, UK
[g] Mid Yorkshire NHS Trust

ARTICLE INFO

Article history:
Received 8 March 2022
Revised 21 June 2022
Accepted 3 October 2022
Available online 15 October 2022

ABSTRACT

Background: Studies have demonstrated a higher risk of developing colorectal cancer (CRC) in individuals with Cystic Fibrosis (CF), and also a potentially increased risk in carriers of cystic fibrosis transmembrane conductance regulator (CFTR) mutations. Life expectancy for those with CF is rising, increasing the number at risk of developing CRC.

Methods: The incidence of CRC amongst individuals with CF was calculated using data from CORECT-R and linked UK CF Registry and Secondary User Services (SUS) data. Crude, age-specific and age-standardised rates were compared to those without CF. The presence of CFTR mutations in individuals with CRC was assessed using 100,000 Genomes Project data.

Findings: The crude incidence rate of CRC in the CF population was 0.29 per 1,000 person-years (28 cases). The CF population were significantly younger than those without (median age at CRC diagnosis 52 years versus 73 years; $p<0{\cdot}01$). When age-adjusted, there was a 5-fold increased CRC incidence amongst individuals with CF compared to those without (SIR 5.0 95%CI 3.2–6.9). When compared to other population studies the overall prevalence of CFTR mutations in the CRC population was significantly higher than expected ($p<0{\cdot}01$).

Interpretation: CF is linked to an increased risk of CRC. The incidence of CFTR mutations in the CRC population is higher than would be expected, suggesting an association between CFTR function and CRC risk. Further research is needed to develop effective screening strategies for these populations.

Funding: Cancer Research UK (grants C23434/A23706 & C10674/A27140)

1. Introduction

In the UK, there are over 10,500 adults and children with Cystic Fibrosis (CF) and the carrier frequency is one in 25 [1]. The median age of the UK population with CF continues to increase (18 in 2010 to 21 in 2020 [2]) as does median predicted survival which has risen, from 43·5 years in 2007–2011 to 50·6 years in 2016–20,20[2].

* Corresponding author.
E-mail address: r.j.birch@leeds.ac.uk (R.J. Birch).

Ambrose, J. C.1; Arumugam, P.1; Bevers, R.1; Bleda, M.1; Boardman-Pretty, F.1,2; Boustred, C. R.1; Brittain, H.1; Caulfield, M. J.1,2; Chan, G. C.1; Fowler, T.1; Giess A.1; Hamblin, A.1;Henderson, S.1,2; Hubbard, T. J. P.1; Jackson, R.1; Jones, L. J.1,2; Kasperaviciute, D.1,2; Kayikci, M.1; Kousathanas, A.1; Lahnstein, L.1; Leigh, S. E. A.1; Leong, I. U. S.1; Lopez, F. J.1;Maleady-Crowe, F.1; McEntagart, M.1; Minneci F.1; Moutsianas, L.1, 2; Mueller, M.1, 2;Murugaesu, N.1; Need, A. C.1, 2; O'Donovan P.1; Odhams, C. A.1; Patch, C.1, 2; Perez-Gil, D.1; Pereira, M. B.1; Pullinger, J.1; Rahim, T.1; Rendon, A.1; Rogers, T.1; Savage, K.1; Sawant, K.1; Scott, R. H.1; Siddiq, A.1; Sieghart, A.1; Smith, S. C.1; Sosinsky, A.1,2; Stuckey, A.1; Tanguy M.1; Taylor Tavares, A. L.1; Thomas, E. R. A.1,2; Thompson, S. R.1;Tucci, A.1, 2; Welland, M. J.1; Williams, E.1; Witkowska, K.1, 2; Wood, S. M.1, 2. 1Genomics England, London, UK. 2William Harvey Research Institute, Queen Mary University of London, London, EC1M 6BQ, UK .

https://doi.org/10.1016/j.jcf.2022.10.001

In 2017, 21% of individuals with CF in the UK were aged over 40 years and this is expected to increase to 36% by 2030 [3].

