

# International Journal of Population Data Science

Journal Website: [www.ijpds.org](http://www.ijpds.org)



Swansea University  
Prifysgol Abertawe

## Ascertaining Household-Level Exposure to Air Pollution: Socio-Demographic Patterning and Urban-Rural Differences in People with Schizophrenia and Other Psychotic Disorders and the General Population

Ella S. Christoforou<sup>1,2,\*</sup>, Jane Lyons<sup>2</sup>, Ann John<sup>2</sup>, Rich Fry<sup>2</sup>, Martin J.D. Clift<sup>3</sup>, and Amy Mizen<sup>2</sup>

### Submission History

|            |            |
|------------|------------|
| Submitted: | 19/09/2025 |
| Accepted:  | 24/04/2026 |
| Published: | 02/07/2026 |

<sup>1</sup> Institute of Life Science, Swansea University Medical School, Swansea SA2 8PP, UK

<sup>2</sup> Population Data Science, Swansea University Medical School, Swansea SA2 8PP, UK

<sup>3</sup> In Vitro Toxicology Group, Institute of Life Science, Swansea University Medical School, Swansea SA2 8PP, UK

### Abstract

#### Background

People with severe mental disorders, including schizophrenia, experience worse physical health and shorter life expectancy. While socio-demographic factors are well established contributors to these disparities, less is known about how environmental exposures differ between individuals with severe mental disorders and the general population.

#### Objective

To explore socio-demographic and environmental differences between the general population and individuals with schizophrenia and/or other psychotic disorders (OPD).

#### Methods

Anonymised general practice records (2016–2019) for individuals resident in Wales were accessed within the Secure Anonymised Information Linkage (SAIL) Databank. This anonymised health data were linked to small-area level air pollution data for 2016 (PM<sub>10</sub>, PM<sub>2.5</sub>, NO<sub>x</sub>) for individuals who remained at the same residential address during the study period. Individuals with a first relevant diagnosis were identified; age, sex, urban/rural residence, deprivation, and air pollution exposures were compared with the general population.

#### Results

Overall, 0.1% (1,784) of the SAIL population had a first recorded diagnosis of schizophrenia/OPD. Of these, 52.7% (941) were male and 47.3% (843) were female. The psychiatric cohort had a higher mean age at study entry than the general population (46.9 vs 42.3 years). The distribution of diagnoses in urban (72.3%) and rural (27.7%) areas closely matched the total population (urban 70.5%, rural 29.5%). Mean air pollutant levels (PM<sub>10</sub>, PM<sub>2.5</sub>, NO<sub>x</sub>) in 2016 were similar between the psychiatric cohort and the general population. A clear socioeconomic gradient was observed: 29.1% of cases resided in the most deprived quintile compared with 15.0% in the least deprived quintile.

#### Conclusions

Area-level deprivation was associated with higher prevalence of schizophrenia/OPD, whereas no clear differences were observed by urbanicity or air pollution exposure. Linked population-scale health and environmental data provide valuable evidence which could inform service planning and targeted public health interventions.

#### Keywords

schizophrenia; psychotic disorders; air pollution; deprivation; electronic health records

\*Corresponding Author:

Email Address: [991929@swansea.ac.uk](mailto:991929@swansea.ac.uk) (Ella S. Christoforou)

## Introduction

Mental health problems can affect anyone and have a significant effect on the lives of individuals, their families, communities, and wider society. Globally, one in eight people live with a mental disorder [1] and the rates of mental illness have been gradually rising [2, 3]. In the UK, approximately 4.6 million people were diagnosed with a mental disorder (excluding alcohol and drug use disorders) in 2019 [4]. This equates to approximately one in four adults experiencing at least one diagnosable mental health condition annually [5]. In Wales, estimates suggest that severe mental illness (SMI) affects around 2% of the population, equivalent to approximately 62,150 people [6]. Within this group, schizophrenia and other psychotic disorders (OPD) represent a substantial proportion and are associated with significant functional impairment and long-term health inequalities.

Mental illness is closely associated with many forms of inequality. Health inequalities are avoidable and unfair differences in health status between groups of people. These differences arise from variations in social and economic determinants such as demographic, socioeconomic, and geographical factors [7]. Patients suffering from severe mental disorders, such as schizophrenia, have a reduced life expectancy (up to 10–25 years) compared to the general population [8]. Evidence also suggests that the mortality gap is widening [7]. The high mortality rate is a consequence of the simultaneous presence of comorbid physical health problems, such as cardiovascular, respiratory, and infectious diseases in addition to cancer [9, 10].

Environmental and socioeconomic factors may contribute to these inequalities. Air pollution represents a major global health challenge and is often more pronounced in urban environments and in lower income countries [11]. In the most deprived areas there are often higher air pollution levels [12], increased risk of mental illness [13] and decreased life expectancy compared to less deprived areas [14]. Furthermore, the risk of mental disorders is generally higher in more urbanised areas compared to less urbanised or rural areas [15–17]. However, a review [18] revealed complex patterns of urbanicity–psychosis associations with international variation within Europe and between low, middle, and high-income countries. The observed heterogeneity is likely due to multiple interacting risk factors e.g., air pollution and protective factors (e.g., greenness) [18, 19].

Growing epidemiological evidence suggests that exposure to air pollution may contribute to the onset and exacerbation of severe mental disorders, including schizophrenia [20]. Proposed biological mechanisms include neuroinflammation, oxidative stress, altered gene expression, and structural brain changes [20]. Empirical findings support this association. A national cohort study in Denmark found that childhood exposure to nitrogen oxides ( $\text{NO}_x$ ) was associated with an increased risk of schizophrenia after adjustment for deprivation and urbanicity [21]. However, findings were less consistent for particulate matter of different fractions ( $\text{PM}_{10}$  and  $\text{PM}_{2.5}$ ) [21]. Similarly, a Danish population-based cohort study demonstrated that higher childhood nitrogen dioxide ( $\text{NO}_2$ ) exposure and polygenic risk scores were independently associated with increased schizophrenia risk [22]. However, these studies did not describe or compare the environmental

contexts of individuals with and without schizophrenia. Using UK Biobank data, long-term exposure to multiple air pollutants ( $\text{PM}_{2.5}$ ,  $\text{PM}_{10}$ ,  $\text{NO}_x$ , and  $\text{NO}_2$ ) in adulthood has been associated with schizophrenia risk across genetic risk groups [23]. An analysis was also conducted describing urbanicity, deprivation, age, and sex in schizophrenia and non-schizophrenia populations.

Hence, existing studies have largely focused on estimating associations rather than describing and comparing the broader socio-demographic and environmental contexts of individuals with and without psychotic disorders. Moreover, evidence derived from routinely collected, population-level data remains limited [23].

This study builds on existing evidence by using linked, routinely collected population-level data from Wales. The linked data were used to describe and compare socio-demographic and environmental exposures, including deprivation, urbanicity, and air pollution, between individuals with psychotic disorders and the general population. This population-wide descriptive analysis establishes a baseline of inequalities to guide future research, service planning, and policy development [7].

**Aim:** To describe the socio-demographic and environmental differences between the general population and those with schizophrenia/OPD.

**Hypothesis:** Individuals diagnosed with schizophrenia/OPD are more likely to reside in areas with higher air pollution levels and greater deprivation compared to the general population.

