Supplemental Online Content

Atkins JL, Pilling LC, Masoli JAH, et al. Association of hemochromatosis *HFE* p.C282Y homozygosity with hepatic malignancy. *JAMA*. doi:10.1001/jama.2020.21566

eMethods

eTables

eFigures

This supplemental material has been provided by the authors to give readers additional information about their work.

Association of hemochromatosis HFE p.C382Y homozygosity with hepatic malignancy

Atkins et al, JAMA 2020

Supplementary Information

Supplementary Methods

Genotyping of HFE p.C282Y and p.H63D

Genotyping was performed using two near-identical (>95%) Affymetrix microarrays: 49,950 participants involved in the UK Biobank Lung Exome Variant Evaluation (UK BiLEVE) were genotyped using the UK BiLEVE Axiom Array, with the remaining 438,427 participants using the UK Biobank Axiom Array. Data on 805,426 genetic variants were available in 488,377 UK Biobank participants after genotype calling and quality control performed centrally by the UK Biobank team (Bycroft et al., 2018).

Imputation of genetic variants (Bycroft et al., 2018) was performed firstly using data from the Haplotype Reference Consortium (39,235,157 single nucleotide polymorphisms (SNPs)) and secondly using a combined panel from 1000 Genomes Phase 3 and UK10K (87,696,888 bi-allelic markers) using the IMPUTE4 software (https://jmarchini.org/software/). 93,095,623 autosomal SNPs, short indels and large structural variants were available in 487,442 participants after extensive quality controls by the central UK Biobank team (Bycroft et al., 2018). Reference SNP IDs (rsIDs) were assigned to variants using data from the UCSC genome annotation database for the GRCh37 assembly of the human genome (http://hgdownload.cse.ucsc.edu/goldenpath/hg19/database/).

We selected 451,369 participants identified as 'white European' through self-report and verified through principal components (PC) analysis based on genotypes (Pilling et al., 2019). Briefly, PCs were generated in the 1000 Genomes Cohort using high-confidence genetic variants to obtain their individual loadings. These loadings were used to project UK Biobank participants into the same PC space. PCs 1 to 4 were used to identify participants of European descent. Related individuals were identified through kinship analysis.

p.C282Y is a SNP on chromosome 6 (b37 position 26093141, rsID rs1800562) in the *HFE* gene (OMIM #613609). p.H63D is a SNP also in the *HFE* gene (b37 position 26091179, rsID rs1799945). Both are missense variants causing a single amino-acid change in *HFE*. In the UK Biobank participants the SNPs were not correlated (R^2=0.015).

rs1800562 was not directly genotyped, therefore imputed p.C282Y genotypes were used. The imputation INFO quality was 0.997; 445, 462 participants (98.7% of 451,369) were imputed with 100% confidence. We recoded any imprecise imputation values using the following criteria: 5,724 were recoded (i.e. estimated genotype dose between 0 and 0.25 set to 0, values between 0.75 and 1.25 set to 1, and finally between 1.75 and 2 to 2); 183 participants (0.04%) were excluded due to imprecise imputation, yielding 451,186 participants in analyses.

Exome sequence data from the Illumina NovaSeq 6000 platform (75bp paired-end sequencing with 20x coverage of targeted panel) was available on 49,772 of these participants (http://biobank.ctsu.ox.ac.uk/crystal/label.cgi?id=170). *HFE* p.C282Y genotype calls were extracted from the data (FE pipeline (Regier et al., 2018)) and compared to imputed genotype calls, as described above. The correlation coefficient between genotypes from imputation and sequencing was high (R^2=0.998) and only one of 231 (0.4%) imputed p.C282Y homozygotes was incorrectly classified in this subset (the sequenced genotype was p.C282Y heterozygote)

References

Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L. T., 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. (2018). The UK Biobank resource with deep phenotyping and genomic data. *Nature*, *562*(7726), 203–209. https://doi.org/10.1038/s41586-018-0579-z

Pilling, L. C., Tamosauskaite, J., Jones, G., Wood, A. R., Jones, L., Kuo, C. L., Kuchel, G. A., Ferrucci, L., & Melzer, D. (2019). Common conditions associated with hereditary haemochromatosis genetic variants: cohort study in UK Biobank. *BMJ (Clinical Research Ed.)*, *364*, k5222. https://doi.org/10.1136/bmj.k5222

Regier, A. A., Farjoun, Y., Larson, D. E., Krasheninina, O., Kang, H. M., Howrigan, D. P., Chen, B.-J., Kher, M., Banks, E., Ames, D. C., English, A. C., Li, H., Xing, J., Zhang, Y., Matise, T., Abecasis, G. R., Salerno, W., Zody, M. C., Neale, B. M., & Hall, I. M. (2018). Functional equivalence of genome sequencing analysis pipelines enables harmonized variant calling across human genetics projects. *Nature Communications*, *9*(1), 4038. https://doi.org/10.1038/s41467-018-06159-4

Supplementary Tables

Supplementary Table 1. *Baseline characteristics of sample by those included and those not included in the primary care follow-up sample*

	Primary care follow-up	Not included in primary care follow-up	Total	p-value for difference
N (%)	209,811 (46.5)	241,375 (53.5)	451,186	
Age (baseline) - mean (sd)	56.7 (8.0)	56.8 (8.0)	56.8 (8.0)	<0.001
Female, n (%)	114,512 (54.6)	130,322 (54.0)	244,834 (54.3)	<0.001
C282Y +/+ genotype[a]	1,382 (0.7)	1,508 (0.6)	2,890 (0.6)	0.10
Obesity (BMI > 30), n (%)	52,848 (25.2)	57,351 (23.8)	110,199 (24.4)	<0.001
Waist circumference (cm) - mean (sd)	90.3 (13.5)	90.2 (13.5)	90.3 (13.5)	0.009
Diabetes (any type), baseline, n (%)	9,720 (4.6)	11,190 (4.6)	20,910 (4.6)	0.96
Current smoker, n (%)	21,859 (10.4)	25,036 (10.4)	46,895 (10.4)	0.62
Consuming alcohol daily, n (%)	42731 (20.4)	52,687 (21.8)	95,418 (21.2)	<0.001
Hepatitis (any), n (%)	944 (0.5)	1,284 (0.5)	2,228 (0.5)	<0.001
Cirrhosis, n (%)	125 (0.01)	197 (0.01)	322 (0.1)	0.01

[a]*HFE* p.C282Y homozygosity.

Supplementary Table 2. *Baseline characteristics of sample by all study genotypes and sex*