The risk of developing cancer of the digestive tract, and notably colorectal cancer (CRC), appears to be greater in people with CF than the general population [4] with the incidence being accentuated by solid organ, especially lung, transplantation [5,6]. The increased risk of GI malignancy was highlighted in a large CF registry cohort study in the USA. Alongside this increased incidence, CF itself may result in a more aggressive form of CRC due to mutations of the cystic fibrosis transmembrane conductance regulator (CFTR) gene [7].

Over 2000 mutations of the CFTR gene have been identified [8] with around 400 shown to be associated with CF or defined as CF causing [9]. While carriers of the CFTR gene may be predisposed to an increased prevalence of respiratory disease, there is still debate as to whether it alters the risk of malignancy. Earlier studies found no association [10,11], however, more recently studies looking at F508del alone reported it as being elevated in several types of cancer, including CRC [12]. While F508del is the most common CFTR mutation, present in over 70% of CF [13], other mutations are more common in different ethnic groups [14] making it vital to extend the analysis of carriers beyond this mutation alone.

The introduction of highly effective CFTR modulators heralds a new era where treatment can finally target the underlying defect rather than disease complications for 90% of CF genotypes. While this will have the beneficial impact of increasing life expectancy, it may also put individuals with CF at greater risk of developing CRC, as incidence is significantly associated with age. In the UK over 40% of CRC cases occur in individuals aged 75 and above [15].

Due to the relatively small CF population and its low median age there is currently a scarcity of evidence available to quantify cancer risk and, hence, inform optimal screening strategies. This study aims, therefore, to quantify any excess risks of CRC in the CF population in England. It will also assess any association between mutations of the CFTR gene and CRC in carriers and those affected.

2. Methods

2.1. Data

Analyses were undertaken using three different data sources: the COloRECTal cancer data Repository (CORECT-R) [16], Linked Cystic Fibrosis Trust UK CF Registry data and Secondary Use Statistics (SUS) and the version 10 release of the Genomics England 100,000 Genomes Project [17].

2.2. CORECT-R

This resource contains linked information for all individuals diagnosed with a colorectal cancer (ICD10 C18-C20) within the English National Health Service (NHS) between 1st January 2007 and 31st December 2017. Information on age at CRC diagnosis, sex, tumour site within the colorectum (right colon – caecum, appendix, ascending colon, hepatic flexure & transverse colon. Left colon – splenic flexure, descending colon & sigmoid colon. Other – overlapping lesion of colon, colon, unspecified, rectosigmoid & rectum) and year of CRC diagnosis was obtained from the National Cancer Registration and Analysis Service (NCRAS) component of this resource.

Information about the diagnostic reasons for hospital admissions preceding the diagnosis of the relevant colorectal cancer were obtained from Hospital Episode Statistics (HES) within CORECT-R. Individuals were classified as having CF if the relevant ICD10 code (E84) was recorded during an inpatient stay in an NHS hospital at any point (both prior and post CRC diagnosis). Previous studies have demonstrated an over capture of CF when using hospital records alone to identify the condition [18]. The study demonstrated that the false positives were commonly due to a suspicion of CF being reported as a diagnosis, alongside the incorrect classification of CFTR mutation carriers as CF positive and incorrect coding of multiple complex health conditions. To minimise the impact of this and reduce false positives, individuals with fewer than 2 records of CF were classified as not having CF. This was selected as each individual with CF is admitted to hospital once a year on average [19]. The age profile of those with a single record was examined and the median age (67 (IQR 60–88)) was determined to be higher than would be expected for individuals with CF with a single admission to hospital by clinical experts and it was thought that these individuals were likely miscoded (median age for individuals with >1 admission was 52 years (IQR 40–68)). Given the limitations in identifying CF in the available cancer registration data a second data source (UK CF Registry) was used to validate the crude incidence rates obtained.

Solid organ transplant status was determined using relevant ICD10 & OPCS4 codes (Supplementary Table 1) recorded during a hospital inpatient admission at any point prior to the CRC diagnosis.

2.3. UK CF registry

This resource contained information regarding patients who had consented to participate in the UK CF Registry between 1 January 2009 and 31 December 2020 (covering approximately 99% of the CF population of the UK) [20]. Cystic fibrosis information was obtained from UK CF Registry, and cancer information (site of malignancy) was obtained from SUS data. Used in isolation, hospital record data may underestimate cancers in certain cases [21], with overestimation of cases in older individuals, meaning that the incidence rate calculated for the UK CF Registry analysis is expected to be lower than that observed in the CORECT-R analysis.

