13 January 2022

Data analysis report

Title: Background incidence rates of Interstitial Lung Disease (ILD)

Administrative details of the data analysis	
Substance(s)	Enfortumab Vedotin / Bosutinib
Condition/ADR(s)	Interstitial Lung Disease (ILD)
Short title of topic	Background rates of ILD
TDA-DAT lead analyst (and reviewer)	Valentijn De Jong valentijn.de.jong@ema.europa.eu (Maria Clara Restrepo Mendez and Daniel Morales Leaver)

Official address Domenico Scarlattilaan 6 ● 1083 HS Amsterdam ● The Netherlands

Address for visits and deliveries Refer to www.ema.europa.eu/how-to-find-us
Send us a question Go to www.ema.europa.eu/contact **Telephone** +31 (0)88 781 6000

Table of Contents

1. List of abbreviations

MAH	*Marketing Authorisation Holder*
EMA	*European Medicines Agency*
PRAC	*Pharmacovigilance Risk Assessment Committee*
RDA	*Rapid Data Analysis*

2. Rationale and background

Interstitial lung disease (ILD) describes a heterogenous group of respiratory disorders affecting the interstitium of the lungs. [1,2] ILD may occur when an injury to the lungs triggers an abnormal healing response. The repair process is disrupted, and the tissue around the alveoli becomes scarred and thickened. Prolonged ILD may result in pulmonary fibrosis, but this is not always the case.

ILD is not a single disease, it encompasses many different pathological processes including drug-induced ILDs. Drug-induced interstitial lung disease (DI-ILD) is also a large and very heterogeneous group of adverse drug reactions, ranging from mild to progressive and life-threatening disease. The number of drugs associated with the development of ILD continues to rise, mainly due to the use of novel monoclonal antibodies and biologics for neoplastic and rheumatologic diseases, many of which are associated with lung toxicity, and includes, among others, chemotherapeutics, molecular targeting agents, immune checkpoint inhibitors, antibiotics, antiarrhythmics, and conventional or biologic disease-modifying antirheumatic drugs. [3]

ILD is usually diagnosed through chest radiography as a first step, but as the chest radiograph can be normal in up to 10% of patients, high resolution computed tomography of the chest is the preferred modality. A lung biopsy can be required if the clinical history and imaging are not clearly suggestive of a specific diagnosis or malignancy cannot otherwise be ruled out.

Treatment of ILD vary depending on the underlying disease. Early identification and discontinuation of the drug are the priority measures if a drug cause is suspected. If a specific occupational exposure cause is identified, the person should avoid that environment. Many cases due to unknown or connective tissue-based causes are treated with corticosteroids.

During routine signal detection, cases of ILD with potential association to enfortumab vedotin as well as bosutinib were reported from EudraVigilance and literature. It was proposed to generate background incidence rates of ILD from a number of European databases in order to support the PRAC assessment.

3. Research question and objectives

The objective of the study was to describe incidence rates of ILD in the general population and stratified by gender, age, and year.

4. Research methods

4.1. Study design

This was a cohort study describing incidence rates of ILD in the general population.

4.2. Setting and study population

The study population was the general population (UK), patients visiting general practices (France), and patients visiting general or specialized practices (Germany).

4.3. Data sources

The following databases were used: IQVIA™ Medical Research Data (IMRD) UK database, and IQVIA™ Disease Analyzer Germany and France databases. Brief descriptions of these databases were provided in **Annex 1**.

4.4. Study period

The study period varied according to the years of coverage in the different databases. For the IMRD UK database, it was from 2004 to 2021. For the IQVIA™ Disease Analyzer Germany and France databases, the coverage period was from 2016 to 2020.

4.5. Variables

Outcomes: ILDs have been difficult to classify because approximately 180 known individual diseases are characterized by interstitial lung involvement (either primary disease or part of a multiorgan process, e.g., collagen vascular diseases). Therefore, to identify cases that are highly likely to represent the condition of interest, we will follow the **"narrow" scope list of terms** (i.e., very specific preferred terms) documented in the ***Introductory Guide for Standardised MedDRA Queries (SMQs) Version 25.0*** [4].

Detailed list of terms (and their SMQ codes) to be included were shown in **Annex 2**.

We mapped these SMQ codes to SNOMED codes and used their OMOP concept IDs.

4.6. Statistical analysis

4.6.1. Main statistical methods

Incidence rates in the general population: We described the incidence of new onset of ILD diagnoses in patients contributing patient time to the databases listed above. Patients were required to have a minimum observation time of 365 days prior to entering each period in order to establish whether events observed during the period were incident (first-ever) cases. Patients were excluded from the analysis if they had any prior history of the condition in the database.

i. ***Numerator:*** The numerator consists of the number of patients who experience the event of interest during the time period (overall and yearly). Patients with any recorded baseline history of ILD were excluded. Included patients were allowed to contribute only one event each.

ii. ***Denominator:*** The denominator was defined as patient follow-up time. Patients with a baseline history of ILD at the start of each year were excluded. Patient follow-up time were truncated at the earliest date of the following: a) the occurrence of the first event after which they did not contribute to the analysis, b) death of patient, c) end of the observation period (i.e., end of data availability because the patient moved practice, or reached the end of follow-up for their practice), or d) end of study period.

iii. ***Analysis:*** Follow-up time was calculated using the following formula:

Follow-up time (years) = (end date for the period – start date for the period + 1) / 365.25

The **incidence rate** was then calculated as the number of events divided by the total follow up time:

Incidence rate = (number of new onset events) / (total follow up time (years))

The incidence is presented as the number of events per 100,000 person-years and was calculated for the entire population as well as stratified by:

- sex,
- age groups (0-19, 20-29, 30-39, 40-49, 50-59, 60-69, 70-79, 80+), and
- year of recorded diagnosis

N.B.: Results stratified by sex and age group were presented for the last five selected years of each database. For IQVIA™ Disease Analyzer Germany and France, patient observability ends at the last patient visit, therefore, the size of the denominator shrinks towards database end (e.g., last year) whereas the numerator is unaffected, which results in artificially increased incidence rates towards database end. Therefore, we excluded the year of 2021 from the analysis.

Confidence intervals around incidence rates were calculated using the exact method.

Analyses were done using the Instant Health Data (IHD) platform.

4.6.2. Exploratory analysis

Given the broad range of conditions characterised by interstitial lung involvement, we conducted additional analyses in which we explored subgroups of the included conditions, namely:

- Conditions specifically recorded as ILD (ILD-specific codes), and
- ILDs expected to be a drug-related event
- Events recorded as specific conditions: Fibrosis, Pneumonitis or Alveolitis. Note that some events fall under multiple categories.

4.6.3. Sensitivity analysis

None.

4.7. Quality control

The study was conducted according to the ENCePP code of conduct (European Medicines Agency 2018).

Standard operating procedures or internal process guidance were adhered to for the conduct of the study. These procedures include rules for secure and confidential data storage, quality-control procedures for all aspects of the study from protocol development to the reporting of the results.

