Online supplement

Ethical approval

The study was approved by the NHS Health Research Authority (IRAS project ID 203355, REC reference 17/EM/0021) and the sponsor was Imperial College London.

Control clinics

The composition of the control population was as follows:

Centre number (control source clinic)	Cases (N=494)	Controls (N=466)
1 (General Surgery)	42	39
2 (Gastroenterology/Stroke)^1^	13	11
3 (Cardiology)	38	36
4 (Urology)	52	52
5 (Diabetes/Rheumatology)^1^	40	31
6 (Sleep Apnoea)	34	37
7 (Neurology)	15	16
8 (ENT)	40	39
9 (Rheumatology)	31	29
10 (Oncology)	21	73^2^
11 (Urology)	11	11
12 (Haematology)	4	3
13 (Respiratory)	13	14
14 (Cardiology)	20	16
15 (Cardiology)	15	14
16 (Orthopaedics)	39	2^3^
17 (Asthma)	6	6
18 (Hypertension)	15	1^3^
19 (General Surgery)	7	9
20 (Urology)	31	25
21 (Respiratory)	7	2

^1^ The control clinic changed at these two sites because of slow recruitment (defined as fewer than four controls recruited over the course of four clinic attendances).

^2^ Controls were over-recruited at the local participating centre to help achieve the recruitment target.

^3^ Controls were under-recruited because of local research staffing shortage.

The approach of obtaining a list of all outpatient clinics it was possible for the local research team to recruit from and then randomly selecting a clinic from that list to serve as source clinic for the recruitment of controls was taken for two main reasons. Firstly, it is practically convenient to have a single clinic to recruit controls from. There is a single identifiable consultant who can consent to his or her clinic patients being approached and it is easier for the research team to form a relationship with the clinic staff. Secondly, the use of randomisation should, all else being equal, avoid the clinics (and therefore the patients within them) selected being systematically positively or negatively associated with occupational asbestos exposure.

Power calculation

Prior data indicated that the probability of occupational asbestos exposure in UK men aged 37-79 is 0.65.(1) If the true OR for disease in exposed men is 1.5, we calculated we would need to recruit 460 cases and controls, with power of 80% and significance threshold 0.05; our planned sample size included a margin for model stability and incomplete data. In a planned secondary analysis we investigated gene-environment interaction. The global minor allele frequency of MUC5B rs35705950 is 0.05; with an estimated prevalence of IPF of 20/100000 and with ORs of 1.5 for asbestos exposure and

of 6.8 for rs35705950, 460 cases would be required to detect a minimum interaction OR of 5.0.

Genotyping

DNA was extracted from whole blood samples using a Nucleon™ BACC3 Genomic DNA Extraction Kit (GE Healthcare). Genotypes of the *MUC5B* SNP rs35705950 were determined using TaqMan assays (Life Technologies, Carlsbad, CA) in 96-well plates, and fluorescence read using a Viia7 Sequence Detection System (Applied Biosystems).

Analysis code

We undertook statistical analyses using Python, SciPy, Statsmodels, and Stata (StataCorp. 2015. Stata Statistical Software: Release 14. College Station, TX: StataCorp LP).

Analysis code is available online:

https://github.com/drcjar/ipfjes/blob/master/notebooks/8.%20ipfjes_paper_analysis.ipynb

Coding socio-economic class

SOC90 coded jobs were also used to assign National Statistics Socio-Economic analytic classes (NS-SEC). The Office of National Statistics provides a lookup to assign each SOC90 code to one of eight classes:

1. Higher managerial, administrative and professional occupations
2. Lower managerial, administrative and professional occupations
3. Intermediate occupations

4. Small employers and own account workers
5. Lower supervisory and technical occupations
6. Semi-routine occupations
7. Routine occupations
8. Never worked and long-term unemployed

We then assigned each individual to a single code by calculating the median code for all of the jobs they had held.