2.4. Population figures

Total population figures for the CF population were obtained from the Cystic Fibrosis Trust UK CF Registry for the years 2007–2020 [20]. General population figures were obtained from the Office for National Statistics Mid-Year Population estimates [22] for the time period 2007–2017. The data were aggregated across all years in line with the five year age bands (0–4, 5–9, 10–14, 15–19 etc.), used to calculate the European Standard Population, and by sex.

2.5. Genomics analysis

The carrier analysis was performed using whole genome sequencing data from the blood samples of 2023 colorectal cancer patients, obtained from Genomics England 100,000 Genomes Project [17].

2.6. Statistical analysis

2.6.1. Characteristics of the population

Using the CORECT-R data, a nonparametric test of medians was undertaken to compare age at CRC diagnosis between those with and without CF. χ^2 analysis was performed to test for significant differences in age group at CRC diagnosis, sex and tumour site (right sided colon, left sided colon and rectum) between those with and those without CF.

2.7. Crude incidence rates

The populations included are considered dynamic as it was not possible to determine the vital status of the individuals with CF

over the course of the study period. Due to this incidence rates were calculated using the total number of CRC cases diagnosed over the period as the numerator and the mean of the mid-year population estimates multiplied by the number of years of follow up included in the study as the denominator, resulting in an incidence rate per 1000 person years [23]. Given the limitations described around the identification of individuals with CF in the CORECT-R dataset this analysis was replicated using data from the UK CF Registry linked to SUS data (including individuals diagnosed with CRC between 2009 and 2020). Poisson confidence intervals were calculated for the person-years rates using the mid-P exact test methodology.

A standardised incidence ratio (SIR), the ratio of observed to expected cancers, was calculated for those with CF, using the incidence rate of CRC those without CF (both from the CORECT-R data) to calculate the expected rates. The number of expected CRC cases in the CF population was calculated using the age-specific (5-year age-bands) incidence rate amongst those without CF, multiplied by the total person-years during which the population was followed up (in the case of the CORECT-R the mean CF population x 11).

The rate difference, defined as the difference between the rate amongst those exposed (in this case those with CF) and those without exposure (those without CF) was calculated from these figures and is presented per 1000 person-years.

2.8. Age-specific incidence rates

Age specific incidence rates per 1000 person-years were calculated using the same method as the crude incidence rates in the CORECT-R data. The age groupings were created in line with the predicted median survival (49·1 years[2]) for individuals with CF in the UK at the end of the study period and further divided to reflect the age of the commencement of CRC screening in the UK (<40, 40–49, 50–69 & ≥70).

2.9. Genomics analysis

Samples were processed as part of the Genomics England standard pipeline, being sequenced to approximately 30X coverage on the Illumina HiSeq X platform. Reads were aligned to the human genome GRCh38Decoy assembly using Isaac v03.16.02.19 [24], and variants called by Strelka v2.4.7 [25].

All variants in the 31st July 2020 Variant List History from the CFTR2 resource not classed as “non-CF causing” that were present in at least 10% of the reads for each sample were collated.

The proportion of individuals present in the colorectal cancer data who were identified as having one or more of the mutations present in the Variant History List was calculated. A total proportion higher than 4% was considered to demonstrate a higher than expected incidence in this population as 1 in 25 individuals are estimated to be a carrier of a CFTR mutation[1]. The prevalence of specific mutations was compared to that seen in other studies [12,26]. The statistical significance of any difference was assessed using a two-tailed test of significance between independent proportions.

Statistical analyses were undertaken in Stata 16·0.

2.10. Role of the funding source

This study was funded by Cancer Research UK RB and AD were both funded from by grant C23434/A23706. HW and PQ were funded by the CRUK OPTIMISTICC Grand Challenge (C10674/A27140).

The funder had no role in the design of the study, analysis, interpretation of the results or decision to submit for publication.