	Males						
Characteristics	**No variants**	**H63D+/-**	**H63D+/+**	**C282Y+/H63D+**	**C282Y+/-**	**C282Y +/+[a]**	**Total**
N (total sample)	122,860	47,990	4,674	4,955	24,579	1,294	206,352
N (subset with follow-up including primary care records)	56,425	22,351	2,148	2,297	11,466	612	95,299
Age (baseline) - mean (sd)	57.0 (8.1)	57.0 (8.1)	57.0 (8.1)	57.0 (8.1)	57.0 (8.1)	56.9 (8.2)	57.0 (8.1)
Obesity (BMI > 30), n (%)	31,767 (25.9)	12,397 (25.8)	1,219 (26.1)	1,223 (24.7)	6,328 (25.8)	295 (22.8)	53,229 (25.8)
Waist circumference (cm) - mean (sd)	97.0 (11.3)	97.1 (11.4)	97.2 (11.2)	97.0 (11.3)	97.1 (11.3)	96.9 (11.1)	97.1 (11.4)
Current smoker, n (%)	14,953 (12.2)	5,740 (12.0)	579 (12.4)	612 (12.4)	3,043 (12.4)	189 (14.6)	25,116 (12.2)
Consuming alcohol daily, n (%)	32,471 (26.5)	12,630 (26.3)	1,242 (26.6)	1,231 (24.9)	6,313 (25.7)	324 (25.1)	54,211 (26.3)
Polycythemia, n (%) (from baseline blood counts)[b]	2,386 (2.0)	1,073 (2.3)	111 (2.5)	191 (4.0)	627 (2.6)	35 (2.8)	4,423 (2.20
Liver enzymes							
Alanine aminotransferase (>50 U/L), n (%)	7,197 (6.1)	2,801 (6.1)	302 (6.8)	338 (7.1)	1,463 (6.2)	181 (14.7)	12,282 (6.2)
(95% CI)	(6.0-6.3)	(5.9-6.4)	(6.1-7.6)	(6.4-7.9)	(5.9-6.6)	(12.8-16.9)	(6.1-6.4)
Aspartate aminotransferase (>45 U/L), n (%)	4,814 (4.1)	1,861 (4.1)	211 (4.8)	226 (4.8)	1,032 (4.4)	129 (10.6)	8,273 (4.2)
(95% CI)	(4.0-4.2)	(3.9-4.3)	(4.2-5.4)	(4.2-5.4)	(4.2-4.7)	(8.9-12.4)	(4.1-4.3)
Prevalent disease							
Diabetes (any type), n (%)	7,876 (6.4)	3,050 (6.4)	293 (6.3)	311 (6.3)	1,601 (6.5)	113 (8.7)	13,244 (6.4)
Hepatitis (any), n (%)	675 (0.6)	276 (0.6)	28 (0.6)	24 (0.5)	145 (0.6)	5 (0.4)	1153 (0.6)
Cirrhosis, n (%)	85 (0.1)	39 (0.1)	3 (0.1)	2 (<0.1)	22 (0.1)	6 (0.5)	157 (0.1)

	Females						
Characteristics	**No variants**	**H63D+/-**	**H63D+/+**	**C282Y+/H63D+**	**C282Y+/-**	**C282Y +/+[a]**	**Total**
N (total sample)	145,719	57,029	5,584	5,746	29,160	1,596	244,834
N (subset with follow-up including primary care records)	67,916	26,632	2,640	2,754	13,800	770	114,512
Age (baseline) - mean (sd)	56.6 (7.9)	56.6 (7.9)	56.6 (8.1)	56.5 (7.9)	56.5 (8.0)	56.9 (8.0)	56.6 (7.9)
Obesity (BMI > 30), n (%)	33,997 (23.3)	13,193 (23.1)	1,358 (24.3)	1,353 (23.6)	6715 (23.0)	354 (22.2)	56,970 (23.3)
Waist circumference (cm) - mean (sd)	84.6 (12.5)	84.5 (12.4)	84.7 (12.6)	84.9 (12.6)	84.5 (12.5)	84.6 (12.5)	84.6 (12.5)
Current smoker, n (%)	12,837 (8.8)	5,124 (9.0)	537 (9.6)	509 (8.9)	2,634 (9.0)	138 (8.7)	21,779 (8.9)
Consuming alcohol daily, n (%)	24,603 (16.9)	9,675 (17.0)	908 (16.3)	932 (16.2)	4,833 (16.6)	256 (16.0)	41,207 (16.8)
Polycythemia, n (%)(from baseline blood counts)[b]	131 (0.1)	81 (0.2)	6 (0.1)	13 (0.2)	36 (0.1)	1 (0.1)	268 (0.1)
Liver enzymes							
Alanine aminotransferase (>50 U/L), n (%)	3,041 (2.2)	1,177 (2.2)	113 (2.1)	139 (2.5)	645 (2.3)	62 (4.1)	5,177 (2.2)
(95% CI)	(2.1-2.3)	(2.0-2.3)	(1.8-2.6)	(2.1-3.0)	(2.2-2.5)	(3.1-5.2)	(2.2-2.3)
Aspartate aminotransferase (>45 U/L), n (%)	2,792 (2.0)	1,058 (2.0)	107 (2.0)	135 (2.5)	521 (1.9)	54 (3.6)	4,667 (2.01)
(95% CI)	(1.9-2.1)	(1.8-2.1)	(1.7-2.4)	(2.1-2.9)	(1.7-2.1)	(2.7-4.6)	(2.0-2.1)
Prevalent disease							
Diabetes (any type), n (%)	4,478 (3.1)	1,828 (3.2)	185 (3.3)	193 (3.4)	923 (3.2)	59 (3.7)	7,666 (3.1)
Hepatitis (any), n (%)	648 (0.4)	243 (0.4)	29 (0.5)	30 (0.5)	120 (0.4)	5 (0.3)	1075 (0.4)
Cirrhosis, n (%)	93 (0.1)	38 (0.1)	7 (0.1)	3 (0.1)	23 (0.1)	1 (0.1)	165 (0.1)

Total sample (n=451,186).

[a]*HFE* p.C282Y homozygosity.

[b]Polycythemia defined as elevated hemoglobin (>16.5 g/dL for men; >16 g/dL for women) and hematocrit (> 49% for men; > 48% for women)[19].

***Supplementary Table 3**. Incident disease and death during follow-up for the entire cohort and primary care subset, by all study genotypes and sex*[a]

Characteristics	Males						
	No variants	**H63D+/-**	**H63D+/+**	**C282Y+/H63D+**	**C282Y+/-**	**C282Y +/+**	**Total**
N (total sample)[b]	122,860	47,990	4,674	4,955	24,579	1,294	206,352
Hemochromatosis diagnosis, baseline, n (%)	30 (<0.1)	17 (<0.1)	8 (0.2)	29 (0.6)	27 (0.1)	156 (12.1)	267 (0.1)
Hemochromatosis diagnosis, end of follow-up (baseline or incident), n (%)	74 (0.1)	47 (0.1)	24 (0.5)	89 (1.8)	54 (0.2)	327 (25.3)	615 (0.3)
Follow-up time (years), median (IQR)	8.9 (8.3-9.5)	8.9 (8.2-9.5)	8.9 (8.3-9.5)	8.9 (8.3-9.5)	8.9 (8.3-9.5)	8.9 (8.3-9.5)	8.9 (8.3-9.5)
Incident diagnoses, n (%)							
Liver disease (any, non-cancer)	1,598 (1.3)	639 (1.3)	58 (1.3)	80 (1.6)	321 (1.3)	48 (3.7)	2,744 (1.3)
Fibrosis and Cirrhosis	264 (0.2)	109 (0.2)	11 (0.2)	16 (0.3)	67 (0.2)	17 (1.3)	484 (0.2)
Alcoholic liver disease	309 (0.3)	112 (0.2)	13 (0.3)	18 (0.4)	96 (0.4)	8 (0.6)	556 (0.3)
Chronic hepatitis	23 (<0.1)	4 (<0.1)	0 (0)	2 (<0.1)	1 (<0.1)	1 (0.1)	31 (<0.1)
Hepatitis B	44 (<0.1)	16 (<0.1)	0 (0)	3 (0.1)	7 (<0.1)	0 (0)	70 (<0.1)
Hepatitis C	53 (<0.1)	18 (<0.1)	4 (0.1)	2 (<0.1)	11 (<0.1)	0 (0)	88 (<0.1)
Liver cancer	174 (0.1)	82 (0.2)	7 (0.2)	10 (0.2)	37 (0.2)	20 (1.6)	330 (0.2)
Hepatocellular carcinoma	81 (0.1)	43 (0.1)	3 (0.1)	7 (0.1)	16 (0.1)	14 (1.1)	164 (0.1)
Intrahepatic bile duct carcinoma	63 (0.1)	28 (0.1)	2 (<0.1)	3 (0.1)	17 (0.1)	5 (0.4)	118 (0.1)
Deaths, n (%)							
All-cause	6,560 (5.3)	2,633 (5.5)	258 (5.5)	260 (5.3)	1,345 (5.5)	88 (6.8)	11,144 (5.4)
Non-cancer liver disease (any)	141 (0.1)	56 (0.1)	4 (0.1)	5 (0.1)	35 (0.1)	1 (0.1)	242 (0.1)
Liver cancer	102 (0.1)	41 (0.1)	6 (0.1)	6 (0.1)	18 (0.1)	14 (1.1)	187 (0.1)
N (subset with follow-up including primary care records)[c]	56,425	22,351	2,148	2,297	11,466	612	95,299
Incident diagnosis, n (%)							
Liver disease (any, non-cancer)	1,448 (2.6)	550 (2.5)	53 (2.5)	64 (2.8)	276 (2.4)	42 (7.1)	2,433 (2.6)
Fibrosis and Cirrhosis	148 (0.3)	61 (0.3)	9 (0.4)	10 (0.4)	37 (0.3)	14 (2.3)	279 (0.3)
Alcoholic liver disease	272 (0.5)	89 (0.4)	10 (0.5)	12 (0.5)	64 (0.6)	9 (1.5)	456 (0.5)
Liver cancer	75 (0.1)	43 (0.2)	6 (0.3)	3 (0.1)	16 (0.1)	12 (2.0)	155 (0.2
Hepatocellular carcinoma	31 (0.1)	24 (0.1)	3 (0.1)	3 (0.1)	7 (0.1)	6 (1.0)	74 (0.1)
Intrahepatic bile duct carcinoma	31 (0.1)	14 (0.1)	1 (0.1)	0 (0)	5 (<0.1)	4 (0.7)	55 (0.1)