All documents underwent at least one round a review by an experienced reviewer, while the results from the statistical analysis were either reviewed.

The quality control of the data is the responsibility of the data holder. See Annex 1 for more details.

4.8. Protection of human subjects

Patient confidentiality was protected according to the EU General Data Protection Regulation (GDPR) on the protection of individuals.

5. Results

In accordance with database rules on the management of low cell counts, cells with low numbers (<6 in the IMRD database and <10 in IQVIA™ Disease Analyzer France, THIN® Spain, Italy and Romania) will be removed prior to publication of this report. Additional cells may be redacted (events/patients typically being rounded up to the nearest 10) if needed in order to ensure that the aforementioned low cell counts cannot be re-identified. This may include both events/patients and follow-up times.

5.1. Main analysis

5.1.1. Event rates in the general population

Table 1 shows the overall population incidence rates for ILD and stratified by sex, age group and year in the IMRD UK, IQVIA™ Disease Analyzer Germany AND France databases. In UK and Germany databases, the incidence rate of ILD was higher in male than in female subjects, and in all three databases the incidence increased with age.

In IMRD and IQVIA™ Disease Analyzer Germany, the incidence slightly increased over time, except in 2020 where a lower incidence was observed in the UK. In France, incidence rates remained stable over time with a substantial decrease in 2020. The decline in the recorded incidence of ILD in 2020 could be attributed to a change in health seeking behaviours during the COVID-19 pandemic.

Table S1 shows the incidence rate for ILD in the general population in IMRD UK from 2004 to 2016. It shows that, on overall, there is an increase in the incidence rate of ILD over this period, from 3.02 per 100,000 in 2014 to 17.61 per 100,000 in 2016 (**Annex 3**).

Table S2 shows the overall population incidence rates for ILD and stratified by age group for each sex in the IMRD UK, IQVIA™ Disease Analyzer Germany AND France database (**Annex 3**). In the UK and Germany databases, the incidence of ILD increased with age for both sexes. In the France database, the incidence of ILD increases with age for the lower age groups, it peaked in the age group of 50-59, and then decreased with age, though this decrease was nonsignificant for male subjects.

5.2. Exploratory analysis

5.2.1. Including only ILD-specific codes

Table 2 shows the overall population incidence rates for ILD and stratified by sex, age group and year in the IMRD UK, IQVIA™ Disease Analyzer Germany AND France databases using a restricted definition of ILD in which we only include conditions specifically recorded as ILD (**Annex 2**).

The number of records of ILD-specific cases increased each year until and including 2019, and then decreased in 2020 in the UK database. In the Germany database it increased each year until and including 2019, and then a similar number of cases was recorded in 2020. In the France database we observed that in 2018 fewer specific ILD cases were recorded, and the incidence rate subsequently increased.

5.2.2. Including conditions expected to be a drug-related event

Table 3 shows the overall population incidence rates for ILD and stratified by sex and year in the IMRD UK, IQVIA™ Disease Analyzer Germany AND France databases using for the definition of ILD only conditions expected to be a drug-related event (**Annex 2**). The incidence rate was < 1 per

100,000 person-years in the UK and Germany databases, from 2016 to 2020. No cases were identified in IQVIA™ Disease Analyzer France.

5.2.3. Including selected conditions

The estimated incidence rates for the selected conditions are available in **Annex 2**. The overall incidence rate for Fibrosis was 4.92 per 100,000 person-years in the UK database (**Table S3**). No cases were identified in the France and Germany databases.

The overall incidence rate for Pneumonitis was 5.92 per 100,000 person-years in the UK database, 9.78 per 100,000 person-years in the Germany database, and 19.19 per 100,000 person-years in the France database (**Table S4**).

The overall incidence rate for Alveolitis was 1.95 per 100,000 person-years in the UK database, 2.49 per 100,000 person-years in the Germany database and 2.61 per 100,000 person-years in the France database (**Table S5**).

Table 1. *Overall and stratified background rates of ILD*** per 100,000 years of follow-up between 2016 and 2021 in IMRD UK, and between 2016 and 2020 in IQVIA™ Disease Analyzer Germany and France.*

	IMRD UK				IQVIA™ Disease Analyzer Germany				IQVIA™ Disease Analyzer France			
	Background rates				Background rates				Background rates			
	Years of follow-up	No. of outcomes	IR per 100,000 person-years	95% CI	Years of follow-up	No. of outcomes	IR per 100,000 person-years	95% CI	Years of follow-up	No. of outcomes	IR per 100,000 person-years	95% CI
Overall	8.897.852,21	1.978	**22,23**	21.27 - 23.23	35.024.903,86	5.921	**16,91**	16.48 - 17.34	5.740.066,26	1.509	**26,29**	25.00 - 27.65
Gender												
Female	4.443.372,09	881	**19,83**	18.56 - 21.18	20.088.473,61	2.773	**13,80**	13.30 - 14.33	3.053.681,40	830	**27,18**	25.39 - 29.09
Male	4.454.480,12	1.097	**24,63**	23.21 - 26.13	14.936.430,25	3.148	**21,08**	20.35 - 21.83	2.686.384,85	679	**25,28**	23.45 - 27.25
Age (in years)												
0-19	1.947.344,50	38	**1,95**	1.42 - 2.68	4.212.986,93	274	**6,50**	5.78 - 7.32	1.234.514,00	71	**5,75**	4.56 - 7.25
20-29	1.091.616,00	21	**1,92**	1.26 - 2.94	2.946.244,26	150	**5,09**	4.34 - 5.97	600.248,67	90	**14,99**	12.21 - 18.43
30-39	1.382.443,82	69	**4,99**	3.95 - 6.32	3.665.214,18	213	**5,81**	5.08 - 6.65	693.646,57	172	**24,80**	21.36 - 28.79
40-49	1.252.562,53	118	**9,42**	7.87 - 11.28	4.396.740,30	378	**8,60**	7.77 - 9.51	788.811,10	280	**35,50**	31.58 - 39.91
50-59	1.224.205,51	253	**20,67**	18.27 - 23.38	6.415.831,25	774	**12,06**	11.24 - 12.94	810.684,01	364	**44,90**	40.52 - 49.76
60-69	882.793,86	384	**43,50**	39.36 - 48.07	5.600.399,12	991	**17,70**	16.63 - 18.83	737.758,96	270	**36,60**	32.49 - 41.23
70-79	674.902,36	536	**79,42**	72.98 - 86.44	4.791.611,69	1.313	**27,40**	25.96 - 28.93	520.039,53	163	**31,34**	26.89 - 36.54
80+	441.983,61	559	**126,48**	116.42 - 137.41	2.995.876,14	1.828	**61,02**	58.28 - 63.88	354.363,41	99	**27,94**	22.96 - 34.01
Year												
2016	1.618.760,38	285	**17,61**	15.68 - 19.77	7.014.330,67	987	**14,07**	13.22 - 14.98	1.284.000,00	421	**32,79**	29.80 - 36.07
2017	1.693.283,80	330	**19,49**	17.50 - 21.71	7.164.072,36	1.094	**15,27**	14.39 - 16.20	1.262.319,11	349	**27,65**	24.90 - 30.71
2018	1.741.398,96	356	**20,44**	18.43 - 22.68	7.340.820,38	1.220	**16,62**	15.71 - 17.58	1.196.620,82	328	**27,41**	24.60 - 30.54
2019	1.774.235,08	417	**23,50**	21.35 - 25.87	7.041.454,85	1.337	**18,99**	18.00 - 20.03	1.085.554,12	279	**25,70**	22.86 - 28.90
**2020	1.831.413,55	366	**19,98**	18.04 - 22.14	6.464.225,60	1.283	**19,85**	18.79 - 20.96	911.572,21	132	**14,48**	12.22 - 17.17
2021	1.857.520,82	509	**27,40**	25.12 - 29.89	--	--	--	--	--	--	--	--