3. Results

3.1. Characteristics of the population

Using the CORECT-R data, 28 individuals with CF and CRC were identified. The median age at CRC diagnosis was significantly lower in the population with CF (52 years (IQR 40–68)) compared to the population without CF (n = 364,399) (73 years (IQR 63–81; $p<0.01$). The distribution of ages differed between those with and without CF, with a quarter (25.0%) of those with CF and CRC being aged <40, compared to 2.0% of those without CF. The CF population had a higher proportion of females, 60·7% versus 44·5% in the population without CF (Table 1). Half (50·0%) of the tumours identified in those with CF were located in the right side of the bowel (C18·0-C18·4). In the population without CF 26·5% of tumours were located in the right colon (Table 1). Due to the very small numbers (<5 cases) of solid organ transplant no further analysis was undertaken regarding transplant status.

3.2. Crude incidence rates

The identification of 28 individuals with CF and CRC, in a mean CF population of 8894 individuals (mean population for years 2007–2017), represents a crude incidence rate of 0.29 per 1000 person-years (95%CI 0.19–0.41). This compares to a rate of 0.62 per 1000 person-years (95%CI 0.62–0.62) in individuals without CF (Table 2). The results from the UK CF Registry analysis found 28 cases of cancer of the colon, rectosigmoid or rectum (ICD10 C18-C20) present between 1 January 2009 and 31 December 2020 in a mean population of 10,242 individuals with CF included in the Registry at the start of this time period. This represents an incidence of 0.23 per 1000 person-years (95%CI 0.15–0.32) with CF, comparable to that which was observed using the CORECT-R data.

3.3. The SIR demonstrated a significantly increased incidence of CRC amongst individuals with CF when compared to those without (SIR 5.0 (95%CI 3.2–6.9) (Table 2).Age-specific incidence rates

In both those with and without CF the incidence of CRC was higher amongst those aged ≥50 (Table 3). Individuals with CF aged 50–69 had 1.65 extra cases of CRC per 1000 person-years of follow-up compared to those without CF, and for those aged 40–49 there were less than 1 (0·73) extra cases of CRC per 1000 person-years (Table 3).

3.4. Genomics analysis

When compared to the prevalence of CFTR mutations in other population studies [12,26] the overall rate of mutations in the CRC population was significantly higher than expected ($p<0.01$). Of the 2023 individuals with CRC for whom genetic information was available, 141 (6·9%) were heterozygotes for a CFTR mutation with 38 mutations identified (Table 5). Overall 62 (3·1%) carried the F508del mutation which was not significantly different from that observed in other studies including noncancer patients (3.15%[12]). However the rate of R117H mutation was significantly higher ($p<0·01$) and the prevalence of G551D was significantly lower ($p<0·01$) than has been observed in other studies of heterozygotes without cancer [26]. F508del and R117H accounted for over half of all CFTR mutations observed in this cohort (F508del (44·0%) and R117H (12·8%)) (Table 5). The intron-8 polythymidine sequence was not available. Of the 419 mutations listed in the CFTR2 database as being CF causing or of varying clinical consequence, 378 (90·2%) mutations were not identified in the colorectal data.

Table 1
Characteristics of the study population (CORECT-R data).

		Cystic fibrosis		No cystic fibrosis		P value
		n	%	n	%	
Median age at CRC diagnosis (IQR)		52 (40–68)		73 (63–81)		0·001
Age at CRC diagnosis	<40	7	25.0	7316	2.0	0.001
	40–49	6	21.4	13,674	3.8	
	50–69	8	28.6	129,125	35.4	
	≥70	7	25.0	214,287	58.8	
Sex	Male	11	39·3	202,128	55·5	
	Female	17	60·7	162,271	44·5	0·09
Tumour site	Right colon	14	50·0	121,759	33·4	
	Left colon	8	28·6	96,741	26·6	
	Other*	6	21·1	145,899	40·4	0·09
Total		28		364,399		

*Including tumours of the colon, unspecified, rectum and rectosigmoid.

Table 2
Crude incidence rates, per 1000 person-years.

CORECT-R data	Observed cases of CRC	Expected cases of CRC	Crude rate (per 1000 person-years) (95%CI)	SIR (95%CI)
Cystic fibrosis	28	5·57	0.29 (0.19–0.41)	5.0 (3.2–6.9)
No cystic fibrosis	364,399	–	0.62 (0.62–0.62)	

Table 3
Age-specific incidence rates (per 1000 person-years) (CORECT-R data).