### Objectives:

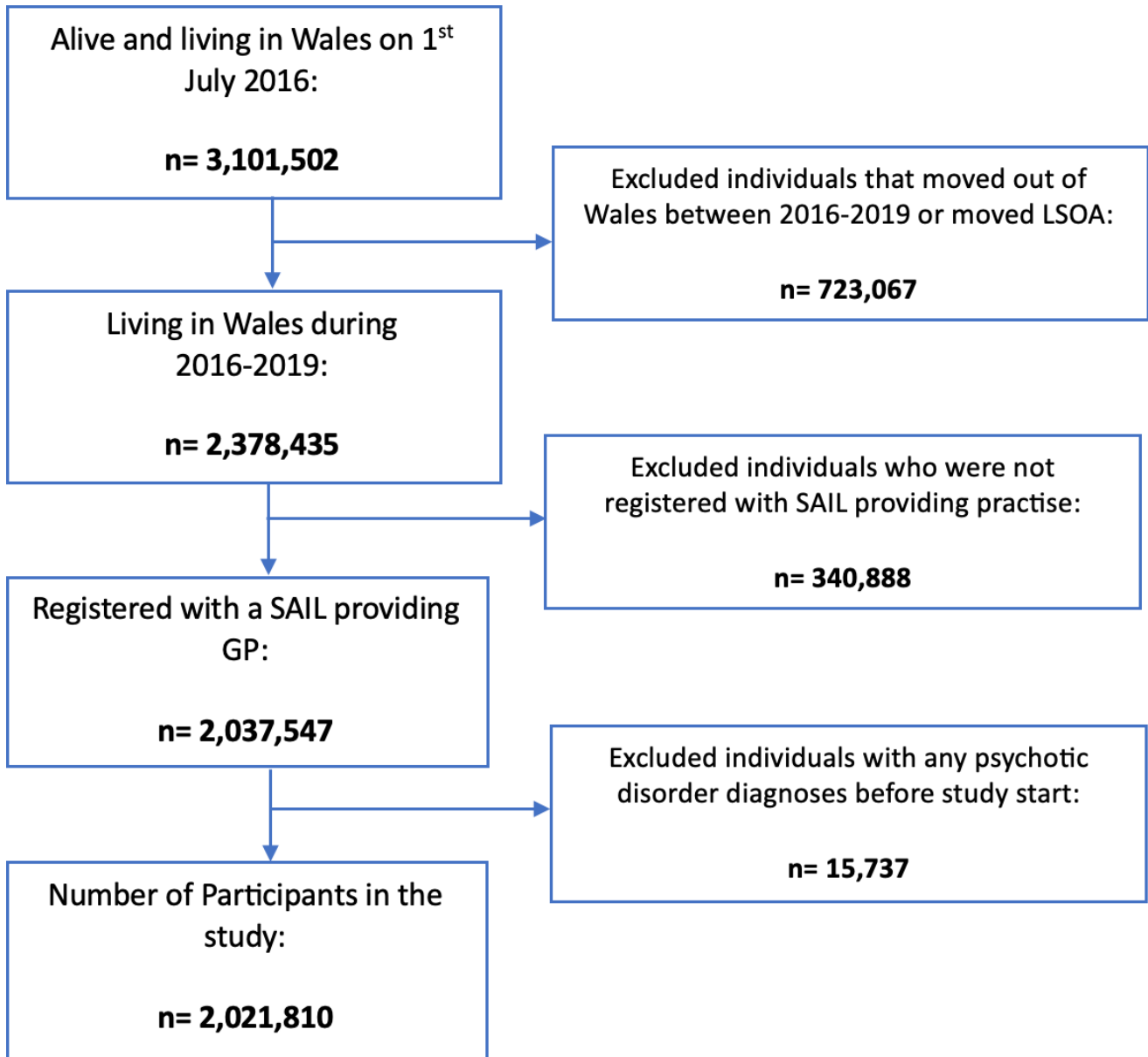
1. Create a longitudinal cohort of individuals living in Wales from 2016–2019 and flag individuals with their first diagnosis of schizophrenia/OPD during this time.
2. Link air pollution data with demographic and health data.
3. Calculate descriptive statistics e.g., age, sex, deprivation level, urban/rural classification, and air pollutants level in the general population and in those with schizophrenia/OPD.

## Methods

### Data Linkage

Previous annual air pollution data from 2016 at the Lower Layer Super Output Area (LSOA) level was linked to individuals registered with a Welsh General Practice (GP) between 1 July 2016 to 30 June 2019. Anonymised, longitudinal information on health, social and environmental data on the Welsh population is contained in the Secure Anonymised Information Linkage (SAIL) Databank [24, 25]. The SAIL databank allows consistent data linkages at the individual and household level. The study population comprised approximately 2 million people living in Wales who were registered with SAIL-contributing GP practices. Eligible participants were those living in Wales and registered with a Welsh GP throughout the full 3-year follow-up period (1 July 2016 to 30 June 2019).

Figure 1: Flow Diagram Illustrating Cohort Derivation and Participant Selection



The psychiatric cohort included individuals with their first schizophrenia/OPD diagnosis during this period. This timeframe was selected to avoid the COVID-19 pandemic, during which GP services were disrupted and patients reported difficulty accessing treatment [26]. To ensure accurate assignment of 2016 air pollution, participants were required to have a stable residential address throughout the study period. No restrictions were applied regarding age, sex, or environmental characteristics. Figure 1 illustrates the derivation of the final cohort.

## Study Area and Subjects

The study was conducted in Wales, a nation of approximately 3.13 million [27]. The population is predominantly concentrated in the south of the country, particularly in the urban centres of Cardiff (~372,000 residents) and Swansea (~241,000 residents) [27]. Air pollution levels vary across Wales. In 2020, South Wales recorded the second-highest annual mean

NO<sub>2</sub> concentration in the UK (64 µg/m<sup>3</sup>), after London (77 µg/m<sup>3</sup>), exceeding the UK and EU legal limit of 40 µg/m<sup>3</sup> [28, 29]. The North Wales zone and Cardiff Urban Area also exceeded the national legal limit (with 44 and 42 µg/m<sup>3</sup> respectively) [28]. Annual mean PM<sub>2.5</sub> concentrations in four out of ten Welsh locations exceeded the recommended WHO guidelines of 10 µg/m<sup>3</sup> [30, 31]. Overall, 88% of the Welsh population resides in built-up areas [32].

## Data Sources

### Mental Health Data

Schizophrenia/OPD were identified using the Welsh Longitudinal General Practice (WLGP) dataset. This dataset contains individual-level primary care records, including Read codes documenting diagnoses, symptoms and treatments for each patient. Read codes are a standardised clinical coding system used within the NHS since 1985 to record patient

information across primary and secondary care settings [33]. Code lists were obtained from the SAIL Databank Concept Library for schizophrenia-related disorders (schizophrenia, schizotypal and delusional disorders) [34]. In addition to code lists for other psychotic disorders such as acute and transient psychotic disorders, depressive episodes/disorders with psychotic symptoms [35]. To identify incident cases, individuals with a recorded diagnosis of schizophrenia or other psychotic disorder prior to the study start date (1 July 2016) were excluded.

### Air Pollution Data

This study was designed as a pilot analysis using the SAIL Databank to assess the feasibility of linking area-level air pollution with mental health outcomes. Data from a single year were used to provide a snapshot of exposure while limiting analytical complexity in this initial investigation. Although multi-year modelled air pollution data are available via the Department for Environment, Food & Rural Affairs (Defra), using multiple years of data was beyond the scope of this pilot study and is being addressed in ongoing work.

Air pollution data were obtained from DataMapWales [36], an open-access platform developed through a partnership between the Welsh Government and Natural Resources Wales. The data are derived from the UK-wide Defra air pollution dataset. For this study, annual mean concentrations of PM<sub>10</sub>, PM<sub>2.5</sub>, and NO<sub>x</sub> for 2016 were aggregated from 1 km<sup>2</sup> grid resolution to the LSOA level. Air pollution values were assigned from the 1km grid where the LSOA centroid was located using QGIS version 3.3. LSOAs are the smallest units of geography at which census data is estimated [37]. There are 1,909 LSOAs in Wales, each containing approximately 400-1,200 households and 1,000 and 3,000 residents [37], with a mean population of approximately 1,500 [38]. LSOAs are designed to be as homogenous as possible in terms of type of residence and urban/rural areas [37]. The 2011 rural-urban classification of LSOAs was obtained from the Office for National Statistics (ONS) [39].

### Demographic and Socioeconomic Data

To assign air pollution exposure and classify urban/rural status the Welsh Demographic Service dataset (WDS) was used. The WDS is a population register that contains administrative information and anonymised historical home addresses of patients registered with a Welsh GP.