*Overall results and results per gender and per age group for IMRD UK in this table include only data from years 2017 to 2021. Data from 2004 to 2016 are presented in **Annex 3**.

** Incidence rate might be falsely reduced due to changes in health seeking behaviour during COVID-19 pandemic.

***** ILD definition based on SMQ terms mapped to SNOMED codes and OMOP concepts**

IR = incidence rates, 95% CI = 95% confidence intervals

Table 2. *Overall and stratified background rates of ILD per 100,000 years of follow-up between 2016 and 2020 in IMRD UK, IQVIA™ Disease Analyzer Germany and France.*

Including only ILD-specific codes***												
	IMRD UK				IQVIA™ Disease Analyzer Germany				IQVIA™ Disease Analyzer France			
	Background rates				Background rates				Background rates			
	Years of follow-up	No. of outcomes	IR per 100,000 person-years	95% CI	Years of follow-up	No. of outcomes	IR per 100,000 person-years	95% CI	Years of follow-up	No. of outcomes	IR per 100,000 person-years	95% CI
Overall	8.680.453,17	903	**10,40**	9.75 - 11.10	35.299.670,87	1.703	**4,82**	4.60 - 5.06	5.813.956,83	246	**4,23**	3.74 - 4.79
Sex												
Female	4.334.417,63	384	**8,86**	8.02 - 9.79	20.247.801,23	837	**4,13**	3.86 - 4.42	3.092.058,73	118	**3,82**	3.19 - 4.57
Male	4.346.035,54	519	**11,94**	10.96 - 13.01	15.051.869,64	866	**5,75**	5.38 - 6.15	2.721.898,11	128	**4,70**	3.96 - 5.59
Age group (in years)												
0-19	-	<6	**0.11**	0.03 - 0.38	4.469.757,34	75	**1.68**	1.34 - 2.10	-	<10	**0.46**	0.22 - 1.00
20-29	-	<6	**0,28**	0.10 - 0.82	2.946.994,13	25	**0,85**	0.58 - 1.25	-	<10	**0,67**	0.27 - 1.71
30-39	1.331.216,22	22	**1,65**	1.10 - 2.50	3.666.351,43	65	**1,77**	1.39 - 2.26	694.523,10	17	**2,45**	1.54 - 3.92
40-49	1.234.971,69	46	**3,72**	2.80 - 4.97	4.398.540,04	126	**2,86**	2.41 - 3.41	790.057,12	22	**2,78**	1.85 - 4.22
50-59	1.193.743,32	92	**7,71**	6.29 - 9.45	6.419.094,68	276	**4,30**	3.82 - 4.84	812.452,74	39	**4,80**	3.52 - 6.56
60-69	865.038,85	185	**21,39**	18.52 - 24.70	5.603.994,34	389	**6,94**	6.29 - 7.67	739.200,62	46	**6,22**	4.67 - 8.30
70-79	655.619,73	306	**46,67**	41.73 - 52.21	4.795.190,64	468	**9,76**	8.92 - 10.69	520.639,43	62	**11,91**	9.30 - 15.27
80+	435.887,92	247	**56,67**	50.03 - 64.19	2.999.748,27	279	**9,30**	8.27 - 10.46	354.588,86	50	**14,10**	10.71 - 18.59
Year												
2016	1.635.487,77	133	**8,13**	6.86 - 9.64	7.070.118,71	275	**3,89**	3.46 - 4.38	1.300.138,75	57	**4,38**	3.39 - 5.68
2017	1.694.373,11	158	**9,32**	7.98 - 10.90	7.220.192,70	296	**4,10**	3.66 - 4.59	1.278.541,14	55	**4,30**	3.31 - 5.60
2018	1.742.540,10	178	**10,21**	8.82 - 11.83	7.398.661,54	348	**4,70**	4.24 - 5.22	1.212.026,48	42	**3,47**	2.57 - 4.68
2019	1.775.418,54	253	**14,25**	12.60 - 16.12	7.097.252,56	411	**5,79**	5.26 - 6.38	1.099.747,05	47	**4,27**	3.22 - 5.68
2020	1.832.633,65	181	**9,88**	8.54 - 11.42	6.513.445,36	373	**5,73**	5.17 - 6.34	923.503,42	45	**4,87**	3.65 - 6.52

***** ILD definition based on conditions specifically recorded as ILD.**

IR = incidence rates, 95% CI = 95% confidence intervals

Table 3. *Overall and stratified background rates of drug-induced ILD per 100,000 years of follow-up between 2016 and 2020 in IMRD UK, IQVIA™ Disease Analyzer Germany and France.*

	Drug-induced ILD***											
	IMRD UK				IQVIA™ Disease Analyzer Germany				IQVIA™ Disease Analyzer France			
	Background rates				Background rates				Background rates			
	Years of follow-up	No. of outcomes	IR per 100,000 person-years	95% CI	Years of follow-up	No. of outcomes	IR per 100,000 person-years	95% CI	Years of follow-up	No. of outcomes	IR per 100,000 person-years	95% CI
Overall	8.684.104,50	6	**0,07**	0.03 - 0.15	35.309.278,25	85	**0,24**	0.19 - 0.30	5.814.924,92	0	**0,00**	0.00 - 0.06
Sex												
Female	-	<6	**0,02**	0.01 - 0.13	20.252.728,94	47	**0,23**	0.17 - 0.31	3.092.562,03	0	**0,00**	0.00 - 0.12
Male	-	<6	**0,11**	0.05 - 0.27	15.056.549,31	38	**0,25**	0.18 - 0.35	2.722.362,90	0	**0,00**	0.00 - 0.14
Year												
2016	-	<6	**0,06**	0.01 - 0.34	7.071.625,06	12	**0,17**	0.10 - 0.30	1.300.302,59	0	**0,00**	0.00 - 0.28
2017	1.694.990,94	0	**0,00**	0.00 - 0.22	7.221.888,39	23	**0,32**	0.21 - 0.48	1.278.734,16	0	**0,00**	0.00 - 0.29
2018	1.743.245,24	0	**0,00**	0.00 - 0.21	7.400.600,36	11	**0,15**	0.08 - 0.27	1.212.235,58	0	**0,00**	0.00 - 0.30
2019	-	<6	**0,11**	0.03 - 0.41	7.099.391,99	22	**0,31**	0.21 - 0.47	1.099.947,02	0	**0,00**	0.00 - 0.34
2020	-	<6	**0,16**	0.06 - 0.48	6.515.772,45	17	**0,26**	0.16 - 0.42	923.705,58	0	**0,00**	0.00 - 0.40