Age group	Incidence rate per 1000 person-years		Rate difference
	Cystic fibrosis	No cystic fibrosis	
<40	0.08	0.02	0.06
40–49	0.90	0.16	0.73
50–69	2.60	0.95	1.65
≥70	24.92	3.06	21.86

Table 5
Mutation frequency, reported where frequency >2 counts below that are available in supplementary information (variant cDNA name, variant protein name and rsID provided in Supplementary Table 2).

Variant legacy name Wild type (no CFTR mutations)	Number of patients	%
Wt	1882	93.0
Heterozygous mutations of CFTR gene	62	7.0
F508del	62	3·1
R117H	18	0·9
L967S	7	0·3
P750L	6	0·3
A349V	4	0·2
621+1G->*T*	3	0·1
296+28A->*G*	3	0·1

4. Discussion

This study demonstrates an increased risk of CRC amongst individuals with CF and a higher than expected incidence of certain CFTR mutations amongst those with CRC. When age-standardised the incidence of CRC was five times higher in individuals with CF. This increase in predisposition is likely to be multifactorial due to a combination of factors such as CFTR dysfunction, increased local and systemic inflammation, altered gut microbiota and the use of high fat and low fibre diets [27–29]. The study also suggests a higher predominance of right sided tumours and an increase incidence in females, compared to the population without CF. This is a significant finding as right sided tumours of the colon have been shown to have worse survival [30] and are frequently diagnosed at a later stage due to their symptom profile [31] making this group key to any interventional strategies around screening. Further work is needed to determine whether this increase in right sided tumours and CRC in females is also seen in the carrier population.

Previous studies have demonstrated an increased risk of malignancy following a solid organ transplant, especially in lung transplant recipients with CF, where the level of immunosuppression is higher [5]. Fewer than 5 of those with CF included in this study had undergone a solid organ transplant at the time of their CRC diagnosis. In the UK over 450 individuals with cystic fibrosis underwent an organ transplant during the time period covered by this study [2,32], this suggests that our finding of fewer than 5 individuals means transplant is not over represented in the CRC population. Further work is required to characterise the true risk in this cohort of patients and optimise screening in those who have or are about to receive a solid organ transplant.

It is only relatively recently that the number of adults with CF (>18 years of age) has exceeded the number of children, but despite this the median age of the UK population in 2020 was still only 21 years. However, survival is likely to increase significantly following the introduction of highly effective CFTR modulators, which target the underlying defect and have resulted in significant clinical improvement and stability. While these drugs may prove beneficial in reducing the risk of CRC in the longer term, by partially correcting CFTR, survival benefits may be negated by the increased risk of CRC associated with an ageing population, previously naïve to modulators in early life. Alternatively, they may not reduce the risk if the microbiome or other associated factors are not corrected.

In people with CF, the development of early onset CRC and precursors to CRC before the age of 50 (the age at which individuals will be eligible for Quantitative Faecal Immunochemical Test (qFIT) screening in the United Kingdom) highlights the need to consider targeted screening at a younger age. Colonoscopy remains the gold standard investigation for CRC and its precursors and along with qFIT was recently estimated to be cost effective in the CF population [33,34]. Despite the gradual increase in the uptake of colonoscopy screening, there are differing views about the benefit and cost effectiveness of screening, the age at which screening should start and the interval between procedures. The CF Founda-

tion has produced comprehensive guidelines and recommend that individuals with CF should be screened from the age of 40 years, with re-screening every five years [33]. In solid organ transplant recipients, screening should start at 30 years [33]. Standard bowel preps in CF, pre colonoscopy, are often inadequate and relatively aggressive regimens are used to maximise bowel clearance and improve diagnostic accuracy. There is growing evidence to suggest that the CF population has the potential to benefit from a tailored screening program but further work is needed to determine what form that should take. This study was unable to determine the impact of screening on the incidence of CRC as only diagnosed CRC was included in the sample, meaning it was not possible to quantify how many cancers were prevented through the screening programmes. Future work should include an analysis of route to cancer diagnosis for individuals with CF compared to those without in order to assess this.