The residential linkages in this dataset allow accurate allocation of individual-level air pollution exposure, creating high spatial resolution data [40]. Furthermore, this also enables inclusion of individuals who continuously resided within the same LSOA during the study period (1<sup>st</sup> July 2016 to 30<sup>th</sup> June 2019). Additionally, the WDS provides key demographic information, including age and sex.

Deprivation at the LSOA level was derived from the Welsh Index of Multiple Deprivation (WIMD) dataset which is the Welsh Government's official measure from 2019 [41]. WIMD identifies areas with the highest concentrations of eight different types of deprivation: income, employment, health, education, housing, access to services, community safety and physical environment [42]. LSOAs are ranked from 1 (most deprived) to 1,909 (least deprived) [41]. An area has a higher deprivation rank than another if the proportion of people living there are classified as deprived is higher. WIMD is a National Statistic produced by statisticians at the Welsh Government. A summary of the data sources and their coverage are presented in Table 1.

### Statistical Analysis

All analyses were conducted in R version 4.1.3 within the SAIL Databank. The dplyr package was used to generate summary statistics. Descriptive statistics were reported for the study population and for individuals with schizophrenia/OPD). The descriptive statistics included: sex (male/female), age at study start (mean, median, interquartile range [IQR]), concentrations of NO<sub>x</sub>, PM<sub>10</sub> and PM<sub>2.5</sub> (mean, minimum, median, maximum), rural urban classification and deprivation

Table 1: Summary of Datasets Used in this Study, Including Data Sources, Derived Variables, and Population or Geographic Coverage

| Dataset name                               | Data source                         | Derived variables                                                                             | Coverage                                       |
|--------------------------------------------|-------------------------------------|-----------------------------------------------------------------------------------------------|------------------------------------------------|
| Air pollution                              | DataMapWales (derived from [Defra]) | Annual mean concentrations of PM <sub>10</sub> , PM <sub>2.5</sub> and NO <sub>x</sub> (2016) | For each LSOA in Wales (n=1,909).              |
| Welsh Longitudinal General Practice (WLGP) | Primary care records.               | Diagnoses of schizophrenia/OPD (Read codes)                                                   | ~80% of GP practices in Wales.                 |
| Welsh Demographic Service (WDS)            | Digital Health and Care Wales       | Age, sex, week of birth, residential history.                                                 | All individuals registered with a GP in Wales. |
| Welsh Index of Multiple Deprivation (WIMD) | Welsh Government.                   | Area-level deprivation quintile (1 = most deprived; 5 = least deprived)                       | LSOA level (2019)                              |
| Rural/urban Classification                 | Office for National Statistics.     | Urban or rural.                                                                               | All LSOAs in Wales (n = 1,909)                 |

Table 2: Number and Percentage of Individuals in the Study Population with Recorded Diagnoses of Schizophrenia/ OPD in Welsh Primary Care Records

| Participants         | Number    | Percentage of all participants (%) |
|----------------------|-----------|------------------------------------|
| Total population     | 2,021,810 | (100.000)                          |
| Schizophrenia only   | 899       | (0.044)                            |
| OPD only             | 776       | (0.038)                            |
| Schizophrenia & OPD  | 109       | (0.005)                            |
| Schizophrenia or OPD | 1,784     | (0.088)                            |

level. The R code used to generate these summary statistics is provided in the supplementary materials.

## Results

Using a combination of open-source environmental and national mapping agency data, the feasibility of creating individual-level longitudinal environmental exposure data across Wales was demonstrated. In total, 2,021,810 individuals were included in the analysis. These individuals were

registered with a Welsh GP practice contributing data to the SAIL Databank and remained at the same residential address between the 1 July 2016 to 30 June 2019. Within this population, 0.1% (1,784) had a diagnosis of schizophrenia/OPD. The number and percentage of individuals with recorded diagnoses of schizophrenia/OPD within the study population is presented in Table 2.

The total cohort had a nearly equal distribution of males (50.2%) and females (49.8%). Among the 1,784 individuals who developed schizophrenia/OPD, a slightly higher proportion were male (52.7%) compared to female (47.3%). The mean age of the total cohort was 42.3 years (median 44, IQR 23–61), whereas individuals with incident schizophrenia/OPD were older on average, with a mean age of 46.9 years (median 45, IQR 28–65). Regarding residential environment, 70.5% of the total cohort resided in cities and towns, while 29.5% lived in rural or fringe areas. A similar pattern was observed among incident cases, with 72.3% living in urban areas and 27.7% in rural/fringe areas. Deprivation, measured using the WIMD, showed that the highest incidence of schizophrenia/OPD was observed among individuals in the most deprived quintile (29.1%), whereas the least deprived quintile accounted for 15.0% of incident cases, indicating a social gradient in the distribution of schizophrenia/OPD diagnoses (Figure 2). This information is presented in Table 3.

Figure 2: Distribution of Schizophrenia/OPD Diagnoses Across Welsh Index of Multiple Deprivation quintiles

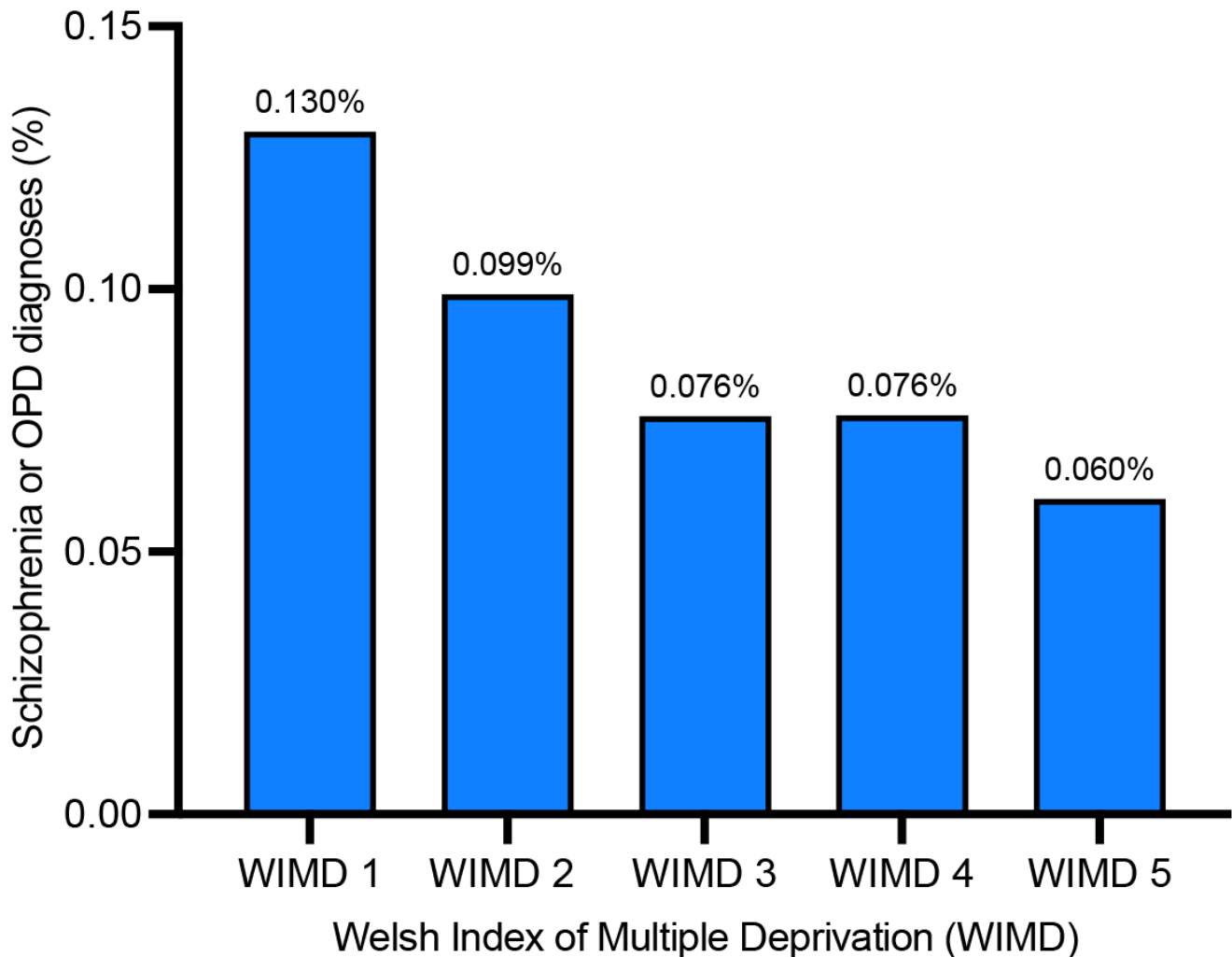


Table 3: Comparison of Demographic, Socioeconomic, and Environmental Characteristics Between the Total Study Population and Individuals with Incident Schizophrenia/OPD

