

**Supplementary Table 1 Other covariates (co-morbid conditions) measured at baseline**

	Exposed cohort (breathlessness symptom)	Unexposed matched cohort	
Number	75,698	75,698	matched
Respiratory symptom, <i>n</i> (%)			
Cough	18,099 (23.9)	6,149 (8.1)	<i>p</i> < 0.001
Sputum	1,206 (1.6)	305 (0.4)	<i>p</i> < 0.001
Inflammatory condition, <i>n</i> (%)			
Rheumatoid arthritis	479 (0.6)	261 (0.3)	<i>p</i> < 0.001
Gout	979 (1.3)	707 (0.9)	<i>p</i> < 0.001
Polymyalgia rheumatic	544 (0.7)	298 (0.4)	<i>p</i> < 0.001
Inflammatory bowel disease	294 (0.4)	183 (0.2)	<i>p</i> < 0.001
Systemic lupus erythematosus	24 (0.03)	11 (0.01)	<i>p</i> = 0.028
Spondyloarthritis*	153 (0.2)	79 (0.1)	<i>p</i> < 0.001
Mental condition, <i>n</i> (%)			
Depression	4,176 (5.5)	2,274 (3.0)	<i>p</i> < 0.001
Anxiety	4,513 (6.0)	2,479 (3.3)	<i>p</i> < 0.001
Diabetes, <i>n</i> (%)	4,780 (6.3)	3,668 (4.9)	<i>p</i> < 0.001
Non-specific chest pain, <i>n</i> (%)	6,335 (8.4)	2,343 (3.1)	<i>p</i> < 0.001
Musculoskeletal pain, <i>n</i> (%)	23,020 (30.4)	15,795 (20.9)	<i>p</i> < 0.001

\*Including ankylosing spondylitis and psoriatic arthritis. *p* value obtained from Chi-square.

**Supplementary Table 2 Incidence of future diagnosis of COPD, asthma or IHD\* in patients with a coded record of breathlessness symptom and in the matched cohort**

Time period	Exposed cohort (breathlessness symptom) <i>n</i> =75,698			Unexposed matched cohort <i>n</i> =75,698			Exposed vs unexposed
	Number with diagnosis (%†)	Time at risk (person-years)	Incidence‡	Number with diagnosis (%†)	Time at risk (person-years)	Incidence‡	Adjusted§ hazard ratio (95% CI)
Diagnostic window (0-6m)	11,150 (14.7)	32,558.2	3,424.6	474 (0.6)	37,100.0	127.8	-
Long term follow-up (>6m)	15,669 (20.7)	345,585.7	453.4	7,440 (9.8)	479,856.2	155.0	-
6–12m	2,411 (3.2)	29,367.1	821.0	532 (0.7)	35,971.7	147.9	5.1 (4.7, 5.7)
12–36m	5,212 (6.9)	100,528.8	518.5	1,959 (2.6)	130,733.8	149.8	3.2 (3.0, 3.3)
>36m	8,046 (10.6)	215,689.8	373.0	4,949 (6.5)	313,150.7	158.0	2.2 (2.1, 2.3)

\*First diagnosed condition if multiple conditions occurred. †Percentage in the total number of patients with a first coded record of breathlessness symptom; or in the total number of the matched cohort. ‡Per 10,000 person-years. §Adjusted hazard ratio, adjusted for index year, age, gender, region, ever smoking, ever alcohol drinking, BMI, deprivation level, comorbidity, cough and sputum, inflammatory conditions, depression and anxiety, diabetes, non-specific chest pain, musculoskeletal pain.

**Supplementary Table 3 Incidence of future diagnosis of COPD, asthma or IHD\* in patients with a coded record of breathlessness symptom and in the matched cohort, stratified by age**

Time period	Exposed cohort (breathlessness symptom) <i>n</i> =75,698			Unexposed matched cohort <i>n</i> =75,698			Exposed vs unexposed
	Number with diagnosis (%†)	Time at risk (person-years)	Incidence‡	Number with diagnosis (%†)	Time at risk (person-years)	Incidence‡	Adjusted§ hazard ratio (95% CI)
Aged 18-59 years		<i>n</i> =37,024			<i>n</i> =37,024		