	Females						
Characteristics	**No variants**	**H63D+/-**	**H63D+/+**	**C282Y+/H63+**	**C282Y+/-**	**C282Y +/+**	**Total**
N (total sample)[b]	145,719	57,029	5,584	5,746	29,160	1,596	244,834
Hemochromatosis diagnosis, baseline, n (%)	8 (<0.1)	6 (<0.1)	2 (<0.1)	12 (0.2)	5 (<0.1)	54 (3.4)	87 (<0.1)
Hemochromatosis diagnosis, end of follow-up (baseline or incident), n (%)	25 (<0.1)	26 (<0.1)	10 (0.2)	35 (0.6)	23 (0.1)	200 (12.5)	319 (0.1)
Follow-up time (years), median (IQR)	8.9 (8.3-9.5)	8.9 (8.3-9.5)	8.9 (8.3-9.5)	8.9 (8.3-9.5)	8.9 (8.3-9.5)	9.0 (8.4-9.5)	8.9 (8.3-9.5)
Incident diagnoses, n (%)							
Liver disease (any, non-cancer)	1,433 (1.0)	583 (1.0)	56 (1.0)	68 (1.2)	310 (1.1)	23 (1.5)	2,473 (1.0)
Fibrosis and Cirrhosis	229 (0.2)	70 (0.1)	7 (0.1)	9 (0.2)	42 (0.1)	4 (0.3)	361 (0.2)
Alcoholic liver disease	77 (0.1)	29 (0.1)	7 (0.1)	2 (<0.1)	25 (0.1)	4 (0.3)	144 (0.1)
Chronic hepatitis	26 (<0.1)	11 (<0.1)	2 (<0.1)	1 (<0.1)	3 (<0.1)	1 (0.1)	44 (<0.1)
Hepatitis B	11 (<0.1)	8 (<0.1)	0 (0)	1 (<0.1)	9 (<0.1)	1 (0.1)	30 (<0.1)
Hepatitis C	14 (<0.1)	17 (<0.1)	1 (<0.1)	0 (0)	5 (<0.1)	1 (0.1)	38 (<0.1)
Liver cancer	129 (0.1)	68 (0.1)	4 (0.1)	4 (0.1)	29 (0.1)	3 (0.2)	237 (0.1)
Hepatocellular carcinoma	30 (<0.1)	14 (<0.1)	2 (<0.1)	0 (0)	8 (<0.1)	1 (0.1)	55 (<0.1)
Intrahepatic bile duct carcinoma	69 (0.1)	39 (0.1)	2 (<0.1)	3 (0.1)	17 (0.1)	0 (0)	130 (0.1)
Deaths, n (%)							
All-cause	4,362 (3.0)	1,655 (2.9)	158 (2.83)	154 (2.7)	904 (3.1)	60 (3.8)	7,293 (3.0)
Non-cancer liver disease (any)	43 (<0.1)	15 (<0.1)	3 (0.1)	2 (<0.1)	10 (<0.1)	0 (0)	73 (<0.1)
Liver cancer	78 (<0.1)	41 (0.1)	4 (0.1)	3 (0.1)	19 (0.1)	3 (0.2)	148 (0.1)
N (subset with follow-up including primary care records)[c]	67,916	26,632	2,640	2,754	13,800	770	114,512
Incident diagnosis, n (%)							
Liver disease (any, non-cancer)	1305 (1.9)	566 (1.1)	60 (2.3)	70 (2.6)	314 (2.3)	22 (2.9)	2337 (2.2)
Fibrosis and Cirrhosis	151 (0.2)	45 (0.2)	6 (0.2)	6 (0.2)	40 (0.3)	4 (0.5)	252 (0.2)
Alcoholic liver disease	72 (0.1)	33 (0.1)	5 (0.2)	4 (0.2)	26 (0.2)	4 (0.5)	144 (0.1)
Liver cancer	62 (0.1)	28 (0.1)	2 (0.1)	1 (<0.1)	13 (0.1)	1 (0.1)	107 (0.1)
Hepatocellular carcinoma	16 (<0.1)	7 (<0.1)	1 (<0.1)	0 (0)	3 (<0.1)	0 (0)	27 (<0.1)
Intrahepatic bile duct carcinoma	31 (0.1)	16 (0.1)	1 (<0.1)	1 (<0.1)	8 (0.1)	0 (0)	57 (0.1)

[a]Excluding individuals with each prevalent diagnosis at baseline only: i.e. the incident liver cancer analysis included participants with baseline non-cancer liver disease.

[b]Total sample (n=451,186), follow-up data from Hospital Episode Statistics, the Cancer Registry and Death Records.

[c]Subset of participants (n=209,811) with additional follow-up including primary care records.

Supplementary Table 4. *Hazards ratios (95% CI) for incident liver outcomes by C282Y-H63D genotype and sex, adjusted for age, assessment center, genotyping array and PCs 1-10*

Incident outcome	Sex	Genotype	Cases	Person years	HR	95% CI	p-value	Incidence rate difference (per 1,000 person years)[a]	95% CI
Liver disease (any, non-cancer)	Females	No variants	1433	1160712	1.00				
		H63D+/-	583	454536	1.04	0.94-1.14	0.44		
		H63D+/+	56	44436	1.02	0.78-1.33	0.90		
		C282Y+/H63D+	68	45883	1.20	0.94-1.53	0.15		
		C282Y+/-	310	232593	1.07	0.95-1.21	0.26		
		C282Y +/+	23	12783	1.42	0.94-2.15	0.09	0.56	-0.17-1.3
Liver disease (any, non-cancer)	Males	No variants	1598	968897	1.00				
		H63D+/-	639	377687	1.02	0.93-1.12	0.62		
		H63D+/+	58	36806	0.95	0.73-1.24	0.72		
		C282Y+/H63D+	80	38989	1.23	0.99-1.55	0.07		
		C282Y+/-	321	193889	0.99	0.88-1.12	0.88		
		C282Y +/+	48	9916	2.90	2.18-3.87	<0.001	3.19	1.82-4.56
Fibrosis and Cirrhosis	Females	No variants	229	1167594	1.00				
		H63D+/-	70	457214	0.78	0.60-1.02	0.07		
		H63D+/+	7	44741	0.79	0.37-1.68	0.55		
		C282Y+/H63D+	9	46139	1.00	0.51-1.94	0.99		
		C282Y+/-	42	234142	0.91	0.66-1.27	0.59		
		C282Y +/+	4	12877	1.55	0.58-4.18	0.38	0.11	-0.19-0.42
Fibrosis and Cirrhosis	Males	No variants	264	975954	1.00				
		H63D+/-	109	380491	1.05	0.84-1.32	0.64		
		H63D+/+	11	37101	1.09	0.60-2.00	0.77		
		C282Y+/H63D+	16	39423	1.46	0.88-2.41	0.14		
		C282Y+/-	67	195430	1.23	0.94-1.61	0.13		
		C282Y +/+	17	10079	5.98	3.65-9.79	0.12	1.42	0.61-2.22