| Characteristic     | Total cohort<br>(n= 2,021,810) | Incident<br>schizophrenia/OPD<br>(n = 1,784) |
|--------------------|--------------------------------|----------------------------------------------|
| <b>Sex</b>         |                                |                                              |
| Male               | 1,014,575 (50.2)               | 941 (52.7)                                   |
| Female             | 1,007,235 (49.8)               | 843 (47.3)                                   |
| <b>Age</b>         |                                |                                              |
| Mean               | 42.3                           | 46.9                                         |
| Median             | 44                             | 45                                           |
| IQR                | 23-61                          | 28-65                                        |
| <b>Environment</b> |                                |                                              |
| Rural and fringe   | 596,594 (29.5)                 | 494 (27.7)                                   |
| City and town      | 1,425,216 (70.5)               | 1,290 (72.3)                                 |
| <b>WIMD</b>        |                                |                                              |
| 1-Most deprived    | 413,529 (20.5)                 | 519 (29.1)                                   |
| 2                  | 411,408 (20.3)                 | 408 (22.9)                                   |
| 3                  | 393,140 (19.5)                 | 298 (16.7)                                   |
| 4                  | 384,453 (19.0)                 | 292 (16.3)                                   |
| 5 - Least deprived | 419,280 (20.7)                 | 267 (15.0)                                   |

Annual mean concentrations of PM<sub>10</sub>, PM<sub>2.5</sub>, and NO<sub>x</sub> in 2016 for the schizophrenia/OPD cohort and for all participants are presented in Tables 4–6. NO<sub>x</sub> encompasses nitric oxide (NO) and nitrogen dioxide (NO<sub>2</sub>), representing total nitrogen oxide emissions [43]. WHO guideline values are provided for PM<sub>10</sub> and PM<sub>2.5</sub> for contextual comparison, while no guideline exists for NO<sub>x</sub> as a combined metric. However, guideline values are available for NO<sub>2</sub> alone.

## Discussion

A cohort of individuals with their first GP diagnosis of schizophrenia/OPD between 2016–2019 was identified, representing 0.1% of the population. Within the psychiatric cohort, 52.7% were male and 47.3% were female, compared with an equal sex distribution (50.2% male; 49.8% female) in the general population. The psychiatric cohort had a higher mean age at study entry (46.9 years) than the general population (42.3 years). Among those with a diagnosis, 72.3% resided in urban areas and 27.7% in rural areas, closely reflecting the distribution of the general population (70.5% urban; 29.5% rural). Mean levels of PM<sub>10</sub>, PM<sub>2.5</sub>, and NO<sub>x</sub> were similar between the psychiatric cohort and the general population. In contrast, a clear socioeconomic gradient was observed: 29.1% of individuals with schizophrenia/OPD lived in the most deprived quintile compared with 15.0% in the least deprived quintile.

Previous research indicates schizophrenia incidence is higher in men than in women, with an approximate ratio of 1.4:1 [44]. However, this pattern is less consistent when considering lifetime prevalence [45]. Gender differences are also observed in age at onset, with men typically diagnosed

Table 4: Annual Mean Outdoor PM<sub>10</sub> Concentrations at the LSOA Level for All Participants and those with Schizophrenia/OPD

| PM <sub>10</sub> (µg/m <sup>3</sup> )                                                                   | All Participants | Participants with Schizophrenia/OPD |
|---------------------------------------------------------------------------------------------------------|------------------|-------------------------------------|
| Mean                                                                                                    | 11.60            | 11.59                               |
| Minimum                                                                                                 | 7.69             | 7.69                                |
| Median                                                                                                  | 11.65            | 11.64                               |
| Maximum                                                                                                 | 16.31            | 16.31                               |
| WHO annual mean guideline values: 20 µg/m <sup>3</sup> (2005–2020); 15 µg/m <sup>3</sup> (2021 update). |                  |                                     |

Table 5: Annual Mean Outdoor PM<sub>2.5</sub> Concentrations at the LSOA Level for All Participants and those with Schizophrenia/OPD

| PM <sub>2.5</sub> (µg/m <sup>3</sup> )                                                                 | All Participants | Participants with Schizophrenia/OPD |
|--------------------------------------------------------------------------------------------------------|------------------|-------------------------------------|
| Mean                                                                                                   | 7.38             | 7.37                                |
| Minimum                                                                                                | 4.83             | 4.83                                |
| Median                                                                                                 | 7.41             | 7.42                                |
| Maximum                                                                                                | 10.64            | 10.64                               |
| WHO annual mean guideline values: 10 µg/m <sup>3</sup> (2005–2020); 5 µg/m <sup>3</sup> (2021 update). |                  |                                     |

Table 6: Annual Mean Outdoor NO<sub>x</sub> Concentrations at the LSOA Level for All Participants and those with Schizophrenia or OPD

| NO <sub>x</sub> (µg/m <sup>3</sup> ) | All Participants | Participants with Schizophrenia/OPD |
|--------------------------------------|------------------|-------------------------------------|
| Mean                                 | 15.92            | 15.75                               |
| Minimum                              | 3.96             | 4.02                                |
| Median                               | 14.16            | 14.32                               |
| Maximum                              | 54.33            | 54.33                               |

No WHO air quality guideline value is available for NO<sub>x</sub>.

between 15 and 24 years and women more commonly after age 40 [45–47].

A population-based linkage study using primary care data found prevalence and incidence of schizophrenia was higher in more deprived areas [16]. A review similarly reported higher risk of psychotic disorders in deprived areas although adjustment for individual factors (age, sex, ethnicity and social class) attenuated this association [48]. The assessment of deprivation in this study included income, employment, health, education, access to services, housing, community safety and physical environment [49]. Bhavsar et al. [50] demonstrated in highly urban areas, an association between area-level deprivation and schizophrenia incidence, was largely explained by age, gender, ethnicity, population density, high crime and low education. Whilst low income, poor housing and living environment did not independently predict incidence [50]. Therefore, the association between area-level deprivation and schizophrenia could be driven, partly, by specific social and urban contextual factors rather than deprivation alone. This research highlights

the complexity of the relationship between deprivation and psychotic disorders.

Evidence shows schizophrenia prevalence is strongly socially patterned, since prevalence is higher in more deprived areas of society [51]. One possible explanation for this is that deteriorating mental state results in lower social position (social drift) [51]. However, evidence has not found people with diagnosed schizophrenia move to more deprived or urban areas [16]. Another explanation is that social position is a contributing factor for mental illness (social causation) [51]. However, Sariaslan et al [52]. found genetic risk for schizophrenia predicts neighbourhood deprivation, which could be evidence against environmental causes for this association. Therefore, overall, individuals in deprived areas could be more at risk of mental disorders [13, 53] however the causation for this is unknown and likely complex.