It is thought that up to 15% of CRC cases are due to mutations or cancer syndromes [35] with the most common of these being Lynch syndrome, which is present in up to 8% of all CRC cancers [36,37]. This study identified an apparent association between CFTR carrier status and CRC was identified, with a higher than expected proportion of individuals with CRC identified as carrying one or more CFTR mutations. Although one in 25 (4%) of the population of the UK are thought to carry a mutation[1], in the present study the incidence of mutations was higher (7%) and comparable to the incidence of Lynch syndrome. With an estimated population of 56·2 million people in England in 2021 this increase potentially has implications for between 1·69 and 2·25 million people. This supports other studies that have suggested CFTR carriers may have an increase incidence of CRC [12,38], Whilst the incidence of F508del in this study was comparable to previous studies [12], and close to the 4% quoted as the background carrier rate, the inclusion of additional mutations alongside F508del, suggests that the relationship may extend beyond this mutation alone. The incidence of R117H and the lower than expected prevalence of G551D in this work demonstrates the need for further investigations into the relationship between CFTR mutations and cancer risk, particularly around a possible association with class of mutations and residual function of the CFTR protein. The impact of R11H7 on CFTR expression largely depends on the polythymidine sequence, with 5T, 7T and 9T types all being present but with 5T being associated with the most severe loss of function. There has been little work to investigate this in the context of cancer risk.

This study is limited by small numbers due to the fact CF is a rare disease. In a small sample size, the approach used to identify CF and CRC can significantly affect the results, by influencing the ability to identify the conditions. To mitigate this, the analyses were undertaken using multiple data sources, one of which contains the gold-standard for cancer registration data (CORECT-R) and one of which is the gold-standard for CF registration data (UK CF Registry). As expected, the crude results from the UK CF Registry and SUS data analysis (Table 2) were lower than the primary analysis, suggesting that the approach taken in this study was appropriate, accurate and not failing to capture a significant number of cases. There was insufficient power to perform any detailed analysis of relationship with transplant status, this has been demonstrated to be a significant factor when assessing cancer risk, and further work is required to quantify this relationship.

Individuals with CF and carriers of the CF gene are at increased risk of CRC. This risk may further increase in the CF population following the recent introduction of highly effective modulators, which will inevitably result in improved survival and a more rapid increase in the age of the population, a key risk factor for CRC. A better understanding of the role of CFTR dysfunction in the development and natural history of CRC cancer is needed to inform screening, and possibly treatment strategies for those with CFTR mutations.

Ethics

The CORECT-R resource, and analyses based upon the data within it, has received approval from the Southwest-Central Bristol research ethics committee (18/SW/0134).

NHS research ethics approval (07/Q0104/2 UK Cystic Fibrosis Registry, AB/AM04/1) has been granted for collection of data into the UK CF Registry. Each patient provided written informed consent for data collection and for use of anonymised data in research. Under the terms of the NHS ethics approval the UK CF Trust steering committee approved the use of anonymised data in this study.

The 100,000 Genomes Project received ethics approval from the Health Research Authority Committee East of England, Cambridge South (Reference: 14/EE/1112). This project is covered by this ethics approval and as such a separate informed consent is not applicable.

DP, RB and AD conceived the study idea. RB undertook the analysis on the data from CORECT-R, HW undertook the analysis using the Genomic England data and RK-H performed the analysis used for the validation using the UK CF Registry and SUS data. DF, RC, KB, RK-H, NB, PQ, HW, AD and RB contributed to the interpretation of the results. RB took the lead in writing the manuscript with input from all the authors. RB, DP, AD, RC, KB, RK-H, NB, HW and PQ all approved the final version for submission.

Author contributions

DP, RB and AD conceived the study idea. RB undertook the analysis on the data from CORECT-R, HW undertook the analysis using the Genomic England data and RK-H performed the analysis used for the validation using the UK CF Registry and SUS data. DF, RC, KB, RK-H, NB, PQ, HW, AD and RB contributed to the interpretation of the results. RB took the lead in writing the manuscript with input from all the authors. RB, DP, AD, RC, KB, RK-H, NB, HW and PQ all approved the final version for submission.