Incident outcome	Sex	Genotype	Cases	Person years	HR	95% CI	p-value	Incidence rate difference (per 1,000 person years)[a]	95% CI
Alcoholic liver disease	Females	No variants	77	1168875	1.00				
		H63D+/-	29	457654	0.96	0.62-1.47	0.84		
		H63D+/+	7	44799	2.33	1.07-5.05	0.03		
		C282Y+/H63D+	2	46182	0.63	0.15-2.55	0.51		
		C282Y+/-	25	234315	1.54	0.98-2.42	0.06		
		C282Y +/+	4	12905	4.21	1.54-11.54	0.01	0.24	-0.06-0.55
Alcoholic liver disease	Males	No variants	309	975393	1.00				
		H63D+/-	112	380327	0.93	0.75-1.15	0.48		
		H63D+/+	13	37048	1.10	0.63-1.91	0.74		
		C282Y+/H63D+	18	39368	1.38	0.86-2.22	0.19		
		C282Y+/-	96	195272	1.48	1.18-1.87	0.001		
		C282Y +/+	8	10221	2.26	1.12-4.56	0.02	0.47	-0.08-1.01
Chronic hepatitis	Females	No variants	26	1168964	1.00				
		H63D+/-	11	457723	1.07	0.53-2.18	0.84		
		H63D+/+	2	44804	1.94	0.46-8.18	0.37		
		C282Y+/H63D+	1	46203	0.96	0.13-7.12	0.97		
		C282Y+/-	3	234423	0.57	0.17-1.88	0.35		
		C282Y +/+	1	12907	3.19	0.43-23.61	0.26	0.06	-0.1-0.21
Chronic hepatitis	Males	No variants	23	977228	1.00				
		H63D+/-	4	381086	0.45	0.15-1.29	0.14		
		H63D+/+	0	37141	No observations				
		C282Y+/H63D+	2	39455	2.33	0.55-9.95	0.25		
		C282Y+/-	1	195772	0.23	0.03-1.74	0.16		
		C282Y +/+	1	10277	4.69	0.62-35.29	0.13	0.07	-0.12-0.26

Incident outcome	Sex	Genotype	Cases	Person years	HR	95% CI	p-value	Incidence rate difference (per 1,000 person years)[a]	95% CI
Hepatitis B	Females	No variants	11	1169150	1.00				
		H63D+/-	8	457783	1.87	0.75-4.65	0.18		
		H63D+/+	0	44824	No observations				
		C282Y+/H63D+	1	46202	2.74	0.35-21.38	0.34		
		C282Y+/-	9	234466	4.54	1.86-11.07	<0.001		
		C282Y +/+	1	12914	8.37	1.06-66.03	0.04	0.07	-0.08-0.22
Hepatitis B	Males	No variants	44	976948	1.00				
		H63D+/-	16	380918	0.92	0.52-1.64	0.79		
		H63D+/+	0	37131	No observations				
		C282Y+/H63D+	3	39484	1.71	0.53-5.52	0.37		
		C282Y+/-	7	195682	0.80	0.36-1.79	0.59		
		C282Y +/+	0	10280	No observations			-0.05	-0.06--0.03
Hepatitis C	Females	No variants	14	1168797	1.00				
		H63D+/-	17	457667	3.14	1.55-6.38	0.002		
		H63D+/+	1	44792	1.91	0.25-14.54	0.53		
		C282Y+/H63D+	0	46208	No observations				
		C282Y+/-	5	234431	1.77	0.63-4.93	0.28		
		C282Y +/+	1	12909	6.56	0.86-50.31	0.07	0.07	-0.09-0.22
Hepatitis C	Males	No variants	53	976467	1.00				
		H63D+/-	18	380825	0.87	0.51-1.49	0.61		
		H63D+/+	4	37108	1.92	0.70-5.32	0.21		
		C282Y+/H63D+	2	39451	0.93	0.23-3.83	0.92		
		C282Y+/-	11	195567	1.05	0.55-2.02	0.88		
		C282Y +/+	0	10271	No observations			-0.05	-0.07--0.04

Incident outcome	Sex	Genotype	Cases	Person years	HR	95% CI	p-value	Incidence rate difference (per 1,000 person years)[a]	95% CI
Liver cancer	Females	No variants	129	1033722	1.00				
		H63D+/-	68	404740	1.36	1.01-1.82	0.04		
		H63D+/+	4	39619	0.81	0.30-2.20	0.68		
		C282Y+/H63D+	4	40867	0.80	0.30-2.16	0.66		
		C282Y+/-	29	207412	1.13	0.75-1.68	0.57		
		C282Y +/+	3	11436	2.06	0.66-6.49	0.22	0.14	-0.16-0.44
Liver cancer	Males	No variants	174	865267	1.00				
		H63D+/-	82	337446	1.20	0.92-1.56	0.17		
		H63D+/+	7	32903	1.06	0.50-2.26	0.88		
		C282Y+/H63D+	10	34965	1.40	0.74-2.65	0.30		
		C282Y+/-	37	173376	1.03	0.72-1.47	0.88		
		C282Y +/+	20	9111	10.47	6.56-16.69	<0.001	1.99	1.03-2.96
Hepatocellular carcinoma	Females	No variants	30	1033754	1.00				
		H63D+/-	14	404761	1.19	0.63-2.25	0.59		
		H63D+/+	2	39619	1.70	0.41-7.13	0.47		
		C282Y+/H63D+	0	40868	No observations				
		C282Y+/-	8	207425	1.33	0.61-2.91	0.48		
		C282Y +/+	1	11436	3.00	0.41-22.06	0.28	0.06	-0.11-0.23
Hepatocellular carcinoma	Males	No variants	81	865309	1.00				
		H63D+/-	43	337457	1.35	0.93-1.95	0.11		
		H63D+/+	3	32909	0.98	0.31-3.09	0.97		
		C282Y+/H63D+	7	34965	2.09	0.97-4.54	0.06		
		C282Y+/-	16	173380	0.94	0.55-1.61	0.83		
		C282Y +/+	14	9115	15.54	8.76-27.57	<0.001	1.44	0.64-2.25

Incident outcome	Sex	Genotype	Cases	Person years	HR	95% CI	p-value	Incidence rate difference (per 1,000 person years)[a]	95% CI
Intrahepatic bile duct carcinoma	Females	No variants	69	1033717	1.00				
		H63D+/-	39	404742	1.46	0.98-2.16	0.06		
		H63D+/+	2	39619	0.76	0.19-3.11	0.70		
		C282Y+/H63D+	3	40867	1.11	0.35-3.52	0.86		
		C282Y+/-	17	207416	1.23	0.72-2.09	0.45		
		C282Y +/+	0	11437	No observations			-0.07	-0.08--0.05
Intrahepatic bile duct carcinoma	Males	No variants	63	865306	1.00				
		H63D+/-	28	337479	1.14	0.73-1.77	0.57		
		H63D+/+	2	32905	0.83	0.20-3.39	0.79		
		C282Y+/H63D+	3	34967	1.15	0.36-3.65	0.82		
		C282Y+/-	17	173381	1.31	0.76-2.24	0.33		
		C282Y +/+	5	9126	7.18	2.87-17.99	<0.001	0.48	-0.01-0.96

Total sample (n=451,186), follow-up data from Hospital Episode Statistics, the Cancer Registry and Death Records.
Hazard ratios adjusted for age, assessment center, genotyping array and population genetics sub-structure using principal components.
Excluding individuals with each prevalent diagnosis at baseline only: i.e. the incident liver cancer analysis included participants with baseline non-cancer liver disease.
[a]Incident rate difference is the difference in incidence between participants with no variants and C282Y homozygotes (C282Y +/+).