Diagnostic window (0–6m)	4,145 (11.2)	16,597.4	2,497.4	80 (0.2)	18,177.0	44.0	
Long term follow-up (>6m)	5,859 (15.8)	200,956.5	291.6	1,839 (5.0)	254,900.2	72.1	
6–12m	858 (2.3)	15,437.0	555.8	102 (0.3)	17,697.6	57.6	8.4 (6.8, 10.4)
12–36m	1,836 (5.0)	54,606.7	336.2	412 (1.1)	65,251.0	63.1	4.5 (4.0, 5.1)
>36m	3,165 (8.5)	130,912.8	241.8	1,325 (3.6)	171,951.6	77.1	2.7 (2.5, 2.8)
Aged 60 years and above		<i>n</i> =38,674			<i>n</i> =38,674		
Diagnostic window (0–6m)	7,005 (18.1)	15,960.8	4,388.9	394 (1.0)	18,923.1	208.2	
Long term follow-up (>6m)	9,810 (25.4)	144,629.2	678.3	5,601 (14.5)	225,048.7	248.9	
6–12m	1,553 (4.0)	13,930.1	1,114.9	430 (1.1)	18,274.1	235.3	4.3 (3.9, 4.8)
12–36m	3,376 (8.7)	45,922.1	735.2	1,547 (4.0)	65,482.8	236.2	2.8 (2.6, 3.0)
>36m	4,881 (12.6)	84,777.0	575.7	3,624 (9.4)	141,291.8	256.5	2.0 (1.9, 2.1)

\*First diagnosed condition if multiple conditions occurred. †Percentage in patients aged 18–59 years/60 years and above with a first coded record of breathlessness symptom; or in matched cohort aged 18–59 years/60 years and above. ‡Per 10,000 person-years. §Adjusted hazard ratio, adjusted for index year, age, gender, region, ever alcohol drinking, BMI, deprivation level, comorbidity, cough and sputum, inflammatory conditions, depression and anxiety, diabetes, non-specific chest pain, musculoskeletal pain.

**Supplementary Table 4 Incidence of future diagnosis of COPD, asthma or IHD\* in patients with a coded record of breathlessness symptom and in the matched cohort, stratified by gender**

Exposed cohort (breathlessness symptom) <i>n</i> =75,698	Unexposed matched cohort <i>n</i> =75,698	Exposed vs unexposed
---	--	----------------------

Time period	Number with diagnosis (%†)	Time at risk (person-years)	Incidence‡	Number with diagnosis (%†)	Time at risk (person-years)	Incidence‡	Adjusted§ hazard ratio (95% CI)
<b>Male</b>		<i>n</i> =32,195		<i>n</i> =32,195			
Diagnostic window (0–6m)	5,451 (16.9)	13,534.0	4,027.6	226 (0.7)	15,767.6	143.3	
Long term follow-up (>6m)	6,951 (21.6)	137,236.0	506.5	3,638 (11.3)	199,002.8	182.8	
6–12m	1,122 (3.5)	12,045.8	931.4	251 (0.8)	15,261.6	164.5	5.2 (4.5, 6.0)
12–36m	2,358 (7.3)	40,637.1	580.3	949 (2.9)	55,298.4	171.6	3.2 (3.0, 3.5)
>36m	3,471 (10.8)	84,553.0	410.5	2,438 (7.6)	128,442.8	189.8	2.1 (2.0, 2.2)
<b>Female</b>		<i>n</i> =43,503		<i>n</i> =43,503			
Diagnostic window (0–6m)	5,699 (13.1)	19,024.2	2,995.7	248 (0.6)	21,332.4	116.3	
Long term follow-up (>6m)	8,718 (20.0)	208,349.7	418.4	3,802 (8.7)	280,946.1	135.3	
6–12m	1,289 (3.0)	17,321.3	744.2	281 (0.6)	20,710.0	135.7	5.1 (4.5, 5.8)
12–36m	2,854 (6.6)	59,891.6	476.5	1,010 (2.3)	75,435.4	133.9	3.1 (2.9, 3.4)
>36m	4,575 (10.5)	131,136.8	348.9	251 (0.6)	184,800.6	135.9	2.3 (2.2, 2.4)

\*First diagnosed condition if multiple conditions occurred. †Percentage in male/female patients with a first coded record of breathlessness symptom; or in matched male/female cohort. ‡Per 10,000 person-years. §Adjusted hazard ratio, adjusted for index year, age, region, ever alcohol drinking, BMI, deprivation level, comorbidity, cough and sputum, inflammatory conditions, depression and anxiety, diabetes, non-specific chest pain, musculoskeletal pain.