N.B.: Rates are not presented stratified by age given the small numbers of events.
IR = incidence rates, 95% CI = 95% confidence intervals ***** ILD based on conditions expected to be a drug-related event.**

6. Discussion

6.1. Key results

Based on a broad definition (SML terms mapped to SNOMED and OMOP concepts), the overall incidence rate of ILD varied between 17 and 26 per 100,000 person-years. In general, incidence rates were higher in male compared to female subjects and there was a pronounced increase in the incidence of ILD with increasing age. Excluding data from 2020, there was an increase in the incidence of ILD over time, particularly in the IMRD UK and IQVIA™ Disease Analyzer Germany databases. **These patterns are broadly in line with published literature** [5].

Similar patterns were obtained when using the **restricted definition of ILD** (including only ILD-specific codes), however, the overall incidence rate of ILD was 10.40, 4.82 and 4.23 per 100,100 person-years for the UK, Germany and France, respectively. **These figures are in line with previous study** on the ILD incidence in the UK and the European Union using data obtained from the Global Burden of Disease Study on residents of the UK and of 27 EU countries [5], in which cases were defined as a combined output of ILD including pulmonary sarcoidosis using the ICD-10 codes J84 and D86, respectively. The male age-standardised incidence rates of ILD in 2017 were 10,91, 5,57, and 5,37 per 100,000 person-years in the UK, Germany and France, respectively. The female age-standardised incidence rates in 2017 were 6,7, 3,14 and 3,8 per 100,000 person-years [5]. Our incidence rates of ILD stratified by sex and using the restricted definition were similar to these previous published figures.

When using terms expected to be a drug-related event for the definition of ILD, the overall incidence rate of ILD was <1 per 100,000 person-years in IQVIA™ Disease Analyzer Germany. We identified a limited number of cases or no cases in IMRD UK and in IQVIA™ Disease Analyzer France, respectively. These very low rates of recorded drug-induced ILD does not necessarily imply that drug-induced ILD is as rare as the reported figures, but it may be a result of physicians not recording the likely cause of the ILD.

Given the multiple aetiological factors and difficulty in diagnosis as well as geographical differences in diagnosis criteria and thresholds, further exploratory analyses for the case definition of ILD using electronic medical record databases is warranted.

6.2. Limitations

Diagnostic coding for ILD is not known to have been validated in the primary care databases available. Given the diagnostic difficulty of ILD, the validity of the coding in primary care databases require further investigation. Therefore, our results should be interpreted cautiously.

It is worth noting that entire patient history may not be included in the data sources, and that there is a risk that a prevalent case may have been misclassified as incident (prevalent pool effect). This may result in overestimation of the incidence rate.

Changes in healthcare utilisation during the COVID-19 pandemic (2020-present) might affect routine clinical practices and information recording, therefore the reported background rates for 2020 and 2021 may be distorted due to potential changes in the way patients interacted with healthcare services in these years.

7. References

[1] Kaul B, Cottin V, Collard HR and Valenzuela C (2021) Variability in Global Prevalence of Interstitial Lung Disease. Front. Med. 8:751181.

[2] Pinheiro L, Blake K, Januskiene J, Yue QY, Arlett P. Geographical variation in reporting Interstitial Lung Disease as an adverse drug reaction: findings from an European Medicines Agency analysis of reports in EudraVigilance. Pharmacoepidemiol Drug Saf. 2016 Jun;25(6):705-12.

[3] Spagnolo P, Bonniaud P, Rossi G, et al. Drug-induced interstitial lung disease. Eur Respir J 2022; in press (https://doi.org/10.1183/13993003.02776-2021).

[4] MedDRA®. Introductory guide version 25.0. March, 2022.

[5] Salciccioli JD, Marshall DC, Goodall R, et al. Interstitial lung disease incidence and mortality in the UK and the European Union: an observational study, 2001–2017. ERJ Open Res 2022;8: 00058-2022 [DOI: 10.1183/23120541.00058-2022].

Annexes

Annex 1 - Information on Databases and Healthcare systems included

IQVIA™ Medical Research Data (IMRD) UK

IQVIA™ Medical Research Data (IMRD) UK is a primary care database from the UK. GPs play a gatekeeper role in the healthcare system in the UK, as they are responsible for delivering primary health care and specialist referrals. Over 98% of the UK-resident population is registered with a GP, so that GP patient records are broadly representative of the UK population in general. Patients are affiliated to a practice, which centralizes the medical information from GPs, specialist referrals, hospitalizations, and tests.

IQVIA™ Disease Analyser Germany

IQVIA™ Disease Analyser Germany collects computerised information from specialised and general primary care practices throughout Germany since 1992. Around 3% of general practitioners (GP) practices are included, which covers all patients consulting a practice. Data from IQVIA™ Disease Analyzer Germany have been shown to be reasonably representative of German healthcare statistics for demographics and certain diseases and is considered one of the largest national medical databases worldwide. IQVIA™ Disease Analyzer Germany includes more than 2,500 practices and 3,100 physicians (13 speciality groups) representing over 15,000,000 patients. This database used to be named IMS® Germany and some use of this terminology may persist.

The quality of IQVIA™ Disease Analyzer data is ensured by a series of continuous QA controls and data refinement. These include checking incoming data for criteria such as completeness and correctness, (e.g. linkage between diagnoses and prescriptions), and standardizing certain data values such as laboratory test results in order to enable reliable analysis.

IQVIA™ Disease Analyzer France

IQVIA™ Disease Analyzer France collects anonymised patient medical records since 1997 through a representative panel of GPs. The physician sample represents approximately 2% of physicians and is weighted by age and gender of the physician, doctor region and the SNIR of the physician (National Official Indicator of the GP volume of activity in terms of visits and consultations). Some 99% of the French population is insured, but there are differences regarding level of coverage. IQVIA™ Disease Analyzer France includes around 1,000 GPs and represents more than 4,000,000 of patients and considered representative for the French population. This database used to be named IMS® France and some use of this terminology may persist.

The quality of IQVIA™ Disease Analyzer data is ensured by a series of continuous QA controls and data refinement. These include checking incoming data for criteria such as completeness and correctness, (e.g. linkage between diagnoses and prescriptions), and standardizing certain data values such as laboratory test results in order to enable reliable analysis.