Supplementary Table 5. *Hazards ratios (95% CI) for incident liver outcomes by C282Y-H63D genotype and sex, adjusted for age, assessment center, genotyping array and PCs 1-10 (in subset with primary care follow-up)*

Incident outcome	Sex	Genotype	Cases	Person years	HR	95% CI	p-value
Liver disease (any, non-cancer)	Females	No variants	1305	534640	1.00		
		H63D+/-	566	209846	1.11	1.00-1.22	0.05
		H63D+/+	60	20730	1.19	0.92-1.54	0.19
		C282Y+/H63D+	70	21777	1.33	1.05-1.69	0.02
		C282Y+/-	314	108807	1.18	1.05-1.34	0.01
		C282Y +/+	22	6124	1.47	0.97-2.25	0.07
Liver disease (any, non-cancer)	Males	No variants	1448	439531	1.00		
		H63D+/-	550	173847	0.96	0.87-1.06	0.39
		H63D+/+	53	16789	0.96	0.73-1.27	0.78
		C282Y+/H63D+	64	17976	1.08	0.84-1.38	0.56
		C282Y+/-	276	89540	0.93	0.82-1.06	0.29
		C282Y +/+	42	4578	2.80	2.06-3.80	<0.001
Fibrosis and Cirrhosis	Females	No variants	151	540199	1.00		
		H63D+/-	45	212177	0.75	0.54-1.05	0.10
		H63D+/+	6	21009	1.01	0.45-2.29	0.97
		C282Y+/H63D+	6	22030	0.98	0.43-2.22	0.97
		C282Y+/-	40	110108	1.29	0.91-1.83	0.15
		C282Y +/+	4	6186	2.28	0.84-6.17	0.10
Fibrosis and Cirrhosis	Males	No variants	148	445246	1.00		
		H63D+/-	61	176047	1.04	0.77-1.40	0.81
		H63D+/+	9	16972	1.59	0.81-3.12	0.18
		C282Y+/H63D+	10	18259	1.60	0.84-3.04	0.15
		C282Y+/-	37	90668	1.18	0.82-1.69	0.38
		C282Y +/+	14	4713	8.14	4.69-14.13	<0.001
Alcoholic liver disease	Females	No variants	72	540835	1.00		
		H63D+/-	33	212385	1.16	0.77-1.76	0.47
		H63D+/+	5	21030	1.79	0.72-4.44	0.21
		C282Y+/H63D+	4	22032	1.35	0.49-3.70	0.56
		C282Y+/-	26	110221	1.77	1.13-2.78	0.01
		C282Y +/+	4	6209	4.85	1.77-13.31	0.002
Alcoholic liver disease	Males	No variants	272	444633	1.00		
		H63D+/-	89	175873	0.82	0.65-1.05	0.11
		H63D+/+	10	16943	0.97	0.52-1.83	0.94
		C282Y+/H63D+	12	18240	1.05	0.59-1.87	0.87
		C282Y+/-	64	90508	1.11	0.84-1.46	0.45
		C282Y +/+	9	4782	2.91	1.49-5.66	0.002

Incident outcome	Sex	Genotype	Cases	Person years	HR	95% CI	p-value
Liver cancer	Females	No variants	62	480508	1.00		
		H63D+/-	28	188686	1.16	0.74-1.82	0.51
		H63D+/+	2	18687	0.83	0.20-3.40	0.80
		C282Y+/H63D+	1	19589	0.39	0.05-2.81	0.35
		C282Y+/-	13	98014	1.03	0.57-1.87	0.93
		C282Y +/+	1	5522	1.44	0.20-10.41	0.72
Liver cancer	Males	No variants	75	396487	1.00		
		H63D+/-	43	156770	1.44	0.99-2.09	0.06
		H63D+/+	6	15117	2.09	0.91-4.81	0.08
		C282Y+/H63D+	3	16272	1.00	0.32-3.18	1.00
		C282Y+/-	16	80735	1.02	0.59-1.75	0.95
		C282Y +/+	12	4314	14.11	7.62-26.12	<0.001
Hepatocellular carcinoma	Females	No variants	16	480540	1.00		
		H63D+/-	7	188692	1.13	0.47-2.76	0.78
		H63D+/+	1	18687	1.52	0.20-11.49	0.68
		C282Y+/H63D+	0	19589	No observations		
		C282Y+/-	3	98021	0.94	0.27-3.22	0.92
		C282Y +/+	0	5522	No observations		
Hepatocellular carcinoma	Males	No variants	31	396507	1.00		
		H63D+/-	24	156778	1.94	1.14-3.31	0.02
		H63D+/+	3	15117	2.61	0.80-8.53	0.11
		C282Y+/H63D+	3	16272	2.38	0.73-7.79	0.15
		C282Y+/-	7	80738	1.07	0.47-2.44	0.87
		C282Y +/+	6	4319	16.78	6.94-40.57	<0.001
Intrahepatic bile duct carcinoma	Females	No variants	31	480517	1.00		
		H63D+/-	16	188686	1.33	0.72-2.42	0.36
		H63D+/+	1	18687	0.83	0.11-6.05	0.85
		C282Y+/H63D+	1	19589	0.80	0.11-5.85	0.82
		C282Y+/-	8	98015	1.29	0.59-2.81	0.52
		C282Y +/+	0	5522	No observations		
Intrahepatic bile duct carcinoma	Males	No variants	31	396505	1.00		
		H63D+/-	14	156796	1.13	0.60-2.13	0.71
		H63D+/+	1	15119	0.83	0.11-6.07	0.85
		C282Y+/H63D+	0	16273	No observations		
		C282Y+/-	5	80734	0.78	0.30-2.01	0.60
		C282Y +/+	4	4316	11.18	3.87-32.24	<0.001

Subset of participants (n=209,811) with additional follow-up including primary care records.

Hazard ratios adjusted for age, assessment center, genotyping array and population genetics sub-structure using principal components.

Excluding individuals with each prevalent diagnosis at baseline only: i.e. the incident liver cancer analysis included participants with baseline non-cancer liver disease.

Supplementary Table 6. *Lifetable estimates for incident liver disease (any, non-cancer) by sex and HFE genotypes (no variants vs C282Y homozygotes)*

	No variants					C282Y homozygotes					Excess proportion (%)
Age interval, years	**Failures**	**Survival Function**	**Incident diagnosis (%)**	**95% CI**		**Failures**	**Survival Function**	**Incident diagnosis (%)**	**95% CI**		
Males											
40-45	25	0.99	1.0	0.6	1.7	0	1.00				
46-50	103	0.98	2.3	1.8	2.9	3	0.97	2.7	0.9	8.3	0.5
51-55	146	0.97	3.5	3.0	4.2	4	0.94	6.2	3.0	12.5	2.7
56-60	231	0.95	5.1	4.6	5.8	7	0.89	10.9	6.6	17.7	5.7
61-65	293	0.93	6.7	6.1	7.4	13	0.83	17.4	12.2	24.5	10.7
66-70	373	0.92	8.4	7.8	9.1	10	0.79	21.2	15.7	28.2	12.8
71-75	233	0.90	10.2	9.5	10.9	3	0.77	22.7	17.1	29.9	12.5
Females											
40-45	15	1.00	0.4	0.2	0.8	0	1.00				
46-50	66	0.99	1.1	0.8	1.5	1	0.99	0.8	0.1	5.9	-0.3
51-55	155	0.98	2.2	1.8	2.5	1	0.99	1.4	0.3	5.7	-0.7
56-60	230	0.97	3.4	3.1	3.8	4	0.97	3.4	1.5	7.5	0.0
61-65	251	0.95	4.5	4.1	4.9	5	0.95	5.5	3.0	9.8	1.0
66-70	329	0.94	5.8	5.4	6.3	4	0.93	6.9	4.1	11.3	1.0
71-75	215	0.93	7.3	6.9	7.8	6	0.90	9.7	6.3	14.7	2.4

Lifetable projections based on observed 5 year age-group specific incidence rates from ages 40 and 75 years e.g. in male p.C282Y homozygotes, 22.7% were projected to develop liver disease by age 75, compared to 10.2% without HFE variants (excess proportion 12.5%).