**Supplementary Table 5** Incidence of future diagnosis of COPD, asthma or IHD\* in patients with a coded record of breathlessness symptom and in the matched cohort, stratified by smoking status

Time period	Exposed cohort (breathlessness symptom) <i>n</i> =75,698			Unexposed matched cohort <i>n</i> =75,698			Exposed vs unexposed
	Number with diagnosis (%†)	Time at risk (person-years)	Incidence‡	Number with diagnosis (%†)	Time at risk (person-years)	Incidence‡	Adjusted§ hazard ratio (95% CI)
Ever-smoking subset		<i>n</i> =24,740			<i>n</i> =19,143		
Diagnostic window (0–6m)	4,493 (18.2)	10,284.6	4368.7	154 (0.8)	9,376.7	164.2	-
Long term follow-up (>6m)	5,821 (23.5)	99,510.7	585.0	2,375 (12.4)	116,093.4	204.6	-
6–12m	955 (3.9)	9,090.2	1050.6	167 (0.9)	9,069.0	184.1	5.58 (4.71, 6.61)
12–36m	2,003 (8.1)	30,364.2	659.7	618 (3.2)	32,711.2	188.9	3.41 (3.10, 3.74)
>36m	2,863 (11.6)	60,056.3	476.7	1,590 (8.3)	74,313.2	214.0	2.18 (2.0, 2.33)
Never-smoking subset		<i>n</i> =40,436			<i>n</i> =42,632		
Diagnostic window (0–6m)	4,995 (12.4)	17,820.8	2802.9	227 (0.5)	20,918.6	108.5	-
Long term follow-up (>6m)	7,481 (18.5)	197,099.7	379.6	3,695 (8.7)	272,711.4	135.5	-
6–12m	1,118 (2.8)	16,285.1	686.5	270 (0.6)	20,336.7	132.8	4.77 (4.16, 5.47)
12–36m	2,429 (6.0)	56,565.6	429.4	1,012 (2.4)	74,332.1	136.1	2.84 (2.63, 3.07)
>36m	3,934 (9.7)	124,249.0	316.6	2,413 (5.7)	178,042.6	135.5	2.18 (2.07, 2.30)

\*First diagnosed condition if multiple conditions occurred. †Percentage in ever-smoking/never-smoking of patients with a first coded record of breathlessness symptom; or in ever-smoking/never-smoking of matched cohort. ‡Per 10,000 person-years. §Adjusted hazard ratio, adjusted for index year, age, gender, region, ever alcohol drinking, BMI, deprivation level, comorbidity, cough and sputum, inflammatory conditions, depression and anxiety, diabetes, non-specific chest pain, musculoskeletal pain. Total number of patients different to 75,698 in each cohort was due to missing data in smoking status.

**Supplementary Table 6 Risk of mortality in patients with a coded record of breathlessness symptom compared to the matched cohort stratified by age group**

	Time period	Exposed cohort (breathlessness symptom) <i>n</i> =75,698			Unexposed matched cohort <i>n</i> =75,698			Exposed vs unexposed Adjusted† hazard ratio‡ (95% CI)
		Number of events (%)	Time at risk (person-years)	Incidence*	Number of events (%)	Time at risk (person-years)	Incidence*	
Aged 18-59 years		<i>n</i> =37,024			<i>n</i> =37,024			
All-cause	0–6m	512 (1.4)	17,896.9	286.1	30 (0.08)	18,001.4	16.7	11.7 (8.1, 17.0)
	>6m	1,467 (4.0)	251,019.6	58.4	685 (1.9)	253,278.6	27.0	1.4 (1.3, 1.6)
COPD-related	>6m	117 (0.3)	251,019.6	4.7	11 (0.03)	253,278.6	0.4	2.8 (1.4, 5.6)
IHD-related	>6m	239 (0.6)	251,019.6	9.5	105 (0.3)	253,278.6	4.1	1.6 (1.3, 2.1)
Aged 60 years and above		<i>n</i> =38,674			<i>n</i> =38,674			
All-cause	0–6m	2,607 (6.7)	18,065.4	1,443.1	393 (1.0)	18,800.5	209.0	5.6 (5.0, 6.2)
	>6m	9,784 (25.3)	211,040.9	463.6	7,041 (18.2)	236,338.3	297.9	1.2 (1.1, 1.3)
COPD-related	>6m	913 (2.4)	211,040.9	43.3	187 (0.5)	236,338.3	7.9	1.6 (1.3, 1.9)
IHD-related	>6m	2,760 (7.1)	211,040.9	130.8	1,520 (3.9)	236,338.3	64.3	1.2 (1.1, 1.3)

\*Per 10,000 person-years. †Adjusted for index year, age, gender, region, ever smoking, ever alcohol drinking, BMI, deprivation level, comorbidity, cough and sputum, inflammatory conditions, depression and anxiety, diabetes, non-specific chest pain, musculoskeletal pain and later diagnosis; Later diagnosis as a time-varying covariate; For all-cause mortality, later diagnosis is the earliest diagnosis of asthma, COPD or IHD. ‡Sub-distribution hazard ratio was used in completing risk models for disease-specific mortality (other death was treated as competing events); and hazard ratio in Cox models for all-cause mortality. Risk of COPD or IHD-related mortality during 0-6 months was not reported due to the CPRD rule around reporting low cell counts (i.e. data less than 5 should not be reported).