Data sharing

The data used as part of this study are available from CORECT-R and the Cystic Fibrosis Trust (www.cysticfibrosis.org.uk/registry). Restrictions apply to the availability of these data, which are used under license in the current study. Data are available on request from the owners of the resources described subject to relevant ethical approval.

Primary data from the 100,000 Genomes Project, which are held in a secure Research Environment, are available to registered users. Please see https://www.genomicsengland.co.uk/about-gecip/for-gecip-members/data-and-data-access for further information.

Acknowledgements

This work uses data that has been provided by patients and collected by the NHS as part of their care and support. The data are collated, maintained and quality assured by the National Cancer Registration and Analysis Service, which is part of NHS Digital. Access to the data was facilitated by the Office for Data Release and the Cancer Research UK funded UK Colorectal Cancer Intelligence Hub (C23424/A23706).

This research was made possible through access to the data and findings generated by the 100,000 Genomes Project. The 100,000 Genomes Project is managed by Genomics England Limited (a wholly owned company of the Department of Health and Social

Care). The 100,000 Genomes Project 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. The 100,000 Genomes Project uses data provided by patients and collected by the National Health Service as part of their care and support.

We would like to thank people with cystic fibrosis and their families whose consent to participate in the UK CF Registry made the validation of this research possible.

References

[1] Cystic Fibrosis Trust. UK distribution of cystic fibrosis. 2020. https://www.cysticfibrosis.org.uk/the-work-we-do/uk-cf-registry/reporting-and-resources/uk-distribution-of-cystic-fibrosis.
[2] UK Cystic Fibrosis Registry. 2020 annual data report - Version 2, 2021.
[3] Keogh RH, Tanner K, Simmonds NJ, Bilton D. The changing demography of the cystic fibrosis population: forecasting future numbers of adults in the UK. Sci Rep 2020;10(1):10660.
[4] Yamada A., Komaki Y., Komaki F., Micic D., Zullow S., Sakuraba AJTLO. Risk of gastrointestinal cancers in patients with cystic fibrosis: a systematic review and meta-analysis. 2018; 19(6): 758–67.
[5] Fink AK, Yanik EL, Marshall BC, et al. Cancer risk among lung transplant recipients with cystic fibrosis. J Cyst Fibros 2017;16(1):91–7.
[6] Maisonneuve P, Lowenfels AB. Cancer in cystic fibrosis: a narrative review of prevalence, risk factors, screening, and treatment challenges: adult cystic fibrosis series. Chest 2022;161(2):356–64.
[7] Scott P, Anderson K, Singhania M, Cormier R. Cystic fibrosis, C.F.T.R., and colorectal cancer. Int J Mol Sci 2020;21(8):2891.
[8] Sosnay PR, Siklosi KR, Van Goor F, et al. Defining the disease liability of variants in the cystic fibrosis transmembrane conductance regulator gene. Nat Genet 2013;45(10):1160–7.
[9] Clinical and functional translation of CFTR. CFTR2 variant list history. https://cftr2.org/mutations_history. Accessed: 16/12/2021.
[10] Johannesson M, Askling J, Montgomery SM, Ekbom A, Bahmanyar S. Cancer risk among patients with cystic fibrosis and their first-degree relatives. Int J Cancer 2009;125(12):2953–6.
[11] Padua RA, Warren N, Grimshaw D, et al. The cystic fibrosis delta F508 gene mutation and cancer. Hum Mutat 1997;10(1):45–8.
[12] Shi Z, Wei J, Na R, et al. Cystic fibrosis F508del carriers and cancer risk: results from the UK Biobank. Int J Cancer 2020;10(10):33431.
[13] Schwarz MJ, Malone GM, Haworth A, et al. Cystic fibrosis mutation analysis: report from 22 UK regional genetics laboratories. Hum. Mutat. 1995;6(4):326–33.
[14] McCormick J, Green MW, Mehta G, Culross F, Mehta A. Demographics of the UK cystic fibrosis population: implications for neonatal screening. Eur J Hum Genet 2002;10(10):583–90.