Liver disease estimates based on subset of participants with primary care data (n=209,811). Excluding individuals with prevalent diagnosis at baseline.

Supplementary Table 7. *Hazards ratios (95% CI) for mortality by C282Y-H63D genotype and sex, adjusted for age, assessment center, genotyping array and PCs 1-10*

Death	Sex	Genotype	Cases	Person years	HR	95% CI	p-value
All-cause	Females	No variants	4349	1286786	1.00		
		H63D+/-	1650	503708	0.97	0.92-1.03	0.35
		H63D+/+	158	49303	0.94	0.80-1.10	0.45
		C282Y+/H63D+	154	50787	0.90	0.76-1.05	0.19
		C282Y+/-	899	257676	1.03	0.95-1.10	0.48
		C282Y +/+	60	14152	1.18	0.91-1.52	0.20
All-cause	Males	No variants	6539	1074517	1.00		
		H63D+/-	2628	418951	1.02	0.98-1.07	0.30
		H63D+/+	258	40833	1.04	0.92-1.17	0.56
		C282Y+/H63D+	258	43330	0.96	0.85-1.09	0.51
		C282Y+/-	1338	214940	1.00	0.94-1.06	0.90
		C282Y +/+	88	11271	1.24	1.00-1.53	0.046
Liver disease (any, non cancer)	Females	No variants	43	1286786	1.00		
		H63D+/-	15	503708	0.89	0.50-1.60	0.70
		H63D+/+	3	49303	1.84	0.57-5.92	0.31
		C282Y+/H63D+	2	50787	1.13	0.27-4.67	0.87
		C282Y+/-	10	257676	1.11	0.55-2.21	0.77
		C282Y +/+	0	14152	No observations		
Liver disease (any, non cancer)	Males	No variants	141	1074517	1.00		
		H63D+/-	56	418951	1.01	0.74-1.37	0.96
		H63D+/+	4	40833	0.73	0.27-1.98	0.54
		C282Y+/H63D+	5	43330	0.81	0.33-1.99	0.65
		C282Y+/-	35	214940	1.16	0.80-1.68	0.44
		C282Y +/+	1	11271	0.59	0.08-4.21	0.60
Liver cancer	Females	No variants	78	1286786	1.00		
		H63D+/-	41	503708	1.36	0.93-1.98	0.12
		H63D+/+	4	49303	1.35	0.49-3.69	0.56
		C282Y+/H63D+	3	50787	0.99	0.31-3.15	0.99
		C282Y+/-	19	257676	1.21	0.74-2.01	0.45
		C282Y +/+	3	14152	3.36	1.06-10.67	0.04
Liver cancer	Males	No variants	102	1074517	1.00		
		H63D+/-	41	418951	1.02	0.71-1.47	0.91
		H63D+/+	6	40833	1.53	0.67-3.48	0.31
		C282Y+/H63D+	6	43330	1.42	0.62-3.23	0.41
		C282Y+/-	18	214940	0.84	0.51-1.39	0.50
		C282Y +/+	14	11271	11.91	6.77-20.97	<0.001

Total sample (n=451,186), follow-up data from Death Records.
Hazard ratios adjusted for age, assessment center, genotyping array and population genetics sub-structure using principal components.

Supplementary Table 8. *Hazards ratios (95% CI) for incident liver outcomes by C282Y-H63D genotype and sex, adjusted for age, assessment center, genotyping array and PCs 1-10 (excluding prevalent haemochromatosis)*

Incident outcome	Sex	Genotype	Cases	Person years	HR	95% CI	p-value
Liver disease (any)	Females	No variants	1432	1160667	1.00		
		H63D+/-	583	454491	1.04	0.94-1.14	0.44
		H63D+/+	56	44421	1.02	0.78-1.33	0.89
		C282Y+/H63+	66	45813	1.16	0.91-1.49	0.23
		C282Y+/-	310	232553	1.07	0.95-1.21	0.26
		C282Y +/+	20	12406	1.28	0.82-1.99	0.27
Liver disease (any)	Males	No variants	1598	968731	1.00		
		H63D+/-	638	377594	1.02	0.93-1.12	0.64
		H63D+/+	57	36741	0.94	0.72-1.22	0.64
		C282Y+/H63D+	78	38820	1.21	0.96-1.52	0.10
		C282Y+/-	319	193730	0.98	0.87-1.11	0.80
		C282Y +/+	41	8948	2.76	2.02-3.76	<0.001
Fibrosis and Cirrhosis	Females	No variants	228	1167543	1.00		
		H63D+/-	70	457169	0.78	0.6-1.02	0.07
		H63D+/+	7	44726	0.80	0.38-1.69	0.55
		C282Y+/H63+	9	46053	1.00	0.51-1.95	1.00
		C282Y+/-	42	234102	0.92	0.66-1.28	0.61
		C282Y +/+	4	12460	1.62	0.60-4.36	0.34
Fibrosis and Cirrhosis	Males	No variants	264	975773	1.00		
		H63D+/-	108	380375	1.04	0.84-1.31	0.70
		H63D+/+	11	37034	1.10	0.60-2.00	0.77
		C282Y+/H63D+	16	39204	1.47	0.88-2.43	0.14
		C282Y+/-	67	195255	1.23	0.94-1.61	0.13
		C282Y +/+	13	9038	5.17	2.96-9.05	0.00
Alcoholic liver disease	Females	No variants	76	1168823	1.00		
		H63D+/-	29	457610	0.97	0.63-1.48	0.88
		H63D+/+	7	44784	2.36	1.09-5.12	0.03
		C282Y+/H63+	2	46103	0.63	0.16-2.59	0.53
		C282Y+/-	25	234275	1.56	0.99-2.46	0.06
		C282Y +/+	3	12473	3.32	1.05-10.57	0.04
Alcoholic liver disease	Males	No variants	308	975209	1.00		
		H63D+/-	111	380219	0.92	0.74-1.14	0.45
		H63D+/+	13	36981	1.11	0.63-1.93	0.72
		C282Y+/H63D+	17	39157	1.31	0.81-2.14	0.27
		C282Y+/-	96	195079	1.49	1.18-1.87	0.00
		C282Y +/+	4	9064	1.28	0.48-3.44	0.62

Incident outcome	Sex	Genotype	Cases	Person years	HR	95% CI	p-value
Chronic hepatitis	Females	No variants	26	1168914	1.00		
		H63D+/-	11	457678	1.07	0.53-2.17	0.84
		H63D+/+	2	44789	1.93	0.46-8.14	0.37
		C282Y+/H63D+	1	46117	0.96	0.13-7.08	0.97
		C282Y+/-	3	234382	0.56	0.17-1.87	0.35
		C282Y +/+	0	12478	No observations		
Chronic hepatitis	Males	No variants	23	977030	1.00		
		H63D+/-	4	380967	0.45	0.15-1.29	0.14
		H63D+/+	0	37074	No observations		
		C282Y+/H63D+	2	39229	2.35	0.55-10.01	0.25
		C282Y+/-	1	195564	0.23	0.03-1.74	0.16
		C282Y +/+	1	9088	5.24	0.70-39.47	0.11
Hepatitis B	Females	No variants	11	1169092	1.00		
		H63D+/-	8	457739	1.87	0.75-4.64	0.18
		H63D+/+	0	44809	No observations		
		C282Y+/H63D+	1	46116	2.81	0.36-21.90	0.32
		C282Y+/-	9	234426	4.61	1.89-11.23	<0.001
		C282Y +/+	0	12486	No observations		
Hepatitis B	Males	No variants	44	976749	1.00		
		H63D+/-	15	380799	0.86	0.48-1.55	0.63
		H63D+/+	0	37064	No observations		
		C282Y+/H63D+	3	39258	1.72	0.53-5.56	0.36
		C282Y+/-	7	195474	0.80	0.36-1.79	0.59
		C282Y +/+	0	9090	No observations		
Hepatitis C	Females	No variants	14	1168739	1.00		
		H63D+/-	17	457631	3.14	1.55-6.38	0.002
		H63D+/+	1	44777	1.91	0.25-14.54	0.53
		C282Y+/H63D+	0	46122	No observations		
		C282Y+/-	5	234390	1.77	0.64-4.93	0.28
		C282Y +/+	1	12482	6.72	0.88-51.50	0.07
Hepatitis C	Males	No variants	53	976268	1.00		
		H63D+/-	18	380699	0.87	0.51-1.49	0.61
		H63D+/+	4	37042	1.93	0.70-5.34	0.21
		C282Y+/H63D+	2	39225	0.95	0.23-3.90	0.94
		C282Y+/-	11	195359	1.05	0.55-2.02	0.88
		C282Y +/+	0	9090	No observations		