Asbestos job exposure matrix

For all participants SOC90 coded jobs from the lifetime occupational history they provided were used to assign them to one of five main categories based on the highest risk job ever held(1):

1. High-risk non-construction
2. High-risk construction
3. Medium risk industrial
4. Low risk industrial
5. Office

Mapping of SOC90 code to occupational asbestos exposure categories was primarily based on occupational proportional mortality ratios (PMRs) for mesothelioma (see Figure E1). The categorisation was developed by staff who were blind to case-control status. In addition to using PMRs for mesothelioma to assist categorisation they used

knowledge of the details of the details of the job and location.(1) Ever exposed was defined as ever having a medium or high risk asbestos exposure category job.

Figure E1

Job category and occupation	*SOC 90 codes & criteria for classification*	*PMR*[1]
Non-construction high risk occupations		***128.4***
Metal plate worker	533, 534	331.8
Coach & vehicle body builders	541	527.5
Asbestos product manufacturer	Hands-on making asbestos products in regulated industry	139.4
Laggers and electrical, energy, boiler attendants	893 plus 896, 921, 929, 990 further categorised on job title	121.8
Docker, shipbuilding or working on board ship	880, 332, 903 plus 169, 173, 174, 239, 385, 463, 620, 621, 630, 900, 930, 952, 953 further categorised on job title & anyone who spent >50% of time on board ship/in a shipyard	135.2
Navy	"Royal Navy", or equivalent, as employer	125.6
Construction		***188.9***
Carpenters	570, 920	381.0
Plumbers	532, plus 913 further categorised on job title	361.8
Electricians	521, plus 913 further categorised on job title	259.0
Painters & decorators	507	170.8
Other construction workers	111, 500-506, 509, 885, 886, 895, 896 (if not classified as lagger, above), 921, 923, 924, 929 plus 990 & 913 further categorised on job title	123.0
Medium risk industrial		***130.1***
Metal working production & maintenance fitters	516 plus 913 further categorised on job title	209.6
Railway worker	881-884, 922	74.2
Chemist or industrial scientist	200, 300-302, 309	159.0
Surveyor or inspector	110, 260, 262, 311	109.7
Metal machining & instrument makers nec.	510-515 517-519	73.0
Electrical & electronic trades nec.	520, 522-529	135.6
Welding, steel erecting & fixing	535-537	173.9
Metal working process operatives	830-844	122.3
Assemblers & routine process operatives	850-869	75.1
Plant & machine operatives nec.	887-892, 894, 897-899	104.8
Low risk industrial		***69.4***
Motor mechanic	540	47.6
Draughtsmen	310	150.3
Engineers & technologists nec.	210-219	140.6
Stores & warehousemen	440-441	72.4
Armed forces nec.	150-151, 600-601	77.5
Drivers & road transport workers	731, 870-875	49.1
Other industrial not elsewhere classified	113, 153, 171, 304, 348, 396, 531, 542-544, 553, 561, 569, 571, 590, 596, 597, 599, 611, 612, 615, 619, 631, 642, 672, 699, 733, 801, 809, 811, 820, 822, 824, 825, 829, 910, 911-913, 919, 923, 924, 931, 933, 934, 940, 941, 955, 956, 958, 990, 999 nec & anyone spent >75% of time in heavy industry (e.g. power station), factory or warehouse	69.8
Office and other low risk		***58.7***
	101, 120-127, 130-132, 139, 154, 155, 160, 169, 170, 172-179, 190, 191, 220-224, 230-235, 239-242, 250-253, 260, 261, 270, 271, 290-293, 320, 340-347, 361-363, 370, 371, 380, 381, 383-387, 390, 392, 399-401, 410-412, 420, 421, 430, 450-452, 459-462, 490, 556, 559, 560, 569, 580-582, 592, 594, 595, 598, 610, 619-622, 630, 640-644, 650-652, 659-661, 670, 671, 673, 690, 691, 699, 700-703, 710, 719-722, 730, 732, 790-792, 900-904, 950-956, 958 & nec	

[1] *Average PMR using all SOC codes contributing to job category*

Figure E1: Classification of job categories with average national mesothelioma PMRs. Table 2.3.2 in Occupational, domestic and environmental mesothelioma risks in Britain. (1)

For analysis of categories of exposure participants were assigned to the highest risk category they ever had a job in.