Urbanicity has also been associated with increased schizophrenia risk [16, 17], particularly during adolescence [15], although findings are mixed internationally [18]. Increased risk of schizophrenia could be a result of elevated exposure to environmental stressors, including air pollution [54, 55]. Higher residential exposure to NO<sub>2</sub>, NO<sub>x</sub>, PM<sub>2.5</sub> and PM<sub>10</sub> has been associated with increased mental health service use and incidence of schizophrenia spectrum disorders [56, 57]. However, in our study, pollutant levels showed minimal variation between the general population and those diagnosed with schizophrenia and other psychotic disorders. This may reflect the geographical and demographic characteristics of Wales. Since, Wales is characterised by smaller, dispersed urban centres and substantial rural areas rather than large metropolitan environments. Hence, the deprivation-

## Strengths and Limitations

Few population-level studies have described the environmental contexts of individuals with severe mental disorders using linked routine health data. Despite, a growing body of evidence investigating the aetiology of ecological determinants and mental disorders [59], research and related interventions have largely focused on biological, psychological, and social determinants [60, 61]. This study contributes to addressing this gap; however, several methodological and data limitations should be acknowledged.

First, no further analyses beyond descriptive statistics were conducted. Therefore, the results are limited to identifying trends, as descriptive statistics cannot be used to establish causal relationships.

Second, air pollution exposure was assigned at the LSOA level (1,500 people approximately per LSOA) rather than at the residential address level. This introduces the potential for ecological fallacy, whereby all individuals within an LSOA are assigned the same air pollution value, despite known variation over short distances. However, increasing the spatial resolution increases the computational complexity of the analyses and was not feasible in the lifetime of this study. Measuring air pollution exposure at a more individual level is a common difficulty [62]. Since participants may attend multiple locations throughout the day, such as work or school. Mental health outcomes are likely influenced by chronic long-term exposure to air pollution [57, 63]. However, this study was intended

as a pilot analysis to assess the feasibility of linking area-level air pollution exposure with mental health outcomes. Accordingly, only a single year of air quality data (2016) was used to provide a snapshot of exposure while limiting analytical complexity. Although multi-year modelled air pollution data are available via DEFRA, cumulative or long-term exposure assessment was beyond the scope of this initial investigation and is being addressed in ongoing work using SAIL data. This likely underestimates the potential impact of air pollution on severe mental disorders and represents a notable limitation.

Regarding mental health data, the use of individual-level GP records provided high-resolution data. However, the percentage of schizophrenia/OPD patients was smaller than expected (0.088% of the cohort). Reliance on GP data alone is a likely explanation for the relatively small number of recorded diagnoses. This may limit the generalisability of the findings. Moreover, individuals who have been diagnosed with schizophrenia/OPD could be more likely to move residence. Lee and colleagues [16] showed individuals with severe mental disorders such as schizophrenia were more likely to move compared to the general population. Therefore, in this study, the exclusion of individuals who moved residence could have contributed to excluding some cases of schizophrenia/OPD. Additionally, all individuals in the study were followed for the full three-year period, resulting in a survivor cohort. Consequently, suicide and premature mortality among people with schizophrenia/OPD may have influenced the size of the psychiatric cohort [64]. An additional limitation is that current diagnoses were not used. As previously justified, 1 July 2016 to 30 June 2019 was selected as the study period to avoid the COVID-19 pandemic (2020–2021). During this period, GP access was difficult and disrupted [26]. A thematic analysis of 10,089 people reported difficulties booking appointments, appointment's not meeting people's needs and a lack of access to regular treatment and medication [26]. Hence, this disruption is likely to have interfered with diagnostic recording and may have introduced bias.

The higher mean age of individuals at the study start date may also be explained by the restriction to individual who did not change LSOA in the 3-year period. Younger age groups, in which residential mobility is higher, particularly those aged 19–29, may therefore be underrepresented.

ONS data indicate that young adults are the most likely to move residence, with a peak at age 19, reflecting transitions into higher education [65].

urbanicity–pollution gradient observed in highly urbanised settings may be less pronounced [58].

## Implications for Future Research

Understanding the environmental contexts of individuals diagnosed with severe mental disorders is crucial for identifying potential risk factors, resilience mechanisms, and protective factors. This is particularly important because this population is at a higher risk of poor physical health and reduced life expectancy compared with the general population [7].

Future research should examine the environments in which people with schizophrenia/OPD reside and work. Factors that could be investigated include pollution levels, greenness, urban versus rural settings, and socioeconomic deprivation. Studies could aim to include a more representative range

of diagnoses by collecting data from various mental health settings, including general practice, inpatient, and outpatient hospital records, ideally at the population level. The use of up-to-date records is recommended wherever possible.

Research would also benefit from improved exposure assessment methods that consider multiple daily locations, including home, work, and school, as well as indoor exposures, to more accurately capture individuals' everyday environmental interactions. It is also recommended future work incorporates multi-year exposure data to better understand long-term environmental impacts on severe mental disorders.

## Conclusions

In this population-wide descriptive study, individuals with schizophrenia or other psychotic disorders were disproportionately represented in more deprived areas. Distributions across urban and rural settings, as well as area-level air pollution exposures, were broadly similar to those in the general population. These findings highlight the importance of socioeconomic context when investigating inequalities in severe mental illness. Linked environmental and healthcare datasets provide opportunities for future longitudinal research. Such studies could examine the relationships between deprivation, urbanicity, environmental exposures, and severe mental disorders.

## Statement of Conflicts of Interest

The authors declare that they have no conflicts of interest with respect to the research, authorship and/or publication of this article.

## Ethics Statement

The use of deidentified data in SAIL complies with National Research Ethics Service (NRES) guidance. Applications to use data held within the SAIL Databank, an ISO: 27001 and UK Statistics Authority (UKSA) Digital Economy Act (DEA) accredited Trusted Research Environment, must first be approved by the independent Information Governance Review Panel (IGRP). This panel contains individuals with expertise in data governance and protection, including the Chair of the Wales NRES Committee, Caldicott Guardians and members of the public. The IGRP approved SAIL project 1413 on March 2023.

## Funding

This research was funded by the UK Research and Innovation (UKRI) Natural Environment Research Council (NERC) 'RESPIRE' study (Grant No. NE/W002264/1) (MJDC) and supported by the *In Vitro* Toxicology Group at Swansea University Medical School. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Author Contributions

EC conducted the data analysis and drafted the manuscript. RF and AJ provided intellectual input throughout the study. MJDC, JL and AM contributed to study conception and design, and supported data analysis and interpretation. JL created the linked cohort which included the Welsh Longitudinal General Practice dataset to identify schizophrenia/OPD diagnoses within the study period. MJDC supported the direct funding of the project. AM provided the air pollution data at the LSOA level. All authors critically reviewed the manuscript and contributed towards iterations prior to submission. All authors have approved the final submission of the article.

## Data Availability Statement

The data used in this study was accessed through Secure Anonymised Information Linkage (SAIL) Databank. Access to SAIL requires approval from the IGRP. Approved researchers can access the data *via* a privacy-protecting safe haven and remote access system, termed SAIL Gateway. The project number was 1413.

## AI Disclosure Statement

The authors used ChatGPT (OpenAI) for minor language editing to improve grammar and sentence clarity. The authors reviewed and verified all changes and take full responsibility for the manuscript.