Incident outcome	Sex	Genotype	Cases	Person years	HR	95% CI	p-value
Liver cancer	Females	No variants	129	1033671	1.00		
		H63D+/-	68	404702	1.36	1.01-1.82	0.04
		H63D+/+	4	39606	0.81	0.30-2.20	0.69
		C282Y+/H63+	4	40792	0.80	0.30-2.17	0.66
		C282Y+/-	29	207377	1.13	0.75-1.68	0.57
		C282Y +/+	3	11051	2.15	0.68-6.77	0.19
Liver cancer	Males	No variants	173	865090	1.00		
		H63D+/-	81	337334	1.20	0.92-1.56	0.19
		H63D+/+	7	32843	1.07	0.50-2.27	0.87
		C282Y+/H63D+	10	34765	1.42	0.75-2.68	0.28
		C282Y+/-	37	173190	1.04	0.73-1.48	0.84
		C282Y +/+	9	8055	5.42	2.77-10.63	<0.001
Hepatocellular carcinoma	Females	No variants	30	1033703	1.00		
		H63D+/-	14	404722	1.19	0.63-2.25	0.59
		H63D+/+	2	39606	1.70	0.41-7.13	0.47
		C282Y+/H63+	0	40793	No observations		
		C282Y+/-	8	207390	1.33	0.61-2.91	0.48
		C282Y +/+	1	11051	3.10	0.42-22.84	0.27
Hepatocellular carcinoma	Males	No variants	80	865131	1.00		
		H63D+/-	42	337345	1.34	0.92-1.94	0.13
		H63D+/+	3	32849	0.99	0.31-3.13	0.99
		C282Y+/H63D+	7	34765	2.14	0.99-4.63	0.05
		C282Y+/-	16	173194	0.96	0.56-1.64	0.88
		C282Y +/+	5	8059	6.38	2.57-15.81	<0.001
Intrahepatic bile duct carcinoma	Females	No variants	69	1033666	1.00		
		H63D+/-	39	404703	1.46	0.98-2.16	0.06
		H63D+/+	2	39606	0.76	0.19-3.11	0.70
		C282Y+/H63+	3	40792	1.11	0.35-3.53	0.86
		C282Y+/-	17	207380	1.23	0.72-2.09	0.45
		C282Y +/+	0	11053	No observations		
Intrahepatic bile duct carcinoma	Males	No variants	63	865128	1.00		
		H63D+/-	28	337366	1.14	0.73-1.77	0.57
		H63D+/+	2	32845	0.83	0.20-3.39	0.80
		C282Y+/H63D+	3	34768	1.15	0.36-3.67	0.81
		C282Y+/-	17	173195	1.31	0.77-2.25	0.32
		C282Y +/+	3	8056	4.96	1.55-15.88	0.01

Sample excluding prevalent hemochromatosis at baseline (n=450,832), follow-up data from Hospital Episode Statistics, the Cancer Registry and Death Records.

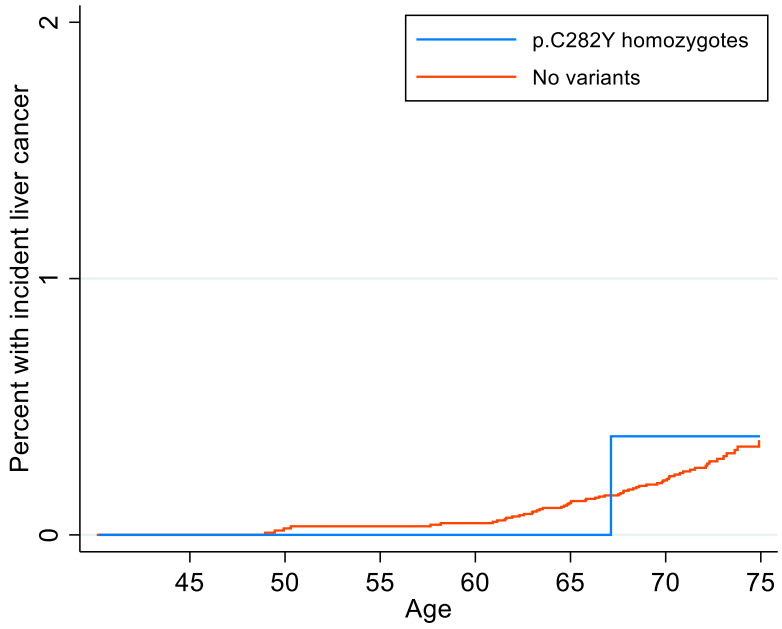
Hazard ratios adjusted for age, assessment center, genotyping array and population genetics sub-structure using principal components.

Excluding individuals with each prevalent diagnosis at baseline only: i.e. the incident liver cancer analysis included participants with baseline non-cancer liver disease.

Supplement Figures

Supplementary Figure 1. Kaplan-Meier failure curve in females for incidence of diagnosed liver cancer by age in HFE p.C282Y homozygotes compared to those with no variants, in subset with primary care follow-up

Incident diagnosis from Hospital Episode Statistics, the Cancer Registry Death Records and primary care data (n=209,811). Excluding participants with prevalent liver cancer diagnosis at baseline, but including those with non-cancer liver disease at baseline.

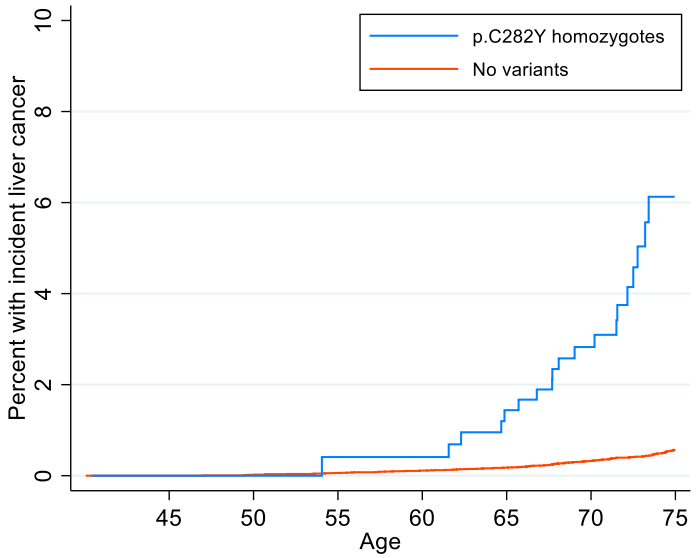


Number at risk							
p.C282Y homozygotes	78	151	176	180	252	232	55
No variants	6424	12095	14688	17460	22432	18162	4694

Supplementary Figure 2. Kaplan-Meier failure curves in a) males and b) females for incidence of diagnosed liver cancer by age in HFE p.C282Y homozygotes compared to those with no HFE variants

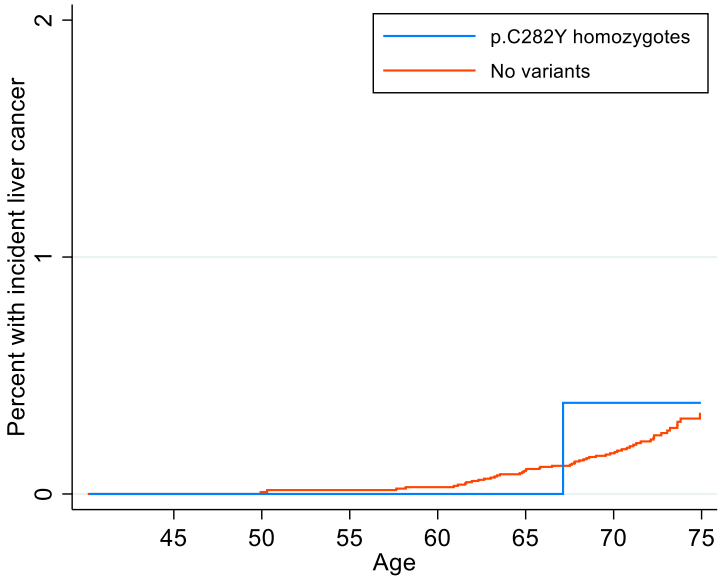
Incident diagnoses from Hospital Episode Statistics, Cancer Registry and Death Records (n=451,186, ignoring primary care data from subset). Excluding participants with prevalent liver cancer diagnosis at baseline, but including those with non-cancer liver disease at baseline.