Asbestos exposure assessment using a source receptor model

For participants who recalled carrying out work with asbestos a detailed assessment of each work task was recorded. A fibre.ml^{-1}.year asbestos exposure (AE) estimate was calculated using a source-receptor model. (2)(3)

First we calculated AE for each task as follows:

AE = E * H * LC

with parameters for the type of asbestos used (substance emission potential, E), what was done with it (activity emission potential, H), and whether there were any local exposure controls, for example wetting (local controls, LC).

AE for each task was then weighted according to the total amount of time spent performing the task and how well ventilated the room the activity was carried out in was (general ventilation parameters, D)(4), to arrive at a task fibre.ml^{-1}.year exposure estimate.

fibre.ml^{-1}.year (job task) = AE * task_duration * (task_frequency / periodicity) * job_duration * D

Task fibre.ml^{-1}.year exposure estimates were then summed at an individual participant level to provide an overall fibre.ml^{-1}.year estimate. A random sample of five high (top 25% of values), five medium (25-75% centile), and five low (bottom 25% of values) estimates (N=15) were independently assessed by a hygiene assessment expert who was blind to participant case status. The independent assessments tended to be lower than study assessments but there was overall acceptable agreement between assessments assessed using the Bland-Altman method. (see Figure E2)

Figure E2

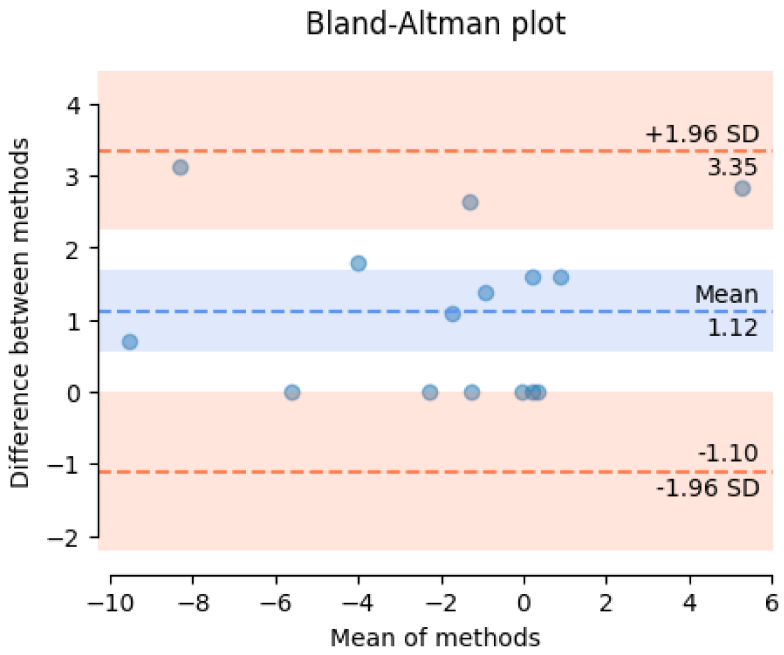


Figure E2: Independent validation of fibre.ml^{-1}.year exposure assessments. A sample of study assessments were repeated independently by an independent assessor. Difference

between methods is study assessment minus the independent assessor assessment.

Mean of methods is the average of the two assessments.

Blinding

Several measures were taken to ensure that, as far as possible, the telephone interviewer was blind to the case status of the research participant. Details of participants, including their case status, were entered into a bespoke web application. The application was used to carry out a telephone interview and by design hid the case status of participants from the interviewer. The interviewer elicited a lifetime occupational history and dynamically assigned a SOC90 code for each job title using the application. Questions regarding asbestos exposure were not asked until the end of the interview unless the participant volunteered working with asbestos. The application of the job exposure matrix and, if the participant recalled asbestos exposure, calculation of the estimated cumulative asbestos exposure was carried out algorithmically using data extracted from the interview application.

Sensitivity analyses

As a result of increasing awareness, and regulation, occupational asbestos exposure was significantly less widespread after 1980.(5) To investigate whether occupational asbestos exposure might be associated with IPF during this period we performed a sensitivity analysis by only including participants' jobs that ended before 1980. We did not observe a significant association. We also performed sensitivity analyses limiting to jobs that started before 1980, participants born prior to 1965, and considering only jobs before

age 45(6); there was no significant association between asbestos exposure and IPF for these.