## References

1. World Health Organisation. Mental disorders [Internet]. 2022 [cited 2024 Jan 9]. Available from: <https://www.who.int/news-room/fact-sheets/detail/mental-disorders>.
2. World Health Organisation. Mental health [Internet]. 2024 [cited 2024 Jan 9]. Available from: <https://www.who.int/health-topics/mental-health>.
3. Patel V, Saxena S, Lund C, Thornicroft G, Baingana F, Bolton P, et al. The Lancet Commission on global mental health and sustainable development. *Lancet*. 2018 Oct 27;392(10157):1553–98. [https://doi.org/10.1016/S0140-6736\(18\)31612-X](https://doi.org/10.1016/S0140-6736(18)31612-X). PubMed PMID: 30314863.
4. Dattani S, Rodés-Guirao L, Ritchie H, Roser M. Mental Health. Our World in Data [Internet]. 2023 Dec 28 [cited 2024 Feb 16]. Available from: <https://ourworldindata.org/mental-health>.
5. NHS. Adult and older adult mental health [Internet]. 2024 [cited 2024 Mar 1]. Available from: <https://www.england.nhs.uk/mental-health/adults/>.
6. Royal College of Psychiatrists. Lost years: The impact of mental ill health on life expectancy and what needs to change [Internet]. 2024. Available

- from: [https://www.rcpsych.ac.uk/docs/default-source/members/devolved-nations/rcpsych-in-wales/rcpsych-wales-smi-5-point-plan.pdf?sfvrsn=456a3b0c\\_5#:~:text=The%20Welsh%20Government%20has%20previously,doubled%20during%20COVID%2D19%20pandemic.](https://www.rcpsych.ac.uk/docs/default-source/members/devolved-nations/rcpsych-in-wales/rcpsych-wales-smi-5-point-plan.pdf?sfvrsn=456a3b0c_5#:~:text=The%20Welsh%20Government%20has%20previously,doubled%20during%20COVID%2D19%20pandemic.)
7. UK Government. GOV.UK [Internet]. 2018 [cited 2024 Jan 22]. Health matters: reducing health inequalities in mental illness. Available from: <https://www.gov.uk/government/publications/health-matters-reducing-health-inequalities-in-mental-illness/health-matters-reducing-health-inequalities-in-mental-illness>
8. Fiorillo A, Sartorius N. Mortality gap and physical comorbidity of people with severe mental disorders: the public health scandal. *Annals of General Psychiatry*. 2021 Dec 13;20(1):52. <https://doi.org/10.1186/s12991-021-00374-y>
9. Nielsen RE, Banner J, Jensen SE. Cardiovascular disease in patients with severe mental illness. *Nat Rev Cardiol*. 2021 Feb;18(2):136–45. <https://doi.org/10.1038/s41569-020-00463-7>. PubMed PMID: 33128044.
10. Eisenberger NI, Moieni M. Inflammation affects social experience: implications for mental health. *World Psychiatry*. 2020 Feb;19(1):109–10. <https://doi.org/10.1002/wps.20724>. PubMed PMID: 31922673; PubMed Central PMCID: PMC6953558.
11. Shaw S, Van Heyst B. An Evaluation of Risk Ratios on Physical and Mental Health Correlations due to Increases in Ambient Nitrogen Oxide (NO<sub>x</sub>) Concentrations. *Atmosphere*. 2022 Jun 1;13:967. <https://doi.org/10.3390/atmos13060967>
12. Brunt H, Barnes J, Jones SJ, Longhurst JWS, Scally G, Hayes E. Air pollution, deprivation and health: understanding relationships to add value to local air quality management policy and practice in Wales, UK. *Journal of Public Health*. 2017 Sep 1;39(3):485–97. <https://doi.org/10.1093/pubmed/fdw084>
13. Knifton L, Inglis G. Poverty and mental health: policy, practice and research implications. *BJPsych Bull*. 2020;44(5):193–6. <https://doi.org/10.1192/bjb.2020.78>. PubMed PMID: 32744210; PubMed Central PMCID: PMC7525587.
14. Currie J, Boyce T, Evans L, Luker M, Senior S, Hartt M, et al. Life expectancy inequalities in Wales before COVID-19: an exploration of current contributions by age and cause of death and changes between 2002 and 2018. *Public Health*. 2021 Apr 1;193:48–56. <https://doi.org/10.1016/j.puhe.2021.01.025>
15. Gruebner O, A. Rapp M, Adli M, Kluge U, Galea S, Heinz A. Cities and Mental Health. *Dtsch Arztebl Int*. 2017 Feb;114(8):121–7. <https://doi.org/10.3238/arztebl.2017.0121>. PubMed PMID: 28302261; PubMed Central PMCID: PMC5374256.
16. Lee SC, DelPozo-Banos M, Lloyd K, Jones I, Walters JTR, Owen MJ, et al. Area deprivation, urbanicity, severe mental illness and social drift — A population-based linkage study using routinely collected primary and secondary care data. *Schizophrenia Research*. 2020 Jun 1;220:130–40. <https://doi.org/10.1016/j.schres.2020.03.044>
17. Ventriglio A, Torales J, Castaldelli-Maia JM, Berardis DD, Bhugra D. Urbanization and emerging mental health issues. *CNS Spectrums*. 2021 Feb;26(1):43–50. <https://doi.org/10.1017/S1092852920001236>
18. Fett AKJ, Lemmers-Jansen ILJ, Krabbendam L. Psychosis and urbanicity: a review of the recent literature from epidemiology to neurourbanism. *Curr Opin Psychiatry*. 2019 May;32(3):232–41. <https://doi.org/10.1097/YCO.0000000000000486>. PubMed PMID: 30724751; PubMed Central PMCID: PMC6493678.
19. Lauwers L, Trabelsi S, Pelgrims I, Bastiaens H, Clercq ED, Guilbert A, et al. Urban environment and mental health: the NAMED project, protocol for a mixed-method study. *BMJ Open*. 2020 Feb 1;10(2):e031963. <https://doi.org/10.1136/bmjopen-2019-031963>. PubMed PMID: 32086354.
20. Catapano P, Luciano M, Cipolla S, D'Amico D, Cirino A, Della Corte MC, et al. What is the relationship between exposure to environmental pollutants and severe mental disorders? A systematic review on shared biological pathways. *Brain, Behavior, & Immunity - Health*. 2025 Feb 1;43:100922. <https://doi.org/10.1016/j.bbih.2024.100922>
21. Antonsen S, Mok PLH, Webb RT, Mortensen PB, McGrath JJ, Agerbo E, et al. Exposure to air pollution during childhood and risk of developing schizophrenia: a national cohort study. *The Lancet Planetary Health*. 2020 Feb 1;4(2):e64–73. [https://doi.org/10.1016/S2542-5196\(20\)30004-8](https://doi.org/10.1016/S2542-5196(20)30004-8)
22. Horsdal HT, Agerbo E, McGrath JJ, Vilhjálmsdóttir BJ, Antonsen S, Cløster AM, et al. Association of Childhood Exposure to Nitrogen Dioxide and Polygenic Risk Score for Schizophrenia With the Risk of Developing Schizophrenia. *JAMA network open*. 2019 Nov 1;2(11):e1914401. Located at: cmedm; 31675084. <https://doi.org/10.1001/jamanetworkopen.2019.14401>
23. Liu R, Li D, Ma Y, Tang L, Chen R, Tian Y. Air pollutants, genetic susceptibility and the risk of schizophrenia: large prospective study. *The British Journal of Psychiatry*. 2024 Oct;225(4):427–35. <https://doi.org/10.1192/bjp.2024.118>
24. Ford DV, Jones KH, Verplancke JP, Lyons RA, John G, Brown G, et al. The SAIL Databank: building a national architecture for e-health research and evaluation. *BMC Health Services Research*. 2009 Sep 4;9(1):157. <https://doi.org/10.1186/1472-6963-9-157>