Annex 2 – Code lists

SMQ Codes for ILD *(which were subsequently mapped to SNOMED codes and OMOP concepts)*

Preferred term	SMQ codes
Acute interstitial pneumonitis	10066728
Alveolar lung disease	10073344
Alveolar proteinosis	10001881
Alveolitis	10001889
Alveolitis necrotising	10050343
Autoimmune lung disease	10080701
~~Bronchiolitis~~***	~~10006448~~
~~Bronchiolitis obliterans syndrome~~***	~~10006448~~
Chronic graft versus host disease in lung	10086041
Combined pulmonary fibrosis and emphysema	10076515
Confirmed e-cigarette or vaping product use associated lung injury	10085189
Diffuse alveolar damage	10060902
~~Eosinophilia myalgia syndrome~~***	10014952
~~Eosinophilic granulomatosis with polyangiitis~~***	10078117
Eosinophilic pneumonia	10014962
Eosinophilic pneumonia acute	10052832
Eosinophilic pneumonia chronic	10052833
Hypersensitivity pneumonitis	10081988
Idiopathic interstitial pneumonia	10078268
Idiopathic pneumonia syndrome	10063725
Idiopathic pulmonary fibrosis	10021240
Immune-mediated lung disease	10085352
Interstitial lung abnormality	10087834
Interstitial lung disease	10022611
Low lung compliance	10086117
Lung infiltration	10025102
Lung opacity	10081792
Necrotising bronchiolitis	10070831

Obliterative bronchiolitis	10029888
Pleuroparenchymal fibroelastosis	10084305
Pneumonitis	10035742
Probable e-cigarette or vaping product use associated lung injury	10085188
Progressive massive fibrosis	10036805
Pulmonary fibrosis	10037383
~~Pulmonary necrosis~~***	10058824
Pulmonary radiation injury	10061473
~~Pulmonary toxicity~~***	10061924
Pulmonary vasculitis	10037457
Radiation alveolitis	10037754
Radiation bronchitis	10085628
Radiation fibrosis - lung	10037758
Radiation pneumonitis	10037765
Rheumatoid arthritis-associated interstitial lung disease	10085517
Small airways disease	10080547
Transfusion-related acute lung injury	10052235

*** These terms were excluded.

Selected ILD based on SMQ code list (including fibrosis, pneumonitis, alveolitis)

OMOP concept ID	Description
4170900	Asthmatic pulmonary alveolitis
444084	Extrinsic allergic alveolitis
37110889	House allergic alveolitis
4243675	Hypersensitivity alveolitis in lungworm infection
4215594	Prolonged pulmonary alveolitis
4119785	Pyrethrum alveolitis
4184896	Rheumatoid fibrosing alveolitis
4243523	Simple pulmonary alveolitis
4008132	Tropical pulmonary alveolitis
37110291	Idiopathic acute eosinophilic pneumonia
37110292	Idiopathic chronic eosinophilic pneumonia
45769390	Idiopathic eosinophilic pneumonia
4213867	Hypersensitivity pneumonitis panel measurement

4119428	Summer-type hypersensitivity pneumonitis
45763749	Idiopathic interstitial pneumonia
36712839	Idiopathic pneumonia syndrome
762964	Chronic interstitial lung disease
37116655	Congenital nephrotic syndrome, interstitial lung disease, epidermolysis bullosa syndrome
4119786	Interstitial lung disease
37207520	Interstitial lung disease annual review
4140605	Interstitial lung disease due to collagen vascular disease
46272927	Interstitial lung disease due to connective tissue disease
42539687	Interstitial lung disease due to granulomatous disease
3655634	Interstitial lung disease due to juvenile polymyositis
42537658	Interstitial lung disease due to metabolic disease
42537657	Interstitial lung disease due to systemic disease
46270493	Interstitial lung disease of childhood
42539090	Interstitial lung disease with systemic vasculitis
37208102	PF-ILD-progressive fibrosing interstitial lung disease
44783629	Pulmonary hypertension due to interstitial lung disease
4045227	Respiratory bronchiolitis associated interstitial lung disease
4112678	Acute pneumonitis due to chemical fumes
260434	Acute radiation pneumonitis
4306082	Aspiration pneumonitis
4112809	Aspiration pneumonitis due to anesthesia during labor and delivery
37018497	Aspiration pneumonitis of fetus
439298	Bronchitis and pneumonitis due to chemical fumes
4291799	Cadmium pneumonitis
46274046	Chemical pneumonitis caused by anesthesia
45769386	Chronic pneumonitis of infancy
4188480	Congenital rubella pneumonitis
4140472	Drug-induced pneumonitis
4103099	GIP - Giant cell interstitial pneumonitis
256036	Hemorrhagic varicella pneumonitis
4218175	Lipoid pneumonitis
4229303	Manganese pneumonitis
4315046	Meconium pneumonitis
253506	Pneumonitis
3655110	Pneumonitis caused by fumes
3655088	Pneumonitis caused by inhalation of oil
3655111	Pneumonitis caused by vapors
444099	Pneumonitis due to acquired toxoplasmosis
42536542	Pneumonitis due to aspiration of blood
36712850	Pneumonitis due to Herpes zoster

4311410	Pneumonitis due to inhalation of essence
4112683	Pneumonitis due to inhalation of milk
4112839	Pneumonitis due to inhalation of regurgitated food
4110180	Pneumonitis due to inhalation of vomitus
442297	Pneumonitis due to inhaled liquid
256721	Pneumonitis due to inhaled solid
4187218	Pneumonitis due to inhaled substance
4226132	Post-radiotherapy pneumonitis
4221865	Radiation pneumonitis
4119796	Toxic pneumonitis
4089507	Toxoplasma pneumonitis
45768909	Autoimmune pulmonary alveolar proteinosis
45772936	Congenital pulmonary alveolar proteinosis
435853	Pulmonary alveolar proteinosis
45771023	Secondary pulmonary alveolar proteinosis
37312199	Acute exacerbation of idiopathic pulmonary fibrosis
4140134	Familial idiopathic pulmonary fibrosis
45763750	Idiopathic pulmonary fibrosis
3655115	Chronic pulmonary fibrosis caused by chemical vapors
4112681	Chronic pulmonary fibrosis due to chemical fumes
4025168	Diffuse interstitial pulmonary fibrosis
4119793	Drug-induced diffuse interstitial pulmonary fibrosis
37017059	Drug induced pulmonary fibrosis
36675042	Hereditary fibrosing poikiloderma, tendon contractures, myopathy, pulmonary fibrosis syndrome
4236182	Interstitial pulmonary fibrosis of prematurity
4120272	Localized pulmonary fibrosis
253797	Post-inflammatory pulmonary fibrosis
4236725	Pulmonary fibrosis due to and following radiotherapy
45768903	Pulmonary fibrosis due to Hermansky-Pudlak syndrome
37109889	Pulmonary fibrosis, hepatic hyperplasia, bone marrow hypoplasia syndrome
4119792	Toxic diffuse interstitial pulmonary fibrosis
4112813	Drug-induced interstitial lung disorder
4110182	Acute drug-induced interstitial lung disorder
4112814	Chronic drug-induced interstitial lung disorders