Sensitivity analysis (limited to jobs that ended before 1980, N=771)

Exposure*	Adjusted OR† (95%CI;p-value)
ever asbestos exposed	0.9(0.7-1.3;0.66)
high-risk non-construction	1.1(0.6-1.9;0.79)
high-risk construction	1(0.6-1.5;0.84)
medium risk industrial	0.8(0.5-1.3;0.43)
low risk industrial	1(0.6-1.6;0.89)
office	1
ever smoked	1.3(0.9-1.9;0.13)
interaction model (asbestos*smoking)	
ever asbestos exposed	0.6(0.3-1.2;0.14)
ever smoked	1(0.6-1.7;0.98)
ever asbestos exposed and ever smoked interaction	1.7(0.8-3.5;0.15)

*Categories of occupational asbestos exposure risk were defined on the basis of occupational proportional mortality ratios for mesothelioma(1) and ever asbestos exposed was defined as ever having had a high or medium asbestos exposure risk job.

†Adjusted for age, centre, and smoking; smoking was not adjusted for when it was the exposure under consideration.

Sensitivity analysis (limited to jobs that participants spent 5 or more years in, N=957)

Exposure*	Adjusted OR† (95%CI;p-value)
ever asbestos exposed	0.9(0.7-1.2;0.59)
high-risk non-construction	0.7(0.4-1.2;0.2)

high-risk construction	0.9(0.6-1.4;0.77)
medium risk industrial	0.7(0.5-1.1;0.89)
low risk industrial	0.7(0.5-1.1;0.1)
office	1
ever smoked	1.4(1-1.9;0.03)
interaction model (asbestos*smoking)	
ever asbestos exposed	0.7(0.4-1.1;0.11)
ever smoked	1.1(0.7-1.7;0.58)
ever asbestos exposed and ever smoked interaction	1.6(0.9-3;0.12)

*Categories of occupational asbestos exposure risk were defined on the basis of occupational proportional mortality ratios for mesothelioma(1) and ever asbestos exposed was defined as ever having had a high or medium asbestos exposure risk job.

†Adjusted for age, centre, and smoking; smoking was not adjusted for when it was the exposure under consideration.

Sensitivity analysis (limited to participants within 10km of the recruiting hospital, N=426)

To investigate the importance of distance from recruiting hospital we estimated participants distance from the hospital by measuring the distance between their registered primary care provider (assumed to approximate residential address) and their recruiting hospital in kilometres(km) using postcode centroid data and Vincenty's formulae, we then analysed participants for whom this distance was less than 10km.

Exposure*	Adjusted OR† (95%CI;p-value)
ever asbestos exposed	1.3(0.8-2;0.37)

high-risk non-construction	1(0.4-2.3;0.96)
high-risk construction	1.1(0.5-2.2;0.86)
medium risk industrial	0.9(0.4-1.9;0.71)
low risk industrial	0.7(0.3-1.5;0.29)
office	1
ever smoked	1.3(0.8-2.1;0.35)
interaction model (asbestos*smoking)	
ever asbestos exposed	1 (0.4-2.4;0.95)
ever smoked	1(0.4-2.3;0.99)
ever asbestos exposed and ever smoked interaction	1.4(0.5-4.1;0.5)

*Categories of occupational asbestos exposure risk were defined on the basis of occupational proportional mortality ratios for mesothelioma(1) and ever asbestos exposed was defined as ever having had a high or medium asbestos exposure risk job.

†Adjusted for age, centre, and smoking; smoking was not adjusted for when it was the exposure under consideration. One centre was excluded from the analysis because it recruited no control participants within 10km.

Cumulative 'dose' based on occupational asbestos exposure (inferred by job title)