25. Lyons RA, Jones KH, John G, Brooks CJ, Verplancke JP, Ford DV, et al. The SAIL databank: linking multiple health and social care datasets. *BMC Medical Informatics and Decision Making*. 2009 Jan 16;9(1):3. <https://doi.org/10.1186/1472-6947-9-3>
26. Healthwatch. GP access during COVID-19 | Healthwatch [Internet]. 2021 [cited 2024 Oct 22]. Available from: <https://www.healthwatch.co.uk/report/2021-03-22/gp-access-during-covid-19>.
27. ONS. Estimates of the population for the UK, England, Wales, Scotland and Northern Ireland - Office for National Statistics [Internet]. 2022 [cited 2024 Mar 15]. Available from: <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/populationestimatesforukenglandandwalesscotlandandnorthernireland>.
28. ClientEarth. UK Air Pollution: How clean is the air you breathe? | ClientEarth [Internet]. 2022 [cited 2024 Jan 8]. Available from: <https://www.clientearth.org/latest/news/uk-air-pollution-how-clean-is-the-air-you-breathe/>.
29. Francesca Howorth. Clean air for Wales [Internet]. 2021 [cited 2024 Jan 8]. Available from: <https://research.senedd.wales/research-articles/clean-air-for-wales/>.
30. IQAir. REPORT: Over 90% of global population breathes dangerously polluted air [Internet]. 2019 [cited 2024 Jan 8]. Available from: <https://www.iqair.com/gb/newsroom/report-over-90-percent-of-global-population-breathes-dangerously-polluted-air>.
31. Welsh Government. Fine particulate matter air quality targets | GOV.WALES [Internet]. 2025 [cited 2026 Jan 22]. Available from: <https://www.gov.wales/fine-particulate-matter-air-quality-targets>.
32. Office for National Statistics. Towns and cities, characteristics of built-up areas, England and Wales - Office for National Statistics [Internet]. 2023 [cited 2024 Mar 15]. Available from: <https://www.ons.gov.uk/peoplepopulationandcommunity/housing/articles/townsandcitiescharacteristicsofbuiltupareasenglandandwales/census2021>.
33. NHS Digital. NHS Digital [Internet]. 2022 [cited 2024 Jan 8]. Read Codes. Available from: <https://digital.nhs.uk/services/terminology-and-classifications/read-codes>.
34. John, A, McGregor, J., Jones, I., Lee, S. C., Walters, J. T. R., Owen, M. J., O'Donovan, M., DelPozo-Banos, M., Berridge, D., & Lloyd, K. SAIL databank concept library [Internet]. 2024a [cited 2024 Mar 15]. Schizophrenia Primary Care. Available from: <http://conceptlibrary.saildatabank.com/phenotypes/PH939/version/1957/detail/>.
35. John, A, McGregor, J., Jones, I., Lee, S. C., Walters, J. T. R., Owen, M. J., O'Donovan, M., DelPozo-Banos, M., Berridge, D., & Lloyd, K. SAIL databank concept library [Internet]. 2024 [cited 2024 Mar 15]. Other Psychotic Disorders Primary Care. Available from: <http://conceptlibrary.saildatabank.com/phenotypes/PH938/version/1955/detail/>.
36. Welsh Government. Home | DataMapWales [Internet]. 2024 [cited 2024 Jul 29]. Available from: <https://datamap.gov.wales/>.
37. ONS. Census 2021 geographies - Office for National Statistics [Internet]. 2024 [cited 2024 Jan 8]. Available from: <https://www.ons.gov.uk/methodology/geography/ukgeographies/censusgeographies/census2021geographies>.
38. UK Government. GOV.UK [Internet]. 2018 [cited 2024 Mar 1]. Methods, data and definitions. Available from: <https://www.gov.uk/government/publications/health-profile-for-england-2018/methods-data-and-definitions>.
39. ONS. 2011 rural/urban classification - Office for National Statistics [Internet]. 2016 [cited 2024 Jan 8]. Available from: <https://www.ons.gov.uk/methodology/geography/geographicalproducts/ruralurbanclassifications/2011ruralurbanclassification>.
40. Mizen A, Johnson R, Dearden L, Sallakh MA, Mavrogianni A, Davies G, et al. Creating individual level air pollution exposures in an anonymised data safe haven: a platform for evaluating impact on educational attainment. *International Journal of Population Data Science*. 2018 Aug 21;3(1):1. <https://doi.org/10.23889/ijpds.v3i1.412>
41. Welsh Government. Welsh Index of Multiple Deprivation [Internet]. 2024 [cited 2024 Jan 8]. Available from: <https://stats.wales.gov.wales/Catalogue/Community-Safety-and-Social-Inclusion/Welsh-Index-of-Multiple-Deprivation>.
42. Welsh Government. Welsh Index of Multiple Deprivation 2019: analysis relating to areas of deep-rooted deprivation | GOV.WALES [Internet]. 2022 [cited 2024 Jan 8]. Available from: <https://www.gov.wales/welsh-index-multiple-deprivation-2019-analysis-relating-areas-deep-rooted-deprivation-html>.
43. Defra. GOV.UK [Internet]. 2025 [cited 2025 Oct 16]. Emissions of air pollutants in the UK - Nitrogen oxides (NOx). Available from: <https://www.gov.uk/government/statistics/emissions-of-air-pollutants/emissions-of-air-pollutants-in-the-uk-nitrogen-oxides-nox>.
44. Li X, Zhou W, Yi Z. A glimpse of gender differences in schizophrenia. *General Psychiatry*. 2022;35(4). <https://doi.org/10.1136/gpsych-2022-100823>. PubMed PMID: 36118418.
45. McGrath J, Saha S, Welham J, El Saadi O, MacCauley C, Chant D. A systematic review of the

- incidence of schizophrenia: the distribution of rates and the influence of sex, urbanicity, migrant status and methodology. *BMC Medicine*. 2004 Apr 28;2(1):13. <https://doi.org/10.1186/1741-7015-2-13>
46. Chan V. Schizophrenia and Psychosis: Diagnosis, Current Research Trends, and Model Treatment Approaches with Implications for Transitional Age Youth. *Child and Adolescent Psychiatric Clinics*. 2017 Apr 1;26(2):341–66. <https://doi.org/10.1016/j.chc.2016.12.014>. PubMed PMID: 28314460.
47. Giordano GM, Bucci P, Mucci A, Pezzella P, Galderisi S. Gender Differences in Clinical and Psychosocial Features Among Persons With Schizophrenia: A Mini Review. *Front Psychiatry*. 2021;12:789179. <https://doi.org/10.3389/fpsyt.2021.789179>. PubMed PMID: 35002807; PubMed Central PMCID: PMC8727372.
48. O'Donoghue B, Roche E, Lane A. Neighbourhood level social deprivation and the risk of psychotic disorders: a systematic review. *Soc Psychiatry Psychiatr Epidemiol*. 2016 Jul 1;51(7):941–50. <https://doi.org/10.1007/s00127-016-1233-4>
49. Welsh Government. Welsh Index of Multiple Deprivation (WIMD) 2019: results report. 2019.
50. Bhavsar V, Boydell J, Murray R, Power P. Identifying aspects of neighbourhood deprivation associated with increased incidence of schizophrenia. *Schizophrenia Research*. 2014 Jun 1;156(1):115–21. <https://doi.org/10.1016/j.schres.2014.03.014>
51. Gage SH, Davey Smith G, Munafò MR. Schizophrenia and neighbourhood deprivation. *Transl Psychiatry*. 2016 Dec;6(12):e979. <https://doi.org/10.1038/tp.2016.244>. PubMed PMID: 27959332; PubMed Central PMCID: PMC5290338.
52. Sariaslan A, Fazel S, D'Onofrio BM, Långström N, Larsson H, Bergen SE, et al. Schizophrenia and subsequent neighborhood deprivation: revisiting the social drift hypothesis using population, twin and molecular genetic data. *Transl Psychiatry*. 2016 May;6(5):5. <https://doi.org/10.1038/tp.2016.62>
53. Bernardini F, Attademo L, Trezzi R, Gobbi C, Balducci PM, Bello VD, et al. Air pollutants and daily number of admissions to psychiatric emergency services: evidence for detrimental mental health effects of ozone. *Epidemiology and Psychiatric Sciences*. 2020 Jan;29:e66. <https://doi.org/10.1017/S2045796019000623>
54. Zhan C, Xie M, Lu H, Liu B, Wu Z, Wang T, et al. Impacts of urbanization on air quality and the related health risks in a city with complex terrain. *Atmospheric Chemistry and Physics*. 2023 Jan 17;23(1):771–88. <https://doi.org/10.5194/acp-23-771-2023>
55. Liang L, Gong P. Urban and air pollution: a multi-city study of long-term effects of urban landscape patterns on air quality trends. *Sci Rep*. 2020 Oct 29;10(1):1. <https://doi.org/10.1038/s41598-020-74524-9>
56. Newbury JB, Stewart R, Fisher HL, Beevers S, Dajnak D, Broadbent M, et al. Association between air pollution exposure and mental health service use among individuals with first presentations of psychotic and mood disorders: retrospective cohort study. *Br J Psychiatry*. 2021;219(6):678–85. <https://doi.org/10.1192/bjp.2021.119>. PubMed PMID: 35048872; PubMed Central PMCID: PMC8636613.
57. Nobile F, Forastiere A, Michelozzi P, Forastiere F, Stafoggia M. Long-term exposure to air pollution and incidence of mental disorders. A large longitudinal cohort study of adults within an urban area. *Environment International*. 2023 Nov 1;181:108302. <https://doi.org/10.1016/j.envint.2023.108302>
58. Fecht D, Fischer P, Fortunato L, Hoek G, De Hoogh K, Marra M, et al. Associations between air pollution and socioeconomic characteristics, ethnicity and age profile of neighbourhoods in England and the Netherlands. *Environmental Pollution*. 2015 Mar;198:201–10. <https://doi.org/10.1016/j.envpol.2014.12.014>
59. Braithwaite I, Zhang S, Kirkbride JB, Osborn DPJ, Hayes JF. Air Pollution (Particulate Matter) Exposure and Associations with Depression, Anxiety, Bipolar, Psychosis and Suicide Risk: A Systematic Review and Meta-Analysis. *Environmental Health Perspectives*. 2019 Dec;127(12):126002. <https://doi.org/10.1289/EHP4595>
60. Lund C, Brooke-Sumner C, Baingana F, Baron EC, Breuer E, Chandra P, et al. Social determinants of mental disorders and the Sustainable Development Goals: a systematic review of reviews. *The Lancet Psychiatry*. 2018 Apr 1;5(4):357–69. [https://doi.org/10.1016/S2215-0366\(18\)30060-9](https://doi.org/10.1016/S2215-0366(18)30060-9)
61. Zwicker A, MacKenzie LE, Drobinin V, Bagher AM, Howes Vallis E, Propper L, et al. Neurodevelopmental and genetic determinants of exposure to adversity among youth at risk for mental illness. *Journal of Child Psychology and Psychiatry*. 2020;61(5):536–44. <https://doi.org/10.1111/jcpp.13159>
62. Hahad O, Lelieveld J, Birklein F, Lieb K, Daiber A, Münzel T. Ambient Air Pollution Increases the Risk of Cerebrovascular and Neuropsychiatric Disorders through Induction of Inflammation and Oxidative Stress. *International Journal of Molecular Sciences*. 2020 Jan;21(12):12. <https://doi.org/10.3390/ijms21124306>
63. Bhui K, Newbury JB, Latham RM, Ucci M, Nasir ZA, Turner B, et al. Air quality and mental health: evidence, challenges and future directions. *BJPsych Open*. 2023 Jul;9(4):e120. <https://doi.org/10.1192/bjo.2023.507>
64. Hor K, Taylor M. Suicide and schizophrenia: a systematic review of rates and risk factors. *J Psychopharmacol*. 2010 Nov;24(4\_supplement):81–90. <https://doi.org/10.1177/1359786810385490>. PubMed PMID: 20923923; PubMed Central PMCID: PMC2951591.