Selected ILD-specific codes

OMOP concept ID	Description
762964	Chronic interstitial lung disease
37116655	Congenital nephrotic syndrome, interstitial lung disease, epidermolysis bullosa syndrome

4119786	Interstitial lung disease
37207520	Interstitial lung disease annual review
4140605	Interstitial lung disease due to collagen vascular disease
46272927	Interstitial lung disease due to connective tissue disease
42539687	Interstitial lung disease due to granulomatous disease
3655634	Interstitial lung disease due to juvenile polymyositis
42537658	Interstitial lung disease due to metabolic disease
42537657	Interstitial lung disease due to systemic disease
46270493	Interstitial lung disease of childhood
42539090	Interstitial lung disease with systemic vasculitis
37208102	PF-ILD-progressive fibrosing interstitial lung disease
44783629	Pulmonary hypertension due to interstitial lung disease
4045227	Respiratory bronchiolitis associated interstitial lung disease

Selected Drug-induced ILD

OMOP concept ID	Description
4112809	Aspiration pneumonitis due to anesthesia during labor and delivery
46274046	Chemical pneumonitis caused by anesthesia
4140472	Drug-induced pneumonitis
4119793	Drug-induced diffuse interstitial pulmonary fibrosis
37017059	Drug induced pulmonary fibrosis
4112813	Drug-induced interstitial lung disorder
4110182	Acute drug-induced interstitial lung disorder
4112814	Chronic drug-induced interstitial lung disorders

Selected Fibrosis ILD codes

OMOP concept ID	Description
4184896	Rheumatoid fibrosing alveolitis
37312199	Acute exacerbation of idiopathic pulmonary fibrosis
4140134	Familial idiopathic pulmonary fibrosis
45763750	Idiopathic pulmonary fibrosis
37312199	Acute exacerbation of idiopathic pulmonary fibrosis
3655115	Chronic pulmonary fibrosis caused by chemical vapors
4112681	Chronic pulmonary fibrosis due to chemical fumes
4025168	Diffuse interstitial pulmonary fibrosis

4119793	Drug-induced diffuse interstitial pulmonary fibrosis
37017059	Drug induced pulmonary fibrosis
4140134	Familial idiopathic pulmonary fibrosis
36675042	Hereditary fibrosing poikiloderma, tendon contractures, myopathy, pulmonary fibrosis syndrome
45763750	Idiopathic pulmonary fibrosis
4236182	Interstitial pulmonary fibrosis of prematurity
4120272	Localized pulmonary fibrosis
253797	Post-inflammatory pulmonary fibrosis
4236725	Pulmonary fibrosis due to and following radiotherapy
45768903	Pulmonary fibrosis due to Hermansky-Pudlak syndrome
37109889	Pulmonary fibrosis, hepatic hyperplasia, bone marrow hypoplasia syndrome
4119792	Toxic diffuse interstitial pulmonary fibrosis

Selected Pneumonitis ILD codes

OMOP concept ID	Description
37110291	Idiopathic acute eosinophilic pneumonia
37110292	Idiopathic chronic eosinophilic pneumonia
45769390	Idiopathic eosinophilic pneumonia
4213867	Hypersensitivity pneumonitis panel measurement
4119428	Summer-type hypersensitivity pneumonitis
45763749	Idiopathic interstitial pneumonia
36712839	Idiopathic pneumonia syndrome
4112678	Acute pneumonitis due to chemical fumes
260434	Acute radiation pneumonitis
4306082	Aspiration pneumonitis
4112809	Aspiration pneumonitis due to anesthesia during labor and delivery
37018497	Aspiration pneumonitis of fetus
439298	Bronchitis and pneumonitis due to chemical fumes
4291799	Cadmium pneumonitis
46274046	Chemical pneumonitis caused by anesthesia
45769386	Chronic pneumonitis of infancy
4188480	Congenital rubella pneumonitis
4140472	Drug-induced pneumonitis
4103099	GIP - Giant cell interstitial pneumonitis
256036	Hemorrhagic varicella pneumonitis
4213867	Hypersensitivity pneumonitis panel measurement
4218175	Lipoid pneumonitis
4229303	Manganese pneumonitis

4315046	Meconium pneumonitis
253506	Pneumonitis
3655110	Pneumonitis caused by fumes
3655088	Pneumonitis caused by inhalation of oil
3655111	Pneumonitis caused by vapors
444099	Pneumonitis due to acquired toxoplasmosis
42536542	Pneumonitis due to aspiration of blood
36712850	Pneumonitis due to Herpes zoster
4311410	Pneumonitis due to inhalation of essence
4112683	Pneumonitis due to inhalation of milk
4112839	Pneumonitis due to inhalation of regurgitated food
4110180	Pneumonitis due to inhalation of vomitus
442297	Pneumonitis due to inhaled liquid
256721	Pneumonitis due to inhaled solid
4187218	Pneumonitis due to inhaled substance
4226132	Post-radiotherapy pneumonitis
4221865	Radiation pneumonitis
4119428	Summer-type hypersensitivity pneumonitis
4119796	Toxic pneumonitis
4089507	Toxoplasma pneumonitis
260434	Acute radiation pneumonitis
4221865	Radiation pneumonitis

Selected Alveolitis ILD codes

OMOP concept ID	Description
4170900	Asthmatic pulmonary alveolitis
444084	Extrinsic allergic alveolitis
37110889	House allergic alveolitis
4243675	Hypersensitivity alveolitis in lungworm infection
4215594	Prolonged pulmonary alveolitis
4119785	Pyrethrum alveolitis
4184896	Rheumatoid fibrosing alveolitis
4243523	Simple pulmonary alveolitis
4008132	Tropical pulmonary alveolitis
37110291	Idiopathic acute eosinophilic pneumonia
37110292	Idiopathic chronic eosinophilic pneumonia
45769390	Idiopathic eosinophilic pneumonia
4213867	Hypersensitivity pneumonitis panel measurement

4119428	Summer-type hypersensitivity pneumonitis
45763749	Idiopathic interstitial pneumonia
36712839	Idiopathic pneumonia syndrome

Annex 3 – Additional incidence rates

Table S1. *Yearly background rates of ILD per 100,000 years of follow-up between 2004 and 2016 in IMRD UK.*