[15] Cancer research UK. Bowel cancer incidence by age. 2021. https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/incidence#heading-One.
[16] University of Oxford. UK Colorectal Cancer Intelligence Hub. 2020. https://www.ndph.ox.ac.uk/corectr/about.
[17] Turnbull C, Scott RH, Thomas E, et al. The 100 000 genomes project: bringing whole genome sequencing to the NHS. BMJ 2018:361.
[18] Griffiths R, Schlüter DK, Akbari A, Cosgriff R, Tucker D, Taylor-Robinson D. Identifying children with cystic fibrosis in population-scale routinely collected data in wales: a retrospective review. Int J Popul Data Sci 2020;5(1) 1346-.
[19] British Lung Foundation. Emergency Hospital Admissions. 2021. https://statistics.blf.org.uk/cystic-fibrosis.
[20] Cystic Fibrosis Trust. UK cystic fibrosis registry. 2019.
[21] Strongman H, Williams R, Bhaskaran K. What are the implications of using individual and combined sources of routinely collected data to identify and characterise incident site-specific cancers? a concordance and validation study using linked English electronic health records data. BMJ Open 2020;10(8):e037719.
[22] Statistics OfN Estimates of the population for the UK, England and Wales. Scotland and Northern Ireland; 2019.
[23] Vandenbroucke JP, Pearce N. Incidence rates in dynamic populations. Int J Epidemiol 2012;41(5):1472–9.
[24] Raczy C, Petrovski R, Saunders CT, et al. Isaac: ultra-fast whole-genome secondary analysis on Illumina sequencing platforms. Bioinformatics 2013;29(16):2041–3.
[25] Kim S, Scheffler K, Halpern AL, et al. Strelka2: fast and accurate calling of germline and somatic variants. Nat Methods 2018;15(8):591–4.
[26] Brock DJ, Gilfillan A, Holloway S. The incidence of cystic fibrosis in Scotland calculated from heterozygote frequencies. Clin Genet 1998;53(1):47–9.
[27] Jarosz-Griffiths HH, Scambler T, Wong CH, et al. Different CFTR modulator combinations downregulate inflammation differently in cystic fibrosis. Elife 2020;9.
[28] Amaral MD, Quaresma MC, Pankonien I. What role does CFTR play in development, differentiation, regeneration and cancer? Int J Mol Sci 2020;21(9).
[29] Dayama G, Priya S, Niccum DE, Khoruts A, Blekhman R. Interactions between the gut microbiome and host gene regulation in cystic fibrosis. Genome Med 2020;12(1):12.
[30] Mangone L., Pinto C., Mancuso P., et al. Colon cancer survival differs from right side to left side and lymph node harvest number matter. 2021; 21(1): 1–10.
[31] Nawa T., Kato J., Kawamoto H., et al. Differences between right-and left-sided colon cancer in patient characteristics, cancer morphology and histology. 2008; 23(3): 418–23.
[32] UK cystic fibrosis registry. Annual data report 2012, 2013.
[33] Hadjiliadis D, Khoruts A, Zauber AG, Hempstead SE, Maisonneuve P, Lowenfels AB. Cystic fibrosis colorectal cancer screening consensus recommendations. Gastroenterology 2018;154(3):736–45.
[34] Gini A, Zauber AG, Cenin DR, et al. Cost effectiveness of screening individuals with cystic fibrosis for colorectal cancer. Gastroenterology 2018;154(3):556–67.
[35] Lynch HT ST, Watson P, Lanspa SJ, Lynch JF, Lynch PM, Cavalieri RJ, Boland CR. Genetics, natural history, tumor spectrum, and pathology of hereditary nonpolyposis colorectal cancer: an updated review. Gastroenterology 1993;104(5):1535–49.
[36] Lynch HT DlCAGenetic susceptibility to non-polyposis colorectal cancer. J Med Genet 1999;36(11):801–18.
[37] Lynch HT BC, Gong G, Shaw TG, Lynch PM, Fodde R, LynchJF de la Chappelle A. Phenotypic and genotypic heterogeneity in the Lynch syndrome: diagnostic, surveillance and management implications. Eur J Hum Genet 2006;14:390–402.
[38] Miller AC, Comellas AP, Hornick DB, et al. Cystic fibrosis carriers are at increased risk for a wide range of cystic fibrosis-related conditions. *Proc Natl Acad Sci U S A* 2020;117(3):1621–7.