(a)



Number at risk							
p.C282Y homozygotes	133	220	245	308	410	375	99
No variants	12019	20946	24226	29524	40111	36223	9461

(b)

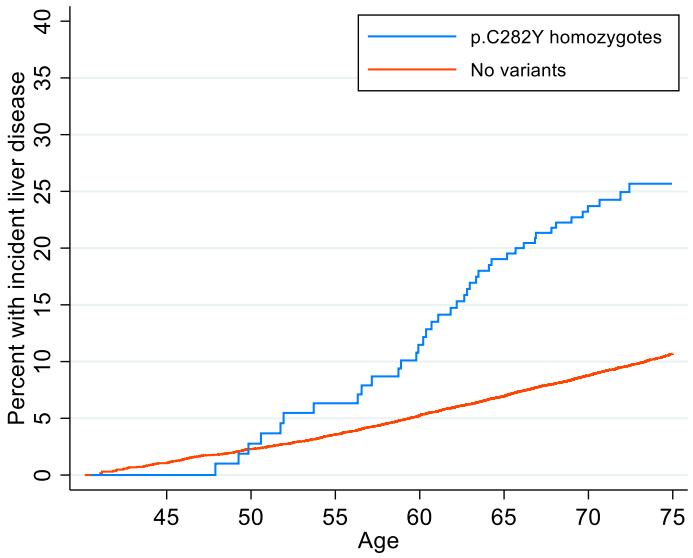


Number at risk							
p.C282Y homozygotes	78	151	176	180	252	232	55
No variants	6424	12095	14688	17460	22433	18163	4694

Supplementary Figure 3. Kaplan-Meier failure curves in a) males and b) females for incidence of diagnosed liver disease by age in HFE p.C282Y homozygotes compared to those with no variants, in subset with primary care follow-up

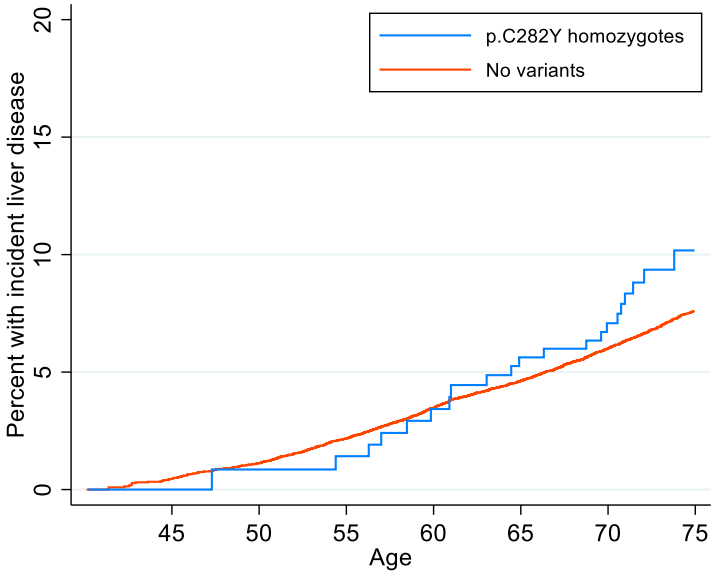
Incident diagnosis from Hospital Episode Statistics, the Cancer Registry and Death Records and primary care data (n=209,811). Excluding participants with prevalent non-cancer disease diagnoses at baseline.

(a)



Number at risk							
p.C282Y homozygotes	66	109	127	141	201	203	66
No variants	5485	10704	12263	14856	19834	18950	5881

(b)

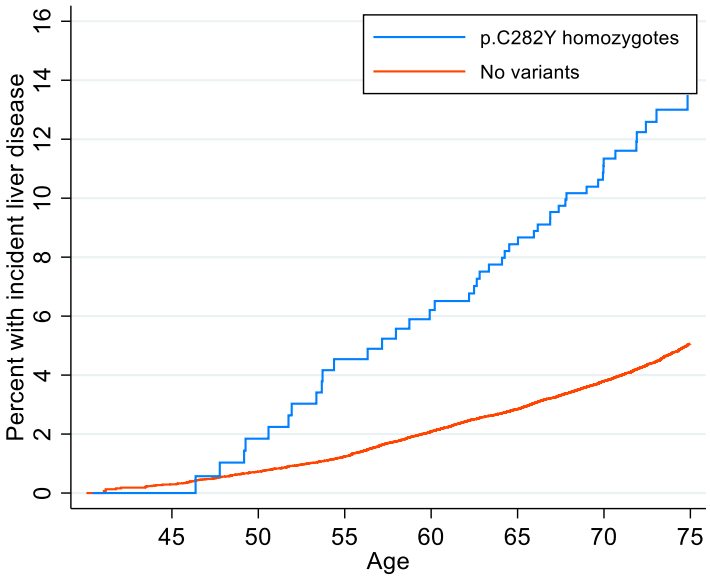


Number at risk							
p.C282Y homozygotes	78	160	188	192	269	265	85
No variants	6399	13332	16055	19190	24361	21285	6455

Supplementary Figure 4. Kaplan-Meier failure curves in a) males and b) females for incidence of diagnosed (non-cancer) liver disease by age in *HFE* p.C282Y homozygotes compared to those with no variants

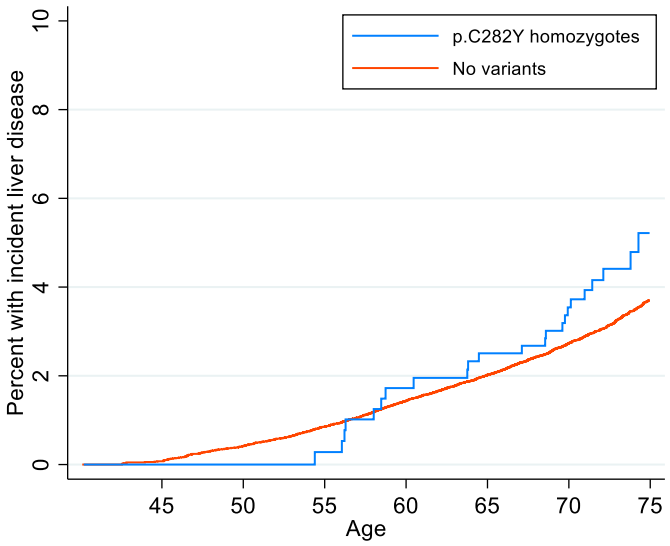
Incident diagnosis from Hospital Episode Statistics, the Cancer Registry and Death Records (n=451,186, ignoring primary care data from subset). Excluding participants with prevalent non-cancer liver disease diagnoses at baseline.

(a)



Number at risk							
p.C282Y homozygotes	131	249	274	317	432	416	131
No variants	11979	23345	26765	32413	43706	42171	13471

(b)



Number at risk							
p.C282Y homozygotes	153	306	372	437	580	554	177
No variants	13776	28622	34294	41282	52859	46710	14472