	IMRD UK			
	Background rates			
	Years of follow-up	No. of outcomes	IR per 100,000 person-years	95% CI
Year				
2004	696.172,61	21	**3,02**	1.98 - 4.61
2005	741.384,96	34	**4,59**	3.29 - 6.41
2006	789.830,05	37	**4,68**	3.40 - 6.46
2007	842.996,66	60	**7,12**	5.54 - 9.16
2008	904.063,18	64	**7,08**	5.55 - 9.04
2009	963.414,03	79	**8,20**	6.58 - 10.22
2010	1.028.347,03	84	**8,17**	6.60 - 10.11
2011	1.115.158,85	100	**8,97**	7.38 - 10.91
2012	1.216.821,67	102	**8,38**	6.91 - 10.18
2013	1.327.788,61	168	**12,65**	10.88 - 14.72
2014	1.446.309,74	176	**12,17**	10.50 - 14.11
2015	1.546.869,27	241	**15,58**	13.74 - 17.68
2016	1.618.760,38	285	**17,61**	15.68 - 19.77

***** ILD definition based on SMQ terms mapped to SNOMED codes and OMOP concepts**

IR = incidence rates, 95% CI = 95% confidence intervals

Table S2. *Age-stratified background rates of ILD*** per 100,000 years of follow-up for female and male subjects between 2017 and 2021 in IMRD UK, and between 2016 and 2020 in IQVIA™ Disease Analyzer Germany and France.*

	IMRD UK				IQVIA™ Disease Analyzer Germany				IQVIA™ Disease Analyzer France			
	Background rates				Background rates				Background rates			
	Years of follow-up	No. of outcomes	IR per 100,000 person-years	95% CI	Years of follow-up	No. of outcomes	IR per 100,000 person-years	95% CI	Years of follow-up	No. of outcomes	IR per 100,000 person-years	95% CI
Age (in years)												
Female												
0-19	944.051,84	19	**2,01**	1.29 - 3.14	2.089.907,42	114	**5,45**	4.54 - 6.55	589.426,16	40	**6,79**	4.99 - 9.24
20-29	552.603,97	11	**1,99**	1.12 - 3.56	1.799.298,66	55	**3,06**	2.35 - 3.98	335.662,24	54	**16,09**	12.34 - 20.99
30-39	695.060,87	39	**5,61**	4.11 - 7.67	2.287.728,02	113	**4,94**	4.11 - 5.94	394.899,64	92	**23,30**	19.01 - 28.57
40-49	602.553,11	54	**8,96**	6.88 - 11.69	2.680.604,41	185	**6,90**	5.98 - 7.97	425.347,10	152	**35,74**	30.50 - 41.89
50-59	592.571,87	128	**21,60**	18.18 - 25.68	3.702.489,18	388	**10,48**	9.49 - 11.58	426.796,56	221	**51,78**	45.40 - 59.08
60-69	442.171,43	189	**42,74**	37.08 - 49.29	3.124.965,52	437	**13,98**	12.73 - 15.36	387.876,43	142	**36,61**	31.07 - 43.15
70-79	354.480,86	198	**55,86**	48.61 - 64.20	2.647.692,42	559	**21,11**	19.43 - 22.94	278.707,30	78	**27,99**	22.44 - 34.93
80+	259.878,14	243	**93,51**	82.48 - 106.03	1.755.787,99	922	**52,51**	49.23 - 56.01	214.965,98	51	**23,72**	18.07 - 31.19
Male												
0-19	1.003.292,66	19	**1,89**	1.22 - 2.96	2.123.079,51	160	**7,54**	6.46 - 8.80	645.087,84	31	**4,81**	3.39 - 6.82
20-29	539.012,04	10	**1,86**	1.02 - 3.41	1.146.945,60	95	**8,28**	6.78 - 10.13	264.586,43	36	**13,61**	9.85 - 18.84
30-39	687.382,96	30	**4,36**	3.06 - 6.23	1.377.486,16	100	**7,26**	5.97 - 8.83	298.746,93	80	**26,78**	21.53 - 33.33
40-49	650.009,42	64	**9,85**	7.72 - 12.57	1.716.135,89	193	**11,25**	9.77 - 12.95	363.464,00	128	**35,22**	29.63 - 41.87
50-59	631.633,64	125	**19,79**	16.62 - 23.58	2.713.342,07	386	**14,23**	12.88 - 15.72	383.887,45	143	**37,25**	31.63 - 43.88
60-69	440.622,43	195	**44,26**	38.47 - 50.92	2.475.433,61	554	**22,38**	20.59 - 24.32	349.882,53	128	**36,58**	30.78 - 43.50
70-79	320.421,50	338	**105,49**	94.83 - 117.35	2.143.919,27	754	**35,17**	32.75 - 37.77	241.332,23	85	**35,22**	28.50 - 43.55
80+	182.105,48	316	**173,53**	155.44-193.75	1.240.088,15	906	**73,06**	68.46 - 77.98	139.397,43	48	**34,43**	26.01 - 45.65

IR = incidence rates, 95% CI = 95% confidence intervals

***** ILD definition based on SMQ terms mapped to SNOMED codes and OMOP concepts.**

Table S3. *Background rates of fibrosis ILD per 100,000 years of follow-up between 2016 and 2020 in IMRD UK.*

	IMRD UK			
	Background rates			
	Years of follow-up	No. of outcomes	IR per 100,000 person-years	95% CI
Overall	8,681,129.98	427	**4.92**	4.47 - 5.41
Sex				
Female	4,334,900.80	152	**3.51**	2.99 - 4.11
Male	4,346,229.18	275	**6.33**	5.62 - 7.12
Age group (in years)				
0-19	-	<6	**0.05**	0.01 - 0.29
20-29	-	<6	**0.09**	0.02 - 0.52
30-39	-	<6	**0.15**	0.05 - 0.54
40-49	-	<6	**0.24**	0.09 - 0.71
50-59	1,193,943.63	21	**1.76**	1.15 - 2.69
60-69	865,331.31	67	**7.74**	6.10 - 9.83
70-79	655,750.57	155	**23.64**	20.20 - 27.66
80+	435,785.06	178	**40.85**	35.28 - 47.31
Year				
2016	1,635,427.05	86	**5.26**	4.26 - 6.49
2017	1,694,400.59	90	**5.31**	4.32 - 6.53
2018	1,742,641.83	85	**4.88**	3.95 - 6.03
2019	1,775,664.75	84	**4.73**	3.82 - 5.86
2020	1,832,995.75	82	**4.47**	3.61 - 5.55

Note: Zero ILD-fibrosis codes were recorded for IQVIA DA Germany and IQVIA DA France.