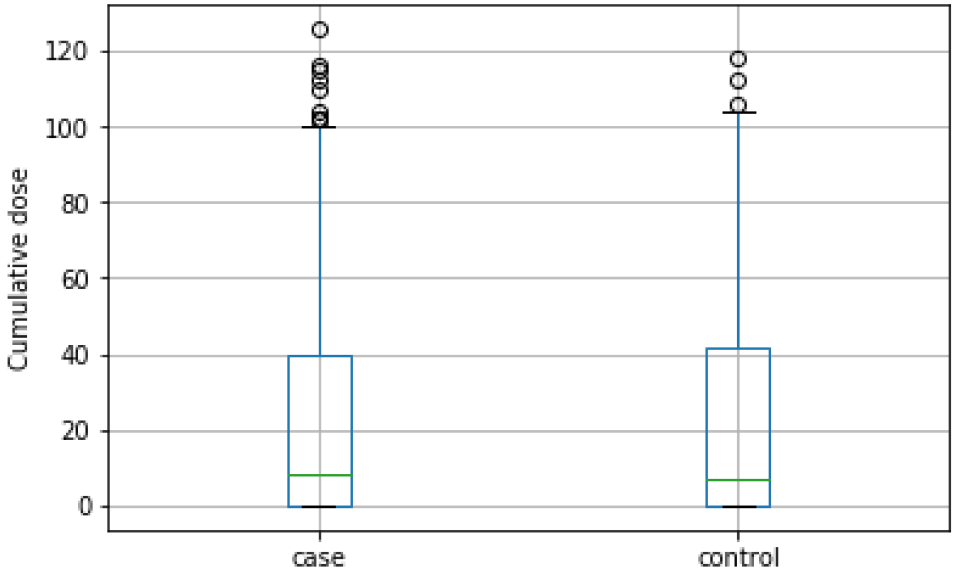
To investigate cumulative 'dose' of exposure based on job title a score was assigned based on asbestos exposure risk category of each job as follows:

- high-risk non-construction : 2
- high-risk construction : 2
- medium risk industrial : 1

- low risk industrial : 0
- office : 0

Scores were then multiplied for each job by the duration in years of the job and then summed at participant level.

	N	mean	std	min	25%	50%	75%	max
cases	494	24	30.6	0	0	8.5	39	126
controls	466	24	30.4	0	0	6.5	42	118



Boxplot of cumulative asbestos exposure estimates (inferred from job title) for cases and controls (N=960)

MUC5B rs35705950 minor allele frequency for genotyped cases, case subsets, and controls (N)*

	IPF (464)	IPF smoker (352)	IPF asbestos exposed (309)	IPF ≥25 fml.yrs (39)	IPF asbestos exposed AND smoker (253)	IPF ≥25 fml.yrs AND smoker (30)	Hospital controls (438)

GG	183	135	121	13	94	11	336
GT	248	189	164	21	137	15	97
TT	33	28	24	5	22	4	5
MAF	34	35	34	40	36	38	12

*Genotype of *MUC5B* rs35705950, T is the minor allele. MAF is minor allele frequency (%). fml-yrs is cumulative fibre.ml^{-1}.years of asbestos exposure. Smoker is defined as ever smoked. Ever asbestos exposed is defined as ever having had a high or medium asbestos exposure risk job, defined on the basis of occupational proportional mortality ratios for mesothelioma.(1)

Further information, including the full study protocol, all analysis code, and all other study documentation, is available from the study website https://ipfjes.org/

References

1. Peto J, Rake C, Gilham C, Hatch J. Occupational, domestic and environmental mesothelioma risks in britain: A case-control study. *Health and Safety Executive: Norwich, UK* 2009

2. Cherrie JW, Schneider T. Validation of a new method for structured subjective assessment of past concentrations. *Ann Occup Hyg* 1999;43:235–245.

3. Cherrie JW, McElvenny D, Blyth KG. Estimating past inhalation exposure to asbestos: A tool for risk attribution and disease screening. *International journal of hygiene and environmental health* 2018;221:27–32.

4. Cherrie, J., Maccalman, L., Fransman, W., Tielemans, E., Tischer, M., van Tongeren, M., 2011. Revisiting the effect of room size and general ventilation on the relationship

between near- and far-field air concentrations. Ann Occup Hyg 55, 1006–1015. doi:10.1093/annhyg/mer092

5. Gilham C, Rake C, Hodgson J, Darnton A, Burdett G, Peto Wild J, Newton M, Nicholson AG, Davidson L, Shires M, others. Past and current asbestos exposure and future mesothelioma risks in britain: The inhaled particles study (tips). *International Journal of Epidemiology* 2018;

6. Darnton A, Hodgson J, Benson P, Coggon D. Mortality from asbestosis and mesothelioma in britain by birth cohort. *Occup Med* 2012;62:549–552.