65. Office for National Statistics. Internal migration, England and Wales: Year Ending June 2015 [Internet]. 2016. Available from: <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/migrationwithintheuk/bulletins/internalmigrationbylocalauthoritiesinenglandandwales/yearendingjune2015>.



## Appendix A: Code used in SAIL

### A.1 Installation

```
install.packages("sailr");
library(sailr);
install.packages("tidyverse")
library(tidyverse)

conn <- sail_open();

q <- "SELECT * FROM SAILW1413V.JL_COHORT";
```

### A.2 Descriptive statistics for all participants

```
#Number of all participants
nrow(data)

#Male (1)
sum(data$gndr_cd==1)

#Female (2)
sum(data$gndr_cd==2)

#Min, Q1, Median, Mean, Q3, Max Age
summary(data$age)

#Environment
unique(data$urban_rural_inception)
#Rural and fringe
sum(data$urban_rural_inception=="Rural town and fringe")
sum(data$urban_rural_inception== "Rural village and dispersed in a sparse setting")
sum(data$urban_rural_inception=="Rural town and fringe in a sparse setting")
sum(data$urban_rural_inception=="Rural village and dispersed")
#City and town
sum(data$urban_rural_inception=="Urban city and town")
sum(data$urban_rural_inception=="Urban city and town in a sparse setting")

#Deprivation
unique(data$wimd2019_quintile_desc_inception)
sum(data$wimd2019_quintile_desc_inception=="1. Most deprived ")
sum(data$wimd2019_quintile_desc_inception=="2 ")
sum(data$wimd2019_quintile_desc_inception=="3 ")
sum(data$wimd2019_quintile_desc_inception=="1. Least deprived ")
```

### A.3 Descriptive statistics of the schizophrenia and OPD cohort

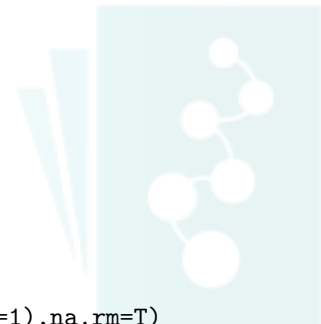
```
#All
sum((data$opd==1|data$schiz==1),na.rm=T)

# People with only schiz
sum((data$schiz==1&data$opd!=1),na.rm=T)
# People with only OPD
sum((data$opd==1&data$schiz!=1),na.rm=T)
# People with both
sum((data$opd==1&data$schiz==1),na.rm=T)

#Sex
sum(((data$opd==1|data$schiz==1)&(data$gndr_cd==2)),na.rm=T)

#Min, Q1, Median, Mean, Q3, Max Age
summary(data$age, data$opd==1|data$schiz==1)

#Rural and fringe
sum((data$urban_rural_inception=="Rural town and fringe") &(data$opd==1|data$schiz==1),na.rm=T)
sum((data$urban_rural_inception== "Rural village and dispersed in a sparse setting")&(data$opd=
=1|data$schiz==1),na.rm=T)
```



```

sum((data$urban_rural_inception=="Rural town and fringe in a sparse setting")&(data$opd==1|data$
schiz==1),na.rm=T)
sum((data$urban_rural_inception=="Rural village and dispersed")&(data$opd==1|data$schiz==1),na.rm=T)

#City and town
Sum((data$urban_rural_inception=="Urban city and town")&(data$opd==1|data$schiz==1),na.rm=T)
sum((data$urban_rural_inception=="Urban city and town in a sparse setting")&(data$opd==1|data$schiz
==1),na.rm=T)

#Deprivation
sum((data$wimd2019_quintile_desc_inception=="1. Most deprived ")&(data$opd==1|data$schiz==1),na.rm=T)
sum((data$wimd2019_quintile_desc_inception=="2 ")&(data$opd==1|data$schiz==1),na.rm=T)
sum((data$wimd2019_quintile_desc_inception=="3 ")&(data$opd==1|data$schiz==1),na.rm=T)
sum((data$wimd2019_quintile_desc_inception=="1. Least deprived ")&(data$opd==1|data$schiz==1),na.rm=T)

```

#### A.4 Air pollution values

```

#Air pollutants for All Participants
summary(data$pm10_mean)
summary(data$pm25_mean)
summary(data$nox_mean)

#Air pollutants Schizophrenia and Other Psychotic disorders Cohort
summary(subset(data, data$opd==1|data$schiz==1))

```

#### A.5 Code used to create Boxplots and histograms

```

#Type of distribution
#PM10
boxplot(data$pm10_mean)
hist(data$pm10_mean)

#PM2.5
boxplot(data$pm25_mean)
hist(data$pm25_mean)

#NOx
boxplot(data$nox_mean)
hist(data$nox_mean)

```



## Appendix B: Distribution of the variables

Figure B.1: Boxplots and histogram of the  $\text{NO}_x$  (A and B),  $\text{PM}_{2.5}$  (C and D), and  $\text{PM}_{10}$  (E and F) means