Table S4. *Background rates of pneumonitis ILD per 100,000 years of follow-up between 2016 and 2020 in IMRD UK, IQVIA™ Disease Analyzer Germany and France.*

	IMRD UK				IQVIA™ Disease Analyzer Germany				IQVIA™ Disease Analyzer France			
	Background rates				Background rates				Background rates			
	Years of follow-up	No. of outcomes	IR per 100,000 person-years	95% CI	Years of follow-up	No. of outcomes	IR per 100,000 person-years	95% CI	Years of follow-up	No. of outcomes	IR per 100,000 person-years	95% CI
Overall	8,681,824.33	514	**5.92**	5.43 - 6.46	35,297,192.34	3,452	**9.78**	9.46 - 10.11	5,808,988.65	1,115	**19.19**	18.10 - 20.35
Sex												
Female	4,335,114.39	236	**5.44**	4.79 - 6.18	20,247,480.25	1,533	**7.57**	7.20 - 7.96	3,089,221.17	629	**20.36**	18.83 - 22.02
Male	4,346,709.94	278	**6.4**	5.69 - 7.19	15,049,712.09	1,919	**12.75**	12.19 - 13.33	2,719,767.48	486	**17.87**	16.35 - 19.53
Age group (in years)												
0-19	1,891,289.02	22	**1.16**	0.77 - 1.76	4,468,692.76	160	**3.58**	3.07 - 4.18	1,301,808.53	37	**2.84**	2.07 - 3.92
20-29	1,072,300.40	12	**1.12**	0.65 - 1.95	2,946,651.55	77	**2.61**	2.09 - 3.27	600,353.25	67	**11.16**	8.80 - 14.17
30-39	1,331,102.52	19	**1.43**	0.92 - 2.23	3,665,999.90	77	**2.1**	1.68 - 2.63	693,824.43	139	**20.03**	16.97 - 23.65
40-49	1,234,948.00	41	**3.32**	2.45 - 4.50	4,398,337.88	141	**3.21**	2.72 - 3.78	789,065.28	229	**29.02**	25.50 - 33.03
50-59	1,193,822.20	66	**5.53**	4.35 - 7.03	6,419,165.76	316	**4.92**	4.41 - 5.50	810,967.51	308	**37.98**	33.97 - 42.47
60-69	865,419.09	89	**10.28**	8.36 - 12.66	5,604,505.09	431	**7.69**	7.00 - 8.45	738,053.62	202	**27.37**	23.85 - 31.42
70-79	656,457.37	116	**17.67**	14.74 - 21.19	4,795,503.21	748	**15.6**	14.52 - 16.76	520,325.34	87	**16.72**	13.56 - 20.62
80+	436,485.73	149	**34.14**	29.09 - 40.08	2,998,336.19	1,502	**50.09**	47.63 - 52.69	354,590.69	46	**12.97**	9.74 - 17.30
Year												
2016	1,635,641.27	77	**4.71**	3.77 - 5.88	7,069,702.37	571	**8.08**	7.44 - 8.77	1,299,240.95	318	**24.48**	21.93 - 27.32
2017	1,694,575.85	92	**5.43**	4.43 - 6.66	7,219,698.30	649	**8.99**	8.32 - 9.71	1,277,527.70	267	**20.90**	18.54 - 23.56
2018	1,742,793.39	99	**5.68**	4.67 - 6.92	7,398,116.00	734	**9.92**	9.23 - 10.67	1,210,943.87	250	**20.65**	18.24 - 23.37
2019	1,775,768.93	110	**6.19**	5.14 - 7.47	7,096,719.12	756	**10.65**	9.92 - 11.44	1,098,691.89	208	**18.93**	16.53 - 21.69
2020	1,833,044.89	136	**7.42**	6.27 - 8.78	6,512,956.56	742	**11.39**	10.60 - 12.24	922,584.24	72	**7.8**	6.20 - 9.83

Table S5. *Background rates of Alveolitis ILD per 100,000 years of follow-up between 2016 and 2020 in IMRD UK, IQVIA™ Disease Analyzer Germany and France.*

	IMRD UK				IQVIA™ Disease Analyzer Germany				IQVIA™ Disease Analyzer France			
	Background rates				Background rates				Background rates			
	Years of follow-up	No. of outcomes	IR per 100,000 person-years	95% CI	Years of follow-up	No. of outcomes	IR per 100,000 person-years	95% CI	Years of follow-up	No. of outcomes	IR per 100,000 person-years	95% CI
Overall	8,682,889.50	169	**1.95**	1.67 - 2.26	35,302,291.34	878	**2.49**	2.33 - 2.66	5,814,065.37	152	**2.61**	2.23 - 3.06
Sex												
Female	4,335,370.42	105	**2.42**	2.00 - 2.93	20,248,907.87	461	**2.28**	2.08 - 2.49	3,092,055.30	85	**2.75**	2.22 - 3.40
Male	4,347,519.09	64	**1.47**	1.15 - 1.88	15,053,383.47	417	**2.77**	2.52 - 3.05	2,722,010.08	67	**2.46**	1.94 - 3.13
Age group (in years)												
0-19	1,891,509.76	0	**0.00**	0.00 - 0.20	4,469,797.28	39	**0.87**	0.64 - 1.19	1,301,755.71	28	**2.15**	1.49 - 3.11
20-29	1,072,516.37	0	**0.00**	0.00 - 0.34	2,946,823.97	49	**1.66**	1.26 - 2.20	600,566.03	19	**3.16**	2.03 - 4.94
30-39	1,331,219.33	6	**0.45**	0.21 - 0.98	3,666,222.24	73	**1.99**	1.58 - 2.50	694,469.22	17	**2.45**	1.54 - 3.92
40-49	1,235,048.87	17	**1.38**	0.86 - 2.20	4,398,474.82	113	**2.57**	2.14 - 3.09	790,021.14	29	**3.67**	2.56 - 5.27
50-59	1,193,962.47	43	**3.60**	2.68 - 4.85	6,419,399.96	204	**3.18**	2.77 - 3.65	812,495.75	18	**2.22**	1.41 - 3.50
60-69	865,424.17	42	**4.85**	3.60 - 6.56	5,604,660.84	204	**3.64**	3.17 - 4.18	739,240.81	23	**3.11**	2.08 - 4.67
70-79	656,486.28	41	**6.25**	4.61 - 8.47	4,796,308.88	126	**2.63**	2.21 - 3.13	520,752.14	15	**2.88**	1.76 - 4.75
80+	436,722.26	20	**4.58**	2.98 - 7.07	3,000,603.36	70	**2.33**	1.85 - 2.95	-	<10	**0.85**	0.31 - 2.47
Year												
2016	1,635,818.05	31	**1.90**	1.34 - 2.69	7,070,395.08	163	**2.31**	1.98 - 2.69	1,300,137.75	46	**3.54**	2.66 - 4.72
2017	1,694,772.11	32	**1.89**	1.34 - 2.67	7,220,568.78	167	**2.31**	1.99 - 2.69	1,278,552.74	28	**2.19**	1.52 - 3.17
2018	1,743,002.35	40	**2.29**	1.69 - 3.12	7,399,182.24	167	**2.26**	1.94 - 2.63	1,212,051.04	38	**3.14**	2.29 - 4.30
2019	1,775,995.60	37	**2.08**	1.51 - 2.87	7,097,909.50	195	**2.75**	2.39 - 3.16	1,099,773.86	25	**2.27**	1.54 - 3.36
2020	1,833,301.38	29	**1.58**	1.10 - 2.27	6,514,235.73	186	**2.86**	2.47 - 3.30	923,549.98	15	**1.62**	0.99 - 2.68